Topics in Empirical Political Economy

by

ELEONORA ALABRESE

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy in Economics

Department of Economics

DECEMBER 2022

Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank all people who contributed to the work described in this thesis. First and foremost, I thank my academic advisors, Thiemo Fetzer, Sascha Becker, and Andreas Stegmann, for their dedicated support. During my doctorate, they actively contributed to a rewarding graduate school experience by sharing precious insights into the life of an academic, giving me intellectual freedom in my work, engaging me in new ideas, and demanding a high quality of work in all my endeavours. Additionally, I would like to thank my examiners, Mirko Draca and Andrea Tesei, for their interest in my work.

Every result described in this thesis was accomplished with the help and support of the active research community of old and new members of the Department of Economics at Warwick Uni. Each one provided a friendly and cooperative atmosphere and insightful comments on my work. Firstly, Vera Troeger had the merit of wholeheartedly supporting me as a woman in academia. Secondly, I greatly benefited from the human and scientific insights of Wiji Arulampalam, Manuel Bagues, Luis Candelaria, Emma Duchini, James Fenske, Ludovica Gazze, Amrita Kulka, Arianna Ornaghi, Roberto Pancrazi, Kirill Pogorelskiy, and Natalia Zinovyeva. Everyone has been extremely generous with their time and attention, allowing me both personal and academic growth. Lastly, I thank Natalie Deven and Colin Ellis, who deserve credit for assisting with administrative tasks and for immediate support for any problems encountered.

I am much indebted to my co-authors Francesco Capozza, Federica Liberini, Dennis Novy, Francesco Porcelli, Michela Redoano, Antonio Russo, and Carlo Schwarz, whose intellectual and personal interaction enrich my life as a researcher daily.

I am blissful for sharing these years with my fellow PhD students, especially: Ashish, Apurav, Carolina, Mary, Andrea, Giulia and all the *Rushmorers*. Everyone provided me with endless learning opportunities, big ears and lots of fun.

I am deeply grateful to my non-academic network. Mom and Dad harbour the responsibility of providing me with lifelong support. Giulia has the virtue to *get me* from our first day of undergrad. Mauro and Ale have the character of hearing me rambling about academia and providing the much-needed shoulder. Last but not least, my friends from home — Analcoholic, Molise, and Casa Merini — have the merit of making me feel grounded despite time and distance.

Declarations

This thesis is submitted to the University of Warwick in accordance with the requirements of the degree of Doctor of Philosophy. I declare that any material contained in this thesis has not been submitted for a degree to any other university. Chapter 1 is my own, single-authored work. Chapter 2 is my own, single-authored work. Chapter 3 is collaborative and published work with Sascha O. Becker, Thiemo Fetzer, and Dennis Novy (see Alabrese et al., 2019).

Eleonora Alabrese

Abstract

This work focuses on misinformation and on voters' behaviour and addresses questions that have direct implication for policy.

Chapter 1: *How does media coverage of research articles shape their retraction process?*

Flawed research can be harmful both within and outside of academia. The media can play an important role in drawing broader attention to research, but may also ensure that research, once retracted, ceases to feature in popular discourse. Yet, there is little evidence on whether media reporting influences the retraction process and authors' careers. This chapter shows that the salience of a research article at publication amplifies the impact of a later retraction on its citations and the research output of its authors.

Chapter 2: *What is the relationship between voters' participation, opinion polls, and the electoral system?*

A central challenge for social scientists consists in explaining why people vote and what are the consequences of their behaviour. In this chapter I study one of the most contested drivers of voters' participation which is the role of opinion polls. Voters may use polls information when deciding whether or not to vote, but the relevance of this information may depend on the electoral system. Looking at UK general elections I find evidence that polls predictions interact with the recent local electoral preferences of a constituency, and significantly impact voters' participation, concentration of vote shares, and local parties' performances.

Chapter 3: *What qualifies populist attitudes? Can we infer individual associations from aggregate data?*

Early analyses of the 2016 Brexit referendum used region-level data or small samples based on polling data. The former might be subject to ecological fallacy and the latter might suffer from small-sample bias. Using individual-level data on thousands of respondents in *Understanding Society*, we find that voting Leave is associated with older age, white ethnicity, low educational attainment, infrequent use of smartphones and the internet, receiving benefits, adverse health and low life satisfaction. These results coincide with corresponding patterns at the aggregate level of voting areas. We therefore do not find evidence of ecological fallacy.

Chapter 1

BAD SCIENCE: RETRACTIONS AND MEDIA COVERAGE

I am greatful to Thiemo Fetzer, Sascha Becker and Andreas Stegmann for the precious mentoring and guidance. I want to thank Arun Advani, Wiji Arulampalam, Manuel Bagues, Luis Candelaria, Gabriella Conti, Mirko Draca, James Fenske, Ludovica Gazze, Vardges Levonyan, Andrew Oswald, Kirill Pogorelskiy, Carlo Schwarz, Paula Stephan, Vera Troeger, Natalia Zinovyeva and participants of European Winter Meeting of the Econometric Society 2022 (Berlin); CESifo 16th Workshop on Political Economy (Dresden); NSEF Workshop 2022 (Napoli Federico II); General Online Research Conference 2022 (Berlin); European Economic Science Association 2022 (Bologna); 37th Congress of the European Economic Association (Bocconi); SMSociety 2022 International Conference on Social Media and Society (online); CAGE Summer School 2022 (Warwick); PaCSS Politics and Computational Social Science (online); PolMeth Europe meeting (Hamburg); 10th Warwick Economics PhD Conference 2022 (Warwick); OFCOM Workshop on the Economics and Regulation of Social Media (London); 4th QMUL Economics and Finance Workshop (Queen Mary); Applied Young Economist Webinar (online); Workshop on the Organisation, Economics, and Policy of Scientific Research 2022 (KU Leuven); Monash-Warwick-Zurich Text-As-Data Workshop (online); RES 2022 Annual Conference (online); RGS Doctoral Conference in Economics (online); Workshop on Political Economy (UniBZ-Brunek), the Political Economy and Public Economics Research Away Day (Warwick) and Warwick internal seminars for insightful comments and suggestions. All errors remain my own.

Abstract

Flawed research can be harmful both within and outside of academia. Even when published research has been retracted and refuted by the scientific community, it may continue to be a source of misinformation. The media can play an important role in drawing broader attention to research, but may also ensure that research, once retracted, ceases to feature in popular discourse. Yet, there is little evidence on whether media reporting influences the retraction process and authors' careers. Using a conditional difference-in-differences strategy, this paper shows that articles that gained popularity in the media at publication and were later retracted face heavy citation losses, while subsequent citations become more *accurate*. Further, authors of such papers see a permanent decline in research output. Lastly, the paper provides evidence that media can influence both the likelihood of retraction and its timing, highlighting that the media can play an important role in contributing to the integrity of the research process.

Keywords: Science; Retractions; Media Coverage; Misinformation; Altmetric; Citations; Career impact.

1.1 Introduction

Poor quality research, even when retracted, often persists and perpetuates misinformation among academic and non-academic audiences (Lewandowsky et al., 2012; The Economist, 2021; Peng et al., 2022). A recent case study illustrates how research continues to be cited positively and uncritically in support of a medical nutrition intervention, without mention of its retraction for falsifying data (Schneider et al., 2020). Inattention (Woo and Walsh, 2021) or failure to update beliefs (Goncalves et al., 2021) are two possible channels that could explain the ongoing use of retracted work. For a flawed study to be quashed, the literature on retractions suggests the visibility and the accessibility of retractions and their associated retraction notices are decisive (Bar-Ilan and Halevi, 2017; Teixeira da Silva and Bornemann-Cimenti, 2017; Cox et al., 2018; Bordignon, 2020). Yet, there is limited evidence on the potential role of media reporting of science.[1] The media can play an important role in drawing broader attention to research, but may as well ensure that research, once retracted, ceases to feature in popular discourse. This work hence studies whether and how media attention at the time of publication impacts the survival of retracted work (i.e. future citations) and the future career outcomes of its authors (i.e. future publication rate).

A recent example of well-published fraudulent Covid-19 research (Mehra et al., 2020) illustrate the role of media raising attention to research, while potentially interacting with the retraction process. These two studies attracted wide media coverage right after publication and led to a global halt of hydroxychloroquine trials. But high scrutiny from the scientific community together with an investigation from the Guardian[2] led to a prompt retraction, as the underlining data were fabricated. However, it is unclear whether this interaction with the media helps reduce the amount of misinformation (proxied by citations) or increases the cost of retraction (proxied by authors' productivity). This paper sheds light on the empirical relevance of this media attention channel.

Studying the media attention channel is not straightforward. There are concerns related to the timing and the content of the coverage. Media attention may increase the salience of a study, leading to higher future citations (Phillips et al., 1991; Fanelli, 2013), or the media could

[1]See Hesselmann et al. (2017) for a review on scientific retractions.
[2]Guardian article.

exert a monitoring role, criticizing a study and updating the public about the shortcomings of a paper (Peng et al., 2022), potentially reducing future citations (see Whitely, 1994, for an example).[3] Therefore, to test whether media attention leads to differential scrutiny and differential punishment, I contrast research articles that appear in the media in a tight window around publication, to those that did not. These *early mentions* (i.e. appearances in either newspapers or blogs within two weeks from publication) are assumed to advertise the original research findings. This approach is suitable given that early media coverage is informative about later coverage (see Serghiou et al., 2021; Peng et al., 2022).

Furthermore, the endogeneity of media creates potential issues of selection into retraction. Some authors may focus on research questions which are more likely to attract media attention or manipulate their findings to attract more coverage. However, due to their relevance, these questions may receive differential scrutiny that, in turn, could impact their likelihood of retraction and their retraction timing. In addition, popular studies featuring in the media often involve eminent authors (e.g. Ivanova et al., 2013) who might be less impacted by retractions (Azoulay et al., 2017; Jin et al., 2019). I use a refined conditional difference-in-differences estimation strategy to address these issues (Heckman et al., 1997, 1998; Buscha et al., 2012). This approach compares the evolution of citations of retracted papers to that of closely matched control group papers, before and after retraction — while contrasting papers with or without *early mentions*. Control group papers are selected to mimic multiple ex-ante characteristics of retracted papers, those that best simulate the citation path of retracted papers absent the retraction. This approach enables a test of the common-trends assumption inherent to such a research design. Papers in the control group are (a) published in the same journal and year, (b) have similar citation trends in the years prior the retraction, and (c) have attracted comparable media coverage as their retracted counterpart at publication.[4] Another reason to match on *early mentions* is further justified by recent findings that retracted papers experience more coverage than non-retracted papers published in the same journal, similar publication year and author characteristics (see Peng et al., 2022).

[3]However, Serghiou et al. (2021) finds that retracted articles may receive high coverage, but pre-retraction coverage far outweighs post-retraction coverage.

[4]From a methodological point of view, Furman et al. (2012); Lu et al. (2013); Azoulay et al. (2015); Mongeon and Larivière (2016); Azoulay et al. (2017); Jin et al. (2019) all adopt similar difference-in-differences strategies to estimate the effect of a retraction shock, with the exception of the matching on early mentions.

Using the same matching strategy, I conduct a difference-in-differences estimation at the author level. This estimation compares the publication record of authors of retracted papers to that of authors of matched control papers before and after their first observed retraction — distinguishing cases where the original publication had *early mentions* or not. Crucially, this analysis allows to capture heterogeneous effects by authors' order of appearance in the coauthorship roster and by authors' seniority (i.e. based on H-index pre-retraction).

There could be other challenges to the validity of the estimates. For example, there could be unobservable paper- and author-specific factors that interact with time to confound these estimates. To allay such concerns, both paper- and author-level exercises incorporate a large set of additional control variables. All paper-level specifications include age, calendar-year, and paper-specific effects. All author-level specifications include career length, calendar-year and author-specific effects.

Despite this rich econometric framework, there may be residual concerns that the endogeneity of media attention — at the publication stage — may not be adequately accounted for. I employ two more strategies to tackle these concerns. First, using methods from computational linguistics, I build a predictive model of media coverage based on words in paper titles. I then use this prediction to separate (i) the impact of the arguably exogenous *excess* coverage on the likelihood of retraction, from (ii) the media coverage that a study may receive due to its authors self-selecting into a general interest topic. Second, I study the relationship between the average coverage of non-retracted articles within a journal-year (*journal visibility*) and the timing of detection of the retracted articles that appeared in those journals.

This paper draws on a comprehensive data collection effort to measure (i) the representation of research in broader media through *Altmetrics*, along with (ii) papers' citation history and (iii) authors' publication history. Media coverage data are recent, I therefore select papers that were published after 2010 and later retracted. Furthermore, I balance relevance and manageability of sample size by focusing on retractions published in highly ranked journals of each available discipline[5] and listed in the *RetractionWatch* database. Notice that publications in top journals receive more media coverage (Yin et al., 2022), which means that a considerable number of

[5]I retain retractions that either appear in *Scimago* top ten journals or in *Google scholar* top journals for each available category.

papers receive some coverage. This could introduce a bias as Woo and Walsh (2021) find that top journal citations to retracted articles are more likely to occur outside one's field and are potentially more prone to misinformation. I argue that investigating the relationship between media and punishment after retraction is more compelling for high-impact research. Yet, the average citation penalty for the retracted papers I select (≃65% of forward citations) is aligned with previous literature that does not specifically focus on high rank journals (Furman et al., 2012; Azoulay et al., 2015).

Using this sample, I first show descriptive evidence that retracted papers attract some media coverage, although publication and retraction events feature differently across outlets. Newspapers are more likely to cover the publication of a paper, while blogs are more likely to cover its retraction.[6] However, blogs have a more scientific target audience.[7]

Moreover, difference-in-differences estimates at the paper level show that retractions harm citations of retracted papers, and media coverage amplifies this effect (on average, media contributes to a ≃20-28% further reduction in forward citations). This aggravating effect is present only in hard sciences.[8] The media effect is also stronger for severe cases of misconduct.[9] To prove that this exacerbating effect of media on future citations is robust, I first show that media is not a proxy for other paper features. Indeed, papers with high cumulative citations (ex-ante) do not drive nor confound media estimates. Estimates are not sensitive to including or excluding the year of retraction in the *Post* indicator. Excluding non-actively cited papers or excluding more recent publications does not impact estimates. Finally, alternative strategies that look at *citation statements*[10] or study the relationship between citation penalty and *journal visibility*[11] confirm that media intensifies the citation penalty after retraction.

I propose two mechanisms that might explain the effect of media on citations: (a) higher scrutiny

[6]Visual differences emerge when looking at the leaning of (US) news coverage, based on the Gentzkow and Shapiro (2010) index of media slant, which deserves further investigation in future work.

[7]RetractionWatch alone represents about one third of all blog mentions for the selected retractions.

[8]The heterogeneity for hard versus social sciences suggests distinct publication practices may impact the visibility of a retraction. The difference across disciplines may also reflect the public perceptions that social science research is less "absolute" as the object of study is more volatile.

[9]The separation between severe and non-severe cases of misconduct comes from the classification developed by Woo and Walsh (2021).

[10]The *citation statements* are textual paragraphs where a citation appears.

[11]The *journal visibility* is calculated by averaging the media coverage of non-retracted articles published in a specific journal and year.

by the scientific community of a paper that gained popularity; (b) additional information provided to some part of the scientific community which would have otherwise remained unaware of the retraction. For (b), I check whether the textual content of post-retraction citations significantly differs with media coverage. The presumption is that absent the information mechanism, retracted papers with media coverage would experience fewer citations after retraction but with no relative change in their level of *accuracy*. In contrast, I find that with media coverage, new citations mention more often that the paper is retracted. This finding suggests that, with media coverage, scientists become more aware of a retraction and correctly acknowledge it when citing the original paper, reducing potential misinformation.

Consistently, difference-in-differences estimates at the author level show that retractions have a negative impact on authors' future *productivity*. However, this negative impact is large and permanent only if the original publication had some media exposure ($\simeq$10% larger reduction in future publication rate relative to a case with no media). This differential effect of media is evident for first authors, senior authors, and in cases of severe misconduct.

Furthermore, I show evidence that suggests there is some selection into retraction. On the one hand, papers that attract predictable (endogenous) media coverage are less likely to be retracted.[12] On the other hand, papers with exogenous *excess* coverage get retracted more often. However, both effects are modest. Finally, I show that journals that publish articles that are popular in the media are journals that retract faster (one standard deviation increase in journal visibility implies a reduction in the timing for detection of 15%).[13]

Overall, this paper investigates whether media coverage of scientific articles influences the auto-correcting process of science by observing media coverage of research articles across outlets and time. Media coverage — at the publication stage — amplifies the penalty for flawed research in terms of both future citations and authors' future publication rate. Although media coverage seems to help the auto-correcting process of science, this implies that (a) plenty of *wrong* science remains unnoticed and (b) that academia needs better strategies to raise the level of scrutiny and reduce incentives for poor-quality research.

[12] Words in titles such as climate, stem, meta-analysis, and trial are predictive of media around publication.

[13] In these journals, citation penalties for retracted papers are also sizable, which corroborates the main media effect on yearly citations.

The paper contributes to the large body of research on scientific retractions. It is closest to studies that estimate the causal effect of retraction on citations of retracted papers (Furman et al., 2012), on authors' previous publications (Lu et al., 2013; Azoulay et al., 2017; Jin et al., 2019), or their future research output (Mongeon and Larivière, 2016) and potential spillover to the related field (Azoulay et al., 2015). The main contribution is to show that media coverage — at publication — amplifies the causal effect of retraction on citations of retracted papers, and substantially explains the negative and persistent impact on the future research output of retracted authors. Further, I show that media attention may impact the likelihood of retraction and its timing. To address the risk of selection into treatment in this result, I improve the methodological approach by matching on early media exposure of papers. Only a few studies on retractions mention the role of media (Sugawara et al., 2017; Sarathchandra and McCright, 2017; Serghiou et al., 2021). Among these, this paper is closest to Peng et al. (2022) who use the similar data to show that retractions are ineffective at reducing online attention. They find that retracted papers receive more coverage after publication than non-retracted control papers from the same journals with similar publication years, number of coauthors, and authors' impact. The current paper addresses a complementary question of whether media attention intensifies the effect of retraction on papers' citations and authors' careers, thus reducing misinformation and increasing authors' cost of retraction. As the question differs, the matching strategy also differs. In fact, I compare retracted papers to never-retracted control papers from the same journal and year, with similar pre-trends in citations, and with similar *salience* at publication (i.e. similar early mentions).[14]

The paper also relates to the literature investigating the relationship between science and the media (Weingart, 1998; Phillips et al., 1991; Fanelli, 2013; Ivanova et al., 2013; Sumner et al., 2014; Dumas-Mallet et al., 2020; Ziegler, 2021) to which I contribute by showing that media coverage of subsequently retracted papers can influence the reputation of papers and authors, within science. This work further contributes to the literature on factors influencing citation rates (for example see Card and Dellavigna, 2020; Card et al., 2020; see also Tahamtan et al., 2016, for a review of the literature), to which I add that (ex-ante) salience impacts the citations of a paper and its authors' careers in case of a negative event (such as a retraction). Finally,

[14]This choice is motivated by the fact that media may impact selection into retraction. Therefore, control papers should be equally likely to be detected (if wrong).

this work relates to the broad literature on misinformation and how media channels influence politics and public policies (for an example see Allcott and Gentzkow, 2017; Lazer et al., 2018; see also Prat and Strömberg, 2013, for a review of the literature). I contribute to this literature by showing that media coverage attenuates misinformation within academia. At the same time, I illustrate that newspapers cover more the publication of a paper rather than its retraction, which creates the potential for disseminating misinformation to a larger audience.

In the remainder of the paper, Section 1.2 illustrates the institutional context of retractions. Section 1.3 describes the data, the sample selection and the main empirical strategy. Section 1.4 presents descriptive results on the media coverage of retraction. Section 1.5 and Section 1.6 provide a detailed presentation of results at the paper- and author-level respectively. Section 1.7 concludes.

1.2 Background

Understanding the incentives and governance regulating scientific knowledge production, dissemination and accumulation is crucial to this work. In what follows I discuss relevant aspects of the publication system.

One of the most discussed institutional setting is the peer review system. Articles are submitted and reviewed by independent experts before being accepted for publication. This feature is used to maintain high quality standards while allowing a suitable publication timing, even though practices vary greatly across disciplines and journals. This system eventually provides only limited guarantee against bad science.

Another aspect is the practice of citing related literature which is crucial for scientific communication. It allows to effectively contextualise a research article with respect to pre-existing literature while acknowledging original contributions from previous authors. Citations are regarded as an indicator of the importance of scientific findings and of their creators and can be negatively impacted by a retraction (Furman et al., 2012; Lu et al., 2013; Azoulay et al., 2015; Mongeon and Larivière, 2016; Azoulay et al., 2017; Jin et al., 2019).

In academic publishing, a retraction is the result of a procedure used by journals to alert readers that a published article should be removed from the literature. A retraction may occur when a major error (e.g. in the analysis or methods) invalidates the conclusions of the article, or in presence of misconduct (e.g. fabricated data, manipulated images, plagiarism, duplicate publication, research without required ethical approvals etc). It differs from a correction issued in case of an error or omission which can impact the interpretation of the article, but where the scholarly integrity remains intact. The surge in the absolute number of retractions across all disciplines has alarmed many in the scientific community (see Figure 1.1). Nonetheless, retractions remain relatively rare involving 4 in every $10,000$ published papers of which 60% due to some type of misconduct, though both rates have been rising steadily over time (Brainard, 2018).

A retraction can be initiated by the editors of a journal, by some or all the authors or their institution and are typically complemented by a notice meant to clarify the reason of such decision. But, the information contained in notices vary significantly, some explain the details which lead to the retraction outcome and inform on whether an article results and conclusion should be disregarded entirely or in part, others are rather succit and vague.

A further element of discussion is therefore the visibility and accessibility of both retractions and notices. "Authors are responsible for checking that none of the references cite retracted articles except in the context of referring to the retraction" (International Committee of Medical Journal Editors 2019). Awareness of readers is therefore decisive, yet the current institutional setting does not suffice. Retractions are usually published and linked to the original publication and can be often identified via different sources (e.g. libraries, databases and search engines) but inaccurate citations still remain. Schneider et al. (2020) found that in the case of an infamous clinical trial (Matsuyama et al., 2005), in which data were falsified leading to a retraction in 2008, the retraction is not mentioned by 96% of post-retraction citations and 41% of these inaccurate citations describe the paper in detail leading to possible disinformation. On the other hand, Piller (2021) looked at the recent case of high-profile Covid-19 retraction (Mehra et al., 2020) and finds that 52.5% of the citations do not correctly mention the paper status. In what follows I will illustrate that the media attention attracted by the latter case could be a relevant factor behind the difference in the two examples just discussed.

In this respect, recent efforts make use of media platforms to alert scientists of retractions, as in the case of the specialised blog *RetractionWatch* which reports on retractions and gathers information surronding specific retraction events, such as which of the authors is responsible for the article ultimate fate. Information which is usually hard to acquire based on the notice alone. New tools are also emerging as in the case of *Scite.ai*, a recently launched platform which categorises references, monitors retracted papers by searching through Crossref, PubMed, and the RetractionWatch database, and flags both citing and retracted papers on Twitter.

1.3 Data and method

1.3.1 Data

This study combines multiple data sources on scientific publications and their authors which I hereby list in details.

Retractions. The treatment sample is extracted from the *RetractionWatch* database[15] which Brainard (2018) defines as "the largest-ever database of retracted articles". The dataset contains a list of retracted research articles[16] together with the following information: title, doi, date of publication, date of retraction, journal, name of authors and their institutions, list of reasons for retraction, and when available, a link to the associated blog post reporting on the paper background story.

Journal ranking. I further select papers featuring in either *Scimago* or *Google scholar* rankings. The selected journals appear either as one of the ten highest ranked in Scimago in any of the available subjects or among those listed in Google scholar top publications in any of the existing categories.

Citations and authors' publications. Yearly citations and authors publications are the main outcome of this study and are collected for each article and author using *Scopus*, one of the two largest abstract and citation database of peer-reviewed literature.[17]

[15] Version obtained in July 2020.

[16] Dense since the '80s.

[17] For the period considered in the analysis, there exists little difference between Scoups and WoS in terms of coverage (see: Scopus vs. WoS). Scopus though has the advantage of having an API easily accessible via *rscopus*, a library by John Muschelli available on R.

Media coverage. Data on online mentions are gathered thanks to *Altmetric*, a company that since 2011 tracks the attention that research outputs receive online.[18] For each paper I retrieve the *Altscore*, an aggregate measure of online mentions (i.e. it combines all mentions across outlets giving a higher weight to outlets such as newspapers, see appendix Table A.1), and details about single mentions (e.g. date, url, author, title, summary).

Citation textual content. I obtained data on the content of citation statements quoting the research articles in the sample with the support of *Scite.ai*, a recently launched start-up that uses text analysis to categorize reference statements. For each pair of citing and cited study, statements are categorised as "mentioning", "contrasting" and "supportive".[19] In addition, access is gathered for any statement containing the words "*etract*" or "*ithdraw*".[20]

1.3.2 Empirical strategy

This work investigates the possibility that media coverage influences scientists' awareness and assessment of research findings (looking at citations of retracted papers) and authors' careers (looking at the publication rate of retracted authors, see Section 1.6 for author-level analysis). Holding other factors constant, a loss in citations and lower authors rate of publication, reflects an erosion of trust in the authors' work by the scientific community.

To understand the interplay between the retraction of a paper and the information available online one needs to consider how scientific publications feature in the media and what challenges this poses in terms of identification.

A research article that is accepted for publication may endogenously attract media coverage. Online attention may depend on factors such as the salience of a topic, the importance of the findings, the prestige of authors and publishers, the presence of a press release (Sumner et al., 2014, 2016). Media coverage can therefore bring publicity to a paper increasing future citations (Phillips et al., 1991; Fanelli, 2013) as well as prompt higher scrutiny from the scientific community making any fault more likely to emerge. Online attention can finally inform about

[18] I here focus on sources with the highest number of mentions (i.e. newspapers, blogs and Twitter) though Altmetric collects mentions from numerous additional outlets (e.g. Pubpeer, Wikipedia).

[19] According to Rosati (2021)

[20] Manually checked to exclude any false positive.

the fate of an article as in the case of an expression of concern or a retraction (Serghiou et al., 2021), information that could reach unaware scientists that would otherwise incorrectly cite a flawed article.

Therefore media endogeneity, together with observables and unobservable characteristics of papers and authors can create issues of selection into retraction (treatment). I tackle this challenge using a conditional difference-in-differences strategy (Heckman et al., 1997, 1998; Buscha et al., 2012) which compares retracted papers to matched (never-retracted) controls, before and after the retraction, while contrasting papers with or without *early mentions* (i.e. apperances in either newspapers or blogs within two weeks from publication). Smith and Todd (2005) shows that the difference-in-differences matching estimator performs the best among nonexperimental matching based estimators.

Early mentions are the preferred measure of media coverage as this facilitates identification. These mentions are assumed to broadly advertise the original research findings and are therefore virtually independent from the retraction.

Control papers are choosen to mimic multiple (ex-ante) characteristics of retracted papers, such that they could best simulate the citation path of retracted papers absent the retraction. Specifically, controls are (a) published in the same journal and year, (b) have similar citation trends in the years prior the retraction, and (c) attracted comparable media coverage at publication, as their retracted counterpart.

The rest of this session explains in details the process determining the sample of treated papers, the matching strategy employed to choose control papers, and the main regression model.

1.3.3 Treatment group

The full RetractionWatch database counts $N_r = 21,968$ retractions starting from 1980. Provided that data availability on online mentions is only relatively recent, I select retracted papers both published and eventually retracted after 2010 ($N_r = 11,258$). Only research articles[21] with non missing paper DOI and retraction notice DOI are maintained.[22] To balance relevance and

[21]Excluding for examples: conference abstracts and clinical studies, $N_r = 6,676$.
[22]$N_r = 6,189$.

manageability of sample size, I focus on articles published in journals featuring in either Google scholar top journals by field or among the ten highest ranked journals in Scimago per subject category.[23] Remaining papers are certainly relevant for the scientific community, hence, it is important to study whether in this case disinformation is halted or fostered by media coverage.[24] In addition, publications in reputable journals may be more likely to attract media coverage (Yin et al., 2022), thus helping identification.[25] Next, I exclude articles for which I cannot find any author with at least one publication in the 9 years before the retraction event.[26] Of these I retain cases for which I can find an appropriate control, leading to a final sample of $N_r = 990$.

1.3.4 Control group

Trends in citations vary across disciplines, age and media coverage, hence, control publications were selected to mimic pre-retraction characteristics of the treated. This strategy draws from the approach first used in the literature on retractions by Furman et al. (2012) and further developed by Lu et al. (2013) and Jin et al. (2019). The main assumption is that treated papers would continue to perform similarly to control ones in absence of a retraction event.

The selection of the control group proceeds in steps. For each retracted paper I search in Scopus for studies[27] published in the same journal and year of the treated.[28] For each retracted i and potential control pair j I compute the measures listed below.

- *Absolute arithmetic distance* in citations.

$$|AD| = \left| \sum_{t=pub}^{retr-1} (c_{it} - c_{jt}) \right| ;$$

[23] $N_r = 1,163$.

[24] Woo and Walsh (2021) find that top journal citations to retracted articles are more likely to occur outside one's field and are potentially more prone to disinformation.

[25] This could potentially introduce a bias. Yet, the reader may be reassured that I observe an average penalty for selected papers which is alligned with previous literature (Furman et al., 2012; Azoulay et al., 2015).

[26] This is important to study the career impact for retracted authors', hence I need at least some authors with a minimum reputation ex-ante ($N_r = 1,008$).

[27] Articles or reviews.

[28] $N_c = 586.281$ overall results.

- *Euclidean distance* in citations.

$$ED = \left[\sum_{t=pub}^{retr-1} (c_{it} - c_{jt})^2\right]^{1/2} ;$$

where c_i indicate the citations paper i receives in year t in the time span between the year of publication *pub* and the year of retraction $retr$. These measures capture the disparity in citation trends in different ways. AD allows for positive and negative yearly differences to balance over time while any discrepancy is accumulated over time in the case of ED.

- *Early mentions absolute distance* (MD) of blog b and newspaper n mentions whitin two weeks from publication.

$$MD_b = |(b_{i,2w} - b_{j,2w})| \quad \& \quad MD_n = |(n_{i,2w} - n_{j,2w})|$$

The reason for choosing a cutoff close to the day of publication draws from observing that notable studies attract most online publicity around the publication date as suggested by Figure 1.3 for treated papers and more evidently by Figure A.4 for control papers. Matching media mentions becomes more and more challenging the further away from publication as flawed articles may later prompt additional critical mentions.[29] To capture whether a paper is newsworthy without including mentions related to its misfortune, I focus on a two weeks cutoff from publication date. This threshold is also less sensitive to imprecisions in the publication date compared to a shorter cutoff.

I then retain for each i all j with $|AD| \leq 10$; $MD_b \leq 10$; and $MD_n \leq 10$. These cut offs allow to maximise the number of matches while limiting the maximum conceded distance in either citations or media mentions. These thresholds lie at the extremes of the distribution of distances and improve the quality of matches without affecting results. I rank the remaining j in terms of smallest $MD_b + MD_n$ and select two controls (or one depending on availability) with the minimum ED among those. This final selection leads to a sample of $N_c = 1,969$ control articles.

[29] Note that for either news or blogs the bulk of mentions appears in week one, grows at a progressively smaller rate in week two and three, and flattens out afterwards.

The quality of selected controls is assessed in Figure A.1 and A.2 of the appendix. The Euclidean distance between the selected controls and the treated paper is dense around zero (in over 68% of the cases this selection yields a perfect match), and the arithmetic distance is fairly centred around zero. No significant difference emerges when comparing treated and control distributions of cumulative citations pre-retraction. Similarly, there is no significant difference in the distribution of early mentions across treated and control groups for either newspaper articles or blog posts.[30] Control papers are marginally more likely to have little citations and no media mentions pre-retraction. In general, the vast majority of published articles have little citations and no mentions in either media outlets at publication.

1.3.5 Selected summary statistics

Table 1.1 illustrates a set of distinct summary statistics for treatment and control group. The top of the table looks at variables which should be similar across the two groups for the identification strategy to be successful. ED and AD are on average somewhat close to zero (0.93 and 0.17 respectively) and both groups of papers attracted an average of about 7 citations in the pre-retraction period, substantially confirming the finding reported in Figures A.1 and A.2. Within two weeks from publication papers experience comparable online mentions on newspapers and blogs, even though eventually retracted papers have on average moderately higher coverage (1.04 vs. 0.79 news articles, and 0.24 vs 0.15 blog posts). The age for the two groups of papers is almost identical by construction. Moving to the bottom of the table one can observe that papers take on average two years to be retracted. Furthermore, yearly citations have a distribution that is very skewed, with 32.2% observations actually equal to 0, a Poisson model would therefore better approximate the distribution of the dependent variable. Unsurprisingly, treated papers cumulate substantially less citations over the years as compared to controls (16.8 vs. 33.9 respectively), but attract generally higher online attention with an *Altscore* of 37.5 for retracted papers and 19.1 for controls. In general, a non negligible share of articles experiences some online coverage, most articles are mentioned on social media (60% of retracted papers and 44% of controls) while only a limited fraction appears in newspaper articles (13% and 12% respectively), in addition blogs actively mention over one third of retracted papers while significantly less attention is devoted to controls. Finally, about one tenth of papers in either group appears in either newspapers or

[30] This remains true when removing observation with no mentions, as shown in Figure A.3.

blogs around the publication date.

1.3.6 Estimating specification

The study employs a difference-in-differences strategy that allows to compare the evolution of citations of retracted papers before and after retraction relative to citations of a control group of non-retracted studies published in the same journal and year and with a comparable trend in yearly citations before retraction. Treatment and control papers also have similar number of online mentions (on blogs and newspapers) within two weeks from the day of publication (i.e. *early mentions*) to account for unobservable characteristics which make a study newsworthy and could therefore create a problem of selection into retraction.

Therefore, the regression model is the following:

$$E[Y_{igt}|X_{igt}] = exp[\alpha + \gamma_1 Post_{igt} + \beta_1 R_i * Post_{igt} + \beta_2 Post_{igt} * Media_i + \\ + \beta_3 R_i * Post_{igt} * Media_i + \delta_i + f(age_{it}) + \delta_\tau] \tag{1}$$

where i is the treatment (or control) paper, g is the case-level group and includes the retracted paper and its respective controls, t are years relative to the retraction. The dependent variable Y represent a paper yearly citation count and exclude self-citations, as the estimation wants to capture the reaction of the scientific community other than that of the authors involved. $Post$ is an indicator variable equal to one for all years after retraction, R is an indicator for retracted articles, and $Media$ captures the exposure of an article to online coverage. Different media dummies will be used to indicate articles with or without media mentions. Due to previously discussed issues related to the timing and the content of coverage, the media indicator which is best identified is equal to one if a paper receives at least one online mention within two weeks from publication in either newspapers or blogs and zero otherwise (i.e. $\mathbb{1}[Early\ Mentions > 0]$). Early mentions are assumed to broadly advertise the original research findings and are generally balanced across treatment and control papers. Other media indicators equal one for research papers that receive at least one overall mention in any of the media outlet analysed (i.e. socialmedia, newspaper articles or blog posts). In order to look at different levels of media exposure of each paper, indicators are also derived from the distribution of *Altscore*, an aggregate measure of weighted online mentions.[31] The coefficient β_1 captures the effect of a retraction shock on citations of retracted papers as compared to similar control papers. The coefficient β_3 captures any difference in the effect of the shock for papers that received online attention. Fixed effects are included for each paper δ_i and each calendar year δ_τ while $f(age_{it})$ represents a full set of dummies for years since publication (age) and is meant to flexibly control for the age of the articles.[32] To look at the dynamics of the differential effect of *Media*, estimates will be presented for a model that replace the indicator *Post* with a full set of dummies for each year relative to the year of retraction.[33] Given the skewed nature of the dependent variable, I follow a long-standing tradition in bibliometric studies, hence I use a pseudo Poisson regression model developed by Correia et al. (2020)[34] where consistency is achieved under the only assumption that the conditional mean of the dependent variable is correctly specified (Gourieroux et al., 1984). Finally, standard errors are clustered at the case g level.

1.4 Descriptives results: coverage of retractions

Popular online media like newspapers, blogs and social media, whose target is often beyond the scientific community, have been recently active in advertising retracted articles (see Figure 1.2).[35] In general, media platforms seem to cover both original publications and retractions, but the two events feature to a different extent across outlets, giving raise to potential disinformation. Indeed, Figure 1.3 shows that mentions in newspaper articles appear predominantly close to the publication date of a study and generally inform the public about its discovery, less often this information is updated with a new mention at the time of retraction. On the other hand, blog posts occur mostly around the retraction event. These blogs are often specialized and directly target academics[36] while a wider audience is exposed to information which is not always complete. This could lead to unintended consequences that deserve further work.

[31] See Table A.1 for details about *Altscore* weights across outlets.

[32] Note that the interaction term $R_i * Media_i$ is absorbed by the paper fixed effect.

[33] $E[Y_{igt}|Xigt] = exp[\sum_{t=r-4}^{r-2} \gamma_{1t} * d_t + \sum_{t=r}^{r+6} \gamma_{1t} * d_t + \sum_{t=r-4}^{r-2} \beta_{1t} * d_t * R_i + \sum_{t=r}^{r+6} \beta_{1t} * d_t * R_i + \sum_{t=r-4}^{r-2} \beta_{2t} * d_t * Media_i + \sum_{t=r}^{r+6} \beta_{2t} * d_t * Media_i + \sum_{t=r-4}^{r-2} \beta_{3t} * d_t * R_i * Media_i + \sum_{t=r}^{r+6} \beta_{3t} * d_t * R_i * Media_i + \delta_i + f(age_{it}) + \delta_\tau]$

[34] http://scorreia.com/software/ppmlhdfe/

[35] Notice that recent years are likely underreported given retractions take some time to arise and hence feature in the database.

[36] Around a third of blog coverage is from RetractionWatch, the single outlet most committed to inform about scientific retractions (Figure A.5 exclude RetractionWatch mentions).

To shed some light into factors that could shape an outlet decision to acover a retraction event or not, I look at US news coverage classified based on Gentzkow and Shapiro (2010) measure of media slant.[37] Figures A.6 to A.9 contrast the observed mentions for relatively left- or right-leaning outlets. Limited differences seem to emerge as left leaning news show a somewhat more balanced reporting which deserve to be further studied.

In essence, the rise of the internet and the appereance of new platforms has the potential to direct scientists (and non-scientists) attention towards "interesting" contributions which in some cases prove to be less reliable (Serra-Garcia and Gneezy, 2021). It is therefore important to investigate whether positive remaining citations, and the retraction process more in general, relate to the media visibility of a research paper and its retraction.

1.5 Paper-level results

Table A.3 shows results for a simple difference-in-differences analysis for the pooled sample of retracted papers and selected controls. Estimates imply that relative to controls, retracted papers experience a 65% (i.e. $1 - exp(-1.06) = 0.65$) loss in yearly citations after the shock and the magnitude is comparable to previous studies (Furman et al., 2012; and Azoulay et al., 2015) which rely on different samples, disciplines and time periods.[38] Figure A.10 illustrates the dynamic of the effect of a retraction. The post-retraction loss in citations increases over time and there is no evidence of pre-trends.[39]

1.5.1 Main results

Table 1.5 to A.5 report results from the main specification. The tables differ by measures of media coverage, using indicators for papers with at least one mention within two weeks from publication (early mentions), papers with at least one mention overall in a certain online outlet (any news, blog or social media) or papers that fall in some part of the Altmetric score (Altscore) distribution. Tables highlight the difference-in-differences coefficient $Post * Treatment$, accord-

[37] To maximise the number of observable US press mentions, I take the sample of research articles published and retracted after 2010 with non missing DOI and whose DOI is different from that of its retraction notice ($N_r = 4,763$) and retain only mentions matching the list of outlets classified by Gentzkow and Shapiro (2010). I remain with 53 retracted papers for which I observe at least one news mention with measurable slant.

[38] Estimates are similar when using an IHS (Inverse hyperbolic sine) transformation of the dependent variable.

[39] Note that effects in the year of retraction are also small due to the fact that papers in the sample get retracted at different points within the year.

ing to which the average citation penalty of a paper after its retraction amount to 55-62% across all specifications. The relative effect for papers that experienced some media coverage is estimated by the coefficient of the triple interaction $Post * Treatment * Media$.[40] Retracted papers with media coverage experience a penalty in post-citations of about 75% (i.e. $1 - exp(-0.96 - 0.45) = 0.76$). Across specifications in Table 1.5 and Table A.4 the loss in forward citations for retracted papers with media varies between 68-76%, corresponding to a difference of 12.3-15.8 p.p. (or 19.7-28.7%) with respect to retracted papers without media exposure. Furthermore, the effect seems monotonically increasing in the amount of coverage received (see Table A.5). The almost entirety of these estimates is highly significant. Figure 1.4 represents the dynamics of the additional penalty in presence of (alternative measures of) media coverage. The loss in yearly citations becomes progressively more evident over time without any sign of recovery, and I find no evidence of pre-trends.

1.5.2 Robustness checks

Highly cited papers differencial

In what follow I intend to increase confidence that the exacerbating effect I showed, is solely due to the presence of online attention. It could be that media exposure is actually capturing some alternative paper features, related to media presence, and confund my estimates. To address this concern I repeat the main excersise looking at whether the effect of retraction on forward citations differs for papers which are highly cited *ex-ante*. The rational being that influencial papers may face higher scrutiny as well as higher chances of featuring in the media (Yin et al., 2022). Reassuringly, Tables A.6 to A.8 show that papers with high cumulative citations before the year of retraction do not drive nor confound media estimates.

Including retraction year into *Post* indicator

Previous estimates illustrate effects on citations for all years strictly after the one of retraction (i.e. excluding the year of retraction). The rationale behind this choice is the fact that papers can get retracted at any point during the year and this can therefore act as a confounder.[41]

[40] Where the *Media* variable is defined in alternative ways across specifications as described at the top of this paragraph.

[41] Figure A.10 and Figure 1.4 show smaller or insignifican effects in the year of retraction relative to the previous year.

Nonetheless, Tables A.9 to A.11 show that the main results are not sensitive to this decision. If anything, the additional effect of early mentions is smaller in case of blog mentions (see Table A.9 column (2)). This difference may relate to the fact that most blog mentions appear later when the paper gets discredited. In addition, the fact that effect of early mentions are less significant, may speak to a possible information effect of media which emerges more clearly at a later stage, as captured by overall measures of online coverge in A.10 and A.11.

Actively cited papers

The algorithm for selecting controls attempts to choose papers that could likely mimic the citation path of retracted papers absent the retraction shock. Finding good controls for retracted papers that are not actively cited soon after publication may be challenging and could bias estimates. For this reason I here exclude all retracted papers (and respective controls) with zero citations in any year before retraction. This exercise halves the original sample.[42] Even so, Tables A.12 to A.14 confirm the results all remain robust.

Excluding late published papers

One concern is that for more recently published papers there may not be the sufficient time frame to display changes in citation patterns. To this respect I repeat the exercise retaining only older publications. Specifically, I retain only retracted papers (and associated controls) that were published between 2011 and 2017. Tables A.15 to A.17 confirm the results remain virtually unchanged.

Citation textual content

One additional exercise is that of looking directly at the textual content of citations. *Scite.ai* (a newly launched platform featuring in Nature)[43] scans article PDFs for references to papers and categorises these references as mentioning, contrasting or supporting.[44] With the platform support, I built a dataset of yearly citation statements for each classification, paper and year and performed an exercise equivalent to that of Section 1.5.1. Tables A.29 to A.31 substantially corroborate the main findings. Retracted papers experience a penalty in all type of citation

[42] 48% observations left in either treated or control group.
[43] Nature article on Scite.ai.
[44] The classification is according to Rosati (2021)

statements after the retraction shock, and for citation statements that only mention the study, this penalty is aggravated in presence of media coverage. No additional change is detected for either contrasting or supporting references. One caveat is that almost the entirety of the citation statements is classified as mentioning.

Journal visibility and loss in citation

In a last robustness exercise I relate the individual loss in citations obtained comparing each retracted paper to its selected controls for different pre- and post- time windows ($DiD = [E(cit_1^T) - E(cit_1^C)] - [E(cit_0^T) - E(cit_0^C)]$) to the average media coverage of (non-retracted) articles published in the same journal and year (*journal visibility*).[45] Figure A.22 shows that a retracted paper experiences a significantly larger loss in citations if published in a journal with higher average visibility. Notice this negative relationship becomes stronger when looking at wider time windows around the year of retraction. These same conclusions are evident in Table A.28 were alternative measures of media exposure are also used. These findings substantially confirm the main results presented in section 1.5.1.

1.5.3 Heterogeneity

Hard vs. Social sciences

Various disciplines have distinctive publication practices which could create different incentives at publication and therefore lead to heterogeneous effect. Table 1.6 and A.18 together with Figure 1.5 illustrate that this may indeed be the case. What consistently emerges across specifications is that, in the case of social sciences,[46] there is no additional penalty associated to retracted papers with media attention. Perhaps one interpretation is that the timing of publications in hard sciences is generally fast; while working papers in social sciences may circulate for longer inside the scientific community. In the latter case, media coverage may therefore offer little room for update on the validity of the study as compared to the former case where online

[45] Controlling for year of publication effects, age of paper at retraction effects, number of non-retracted articles within same journal and year, the average Euclidean distance of those non-retracted articles, and the level of (non-self) cumulative citations of the retracted paper before retraction. See Section 1.5.6 for further details.

[46] Disciplines are identified using Scopus journal classification. Social sciences are: business and technology, humanities and other; while hard sciences are: life sciences, environment, health and physical sciences.

attention may further stimulate the academic discussion around a paper.[47] One caveat is that the subsamples of disciplines are quite small in particular in the case of social sciences.[48]

Cause of retraction

One aspect I investigate is whether the aggravating effect of online attention differs depending on the reason behind the retraction. Retractions can occur because of honest mistakes or actual misconduct of the authors. Distinguishing the two is relevant as original findings should be entirely discarded in cases of severe misconduct, leading to higher concerns over the spread of disinformation. On the other hand, cases of misconduct may be newsworthy and online discussion may play a special role by circulating detailed information on the case. I therefore divide reasons for retraction in minor, moderate and severe cases of misconduct using the classification developped by Woo and Walsh (2021) (see Table A.20).[49] Table 1.7 (and A.19) suggests media attention plays a bigger role in presence of severe cases of misconduct and there is little additional penalty associated to minor cases with media attention.

1.5.4 Information mechanism

This works has so far shown that retractions disappear from the literature at a faster pace in presence of media coverage. This additional effect of media may be derived by two different mechanisms: (a) higher scrutiny by the scientific community to a paper that gained publicity; (b) additional information provided to some part of the scientific community which would have otherwise remained unaware of the retraction. Although difficult to distinguish, one way to corroborate the information mechanism is to check whether the content of remaining ex-post citations is more "accurate" in presence of media coverage. With the help of *Scite.ai*, I collected for each retracted paper all yearly citation statements that mentioned the retraction. Citation statements were searched for the terms "*etract*" or "*ithdraw*", manually excluding false

[47] Related to this, Wohlrabe and Bürgi (2021) suggests that in the case of economics, the practice of releasing working papers before their publication in a journal has a positive impact on citations.

[48] Over 80% of retraction appears in hard sciences (809) of which 12% (95) with early visibility and 59% (475) with Altscore above median. Of the 179 retraction in social sciences 8% (15) have early visibility and 53% (95) have Altscore above median.

[49] The selected sample is divided in 30% (301) cases of minor misconduct, 26% (261) of moderate misconduct and 43% (428) of severe misconduct.

positives. I then estimate the following regression model:

$$E[Y_{it}|X_{it}] = exp[\alpha + \beta_1 Post_{it} * Media_i + \delta_i + \delta_t + f(age_{it}) + \delta_\tau] \tag{2}$$

where for each retracted paper i and year relative to retraction t, Y represents the number of citation statements mentioning the paper is retracted, $Post$ is an indicator for year strictly after retraction, $Media$ is an indicator for whether a paper gained some kind of online coverage. Estimates of β_1 capture the differential change in number of citations "correctly" mentioning the retraction (after the shock) in presence of media coverage. Fixed effects are included for each paper δ_i, each year relative to retraction δ_t, each year since publication $f(age_{it})$ and each calendar year δ_τ. Standard errors are clustered at the retraction level.

Table 1.9 confirms that the number of references correctly mentioning the cited paper is retracted increases significantly in presence of media coverage. This result support the hypothesis that media coverage provides additional *information* on retractions, hence favouring the belief update of part of the scientific community which would have otherwise remained unaware. One caveat to consider is the small sample of retractions for which an "accurate" yearly-citation is indeed observed (slightly less than 10% of the treated sample).[50]

1.5.5 Media and likelihood of retraction

In section 1.3.2 I argued that a challenge one faces when trying to understand the interaction between the retraction process and media coverage, arises from the endogeneity of the latter. To circumvent this issue to some extent and study the relationship between media and likelihood of retraction, I turn to the text analysis of titles of research articles. This in turn allows me to use the presence of specific words to control for papers' endogenous coverage.

More specifically, I start with the full sample of eventually retracted articles published (and retracted) after 2010 and for each of these articles I add to the sample twenty randomly selected articles that appear in the same journal and year but were never retracted.[51] I then use

[50] This is consistent with previous work by Schneider et al. (2020) which finds that, for the case considered, the retraction is not mentioned in 96% of direct post-retraction citations.

[51] This selection facilitate a speedy computation without restricting the corpus of titles. Among the 1008 retracted papers in the sample, 44 have less than 20 associated random controls due to the respective scarsity of

the titles of these papers as corpus of analysis.[52] After cleaning the text according to Porter (1980) algorithm, Figure 1.6 shows the most frequent words present in the titles of papers that experience some (Panel A) or no (Panel B) online coverage (in newspapers or blogs) within two weeks from publication. On the one hand, popular papers mention more often words shuch as "cancer", "patient" and "disease", on the other, articles that did not feature in the media often quote different words such as "model" or "system". In what follow I try using this differences to predict articles coverage.

After building the document-term matrix of words (unigrams and bigrams) that appear in at least 100 titles I randomly split the observations into 90% training and 10% testing subsample. The training sample is used to select words with some predictive power for papers' media coverage based on lasso selection procedure. The testing sample is then used to compute the out-of-sample performance of the predicted media coverage based on the selection.[53]

The lasso estimates and the set of selected variables (words) depends on the penalty level λ. I obtained alternative lists of selected words using different procedures that choose the optimal penalty level using: (a) EBIC information criteria; (b) AICC information criteria; (c) K-fold cross-validation and (d) Rigourous (theory-driven) penalty levels. These procedures are then repeated including a full set of subject fixed effects, publication year fixed effects and excluding retracted articles from the sample. This strategy allows to estimate the following model:

$$Retraction_{ijp} = \beta_1 Media_{ijp} + \beta_2 \widehat{Media}_{ijp} + \delta_j + \delta_p + \epsilon_{ijp} \qquad (3)$$

where for each article i published in year p and journal j, $Retraction$ is an indicator for whether the article was retracted, $Media$ is a dummy taking value one if the article gained any online coverage (in either newspapers or blogs) within the first two weeks from publication, while δ_j and δ_p absorb journal fixed effects and publication year fixed effects respectively. Estimating the $Media$

potential controls found in Scopus.

[52] $N = 20755$

[53] The lasso estimation minimizes the mean squared error subject to a penalty on the absolute size of coefficient estimates and where λ controls the overall penalty level.

$$\hat{\beta}_{lasso} = argmin \frac{1}{n} \sum_{i=1}^{n} (Media_i - \sum_{j=1}^{p} \beta_j Word_{ij})^2 + \frac{\lambda}{n} \sum_{j=1}^{p} \psi_j \mid \beta_j \mid$$

Due to the nature of the penalty, the lasso sets some coefficients exactly to zero and in doing so removers some predictors from the model.

impact on the likelihood of a retraction (β_1) is challenging as it is difficult to exclude that researchers may choose to investigate salients topics that, given their relevance, are scrutinized differently from the scientific community (see for example Serra-Garcia and Gneezy, 2021) leading to different retractions rates, despite the fact that these topics may be of interest to the general public and hence attract media coverage. The inclusion of $\widehat{Media} = \sum_s \hat{\beta}_{s,lasso} SelectedWord_s$ as predicted from the lasso procedure, where $SelectedWord$ represents the number of times a selected n-gram appears in the title of a paper i, allows to control for endogenous topic selection that could otherwise lead to bias. Given that $\widehat{Media}$ is derived from separate estimates, standard errors are bootstrapped and clustered around retraction cases.[54]

Table A.22 shows the correlation between some of the most powerful lasso selected predictors and the $Media$ indicator variable. The n-grams with the largest coefficients provide insights into which articles receive media coverage. For example, the word "climate" appears. Similarly, the n-grams "brain", "graphen", "genom" and "stem" all represent research topics of large interest. Also, some research methodologies seem popular as suggested from the n-gram "meta analysis" and "trial". Accuracy ranges between 60 and 76% across procedures and more parsimonious lasso (and logit lasso) seem to provide better performing selections. The fraction of correctly classified observations reaches up to 86% when a full set of subject and year fixed effects are included and when retracted papers are excluded.[55] Accuracy is calculated after estimating the optimal positive cutoff threshold using the Matthews Correlation Coefficient (MCC). In the area of machine learning with binary classification the MCC is the preferred single metric, especially for imbalanced data (Chicco and Jurman, 2020). The metric ranges $[-1, 1]$ and takes on the value of zero if the prediction is the same as a random guess. Table A.22 shows MCC ranging between 0.37 and 0.45 across different selection procedure.

Equation (3) estimates are reported in Table 1.8 (Panel A) where despite the differences in n-gram selection and predictive accuracy across models, very similar results emerge across specifications. Evidence suggests that articles with higher *predicted* media coverage are less likely to experience a retraction. The interpretation of this result is twofold. On the one hand, the fact that popular articles are retracted less often seems reassuring and could be due to experienced

[54] Summary statistics of main variables and a selection of n-grams are displayed in Table A.21.

[55] The most powerfull predictors selected with these alternative strategies remain fairly similar (not shown and available upon request).

academics answering salient research questions.[56] On the other, it could indicate that "interesting" research articles may be reviewed with a laxer standard (as suggested in Serra-Garcia and Gneezy, 2021). Under the assumption that predicted media coverage effectively controls for endogenous topic selection, the remaining variation in media coverage is arguably exogenous and therefore allows to estimate the impact of additional attention on the likelihood of being retracted. Estimates show that wider media coverage at publication leads to higher chances of retraction, but the magnitude of this effect remain small. Note that the media variables (observed or predicted) capture very limited variation in the outcome variable. Equivalent results are displayed in Table A.23 for logit estimations, in Table A.24 for direct estimates of residual coverage, in Table A.25 for lasso procedures trained with TF-IDF word scores, and in Table A.26 for lasso procedures trained within subjects and years and excluding retracted articles.[57]

These findings justify selecting controls with early media presence similar to that of their retracted counterpart as allowing the however small selection into treatment of more popular articles could otherwise bias the main results reported in section 1.5.1. Finally, one could be concerned about the common inclusion of both the media indicator and its text-based prediction due to their positive correlation ($\rho \approx 0.3$). To this respect, Table 1.8 additionally reports the impact of the two regressors separately (see Panel A column (1-2) and Panel B respectively), the magnitudes of coefficients varies only slightly in this case, ressuring us against a collinearity issue.

1.5.6 Journal visibility and retraction timing

In the following section I offer one way to circumvent media endogeneity and study the relationship between coverage and timing of retractions. In what follows I argue that non-retracted articles published in the same journal and year as a retracted one, attract online coverage which is arguably *exogenous* to the retracted article own coverage. Based on this, a good proxy for online *visibility* of a specific journal and year is the average coverage of all non-retracted papers

[56] Notice that predicted coverage is endogenous.

[57] Non reported estimates reveal equivalent results when selecting 50 or 100 random controls per retracted paper.

published in there.[58]

$$JVisibility_{jp} = \frac{1}{n}\sum_{k \neq i} Altscore_{kjp} \tag{4}$$

where k are non-retracted papers published in same journal j year p as the retracted paper i. Alternatively I use the average share of $k \neq i$ published in j and p with some media mentions. Hence, I can study the following relationship using an OLS regression in a cross-sectional context:

$$Y_{ijp} = \beta JVisibility_{jp} + \delta_p + \nu X_{ijp} + \epsilon_{ijp} \tag{5}$$

where for each retracted paper i published in j in year p, Y represents either one of the dependent variables: $Time\ to\ retract = (Retraction\ date - Publication\ date) \times \frac{12}{365}$ or $DiD = [E(cit_1^T) - E(cit_1^C)] - [E(cit_0^T) - E(cit_0^C)]$ the individual loss in citations obtained comparing each retracted paper to its selected controls for different pre- and post- time windows (see Section 1.5.2 for results on loss in citations). In addition, δ_p indicate publication year fixed effects, while X_{ijp} control for $N_{jpk \neq i}$ number of non-retracted papers in same journal-year as retrieved from Scopus, $\frac{1}{n}\sum_{k \neq i} ED_{kjp}$ their average Euclidean distance in citation from the retracted one, and $\sum_{p \leq t < r} cit_{ijpt}$ cumulative citation of i before retraction year r. Standard errors are clustered at the journal level.

Figure 1.7 shows that papers are retracted faster when published in journals where the average article attracts higher online coverage. Table A.27 (column (1)) illustrates that one standard deviation increase in journal visibility (measured as the average Altscore of non-retracted articles in a journal-year) reduces time to retraction by approximately 15% of its average. Looking across the remaining columns, the relationship is robust to different measures of visibility. The negative association between journal visibility and retraction timing could be driven by higher scrutiny from the scientific community for salient research, together with better detection technologies for visible outlets.

[58] The measure is based on the entire pool of papers published in same year and journal as the retracted ones (and excluding the retracted ones).

1.6 Author-level analysis

Previous sections have shown how media coverage impacts the scientific recognition of retracted papers (i.e. citations) as well as the chances and pace of discovery of faulted research. I now turn to investigate the potential impact of online attention on the subsequent research output of the authors of retracted papers. Low authors output after retraction may reflect a combination of factors: (a) erosion of trust in the authors' work by the scientific community, (b) loss of individual resources for research, or (c) any other direct consequence in terms of academic employment. Studying retracted authors' output is important to measure the individual overall cost of "bad science" and allows to discern the potential role of media on authors' future careers.

1.6.1 Selected sample

Using the same selection of papers illustrated in Section 1.3.3 and 1.3.4, I search all authors of retracted and control papers available in the Scopus library, the retrieved list contains $N_a = 17,991$ distinct authors with any prior publication.[59] For each of these authors, I retrieve their (non-retracted) publications and their corresponding yearly citations and I compute individual output measures focusing on a window of 5 years around the retraction.

Before proceding to the analysis, I select my sample as follows. First, I observe that some retracted authors appear multiple times, therefore I retain only observations relative to their first (in-sample) retraction.[60] Second, I keep cases for which I observe at least one retracted and one control author.[61] Finally, when available, I retain a maximum of three authors per paper (first, mid and last)[62] which provide a final sample of $N_a = 6,718$ authors of which $N_{a,r} = 2,047$ retracted authors, $N_{a,c} = 4,671$ control authors, corresponding to $N_r = 874$ undelining retractions.

In this analysis I intend to study hetherogeneus effects by ranking of appereace in the authorship list, by seniority (based on H-index prior retraction), and by severity of misconduct. Table 1.2 and Table 1.3 provide all relative sub-sample sizes.

[59] $N_{a,r} = 4,105$ retracted authors and $N_{a,c} = 13,886$ control authors, corresponding to the original sample of retractions $N_r = 900$.

[60] Reducing the underlining sample of retractions, $N_r = 987$.

[61] $N_a = 17,148$, of which $N_{a,r} = 4,077$ and $N_{a,c} = 13,071$ corresponding to $N_r = 874$.

[62] Note these three categories are mutually exclusive and sigle authors are confidered as first authors.

1.6.2 Summary statistics

To measure authors' research output I look at: (a) the number of papers published per author per year, (b) the number of papers published per author per year mentioning any source of funding, and (c) the average number of authors across all papers published within the same year. The first is a measure of output *productivity*, the second is an imperfect measure of access to funding, while the last is an indicator for individual collaboration practices. Despite being impefect, these measures allow to assess, in a within author analysis, for the presence of career effects of retractions due to media exposure.

Table 1.4 illustrates the average outputs for this selected sample of academics. In general, authors publish 5.5 papers per year, of which 2.8 with some declared funding support and with 5.8 authors per paper. First authors seem relatively more junior while last authors are generally more senior (with 3.3 vs. 8.3 publications per year). In medical research, last authors are usually senior researchers with stable careers, whereas first and middle authors can be transient authors who may not pursue a scientific career. Large differences also emerge when comparing authors with high and low (ex-ante) H-index. The formers indeed publish more papers (8.3 vs 2 per year) and with a larger set of authors (7 vs. 4 per paper). Output measures of authors associated to different causes of misconduct are generally balanced. Finally, Table A.2 compares average outputs across all authors whose original publication gained initial online attention (or not). Across all categories, authors of newsworthy research have higher publication rates, higher funding support and a larger set of coauthors.

1.6.3 Estimating specification

This section investigates whether media coverage influences authors' careers after the reputational shock of a retraction. The worsening of authors' output may reflect a combination of destruction in access to resources for research as well as an erosion of trust in authors' work by the scientific community. To study this I employ a difference-in-differences strategy that compares output measures of retracted authors to that of control authors of similar never retracted studies,[63] before and after their first observed retraction [64]. Crucially, I further contrast authors

[63]See Section 1.3.4 for details on the selection of control papers.
[64]See details on sample selection in Section 1.6.1

whose original publication obtained any media exposure at publication (i.e. *early mentions > 0*) to those who did not.

Therefore, the regression model is the following:

$$E[Y_{aigt}|X_{aigt}] = exp[\alpha + \gamma_1 Post_{aigt} + \beta_1 R_{aigt} * Post_{aigt} + \beta_2 Post_{aigt} * Media_{ai} + \\ + \beta_3 R_{ai} * Post_{aigt} * Media_{ai} + \delta_a + f(CareerLenght_{at}) + \delta_\tau] \quad (6)$$

where a are authors of treatment (or control) paper i of case-level group g[65] in the years t relative to the retraction. The dependent variable Y is either one of the measures of author outputs: (a) the number of publications in each year, (b) the number yearly publications with grant support, and (c) the average number of authors across all publications of each year. R is an indicator for retracted authors, $Post$ is an indicator variable equal to one for all years after retraction, and $Media$ captures the exposure of the original (retracted or control) paper to online coverage at the time of publication. Specifically, the indicator is equal to one if the paper received at least one online mention within two weeks from publication in either newspapers or blogs and zero otherwise (i.e. $\mathbb{1}[Early\ Mentions > 0]$). Notice that early mentions are assumed to broadly advertise the original research findings and are generally balanced across treatment and control papers. The coefficient β_1 captures the effect of a retraction shock on retracted authors careers as compared to control authors. The coefficient β_3 captures any difference in the effect of the shock for authors whose papers received online attention. Fixed effects are included for each author δ_a and each calendar year δ_τ while $f(CareerLenght_{at})$ represents a full set of dummies for years since the author first publication (ever observed in the Scopus library) and is meant to flexibly control for the academic experience of authors.[66] To look at the dynamics of the differential effect of $Media$, estimates will be presented splitting the sub-samples of authors exposed or not to media and replacing the $Post$ indicator with a full set of dummies for each year relative to the year of retraction. [67] Given the skewed nature of the dependent variables (e.g. almost 20% of yearly observations see 0 published papers), I use a pseudo Poisson regression model

[65] Notice that a group is composed by all authors of a retracted paper and its paired control papers.

[66] Note that the interaction term $R * Media$ is absorbed by author fixed effect.

[67] $E[Y_{aigt\tau}|X_{aigt}] = exp[\sum_{t=r-5}^{r-2} \gamma_{1t} * d_t + \sum_{t=r}^{r+5} \gamma_{1t} * d_t + \sum_{t=r-5}^{r-2} \beta_{1t} * d_t * R_{ai} + \sum_{t=r}^{r+5} \beta_{1t} * d_t * R_{ai} + \sum_{t=r-5}^{r-2} \beta_{2t} * d_t * Media_{ai} + \sum_{t=r}^{r+5} \beta_{2t} * d_t * Media_{ai} + \sum_{t=r-5}^{r-2} \beta_{3t} * d_t * R_{ai} * Media_{ai} + \sum_{t=r}^{r+5} \beta_{3t} * d_t * R_{ai} * Media_{ai} + \delta_i + f(age_{it}) + \delta_\tau]$ separately for the subsample of original papers with $Media = 1$ (or $Media = 0$).

developed by Correia et al. (2020)[68] where consistency is achieved under the only assumption that the conditional mean of the dependent variable is correctly specified (Gourieroux et al., 1984). Finally, standard errors are clustered at the case g level.

1.6.4 Results

Author-level estimates are presented in Table 1.10. On average, authors' future outputs are negatively impacted by a retraction (about 8.6% loss in forward yearly publications and number of collaborators, corresponding to about half publication per year and half author per paper per year.) and more so for senior authors and severe cases of misconduct. Media exposure adds a further loss in output which is never significant.

Looking more closely into sub-groups, estimates suggests that authors appearing first in the co-authorship list are the ones whose *productivity* is differentially and significantly impacted by media ($\simeq$45% loss in yearly publications and grant supported publications, corresponding to 1.5 less papers published per year, half of which with grant support.). Senior authors and severe cases of misconduct also display a further loss with media which is never significant. This could explain previous findings by Mongeon and Larivière (2016) that first authors are most punished after retractions.[69]

However studying the dynamics of this effect, Figure 1.8 (together with Figures A.14 and A.18) shows that the negative impact of retraction is large and permanent only if the original publication had some media exposure ($\simeq$10% larger reduction in future publication rate relative to a case with no media). Absent media coverage author outputs are only moderately impacted and may even fully recover by the end of the 5 year window. The differential impact corresponds to 1 less publication per year for authors with media exposure, against half publication less per year, compared to their respective averages. This differential effect of media is evident for first authors (see Figures A.11, A.15, A.19), for authors with high ex-ante H-index (see Figures A.12, A.16, A.20) and authors whose paper was retracted for severe misconduct (see A.13, A.17, A.21). These figures are all based on split regressions illustrated in Table A.32.

[68] http://scorreia.com/software/ppmlhdfe/
[69] Excluding few single authors cases does not change the results (not shown).

1.7 Conclusion

Flawed research can be harmful both within and outside of academia. The literature document that scientific publications lose significant citations after a retraction. Worryingly though, studies also show that retracted publications still get cited long after they are removed from the literature, potentially disseminating misinformation. In the context of scientific retractions, their visibility is a crucial factor, yet there is little evidence on how media reporting may influence the retraction process and authors' careers. This paper shows that media coverage shapes the auto-correcting process of science by reducing the amount of misinformation and increasing punishment for retracted authors.

I use a conditional difference-in-differences strategy to show that articles that gained popularity in the media — at the time of publication — face heavy citation losses after their retraction while remaining citations become more *accurate* in acknowledging the retraction. This differential effect is considerable for cases of severe misconduct, and it is present only for publications in hard sciences, suggesting distinct publication practices or different topic salience may impact the visibility of a retraction. In addition, retracted authors' future research output is permanently reduced, but only with media coverage (specifically for first authors). I also produce evidence that media can influence the likelihood of retraction and its timing.

Overall, the media seems to help the auto-correcting process of science. At the same time, this implies that plenty of wrong science goes unnoticed. The scientific community, thus, needs better strategies to increase the level of scrutiny and lower incentives for *bad* science. For example, journals could increase transparency at submission and systematically check references of newly accepted papers before publication. This research also proves that media platforms can be a useful communication tool, as in the case of *RetractionWatch* and, more recently, the Twitter bot from *Scite.ai*.[70]

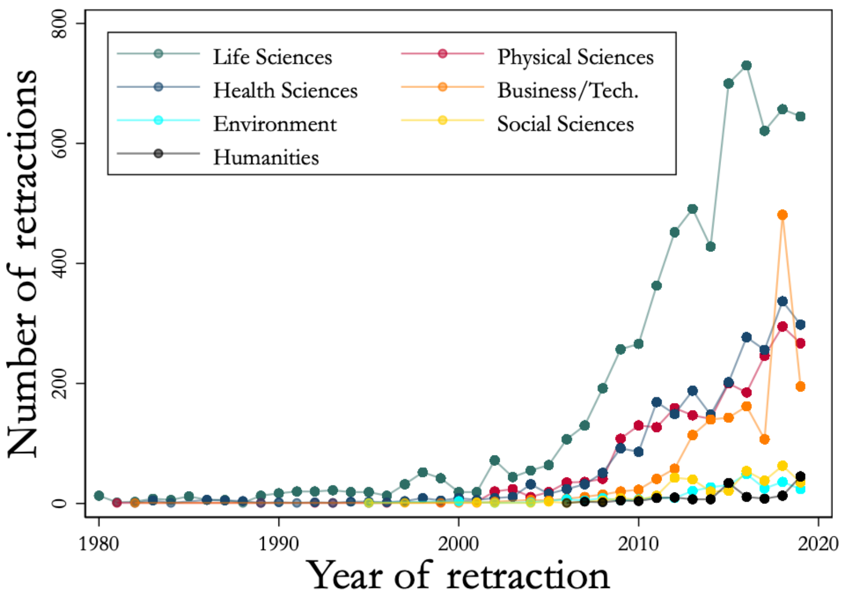
Yet, the scientific information that appears in the media spreads beyond the scientific community. Indeed, while media helps scientists to update beliefs about the credibility of a study and its authors, one question remains about whether this could generate unintended consequences for the main audience of mainstream media: the general public. I show that newspapers, as opposed

[70] See: Sciete.ai Twitter bot.

to blogs, are more likely to advertise the publication of a paper rather than inform about its later retraction. This possible misinformation can impact public perceptions and behaviour, therefore, deserves further research.

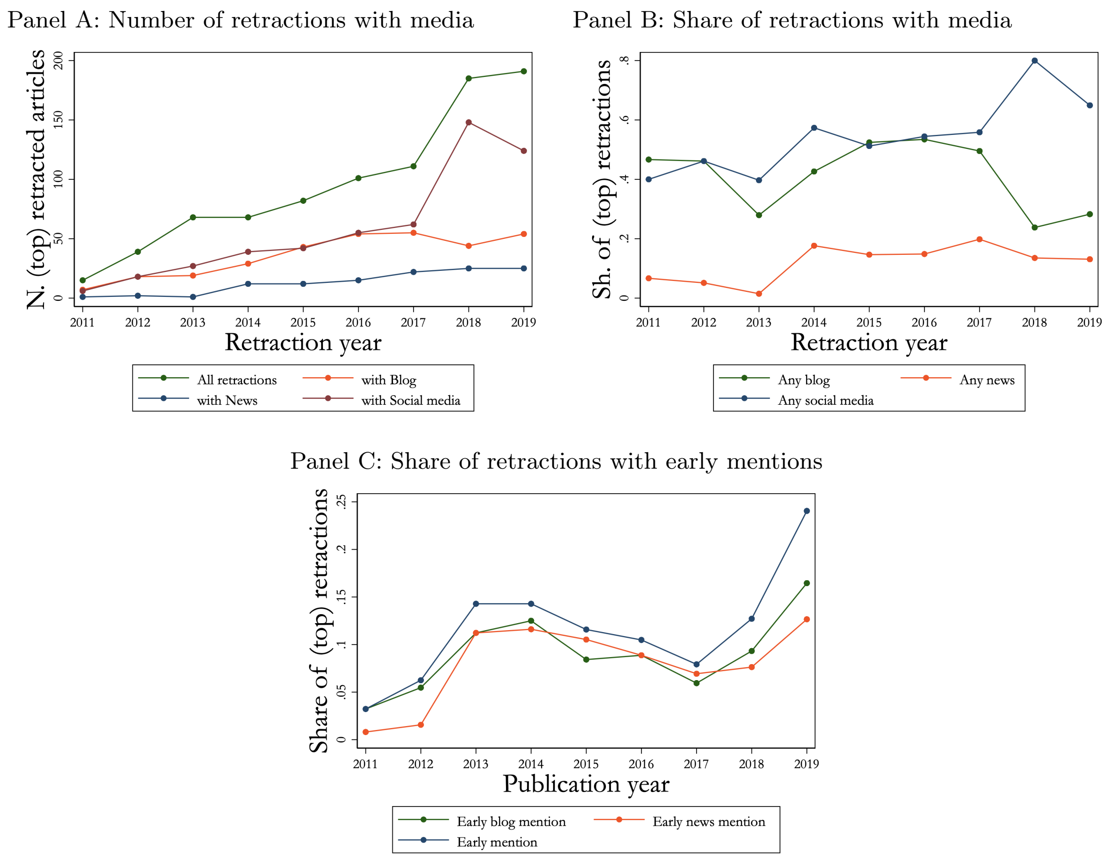
1.8 Figures

Figure 1.1: Retractions over time and across subjects.



Note: Numbers reflect the full *RetractionWatch* database as of July 2020, for visual purposes one outlier publisher (e.g. IEEE) was excluded.

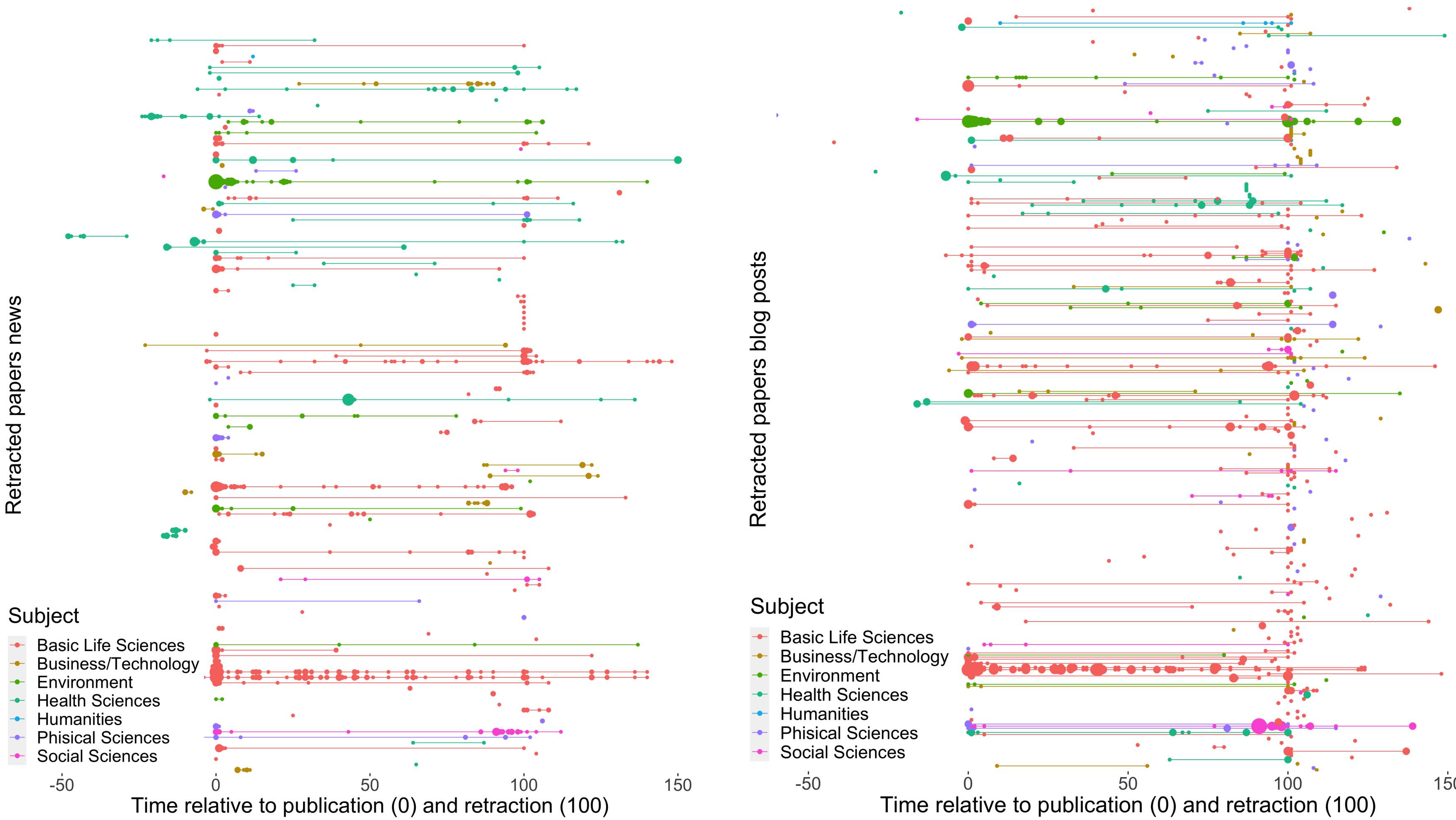
Figure 1.2: Media coverage of sample of retracted papers.



Note: Panel A shows the absolute number of retracted articles in the sample (green) which ever featured in blogs (orange), newspapers (blue), or social media (red), ordered by the year when the retraction occurred. Panel B shows the share of retracted papers that ever appeared in blogs (green), newspapers (orange), or social media (blues), again ordered by year of retraction. Panel C represents the share of retracted articles that were ever mentioned in blogs (green), newspapers (orange) or at least one of the two (blue) within two weeks from publication (i.e. *early mentions*), ordered by year of publication.

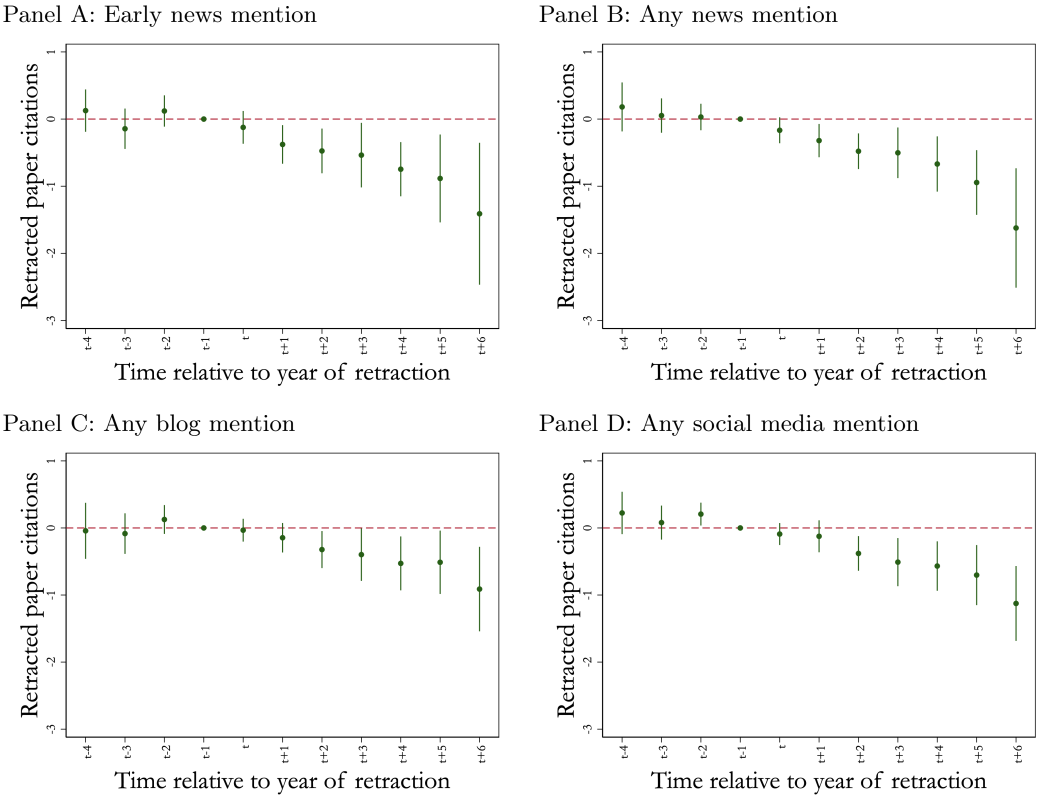
Figure 1.3: Newspaper and blog mentions of retracted articles.

Panel A: News mentions (N=135 retractions) Panel B: Blog mentions (N=365 retractions)



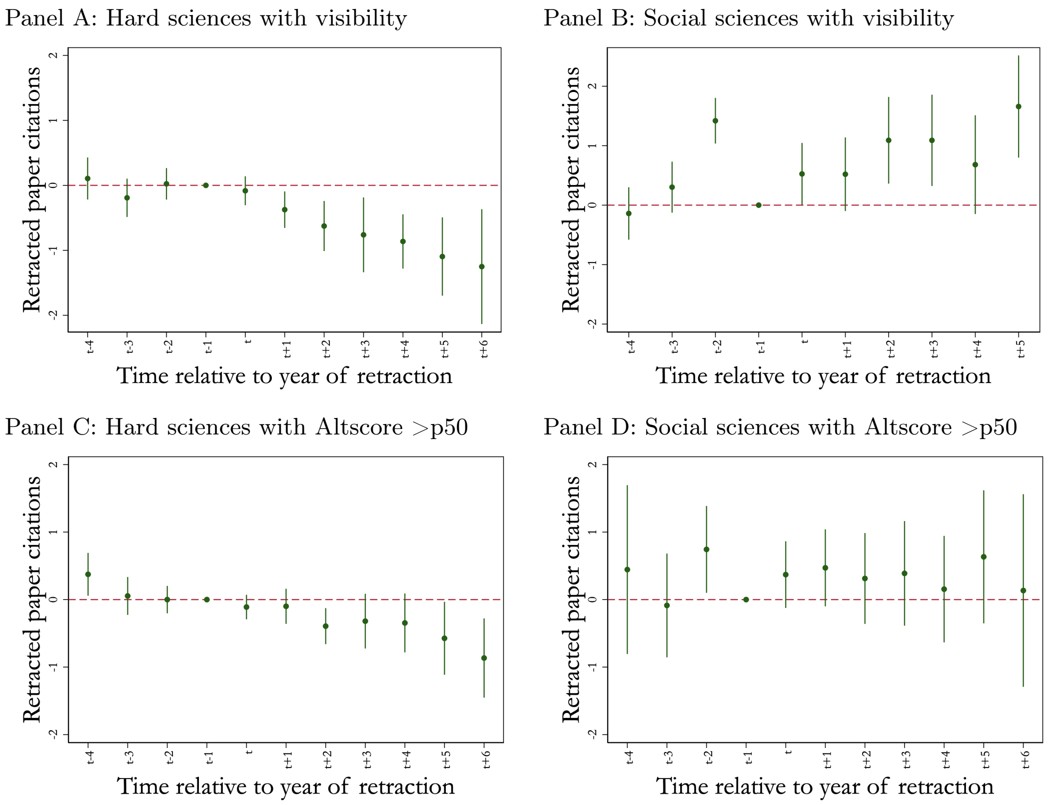
Note: Each line connects the first to the last mention of a single research article on either newspapers (Panel A) or blogs (Panel B) within the considered time window. Dots represent the number of mentions at a certain point in time. The window of analysis focuses on two events: the paper publication date (indexed with 0) and the paper retraction date (indexed with 100). The time score is allocated following the formula $\frac{(t_{mentionposted} - t_{publication})}{(t_{publication} - t_{retraction})} * 100$. The sources of publication date and retraction date are *Altmetric* and *RetractionWatch* respectively.

Figure 1.4: Dynamics of retracted papers penalty with media coverage



Note: Estimates replicate the following models: Table 1.5 column (3) for Panel A; Table A.4 column (3)-(4)-(1) respectively for Panel B, Panel C and Panel D. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). The coefficients displayed are that of the interaction between time dummies, a treatment indicator and a media indicator while vertical lines represent 95% CI.

Figure 1.5: Dynamics of retracted papers penalty with media coverage by discipline



Note: Hard sciences: life sciences, environment, health and physical sciences. Social sciences: business and technology, humanities, other social sciences. Estimates replicate the following models: Table 1.6 column (3)-(4) for Panel A and Panel B; Table A.18 column (3)-(4) respectively for Panel C and Panel D. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). The coefficients displayed are that of the interaction between time dummies, a treatment indicator and a media indicator, for different subsamples of discipline, while vertical lines represent 95% CI.

Figure 1.6: Papers' titles wordclouds

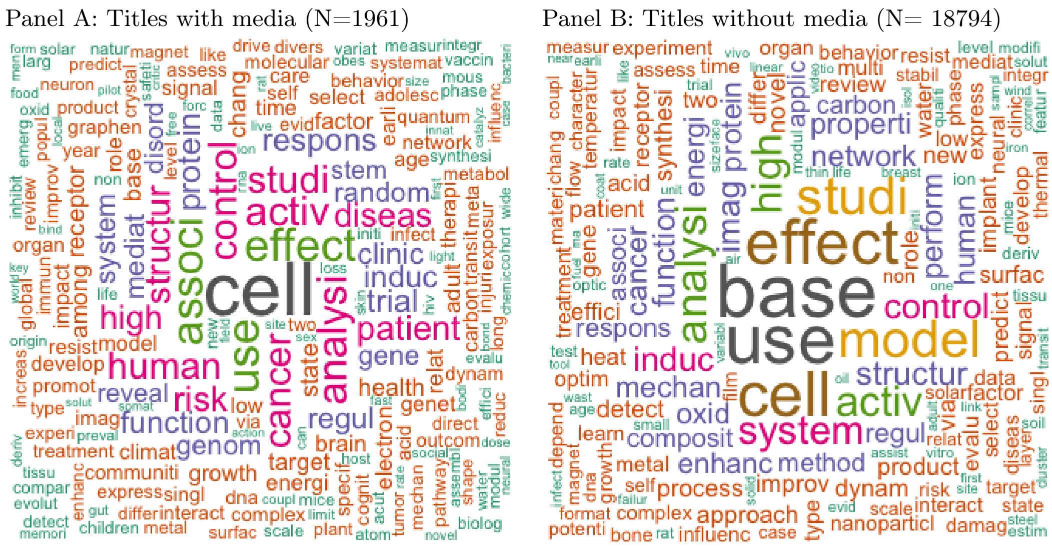
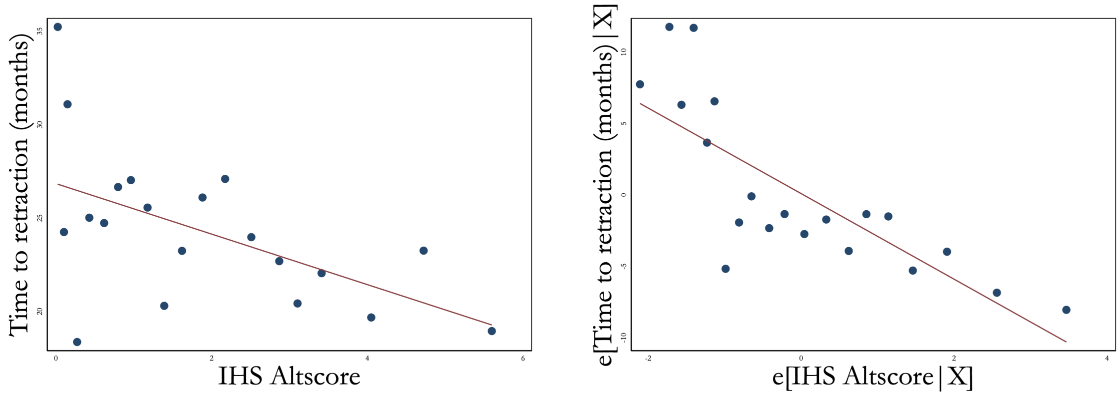


Figure 1.7: Months to retraction and Journal-year average visibility

Panel A: Raw data

Panel B: Absorb Controls + FE

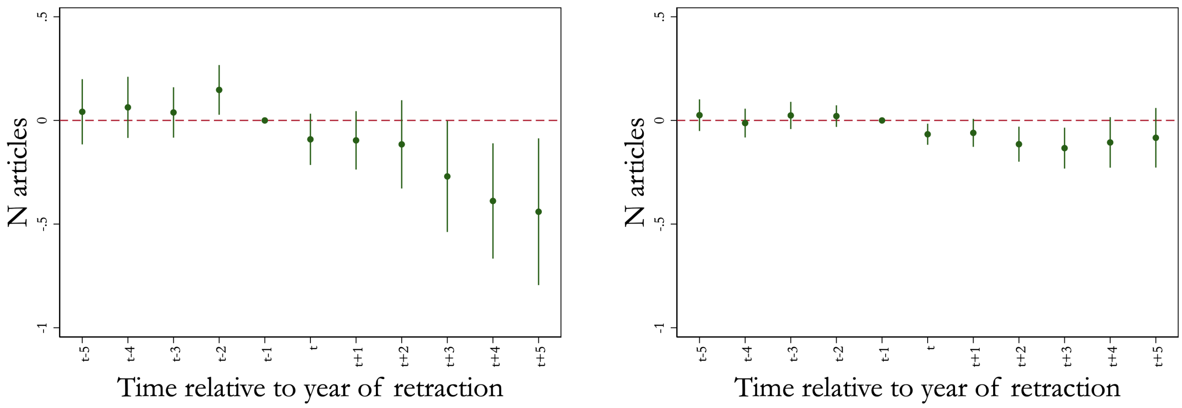


Note: The vertical axis represents the time intercurring between an article publication and its retraction, expressed in months. The orixontal axis represents the inverse hyperbolic sine transformation of journal visibility, measured as the average Altscore of non-retracted papers that appear in the same yournal and year of the retracted one. Controls include the number of non-retracted articles within same journal and year of the treated, the average Euclidean distance of those from the retracted paper, and the level of (non-self) cumulative citations of the retracted paper before retraction. Publication year fixed effects are included. * p<0.1, ** p<0.05, *** p<0.01.

Figure 1.8: Dynamics of Author "productivity" with media coverage

Panel A: With Media

Panel B: Without Media



Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32 Panel A column (1) and (9) respectively. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

1.9 Tables

Table 1.1: Selected summary statistics

	Mean	Sd	Min	Max	Mean	Sd	Min	Max	Diff.
	TREATMENT (N=990)				CONTROL papers (N=1969)				P-value
BALANCING VARIABLES									
Euclidean distance					0.937	2.661	0	45.51	
Arithmetic distance					0.171	1.699	-10	10	
Cum. (no self) citations ($t-1$)	7.103	19.11	0	254	6.807	18.65	0	258	0.6862
Early news mentions	1.037	7.844	0	134	0.787	5.790	0	127	0.3259
Early blog mentions	0.240	1.511	0	28	0.152	0.862	0	21	0.0429
Early mentions	1.278	9.194	0	155	0.939	6.493	0	139	0.2461
Age	5.138	2.612	0	9	5.140	2.612	0	9	0.9860
ADDITIONAL VARIABLES									
Time to retract	2.067	2.021	-0.504	9.353					
Yearly citations (no self)	2.628	4.442	0	56	5.259	12.27	0	354.8	
Cum. (no self) citations	16.83	30.50	0	418	33.91	75.29	0	1,774	
Altscore	37.54	274.5	0	7,128	19.09	130.2	0	3,728	
Tweeters count	32.12	349.9	0	10,105	13.55	165.4	0	5,100	
News count	1.459	8.589	0	122	1.068	6.158	0	113	
Blog count	0.871	3.283	0	65	0.287	1.333	0	27	
Any social media mention	0.597	0.491	0	1	0.443	0.497	0	1	
Any news mention	0.136	0.343	0	1	0.119	0.324	0	1	
Any blog mention	0.369	0.483	0	1	0.110	0.313	0	1	
Any early mentions	0.111	0.314	0	1	0.0945	0.293	0	1	

Note: Self-citations are excluded from citation count. *Early mentions* include all news and/or blog posts published within 2 weeks from publication. *Altscore* is a weighted average of all online mentions across outlets. Media *counts* are the number of outlets/accounts referring to a paper at any point in time. All papers are published/retracted between 2011 and 2020.

Table 1.2: Author level sample size

	First author	Mid author	Last author	H-index >p50	H-index <=p50	Media	Not media	Severe misconduct	Non-Severe misconduct	Total
Treatment	708 (35%)	650 (32%)	689 (34%)	922 (45%)	1125 (55%)	265 (13%)	1782 (87%)	851 (42%)	1196 (58%)	2047
Control	1639 (35%)	1437 (31%)	1595 (34%)	2412 (52%)	2259 (48%)	500 (11%)	4171 (89%)	1974 (42%)	2697 (58%)	4671
Total	2347	2087	2284	3334	3384	765	5953	2825	3893	**6718**

Note: The sample includes authors of retracted (treatment) papers and authors of matched control papers after the first observed retraction. It includes a maximum of 3 authors per paper (ranked as first, mid or last as per order of appereance) which have at least one publication in the 5 years before the first observed retraction. The H-index is calculated based on pre-retraction publications. Media is an indicator for whether the original publication gained any early popularity in the media. Causes of retractions are classified as Severe based on Woo and Walsh (2021).

Table 1.3: Author level sub-sample size

	First author	Mid author	Last author	H-index >p50	H-index <=p50	Severe misconduct	Non-Severe misconduct	Sub total
Panel A: With (early) media								
Treatment	97 (37%)	82 (31%)	86 (32%)	154 (58%)	111 (42%)	138 (52%)	127 (48%)	265
Control	172 (34%)	161 (32%)	167 (33%)	300 (60%)	200 (40%)	279 (56%)	221 (44%)	500
Sub-total	269	243	253	454	311	417	348	765
Panel B: Without (early) media								
Treatment	611 (34%)	568 (32%)	603 (34%)	768 (43%)	1014 (57%)	713 (40%)	1069 (60%)	1782
Control	1467 (35%)	1276 (31%)	1428 (34%)	2112 (51%)	2059 (49%)	1695 (41%)	2476 (59%)	4171
Sub-total	2078	1844	2031	2880	3073	2408	3545	5953
Total	2347	2087	2284	3334	3384	2825	3893	**6718**

Note: The sample includes authors of retracted (treatment) papers and authors of matched control papers after the first observed retraction. It includes a maximum of 3 authors per paper (ranked as first, mid or last as per order of appereance) which have at least one publication in the 5 years before the first observed retraction. The H-index is calculated based on pre-retraction publications. Media is an indicator for whether the original publication gained any early popularity in the media. Causes of retractions are classified as Severe based on Woo and Walsh (2021).

Table 1.4: Author level summary statistics

	(1) N	(2) mean	(3) sd	(4) min	(5) max	(6) N	(7) mean	(8) sd	(9) min	(10) max
	Panel A: All					Panel B: First author				
N articles	55,116	5.465	8.413	0	170	18,792	3.366	5.689	0	149
N articles with grant	55,116	2.781	5.617	0	156	18,792	1.651	4.069	0	149
N of coauthors	55,116	5.818	7.460	0	100	18,792	5.166	7.271	0	100
	Panel C: Mid author					Panel D: Last author				
N articles	16,798	4.501	7.169	0	129	19,526	8.315	10.55	0	170
N articles with grant	16,798	2.307	4.817	0	129	19,526	4.277	7.036	0	156
N of coauthors	16,798	5.945	8.090	0	100	19,526	6.336	7.015	0	100
	Panel E: H-index above median					Panel F: H-index below median				
N articles	28,822	8.316	10.18	0	170	26,294	2.341	4.023	0	149
N articles with grant	28,822	4.386	6.988	0	156	26,294	1.021	2.585	0	149
N of coauthors	28,822	7.184	8.196	0	100	26,294	4.320	6.224	0	100
	Panel G: Severe cases of misconduct					Panel H: Non-severe cases of misconduct				
N articles	23,614	5.227	8.040	0	129	31,502	5.644	8.677	0	170
N articles with grant	23,614	2.843	5.784	0	129	31,502	2.735	5.488	0	156
N of coauthors	23,614	6.085	7.676	0	100	31,502	5.618	7.287	0	100

Note: All statistics are reported by year. *N of articles* are yearly publications per author. *N of articles with grant* are yearly publications per author that mention any source of funding. *N of coauthors* is the average number of authors across papers published by an author within a year. The sample includes authors of retracted (treatment) papers and authors of matched control papers after the first observed retraction. It includes a maximum of 3 authors per paper (ranked as first, mid or last as per order of appeerance) which have at least one publication in the 5 years before the first observed retraction. The H-index is calculated based on pre-retraction publications. Media is an indicator for whether the original publication gained any early popularity in the media. Causes of retractions are classified as Severe based on Woo and Walsh (2021).

Table 1.5: Retracted papers penalty with early mentions

	(1) Citations	(2) Citations	(3) Citations
Post * Treatment	-0.959***	-0.983***	-0.977***
	(0.059)	(0.058)	(0.058)
Post * Treatment * Early mentions	-0.449***		
	(0.158)		
Post * Treatment * Early blog mentions		-0.396**	
		(0.185)	
Post * Treatment * Early news mentions			-0.418***
			(0.149)
Article FE	Y	Y	Y
Age FE	Y	Y	Y
Year FE	Y	Y	Y
Pseudo R2	0.709	0.709	0.709
N	15438	15438	15438
N clusters	966	966	966
N full	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Early mentions is an indicator for papers with at least one mention (in newspapers and/or blogs) within two weeks from publication. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table 1.6: Retracted papers penalty and early mentions by discipline

	(1)	(2)	(3) Hard sciences	(4) Social sciences
	Citations	Citations	Citations	Citations
Post * Treatment * Early mentions	0.592	0.364	-0.520***	0.366
	(0.396)	(0.262)	(0.163)	(0.264)
Post * Treatment * Early mentions * Business/Technology	-0.155			
	(0.522)			
Post * Treatment * Early mentions * Life sciences	-1.169**			
	(0.458)			
Post * Treatment * Early mentions * Environment	-0.917**			
	(0.457)			
Post * Treatment * Early mentions * Health	-1.195**			
	(0.529)			
Post * Treatment * Early mentions * Physics	-0.938**			
	(0.455)			
Post * Treatment * Early mentions * Hard sciences		-0.887***		
		(0.309)		
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.711	0.710	0.718	0.595
N	15399	15438	12980	2419
N clusters	964	966	798	166
N full	16672	16711	13837	2835

Note: Estimates derive from pseudo Poisson specifications. Hard sciences: life sciences, environment, health and physical sciences. Social sciences: business and technology, humanities, other social sciences. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Early mentions is an indicator for papers with at least one mention (in newspapers and/or blogs) within two weeks from publication. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table 1.7: Retracted papers penalty and early mentions by severity of misconduct

	All	Minor	Moderate	Severe
	Citations	Citations	Citations	Citations
Post * Treatment * Early mentions	-0.158	-0.142	0.096	-0.566***
	(0.335)	(0.360)	(0.196)	(0.193)
P * T * Early mentions * Moderate misconduct	0.239			
	(0.389)			
P * T * Early mentions * Severe misconduct	-0.418			
	(0.389)			
Article, Age & Year FE	Y	Y	Y	Y
Pseudo R2	0.710	0.593	0.694	0.739
N (N clusters)	15438 (966)	4312 (295)	3857 (256)	7269 (415)
N full	16711	4859	4157	7695

Note: Estimates derive from pseudo Poisson specifications. Causes of retractions are classified based on Woo and Walsh (2021). The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Early mentions is an indicator for papers with at least one mention (in newspapers and/or blogs) within two weeks from publication. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table 1.8: Likelihood of retraction and media coverage

	Panel A: Retraction									
	OLS		EBIC		AICC		CV		Rigourous	
Media coverage	0.009**	0.019***	0.014***	0.021***	0.016***	0.021***	0.016***	0.022***	0.013***	0.020***
	(0.004)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)
Predicted media			-0.078***	-0.074***	-0.067***	-0.068***	-0.066***	-0.068***	-0.135***	-0.140***
			(0.027)	(0.031)	(0.018)	(0.023)	(0.018)	(0.022)	(0.045)	(0.058)
R-squared	0.000	0.001	0.001	0.002	0.001	0.002	0.001	0.002	0.001	0.002
	Panel B: Retraction									
			EBIC		AICC		CV		Rigourous	
Predicted media			-0.059***	-0.065**	-0.048***	-0.060***	-0.048***	-0.068***	-0.104***	-0.127**
			(0.024)	(0.031)	(0.016)	(0.022)	(0.015)	(0.022)	(0.041)	(0.058)
R-squared			0.000	0.001	0.000	0.001	0.000	0.002	0.000	0.001
Pub. year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Journal FE	N	Y	N	Y	N	Y	N	Y	N	Y
N	20755	20755	20755	20755	20755	20755	20755	20755	20755	20755
N clusters	1008	1008	1008	1008	1008	1008	1008	1008	1008	1008

Note: Estimates from OLS regression. The dependent variable *Retraction* is an indicator for whether a paper was retracted. *Media coverage* is an indicator for whether a paper attracted online coverage at publication. *Predicted media* is media coverage as predicted from the respective *lasso* procedures. Boostrap standard errors in parentheses, clustered around retraction cases. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 1.9: Citation statements mentioning paper is retracted

	(1)	(2)	(3)
VARIABLES		Statements mentioning paper is retracted	
Post * Early mentions	2.077***		
	(0.457)		
Post * Altscore >p50		1.562**	
		(0.670)	
Post * Altscore 3rd quintile			-0.101
			(1.344)
Post * Altscore 4th quintile			1.137
			(1.223)
Post * Altscore 5th quintile			2.305**
			(1.155)
Article FE	Y	Y	Y
Age FE	Y	Y	Y
Year FE	Y	Y	Y
Relative yr FE	Y	Y	Y
Pseudo R2	0.361	0.341	0.355
N	531	531	531
N clusters	95	95	95
N full	5591	5591	5591

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations statements received by each paper in a particular year which explicitly mention the retraction. Early mentions is an indicator for papers with at least one mention (in newspapers and/or blogs) within two weeks from publication. Altscore is an aggregate measure of weighted online mentions. All models incorporate article fixed effects, a full suite of calendar-year effects, article age indicator variables and dummies for each year relative to the retraction. Standard errors in parentheses, clustered around retraction cases. *p < 0.10, **p < 0.05, ***p < 0.01.

Table 1.10: Impact on authors' careers (interaction)

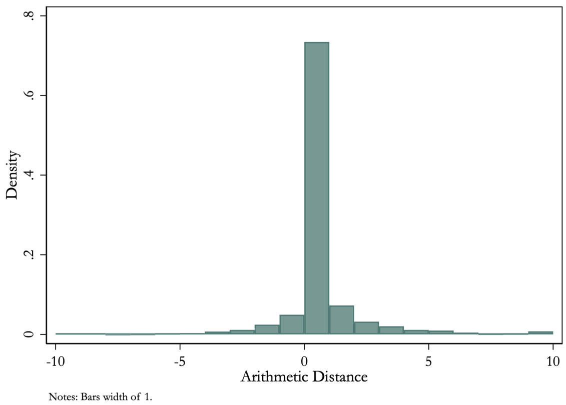
	(1) All	(2) First author	(3) Mid author	(4) Last author	(5) H-index >p50	(6) H-index <=p50	(7) Severe misconduct	(8) Non-Severe misconduct
	Panel A: N articles							
Post * Treatment	-0.093***	-0.090	-0.071	-0.099**	-0.116***	-0.070	-0.225***	-0.019
	(0.035)	(0.059)	(0.055)	(0.046)	(0.041)	(0.054)	(0.051)	(0.046)
Post * Treatment * Media	-0.121	-0.513***	0.174	-0.110	-0.136	0.032	-0.153	-0.030
	(0.095)	(0.174)	(0.128)	(0.124)	(0.102)	(0.154)	(0.131)	(0.125)
Pseudo R2	0.595	0.495	0.578	0.592	0.566	0.411	0.593	0.598
N	54732	18633	16645	19448	28761	25966	23392	31336
N clusters	874	872	851	870	848	856	367	507
N authors	6666	2327	2065	2273	3326	3340	2795	3870
	Panel B: N articles with grant							
Post * Treatment	-0.061	-0.019	-0.022	-0.089	-0.105**	0.007	-0.165**	-0.014
	(0.045)	(0.075)	(0.068)	(0.054)	(0.049)	(0.071)	(0.069)	(0.054)
Post * Treatment * Media	-0.152	-0.563***	0.036	-0.099	-0.160	0.146	-0.181	-0.020
	(0.110)	(0.194)	(0.149)	(0.156)	(0.117)	(0.181)	(0.160)	(0.130)
Pseudo R2	0.562	0.479	0.540	0.567	0.541	0.386	0.566	0.561
N	50243	16642	15120	18472	28134	22108	21753	28486
N clusters	871	859	832	859	845	833	367	504
N authors	6070	2066	1857	2146	3252	2818	2578	3491
	Panel C: Avg. n collaborators							
Post * Treatment	-0.089***	-0.113**	-0.083*	-0.078**	-0.089***	-0.081*	-0.168***	-0.028
	(0.028)	(0.056)	(0.046)	(0.034)	(0.033)	(0.044)	(0.040)	(0.038)
Post * Treatment * Media	-0.002	-0.118	0.158	-0.024	-0.027	0.041	-0.040	0.080
	(0.085)	(0.154)	(0.148)	(0.090)	(0.089)	(0.181)	(0.122)	(0.104)
Pseudo R2	0.383	0.373	0.386	0.385	0.398	0.312	0.371	0.393
N	54732	18633	16645	19448	28761	25966	23392	31336
N clusters	874	872	851	870	848	856	367	507
N authors	6666	2327	2065	2273	3326	3340	2795	3870
Author FE	Y	Y	Y	Y	Y	Y	Y	Y
Career lenght FE	Y	Y	Y	Y	Y	Y	Y	Y
Calendar Year FE	Y	Y	Y	Y	Y	Y	Y	Y
N full	55116	18792	16798	19526	28822	26294	23614	31502

Note: Estimates derive from pseudo Poisson specifications. The dependent variables are: N. published articles x author x year; N. published articles with grant support x author x year; or Avg. n collaborators across all author's publications x year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Media is an indicator for cases where the original publication (either retracted or control papers) had at least one mention (in newspapers and/or blogs) within two weeks from publication. All models incorporate author fixed effects, a full suite of calendar-year effects and carreer lenght indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (e.g. $x\%$ loss in publication rate). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

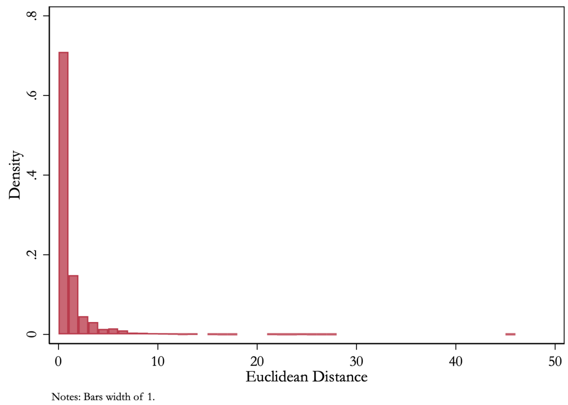
A.10 Appendix Figures

Figure A.1: Control quality: citations

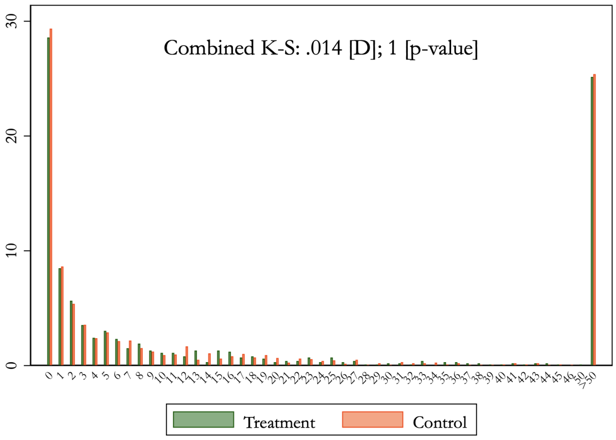
Panel A: AD



Panel B: ED

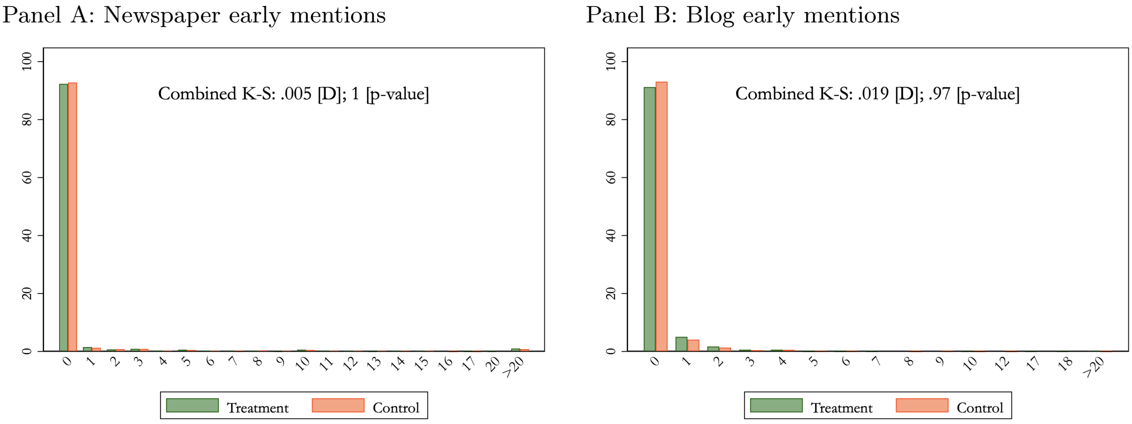


Panel C: Cumulative citations



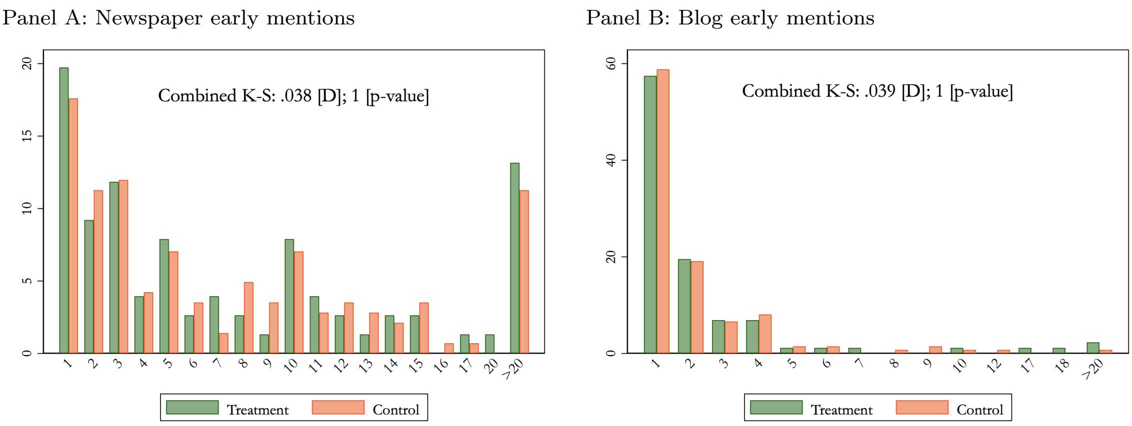
Note: All panels refer to pre-retraction measures. The year of retraction in that of the corresponding treated paper. Panel A shows the distribution of arithmetic distance (AD), panel B shows the distribution of Euclidean distance (ED), and panel C shows the distribution of cumulative citations from publication to the year before retraction and display the result of the Kolmogorov-Smirnov test of the equality of distributions between treatment and control group.

Figure A.2: Control quality: early mentions



Note: Panels display the distribution of online mentions within two weeks from publication in newspapers (panel A) and blogs (panel B) across treated (green) and control (orange) papers. Both graph report the result of the Kolmogorov-Smirnov test of the equality of distributions across groups.

Figure A.3: Control quality: early mentions (Mentions > 0)

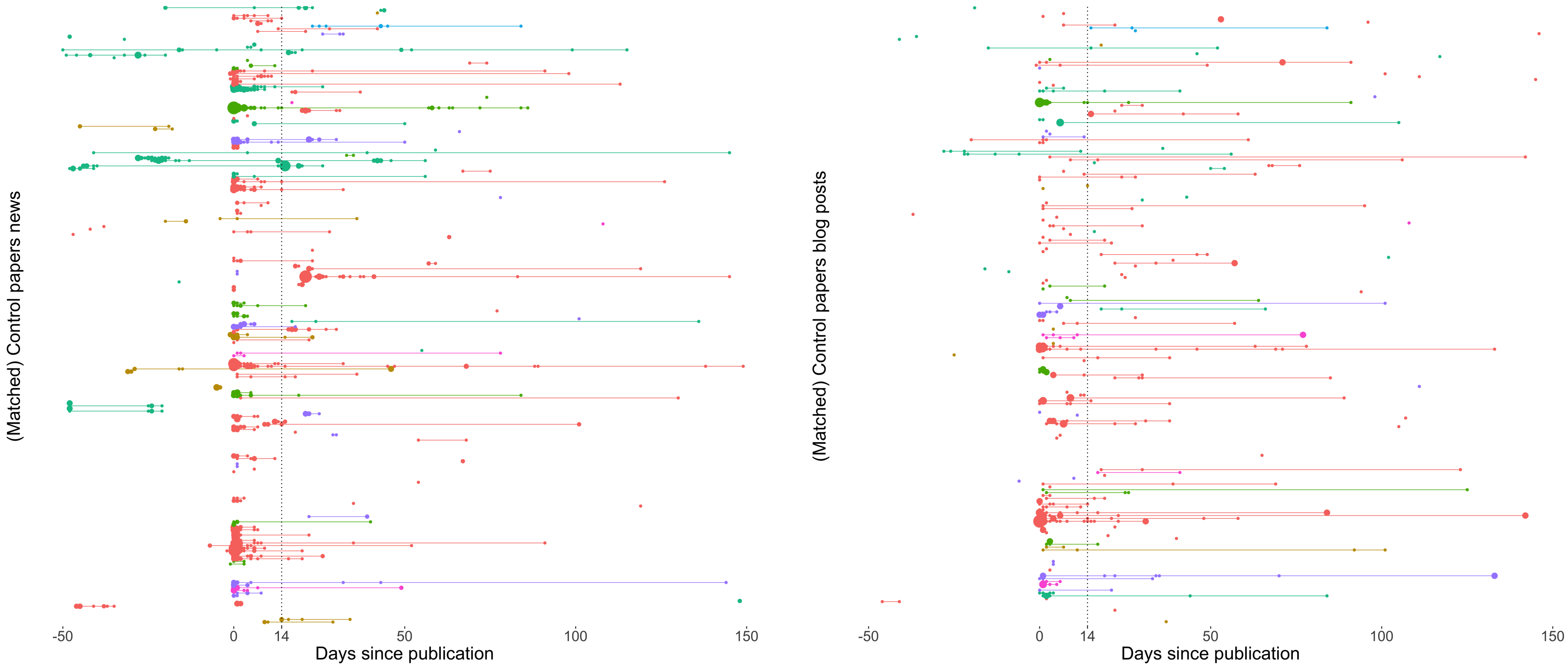


Note: Panels display the distribution of online mentions within two weeks from publication in newspapers (panel A) and blogs (panel B) across treated (green) and control (orange) papers. Publications with no mentions are excluded. Both graph report the result of the Kolmogorov-Smirnov test of the equality of distributions across groups.

Figure A.4: Newspaper and blog mentions of selected control articles.

Panel A: News mentions (N=235 controls)

Panel B: Blog mentions (N=216 controls)

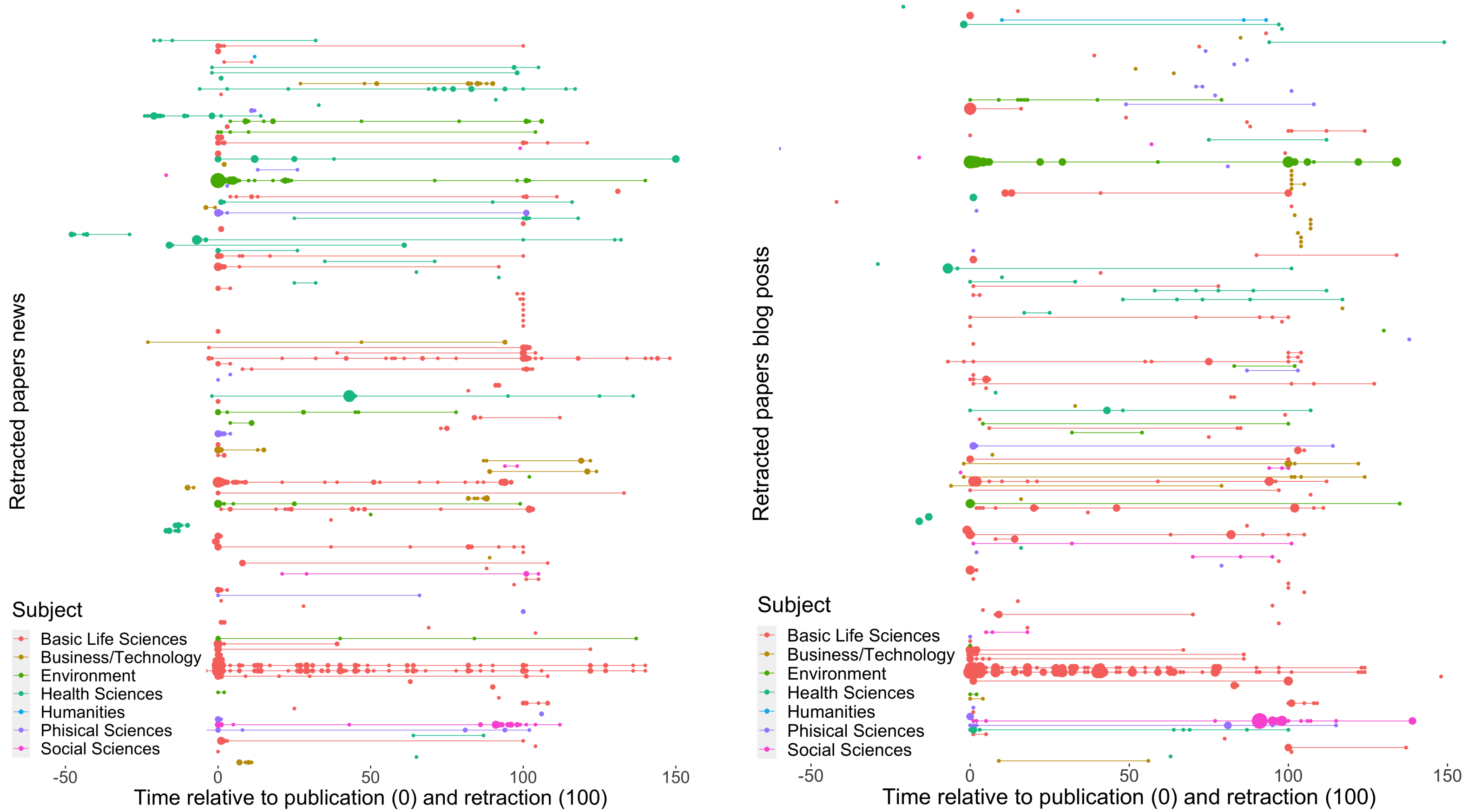


Note: Each line connects the first to the last mention of a single research article on either newspapers (Panel A) or blogs (Panel B) within the considered time window. Dots represent the number of mentions at a certain point in time. The source of publication date is *Altmetric*.

Figure A.5: Newspaper and blog mentions of retracted articles (excluding Retraction Watch from blogs)

Panel A: News mentions (N=135)

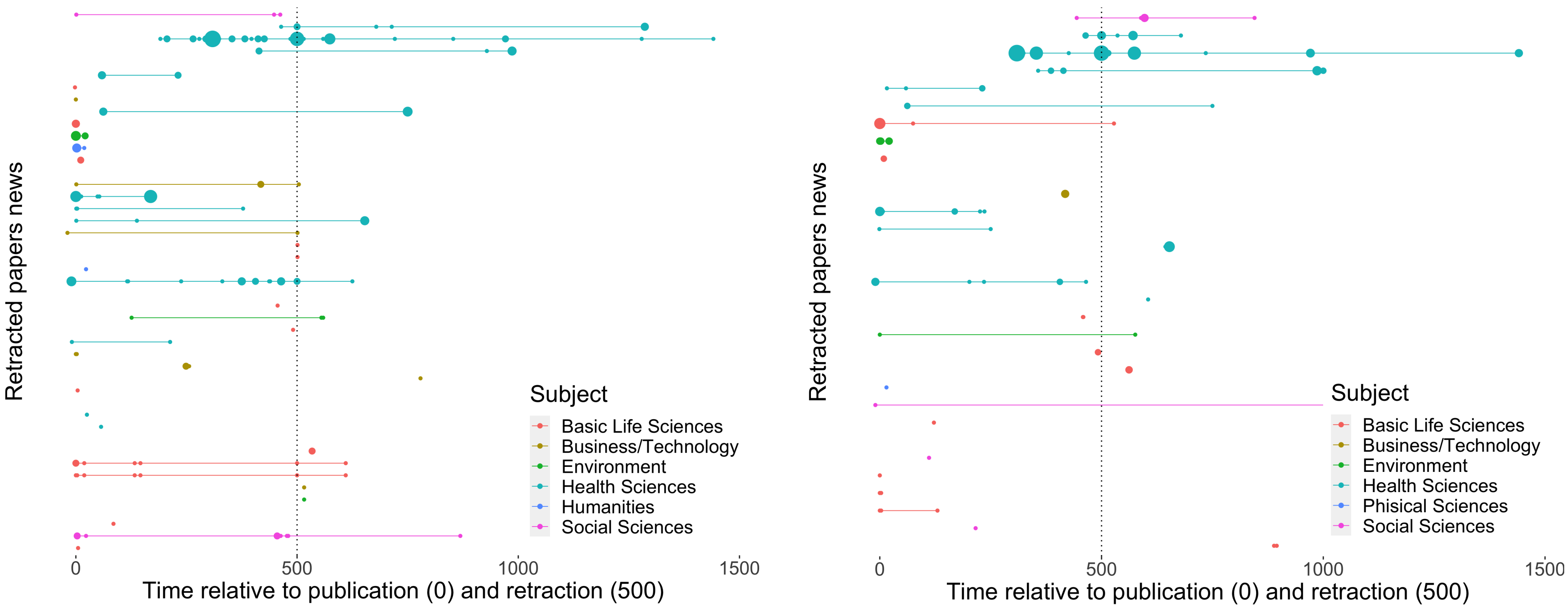
Panel B: Blog mentions (N=171)



Note: Each line connects the first to the last mention of a single research article on either newspapers (Panel A) or blogs (Panel B) within the considered time window. Dots represent the number of mentions at a certain point in time. Blog mention from *RetractionWatch* are excluded. Source of publication date and retraction date: *Altmetric* and *RetractionWatch* respectively.

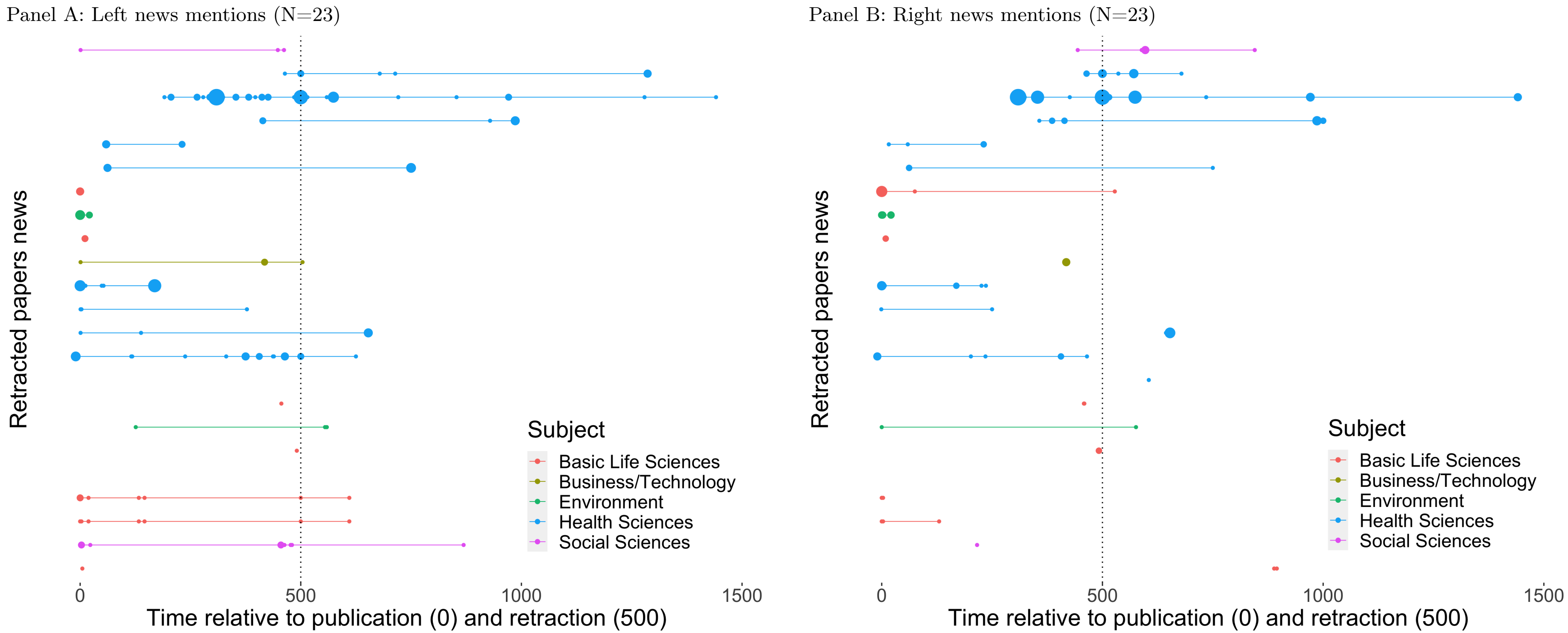
Figure A.6: US News media coverage of retracted papers by slant (within sample median).

Panel A: Left news mentions (N=45) Panel B: Right news mentions (N=31)



Note: Each line connects the first to the last mention of a single research article on left leaning newspapers (Panel A) or right leaning (Panel B) within the considered time window. Right (left) leaning newspapers have a slam index (GS10) above median. The sample includes retractions in lower ranked journals. Dots represent the number of mentions at a certain point in time. Source of publication date and retraction date: *Altmetric* and *RetractionWatch* respectively.

Figure A.7: US News media coverage of retracted papers by slant (within sample median and balanced sample).

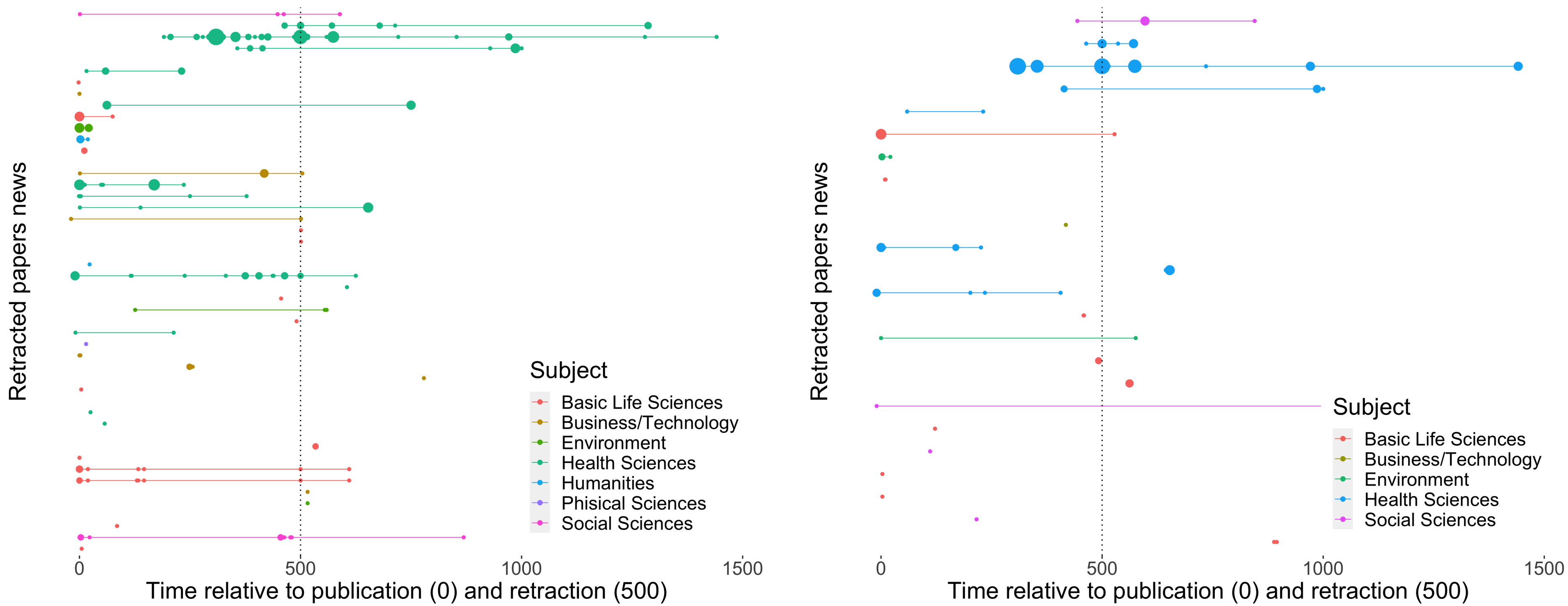


Note: Each line connects the first to the last mention of a single research article on left leaning newspapers (Panel A) or right leaning (Panel B) within the considered time window. Right (left) leaning newspapers have a slam index (GS10) above median. The sample includes retractions in lower ranked journals. Dots represent the number of mentions at a certain point in time. Source of publication date and retraction date: *Altmetric* and *RetractionWatch* respectively.

Figure A.8: US News media coverage of retracted papers by slant (GS10 median).

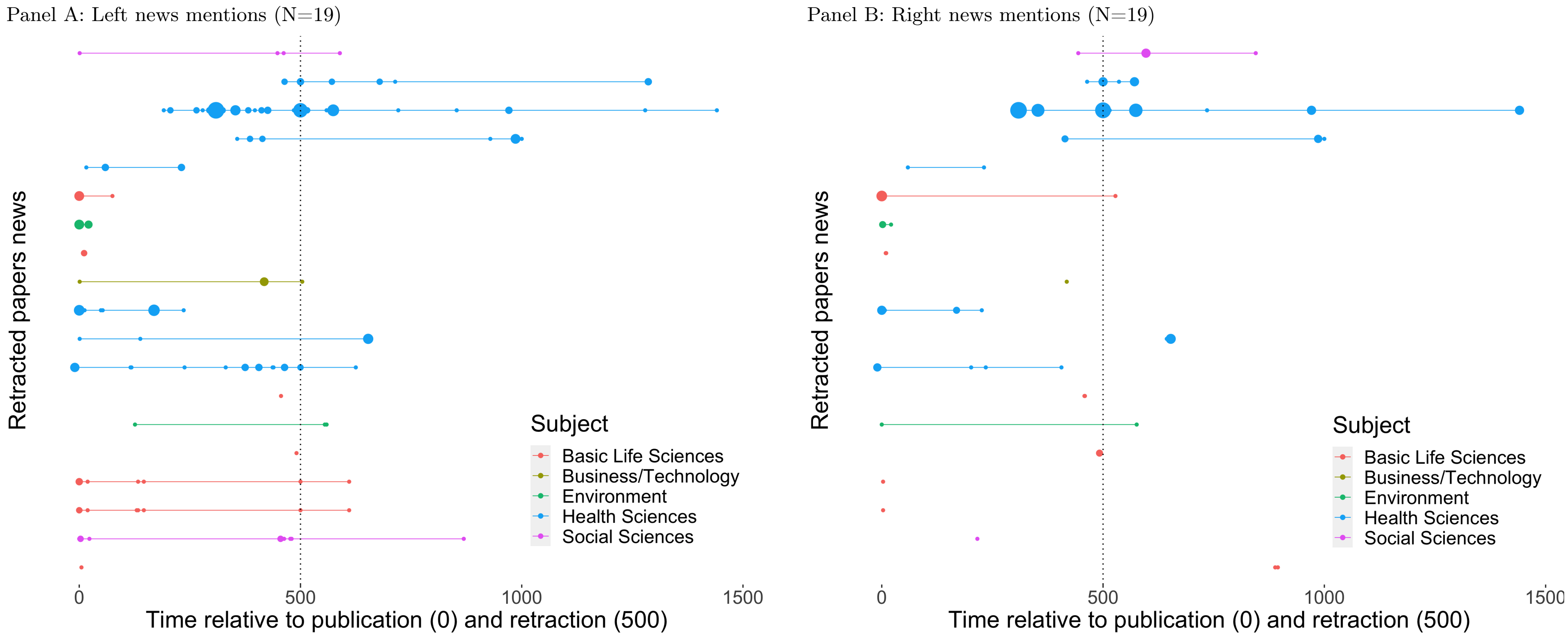
Panel A: Left news mentions (N=48)

Panel B: Right news mentions (N=24)



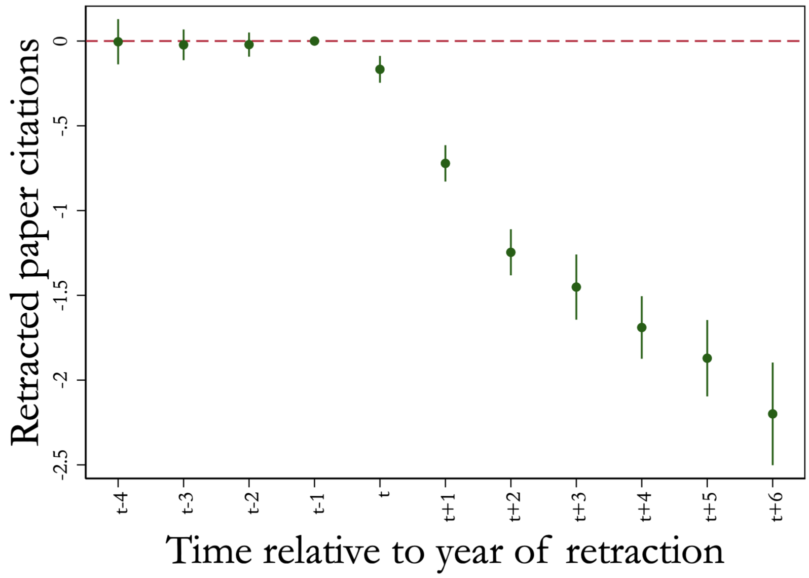
Note: Each line connects the first to the last mention of a single research article on left leaning newspapers (Panel A) or right leaning (Panel B) within the considered time window. Right (left) leaning newspapers have a slam index (GS10) above median. The sample includes retractions in lower ranked journals. Dots represent the number of mentions at a certain point in time. Source of publication date and retraction date: *Altmetric* and *RetractionWatch* respectively.

Figure A.9: US News media coverage of retracted papers by slant (GS10 median and balanced sample).



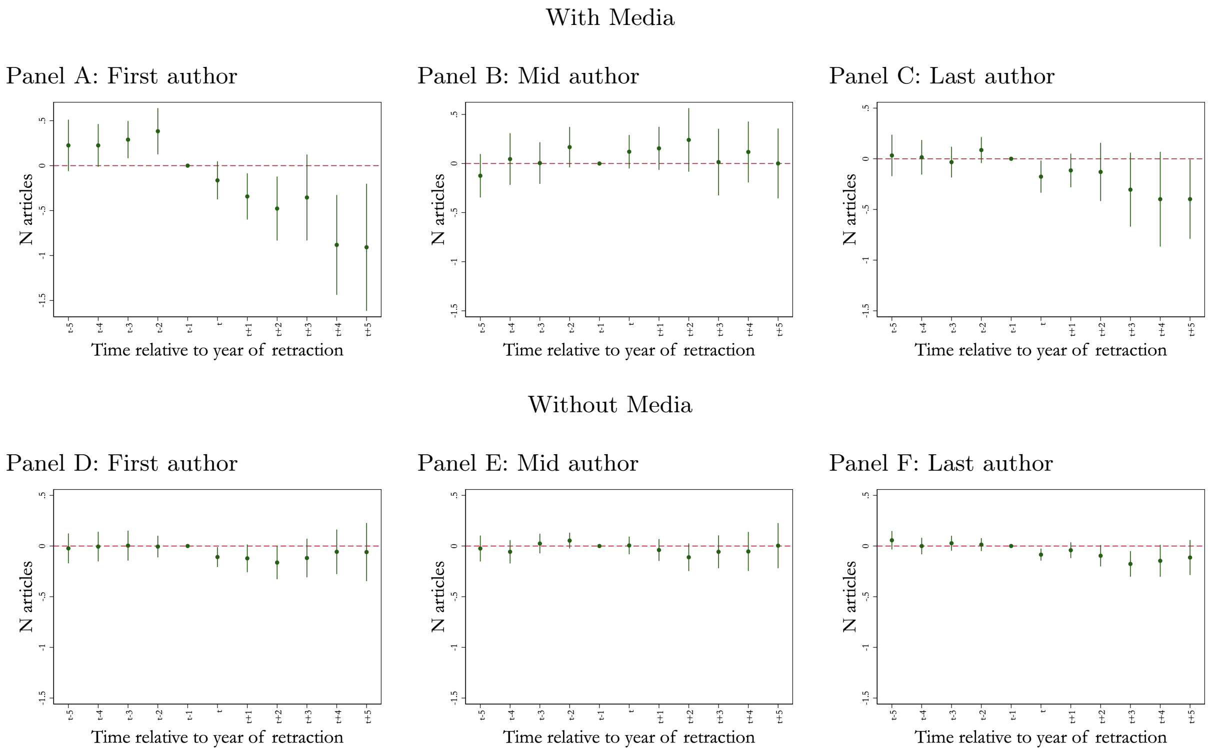
Note: Each line connects the first to the last mention of a single research article on left leaning newspapers (Panel A) or right leaning (Panel B) within the considered time window. Right (left) leaning newspapers have a slam index (GS10) above median. The sample includes retractions in lower ranked journals. Dots represent the number of mentions at a certain point in time. Source of publication date and retraction date: *Altmetric* and *RetractionWatch* respectively.

Figure A.10: Dynamics of retracted papers penalty



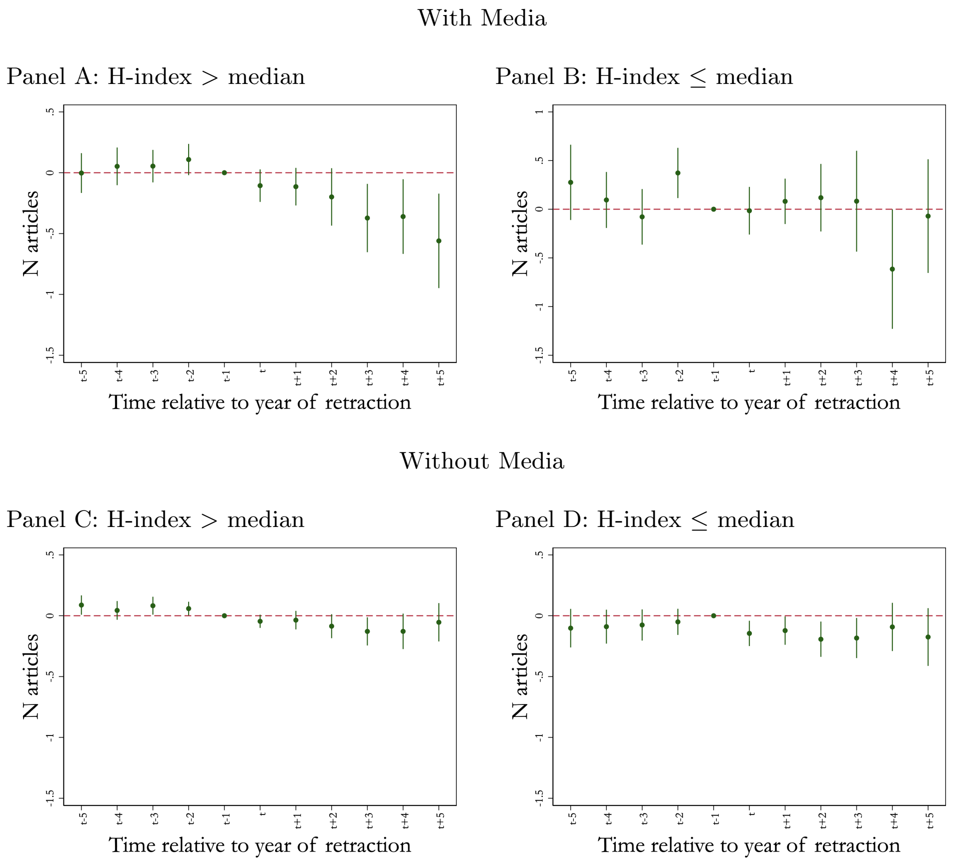
Note: Estimates replicate the model in Table A.3 column (2) but replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t - 1$ excluded). The coefficient displayed are that of the interaction between time dummies and a treatment indicator while the vertical lines represent 95% CI.

Figure A.11: Author "productivity" by rank (with and without media)



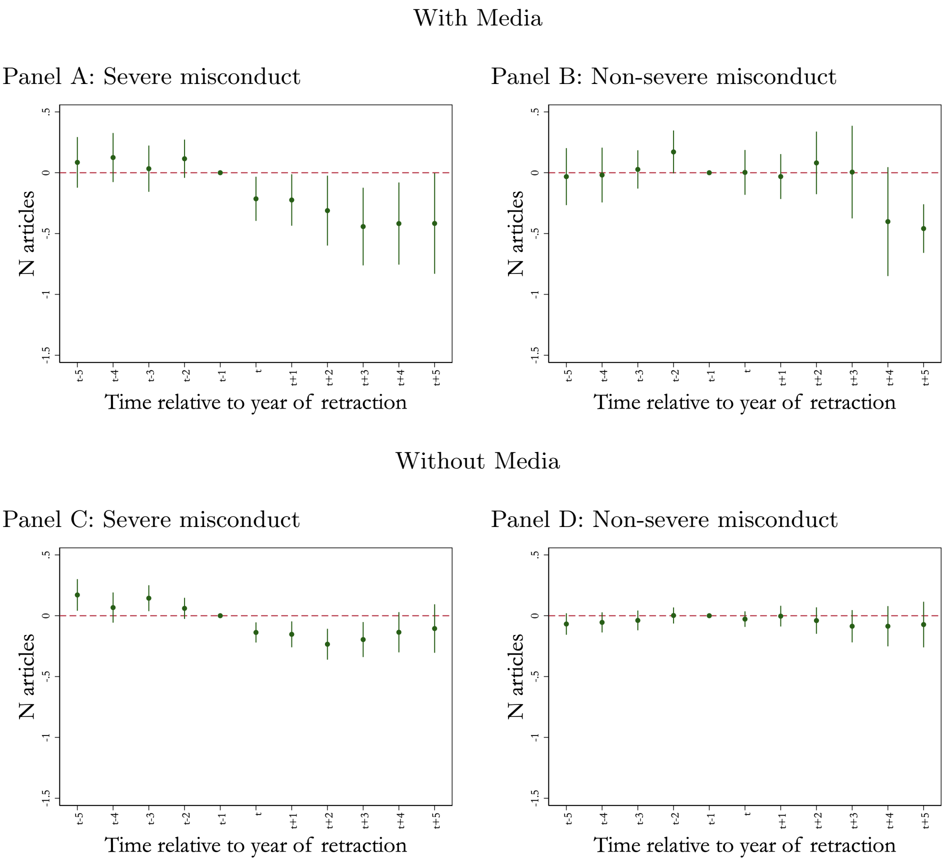
Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

Figure A.12: Author "productivity" by seniority (with and without media)



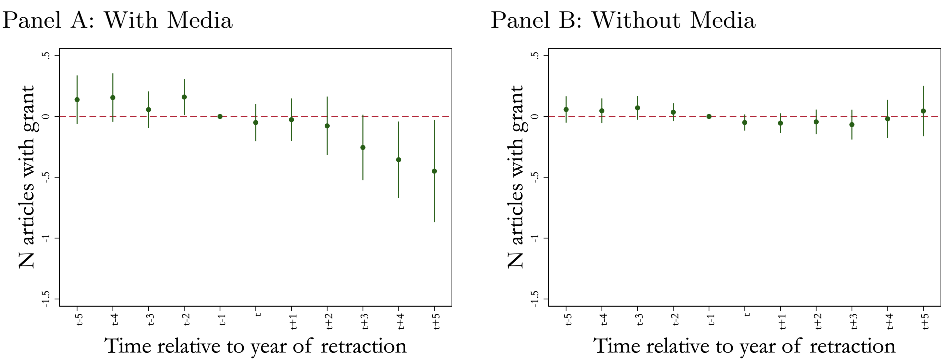
Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

Figure A.13: Author "productivity" by cause of retraction (with and without media)



Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

Figure A.14: Author grant supported "productivity" (with and without media)

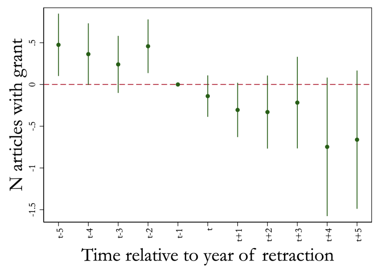


Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

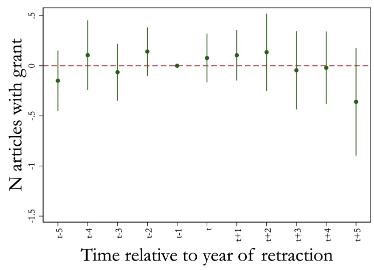
Figure A.15: Author grant supported "productivity" by rank (with and without media)

With Media

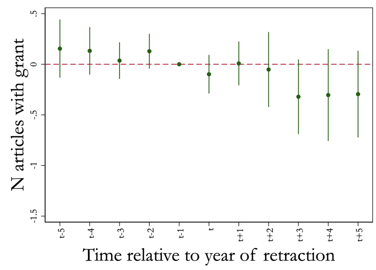
Panel A: First author



Panel B: Mid author

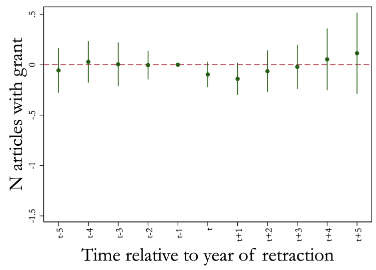


Panel C: Last author

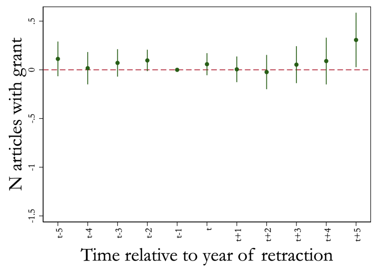


Without Media

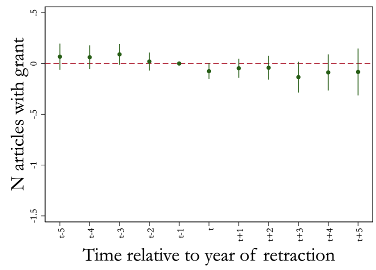
Panel D: First author



Panel E: Mid author

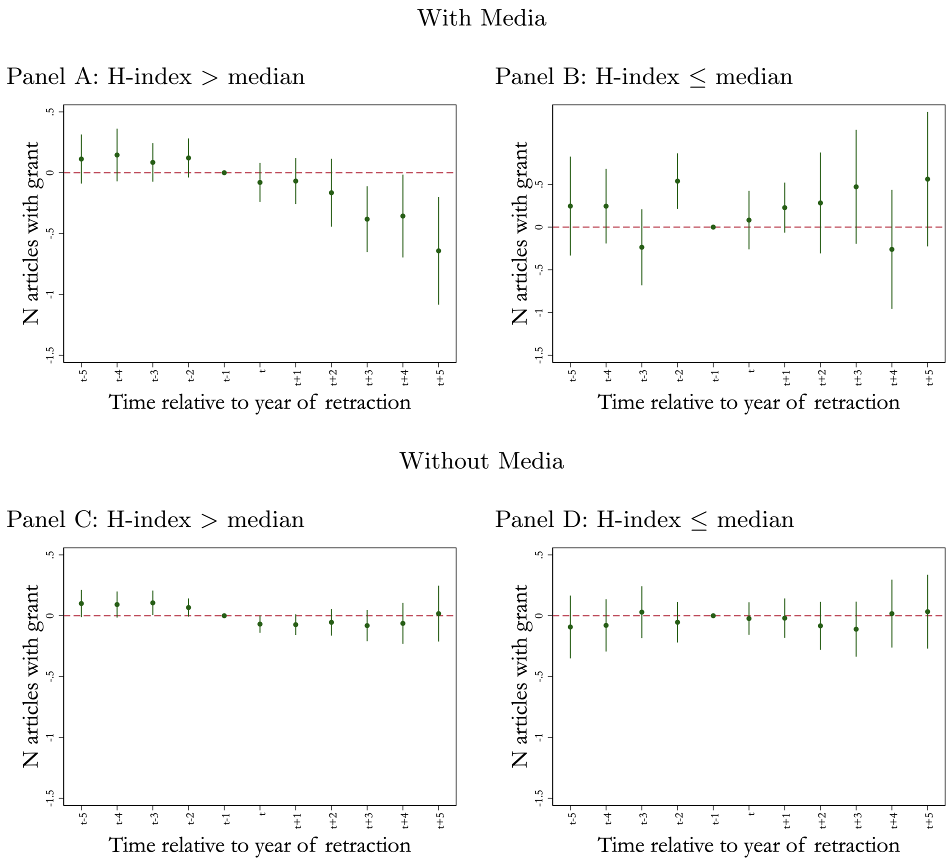


Panel F: Last author



Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

Figure A.16: Author grant supported "productivity" by seniority (with and without media)

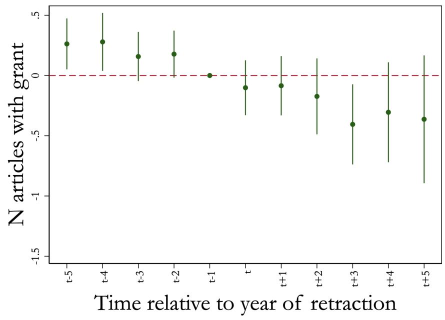


Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

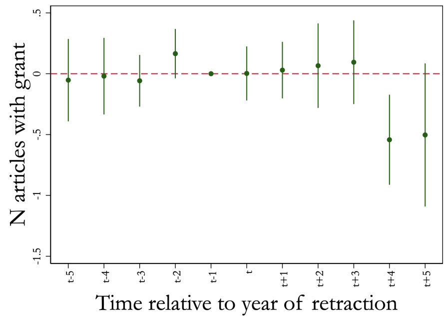
Figure A.17: Author grant supported "productivity" by cause of retraction (with and without media)

With Media

Panel A: Severe misconduct

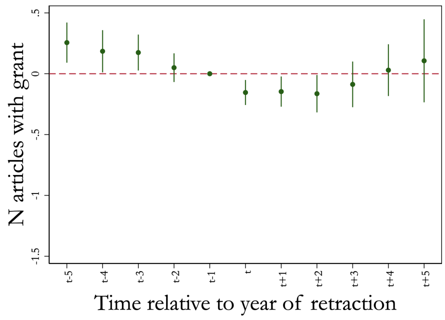


Panel B: Non-severe misconduct

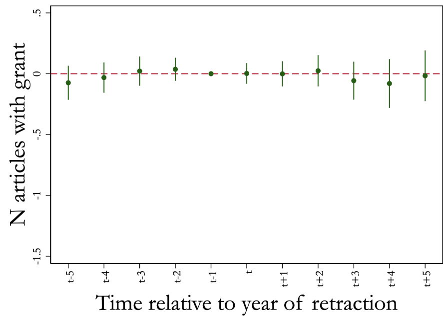


Without Media

Panel C: Severe misconduct



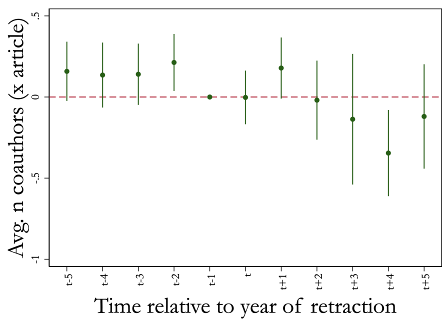
Panel D: Non-severe misconduct



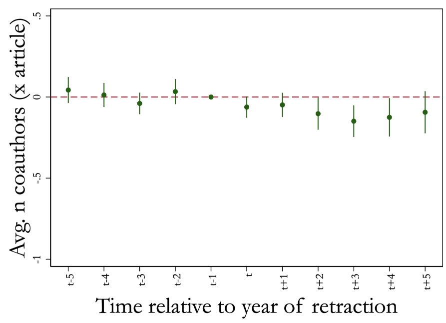
Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

Figure A.18: Author n. of coauthors (with and without media)

Panel A: With Media

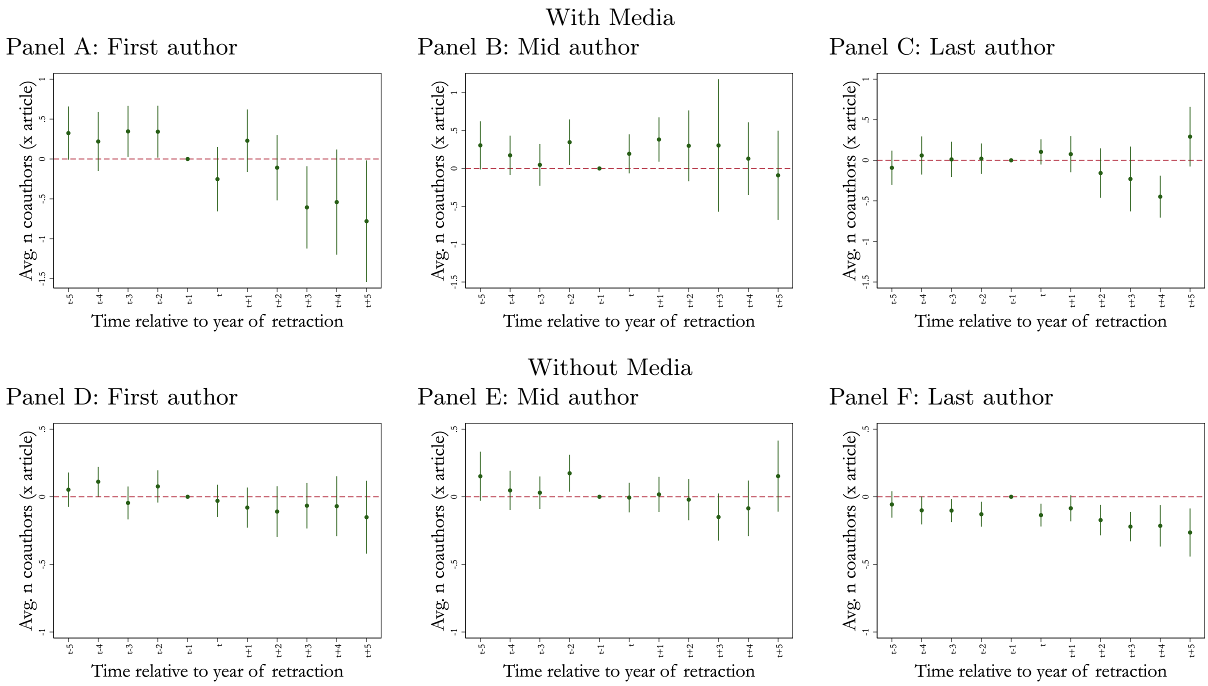


Panel B: Without Media



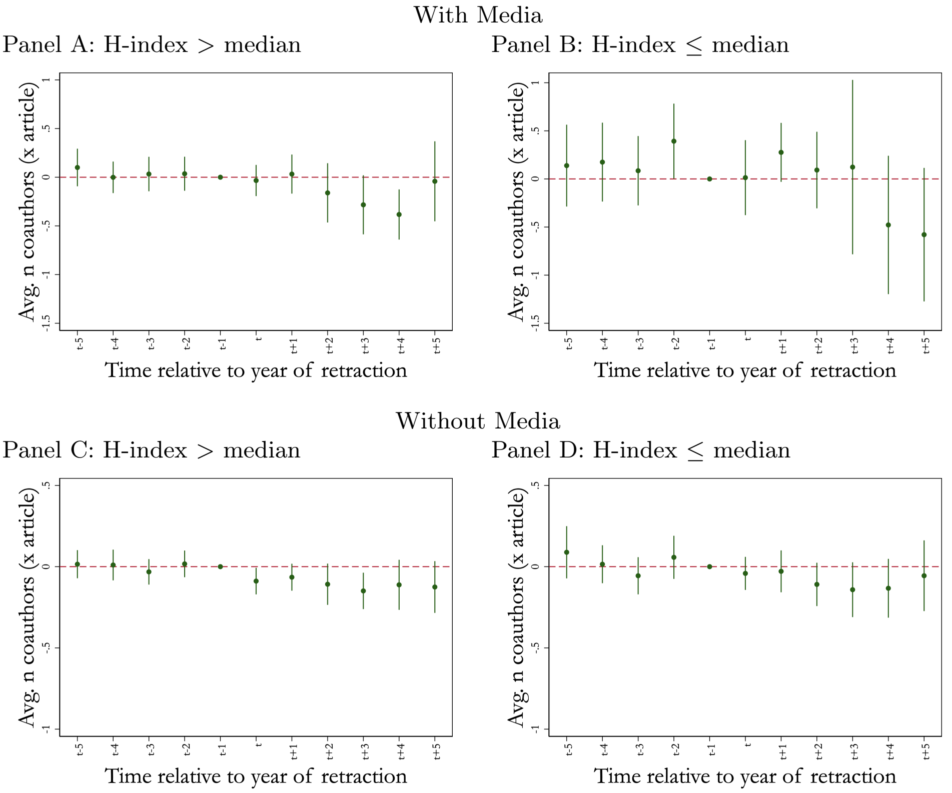
Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

Figure A.19: Author n. of coauthors by rank (with and without media)



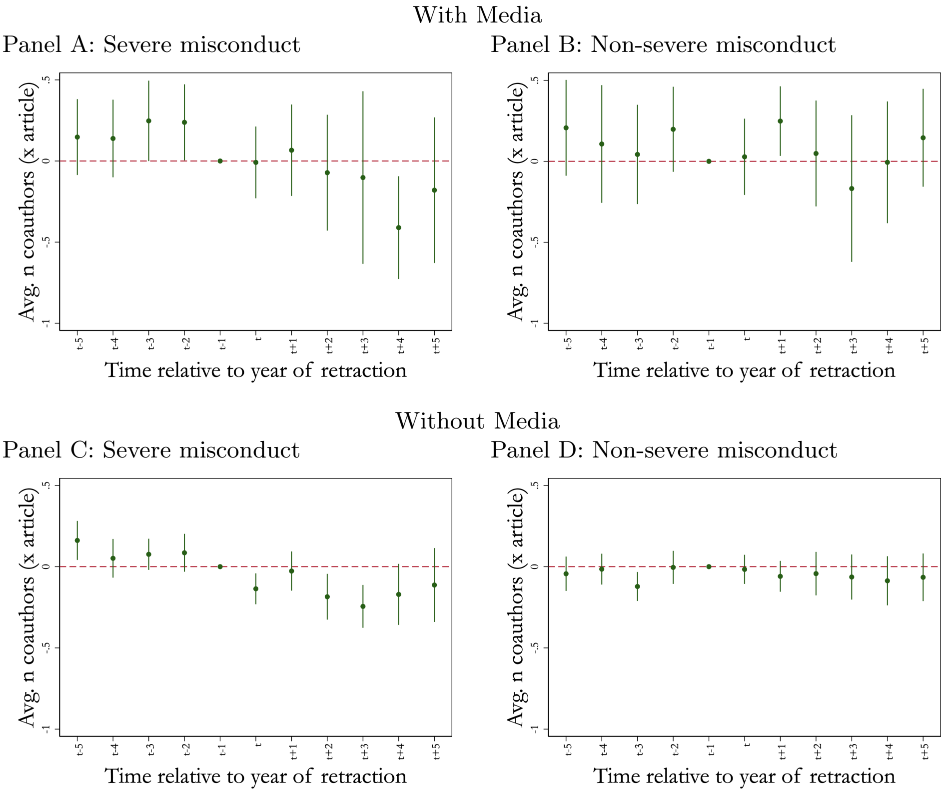
Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

Figure A.20: Author n. of coauthors by seniority (with and without media)



Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

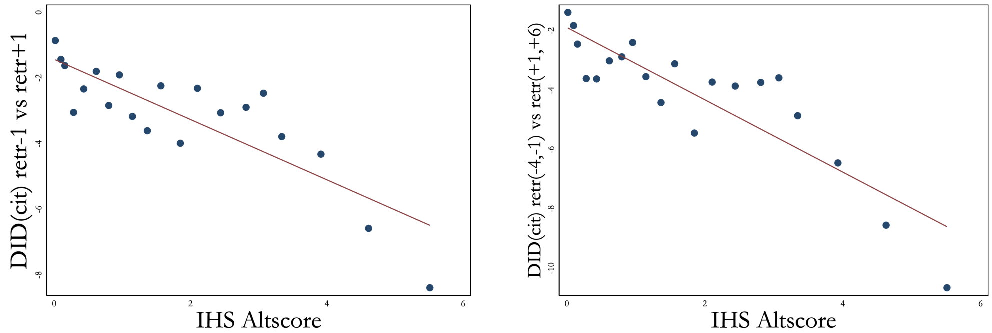
Figure A.21: Author n. of coauthors by cause of retraction (with and without media)



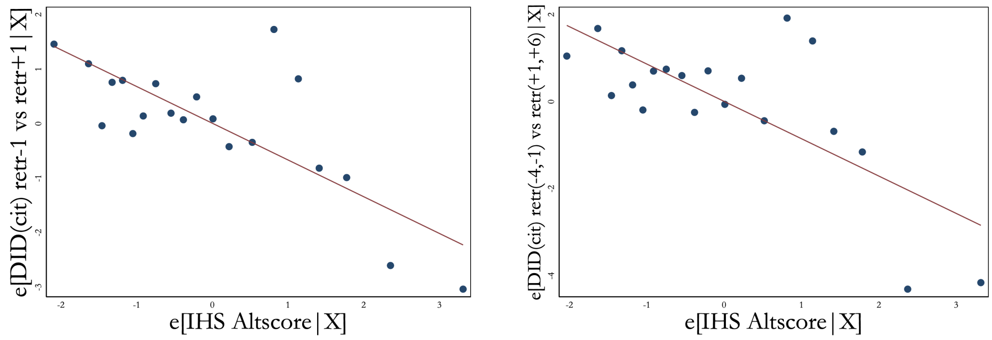
Note: Estimates compare publication rate for authors of retracted papers vs those of control papers before/after their first (in sample) retraction. Estimates replicate the models from Table A.32. Models are estimated replacing the *Post* indicator with a full set of dummies for each year relative to the retraction ($t-1$ excluded). Vertical lines represent 95% CI.

Figure A.22: Loss in citation and Controls average mentions

Panel A: Raw data



Panel B: Absorb Controls + FE



Note: The vertical axis represents the individual loss in citations obtained comparing each retracted paper to its selected controls for different pre- and post- time windows. The time window around retraction become larger moving left to right. The orixontal axis represents the inverse hyperbolic sine transformation of journal visibility, measured as the average Altscore of non-retracted papers that appear in the same yournal and year of the retracted one. Controls include the number of non-retracted articles within same journal and year of the treated, the average Euclidean distance of those from the retracted paper, and the level of (non-self) cumulative citations of the retracted paper before retraction. Publication year fixed effects are included.

A.11 Appendix Tables

Table A.1: Altscore weights

News	8
Blog	5
Policy document (per source)	3
Patent	3
Wikipedia	3
Twitter (tweets and retweets)	1
Peer review (Publons, Pubpeer)	1
Weibo (not trackable since 2015, but historical data kept)	1
Google+ (not trackable since 2019, but historical data kept)	1
F1000	1
Syllabi (Open Syllabus)	1
LinkedIn (not trackable since 2014, but historical data kept)	0.5
Facebook (only a curated list of public Pages)	0.25
Reddit	0.25
Pinterest (not trackable since 2013, but historical data kept)	0.25
Q&A (Stack Overflow)	0.25
Youtube	0.25
Number of Mendeley readers	0
Number of Dimensions and Web of Science citations	0

Table A.2: Author level statistics (within sub-samples)

	Media					No Media				
	(1) N	(2) mean	(3) sd	(4) min	(5) max	(6) N	(7) mean	(8) sd	(9) min	(10) max
Panel A: All										
N articles	6,100	5.932	7.770	0	69	49,016	5.407	8.488	0	170
N articles with grant	6,100	3.560	5.363	0	61	49,016	2.684	5.640	0	156
N of coauthors	6,100	7.827	9.041	0	100	49,016	5.568	7.200	0	100
Panel B: First author										
N articles	2,111	3.656	5.075	0	69	16,681	3.330	5.761	0	149
N articles with grant	2,111	2.147	3.537	0	47	16,681	1.588	4.127	0	149
N of coauthors	2,111	7.082	9.513	0	100	16,681	4.923	6.899	0	100
Panel C: Mid author										
N articles	1,877	4.977	7.378	0	69	14,921	4.442	7.141	0	129
N articles with grant	1,877	3.058	5.200	0	52	14,921	2.212	4.758	0	129
N of coauthors	1,877	7.920	9.353	0	100	14,921	5.696	7.882	0	100
Panel D: Last author										
N articles	2,112	9.055	9.195	0	68	17,414	8.225	10.70	0	170
N articles with grant	2,112	5.417	6.394	0	61	17,414	4.138	7.097	0	156
N of coauthors	2,112	8.488	8.184	0	82	17,414	6.075	6.814	0	100
Panel E: H-index above median										
N articles	3,813	8.290	8.822	0	69	25,009	8.320	10.37	0	170
N articles with grant	3,813	5.053	6.189	0	61	25,009	4.285	7.096	0	156
N of coauthors	3,813	9.037	9.485	0	100	25,009	6.902	7.943	0	100
Panel F: H-index below median										
N articles	2,287	2	2.560	0	27	24,007	2.374	4.134	0	149
N articles with grant	2,287	1.070	1.713	0	20	24,007	1.016	2.654	0	149
N of coauthors	2,287	5.808	7.845	0	100	24,007	4.178	6.028	0	100
Panel G: Severe cases of misconduct										
N articles	3,496	5.711	7.724	0	69	20,118	5.143	8.091	0	129
N articles with grant	3,496	3.632	5.490	0	61	20,118	2.706	5.823	0	129
N of coauthors	3,496	7.986	9.226	0	100	20,118	5.754	7.324	0	100
Panel H: Severe cases of misconduct										
N articles	2,604	6.228	7.822	0	56	28,898	5.591	8.748	0	170
N articles with grant	2,604	3.463	5.186	0	44	28,898	2.669	5.509	0	156
N of coauthors	2,604	7.613	8.785	0	100	28,898	5.438	7.109	0	100

Note: All statistics are reported by year. *N of articles* are yearly publications per author. *N of articles with grant* are yearly publications per author that mention any source of funding. *N of coauthors* is the average number of authors across papers published by an author within a year.

Table A.3: Retracted papers penalty

	Exponential	Exponential	OLS
	Citations	Citations	IHS(Citations)
Post	0.122***	0.115***	0.150***
	(0.024)	(0.025)	(0.025)
Post * Treatment	-1.067***	-1.064***	-0.830***
	(0.062)	(0.060)	(0.030)
Article FE	Y	Y	Y
Age FE	Y	Y	Y
Year FE	N	Y	Y
Pseudo R2 / R2	0.708	0.708	0.772
N	15438	15438	16679
N clusters	966	966	979
N full	16711	16711	16711

Note: First two columns show estimates of pseudo Poisson specifications while third column shows OLS estimation with IHS transformed dependent variable. The dependent variable is the total number of citations (exclusive of self-citations) received by each article in a particular year. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p < 0.10, **p < 0.05, ***p < 0.01.

Table A.4: Retracted papers penalty with any media coverage

	(1) Citations	(2) Citations	(3) Citations	(4) Citations
Post * Treatment	-0.831***	-0.798***	-0.944***	-0.840***
	(0.083)	(0.072)	(0.059)	(0.077)
Post * Treatment * Any social media	-0.325***			
	(0.119)			
Post * Treatment * Any news-blog		-0.434***		
		(0.119)		
Post * Treatment * Any news			-0.377***	
			(0.127)	
Post * Treatment * Any blog				-0.392***
				(0.129)
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.709	0.709	0.709	0.709
N	15438	15438	15438	15438
N clusters	966	966	966	966
N full	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Any socialmedia/news/blog is an indicator for papers with at least one overall mention in any of the indicated outlets. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p < 0.10, **p < 0.05, ***p < 0.01.

Table A.5: Retracted papers penalty and attention score

	(1) Citations	(2) Citations	(3) Citations	(4) Citations
Post * Treatment	-0.841***	-0.815***	-0.921***	
	(0.090)	(0.068)	(0.051)	
Post * Treatment * Altscore >p50	-0.283**			
	(0.117)			
Post * Treatment * Altscore >p75		-0.433***		
		(0.118)		
Post * Treatment * Altscore >p90			-0.488***	
			(0.150)	
Post * Treatment * Altscore 3rd quintile				-0.732***
				(0.091)
Post * Treatment * Altscore 4th quintile				-0.918***
				(0.083)
Post * Treatment * Altscore 5th quintile				-1.267***
				(0.088)
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.709	0.709	0.709	0.709
N	15438	15438	15438	15438
N clusters	966	966	966	966
N full	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Altscore is an aggregate measure of weighted online mentions. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.6: Retracted papers penalty with high cum. citations (pre-retraction)

	(1) Citations	(2) Citations	(3) Citations	(4) Citations	(5) Citations	(6) Citations
Post * Treatment	-1.121***	-1.057***	-0.996***	-1.008***	-0.954***	-0.900***
	(0.119)	(0.074)	(0.058)	(0.107)	(0.071)	(0.058)
Post * Treatment * Pre-retraction citations >p50	0.029			0.026		
	(0.140)			(0.126)		
Post * Treatment * Pre-retraction citations >p75		-0.045			-0.035	
		(0.122)			(0.115)	
Post * Treatment * Pre-retraction citations >p90			-0.226			-0.204
			(0.148)			(0.137)
Post * Treatment * Early mentions				-0.482***	-0.479***	-0.456***
				(0.163)	(0.160)	(0.155)
Article FE	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Pseudo R2	0.709	0.709	0.709	0.709	0.709	0.709
N	15438	15438	15438	15438	15438	15438
N clusters	966	966	966	966	966	966
N full	16711	16711	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Early mentions is an indicator for papers with at least one mention (in newspapers and/or blogs) within two weeks from publication. Pre-retraction citations are indicators for papers with relatively higher cumulative citations before the year of retraction. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.7: Retracted papers penalty with high cum. citations (pre-retraction)

	(1) Citations	(2) Citations	(3) Citations	(4) Citations	(5) Citations	(6) Citations
Post * Treatment	-0.921***	-0.838***	-0.752***	-0.896***	-0.842***	-0.767***
	(0.112)	(0.077)	(0.073)	(0.126)	(0.093)	(0.087)
Post * Treatment * Pre-retraction citations >p50	0.117			0.050		
	(0.134)			(0.134)		
Post * Treatment * Pre-retraction citations >p75		0.048			-0.006	
		(0.116)			(0.116)	
Post * Treatment * Pre-retraction citations >p90			-0.178			-0.220
			(0.129)			(0.136)
Post * Treatment * Any news-blog	-0.447***	-0.455***	-0.436***			
	(0.123)	(0.124)	(0.120)			
Post * Treatment * Any social media				-0.335***	-0.328***	-0.319***
				(0.121)	(0.120)	(0.121)
Article FE	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Pseudo R2	0.709	0.709	0.709	0.709	0.709	0.709
N	15438	15438	15438	15438	15438	15438
N clusters	966	966	966	966	966	966
N full	16711	16711	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Any socialmedia/news/blog is an indicator for papers with at least one overall mention in any of the indicated outlets. Pre-retraction citations are indicators for papers with relatively higher cumulative citations before the year of retraction. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.8: Retracted papers penalty with high cum. citations (pre-retraction)

	(1) Citations	(2) Citations	(3) Citations	(4) Citations	(5) Citations	(6) Citations
Post * Treatment	-0.924***	-0.854***	-0.771***	-1.009***	-0.948***	-0.885***
	(0.110)	(0.076)	(0.068)	(0.103)	(0.068)	(0.055)
Post * Treatment * Pre-retraction citations >p50	0.102			0.072		
	(0.133)			(0.128)		
Post * Treatment * Pre-retraction citations >p75		0.052			0.020	
		(0.115)			(0.112)	
Post * Treatment * Pre-retraction citations >p90			-0.173			-0.150
			(0.131)			(0.127)
Post * Treatment * Altscore >p75	-0.448***	-0.453***	-0.429***			
	(0.121)	(0.121)	(0.120)			
Post * Treatment * Altscore >p90				-0.508***	-0.503***	-0.480***
				(0.153)	(0.153)	(0.148)
Article FE	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Pseudo R2	0.709	0.709	0.709	0.709	0.709	0.709
N	15438	15438	15438	15438	15438	15438
N clusters	966	966	966	966	966	966
N full	16711	16711	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Altscore is an aggregate measure of weighted online mentions. Pre-retraction citations are indicators for papers with relatively higher cumulative citations before the year of retraction. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p < 0.10, **p < 0.05, ***p < 0.01.

Table A.9: Retracted papers penalty with early mentions ($Post\ t \geq 0$)

VARIABLES	(1) Citations	(2) Citations	(3) Citations
Post * Treatment	-0.653***	-0.668***	-0.657***
	(0.061)	(0.061)	(0.060)
Post * Treatment * Early mentions	-0.327**		
	(0.138)		
Post* Treatment * Early blog mentions		-0.268	
		(0.170)	
Post * Treatment * Early news mentions			-0.347**
			(0.139)
Article FE	Y	Y	Y
Age FE	Y	Y	Y
Year FE	Y	Y	Y
Pseudo R2	0.701	0.701	0.701
N	15438	15438	15438
N clusters	966	966	966
N full	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for the year of retraction and all subsequent years. Early mentions is an indicator for papers with at least one mention (in newspapers and/or blogs) within two weeks from publication. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p < 0.10, **p < 0.05, ***p < 0.01.

Table A.10: Retracted papers penalty with any media coverage ($Post\ t \geq 0$)

VARIABLES	(1) Citations	(2) Citations	(3) Citations	(4) Citations
Post * Treatment	-0.465***	-0.465***	-0.608***	-0.527***
	(0.105)	(0.092)	(0.065)	(0.088)
Post * Treatment * Any social media	-0.346***			
	(0.128)			
Post * Treatment * Any news-blog		-0.414***		
		(0.126)		
Post * Treatment * Any news			-0.355***	
			(0.115)	
Post * Treatment * Any blog				-0.339***
				(0.129)
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.701	0.702	0.701	0.701
N	15438	15438	15438	15438
N clusters	966	966	966	966
N full	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for the year of retraction and all subsequent years. Any socialmedia/news/blog is an indicator for papers with at least one overall mention in any of the indicated outlets. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p < 0.10, **p < 0.05, ***p < 0.01.

Table A.11: Retracted papers penalty and attention score ($Post\ t \geq 0$)

VARIABLES	(1) Citations	(2) Citations	(3) Citations	(4) Citations
Post * Treatment	-0.452***	-0.467***	-0.572***	
	(0.118)	(0.083)	(0.058)	
Post * Treatment * Altscore >p50	-0.339**			
	(0.133)			
Post * Treatment * Altscore >p75		-0.439***		
		(0.118)		
Post * Treatment * Altscore >p90			-0.554***	
			(0.127)	
Post * Treatment * Altscore 3rd quintile				-0.521***
				(0.089)
Post * Treatment * Altscore 4th quintile				-0.607***
				(0.082)
Post * Treatment * Altscore 5th quintile				-0.919***
				(0.072)
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.701	0.702	0.702	0.702
N	15438	15438	15438	15438
N clusters	966	966	966	966
N full	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for the year of retraction and all subsequent years. Altscore is an aggregate measure of weighted online mentions. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.12: Retracted papers penalty with early mentions (actively cited papers)

	(1) Citations	(2) Citations	(3) Citations
Post * Treatment	-1.036***	-1.078***	-1.064***
	(0.089)	(0.086)	(0.090)
Post * Treatment * Early mentions	-0.428**		
	(0.198)		
Post * Treatment * Early blog mentions		-0.362	
		(0.238)	
Post * Treatment * Early news mentions			-0.399**
			(0.165)
Article FE	Y	Y	Y
Age FE	Y	Y	Y
Year FE	Y	Y	Y
Pseudo R2	0.733	0.733	0.733
N	7308	7308	7308
N clusters	466	466	466
N full	7662	7662	7662

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Early mentions is an indicator for papers with at least one mention (in newspapers and/or blogs) within two weeks from publication. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.13: Retracted papers penalty with any media coverage (actively cited papers)

	(1) Citations	(2) Citations	(3) Citations	(4) Citations
Post * Treatment	-0.739***	-0.785***	-1.010***	-0.892***
	(0.124)	(0.113)	(0.093)	(0.113)
Post * Treatment * Any social media	-0.577***			
	(0.162)			
Post * Treatment * Any news-blog		-0.618***		
		(0.162)		
Post * Treatment * Any news			-0.387***	
			(0.139)	
Post * Treatment * Any blog				-0.494***
				(0.171)
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.733	0.734	0.733	0.733
N	7308	7308	7308	7308
N clusters	466	466	466	466
N full	7662	7662	7662	7662

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Any socialmedia/news/blog is an indicator for papers with at least one overall mention in any of the indicated outlets. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.14: Retracted papers penalty and attention score (actively cited papers)

	(1) Citations	(2) Citations	(3) Citations	(4) Citations
Post * Treatment	-0.704***	-0.805***	-0.967***	
	(0.132)	(0.106)	(0.080)	
Post * Treatment * Altscore >p50	-0.583***			
	(0.163)			
Post * Treatment * Altscore >p75		-0.590***		
		(0.157)		
Post * Treatment * Altscore >p90			-0.544***	
			(0.175)	
Post * Treatment * Altscore 3rd quintile				-0.769***
				(0.149)
Post * Treatment * Altscore 4th quintile				-1.123***
				(0.119)
Post * Treatment * Altscore 5th quintile				-1.317***
				(0.102)
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.733	0.733	0.734	0.733
N	7308	7308	7308	7308
N clusters	466	466	466	466
N full	7662	7662	7662	7662

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Altscore is an aggregate measure of weighted online mentions. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.15: Retracted papers penalty with early mentions (published in 2011-2017)

VARIABLES	(1) Citations	(2) Citations	(3) Citations
Post*Treatment	-0.989***	-1.015***	-1.008***
	(0.060)	(0.059)	(0.060)
Post*Treatment*Early mentions	-0.438***		
	(0.165)		
Post*Treatment*Early blog mentions		-0.368*	
		(0.194)	
Post*Treatment*Early news mentions			-0.399**
			(0.155)
Article FE	Y	Y	Y
Age FE	Y	Y	Y
Year FE	Y	Y	Y
Pseudo R2	0.710	0.710	0.710
N	14194	14194	14194
N clusters	776	776	776
N full	15146	15146	15146

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for the year of retraction and all subsequent years. Early mentions is an indicator for papers with at least one mention (in newspapers and/or blogs) within two weeks from publication. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.16: Retracted papers penalty with any media coverage (published in 2011-2017)

VARIABLES	(1) Citations	(2) Citations	(3) Citations	(4) Citations
Post*Treatment	-0.837***	-0.803***	-0.959***	-0.878***
	(0.086)	(0.078)	(0.062)	(0.081)
Post*Treatment*Any social media	-0.359***			
	(0.123)			
Post*Treatment*Any news-blog		-0.459***		
		(0.123)		
Post*Treatment*Any news			-0.415***	
			(0.126)	
Post*Treatment*Any blog				-0.360***
				(0.134)
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.710	0.710	0.710	0.710
N	14194	14194	14194	14194
N clusters	776	776	776	776
N full	15146	15146	15146	15146

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for the year of retraction and all subsequent years. Any socialmedia/news/blog is an indicator for papers with at least one overall mention in any of the indicated outlets. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.17: Retracted papers penalty and attention score (published in 2011-2017)

VARIABLES	(1) Citations	(2) Citations	(3) Citations	
Post*Treatment	-0.835*** (0.095)	-0.820*** (0.073)	-0.936*** (0.054)	
Post*Treatment*Altscore >p50	-0.325*** (0.122)			
Post*Treatment*Altscore >p75		-0.463*** (0.121)		
Post*Treatment*Altscore >p90			-0.533*** (0.150)	
Post*Treatment*Altscore 3rd quintile				-0.681*** (0.092)
Post*Treatment*Altscore 4th quintile				-0.940*** (0.086)
Post*Treatment*Altscore 5th quintile				-1.308*** (0.087)
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.710	0.710	0.710	0.711
N	14194	14194	14194	14194
N clusters	776	776	776	776
N full	15146	15146	15146	15146

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for the year of retraction and all subsequent years. Altscore is an aggregate measure of weighted online mentions. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.18: Retracted papers penalty and attention score by discipline

	(1)	(2)	(3) Hard sciences	(4) Social sciences
	Citations	Citations	Citations	Citations
Post * Treatment * Altscore >p50	1.333**	0.167	-0.321**	0.150
	(0.622)	(0.254)	(0.127)	(0.264)
Post * Treatment * Altscore >p50 * Business/Technology	-1.307*			
	(0.680)			
Post * Treatment * Altscore >p50 * Life sciences	-1.954***			
	(0.650)			
Post * Treatment * Altscore >p50 * Environment	-1.197*			
	(0.717)			
Post * Treatment * Altscore >p50 * Health	-0.948			
	(0.692)			
Post * Treatment * Altscore >p50 * Physics	-1.485**			
	(0.649)			
Post * Treatment * Altscore >p50 * Hard sciences		-0.494*		
		(0.283)		
Article FE	Y	Y	Y	Y
Age FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Pseudo R2	0.710	0.709	0.718	0.595
N	15399	15438	12980	2419
N clusters	964	966	798	166
N full	16672	16711	13837	2835

Note: Estimates derive from pseudo Poisson specifications. Hard sciences: life sciences, environment, health and physical sciences. Social sciences: business and technology, humanities, other social sciences. The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Altscore is an aggregate measure of weighted online mentions. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.19: Retracted papers penalty and attention score by severity of misconduct

	All	Minor	Moderate	Severe
	Citations	Citations	Citations	Citations
P * T * Altscore >p50	0.013	-0.023	0.381**	-0.519***
	(0.174)	(0.165)	(0.182)	(0.172)
P * T * Altscore >p50 * Moderate misconduct	0.361			
	(0.249)			
P * T * Altscore >p50 * Severe misconduct	-0.547**			
	(0.245)			
Article, Age & Year FE	Y	Y	Y	Y
Pseudo R2	0.710	0.594	0.694	0.738
N (N clusters)	15438 (966)	4312 (295)	3857 (256)	7269 (415)
N full	16711	4859	4157	7695

Note: Estimates derive from pseudo Poisson specifications. Causes of retractions are classified based on Woo and Walsh (2021). The dependent variable is the total number of citations (exclusive of self-citations) received by each paper in a particular year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Altscore is an aggregate measure of weighted online mentions. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. x% loss in yearly citations). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.20: Misconduct classification from Woo and Walsh (2021)

Classification	Coded Retracted Reasons from Retraction Watch database
Minor misconducts	'Withdrawal', 'Author Unresponsive', 'Breach of Policy by Author', 'Breach of Policy by Third Party', 'Complaints about Author', 'Complaints about Company/Institution', 'Complaints about Third Party', 'Civil Proceedings', 'Concerns/Issues about Referencing/Attributions', 'Concerns/Issues about Third Party Involvement', 'Concerns/Issues About Authorship', 'Conflict of Interest', 'Copyright Claims', 'Criminal Proceedings', 'Date of Retraction/Other Unknown', 'Doing the Right Thing', 'Duplication of Article', 'Duplication of Data', 'Duplication of Image', 'Duplication of Text', 'Error by Journal/Publisher', 'Error by Third Party', 'Ethical Violations by Author', 'Ethical Violations by Third Party', 'Euphemisms for Duplication', 'Euphemisms for Misconduct', 'Euphemisms for Plagiarism', 'Informed/Patient Consent – None/Withdrawn', 'Lack of Approval from Author', 'Lack of Approval from Company/Institution', 'Lack of Approval from Third Party', 'Lack of IRB/IACUC Approval', 'Legal Reasons/Legal Threats', 'No Further Action', 'Nonpayment of Fees/Refusal to Pay', 'Notice – Lack of', 'Notice – Limited or No Information', 'Notice – Unable to Access via current resources', 'Objections by Author(s)', 'Objections by Company/Institution', 'Objections by Third Party', 'Plagiarism of Article', 'Plagiarism of Data', 'Plagiarism of Image', 'Plagiarism of Text', 'Publishing Ban', 'Retract and Replace', 'Salami Slicing', 'Temporary Removal', 'Updated to Correction', 'Updated to Retraction', 'Upgrade/Update of Prior Notice', 'Cites Prior Retracted Work'
Moderate misconducts	'Concerns/Issues About Authorship', 'Concerns/Issues About Data', 'Concerns/Issues About Image', 'Concerns/Issues About Results', 'Contamination of Cell Lines/Tissues', 'Contamination of Materials (General)', 'Contamination of Reagents', 'Error in Analyses', 'Error in Cell Lines/Tissues', 'Error in Data', 'Error in Image', 'Error in Materials (General)', 'Error in Methods', 'Error in Results and/or Conclusions', 'Error in Text', 'Investigation by Company/Institution', 'Investigation by Journal/Publisher', 'Investigation by ORI', 'Investigation by Third Party', 'Lack Of Balance/Bias Issues', 'Miscommunication by Author', 'Miscommunication by Company/Institution', 'Miscommunication by Journal/Publisher', 'Miscommunication by Third Party', 'Misconduct by Third Party'
Severe misconducts	'Fake Peer Review', 'Falsification/Fabrication of Data', 'Falsification/Fabrication of Image', 'Falsification/Fabrication of Results', 'Forged Authorship', 'Hoax Paper', 'Manipulation of Images', 'Manipulation of Results', 'Misconduct – Official Investigation/Finding', 'Misconduct by Author', 'Misconduct by Company/Institution', 'Results Not Reproducible', 'Sabotage of Materials', 'Sabotage of Methods', 'Unreliable Data', 'Unreliable Image', 'Unreliable Results'

Table A.21: Selected summary statistics: title ngrams

Most frequent (selected) ngrams					Relevant selected ngrams				
	Mean	Sd	Min	Max		Mean	Sd	Min	Max
Treatment	0.0477	0.213	0	1	# of adult	0.0070	0.084	0	2
Media	0.0945	0.293	0	1	# of algorithm	0.0113	0.107	0	2
Citations $year_p$	0.901	3.217	0	249	# of brain	0.009	0.097	0	1
Total number of words	14.28	5.047	1	54	# of climat	0.005	0.071	0	2
# of base	0.0773	0.272	0	2	# of commun	0.0080	0.091	0	2
# of effect	0.0666	0.252	0	2	# of composit	0.0201	0.142	0	2
# of studi	0.0543	0.228	0	2	# of disord	0.0064	0.083	0	2
# of model	0.0517	0.225	0	2	# of earli	0.0075	0.086	0	1
# of analysi	0.0453	0.209	0	2	# of genom	0.0100	0.101	0	2
# of system	0.0380	0.195	0	2	# of global	0.0065	0.082	0	2
# of induc	0.0314	0.176	0	2	# of graphen	0.0093	0.101	0	3
# of imag	0.0291	0.172	0	3	# of meta_analysi	0.0050	0.070	0	1
# of human	0.0270	0.165	0	2	# of model	0.0517	0.225	0	2
# of perform	0.0256	0.159	0	2	# of network_ETX	0.0082	0.090	0	1
# of mechan	0.0252	0.158	0	2	# of neuron	0.0052	0.075	0	2
# of properti	0.0244	0.155	0	2	# of reveal	0.0092	0.096	0	1
# of oxid	0.0242	0.163	0	3	# of risk	0.0154	0.127	0	3
# of enhanc	0.0238	0.153	0	2	# of stem	0.0095	0.100	0	2
# of regul	0.0238	0.154	0	2	# of STX_structur	0.0057	0.076	0	1
# of associ	0.0236	0.155	0	2	# of trial	0.0108	0.104	0	2
# of respons	0.0233	0.153	0	2	# of vitro	0.0064	0.080	0	1

Note: N-grams represent the number of times the selected espression appears in the title of a research article. All n-grams in the table were selected by one of the lasso procedures. N=20755.

Table A.22: Selected words and media coverage

	Media coverage							
	Linear				Logit			
	EBIC	AICC	CV	Rigourous	EBIC	AICC	CV	Rigourous
# of adult	0.110***	0.093***	0.092***		0.067***	0.044**	0.045**	0.073***
	(0.036)	(0.035)	(0.035)		(0.018)	(0.017)	(0.017)	(0.017)
# of algorithm		-0.047***	-0.048***	-0.049***		-0.186**	-0.185**	
		(0.008)	(0.008)	(0.007)		(0.076)	(0.076)	
# of brain	0.081***	0.068**	0.066**		0.062***	0.041**	0.041**	
	(0.026)	(0.027)	(0.027)		(0.015)	(0.017)	(0.017)	
# of climat	0.228***	0.221***	0.219***		0.134***	0.120***	0.126***	0.132***
	(0.054)	(0.050)	(0.050)		(0.022)	(0.020)	(0.020)	(0.021)
# of commun	0.103***	0.091***	0.089***		0.073***	0.061***	0.061***	0.074***
	(0.034)	(0.034)	(0.035)		(0.019)	(0.020)	(0.020)	(0.019)
# of composit	-0.046***	-0.036***	-0.035***	-0.051***	-0.124***	-0.098***	-0.099***	-0.127***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.034)	(0.033)	(0.033)	(0.034)
# of disord	0.170***	0.162***	0.162***		0.089***	0.080***	0.081***	0.094***
	(0.058)	(0.057)	(0.057)		(0.022)	(0.021)	(0.021)	(0.023)
# of earli	0.090***	0.078**	0.077**		0.066***	0.057***	0.056***	
	(0.034)	(0.033)	(0.033)		(0.019)	(0.019)	(0.019)	
# of genom	0.076**	0.065**	0.065**		0.045**	0.037**	0.036**	0.044**
	(0.033)	(0.033)	(0.033)		(0.017)	(0.017)	(0.017)	(0.017)
# of global		0.070**	0.071**			0.049***	0.049***	
		(0.033)	(0.033)			(0.019)	(0.019)	
# of graphen		0.076**	0.075**			0.087***	0.087***	
		(0.031)	(0.031)			(0.022)	(0.022)	
# of meta_analysi	0.098**	0.120***	0.120***		0.061***	0.080***	0.061***	
	(0.039)	(0.040)	(0.042)		(0.019)	(0.027)	(0.022)	
# of model	-0.043***	-0.036***	-0.037***	-0.044***	-0.074***	-0.065***	-0.064***	-0.075***
	(0.006)	(0.007)	(0.007)	(0.007)	(0.016)	(0.016)	(0.016)	(0.016)
# of network_ETX		-0.063***	-0.063***	-0.061***	-0.183**	-0.181**	-0.181**	
		(0.011)	(0.011)	(0.010)	(0.081)	(0.080)	(0.080)	
# of neuron	0.082**	0.070*	0.069*		0.056***	0.041*	0.039*	
	(0.038)	(0.038)	(0.038)		(0.020)	(0.021)	(0.021)	
# of reveal	0.110***	0.102***	0.102***		0.070***	0.065***	0.059***	0.070***
	(0.029)	(0.029)	(0.029)		(0.016)	(0.015)	(0.015)	(0.016)
# of risk	0.080***	0.068**	0.067**		0.054***	0.034**	0.034**	0.056***
	(0.028)	(0.027)	(0.028)		(0.015)	(0.015)	(0.015)	(0.015)
# of stem	0.092***	0.081***	0.084***		0.060***	0.051***	0.050***	0.059***
	(0.030)	(0.030)	(0.030)		(0.019)	(0.019)	(0.019)	(0.019)
# of STX_structur	0.118***	0.113***	0.112***		0.079***	0.081***	0.079***	
	(0.035)	(0.035)	(0.034)		(0.019)	(0.018)	(0.018)	
# of trial	0.105**	0.074	0.075		0.074***	0.048	0.048	0.078***
	(0.050)	(0.054)	(0.054)		(0.027)	(0.032)	(0.032)	(0.027)
# of vitro		-0.052***	-0.054***	-0.069***		-0.182***	-0.180***	
		(0.009)	(0.009)	(0.008)		(0.058)	(0.058)	
Total # of words	-0.006***	-0.005***	-0.004***	-0.004***	-0.006***	-0.004***	-0.004***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Citations $year_p$	0.018***	0.017***	0.017***	0.018***	0.014***	0.013***	0.013***	0.014***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)
N	20755	20755	20755	20755	20755	20755	20755	20755
Out-of-sample R2	0.095	0.101	0.100	0.077				
R-squared	0.088	0.114	0.115	0.074				
Out-of-sample accuracy	63.47	61.88	60.10	71.08	66.84	63.76	67.37	73.49
Overall accuracy	70.50	63.92	62.25	65.26	76.37	69.36	62.53	71.68
Best cutoff	0.091	0.094	0.091	0.105	0.088	0.080	0.090	0.097
Matthew corr. coeff.	0.373	0.416	0.408	0.378	0.386	0.430	0.449	0.414

Note: Estimates from OLS (columns 1-4) or Logit regression (columns 5-8). The dependent variable *Media coverage* is an indicator for whether a paper attracted online coverage at publication. In column (1)-(4) predictors are selected based on *lasso* while column (5)-(8) predictors are selected based on *lassologit*. Matthews Correlation Coefficient (MCC) is the preferred single metric in the area machine learning with binary classification, especially for imbalanced data. The metric ranges $[-1, 1]$ and takes on the value of zero if the prediction is the same as a random guess. Best cutoff pproximates the optimal positive cutoff using MCC. Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.23: Likelihood of retraction and media coverage (Logit)

	Retraction									
	Logit		EBIC		AICC		CV		Rigourous	
Media coverage	0.009**	0.011***	0.014***	0.016***	0.015***	0.017***	0.015***	0.017***	0.014***	0.016***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)
Predicted media, Prob.			-0.089***	-0.083**	-0.054***	-0.052***	-0.058***	-0.055***	-0.095**	-0.089**
			(0.034)	(0.034)	(0.017)	(0.017)	(0.018)	(0.019)	(0.038)	(0.039)
Publication year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Subject FE	N	Y	N	Y	N	Y	N	Y	N	Y
N	20755	20393	20755	20393	20755	20393	20755	20393	20755	20393

Note: Estimates from Logit equation. The dependent variable *Retraction* is an indicator for whether a paper was retracted. *Media coverage* is an indicator for whether a paper attracted online coverage at publication. *Predicted media* is media coverage as predicted from the respective *lassologit* procedures. Bootstrap standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.24: Likelihood of retraction and media coverage (Residuals)

	Retraction							
	EBIC		AICC		CV		Rigourous	
Predicted media, Resid.	0.013***	0.022***	0.015***	0.024***	0.015***	0.024***	0.011**	0.021***
	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)
R-squared	0.000	0.002	0.000	0.002	0.000	0.002	0.000	0.001
Publication year FE	Y	Y	Y	Y	Y	Y	Y	Y
Journal FE	N	Y	N	Y	N	Y	N	Y
N	20755	20755	20755	20755	20755	20755	20755	20755
N clusters	1008	1008	1008	1008	1008	1008	1008	1008

Note: Estimates from OLS regression. The dependent variable *Retraction* is an indicator for whether a paper was retracted. *Media coverage*, *Resid.* is the residual of the predicted coverage according to different *lassologit* procedures. Bootstrap standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.25: Likelihood of retraction and media coverage (tf-idf)

	Retraction							
	EBIC		AICC		CV		Rigourous	
Media coverage	0.011**	0.020***	0.013***	0.020***	0.013***	0.020***	0.011**	0.020***
	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)
Predicted media	-0.040	-0.032	-0.039**	-0.035*	-0.039**	-0.035*	-0.072	-0.061
	(0.027)	(0.030)	(0.016)	(0.020)	(0.017)	(0.020)	(0.044)	(0.055)
R-squared	0.000	0.001	0.000	0.002	0.000	0.001	0.000	0.001
Publication year FE	Y	Y	Y	Y	Y	Y	Y	Y
Journal FE	N	Y	N	Y	N	Y	N	Y
N	20755	20755	20755	20755	20755	20755	20755	20755
N clusters	1008	1008	1008	1008	1008	1008	1008	1008

Note: Estimates from OLS regression. The dependent variable *Retraction* is an indicator for whether a paper was retracted. *Media coverage* is an indicator for whether a paper attracted online coverage at publication. *Predicted media* is media coverage as predicted from the respective *lasso* procedures. Bootstrap standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.26: Likelihood of retraction and media coverage (selection within subjects, publication years and excluding retractions)

	Retraction							
	EBIC		AICC		CV		Rigourous	
Media coverage	0.016***	0.021***	0.018***	0.022***	0.018***	0.022***	0.014***	0.020***
	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)
Predicted media	-0.056***	-0.071**	-0.061***	-0.074***	-0.062***	-0.074***	-0.047***	-0.084**
	(0.018)	(0.032)	(0.016)	(0.025)	(0.016)	(0.025)	(0.015)	(0.041)
R-squared	0.001	0.001	0.001	0.001	0.001	0.002	0.000	0.001
Publication year FE	Y	Y	Y	Y	Y	Y	Y	Y
Journal FE	N	Y	N	Y	N	Y	N	Y
N	20393	20393	20393	20393	20393	20393	20393	20393
N clusters	988	988	988	988	988	988	988	988

Note: Estimates from OLS regression. The dependent variable *Retraction* is an indicator for whether a paper was retracted. *Media coverage* is an indicator for whether a paper attracted online coverage at publication. *Predicted media* is media coverage as predicted from the respective *lasso* procedures. Bootstrap standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.27: Months to retraction and Journal-year average visibility

	(1) Time to Retract	(2) Time to Retract	(3) Time to Retract	(4) Time to Retract	(5) Time to Retract	(6) Time to Retract
Altscore	-3.579***					
	(0.565)					
Sh. Blog		-3.376				
		(2.413)				
Blog count		-1.662				
		(1.850)				
Sh. news			-6.287***			
			(1.937)			
News count			2.105			
			(1.384)			
Sh. Tweets				-3.127***		
				(1.117)		
Tweets count				-1.688***		
				(0.645)		
Sh. early blog					-3.331*	
					(1.995)	
Early blog count					-0.423	
					(1.675)	
Sh. early news						-5.453***
						(1.666)
Early news count						1.931
						(1.262)
Observations	967	961	962	968	962	967
R-squared	0.455	0.468	0.468	0.460	0.459	0.455
Publication year FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Mean	24.21	24.21	24.21	24.21	24.21	24.21

Note: The dependent variable is the time intercurring between an article publication and its retraction, expressed in months. Covariates represents different measures of journal visibility, measured as the average of non-retracted papers that appear in the same yournal and year of the retracted one. All covariates are standardized and outliers trimmed. Controls include the number of non-retracted articles within same journal and year of the treated, the average Euclidean distance of those from the retracted paper, and the level of (non-self) cumulative citations of the retracted paper before retraction. Publication year fixed effects are included. Journal clustered standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A.28: Loss in citation and Journal-year average visibility

	(1) DID (-1,1)	(2) DID (-1,1)	(3) DID (-2,2)	(4) DID (-2,2)	(5) DID (-4,4)	(6) DID (-4,4)	(7) DID (-4,6)	(8) DID (-4,6)
Altscore	-1.362***	-1.111***	-1.699***	-1.327***	-1.830***	-1.399***	-1.877***	-1.439***
	(0.187)	(0.221)	(0.201)	(0.202)	(0.224)	(0.221)	(0.236)	(0.232)
Observations	840	840	840	840	840	840	840	840
R-squared	0.120	0.171	0.165	0.230	0.158	0.236	0.152	0.231
Sh. Blog	-1.698**	-1.643**	-1.598**	-1.440**	-1.746**	-1.548**	-1.658**	-1.449**
	(0.738)	(0.686)	(0.740)	(0.677)	(0.767)	(0.687)	(0.764)	(0.686)
Blog count	0.078	0.268	-0.327	-0.088	-0.334	-0.066	-0.467	-0.198
	(0.651)	(0.627)	(0.601)	(0.562)	(0.614)	(0.562)	(0.615)	(0.567)
Observations	831	831	831	831	831	831	831	831
R-squared	0.139	0.188	0.182	0.241	0.175	0.246	0.167	0.239
Sh. News	-1.611***	-1.365***	-1.795***	-1.325**	-1.952***	-1.384**	-1.995***	-1.406**
	(0.519)	(0.501)	(0.582)	(0.559)	(0.633)	(0.599)	(0.650)	(0.611)
News count	0.129	0.107	-0.018	-0.128	-0.051	-0.194	-0.059	-0.214
	(0.446)	(0.421)	(0.480)	(0.461)	(0.515)	(0.489)	(0.525)	(0.495)
Observations	835	835	835	835	835	835	835	835
R-squared	0.125	0.173	0.169	0.228	0.167	0.238	0.161	0.233
Sh. Tweets	-0.650**	-0.540**	-0.764***	-0.558**	-0.856***	-0.601**	-0.894***	-0.627**
	(0.268)	(0.262)	(0.275)	(0.257)	(0.295)	(0.270)	(0.298)	(0.273)
Tweets count	-0.965***	-0.788***	-1.185***	-0.944***	-1.233***	-0.966***	-1.254***	-0.986***
	(0.222)	(0.238)	(0.251)	(0.235)	(0.286)	(0.261)	(0.302)	(0.275)
Observations	842	842	842	842	842	842	842	842
R-squared	0.123	0.173	0.164	0.229	0.156	0.234	0.149	0.228
Pub. year FE	Y	Y	Y	Y	Y	Y	Y	Y
Age at retraction FE	N	Y	N	Y	N	Y	N	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Mean	-3.149	-3.149	-3.707	-3.707	-4.274	-4.274	-4.158	-4.158

Note: The dependent variable is the individual loss in citations obtained comparing each retracted paper to its selected controls for different pre- and post- time windows. Covariates represents different measures of journal visibility, measured as the average of non-retracted papers that appear in the same yournal and year of the retracted one. All covariates are standardized and outliers trimmed. Controls include the number of non-retracted articles within same journal and year of the treated, the average Euclidean distance of those from the retracted paper, and the level of (non-self) cumulative citations of the retracted paper before retraction. Publication year fixed effects are included. Fixed effects for age of the article at retraction are added in even comuns. Journal clustered standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A.29: Citation statements and early mentions

	(1) Cit. statements	(2) Cit. statements	(3) Mentioning	(4) Mentioning	(5) Contrasting	(6) Contrasting	(7) Supporting	(8) Supporting
Post	0.033	0.038	0.034	0.040	-0.108	-0.132	0.035	0.031
	(0.038)	(0.040)	(0.038)	(0.040)	(0.236)	(0.263)	(0.100)	(0.109)
Post * Treatment	-1.215***	-1.165***	-1.231***	-1.178***	-1.147***	-1.184***	-1.055***	-1.061***
	(0.096)	(0.093)	(0.095)	(0.093)	(0.339)	(0.348)	(0.197)	(0.209)
Post * Early mentions	0.245***	0.171***	0.238***	0.167**	0.738**	0.745**	0.417***	0.306*
	(0.063)	(0.066)	(0.065)	(0.068)	(0.317)	(0.347)	(0.158)	(0.170)
Post * Treatment * Early mentions	-0.411**	-0.486***	-0.396**	-0.478***	-0.595	-0.416	-0.456	-0.455
	(0.164)	(0.170)	(0.167)	(0.173)	(0.684)	(0.730)	(0.414)	(0.440)
Self cit. excluded	N	Y	N	Y	N	Y	N	Y
Article FE	Y	Y	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Pseudo R2	0.713	0.717	0.711	0.715	0.138	0.138	0.267	0.254
N	14594	14158	14536	14089	1701	1421	6370	5761
N clusters	957	948	956	946	201	170	567	521
N full	16711	16711	16711	16711	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations statements received by each paper in a particular year, even columns exclude self citations. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Early mentions is an indicator for papers with at least one mention (in newspapers and/or blogs) within two weeks from publication. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1 - exp[\beta]) * 100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citation statements). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.30: Citation statements and attention score

	(1) Cit. statements	(2) Cit. statements	(3) Mentioning	(4) Mentioning	(5) Contrasting	(6) Contrasting	(7) Supporting	(8) Supporting
Post	-0.023	0.024	-0.022	0.024	0.050	0.352	0.045	0.115
	(0.052)	(0.055)	(0.053)	(0.056)	(0.344)	(0.360)	(0.131)	(0.142)
Post * Treatment	-1.001***	-0.941***	-0.992***	-0.923***	-1.696***	-1.972***	-1.218***	-1.345***
	(0.121)	(0.119)	(0.121)	(0.119)	(0.634)	(0.681)	(0.351)	(0.373)
Post * Altscore >p50	0.173***	0.090	0.170***	0.091	0.048	-0.367	0.149	0.009
	(0.060)	(0.063)	(0.060)	(0.064)	(0.350)	(0.368)	(0.151)	(0.159)
Post * Treatment * Altscore >p50	-0.415***	-0.441***	-0.441***	-0.477***	0.480	0.808	0.030	0.190
	(0.155)	(0.153)	(0.155)	(0.153)	(0.704)	(0.756)	(0.403)	(0.428)
Self cit. excluded	N	Y	N	Y	N	Y	N	Y
Article FE	Y	Y	Y	Y	Y	Y	Y	Y
Age FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Pseudo R2	0.713	0.717	0.711	0.715	0.136	0.136	0.266	0.253
N	14594	14158	14536	14089	1701	1421	6370	5761
N clusters	957	948	956	946	201	170	567	521
N full	16711	16711	16711	16711	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations statements received by each paper in a particular year, even columns exclude self citations. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Altscore is an aggregate measure of weighted online mentions. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citation statements). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.31: Citation statements and attention score extremes

	(1) Cit. statements	(2) Cit. statements	(3) Mentioning	(4) Mentioning	(5) Contrasting	(6) Contrasting	(7) Supporting	(8) Supporting
Post	0.043	0.074	0.046	0.074	0.262	0.567	0.070	0.177
	(0.056)	(0.059)	(0.056)	(0.060)	(0.373)	(0.355)	(0.146)	(0.157)
Post * Treatment	-1.103***	-1.002***	-1.097***	-0.984***	-1.940**	-2.203***	-1.226***	-1.385***
	(0.156)	(0.153)	(0.155)	(0.150)	(0.769)	(0.820)	(0.467)	(0.484)
Post * Treatment * Altscore 3rd qntl	0.428**	0.287	0.438**	0.286	1.369	1.523	0.203	0.315
	(0.211)	(0.217)	(0.212)	(0.218)	(1.254)	(1.342)	(0.654)	(0.684)
Post * Treatment * Altscore 4th qntl	-0.047	-0.105	-0.087	-0.160	0.249	0.468	0.476	0.743
	(0.205)	(0.205)	(0.204)	(0.204)	(1.288)	(1.338)	(0.546)	(0.568)
Post * Treatment * Altscore 5th qntl	-0.456**	-0.510***	-0.472**	-0.538***	0.513	0.899	-0.224	-0.024
	(0.197)	(0.193)	(0.196)	(0.191)	(0.849)	(0.907)	(0.534)	(0.554)
Self cit. excluded	N	Y	N	Y	N	Y	N	Y
Age FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Pseudo R2	0.714	0.717	0.712	0.716	0.142	0.140	0.268	0.255
N	14594	14158	14536	14089	1701	1421	6370	5761
N clusters	957	948	956	946	201	170	567	521
N full	16711	16711	16711	16711	16711	16711	16711	16711

Note: Estimates derive from pseudo Poisson specifications. The dependent variable is the total number of citations statements received by each paper in a particular year, even columns exclude self citations. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Altscore is an aggregate measure of weighted online mentions. All models incorporate article fixed effects, a full suite of calendar-year effects and article age indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (i.e. $x\%$ loss in yearly citation statements). Standard errors in parentheses, clustered around retraction cases. *p $<$ 0.10, **p $<$ 0.05, ***p $<$ 0.01.

Table A.32: Impact on authors' careers (split samples)

	Media								No Media							
	(1) All	(2) First author	(3) Mid author	(4) Last author	(5) H-index >p50	(6) H-index <=p50	(7) Severe misconduct	(8) Non-Severe misconduct	(9) All	(10) First author	(11) Mid author	(12) Last author	(13) H-index >p50	(14) H-index <=p50	(15) Severe misconduct	(16) Non-Severe misconduct
Panel A: N articles																
Post * Treatment	-0.212**	-0.612***	0.102	-0.187*	-0.246***	-0.029	-0.353***	-0.064	-0.092***	-0.090	-0.072	-0.098**	-0.114***	-0.069	-0.222***	-0.017
	(0.085)	(0.159)	(0.116)	(0.105)	(0.091)	(0.133)	(0.113)	(0.106)	(0.035)	(0.059)	(0.055)	(0.046)	(0.041)	(0.054)	(0.051)	(0.046)
Pseudo R2	0.583	0.479	0.617	0.542	0.543	0.301	0.590	0.578	0.597	0.498	0.573	0.597	0.570	0.419	0.594	0.600
N	6055	2082	1867	2102	3788	2266	3459	2595	48672	16549	14776	17344	24968	23699	19931	28740
N clusters	105	105	99	102	100	101	55	50	783	779	761	778	757	766	315	468
N authors	760	266	242	252	451	309	413	347	5905	2061	1823	2021	2874	3031	2382	3523
Panel B: N articles with grant																
Post * Treatment	-0.207**	-0.543***	0.012	-0.165	-0.262***	0.165	-0.332***	-0.019	-0.060	-0.020	-0.022	-0.089	-0.103**	0.008	-0.160**	-0.015
	(0.091)	(0.186)	(0.134)	(0.122)	(0.098)	(0.169)	(0.128)	(0.118)	(0.045)	(0.075)	(0.068)	(0.054)	(0.049)	(0.071)	(0.069)	(0.054)
Pseudo R2	0.552	0.467	0.586	0.518	0.516	0.297	0.558	0.550	0.563	0.481	0.533	0.573	0.546	0.394	0.568	0.562
N	5893	2008	1795	2082	3765	2125	3362	2530	44345	14630	13321	16388	24364	19980	18389	25955
N clusters	104	103	99	101	100	98	55	49	780	766	741	766	754	744	315	465
N authors	735	255	231	249	448	287	399	336	5334	1811	1626	1897	2803	2531	2179	3155
Panel C: Avg. n coauthors (x article)																
Post * Treatment	-0.090	-0.233	0.138	-0.092	-0.119	-0.034	-0.204*	0.046	-0.090***	-0.113**	-0.082*	-0.077**	-0.089***	-0.082*	-0.167***	-0.029
	(0.080)	(0.152)	(0.143)	(0.082)	(0.083)	(0.178)	(0.119)	(0.090)	(0.028)	(0.056)	(0.046)	(0.034)	(0.033)	(0.044)	(0.040)	(0.038)
Pseudo R2	0.384	0.392	0.384	0.395	0.419	0.281	0.392	0.383	0.378	0.365	0.383	0.378	0.390	0.313	0.360	0.391
N	6055	2082	1867	2102	3788	2266	3459	2595	48672	16549	14776	17344	24968	23699	19931	28740
N clusters	105	105	99	102	100	101	55	50	783	779	761	778	757	766	315	468
N authors	760	266	242	252	451	309	413	347	5905	2061	1823	2021	2874	3031	2382	3523
Author FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Career lenght FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Calendar Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N full	6100	2111	1877	2112	3813	2287	3496	2604	49016	16681	14921	17414	25009	24007	20118	28898

Note: Estimates derive from pseudo Poisson specifications. The dependent variables are: N. published articles x author x year; N. published articles with grant support x author x year; or Avg. n collaborators across all author's publications x year. Treatment is an indicator for retracted papers. Post is an indicator for all years strictly after the year of retraction. Media is an indicator for cases where the original publication (either retracted or control papers) had at least one mention (in newspapers and/or blogs) within two weeks from publication. All models incorporate author fixed effects, a full suite of calendar-year effects and carreer lenght indicator variables. Using the following transformation $(1-exp[\beta])*100 = x$ coefficients can be interpreted as elasticities (e.g. x% loss in publication rate). Standard errors in parentheses, clustered around retraction cases. *p < 0.10, **p < 0.05, ***p < 0.01.

Chapter 2

NATIONAL POLLS, LOCAL PREFERENCES AND VOTERS' BEHAVIOUR: EVIDENCE FROM THE UK GENERAL ELECTIONS

I thank Thiemo Fetzer and Sascha Becker for the precious mentoring and guidance, Ceci Cruz, Maurice Dunaiski, Teresa Esteban-Casanelles, Helios Herrera, Vardges Levonyan, Shanker Satyanath, Vera Troeger as well as the participants of GRAPE meeting at IRCrES-CNR, the Sixth LSE-Oxford graduate student conference in political economy, the 25th Spring Meeting of Young Economists, and Warwick internal seminars for insightful comments and suggestions. Remaining errors are my own.

Abstract

A central challenge for social scientists consists in explaining why people vote and what are the consequences of their behaviour. Exploiting variation in national opinion polls across UK general elections, and in the degree of safeness of British constituencies over time, I provide evidence of a significant impact of pre-election polls on electoral outcomes and shed light on a novel mechanism. I find that opinion polls affect voters' behaviour via their interaction with the recent electoral history of a constituency: first, turnout decreases when the polls predict non-competitive elections, and this effect is stronger in safe seats. Second, the composition of local vote shares and parties' performance is also impacted by anticipated election closeness and the effects vary heterogeneously depending on whether poll predictions are aligned with the past electoral outcomes of a constituency. Finally, the causal impact on voters' participation is confirmed with consistent individual-level evidence.

Keywords: Opinion Polls, Closeness, Voters' Behaviour, First-past-the-post, UK general elections

JEL Classification: D72, P16

Word count: 11,998

2.1 Introduction

Why do citizens decide to turn out to vote? How does this affect electoral outcomes? Understanding voters' participation is a challenge that social scientists have been trying to solve for many decades (Blais, 2006).[1] In recent years scholars have made considerable improvements in shedding light on what drives electorate's choices.[2] Yet, one of the most empirically contested driver is the role of pre-election polls. Following the unexpected Brexit vote and Trump's electoral success, pre-election polls have been the object of heated debates regarding their capability to predict the electoral results or to directly influence voters' behaviour.

Canonical rational choice models (see seminal contribution by Downs, 1957) predict that the smaller the predicted margin of victory, the higher voters' participation. The mechanism generating this observation is that, the more competitive the electoral race, the higher a voter's perception of the importance of her voting decision.[3] Polls give voters an indication about the expected closeness of elections (Palfrey and Rosenthal, 1985) and voters may use this information when deciding whether or not to vote. Anecdotal evidence supports this statement: in the UK general election of 2001, an expected high margin of victory for Labour resulted in very low turnout. The BBC surveyed those who abstained from voting and found out that a vast majority reported there being no point in voting as their vote would not change the result. Similarly, just over half of the respondents said it was obvious that Labour would win.[4] However, the constituency of Arundel and South Down experienced the victory of a Conservative candidate and a turnout well above the national average. These observations raise concerns on the existence of alternative mechanisms through which opinion polls may affect voters' behaviour. Indeed, the often heard cry of "every vote matters" may depend considerably on the electoral system. For instance, in the UK it is a widespread belief that Members of Parliament (MPs) being elected via first-past-the-post (FPTP) system may be the cause of what are commonly referred to as safe

[1]Partecipation is a puzzle even in context where it is far more likely an individual voter is pivotal (see Coate et al., 2008; Farber, 2010).

[2]Among the factors under study there are: habits (Fujiwara et al., 2016), personality traits (Ortoleva and Snowberg, 2015), social considerations (Gerber et al., 2015; Funk, 2010; Dellavigna et al., 2017), political movements (Madestam et al., 2013), media content (Strömberg, 2004; DellaVigna and Kaplan, 2007; Gentzkow, 2006; Enikolopov et al., 2011; Gentzkow et al., 2011), and compulsory voting laws (León, 2017; Hoffman et al., 2017).

[3]Increases in turnout may be induced by alternative mechanisms: for instance, election closeness may interact with social preferences (e.g., Dellavigna et al., 2017) or with the intrinsic utility from voting (e.g., Riker and Ordeshook, 1968; Brennan and Buchanan, 1984; Schuessler, 2000; Feddersen and Sardoni, 2006; Ali and Lin, 2013).

[4]Source: BBC - "Turnout at 80-year low".

seats. According to the Electoral Reform Society (ERF) 192 constituencies have not changed hands electorally since WWII.[5] For instance, North Shropshire has been a Tory seat ever since 1835. It is not a surprise that in those safe seats many voters may feel discouraged to vote and mobilization efforts may be lower (see for example Cox, 1999; Franklin et al., 2004; Selb, 2009; Herrera et al., 2014).

What would then be the effect of the polls predicting a Labor victory in a constituency such as Shropshire? It appears likely that the joint presence of safe seats and polls predictions could play a key role in explaining electoral outcomes. Thus, the aim of this paper is to shed light on this so far unexplored mechanism through which anticipated election closeness may affect voters' behaviour.

Exploiting a panel of constituencies across UK general elections from 1983 to 2017, I find strong evidence that polls predictions, interacted with the recent historical preferences of a constituency electorate, significantly impact voters' participation, concentration of vote shares, as well as local parties' shares and chances of victory. I also show that pollsters' predictions matter more as the election becomes closer. In addition, since I measure the extent to which a seat can be considered safe, results suggest the effect of polls is not homogeneous along the safeness distribution. Furthermore, findings indicate that polls could have different effects on a party performance depending on whether the information they provide is aligned with the electoral history of a constituency. Finally, I use quasi-random variation in individual-level exposure to opinion polls, to corroborate that the interaction between polls predictions and past local electoral preferences influence voters' political engagement. Importantly, this relationship emerges only when the opinion polls information is relevant for voters, i.e. before a general election.

Previous empirical efforts aimed at measuring the causal effect of anticipated election closeness can be categorized in three broad groups providing mixed evidence. A first group of contributions, reviewed in the meta-analysis by Cancela and Geys (2016), exploits observational data and find suggestive evidence that turnout tends to increase in measures of actual (e.g., Barzel and Silberberg, 1973; Cox and Munger, 1989; Matsusaka, 1993) or predicted closeness (e.g., Shachar and Nalebuff, 1999) across elections. However, these efforts have been plagued by re-

[5]Source: ERF - "The 2019 General Election: Voters Left Voiceless".

verse causality (realised closeness) and omitted variables bias (predicted closeness). On the one hand, ex-post electoral results could endogenously depend on the realized turnout. On the other hand, turnout could be affected by factors which may also make the electoral race more competitive such as the importance of a certain election, the intensity of the campaign and campaign advertisement, or news coverage. For instance, tight races have been shown to be correlated with more campaign spending (Cox and Munger, 1989; Matsusaka, 1993; Ashworth and Clinton, 2006), more party contact (Shachar and Nalebuff, 1999; Gimpel et al., 2007), more campaign appearances (Althaus et al., 2002), and more news coverage (Banducci and Hanretty, 2014). Furthermore, social pressure to vote may be enhanced by elites as a result of close elections (Cox et al., 1998). Some recent contributions started addressing these concerns seriously. Morton et al. (2015) show that the availability of exit poll results in French elections reduces turnout in late-voting constituencies, though these constituencies are far from being pivotal. Bursztyn et al. (2020) rigorously analyse the impact of ex-ante closeness of a race by exploiting naturally occurring variation in the existence, closeness, and dissemination of Swiss pre-election polls, finding that anticipated election closeness increases turnout significantly more in areas where newspapers report on them most. Yet, the referenda setting is not the best suited to exploit naturally occurring variation in the political composition of local preferences (safeness of a constituency), which I believe to be a powerful factor interacting with the polls and thus determining voters' behaviour.

A second stream of literature uses lab experiments (see Levine and Palfrey, 2007; Duffy and Tavits, 2008; Großer and Schram, 2010; Agranov et al., 2018) to provide strong evidence that increased predicted tightness of an electoral race is associated with enhanced voters' participation.[6] However, external validity remains an unresolved issue as lab experiments are by definition unable to capture the context of real-life elections. Thus, one would ideally like to identify similar results in the field.

A third group of scholars implemented field experiments providing information treatments to potential voters (Gerber and Green, 2000; Bennion, 2005; Dale and Strauss, 2009; Enos and Fowler, 2014; Gerber et al., 2020), eventually finding little or no evidence of a link between

[6] Nonetheless, participants' behaviour is not always consistent with the full set of predictions arising from the pivotal voter model.

closeness and turnout. Yet, in such settings it is difficult to control for voters' access to outside information. The weak relationship may in fact result from voters recovering additional common information outside of the experiment.

Compared to the existing empirical works I make four unique contributions. First, I provide evidence of a previously neglected mechanism: anticipated election closeness interacts with the local history of a constituency. Second, I show that polls predictions not only affect voters' participation, but also the composition of local vote shares and parties' performances. Third, I exploit a rich setting of elections across thirty-five years which makes results easier to interpret and compare. Forth, I provide a robust validation of the main results using quasi-random individual level variation.

This setting allows to estimate models with election fixed effects, exploiting within-election, cross-constituency variation in historical preferences which may or may not be aligned with pre-election polls. Therefore, I can seriously address concerns related not only to reverse causality, but also related to presence of potential confounders. Furthermore, individual level data offer an important feature for analysis as interview dates are randomly assigned. Survey respondents are hence exposed to a quasi-random polling information at the start of their interview whose timing is exogenous to their political engagement and therefore allows to credibly address the identification issues highlighted above.

The paper is structured as follows: Section 2.2 describes the institutional settings, the data at hand and discusses the empirical design; Section 2.3 reports results of the aggregate level analysis; Section 2.3.4 describes the individual-level analysis; Section 2.4 provides conclusive remarks.

2.2 Background, data and empirical approach

The focus of this work is on the UK's general elections for two reasons. First, despite their national nature, voters express electoral preferences for their local MP. This makes it possible to set up an empirical design that exploits national level polling with local level historical electoral information. Second, the stability of the UK's electoral system allows to study the evolution of the impact of electoral polls in a wide range of elections.

2.2.1 UK general elections

General elections provide an opportunity for UK citizens to elect MPs forming the House of Commons of the UK Parliament. Each MP is the winner of the electoral race at the constituency level. A key feature is that every constituency elects its MP via a FPTP system (i.e. voters can only name one candidate, and the one who obtains most votes becomes MP). Upon election, MPs will represent their local area for up to five years. In terms of party membership, local candidates can either belong to a political party or stand as independents. Historically, few independent MPs ever got elected. At the national level, the party that obtains more seats than all the other parties combined (i.e. the one with the overall parliamentary majority) is appointed the formation of the government. In the absence of an outright majority, parties usually seek to form coalitions.

An additional remark concerns the rules governing shape and formation of parliamentary constituencies. The UK is currently divided into 650 constituencies (corresponding to 650 MPs), but number and boundaries changed repeatedly. Following the Parliamentary Constituencies Act of 1986, boundaries have been subject to periodic reviews by four Boundary Commissions (one per country). These Commissions update boundaries in accordance with rules which set out both the number of constituencies and the extent to which the size of the electorate in each constituency can differ from the electoral quota (i.e. average size of a constituency). That said, under the assumption that constituencies retaining the same name over time have been subject to little or no change in boundaries, the analysis is based on a panel of different constituency-names over time.[7]

This work considers all general elections between 1983 and 2017, with electoral outcomes reported at the constituency level.[8] A summarizing picture of these past elections is presented in Figure B.1. The bar chart illustrates that, considering different seats in each general election as a distinct observation, roughly 88 percent were won by either a Conservative or a Labour candi-

[7] For example consider the constituency of Basildon, which in 2010 was divided in the two constituencies of Basildon and Billericay, and South Basildon and East Thurrock. In this case the three uniquely named areas figure in the data as separate observations in different general elections.

[8] General election years are the following: 1983; 1987; 1992; 1997; 2001; 2005; 2010; 2015; 2017.

date (over 90 percent when excluding Northern Ireland) with a slight supremacy of Conservative seats. Given the widespread prevalence of victories by the two major UK parties, I restrict the attention to those constituencies where both a Conservative and a Labour candidate competed at least once.[9]

Despite a similar proportion of constituencies held by the two main parties over time, electoral results vary considerably across time and space, and this will be fundamental for the analysis. To exploit such variation I build a measure of electoral competitiveness between Conservative and Labour party:

$$Adj.margin_{c,t} = \frac{|shareCon_{c,t} - shareLab_{c,t}|}{shareCon_{c,t} + shareLab_{c,t}}$$

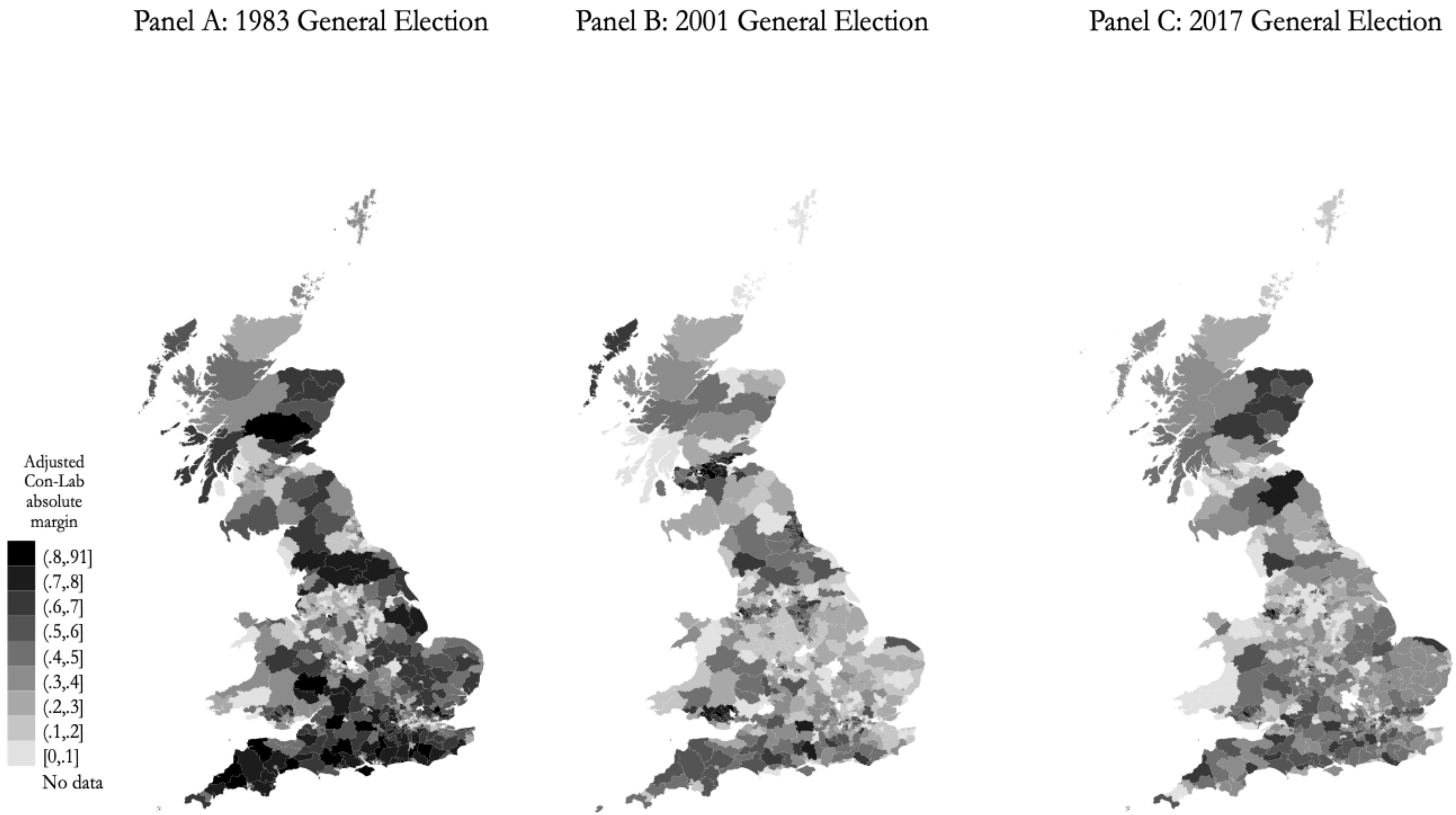
where *share* is the proportion of votes obtained by the party in the local race, subscript c indicates a constituency and t refers to a given general election. Note that the electoral margin is adjusted to the local relevance of the two parties combined (i.e. the denominator in the formula).

Figure 2.1 depicts *Adj.margin* across the UK for three different elections: the furthest in time, the most recent, and the mid 2001 election. The figure helps visualize the presence of constituencies with a solid and persistent support for one of the two parties (often named safe seats), as opposed to those generally more competitive (in lighter shades).

As the objective of this paper is to study whether being a *safe* Conservative or Labour constituency is a fundamental factor interacting with opinion polls which may contrast or reinforce the predicted result, I will use (one period lagged) *Adj.margin* as a measure of *safeness* of a seat, thus taking advantage of the variation just presented.

[9] This cleaning process eliminates the constituencies of Northern Ireland (60 percent of the dropped observations, i.e. 17 or 18 yearly seats) and few additional ones.

Figure 2.1: Adjusted margin of victory across general elections (Conservative - Labour) in absolute terms



Note: Shades map the variation in absolute vote share margin between Conservative and Labour parties across general elections, adjusted dividing by the sum of the two party shares.

2.2.2 Opinion polls in the UK

Great Britain has a long history of surveys on voting intentions. First was Gallup in 1937, just two years after its American counterpart. However, at the dawn of their diffusion, polls were largely ignored by politicians. This attitude changed in the 1950s, when the appearance of new pollsters led parties members to pay greater attention to this tool. As a consequence, the following years witnessed a rapid rise in the number of commissioned polls by parties. New companies entered the market and traditional media began to devote greater consideration to the polls. In the 1970s, following the abandonment of exclusive publication, polls became accessible to an enormously enhanced audience. Not surprisingly, during this period both Conservative and Labour party initiated substantial private polling programs. Ever since, pre-election polls have been dominating campaign reporting (Worcester, 1980). Nowadays, various organisations carry out opinion polling to gauge voting intention and most of the polling companies are members of the British Polling Council (BPC) and abide by its disclosure rules. Predicted support for political parties out of the electoral campaign periods is frequently and widely reported in the news.

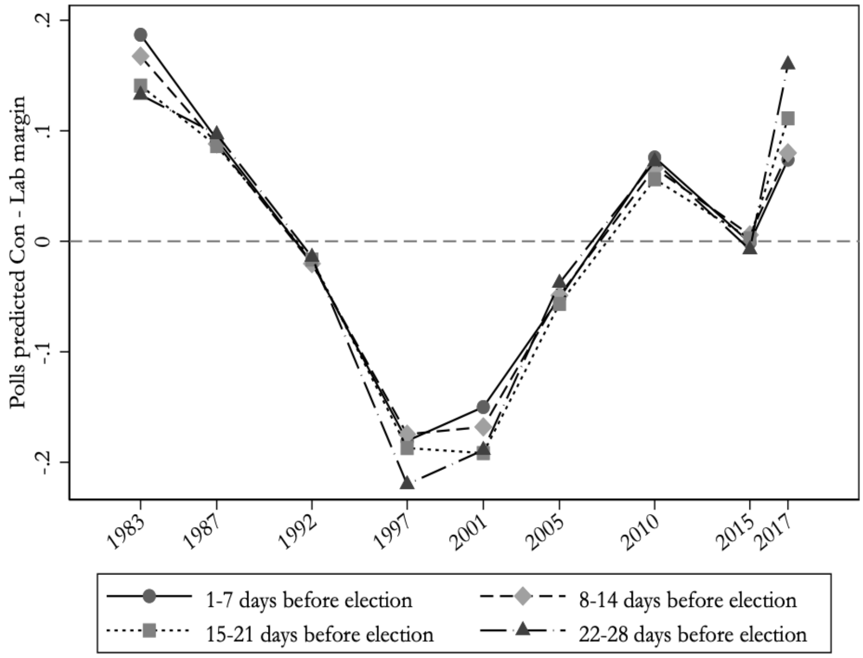
For the analysis, I focus on national polls produced within four weeks from the general election day. In the data, I condense polling information in each year, starting with the existing six pollsters of 1983 and finishing with the ten polling companies active during the 2017 general election campaign.[10] The number of pollsters I observe ranges from 5 in 1997 to 11 in 2015. As mentioned above, I am interested in studying the impact of predicted closeness on election outcomes. Thus, given that Conservative and Labour parties were the top competing forces during all the general elections in the sample (see also Figure B.1), I measure ex-ante closeness of the race as follows:

$$Pollmargin_w = |\widehat{shareCon}_w - \widehat{shareLab}_w| \;\; with \;\; w = 1, 2, 3, 4$$

where $\widehat{shareParty}_w = \frac{1}{Nr\ Pollsters} \sum_j \widehat{shareParty}_{jw}$ and $\widehat{shareParty}_{jw}$ is the predicted vote share for a given $Party$ (either Con or Lab) by a given pollster j, in a given week w preceding the election.

[10] For the individual-level analysis I use data on all opinion polls produced within four weeks from the start of respondents' interviews (see section 2.2.3).

Figure 2.2: Yearly variation in average opinion polls margin



Note: Estimates map variation in average opinion poll margin between Conservative and Labour parties across general elections. The margin is calculated averaging the differece in party vote shares across all national pollsters released in a given week before the election date. Positive margin refer to a predicted conservative advantage and viceversa.

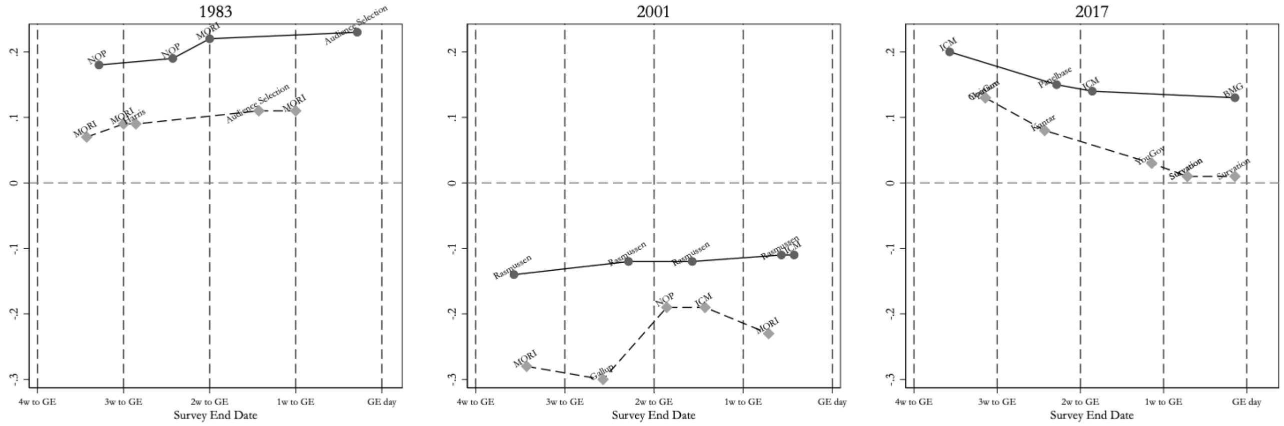
Figure 2.2 displays the trends in (national) predicted polls margin across all general elections from 1983 to 2017. For illustrative purposes I use positive margins for a predicted Tory lead and negative otherwise. Two features emerge from this graph. First, $Pollmargin$ varies considerably across the years. The sample contains both competitive and non-competitive elections with either party leading the polls at least three times. Second, there seems to be variation in the polls margin reported at different points in time along the electoral campaign (comparing the different line colours). For instance, in 1983, as the election day became closer, pollsters predicted a larger Conservative victory. Conversely, in 1997 or 2017, approaching the election day the margins reported by the pollsters became increasingly small. This variation is also well presented in Figure B.2. which displays the distributions of residuals of all (absolute) polls margins published in a certain period of time (i.e. from the last to the fourth week preceding elections) after accounting for election fixed effects. Densities are all bell-shaped but the dispersion changes systematically across weeks as poll estimates get generally more similar the closer the election date.

Polls margins vary depending on the polling institution which produce them (Panel A of Figure 2.3 and B.4) and, as a consequence, on the related publisher (Panel B of Figure 2.3 and B.4).[11] Looking at reported minimum and maximum margins by pollsters, one can notice some interesting features. First, while in 1983 the difference between the minimum and the maximum remains almost constant across the four weeks preceding the elections, the gap seems to widen in 2017, indicating that variance of the polls differs across years (the same emerges from the graphs in Figure 2.3 (Panel B)). Second, in 2001 it is notable that the margin closest to zero is always reported by the same pollster, i.e. Rasmussen, suggesting the presence of a systematic prediction bias by some polling companies. Related to this second point, the graphs in Figure 2.3 (Panell B) show an almost equal picture, with some minor differences. For instance, looking at the 2001 general elections, one can see that the Sunday Telegraph chose to report polls from different firms, which however both coincide with those that predicted the largest margin in favour of the Labour party, suggesting the presence of a publication bias.

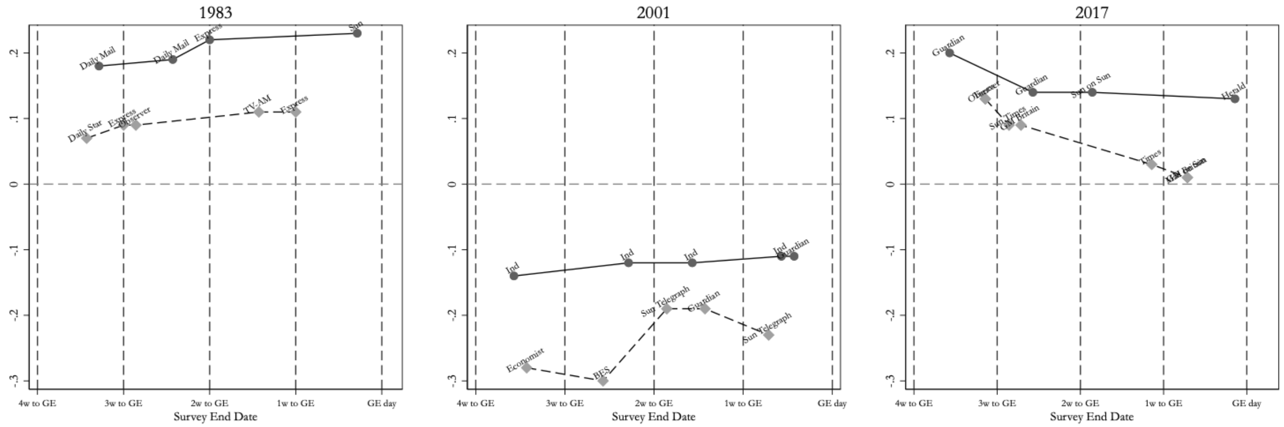
[11] Panel B of Figure B.4 focuses on the top ten publishers across the period under study.

Figure 2.3: Weekly variation in (min and max) polls margins

Panel A: **Variation by pollster**



Panel B: **Variation by publisher**



Note: Estimates show the maximum (solid) and the minimum (dashed) opinion poll margin between Conservative and Labour parties in a given week before the general election date and across general elections. Color labels name the pollster associated to each estimated margin (panel A) or its publisher (panel B).

Motivated by the features just described, I dug more deeply into the opinion polling panel looking for regularities. For each reported opinion poll in the last four weeks preceding elections, the panel lists: the predicted party shares, the margin, the end date of poll, the associated polling house and the (first) publisher.[12] The following tables suggest systematic differences in reported opinion polls. Table B.11 displays results of a simple pollsters fixed effects regression:

$$y_{j,t,w} = \sum_j \beta_j Pollster_j + \gamma' X_{t,w} + \epsilon_{j,t,w} \;\; with \;\; w = 1, 2, 3, 4 \tag{1}$$

where y are either the Conservative or Labour share or the absolute difference between the two, as reported by pollster j in week w preceding general election t. X represents week-by-year fixed effects.

Assuming that the sampling methodologies used and the analysis performed by the different polling houses are comparable, there should be no systematic difference across polls. However, the fact that some of the pollsters fixed effects in Table B.11 are significantly different from zero suggests otherwise. Take the example of Rasmussen, results suggest this polling house systematically reports higher Conservative shares and lower Labour shares thus lower poll margins than the excluded pollster MORI.

One interesting avenue for future research is to explore causes behind these differences. One possibility is that since media outlets select their pollsters, they may release pre-election poll estimates that are distorted based on their political leaning.[13] The awareness of a feedback between opinion polling and turnout may be the reason for this behaviour, possibly aimed at mobilizing (or discouraging) readers' participation. Table 2.1 displays results for a preliminary test for this assumption. More specifically, I perform the following regression:

$$y_{j,t,w} = \beta I_{j,t} + \gamma' X_{t,w} + \epsilon_{j,t,w} \;\; with \;\; w = 1, 2, 3, 4 \tag{2}$$

where y are again either party shares or poll margins, as reported by pollster j in week w preceding general election t.

[12] There are very few cases where two publishers are listed, I ignore those second publishers for simplicity.

[13] In the context of the Brexit referendum, Cipullo and Reslow (2019) find evidence of bias in macroeconomic forecasts released by institutions with stakes and influence.

Table 2.1: Reported opinion poll shares and margin by publisher orientation

	Panel A - Dep. var.:					
	share Conservative		share Labour		Pollmargin	
	(1)	(2)	(3)	(4)	(5)	(6)
Right	-0.0037*	-0.0033*	0.0059**	0.0053**	-0.0002	0.0012
	(0.0020)	(0.0020)	(0.0026)	(0.0022)	(0.0037)	(0.0035)
Week FE	X		X		X	
Year FE	X		X		X	
Week*Year FE		X		X		X
Observations	345	345	345	345	345	345
R-squared	0.9065	0.9272	0.9217	0.9467	0.8231	0.8583
	Panel B - Dep. var.:					
	share Conservative		share Labour		Pollmargin	
	(7)	(8)	(9)	(10)	(11)	(12)
Endorsing	-0.0027	-0.0022	0.0054*	0.0054**	-0.0044	-0.0035
Conservative	(0.0022)	(0.0021)	(0.0028)	(0.0024)	(0.0039)	(0.0038)
Week FE	X		X		X	
Year FE	X		X		X	
Week*Year FE		X		X		X
Observations	343	343	343	343	343	343
R-squared	0.9039	0.9258	0.9177	0.9428	0.8194	0.8548

Notes: *Pollmargin* is the absolute difference between Conservative and Labour parties' vote shares and $\in (0, 1)$. *Right* is an indicator for whether a publisher (newspaper) is perceived as right or centre-right leaning. *Endorsing conservative* is an indicator for whether a publisher (newspaper) has endorsed the conservative party/candidate in that general election. Robust standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

The variable of interest is now I which is either an indicator for whether the publisher (newspaper) associated to pollster j is perceived as right or centre-right leaning (Panel A)[14] or alternatively an indicator for whether the newspaper associated to pollster j has endorsed the Conservative party or a Conservative candidate in general election t (Panel B).[15] X represents either week and year or week-by-year fixed effects. These indicators are only an approximation of the political position of a newspaper which may well vary across time and voters' readership. However, results across specifications suggest that right leaning newspapers have a tendency to overstate the Labour poll share relative to the Conservative poll share. Although suggestive, there seem to be a publisher bias in line with priors.

2.2.3 Data

Constituency-level analysis

Data come from different sources. Electoral results at constituency level are extracted from the Electoral commission website and from Richard Kimber's www.politicsresources.net. Corresponding opinion polling data covering the electoral campaign of each general election since 1983 were collected from ukpollingreport.co.uk.[16]

The sample is restricted to those constituencies that experienced candidates from both Conservative and Labour party competing at least once in the period considered. In addition, constituencies changing names over time are treated as different observations given that the reference boundaries also change.

The dataset includes variables such as turnout, party shares and the predicted shares from polls which are necessary for the creation of $Adj.margin$ and $Pollmargin$, as already described. In addition, I measure the concentration of vote shares:

$$HHI_{c,t} = \sum_{p} share^2_{p,c,t}$$

where, as before, subscripts c and t indicate respectively the constituency and the election year, while p indicates a party. Thus, $share_{p,c,t}$ is the vote share gained by party p in a given

[14]Source: YouGov survey on perceived newspaper ideology.
[15]Sources: Guardian (a); Guardian (b); Wikipedia (a); Wikipedia (b); Wikipedia (c).
[16]Historical opinion polls are in turn extracted from Mark Pack's online archive.

constituency and year. This measure is inspired by the Herfindahl-Hirschman index, a commonly used measure of market concentration. By construction, HHI can take values between zero and one. The upper limit indicates the case of a single party capturing all cast votes, while zero refers to a scenario with infinitely many parties competing for the seat, each of them obtaining the same share of votes. Like other aggregate level variables, HHI allows to study the general influence of opinion polls on the politics of a constituency. Despite the choice to focus exclusively on the two major parties, this index is computed taking every competing party share into account, which in turn allows to draw more general conclusions. However, the party level analysis focuses on variables related uniquely to Conservative and Labour candidates.

Table 2.2: Summary statistics (main analysis)

	Panel A: National Polls			
	1 week to GE	2 weeks to GE	3 weeks to GE	4 weeks to GE
Pollmargin	0.0776	0.0790	0.0861	0.0970
	(0.0556)	(0.0571)	(0.0659)	(0.0749)
# of polls	16.6236	12.2864	12.2162	12.0505
	(5.2212)	(5.0888)	(4.5047)	(3.8292)

	Panel B: Constituency Level Variables		
	Whole sample	Incumbent = poll leader	Incumbent $\neq$ poll leader
Turnout	0.6814	0.6667	0.6987
	(0.0824)	(0.0866)	(0.0736)
HHI	0.3870	0.3937	0.3790
	(0.0634)	(0.0632)	(0.0627)
Adj. margin$_{t-1}$	0.3704	0.3797	0.3594
	(0.2221)	(0.2217)	(0.2221)
	[4676]	[2530]	[2146]

	Panel C: Party Level Variables		
	Whole sample	Incumbent = poll leader	Incumbent $\neq$ poll leader
Incumbent vote share	0.5100	0.5258	0.4917
	(0.0936)	(0.0911)	(0.0932)
Incumbent prop. victories	0.8938	0.9289	0.8530
	(0.3082)	(0.2571)	(0.3542)
	[4293]	[2306]	[1987]
Follower vote share	0.3115	0.2915	0.3402
	(0.0967)	(0.0933)	(0.0944)
Follower prop. victories	0.1367	0.0800	0.2179
	(0.3436)	(0.2714)	(0.4130)
	[3014]	[1775]	[1239]

Notes: All margins are in absolute terms. Table reports variable means, with standard deviations in parenthesis and number of observations in square brackets.

Table 2.2 reports selected statistics on the variables introduced above. Panel A shows that opinion polls vary substantially depending on the time distance to the election day (in line with Figures 2.2 and B.2). On the one hand, the average prediction becomes more competitive and precise the closer the election (i.e. I observe lower average margin and standard deviation). On the other, polls become more frequent. Panel B displays selected statistics for constituency level variables. Turnout is on average higher when the local incumbent party is not leading in national polls, while the HHI is lower. Moreover, the (lagged) adjusted margin is generally large (with considerable variation, as shown in Figure 2.1) but to a lesser extent in constituencies where polls predictions are not aligned with the previous local results. Finally, Panel C examines party level outcomes. These exhibit some differences in the two sub-samples. Incumbent vote shares are greater when their party is predicted to win in the national race; a similar pattern can be observed when looking at their probability to regain the seat. Conversely, in the same constituencies, follower vote shares and probability of winning are tinier.

Individual-level analysis

The last set of results uses individual-level data from *Understanding society*. The UK's largest panel of representative households covering a wide range of topics among which the following questions on political engagement:

1. Generally speaking do you think of yourself as a supporter of any one political party?
2. Do you think of yourself as a little closer to one particular party than the others?
3. If there were to be a general election tomorrow, which political party do you think you would be more likely to support?

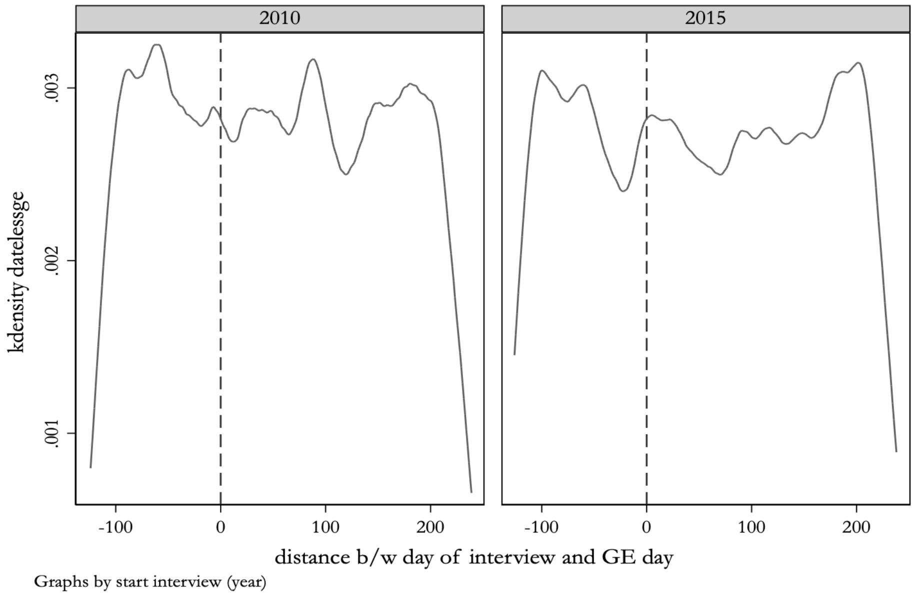
All respondents are asked the first question.[17] Those who reply negatively, are then asked the second, then the third if they keep providing a negative answer. Lastly, individuals are allowed to reply that they would vote for no party in the final question. I use these variables to proxy for respondents' willingness to turnout in general elections.

At the time of the analysis, interviews were conducted in eight semi-overlapping waves, each of 24 months, covering the 2009-2017 period (I disregard the first and last year as the number of respondents interviewed is negligible). Hence, I focus on individuals starting their questionnaire in either 2010, 2015 or both years, which correspond to general elections years. The analysis

[17] I exclude inapplicable respondents, missing answers and those who refuse to reply the first question.

implemented with these data looks separately at responses provided before and after the election date. Figure 2.4 illustrate that the daily frequency of data collection is similar within years.

Figure 2.4: Distribution of survey responses by interview date



Note: Density of respondents by date they started filling-in the USOC questionnaire relative to the general election date (dashed vertical line).

Table 2.3: Summary statistics (individual-level analysis)

	Before general election					After general election				
VARIABLES	N	Mean	Sd	Max	Min	N	Mean	Sd	Max	Min
Do not support any party	30,441	0.683	0.465	1	0	56,542	0.633	0.482	1	0
Do not feel close to any party	20,716	0.723	0.448	1	0	35,675	0.692	0.462	1	0
Would vote for no party tomorrow	12,171	0.400	0.490	1	0	21,371	0.398	0.490	1	0
Pollmargin$_{w1}$	30,437	0.0485	0.0372	0.117	0	51,407	0.0546	0.0322	0.140	0
# of polls$_{w1}$	30,441	12.41	6.152	28	0	56,542	4.831	3.652	23	0
Pollmargin$_{w2}$	30,441	0.0493	0.0374	0.113	0	56,455	0.0568	0.0322	0.140	0.0006
# of polls$_{w2}$	30,441	21.92	10.75	47	2	56,542	9.554	7.166	44	0
Pollmargin$_{w3}$	30,441	0.0500	0.0375	0.112	0	56,542	0.0561	0.0306	0.120	0
# of polls$_{w3}$	30,441	30.57	14.50	67	6	56,542	14.79	11.31	64	1
Pollmargin$_{w4}$	30,441	0.0507	0.0377	0.120	0	56,542	0.0556	0.0295	0.115	0.0003
# of polls$_{w4}$	30,441	38.55	17.57	89	11	56,542	20.44	15.74	83	2
Adj. margin$_{t-1}$	26,353	0.370	0.225	0.883	0.0011	49,012	0.364	0.223	0.883	0.0011

Notes: All margins are in absolute terms.

The panel used in the analysis combines the questions just described with previous election *Adj.margin* and other electoral outcomes for the constituency where the respondent resides, as well as *Pollmargin* and the corresponding number of polls. *Pollmargin* is now constructed averaging all opinion polls individuals were exposed to during a one to four weeks window preceding their interview date. Table 2.3 illustrates descriptive statistics. Looking at the first three indicator variables, there is a significant level of disengagement among respondents, which is more pronounced before elections. On the one hand, opinion polls margins are on average smaller, display higher variability and are more numerous before elections. On the other, polls margins mean and variances are similar across windows of different lenght and, unsurprisingly, the larger the window the higher the number of polls respondents are exposed to.

2.2.4 Empirical approach

To test the hypothesis that opinion poll information interacts with voters' local historical preferences and thus significally impacts electoral outcomes, I consider the following specification:

$$y_{c,t} = \beta Pollmargin_{w,t} * Adj.margin_{c,t-1} + \delta Adj.margin_{c,t-1} + \gamma' X_{c,t} + \epsilon_{c,t} \quad (3)$$

where subscripts indicate constituency c, general election t and a weekly window w before the election day.[18] The dependent variable y is either turnout or HHI; $X_{c,t}$ is a vector of controls that varies by specification (i.e. constituency, year or region-by-year fixed effects). The β coefficient captures the mechanism under investigation. Given that both $Pollmargin_{w,t}$ and $Adj.margin_{c,t-1}$ are measured before the vote is realized, I can exclude issues of reverse causality. Different fixed effects rule-out: (a) time invariant constituency specific factors (e.g. geographic factors); (b) election specific effects (e.g. intensity of national campaign or perceived importance of the election);[19] and (c) relevant circumstances specific of a certain region during a given election (e.g. strength of local parties). This specification cannot exclude that aggregate results may be driven by factors specific to a certain constituency in a given election. However, coherent evidence paired with party level analysis (section 2.3.3) and further individual level evidence (section 2.3.4) corroborate the main strategy.

To test whether polls and previous electoral results have a joint impact on party specific outcomes, I estimate the following model:

$$y_{p,c,t} = \sum_{i \in \{0,1\}} \sum_{j \in \{0,1\}} \beta_{ij} * Adj.margin_{c,t-1} * I_{p,c,t,i,j} + \gamma' X_{p,c,t} + \epsilon_{p,c,t} \quad (4)$$

where y are party vote shares and probability of winning (i.e. an indicator for whether that party candidate becomes the new MP), and subscript p indicates either Labour or Conservative party. $I_{p,c,t,i,j}$ is an indicator for the group a party can belong to (in some constituency for some election). Specifically: (a) $I_{p,c,t,0,0}$ takes value one if the party is neither the incumbent at

[18] *Pollmargin* is calculated respectively in the last, second-to-last, third-to-last or forth-to-last week preceeding the election day.

[19] Year fixed effects are collinear with covariates varing at the national level over time, e.g. *Pollmargin* alone cannot be included in the current specification. Subsequent individual level analysis allows to separately identify *Pollmargin*.

the local level nor is leading national polls; (b) $I_{p,c,t,0,1}$ takes value one if the party is not the incumbent at the local level but is leading national polls; (c) $I_{p,c,t,1,0}$ takes value one if the party is the incumbent at the local level but is predicted to lose at the national level; finally (d) $I_{p,c,t,1,1}$ takes value one if the party is the incumbent at the local level and is also predicted to win at the national level. The coefficients of interest are β_{ij}. $X_{p,c,t}$ is a vector of controls that includes: an indicator for whether the party is the local incumbent; an indicator for whether the party is leading national polls; and an indicator for whether the party is both the local incumbent and the national polls leader. In addition, $X_{p,c,t}$ can here include two more sets of fixed effects than equation (3): party level indicators and constituency-by-year fixed effects. The most demanding specification rules out that results are driven by factors specific to a constituency in a certain general election, e.g. the strength of the local campaign (more on this in section 2.3.3). Holding all these factors fixed, it is difficult to argue that other factors are affecting all outcomes, at different level of analysis, in a similar way. Hence, the coefficients of interest should capture a causal impact of the interaction between polls and local preferences.

To corroborate the main results, I perform an analysis similar to that in equation (3), but exploiting individual level variation in the following model:

$$\begin{aligned} y_{i,c,t} = {} & \beta Pollmargin_{i,w,t} * Adj.margin_{c,t-1} + \\ & + \lambda Pollmargin_{i,w,t} + \delta Adj.margin_{c,t-1} + \gamma' X_{c,t} + \epsilon_{i,c,t} \end{aligned} \tag{5}$$

where y is either an indicator for whether the respondent i answered that she does not support any party; or a dummy taking value one if the interviewee responded that she neither supports, nor feels close, nor would vote for any party tomorrow. These outcome variables proxy individuals willingness to participate in the election. $X_{c,t}$ is a vector of controls that varies across specifications (i.e. constituency effects, year effects or both) and captures time invariant constituency specific factors as well as election specific features. *Pollmargin* is the exposure to a certain time window of opinion polls preceding the interview starting date of each respondent. Each individual is therefore exposed to a quasi-random polling information at the time of the interview, exogenous to her political engagement.[20] *Adj.margin* is a proxy for safeness of each

[20] In the USOC survey, each monthly sample is a representative random sample of the total population.

respondents' constituencies. The individual variation enables to separately identify the impact of the two margins.

2.3 Results

Results from the above specifications are presented in this section. I begin by focusing on how pre-election polls and the electoral history of a constituency affect voters' participation. Next, I show how these factors impact the concentration of vote shares in a constituency. I then report evidence of the link between party level outcomes (i.e. vote share and probability of victory) and the explanatory variables of interest. Finally, I present individual-level evidence supporting the main effect on participation.

2.3.1 Voters' Participation

I start presenting motivating evidence that both the margin predicted by the national opinion polls *and* the margin in the previous general election at the constituency level capture significant variation in voters' participation.

Columns (1) to (4) of Table 2.4 display correlations between turnout and national polls at different points in time (i.e. $w1$ indicates the week preceding the election, etc.). As polls vary at the national level, these specifications can only control for time invariant constituency characteristics. Thus, coefficients should be interpreted with caution. The estimates suggest that the more opinion polls predict a non-competitive election, the lower is voters' participation.[21] Columns (5) and (6) examine the link between turnout and margin in previous elections. Since this explanatory variable is measured at the constituency level, the models can absorb year and region-by-year fixed effects. The reported coefficients indicate that safer seats (i.e. constituencies where previous election margin is large) are associated with lower turnout.[22]

[21] One standard deviation wider predicted margin is associated with a decrease in turnout which varies between 0.11 p.p. and 0.97 p.p.. These results are comparable in sign and magnitude to those found by other scholars, e.g. Bursztyn et al. (2017).

[22] One standard deviation increase in safeness of a constituency is associated with a decrease in turnout between 9.4 p.p. and 10.7 p.p..

Table 2.4: Opinion polls, safeness of a constituency and turnout

	Dep. var.: Turnout					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w1}	-0.0178**					
	(0.0071)					
Pollmargin_{w2}		-0.0849***				
		(0.0073)				
Pollmargin_{w3}			-0.1503***			
			(0.0070)			
Pollmargin_{w4}				-0.0571***		
				(0.0060)		
Adj. margin_{t-1}					-0.0484***	-0.0425***
					(0.0040)	(0.0043)
Constituency FE	X	X	X	X	X	X
Year FE					X	
Region*Year FE						X
Observations	5,599	5,599	5,599	5,599	4,676	4,676
R-squared	0.4286	0.4323	0.4423	0.4310	0.9240	0.9458

Notes: *Turnout* is the ratio between the total number of votes and the number of eligible voters of a constituency. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in a certain time frame, subscripts indicate a specific week before the election date (1=last, ..., 4=fourth to last). *Adj.margin* is the absolute difference between Conservative and Labour constituency-level vote shares in the previous general election, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

While these two sets of results provide evidence of two quite intuitive relationships, they are only partially compelling. In a context like that of the UK general elections, where local MPs are elected via a first-past-the-post system, it is reasonable to hypothesize that the information spread by the opinion polls may affect electoral outcomes differently depending on previous local preferences. To test this hypothesis I now focus on the joint impact of the two factors.

Table 2.5 presents estimates of equation (3), where the dependent variable is local turnout. Odd columns include constituency and year fixed effects, even columns replace year dummies with region-by-year fixed effects. Across specifications coefficients are negative and significant, suggesting that the less competitive the election is predicted to be, the lower is turnout. Even more so in safer constituencies. In addition, the effect of the polls is stronger the closer the election date, i.e. when the information is relevant for the participation decision. The coefficient of $Adj.\ margin_{t-1}$ is also negative and significant across specifications, indicating that participation is lower in safe seats even when polls predict a tight race. Reassuringly, the magnitude of the coefficients is only marginally influenced by different fixed effects.[23]

Table 2.5: Opinion polls, safeness of a constituency and their joint effect on turnout

	Dep. var.: Turnout							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pollmargin$_{w1}$ * Adj. margin$_{t-1}$	-0.1775***	-0.1763***						
	(0.0291)	(0.0275)						
Pollmargin$_{w2}$ * Adj. margin$_{t-1}$			-0.1585***	-0.1716***				
			(0.0291)	(0.0273)				
Pollmargin$_{w3}$ * Adj. margin$_{t-1}$					-0.1112***	-0.1281***		
					(0.0244)	(0.0226)		
Pollmargin$_{w4}$ * Adj. margin$_{t-1}$							-0.0641***	-0.0708***
							(0.0187)	(0.0173)
Adj. margin$_{t-1}$	-0.0343***	-0.0287***	-0.0354***	-0.0288***	-0.0386***	-0.0314***	-0.0422***	-0.0357***
	(0.0050)	(0.0051)	(0.0050)	(0.0051)	(0.0048)	(0.0050)	(0.0046)	(0.0047)
Constituency FE	X	X	X	X	X	X	X	X
Year FE	X		X		X		X	
Region*Year FE		X		X		X		X
Observations	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676
R-squared	0.9247	0.9463	0.9246	0.9463	0.9244	0.9461	0.9242	0.9459

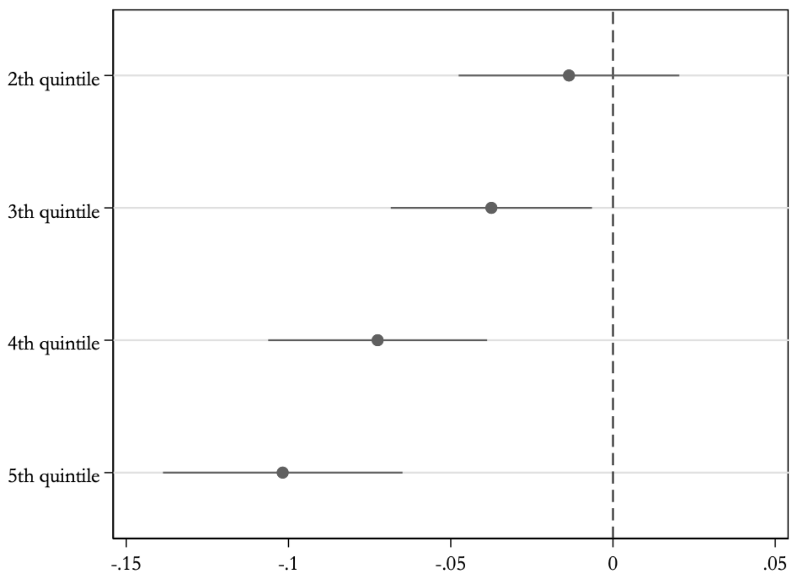
Notes: *Turnout* is the ratio between the total number of votes and the number of eligible voters of a constituency. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in a certain time frame, subscripts indicate a specific week before the election date (1=last, ..., 4=fourth to last). *Adj.margin* is the absolute difference between Conservative and Labour constituency-level vote shares in the previous general election, adjusted by the sum of those vote shares. All margins are $\in (0,1)$. Constituency-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

In terms of magnitudes, a 10 p.p. increase in the previous election margin is associated with a

[23] Results in Table B.2 show that the joint effect is stronger in constituencies where the incumbent party is also the one leading the polls.

decrease in turnout between 0.4 and 0.5 p.p. when polls predict a 10 p.p. (absolute) difference between Conservative and Labour. Instead, a 10 p.p. increase in polls margin in the most contested constituency (in previous election) is associated with a negligible reduction in voters' participation. On the other hand, the same variation in polls margin in the safest constituency leads to a reduction in turnout between 1.6 and 0.6 p.p. depending on whether the polls are released close to or far away from the election day. For this reason, I now focus on the most recent polls margins (i.e. those released in the week preceding the election). Furthermore, the electorate decision to vote vary significantly with the degree of safeness of a constituency: the following figure provides support to this claim.

Figure 2.5: Participation effect by degree of safeness of a constituency



Note: Graph displays estimated coefficients for the interaction between $Pollmargin_{w1}$ and quintiles of $Adj.margin_{t-1}$. Equivalent to the specification in column (2) of Table 2.5 .

Figure 2.5 breaks down the coefficient of the interaction term previously reported in column (2) of Table 2.5. According to the graph, the effects of the polls are (almost) linear in the quintiles of safeness distribution. Specifically, the impact for constituencies in the highest quintiles is significantly stronger compared to constituencies in the lowest quintile.

2.3.2 Vote shares concentration

This section shifts focus towards the concentration of vote shares. This index considers every competing party in a constiutency, therefore allowing more general conclusions.

Table 2.6 displays estimates from equation (3), where the dependent variable is the sum of squares of constituency vote shares. Looking at the whole sample, the reported coefficients in column (1) and (2) indicate that safer seats are associated with greater concentration of votes. However, this effect is significantly reduced the larger the predicted poll margin, remaing positive on average. Yet, the negative coefficient on the interaction term seems to mask evident heterogeneity.

Table 2.6: Opinion polls, safeness of a constituency and their joint effect on HHI

	Dep. var.: HHI							
	All sample		Incumbent party is leading polls		Incumbent party is not leading polls		Follower party is leading polls	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pollmargin_{w1} * Adj. margin_{t-1}	-0.2651***	-0.1602***	0.2829***	0.3549***	-0.5428***	-0.5498***	-0.8185***	-0.6905***
	(0.0514)	(0.0503)	(0.0710)	(0.0747)	(0.0562)	(0.0629)	(0.1248)	(0.1361)
Adj. margin_{t-1}	0.0606***	0.0474***	0.0649***	0.0296**	0.0908***	0.0896***	0.0900***	0.1071***
	(0.0065)	(0.0067)	(0.0121)	(0.0116)	(0.0133)	(0.0123)	(0.0219)	(0.0227)
Constituency FE	X	X	X	X	X	X	X	X
Year FE	X		X		X		X	
Region*Year FE		X		X		X		X
Observations	4,676	4,676	2,306	2,306	2,370	2,370	1,239	1,239
R-squared	0.6747	0.7801	0.8285	0.8831	0.7424	0.8456	0.8920	0.9200

Notes: HHI is the sum of squares of constituency-level vote shares for all parties. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the last week before the election date. *Adj.margin* is the absolute difference between Conservative and Labour constituency-level vote shares in the previous general election, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Incumbent parties are defined at the constituency level. Constituency-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Indeed, repeting the analysis on different sub-samples unveils a more complex picture. Examining constituencies where the incumbent party is also leading the national polls (columns 3 and 4), I observe an increase in the concentration index in safer seats, which is magnified by the national polls predicting a less competitive election. These result could be explained with decreased relative turnout by supporters of the parties opposing the incumbent. Conversely, when the party ahead in the national polls does not coincide with the incumbent one (columns 5 and 6), the coefficient of the interaction is negative and larger in magnitude if compared to the previous case. This indicates that concentration of vote shares in safer seats is diminished when the polls report a larger lead in favour of one of the incumbent's opponents. This may reflect a scenario where the votes cast for parties opposing the incumbent become more fragmented at the local level. Similarly, I observe that larger polls margin reduces the positive effect of safeness on the HHI index also in constituencies where the party that came second in the previous election (the follower) is currently ahead in the national polls (columns 7 and 8).[24]

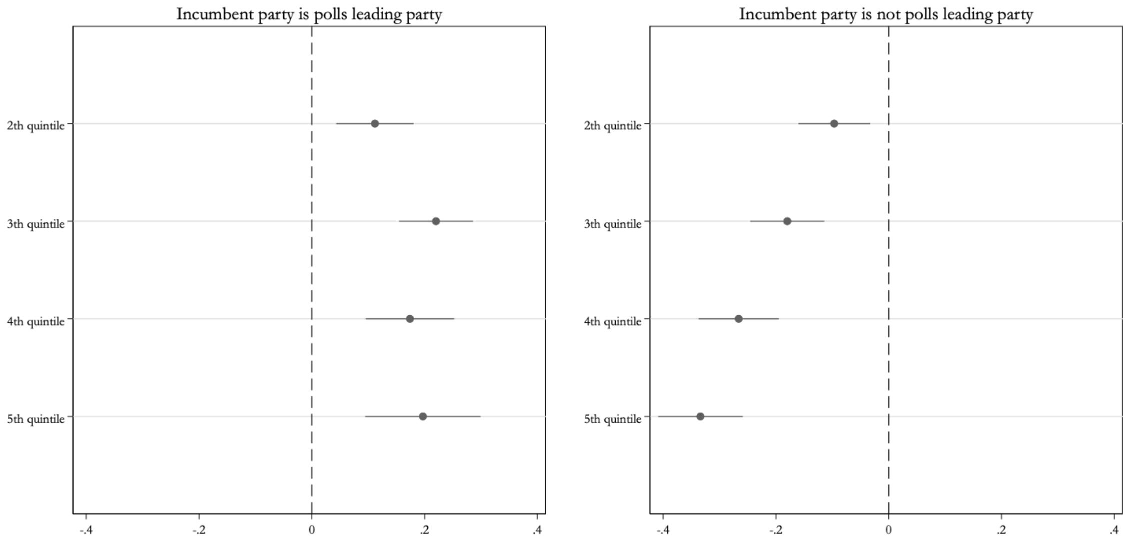
In terms of magnitudes, when referring to the cases reported in columns (3) and (4), I observe that one standard deviation increase in safeness increases the concentration index relative to its mean between 3.2 and 4.9 percent, given average polls margin; on the other hand, one standard deviation increase in the margin reported by the polls raises the concentration index relative to its mean between 1.4 and 1.8 percent, in a constituency with average previous election margin. Moving the attention to columns (5) and (6), I note that an additional standard deviation in safeness, given average polls margin, induces a 3 percent upward shift in HHI, relative to its mean; instead, a one standard deviation increment in polls margin, considering an average level of safeness, is associated with a 3 percent drop in concentration relative to its mean.

In general, the illustrated heterogeneity suggests the following: concentration of votes always increases in safer seats; larger polls margins enhance this effect when the information they provide is coherent with the recent electoral history of a constituency, while they significantly attenuate the impact of safeness otherwise.[25]

[24] Note that as I drop a considerable number of observations in columns (5) to (8), estimates precision is negatively affected; thus I cannot reject the hypothesis that the effect of $Pollmargin_{w1} * Adj.margin_{t-1}$ is the same in the different specifications.

[25] In the latter case, there exist levels of polls margin such that the overall effect of increased safeness becomes negative.

Figure 2.6: HHI effect by degree of safeness of a constituency



Note: Graph displays estimated coefficients for the interaction between $Pollmargin_{w1}$ and quintiles of $Adj.margin_{t-1}$. Equivalent to the specification in column (4) and (6) of Table 2.6 .

Figure 2.6 breaks down the joint effect of polls margin and safeness by quintiles of safeness distribution. Estimates are equivalent to those in columns (4) and (6) of Table 2.6. Constituencies experience a similar impact on vote share concentration when the local incumbent party is leading the national polls. In the opposite scenario, the effect appears significally stronger in safer seats.

2.3.3 Vote shares and probability of winning

The analysis of turnout and HHI only partially explains how votes are redistributed across political forces. In what follows I shed light on how national polls together with electoral history of a constituency affect party level outcomes.

Table 2.7: Previous election margin, vote share and winning probability

	Dep. var.:					
	Vote Share			Pr. of Winning		
	(1)	(2)	(3)	(4)	(5)	(6)
Adj. margin$_{t-1}$ * I_{Inc} * I_{Pl}						
Incumbent=0 & Pollleader=0	-0.2958***	-0.2806***		-0.1579***	-0.1360***	
	(0.0096)	(0.0105)		(0.0267)	(0.0288)	
Incumbent=0 & Pollleader=1	-0.3522***	-0.3588***	-0.2555***	-0.5241***	-0.5514***	-0.6334***
	(0.0090)	(0.0099)	(0.0452)	(0.0334)	(0.0348)	(0.1274)
Incumbent=1 & Pollleader=0	0.2509***	0.2484***	0.3724***	0.5301***	0.5141***	0.4757***
	(0.0090)	(0.0097)	(0.0465)	(0.0386)	(0.0408)	(0.1433)
Incumbent=1 & Pollleader=1	0.2353***	0.2579***	0.5887***	0.1623***	0.2023***	0.4134***
	(0.0090)	(0.0086)	(0.0112)	(0.0321)	(0.0322)	(0.0509)
Controls	X	X	X	X	X	X
Party FE	X	X	X	X	X	X
Constituency FE	X	X		X	X	
Year FE	X			X		
Region*Year FE		X			X	
Constituency*Year FE			X			X
Observations	9,352	9,352	9,352	9,352	9,352	9,352
R-squared	0.8212	0.8299	0.8674	0.6922	0.6997	0.7308

Notes: Dependent variables are: constituency-level party vote shares, and an indicator for whether the party won the constituency race. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the last week before the election date. *Adj.margin* is the absolute difference between Conservative and Labour constituency-level vote shares in the previous general election, adjusted by the sum of those vote shares. All margins are $\in (0,1)$. I_{Inc}=indicator for whether a party is the constituency-level incumbent and I_{Pl}=indicator for whether a party is leading the polls. Controls include: I_{Inc}; I_{Pl}; and their interaction. Constituency-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

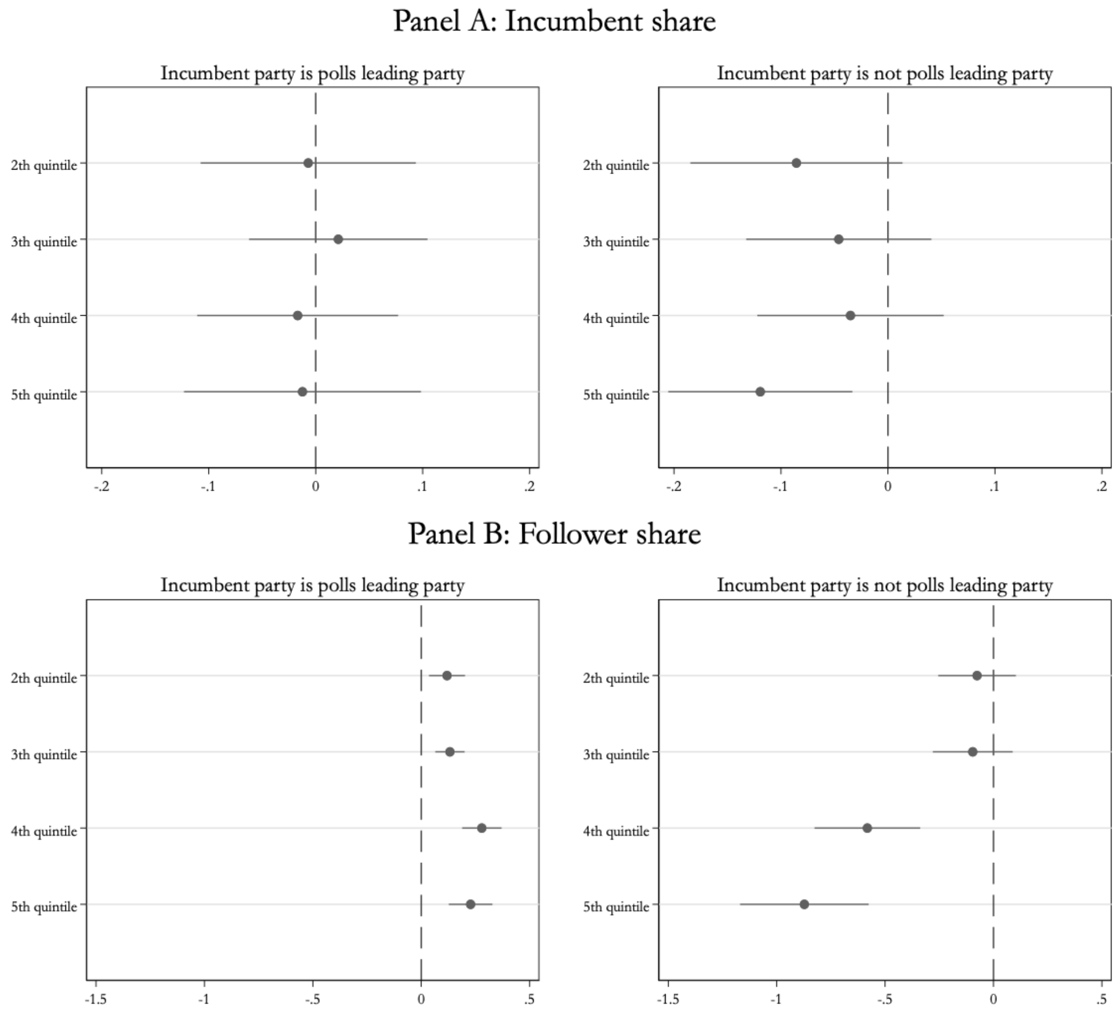
Table 2.7 reports estimates of equation (4), where the dependent variables are either party vote shares or an indicator for the winning party. Given the additional party level variation, I can now include constituency-by-year fixed effects, which allow to control for potential confounders, such as constituency specific intensity of the campaign in a given election, or the presence of a specific candidate for local MP (columns 3 and 6). Note that the incumbent party is the one associated to the constituency MP elected in the previous general election, while either Conservative or Labour are the only parties leading national opinion polls as in Figure 2.2.

What consistently emerges across specifications is the following. First, local non-incumbent parties that are behind in the polls get increasingly lower vote shares and probability of victory, the safer is the constituency. Second, a similar effect is reported for local non-incumbents that are leading the national polls. It appears that, no matter the national trends, if the local incumbent party was strongly favoured in the past, local opponents will revert the order with difficulty. Third, if local incumbents obtained a solid victory in the previous election, their vote shares and chances of victory will increase independently of whether their party is leading the national polls. Note that the increase in chances of victory induced by an equal increase in safeness is systematically higher for incumbent parties that are behind in the polls. Fourth, the enhanced model in columns (3) and (6) does not have a significant impact on the estimated coefficients of interest.

The results just described provide further insights. Cases where the incumbent party and the party leading the polls do not coincide constitute examples of possible upset victories, as polls predictions may not be met at the constituency level. A possible explanation is that voters in a constituency which is safe could fear that another party may win the local race due to the predicted scenario at the national level; the uncertainty may motivate higher relative participation by the supporters of the local incumbent. In addition, results from Table 2.6 suggest this would go hand in hand with a more fragmented opposition. Conversely, when results appear to be quite certain (i.e. incumbent and poll leading party coincide) part of the electorate may think their vote would not make much of a difference and eventually not turn out at the ballots. This may be especially true for supporters of minor parties, consistently with Table 2.6. These results are also aligned with finding a negative effect on turnout in Table 2.5, which is even stronger when analysing this same sub-sample (see Table B.2).

I now present graphical analysis where I display the effect of the polls by quintiles of safeness distribution on these two outcomes of interest, considering distinctly (local) incumbent and follower parties. Estimates underlying the next figures are available in table format in the appendix (see Table B.4).

Figure 2.7: Share effect by degree of safeness of a constituency



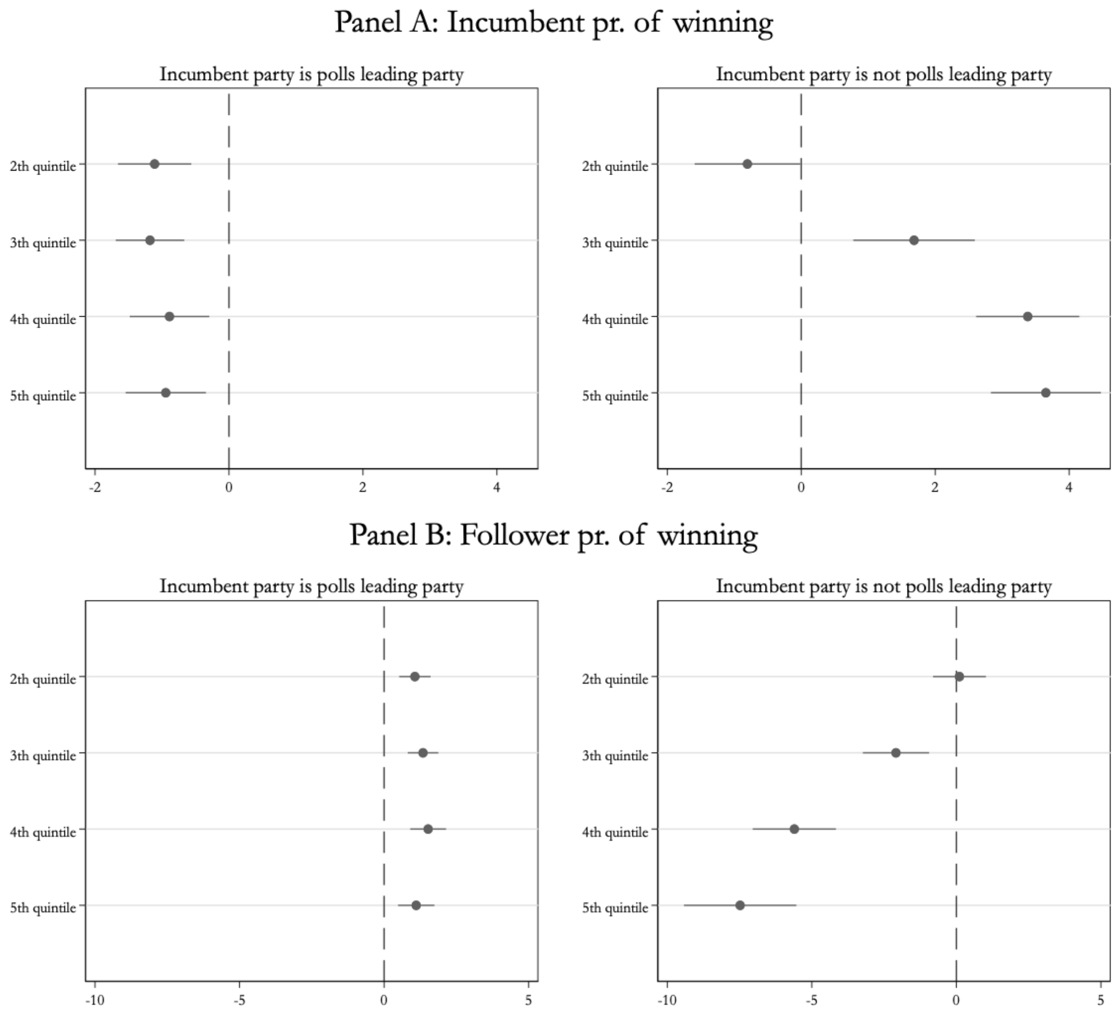
Note: Graph displays estimated coefficients for the interaction between $Pollmargin_{w1}$ and quintiles of $Adj.margin_{t-1}$. Equivalent to the specification in column (1-4) of Table B.4 .

Panel A of Figure 2.7 shows that vote shares for incumbent parties are not statistically affected by variations in polls margin and do not differ systematically across safe and non-safe constituencies independently of whether their party is leading the polls (see also Table B.4, columns 1 and 2).[26] The left graph of Panel B, together with results from column (3) of Table B.4, shows that polls margin has a positive effect on the vote shares of the followers when the incumbent party is leading the polls, and the effect appears to be slightly stronger in safer constituencies. On the right of Panel B (i.e. considering constituencies where the incumbent party is behind in the

[26]With the exception of constituencies in the highest quintile of safeness distribution.

polls) I observe that polls margin has a negative impact on the vote shares of the followers, and that the interquintile difference in the estimated impact is more pronounced, with coefficients being larger in safer seats. However, whether polls margin affect the final results is not clear from looking at vote shares alone. I thus replicate these graphs focusing on the chances of victory.

Figure 2.8: Pr. of winning effect by degree of safeness of a constituency



Note: Graph displays estimated coefficients for the interaction between $Pollmargin_{w1}$ and quintiles of $Adj.margin_{t-1}$. Equivalent to the specification in column (5-8) of Table B.4 .

Figures 2.7 and 2.8 deliver consistent insights which can be interpreted in light of previous sections. First, when the local incumbent party is leading the polls, the reduction in turnout associated with larger polls margin seems detrimental for the incumbent party and beneficial for the follower. While this does not fully emerge by looking at vote shares, it is quite evident in the analysis of the probability of victory. All in all, these figures go along with the findings in Table 2.6, which report enhanced vote shares concentration. Second, when the local incumbent is behind in the polls, it consistently emerges that incumbents in safer seats gain more both in terms of vote shares and probability of victory as the polls predict a larger gap in favour of opposing parties. This could be explained by two complementary factors: on the one hand, supporters of the incumbents may turn out more in response to the rising success of the opposition; on the other hand, the composition of votes cast in favour of opposing parties may change, becoming more fragmented. As a consequence, if the latter effect offsets the former, polls prediction may lead to lower concentration of vote shares, consistently with the results displayed in Table 2.6. For instance, consider the following numerical example, as displayed in Table 2.8. Take the hypothetical scenario presented in column (1), of a constituency where the previously elected MP is Labour and the national polls predict a positive margin in favour of the Conservative party. In column (2) I show how an increase in polls margin in favour of the Conservative party may change the electoral outcomes.

Table 2.8: Numerical example

Case: Incumbent party = Labour; Poll leading party = Conservative		
	Pollmargin ($Con > Lab$)	*Pollmargin* ($Con >> Lab$)
	(1)	(2)
Turnout	71%	68%
Share Lab	52%	53%
Share Con	27%	21%
Share LD	20%	21%
Share UKIP	1%	5%
HHI	0.38	0.37

Notes: The table illustrates a hyphotetical scenario which assumes a constituency with a previously elected Labour MP (constituency-level incumbent) and national polls favouring the Conservative party. The opinion poll margin is more competitive in column (1) and less competitive in column (2). Coherent with the evidence presented above, column (2) thus shows possible changes in the outcome variables listed for an increase in *Pollmargin*.

Consistently with results in Table 2.5, turnout would diminish. Then, as discussed in the party-level analysis, I would observe an increase in the vote share for the incumbent party (Labour in this example) and a reduction in that of the follower. Moreover, in line with findings reported in Table 2.6, I could observe higher fragmentation of vote shares, thus a lower level of concentration.

Summing up, the evidence reported so far highlights the presence of a link between voters' participation, vote shares distribution and outcomes at the party level, as they are all coherently affected by national polls and the electoral history of a constituency.

2.3.4 Individual-level evidence

To this point I used aggregate data to show that electoral history of a constituency and national opinion polls jointly influence voters' behaviour. As a final step, I test the combined influence of these two factors directly looking at their impact on individual variation in political engagement, as a proxy for willingness to participate in general elections.

Table 2.9 presents estimates of equation (5) where the dependent variable is a dummy taking value one if the respondent does not support any party. The coefficient of interest is that of the interaction between the previous election margin for the constituency of the respondent and the national polls margin that she is exposed to 1 week prior her interview. Estimates are generally sensitive to the inclusion of fixed effects and $Pollmargin$ as control, thus magnitudes should be interpreted with caution.

Panel A focuses on individuals interviewed before the general election date. Interaction coefficients are positive and often significant suggesting that the less competitive the election is predicted to be, the higher the chance of voters not supporting any party. More so in safer constituencies. Panel B illustrates estimates for the sample of individuals interviewed after the elections. The interaction term turns now negative or insignificant, suggesting the main impact on participation emerges only when expected, if the information provided by polls is relevant for the voting decision. The negative coefficient may imply some form of ex-post regret from little political engagement.

Table 2.9: Opinion polls margin interacted with previous election margin and political support (1 week window)

	Panel A - Dep. var.: Do not support any party (**pre election**)					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w1} * Adj. margin_{t-1}	0.903***	0.870**	0.479**	0.564	0.591*	0.568
	(0.336)	(0.381)	(0.199)	(0.402)	(0.355)	(0.403)
	[0.007]	[0.023]	[0.017]	[0.161]	[0.096]	[0.158]
Pollmargin_{w1}		0.051		-0.041		0.031
		(0.269)		(0.170)		(0.279)
		[0.851]		[0.810]		[0.910]
Adj. margin_{t-1}	-0.040*	-0.038	0.029	0.025	0.024	0.024
	(0.023)	(0.024)	(0.036)	(0.039)	(0.039)	(0.039)
	[0.079]	[0.111]	[0.431]	[0.516]	[0.545]	[0.539]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	26,352	26,352	26,352	26,352	26,352	26,352
R-squared	0.000	0.000	0.046	0.046	0.046	0.046

	Panel B - Dep. var.: Do not support any party (**post election**)					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w1} * Adj. margin_{t-1}	-0.493**	-0.648*	-0.528**	-0.295	-0.419*	-0.299
	(0.230)	(0.367)	(0.207)	(0.389)	(0.236)	(0.389)
	[0.032]	[0.078]	[0.011]	[0.448]	[0.076]	[0.441]
Pollmargin_{w1}		0.086		-0.114		-0.065
		(0.163)		(0.161)		(0.173)
		[0.598]		[0.480]		[0.708]
Adj. margin_{t-1}	0.031	0.040	-0.002	-0.016	-0.003	-0.011
	(0.020)	(0.026)	(0.037)	(0.042)	(0.037)	(0.042)
	[0.125]	[0.126]	[0.960]	[0.697]	[0.934]	[0.790]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	44,272	44,272	44,272	44,272	44,272	44,272
R-squared	0.000	0.000	0.027	0.027	0.027	0.027

Notes: The dependent variable is an indicator taking value one if the respondent does not support any party. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the last week before the respondent interview date. *Adj.margin* is the absolute difference between Conservative and Labour vote shares in the previous general election in the respontent's constituency, adjusted by the sum of those vote shares. All margins are $\in (0,1)$. Constituency-level clustered standard errors are presented in parentheses, p-values in brackets, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Mirroring results emerge in Table 2.10, where equation (5) is estimated using a different dependent variable, i.e. indicator for whether the respondent does not support nor feel close to a political party and would not vote for any. Results are very similar when expanding the opinion polls window individuals are exposed to (see appendix Table B.5 to B.10). The evidence just presented is coherent to the aggregate level analysis of section 2.3.1.

Table 2.10: Opinion polls margin interacted with previous election margin and political engagement (1 week window)

	Panel A - Dep. var.: Do not support, feel close or vote a party tomorrow **(pre election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w1} * Adj. margin_{t-1}	0.827***	0.478	2.005***	0.273	0.511*	0.243
	(0.290)	(0.319)	(0.161)	(0.315)	(0.279)	(0.315)
	[0.005]	[0.134]	[0.000]	[0.388]	[0.067]	[0.440]
Pollmargin_{w1}		0.536**		0.834***		0.368
		(0.261)		(0.138)		(0.261)
		[0.040]		[0.000]		[0.159]
Adj. margin_{t-1}	-0.060***	-0.044***	-0.084***	-0.016	-0.015	-0.008
	(0.017)	(0.017)	(0.029)	(0.030)	(0.029)	(0.030)
	[0.000]	[0.008]	[0.004]	[0.600]	[0.608]	[0.792]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	26,352	26,352	26,352	26,352	26,352	26,352
R-squared	0.009	0.009	0.058	0.060	0.060	0.060

	Panel B - Dep. var.: Do not support, feel close or vote a party tomorrow **(post election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w1} * Adj. margin_{t-1}	-0.594***	-0.169	-1.685***	0.007	-0.675***	-0.036
	(0.187)	(0.287)	(0.173)	(0.300)	(0.188)	(0.298)
	[0.002]	[0.556]	[0.000]	[0.981]	[0.000]	[0.903]
Pollmargin_{w1}		-0.236*		-0.825***		-0.347***
		(0.121)		(0.123)		(0.127)
		[0.051]		[0.000]		[0.006]
Adj. margin_{t-1}	0.032*	0.008	0.086**	-0.018	0.074***	0.032
	(0.017)	(0.022)	(0.034)	(0.035)	(0.028)	(0.032)
	[0.054]	[0.718]	[0.013]	[0.604]	[0.009]	[0.321]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	44,272	44,272	44,272	44,272	44,272	44,272
R-squared	0.010	0.010	0.037	0.038	0.042	0.042

Notes: The dependent variable is an indicator taking value one if the respondent does not support, feel close nor would vote for any party. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the last week before the respondent interview date. *Adj.margin* is the absolute difference between Conservative and Labour vote shares in the previous general election in the respontent's constituency, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, p-values in brackets, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

2.4 Conclusion

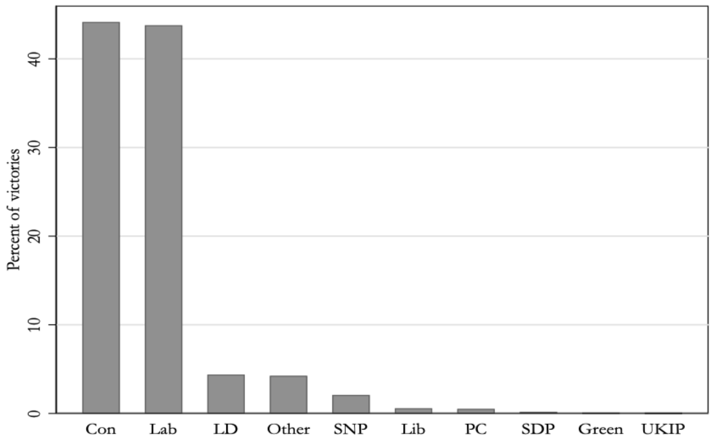
This work expands the literature on the causal effect of anticipated election closeness on voter participation. I specifically investigate the role of opinion polls in the context of UK general elections for two reasons. First, despite their national nature, voters express electoral preferences for their local MP which allows to use constituency-by-year variation in previous election margin. Second, the institutional stability of the electoral system allows to study the impact of polls in a historical context. Findings suggest that individuals decision to vote depends not only on political orientation, but on the combination of the perceived tightness of the race at the national level (as inferred by the polls) and the electoral history of her constituency (as measured by the local margin of the incumbent party in previous elections). The decision to turnout has then repercussions on electoral outcomes being beneficial to some party and detrimental to others.

I first present consistent evidence that polls predictions and local preferences interact with one another. Precisely, the less competitive the election is predicted to be, the lower is turnout and the effect is larger the safer the seat. This further affects the composition of the electorate increasing the concentration of shares when the two information are aligned and reducing it otherwise. Sensing this could shape final results, I dug deeper into local party outcomes. Evidence shows that, when the local incumbent party is leading the polls, the reduction in turnout associated with larger polls margin seems detrimental for the incumbent and beneficial for the follower, which goes along with enhanced vote shares concentration. On the other hand, when the local incumbent is behind in the polls, incumbents in safer seats gain more as the polls predict a larger gap in favour of opposing parties. This could be explained by a non-reduction in incumbent support coupled with a fragmentation of the opposition, leading to a reduction in concentration of shares. Finally, I exploit quasi-random individual-level exposure to opinion polls to corroborate the above findings that the interaction of polls predictions and past local preferences influences voters' political engagement. Relationship which emerges only before an election, when opinion polls are relevant to voters.

In synthesis, the extensive set of findings points coherently in one direction: national opinion polls and the political roots of a constituency play a key role in shaping local electoral results. This underlines the importance of welfare considerations when referring to different polling systems. This is due to opinion polls potential to shape electoral outcomes deviating from more genuine counterfactual results. In addition, it seems that the existence of safe seats, due to its impact on turnout, may result in enlarging the pool of voters who feel disenfranchised and without voice, which may foster more extreme policy positions. This could have repercussions on the quality of elected politicians and possibly lead to radical outcomes which entail strong economic consequences (e.g. the Brexit vote).

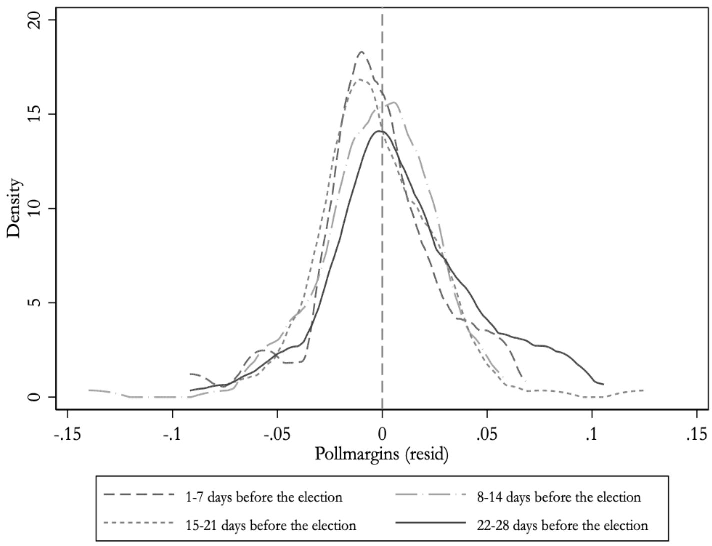
B.5 Appendix

Figure B.1: Party victories across all seats in 1983-2017 general elections



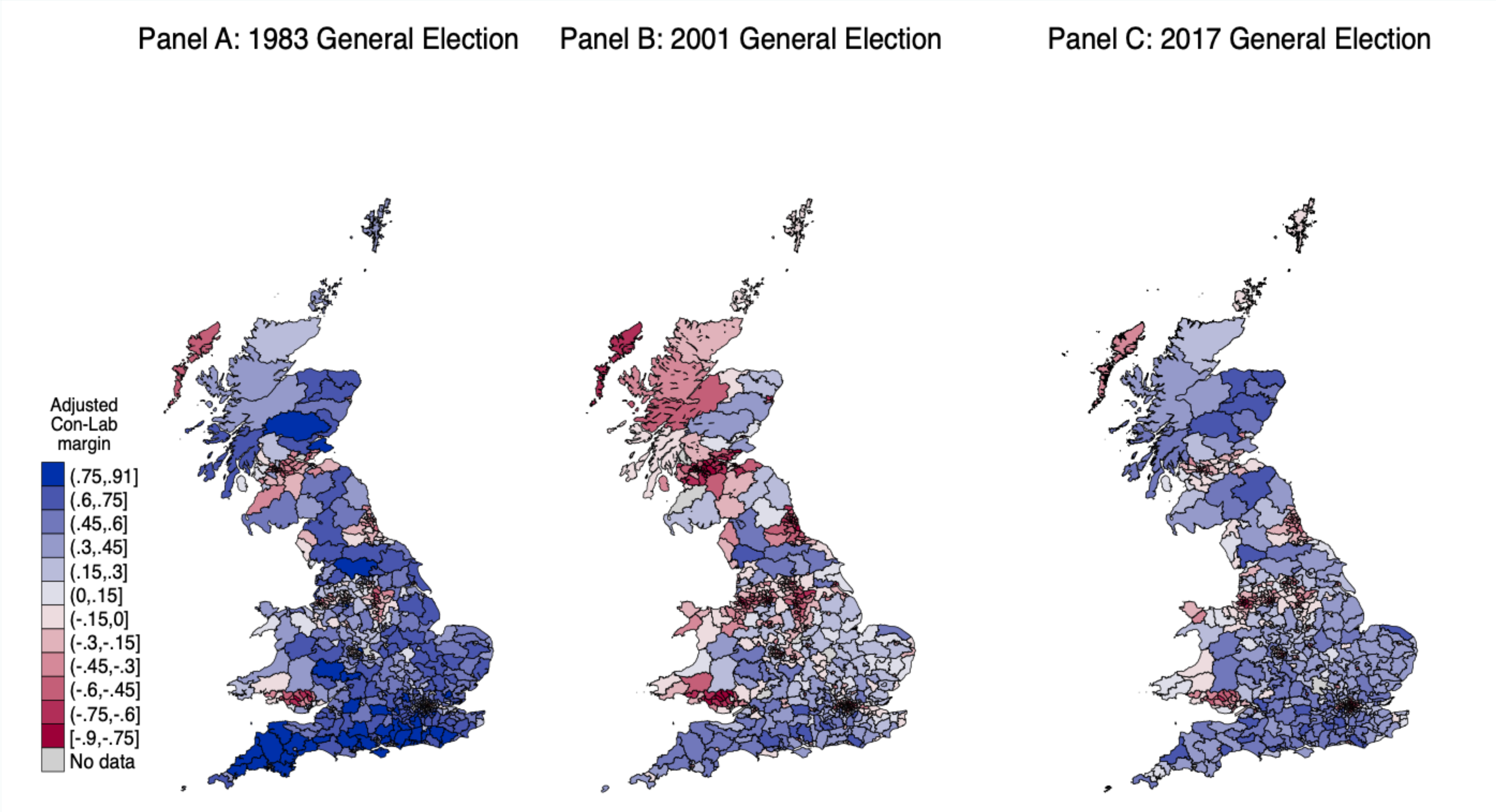
Note: Bars represent the share of winning candidates associated to each party across the full sample of constituencies (seats) across all general elections from 1983 to 2017.

Figure B.2: Variation in polls margins in different weeks preceding the elections



Note: Residual variation in the polls margins after controllinh for election fixed effects. Margins are in absolute terms.

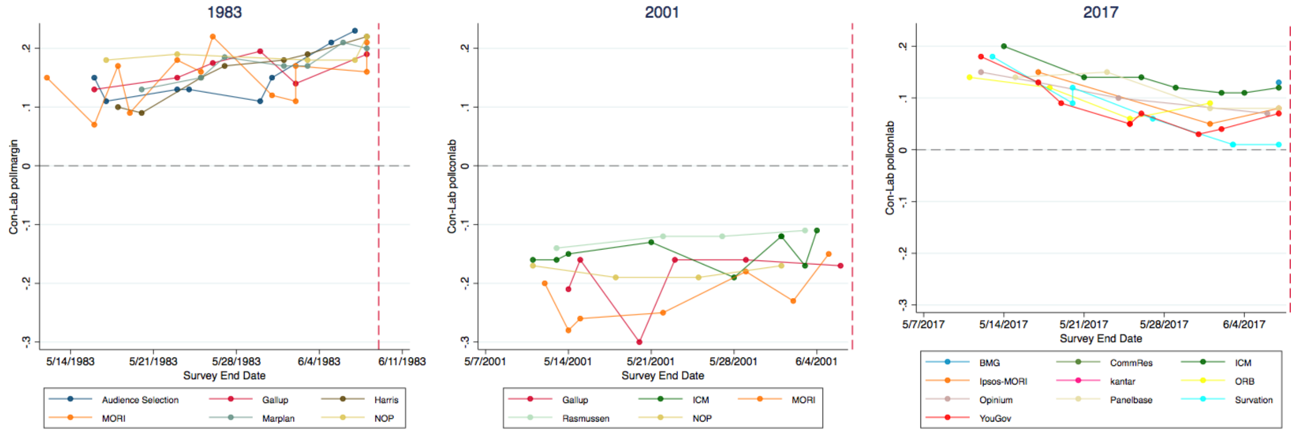
Figure B.3: Adjusted margin of victory across general elections (Conservative - Labour) in absolute terms



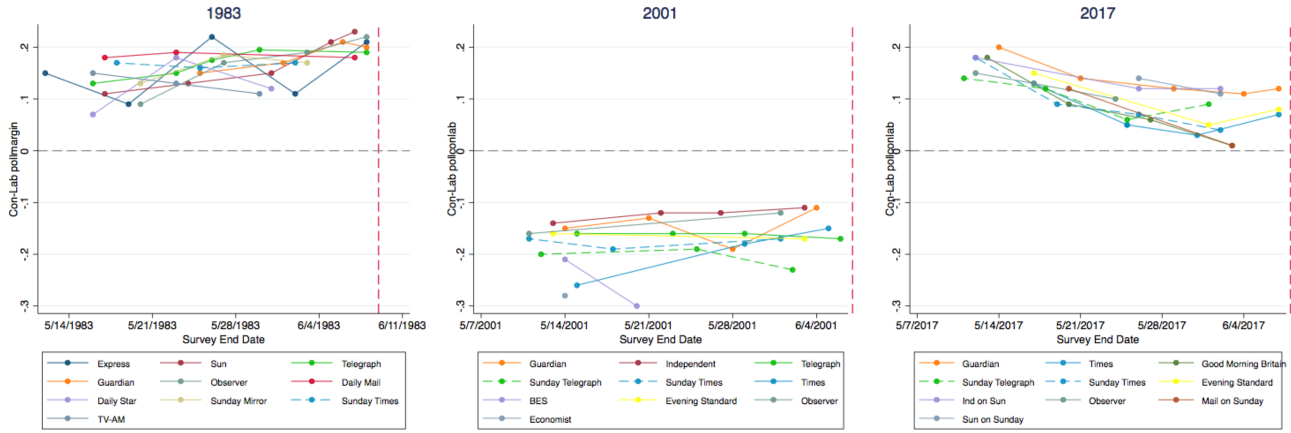
Note: Shades map the variation in vote share margin between Conservative and Labour parties across general elections, adjusted dividing by the sum of the two party shares. Blue shades refer to seats favouring the conservative candidate, red shades refer to seats favouring the labour candidate.

Figure B.4: Weekly variation in polls margins

Panel A: Variation by pollster



Panel B: Variation by publisher



Note: Estimates show opinion poll margins between Conservative and Labour parties in a given week before the general election date and across general elections. Colors represent pollsters (panel A) or publishers (panel B) associated to each estimated margin.

Table B.1: Opinion polls, safeness of a constituency and turnout

	Dep. var.: Turnout									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Pollmargin$_{w1}$	-0.0178**	-0.0024								
	(0.0071)	(0.0070)								
# of polls$_{w1}$		0.0006***								
		(0.0001)								
Pollmargin$_{w2}$			-0.0849***	-0.0522***						
			(0.0073)	(0.0074)						
# of polls$_{w2}$				0.0008***						
				(0.0001)						
Pollmargin$_{w3}$					-0.1503***	-0.1600***				
					(0.0070)	(0.0078)				
# of polls$_{w3}$						-0.0003**				
						(0.0001)				
Pollmargin$_{w4}$							-0.0571***	-0.0893***		
							(0.0060)	(0.0067)		
# of polls$_{w4}$								-0.0022***		
								(0.0002)		
Adj. margin$_{t-1}$									-0.0484***	-0.0425***
									(0.0040)	(0.0043)
Constituency FE	X	X	X	X	X	X	X	X	X	X
Year FE									X	
Region*Year FE										X
Observations	5,599	5,599	5,599	5,599	5,599	5,599	5,599	5,599	4,676	4,676
R-squared	0.4286	0.4295	0.4323	0.4341	0.4423	0.4424	0.4310	0.4400	0.9240	0.9458

Notes: *Turnout* is the ratio between the total number of votes and the number of eligible voters of a constituency. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in a certain time frame, subscripts indicate a specific week before the election date (1=last, ..., 4=fourth to last). *Adj.margin* is the absolute difference between Conservative and Labour constituency-level vote shares in the previous general election, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table B.2: Opinion polls, safeness of a constituency and their joint effect on turnout

	Dep. var.: Turnout					
	All sample		Incumbent party is leading polls		Incumbent party is not leading polls	
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w1} * Adj. margin_{t-1}	-0.1775***	-0.1763***	-0.2392***	-0.2945***	-0.0534*	-0.0717**
	(0.0291)	(0.0275)	(0.0615)	(0.0563)	(0.0302)	(0.0285)
Adj. margin_{t-1}	-0.0343***	-0.0287***	-0.0347***	-0.0476***	-0.0087	-0.0014
	(0.0050)	(0.0051)	(0.0102)	(0.0089)	(0.0085)	(0.0072)
Constituency FE	X	X	X	X	X	X
Year FE	X		X		X	
Region*Year FE		X		X		X
Observations	4,676	4,676	2,306	2,306	2,370	2,370
R-squared	0.9247	0.9463	0.9498	0.9655	0.9313	0.9537

Notes: *Turnout* is the ratio between the total number of votes and the number of eligible voters of a constituency. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the last week before the election date. *Adj.margin* is the absolute difference between Conservative and Labour constituency-level vote shares in the previous general election, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Incumbent parties are defined at the constituency level. Constituency-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table B.3: Opinion polls, safeness of a constituency and their joint effect on HHI

	Dep. var.: HHI							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pollmargin_{w1} * Adj. margin_{t-1}	-0.2651***	-0.1602***						
	(0.0514)	(0.0503)						
Pollmargin_{w2} * Adj. margin_{t-1}			-0.2491***	-0.1517***				
			(0.0522)	(0.0504)				
Pollmargin_{w3} * Adj. margin_{t-1}					-0.2494***	-0.1536***		
					(0.0476)	(0.0449)		
Pollmargin_{w4} * Adj. margin_{t-1}							-0.2495***	-0.1519***
							(0.0410)	(0.0388)
Adj. margin_{t-1}	0.0606***	0.0474***	0.0599***	0.0470***	0.0615***	0.0482***	0.0636***	0.0495***
	(0.0065)	(0.0067)	(0.0066)	(0.0067)	(0.0064)	(0.0064)	(0.0060)	(0.0060)
Constituency FE	X	X	X	X	X	X	X	X
Year FE	X		X		X		X	
Region*Year FE		X		X		X		X
Observations	4,676	4,676	4,676	4,676	4,676	4,676	4,676	4,676
R-squared	0.6747	0.7801	0.6744	0.7800	0.6752	0.7803	0.6762	0.7806

Notes: Notes: HHI is the sum of squares of constituency-level vote shares for all parties. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in a certain time frame, subscripts indicate a specific week before the election date (1=last, ..., 4=fourth to last). *Adj.margin* is the absolute difference between Conservative and Labour constituency-level vote shares in the previous general election, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table B.4: Opinion polls, safeness of a constituency and their joint effect on party shares and pr. of winning

	Dep. var.:							
	Vote Share				Pr. of Winning			
	Incumbent		Follower		Incumbent		Follower	
	$I = P$	$I \neq P$	$I = P$	$I \neq P$	$I = P$	$I \neq P$	$I = P$	$I \neq P$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pollmargin$_{w1}$ * Adj. margin$_{t-1}$	-0.0634	-0.0984	0.5588***	-1.3001***	-1.6799***	7.5263***	3.2762***	-12.0562***
	(0.0821)	(0.0623)	(0.0719)	(0.1808)	(0.5166)	(0.6207)	(0.6366)	(1.1657)
Adj. margin$_{t-1}$	0.2458***	0.2418***	-0.4210***	-0.0623*	0.8579***	0.2531*	-1.3200***	0.2641
	(0.0142)	(0.0188)	(0.0174)	(0.0320)	(0.1072)	(0.1381)	(0.1427)	(0.1760)
Party FE	X	X	X	X	X	X	X	X
Constituency FE	X	X	X	X	X	X	X	X
Region*Year FE	X	X	X	X	X	X	X	X
Observations	2,026	1,706	1,337	1,252	2,026	1,706	1,337	1,252
R-squared	0.9046	0.9217	0.9389	0.8028	0.5018	0.6745	0.5168	0.6270

Notes: Dependent variables are: constituency-level party vote shares, and an indicator for whether the party won the constituency race. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the last week before the election date. *Adj.margin* is the absolute difference between Conservative and Labour constituency-level vote shares in the previous general election, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Incumbent and follower parties are defined at the constituency level. Odd columns refer to constituencies where the incumbent party is polls leading party, even columns the opposite. Constituency-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Results displayed in Table B.4 come from estimates of this model:

$$y_{p,c,t} = \beta Pollmargin_{wi,t} * Adj.margin_{c,t-1} + \delta Adj.margin_{c,t-1} + \gamma' X_{p,c,t} + \epsilon_{p,c,t}$$

.

Table B.5: Opinion polls margin interacted with previous election margin and political support (2 weeks window)

	Panel A - Dep. var.: Do not support any party **(pre election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w2} * Adj. margin_{t-1}	0.780**	0.789**	0.416**	0.455	0.425	0.453
	(0.346)	(0.377)	(0.197)	(0.392)	(0.358)	(0.393)
	[0.024]	[0.037]	[0.036]	[0.247]	[0.236]	[0.250]
Pollmargin_{w2}		-0.018		-0.019		-0.047
		(0.290)		(0.167)		(0.301)
		[0.951]		[0.911]		[0.876]
Adj. margin_{t-1}	-0.035	-0.035	0.029	0.027	0.029	0.028
	(0.023)	(0.024)	(0.036)	(0.039)	(0.039)	(0.039)
	[0.125]	[0.142]	[0.425]	[0.483]	[0.462]	[0.478]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	26,353	26,353	26,353	26,353	26,353	26,353
R-squared	0.000	0.000	0.045	0.045	0.045	0.045

	Panel B - Dep. var.: Do not support any party **(post election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w2} * Adj. margin_{t-1}	-0.234	-0.321	-0.310	-0.025	-0.141	-0.041
	(0.227)	(0.368)	(0.208)	(0.381)	(0.233)	(0.380)
	[0.303]	[0.384]	[0.136]	[0.949]	[0.546]	[0.914]
Pollmargin_{w2}		0.049		-0.142		-0.055
		(0.160)		(0.153)		(0.164)
		[0.761]		[0.353]		[0.738]
Adj. margin_{t-1}	0.016	0.021	-0.004	-0.021	-0.005	-0.011
	(0.020)	(0.026)	(0.035)	(0.040)	(0.035)	(0.040)
	[0.422]	[0.421]	[0.912]	[0.596]	[0.897]	[0.781]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	48,943	48,943	48,943	48,943	48,943	48,943
R-squared	0.000	0.000	0.025	0.026	0.026	0.026

Notes: The dependent variable is an indicator taking value one if the respondent does not support any party. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the second to last week before the respondent interview date. *Adj.margin* is the absolute difference between Conservative and Labour vote shares in the previous general election in the respontent's constituency, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, p-values in brackets, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table B.6: Opinion polls margin interacted with previous election margin and political engagement (2 weeks window)

	Panel A - Dep. var.: Do not support, feel close or vote a party tomorrow **(pre election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w2} * Adj. margin_{t-1}	0.870***	0.501	2.040***	0.324	0.590**	0.301
	(0.288)	(0.313)	(0.157)	(0.311)	(0.277)	(0.310)
	[0.003]	[0.110]	[0.000]	[0.298]	[0.033]	[0.332]
Pollmargin_{w2}		0.672**		0.825***		0.482*
		(0.272)		(0.136)		(0.280)
		[0.014]		[0.000]		[0.086]
Adj. margin_{t-1}	-0.062***	-0.045***	-0.083***	-0.015	-0.018	-0.010
	(0.016)	(0.017)	(0.029)	(0.030)	(0.029)	(0.029)
	[0.000]	[0.007]	[0.005]	[0.620]	[0.546]	[0.737]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	26,353	26,353	26,353	26,353	26,353	26,353
R-squared	0.009	0.009	0.059	0.060	0.060	0.061

	Panel B - Dep. var.: Do not support, feel close or vote a party tomorrow **(post election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w2} * Adj. margin_{t-1}	-0.559***	-0.283	-1.553***	0.050	-0.571***	-0.046
	(0.159)	(0.262)	(0.154)	(0.271)	(0.158)	(0.263)
	[0.000]	[0.281]	[0.000]	[0.854]	[0.000]	[0.861]
Pollmargin_{w2}		-0.155		-0.799***		-0.289**
		(0.120)		(0.116)		(0.118)
		[0.198]		[0.000]		[0.015]
Adj. margin_{t-1}	0.026*	0.009	0.052*	-0.044	0.048**	0.014
	(0.015)	(0.022)	(0.031)	(0.031)	(0.024)	(0.028)
	[0.099]	[0.672]	[0.091]	[0.161]	[0.049]	[0.627]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	48,943	48,943	48,943	48,943	48,943	48,943
R-squared	0.011	0.011	0.036	0.037	0.042	0.042

Notes: The dependent variable is an indicator taking value one if the respondent does not support, feel close nor would vote for any party. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the second to last week before the respondent interview date. *Adj.margin* is the absolute difference between Conservative and Labour vote shares in the previous general election in the respontent's constituency, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, p-values in brackets, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table B.7: Opinion polls margin interacted with previous election margin and political support (3 weeks window)

	Panel A - Dep. var.: Do not support any party **(pre election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w3} * Adj. margin_{t-1}	0.885**	0.796**	0.454**	0.467	0.551	0.475
	(0.350)	(0.377)	(0.199)	(0.393)	(0.366)	(0.394)
	[0.012]	[0.035]	[0.023]	[0.235]	[0.132]	[0.228]
Pollmargin_{w3}		0.170		-0.006		0.136
		(0.295)		(0.165)		(0.301)
		[0.564]		[0.970]		[0.652]
Adj. margin_{t-1}	-0.040*	-0.036	0.029	0.028	0.024	0.027
	(0.023)	(0.024)	(0.036)	(0.039)	(0.039)	(0.039)
	[0.083]	[0.141]	[0.428]	[0.470]	[0.532]	[0.501]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	26,353	26,353	26,353	26,353	26,353	26,353
R-squared	0.000	0.000	0.045	0.045	0.046	0.046

	Panel B - Dep. var.: Do not support any party **(post election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w3} * Adj. margin_{t-1}	-0.173	-0.395	-0.252	-0.108	-0.063	-0.124
	(0.229)	(0.385)	(0.210)	(0.396)	(0.230)	(0.395)
	[0.450]	[0.306]	[0.232]	[0.785]	[0.784]	[0.753]
Pollmargin_{w3}		0.122		-0.071		0.033
		(0.164)		(0.161)		(0.170)
		[0.458]		[0.658]		[0.845]
Adj. margin_{t-1}	0.012	0.025	-0.009	-0.018	-0.010	-0.006
	(0.020)	(0.027)	(0.035)	(0.040)	(0.035)	(0.041)
	[0.540]	[0.350]	[0.790]	[0.656]	[0.782]	[0.887]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	49,012	49,012	49,012	49,012	49,012	49,012
R-squared	0.000	0.000	0.025	0.025	0.026	0.026

Notes: The dependent variable is an indicator taking value one if the respondent does not support any party. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the third to last week before the respondent interview date. *Adj.margin* is the absolute difference between Conservative and Labour vote shares in the previous general election in the respontent's constituency, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, p-values in brackets, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table B.8: Opinion polls margin interacted with previous election margin and political engagement (3 weeks window)

	Panel A - Dep. var.: Do not support, feel close or vote a party tomorrow **(pre election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w3} * Adj. margin_{t-1}	0.929***	0.516	2.061***	0.325	0.664**	0.313
	(0.292)	(0.314)	(0.159)	(0.312)	(0.277)	(0.311)
	[0.002]	[0.100]	[0.000]	[0.298]	[0.017]	[0.315]
Pollmargin_{w3}		0.792***		0.836***		0.623**
		(0.277)		(0.136)		(0.284)
		[0.004]		[0.000]		[0.029]
Adj. margin_{t-1}	-0.065***	-0.046***	-0.084***	-0.013	-0.020	-0.011
	(0.016)	(0.017)	(0.029)	(0.030)	(0.029)	(0.029)
	[0.000]	[0.006]	[0.004]	[0.649]	[0.487]	[0.716]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	26,353	26,353	26,353	26,353	26,353	26,353
R-squared	0.009	0.009	0.059	0.061	0.060	0.061

	Panel B - Dep. var.: Do not support, feel close or vote a party tomorrow **(post election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w3} * Adj. margin_{t-1}	-0.485***	-0.160	-1.517***	0.169	-0.494***	0.089
	(0.171)	(0.282)	(0.166)	(0.292)	(0.170)	(0.280)
	[0.005]	[0.571]	[0.000]	[0.562]	[0.004]	[0.750]
Pollmargin_{w3}		-0.179		-0.836***		-0.316**
		(0.127)		(0.123)		(0.124)
		[0.159]		[0.000]		[0.011]
Adj. margin_{t-1}	0.021	0.002	0.044	-0.056*	0.042*	0.005
	(0.016)	(0.022)	(0.031)	(0.032)	(0.024)	(0.028)
	[0.196]	[0.943]	[0.156]	[0.085]	[0.085]	[0.869]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	49,012	49,012	49,012	49,012	49,012	49,012
R-squared	0.011	0.011	0.035	0.037	0.042	0.042

Notes: The dependent variable is an indicator taking value one if the respondent does not support, feel close nor would vote for any party. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the third to last week before the respondent interview date. *Adj.margin* is the absolute difference between Conservative and Labour vote shares in the previous general election in the respontent's constituency, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, p-values in brackets, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table B.9: Opinion polls margin interacted with previous election margin and political support (4 weeks window)

	Panel A - Dep. var.: Do not support any party **(pre election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w4} * Adj. margin_{t-1}	0.917***	0.810**	0.463**	0.489	0.591	0.499
	(0.351)	(0.378)	(0.198)	(0.393)	(0.367)	(0.394)
	[0.009]	[0.032]	[0.020]	[0.214]	[0.108]	[0.206]
Pollmargin_{w4}		0.210		-0.012		0.170
		(0.309)		(0.166)		(0.317)
		[0.497]		[0.941]		[0.592]
Adj. margin_{t-1}	-0.042*	-0.037	0.029	0.028	0.023	0.026
	(0.023)	(0.024)	(0.036)	(0.039)	(0.039)	(0.039)
	[0.069]	[0.131]	[0.430]	[0.482]	[0.558]	[0.518]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	26,353	26,353	26,353	26,353	26,353	26,353
R-squared	0.000	0.000	0.046	0.046	0.046	0.046

	Panel B - Dep. var.: Do not support any party **(post election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w4} * Adj. margin_{t-1}	-0.070	-0.399	-0.167	-0.187	0.036	-0.204
	(0.235)	(0.397)	(0.212)	(0.404)	(0.232)	(0.402)
	0.766	0.315	0.430	0.644	0.877	0.612
Pollmargin_{w4}		0.179		0.010		0.128
		(0.165)		(0.165)		(0.172)
		0.279		0.953		0.455
Adj. margin_{t-1}	0.006	0.025	-0.016	-0.015	-0.016	-0.001
	(0.020)	(0.027)	(0.035)	(0.040)	(0.035)	(0.041)
	0.761	0.355	0.644	0.706	0.646	0.978
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	49,012	49,012	49,012	49,012	49,012	49,012
R-squared	0.000	0.000	0.025	0.025	0.026	0.026

Notes: The dependent variable is an indicator taking value one if the respondent does not support any party. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the fourth to last week before the respondent interview date. *Adj.margin* is the absolute difference between Conservative and Labour vote shares in the previous general election in the respontent's constituency, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, p-values in brackets, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table B.10: Opinion polls margin interacted with previous election margin and political engagement (4 weeks window)

	Panel A - Dep. var.: Do not support, feel close or vote a party tomorrow **(pre election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w4} * Adj. margin_{t-1}	0.936***	0.513	2.046***	0.315	0.656**	0.304
	(0.294)	(0.313)	(0.157)	(0.309)	(0.276)	(0.308)
	[0.002]	[0.101]	[0.000]	[0.309]	[0.018]	[0.324]
Pollmargin_{w4}		0.829***		0.834***		0.642**
		(0.285)		(0.135)		(0.291)
		[0.004]		[0.000]		[0.028]
Adj. margin_{t-1}	-0.066***	-0.046***	-0.084***	-0.013	-0.020	-0.010
	(0.016)	(0.017)	(0.029)	(0.030)	(0.029)	(0.029)
	[0.000]	[0.006]	[0.004]	[0.671]	[0.485]	[0.727]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	26,353	26,353	26,353	26,353	26,353	26,353
R-squared	0.009	0.010	0.059	0.061	0.060	0.061

	Panel B - Dep. var.: Do not support, feel close or vote a party tomorrow **(post election)**					
	(1)	(2)	(3)	(4)	(5)	(6)
Pollmargin_{w4} * Adj. margin_{t-1}	-0.406**	0.035	-1.458***	0.360	-0.429**	0.287
	(0.171)	(0.282)	(0.169)	(0.296)	(0.170)	(0.280)
	[0.018]	[0.902]	[0.000]	[0.223]	[0.012]	[0.306]
Pollmargin_{w4}		-0.240*		-0.901***		-0.383***
		(0.130)		(0.126)		(0.127)
		[0.064]		[0.000]		[0.003]
Adj. margin_{t-1}	0.016	-0.010	0.038	-0.069**	0.038	-0.008
	(0.016)	(0.022)	(0.031)	(0.033)	(0.024)	(0.028)
	[0.324]	[0.652]	[0.229]	[0.036]	[0.119]	[0.787]
Year FE	X	X			X	X
Constituency FE			X	X	X	X
Observations	49,012	49,012	49,012	49,012	49,012	49,012
R-squared	0.011	0.011	0.035	0.037	0.042	0.042

Notes: The dependent variable is an indicator taking value one if the respondent does not support, feel close nor would vote for any party. *Pollmargin* is the absolute difference between Conservative and Labour vote shares averaged across all national pollsters in the fourth to last week before the respondent interview date. *Adj.margin* is the absolute difference between Conservative and Labour vote shares in the previous general election in the respontent's constituency, adjusted by the sum of those vote shares. All margins are $\in (0, 1)$. Constituency-level clustered standard errors are presented in parentheses, p-values in brackets, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table B.11: Pollster differences in reported opinion poll shares and margin

	Dep. var.:		
	share Conservative	share Labour	Pollmargin
Angus Reid	0.0024	-0.0479***	0.0142
	(0.0050)	(0.0074)	(0.0114)
Ashcroft	0.0011	-0.0318***	-0.0175
	(0.0058)	(0.0068)	(0.0129)
Audience Selection	-0.0177***	-0.0178***	-0.0085**
	(0.0032)	(0.0021)	(0.0038)
BMG	0.0145**	-0.0216***	-0.0056
	(0.0062)	(0.0058)	(0.0143)
BPIX	0.0112**	-0.0184**	-0.0382***
	(0.0046)	(0.0069)	(0.0104)
ComRes	0.0113**	-0.0082	-0.0148
	(0.0047)	(0.0064)	(0.0120)
Gallup	0.0062*	0.0016	-0.0071
	(0.0035)	(0.0040)	(0.0098)
Harris	0.0007	-0.0066**	-0.0074
	(0.0022)	(0.0024)	(0.0058)
ICM	0.0089**	-0.0160***	-0.0191**
	(0.0038)	(0.0022)	(0.0074)
Kantar	-0.0122	-0.0166**	-0.0283
	(0.0077)	(0.0072)	(0.0185)
Marplan	0.0029	-0.0046	-0.0008
	(0.0031)	(0.0050)	(0.0064)
NMR	-0.0071	-0.0211***	-0.0222***
	(0.0065)	(0.0036)	(0.0064)
NOP	-0.0005	-0.0026***	0.0018
	(0.0031)	(0.0008)	(0.0052)
Neilsen	0.0244***	-0.0066*	-0.0318***
	(0.0056)	(0.0032)	(0.0077)
ORB	0.0028	-0.0016	-0.0283
	(0.0077)	(0.0072)	(0.0185)
Opinium	0.0039	-0.0100	-0.0255*
	(0.0053)	(0.0061)	(0.0131)
Panelbase	-0.0071	-0.0034	-0.0142
	(0.0061)	(0.0058)	(0.0148)
Populus	-0.0028	-0.0022	-0.0248**
	(0.0048)	(0.0064)	(0.0119)
Rasmussen	0.0264***	-0.0441***	-0.0686***
	(0.0058)	(0.0060)	(0.0143)
Survation	-0.0130*	-0.0069	-0.0349**
	(0.0063)	(0.0060)	(0.0149)
TNS BMRB	-0.0019	-0.0071	-0.0334***
	(0.0048)	(0.0072)	(0.0118)
YouGov	0.0041	-0.0043	-0.0317**
	(0.0046)	(0.0061)	(0.0120)
Observations	474	474	474
R-squared	0.9322	0.9503	0.8727

Notes: Polls margins are in absolute terms. All dependent variables are $\in (0, 1)$. Covariates represent pollsters' fixed effects. The excluded pollster house is MORI (Ipsos-MORI after 2005 GE) as it covers all general elections considered. All regressions include week-by-year fixed effects. Pollster-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1

Chapter 3

WHO VOTED FOR BREXIT? INDIVIDUAL AND REGIONAL DATA COMBINED

with Sascha O. Becker[1], Thiemo Fetzer[2] and Dennis Novy[3]

[1]Monash University, CAGE, CEPR, CESifo, ifo, IZA and ROA.

[2]University of Warwick, CAGE, SERC.

[3]University of Warwick, CEPR, CESifo and CEP/LSE.

We are grateful to the editor and two referees for constructive comments. We thank the *Understanding Society* team based at the Institute for Social and Economic Research at the University of Essex for access and guidance on the data. Support by the ESRC Centre for Competitive Advantage in the Global Economy (ESRC grant ES/L011719/1) is gratefully acknowledged. *Email addresses:* e.alabrese@warwick.ac.uk, s.o.becker@warwick.ac.uk, t.fetzer@warwick.ac.uk, d.novy@warwick.ac.uk.

Abstract

Previous analyses of the 2016 Brexit referendum used region-level data or small samples based on polling data. The former might be subject to ecological fallacy and the latter might suffer from small-sample bias. We use individual-level data on thousands of respondents in *Understanding Society*, the UK's largest household survey, which includes the EU referendum question. We find that voting Leave is associated with older age, white ethnicity, low educational attainment, infrequent use of smartphones and the internet, receiving benefits, adverse health and low life satisfaction. These results coincide with corresponding patterns at the aggregate level of voting areas. We therefore do not find evidence of ecological fallacy. In addition, we show that prediction accuracy is geographically heterogeneous across UK regions, with strongly pro-Leave and strongly pro-Remain areas easier to predict. We also show that among individuals with similar socio-economic characteristics, Labour supporters are more likely to support Remain while Conservative supporters are more likely to support Leave.

Keywords: Aggregation, Ecological Fallacy, European Union, Populism, Referendum, UK

JEL Classification: D72, I10, N44, R20, Z13

3.1 Introduction

Populism has been on the rise across Europe and the United States in recent years, culminating in the election of Donald Trump as US President and the Brexit vote in the 2016 EU referendum. The Brexit vote came as a shock to many observers and triggered early attempts to understand the voting patterns.[4] These studies relied almost exclusively on aggregate data at the level of voting areas. Regressing vote shares across voting areas on average population characteristics risks falling into the ecological fallacy trap of inferring *individual* associations from *aggregate* data (see Robinson, 1950).

We use detailed individual-level data from the *Understanding Society* survey containing the EU referendum question to address three interrelated questions. First, we investigate the relationship between voters' personal characteristics and their expressed voting intentions. Particularly, we address whether ecological fallacy may be driving the associations documented in the aggregated data. Second, building a predictive model of Leave support we assess which voting determinants have the most power to predict voting behavior out of sample. Third, we investigate the classification errors that this predictive model makes by region and voters' closeness to political parties.

We find that individual and aggregate coefficients point in a similar direction, suggesting that ecological fallacy is of limited concern. Second, we document that the predictive models exhibit a significant gain in accuracy when exploiting both individual and regional variables. Lastly, we document that a predictive model performs best in parts of the UK with the most extreme referendum outcomes: Lincolnshire (highest Leave share) and London (lowest Leave share across mainland Britain). Furthermore, a decomposition of classification errors reveals that closeness to a political party is likely an important omitted variable, suggesting that unobservable traits and identity are further key correlates.

The paper is structured as follows. Section 3.2 lays out the literature background, describes the data and explains our empirical approach. We present graphical summaries of our results in Section 3.3, and we conclude in Section 3.4. Underlying regression results and further details

[4] See Burn-Murdoch (2016) in the Financial Times as an example of various correlation plots; more in-depth work followed, for example Clarke and Whittaker (2016); Darvas (2016); Langella and Manning (2016).

are relegated to an appendix.

3.2 Background, data and empirical approach

3.2.1 Background

This paper builds on Becker et al. (2017) who analyze the Brexit vote shares across UK voting areas, using a wide range of explanatory variables. They show that the Leave vote shares are systematically correlated with older age, lower educational attainment, unemployment, or employment in certain industries such as manufacturing, as well as with a lack of quality of public service provision.

These results fit in with other evidence on the Brexit vote. An early attempt to explain the referendum outcome was made by Ashcroft (2016) whose polling data indicated that the typical Leave voter is white, middle class and lives in the South of England. Sampson (2017) reviews the literature on the likely economic consequences of Brexit on the British economy and other countries.

Our paper also relates to the wider literature on political polarization as well on voting for far-right parties. Ferree et al. (2014) provide an extensive review of academic works which link voting patterns to demographic, economic and political features. Voters' behaviour has also been shown to be strongly associated with individual scepticism towards institutions (e.g. Euroscepticism) or intolerance against foreigners (see Whitaker and Lynch, 2011; Clarke and Whittaker, 2016; Arzheimer, 2009). Additional studies claim that ethnic minorities may engage in 'ethnic' or 'policy' voting depending on the issue they are called to vote upon (see Bratton and Kimenyi, 2008; Tolbert and Hero, 1996).

Polarization has also been related to immigration (see Barone et al., 2016) as well as trade integration (Dippel et al., 2015; Burgoon, 2013; Dorn et al., 2016). In the UK context, Becker et al. (2016) examine immigration from Eastern Europe as a potential driver of support for the UK Independence Party, while Fetzer (2018) explores the role of austerity policies since 2010.

Overall, the voting patterns in the Brexit referendum are complex. One possible – albeit not the only – interpretation of the empirical literature on Brexit so far is that some people who favor

Leave may feel 'left behind', be it economically or culturally (see Hobolt, 2016; Clarke et al., 2017). This is consistent with sociological studies which demonstrate similar patterns for the Tea Party Movement and the 2016 US presidential election, e.g. Hochschild (2018).

3.2.2 Data

Are these aggregate patterns found by Becker et al. (2017) and others a fair reflection of individual-level relationships? The individual-level data from wave 8 of the *Understanding Society* survey makes it possible to investigate this question. Our focus is on individual socio-economic variables for which region-level equivalents are used in Becker et al. (2017). Our approach of combining individual-level and aggregate data allows us (a) to check whether ecological fallacy is an important factor in aggregate analyses of the Brexit vote, and (b) to exploit the combined predictive power of individual-level and aggregate variables. This opens up insights into (c) geographic heterogeneity in predictive power across UK regions.

The *Understanding Society* data cover a wide range of topics, in particular basic demographic data for all household members such as sex, age and ethnicity, place of birth, family background including marital status, educational attainment, current job characteristics, housing characteristics (owning vs. renting), health status and life satisfaction. We describe the sampling design in more detail in an appendix, and how we construct our sample (also see Knies, 2016).

3.2.3 Descriptive statistics

According to the summary statistics in Table 3.1, 42.2% of the 13,136 individuals in our sample indicate that the UK should leave the EU in response to the survey question "Should the UK remain a member of the EU or leave the EU?" This compares to 51.9% of the electorate voting Leave in the referendum. We refer to Becker et al. (2017, section 3.1) for a discussion of the aggregate voting and turnout patterns in the 2016 referendum.

As for demographics, the proportion of males is 45.4% of all individuals in the sample, while just about three out of ten respondents are aged 60 or above. People with no qualification account for about 8% of the sample. Roughly 90% of respondents are born in the UK. Asians are the largest ethnic minority amounting to 5.8% of the sample, followed by blacks (2.5%).[5] Over half

[5] Note that we sourced nationality and ethnicity variables also from earlier waves.

of respondents are married or in a civil partnership. In terms of current employment, roughly four out of ten people declare to be without a paid job or to not have worked in the seven days prior to being questioned.[6]

3.2.4 Understanding Society: Research in progress

We gained access to *Understanding Society* data in the summer of 2017, at the same time as other groups of researchers in a pilot 'early access' project. We briefly summarize related preliminary findings reported by other researchers in short presentations in the summer of 2017. For instance, Creighton and Amaney (2017) find that opposition to immigration played a key role. Martin and Sobolewska (2017) explore racial determinants and find that ethnic minorities are strongly in favor of remaining in the EU. De Vries and Solaz (2017) attempt to explain voters' behavior by analyzing socio-economic determinants such as asset holdings, sources of income and skills, whereas Doebler and Hayes (2017) explore additional potential drivers such as personal economic struggle and regional economic decline.

As far as we are aware, only one other paper using *Understanding Society* data has come out as a working paper so far. Liberini et al. (2017) show that individuals dissatisfied with their own financial situation were more likely to vote Leave and that the very young were most likely to vote Remain. In related work, Pollock (2017) uses the Innovation Panel to argue that the rise in populism and the vote in favor of Brexit can be attributed to generational shifts away from mainstream political parties over the past three decades.

3.2.5 Empirical approach

We start with a simple model where the dependent variable y_{ic} is a dummy for individual i in local authority c which takes on the value 1 if the interviewed person answers "Leave" in response to the question "Should the UK remain a member of the EU or leave the EU?" and 0 if the answer is "Remain":

$$y_{ic} = \mathbf{x}'_{ic}\beta + \mathbf{z}'_{c}\delta + \varepsilon_{ic}. \tag{1}$$

[6] The aggregate variables in Table 3.1 are not standardized for descriptive purposes, but they are in all regressions.

The independent variables in the model are the *Understanding Society* cross-sectional individual covariates $\mathbf{x}_{ic}$ on the one hand, and area-specific aggregate variables $\mathbf{z}_c$ from Becker et al. (2017) on the other. Our overall sample contains 13,136 respondents for our baseline regressions. We also analyze smaller samples and subgroups of variables since not all *Understanding Society* respondents were asked each survey module. As the summary statistics in Table 3.1 show, roughly 42% of respondents are in favor of Leave.

We relegate the details of the underlying regression results to the appendix. For ease of interpretation, throughout the regression tables in the appendix we provide coefficients obtained from a simple linear probability model estimation of equation (1). However, each model is also estimated using the corresponding logistic regression model to provide an estimate of the success rate at the bottom of each table.

Since our interest centers on prediction, we need a metric to assess predictive accuracy of our regression models. We perform a simple validation exercise known from the machine learning literature. Our sample is divided into a random *training set* (2/3 of the sample) and a *validation set*. Logistic regressions are conducted on the training set, and we use the validation set to perform classification. We follow Bayes' optimal decision rule and classify an observation as "Leave" if the predicted posterior probability exceeds 50%. In essence, this rule simple allocates the label ("Leave" or "Remain") to an observation that, conditional on our predictors/features, is most likely. This decision rule minimizes the error rate or maximizes overall accuracy. Yet, it does so putting an equal penality or cost on false-positives versus false-negatives. The comparison of the predicted to the actual assignments allows us to estimate the out-of-sample predictive power and to shed light on the two types of prediction errors (false positives versus false negatives). For instance, individual A in the validation set may, based on her characteristics, look like a typical Remain voter but is in reality a Leave voter, so we have a case of a false negative. Individual B in the validation set may, based on her characteristics, look like a typical Leave voter but is in reality a Remain voter, so we have a case of a false positive.

We stress that causality is beyond the scope of our paper. Instead, our results reflect a broad range of correlation patterns relating voting intentions to fundamental socio-economic features.[7]

[7] In a fascinating paper, Colantone and Stanig (2018) focus on one specific causal factor behind the Leave vote: rising import competition from China. While papers studying causality are extremely important, they give

In our earlier work (Becker et al. (2017), we grouped variables by four topics: (1) EU exposure: immigration, trade and EU transfers; (2) Public service provision and fiscal consolidation; (3) Demography, education and life satisfaction; (4) Economic structure, wages and unemployment. Those groupings follow from prominent hypotheses that have been proposed to explain the EU referendum result. That is, the first grouping looks at the relationship between EU exposure and Leave voting. Here, we follow the same logic and look at groups of variables that correspond to one specific set of explanations for the referendum result. For each variable grouping, we assess its predictive power by itself, and compare this to the joint predictive power of all groups of variables combined. We discuss the different groupings in more detail in the appendix (the regression tables using the groups of variables under discussion are described in sections c.4-c.10 in the appendix).[8]

As Becker et al. (2017) explain, the fundamental difference between prediction, as pursued in this paper, and causal inference is as follows. Causal inference focuses on the internal validity of causally estimated reduced-form (or structural) parameters β. In contrast, prediction is concerned with the external validity of the estimated fitted values $\hat{y}$.[9] Causal inference seeks to obtain a set of estimated parameters $\hat{\beta}$ that are usually studied in isolation. Thus, they often do not render themselves useful for prediction because the out-of-sample model fit is generally poor. Instead, good model fit typically requires a multitude of regressors, and machine learning can often substantially improve out-of-sample predictive performance (Mullainathan and Spiess, 2017). The underlying estimated parameters that yield good model fit are typically of limited interest per se. For this reason, we only show coefficient estimates in appendix tables, while in the main text we focus on graphical representation.

prominence to one factor at a time, an aim different from ours which is to look at the relative predictive power of different variables.

[8]One might wonder whether including region fixed effects above and beyond the individual-level and region-level predictors is beneficial in terms of predition accuracy, but the benefits are very marginal in our case. Since region fixed effects are a 'black box', we refrain from including them given the very limited gains.

[9]While we do not use machine-learning methods in this paper such as best subset selection (BSS) or LASSO, we did so in Becker et al. (2017), i.e. our selection of variables is guided by the (aggregate) variables employed in that earlier paper.

3.3 Predicting the vote

In order to focus on prediction quality, we relegate the discussion of individual regression tables to the appendix. First, we focus on the relative predictive power of individual-level and aggregate variables. Second, we examine the predictive power of our best-performing model across regions and lastly, we investigate the classification error structure.

3.3.1 Individual vs. aggregate variables

Figure 3.1 reports the proportion of correct predictions (success rates) for each variable grouping estimated in the hold-out sample. In particular, Figure 3.1(a) illustrates success rates for (groupings of) aggregate variables and Figure 3.1(b) for individual-level variables. Figure 3.1(c) combines aggregate and individual-level variables. Figure 3.1(d) reports success rates for non-comparable individual variables.

The overall classification success rate when relying on aggregate data in Figure 3.1(a) is 58.8%. In the narrow individual-level sample for which employment and related individual data is collected in the *Understanding Society* sample, the overall accuracy reaches 62.9% using the aggregate level area employment characteristics. The improvement in terms of accuracy relative to a naive classification rule that classifies everyone as Remain (generating a success rate of 57.8%, i.e. one minus the sample 'Leave' share) thus is only modest. When focusing on all comparable individual-level covariates in Figure 3.1(b), we see that individual-level variables have stronger predictive power than aggregate ones. The improvement in accuracy up to 63.4% with all variables included suggests an improvement in prediction accuracy relative to the naive benchmark of 9.7%.

Furthermore, an inspection of the tables in the appendix confirms that the individual-level predictors yield broadly similar sign patterns to their aggregate level equivalents. This suggests that ecological fallacy is not a major concern for the results in Becker et al. (2017).

The combination of individual and aggregate characteristics yields a further slight improvement in prediction accuracy. Relative to the naive classification rule, accuracy can improve up to 64.6% with all covariates included, representing an improvement of 11.7% in relative terms.

Adding further individual-level characteristics that are included in the *Understanding Society* sample (but for which no aggregate proxy measures exist) suggests that overall accuracy is not further improved.

In fact, our best model including all characteristics sees a small drop in the success rate. In terms of the bias-variance trade-off inherent in such predictive models, the improvement in terms of bias are therefore likely offset by an inflation in terms of variance, resulting in worse out-of-sample performance. We refer to Gareth et al. (2013) for a discussion of the bias-variance trade-off.

As explained in the appendix, we explore a number of novel individual determinants. We find that marital status, technology use and dependence on income support and state benefits are all systematically linked to individual voting behavior. In particular, individuals who do not possess smartphones and who use the internet infrequently appear more inclined to support Leave. Those repeatedly seeking health care or receiving income support also tend to be more in favor of Brexit. Similarly, it is also fair to say that Brexit is a predominantly white phenomenon compared to ethnic minorities.

3.3.2 Geographical heterogeneity

An instructive step lies in attempting to decompose in which regions our model does a good job in correctly classifying the voting intentions in the *Understanding Society* sample. Among all NUTS2 regions in Figure 3.2, Inner London displays the lowest error rate (21%) followed by Lincolnshire and North Eastern Scotland (with 23% and 26%, respectively). Lincolnshire and Inner London had among the highest and lowest Leave vote shares in the referendum. Thus, it is hardly surprising that the empirical model performs well in separating voters in these regions.

The model has the lowest performance in Tees Valley and Durham, East Anglia, and Merseyside (with error rates around 43-44%). Generally, the picture that emerges suggests that purely based on the socio-economic characteristics, areas that are more disadvantaged are the ones where it is most difficult to separate Leave from Remain voters. Non-economic factors may therefore be particularly helpful in capturing variation between voters in these areas.

3.3.3 Types of errors

We turn to decomposing errors into false positives and false negatives. The results presented in Figure 3.2 suggest that the regions of Inner and Outer London, Berkshire, Buckinghamshire and Oxford as well as North Eastern Scotland stand out as having the highest rate of false negatives (blue bars). False negatives are cases in which our model identifies an individual as a Remain voter, while in fact they state an intention to vote Leave. The false negatives in Figure 3.2 suggest that there are non-negligible proportions of voters who, based on their socio-economic characteristics, *look like Remain voters* but actually express an intention to vote Leave. In Outer London, 80% of all classification errors are false negatives. The same holds true for many of the other regions in London's wealthy commuter belt.

We next investigate whether classification errors can be related to individual political party preferences. From previous *Understanding Society* survey rounds which asked participants what party they felt closest to, we obtain that measure for 65% of our estimation sample. Figure 3.3 highlights that, while overall accuracy across the stated historical party preferences is similar, the type of classification error is quite heterogeneously distributed. In particular, Labour voters are more likely to contribute to the false positive errors – cases where our model classifies an individual as a Leave voter when in fact they favour Remain – making up 51.27% of the share of all false positives. By contrast, Conservative party supporters make up 44.8% of the share of false negatives – individuals who look like Remain voters but actually intend to vote Leave.

Overall, our findings indicate that Labour voters with observables that put them in the Leave camp – male, older, less educated, less likely to be in employment, etc. – are significantly more likely to express a preference for the status quo of remaining in the EU. Voters with similar socio-economic profiles who identify with the Conservative party are more likely to vote Leave. This suggests the potential importance of other characteristics not in the data set, for instance psychological traits such as openness as well as attitudes towards national identity.

3.4 Conclusion

Individual-level regressors from the British *Understanding Society* survey containing the 2016 EU referendum question give similar results to corresponding aggregate variables at the level of local authority areas analyzed by Becker et al. (2017). We therefore find no evidence of ecological fallacy effects – individuals appear to behave in similar ways as suggested by the aggregate data.

We also shed light on the predictive power of different determinants of the Leave vote. Demographics and employment characteristics are the most relevant covariates for prediction, while the cumulative power of individual-level and aggregate variables shows a non-negligible gain over aggregate data alone. Geographical heterogeneity is also important as our model performs best in more prosperous areas (London in particular).

Finally, we also find that individuals who support the Labour party but have otherwise observables that would put them in the Leave camp are significantly more likely to vote Remain. Vice versa, supporters of the Conservative party with Remain-favouring characteristics are more likely to vote Leave.

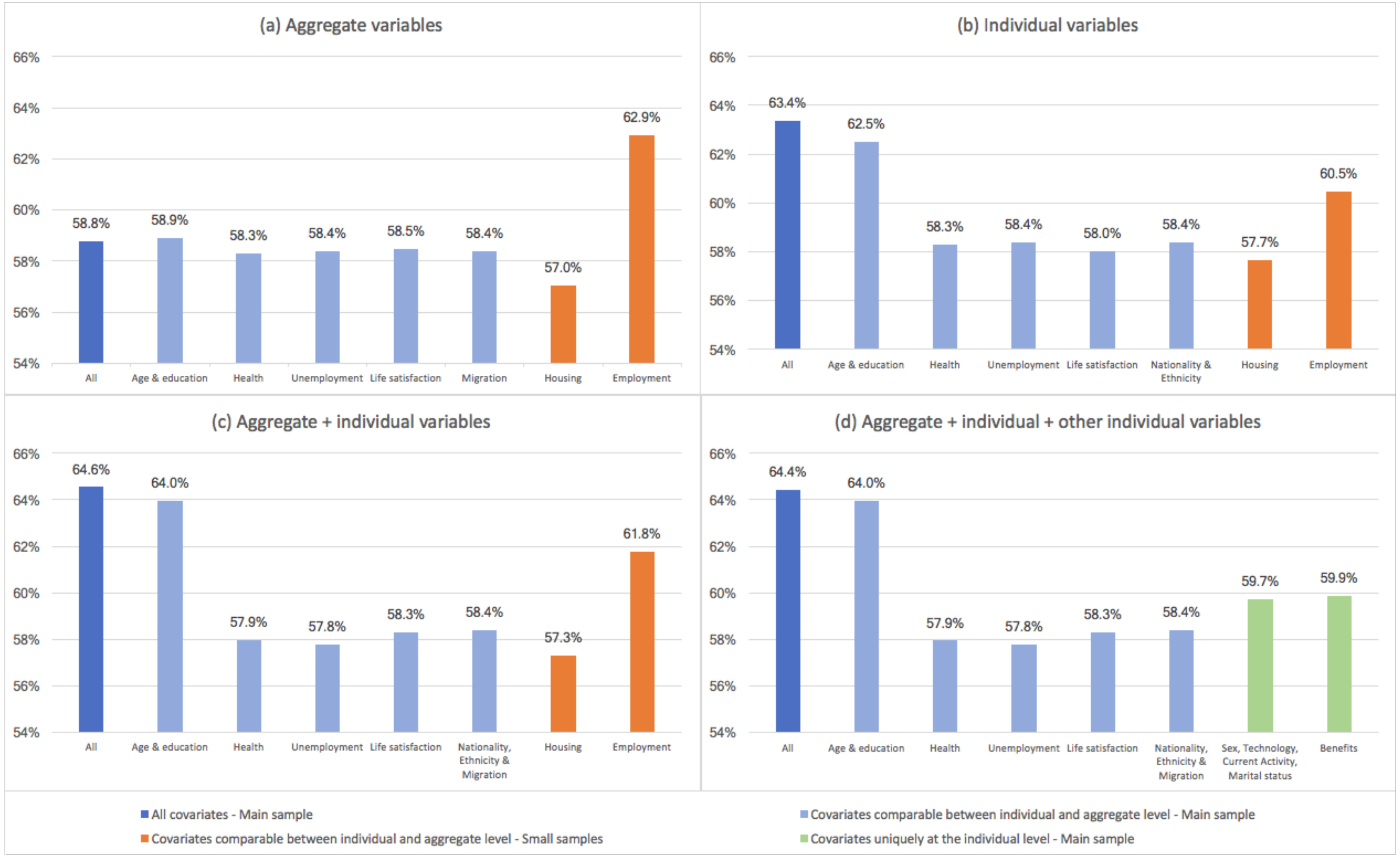
3.5 Table and figures

Table 3.1: Summary statistics

VARIABLES	N	mean	sd	min	max
Dependent variable:					
Should the UK leave the EU	13,136	0.422	0.494	0	1
Individual variables:					
Sex = Male	13,136	0.454	0.498	0	1
Age = 60 or older	13,136	0.305	0.460	0	1
Highest qualification = Other lower qualification	13,136	0.0847	0.278	0	1
Highest qualification = No qualification	13,136	0.0796	0.271	0	1
Frequency using internet = Every day	13,136	0.792	0.406	0	1
Frequency using internet = No access	13,136	0.0153	0.123	0	1
Born in UK	13,136	0.905	0.294	0	1
Ethnic group = White	13,136	0.896	0.306	0	1
Ethnic group = Asian	13,136	0.0579	0.234	0	1
Ethnic group = Black	13,136	0.0245	0.155	0	1
Current legal marital status = Single	13,136	0.287	0.452	0	1
Current legal marital status = Married or civil partner	13,136	0.546	0.498	0	1
Visits GP in 12m = None	13,136	0.214	0.410	0	1
Visits GP in 12m = Over 10	13,136	0.0613	0.240	0	1
Housing tenure = Owned (outright + mortgage)	9,344	0.664	0.472	0	1
No work last week & doesn't have paid job	13,136	0.379	0.485	0	1
Current job sector = Manufacturing	7,950	0.0826	0.275	0	1
Income support	13,136	0.0158	0.125	0	1
Dissatisfied with health	13,136	0.244	0.430	0	1
Dissatisfied with income	13,136	0.214	0.410	0	1
Aggregated variables:					
Unemployment rate (2015)	13,136	5.618	2.159	1.600	12.10
Share of suspected cancer patient treated within 62 Days (2015)	13,136	84.02	7.554	33.30	100
CV life satisfaction APS well-being data (2015)	13,136	0.989	0.394	0.570	3.050
Manufacturing employment share (2001)	13,136	0.153	0.0518	0.0538	0.337
Owned (outright + mortgage) share (2001)	13,136	0.684	0.0975	0.274	0.882

Notes: The table reports the number of observations (N), their mean, standard deviation (sd) as well as the minimum and maximum values. The summary statistics for the aggregate variables are reported based on the raw data, whereas in the regression tables these variables are used in standardized form.

Figure 3.1: Success rates by variable groupings



Notes: The graph plots success rates for different variable groupings. Light blue refers to models for our main sample with covariates comparable at the individual and aggregate levels. Orange relates to variables which are only available for smaller subsamples (individuals answering questions on housing or employment). Dark blue applies to models combining all available covariates in the main sample. Finally, green relates to individual-level variables which do not have a comparable grouping in the aggregate data.

Figure 3.2: Error rates and decomposition into false positives versus false negatives

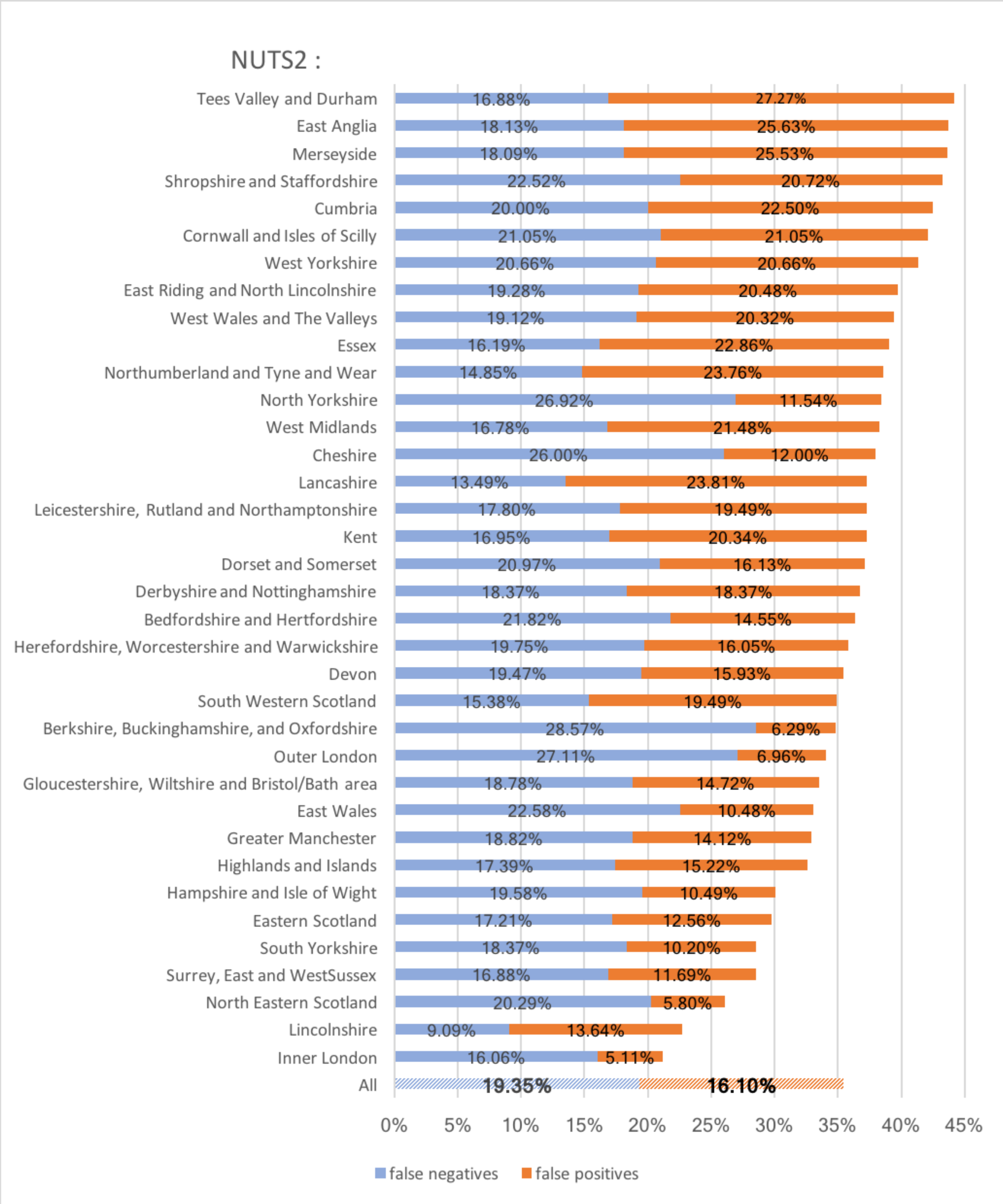
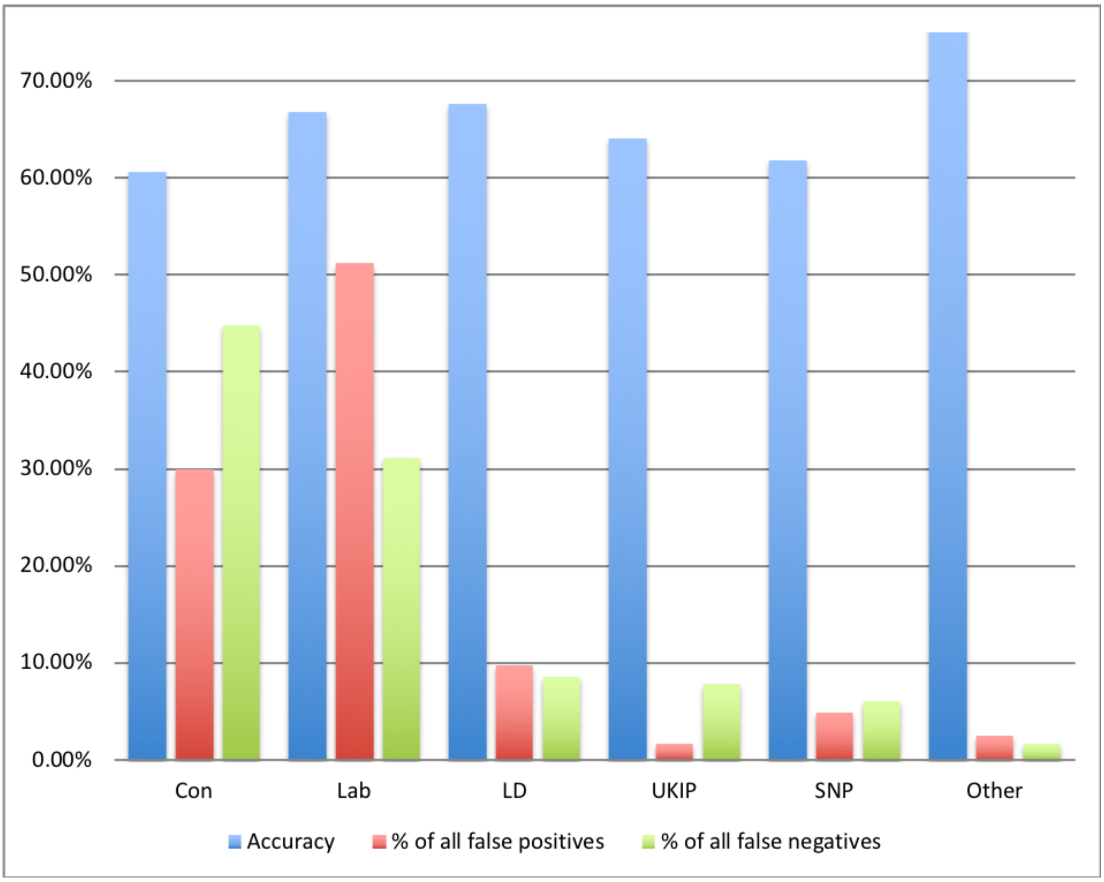


Figure 3.3: Overall accuracy and error decomposition by stated party preference (Conservative, Labour, Liberal Democrats, UKIP, SNP and Others)



C.6 Appendix: data and regression results

In this appendix we present our data and empirical regression results in more detail.

c.1 Sampling design

Concerning the design and data collection of *Understanding Society*, the general population sample is a stratified, clustered, equal probability sample of residential addresses drawn to a uniform design throughout the whole of the UK. For each wave, the data collection is spread over a two-year period, and the overall sample is divided into 24 monthly subsamples, each independently representative of the UK population. Computer assisted personal interviewing (CAPI) was mainly used to collect the data.[10]

c.2 Constructing the sample

The construction of our sample takes place in various steps. Initially, the raw individual survey (wave 8) consists of 21,076 observations. Then, matching the household survey leaves 20,821 individuals. Further matching with local authority codes results in a sample of 17,697 respondents (i.e. over 3000 surveyed individuals get lost because there is no location code associated with their households). Finally, we merge this last sample with the aggregate information used in Becker et al. (2017). In this last step, the number of surveyed individuals is 15,844 across 377 local authorities.

When we consider the initial sample with 21,076 observations, 91% of the individuals provide an answer to the question concerning British EU membership. Among them, the share of those supporting Leave is 35.8%. Of the selected subsample with 15,844 units, 91.4% (14,476 individuals) disclose an answer for the outcome variable, and 42.6% turn out to be Leave supporters.[11]

As a final remark, we want to stress that our estimates come from the analysis of three specific

[10] These details are taken from Understanding Society: Design Overview by Buck and McFall (2011). For further details refer to the Understanding Society User Guide (wave 1-6) by Knies (2016).

[11] In unreported tables (available upon request), we compare the 14,476 individuals individuals who answer the Brexit question to the 1368 non-respondents for each group of covariates (i.e. all regressors in Tables C.1a to C.10 in the appendix) and establish along which dimension the two groups are statistically different. If anything, non-respondents seem to display most of the characteristics of a typical Leave voter. More specifically, non-respondents are significantly older, less used to technology, with lower educational attainments and more frequently unemployed. In addition, they seek more medical attention, their housing status is more often local authority renting, and more of them receive income support. Finally, non-respondents are less often UK natives and more often members of an ethnic minority.

subsamples of the 14,476 selected respondents. The main one contains 13,136 individuals. The sample with housing tenure status contains 6,425 individuals. The subsample on employment characteristics counts 8,434 individuals.

c.3 Regression results

We divide our variables into groupings as follows. The first group of explanatory variables includes basic demographic features such as sex, age, marital status, education and employment. The second group explores data on individuals' use of health services. The third group captures information on housing (ownership vs. renting) drawn from the household questionnaire. The fourth group refers to employment. This is followed by a focus on unearned income and state benefits. The sixth group consists of life satisfaction indicators. The seventh and final group covers nationality and ethnicity.

The results are reported in Tables C.1a to C.10. We present linear probability models as the default, with the exception of logit models in Table C.1b, probit models in Table C.1c and weighted OLS models in Table C.1d and C.1e.[12]

When variables are perfectly comparable at individual and aggregate levels, the first three columns of the tables directly compare those to address the potential ecological fallacy concern.

c.4 Demographics, technology, education and employment

In Tables C.1a to C.3 we present results from regressions based on different types of demographic characteristics. Tables C.1a to C.1e explore the relationship of voting Leave with sex, age and technology use. Table C.1a presents our baseline results estimated with a linear probability model (OLS). Tables C.1b and C.1c use the same explanatory variables but estimated with logistic and probit regressions, respectively, where we report marginal effects. Table C.1d reports weighted OLS regressions, with weights provided by *Understanding Society*. Table C.1e also

[12]We would like to note that sampling weights in Understanding Society, which we use inTable C.1d are quite homogenous. In our main estimation sample, the median sampling weight is 0.956, the 25th percentile is 0.770 and the 75th percentile is 1.237. This explains why weighted and unweighted regression results are so similar. [USOC wave 8 data has a substantive number of observations with missing weights. This is due to the fact that it is a pre-release version. The final version of wave 8 is expected to be released towards the end of 2018 or in early 2019.] In Table C.1e, we mechanically re-weight the sample to align the share of 'Leave' voters with the actual referendum result.

displays weighted OLS regressions, but here we use artificial weights such that the proportion of Leave supporters in the sample matches the actual Brexit vote share. Overall, the coefficient signs and magnitudes are very similar across Tables C.1a to C.1c. They are also similar in comparison to Table C.1d and C.1e despite the weights and the reduced number of observations. We therefore focus our below discussion on Table C.1a.

Columns 1 to 3 of Table C.1a exhibit positive and significant coefficients for the old-age variables at both individual and aggregate levels, showing no evidence of ecological fallacy. Although the coefficient for the aggregate share of the elderly population is lower in magnitude, it presents a predictive power very similar to the individual counterpart. Column 4 indicates that males are 4.7% more likely to vote Leave. Compared to middle-aged respondents, the tendency to support Leave is substantially lower by 12.3% for younger cohorts up to the age of 30 and notably higher by 9.1% for individuals aged 60 or above. Columns 5 and 7 confirm these results in terms of significance even when we control for the share of the population aged 60 or above at the local authority level. In column 6 we focus on technology use. Individuals who do not use a smartphone are substantially more likely to vote Leave. Using the internet every day is associated with a substantially lower probability to vote Leave. These patterns persist even once we control for sex and age in column 7.

In Table C.2 we explore the predictive power of educational attainment. Again, variables on education attainment relate to the referendum outcome in the same way and with matching power at both individual and aggregate levels although aggregate coefficients have lower magnitude and significance. Hence, highly qualified individuals with university and college degrees are considerably less likely to vote Leave by over 20% compared to people with average qualifications. In contrast, having no qualification is a very strong predictor of voting Leave. These results holds up once we control for aggregate characteristics on educational attainment in columns 3 and 5 as well as sex and age in column 6.

Next, in Table C.3 we analyze individuals' current employment and marital status. At the individual level, comparison groups are predominantly retired and divorced respondents, respectively.[13] Here, aggregate rates on employment are indistinguishable from zero (although they

[13] Excluded categories among current activity feature Retired (64.7%), Looking after family or home (10%), Full-time student (14.3%), Long-term sick or disabled (7.5%), Doing something else (2.2%). Excluded categories

have the same predictive power as the individual variables, and self-employment and unemployment coefficients have the 'correct' sign). Column 1 of Table C.3 shows that self-employed and paid employees are more likely to support Remain (relative to mostly retired people). Column 4 shows that single and married people are significantly less likely to vote Leave (compared to divorcees, separated and widowed people). Again, most of these results hold up once we control for aggregate rates in column 3 as well as for age in column 5. Unemployment now also shows up as highly significant.[14]

To sum up our results on demographic variables, we find that individuals are more likely to support Leave if they are male, older, use less technology, are less qualified, retired or unemployed, and divorced, separated or widowed. These findings are consistent with the results by Becker et al. (2017) based on aggregate data who also find that age, low educational attainment and unemployment are key explanatory variables to predict the Leave vote shares across UK voting areas.

c.5 Health

Table C.4 analyzes the relationship between Brexit support and individuals' use of health services. Interestingly, columns 1 and 2 show that individuals who visit their general practitioner (GP) very frequently (over ten times in the previous 12 months) are more likely to support Leave. Those are arguably individuals of poor health or older generations. Conversely, those who did not visited the GP even once have a slightly higher probability to support Remain. Controlling for age in column 2 turns the latter result insignificant (possibly because it is young people who do not go to the doctor) but preserves the former result on frequent GP visits.

A similar picture emerges from columns 3 and 4, focusing on individuals who are never or extremely often classified as out-patients. The same holds for people admitted as in-patients at least once during the preceding 12 months. That is, people of poor health as proxied by

among marital status feature Divorced (57.4%), Separated (10.3%), Widowed (31.6%), Other (0.7%).

[14] To get a sense of whether *changes* in (un)employment status matter, in unreported regressions, we used additional information based on a short employment history (looking at respondents participating in both wave 7 and the pre-release version of wave 8 with the EU question). The results suggest that the preferences for Remain and Leave are quite static or do not respond in a remarkable fashion to individuals switching employment status (by becoming unemployed or employed between wave 7 and wave 8). Rather, the first-order differences in tendencies to support Leave or Remain for our prediction exercise are driven by individuals who are employed or unemployed in both survey waves, implying that looking at only the cross-section is sufficient to capture the role of employment variables.

frequent visits to the GP or hospital are substantially more likely to support Leave. Perhaps it is therefore no coincidence that a key pledge of the pro-Brexit referendum campaign was to invest more in the National Health Service (NHS).

c.6 Housing

Table C.5 explores the role of property values for home owners and housing tenure (owned vs. rented). We note that due to many missing values, we only have 6,425 observations in this table.

When directly comparing individual tenure status to corresponding aggregate shares we see similar paths (columns 1 to 3), in particular with respect to direct ownership which is positively related to Leave support.

In terms of individual housing tenure, owning their own property tends to make individuals more likely to support Leave, although this particular association is barely statistically significant. The omitted category here is renting through a housing association. More importantly, higher property values are significantly related to an increased likelihood of supporting Remain. A one-standard deviation increase in property values increases the Remain likelihood by roughly 4%. Property values are arguably positively linked to individuals' financial status, which would be consistent with earlier evidence on income based on aggregate data (see Becker et al., 2017).

c.7 Employment

This section shifts the focus towards employment-related determinants. For starters, Table C.6 indicates a higher probability of almost 10% to support Leave for individuals who did not work in the week prior to the questionnaire and who did not have a paid job compared to those respondents who were either working or had a paid job (stable across all specifications).

In Table C.7 we narrow our analysis to only those participants who worked or had a paid job. This reduces the number of observations to 8,434. First, columns 1 to 3 compare the individual sector of employment to the respective aggregate controls (manufacturing, construction, retail and finance as used in Becker et al., 2017). Estimates as well as their predictive power are aligned (although aggregate coefficients are lower in magnitude). Indeed, both specifications suggest that workers in the manufacturing, construction and retail industries are significantly more likely to support Leave. Note that individual estimates are fairly stable across all specifications.

In addition, it emerges from column 4 that those with a permanent job compared to those in non-permanent employment have a higher probability of supporting Leave. This result continues to hold qualitatively in column 5 after we control for individuals' age, sex and education as well as the sectoral distribution and growth of employment at the aggregate level in column 6. This result appears surprising, but we note that the subsample in Table C.7 is highly unbalanced in the sense that 90% of the respondents have a permanent job. Still, 60% of individuals with permanent jobs support Remain versus 70% of those with temporary jobs. It also appears likely that the very young respondents, who are overwhelmingly in favor of Remain, are less likely to hold permanent jobs. Our age dummies in column 5 might not pick up these age patterns appropriately. Finally, self-employed respondents are also more likely to support Leave, even though this association is insignificant for most specifications in the table.

Overall, consistent with the aggregate results in Becker et al. (2017) our findings support the view that individuals are more willing to vote for Brexit if they work in sectors such as manufacturing that have arguably been hit relatively hard by trade openness and international competition (also see Colantone and Stanig, 2018). In addition, workers in manufacturing, construction and retail sectors have lower educational attainment on average while the opposite is true for workers in financial sector.

c.8 Unearned income and state benefits

In Table C.8 we highlight the role of unearned income and state benefits. In column 1 we find that respondents who receive core benefits have significantly raised probability of supporting Leave compared to those receiving none. These core benefits are broken down into their various components in column 2. In particular, recipients of income support are substantially more likely to be in favor of Leave (by 20%), whereas job seeker's allowance, child benefit and universal credit do not matter.

Similar results hold for people receiving pensions. This particular finding is likely driven by the overwhelming share of older people amongst pension receivers (see section c.4). The same pattern holds for people on disability benefits, in line with our estimates on health service usage (see section c.5).

Finally, the opposite is true for respondents who receive other sources of income. Those are broken down in column 3. The key income streams are education grants and student loans as well as payments from family members living elsewhere. This suggests a tight link previously with age and education (see section c.4).

In summary, the forms of income and benefits in Table C.8 are likely correlated with more fundamental characteristics such as age and health, as discussed in previous tables.

c.9 Life satisfaction

In Table C.9 we explore the potential link between Brexit support and indices of health, income and life satisfaction. When looking at overall life satisfaction only (columns 1 to 3), the individual coefficients suggest that dissatisfied people are significantly more likely to favor Leave while the aggregate estimate implies that a higher relative dispersion of well-being across voting areas, which can be interpreted as a measure of life satisfaction inequality, has positive predictive power for the Leave support. Success rates of prediction are very similar whichever level of variation is considered.

In addition, people dissatisfied with health and income have a higher probability of supporting Leave by 5.5% and 6.4%, respectively. Once again, we can relate these findings to those in Table C.4 on health and Table C.8 on income and benefits. Interestingly, people dissatisfied with their amount of leisure time are significantly more likely to support Remain by 6.3%. This may be linked to the fact that these respondents have on average higher levels of educational attainment and they are generally younger. Note that when these individual variables are considered (columns 4 and 5) the individual estimate of overall life satisfaction is absorbed and becomes insignificant.

c.10 Nationality and ethnicity

Table C.10 provides insights on the importance of individuals' nationality and ethnicity in shaping their attitudes towards Brexit. Survey participants born in the UK as opposed to elsewhere have a significantly larger probability of supporting Leave by 12.4% (see column 1). It is useful to point out that in the sample, 90% of respondents are born in the UK, and 95% of them are white.

In terms of ethnic minorities compared to whites (see column 2), people of mixed ethnicity, Asians and black respondents all have a significantly larger probability of supporting Remain (in the range of 12% to 23%). These results are in line with the preliminary work by Martin and Sobolewska (2017).

Finally, aggregate controls for migration are insignificant with the exception of the EU share of migrants in 2001, which is positive linked with support for Remain.

C.7 Appendix tables

Table C.1a: Demographics: Sex, Age and Technology Use (OLS)

	Should the UK leave the EU						
	Ecological fallacy						
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)	(6)	(7)
Sex = Male				0.0471***	0.0473***		0.0502***
				(0.0073)	(0.0073)		(0.0072)
Age = 30 or younger				-0.1226***	-0.1192***		-0.1092***
				(0.0130)	(0.0126)		(0.0128)
Age = 60 or older	0.1240***		0.1174***	0.0911***	0.2808***		0.2367***
	(0.0113)		(0.0112)	(0.0116)	(0.0729)		(0.0726)
Share population 60 or older (2001)		0.0339***	0.0280***		0.0365***		0.0364***
		(0.0069)	(0.0069)		(0.0076)		(0.0075)
Age = 60 or older * Share population 60 or older (2001)					-0.0925***		-0.0917***
					(0.0336)		(0.0332)
Use smartphone = No						0.0776***	0.0382***
						(0.0126)	(0.0129)
Has mobile computing device = No						0.0134	0.0214**
						(0.0099)	(0.0098)
Frequency using internet = Every day						-0.1018***	-0.0760***
						(0.0148)	(0.0148)
Frequency using internet = No access						0.0358	0.0320
						(0.0398)	(0.0400)
Constant	0.3845***	0.4221***	0.3864***	0.3954***	0.3363***	0.4817***	0.3907***
	(0.0093)	(0.0083)	(0.0088)	(0.0097)	(0.0237)	(0.0154)	(0.0261)
Observations	13,136	13,136	13,136	13,136	13,136	13,136	13,136
Predictive success rate (from logit)	0.5820	0.5827	0.5834	0.5891	0.5910	0.5887	0.5917

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.1b: Demographics: Sex, Age and Technology Use (Logit)

	Should the UK leave the EU						
	Ecological fallacy						
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)	(6)	(7)
Sex = Male				0.0480***	0.0484***		0.0518***
				(0.0074)	(0.0074)		(0.0074)
Age = 30 or younger				-0.1265***	-0.1238***		-0.1147***
				(0.0134)	(0.0131)		(0.0135)
Age = 60 or older	0.1240***		0.1176***	0.0904***	0.2899***		0.2489***
	(0.0113)		(0.0112)	(0.0115)	(0.0696)		(0.0716)
Share population 60 or older (2001)		0.0340***	0.0284***		0.0382***		0.0382***
		(0.0070)	(0.0071)		(0.0081)		(0.0081)
Age = 60 or older * Share population 60 or older (2001)					-0.0980***		-0.0975***
					(0.0338)		(0.0338)
Use smartphone = No						0.0779***	0.0379***
						(0.0127)	(0.0129)
Has mobile computing device = No						0.0136	0.0221**
						(0.0100)	(0.0101)
Frequency using internet = Every day						-0.1019***	-0.0761***
						(0.0148)	(0.0149)
Frequency using internet = No access						0.0358	0.0319
						(0.0409)	(0.0412)
Observations	13,136	13,136	13,136	13,136	13,136	13,136	13,136

Notes: The table reports results from Logit regressions in terms of marginal effects. Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.1c: Demographics: Sex, Age and Technology Use (Probit)

	Should the UK leave the EU						
	Ecological fallacy						
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)	(6)	(7)
Sex = Male				0.0478***	0.0481***		0.0513***
				(0.0074)	(0.0074)		(0.0074)
Age = 30 or younger				-0.1257***	-0.1233***		-0.1139***
				(0.0133)	(0.0131)		(0.0134)
Age = 60 or older	0.1240***		0.1176***	0.0906***	0.2903***		0.2489***
	(0.0113)		(0.0112)	(0.0116)	(0.0704)		(0.0721)
Share population 60 or older (2001)		0.0341***	0.0285***		0.0382***		0.0383***
		(0.0070)	(0.0071)		(0.0081)		(0.0080)
Age = 60 or older * Share population 60 or older (2001)					-0.0980***		-0.0975***
					(0.0339)		(0.0338)
Use smartphone = No						0.0778***	0.0379***
						(0.0126)	(0.0129)
Has mobile computing device = No						0.0135	0.0219**
						(0.0100)	(0.0101)
Frequency using internet = Every day						-0.1018***	-0.0761***
						(0.0148)	(0.0148)
Frequency using internet = No access						0.0358	0.0321
						(0.0406)	(0.0409)
Observations	13,136	13,136	13,136	13,136	13,136	13,136	13,136

Notes: The table reports results from Probit regressions in terms of marginal effects. Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.1d: Demographics: Sex, Age and Technology Use (Weighted OLS using USOC sampling weights)

	Should the UK leave the EU						
	Ecological fallacy						
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)	(6)	(7)
Sex = Male				0.0359***	0.0359***		0.0390***
				(0.0120)	(0.0120)		(0.0120)
Age = 30 or younger				-0.1265***	-0.1255***		-0.1149***
				(0.0203)	(0.0204)		(0.0201)
Age = 60 or older	0.1393***		0.1351***	0.1030***	0.2986***		0.2608***
	(0.0133)		(0.0131)	(0.0143)	(0.0893)		(0.0903)
Share population 60 or older (2001)		0.0243***	0.0175**		0.0272***		0.0266***
		(0.0085)	(0.0084)		(0.0100)		(0.0100)
Age = 60 or older * Share population 60 or older (2001)					-0.0943**		-0.0951**
					(0.0413)		(0.0414)
Use smartphone = No						0.0818***	0.0355**
						(0.0167)	(0.0178)
Has mobile computing device = No						0.0004	0.0095
						(0.0147)	(0.0146)
Frequency using internet = Every day						-0.1089***	-0.0737***
						(0.0190)	(0.0189)
Frequency using internet = No access						-0.0021	-0.0074
						(0.0464)	(0.0468)
Constant	0.3901***	0.4332***	0.3910***	0.4093***	0.3488***	0.5019***	0.4040***
	(0.0108)	(0.0092)	(0.0106)	(0.0123)	(0.0310)	(0.0201)	(0.0360)
Observations	8,188	8,188	8,188	8,188	8,188	8,188	8,188

Notes: The table reports results from weighted linear probability regressions, using *Understanding Society* sampling weights (weighted OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.1e: Demographics: Sex, Age and Technology Use (Weighted OLS using weights that mechanically align share of 'Leave' voters with referendum result)

	Should the UK leave the EU						
	Ecological fallacy						
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)	(6)	(7)
Sex = Male				0.0480***	0.0481***		0.0510***
				(0.0074)	(0.0074)		(0.0073)
Age = 30 or younger				-0.1288***	-0.1255***		-0.1152***
				(0.0138)	(0.0134)		(0.0137)
Age = 60 or older	0.1254***		0.1188***	0.0916***	0.2916***		0.2478***
	(0.0113)		(0.0112)	(0.0116)	(0.0733)		(0.0731)
Share population 60 or older (2001)		0.0347***	0.0288***		0.0379***		0.0379***
		(0.0071)	(0.0071)		(0.0079)		(0.0078)
Age = 60 or older * Share population 60 or older (2001)					-0.0974***		-0.0966***
					(0.0337)		(0.0334)
Use smartphone = No						0.0782***	0.0380***
						(0.0126)	(0.0128)
Has mobile computing device = No						0.0136	0.0218**
						(0.0101)	(0.0100)
Frequency using internet = Every day						-0.1023***	-0.0760***
						(0.0147)	(0.0146)
Frequency using internet = No access						0.0332	0.0295
						(0.0380)	(0.0382)
Constant	0.4343***	0.4730***	0.4362***	0.4456***	0.3839***	0.5320***	0.4378***
	(0.0096)	(0.0085)	(0.0092)	(0.0100)	(0.0239)	(0.0153)	(0.0261)
Observations	13,136	13,136	13,136	13,136	13,136	13,136	13,136

Notes: The table reports results from weighted linear probability regressions (weighted OLS). Weights mechanically reproduce the Brexit referendum result. Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.2: Demographics: Education

	Should the UK leave the EU					
	Ecological fallacy					
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)	(6)
Highest qualification = Degree	-0.2590***		-0.2407***	-0.2348***	-0.2162***	-0.2394***
	(0.0115)		(0.0119)	(0.0119)	(0.0123)	(0.0123)
Highest qualification = Other higher degree	-0.0842***		-0.0793***	-0.0601***	-0.0556***	-0.0796***
	(0.0155)		(0.0156)	(0.0157)	(0.0157)	(0.0161)
Highest qualification = Other lower qualification				0.1478***	0.1473***	0.1019***
				(0.0177)	(0.0176)	(0.0171)
Highest qualification = No qualification	0.0988***		0.0948***	0.1229***	0.1190***	0.0822***
	(0.0181)		(0.0178)	(0.0186)	(0.0183)	(0.0184)
Share of res. pop. qualification 4+ (2001)		-0.0442***	-0.0292**		-0.0220*	-0.0171
		(0.0134)	(0.0128)		(0.0131)	(0.0127)
Share of res. pop. no qualifications (2001)		0.0244*	0.0180		0.0210*	0.0285**
		(0.0132)	(0.0127)		(0.0123)	(0.0119)
Share population 60 or older (2001)					0.0145**	0.0086
					(0.0065)	(0.0064)
Sex = Male						0.0520***
						(0.0072)
Age = 30 or younger						-0.1637***
						(0.0123)
Age = 60 or older						0.0323***
						(0.0119)
Constant	0.4968***	0.4223***	0.4915***	0.4727***	0.4673***	0.4798***
	(0.0098)	(0.0079)	(0.0092)	(0.0104)	(0.0096)	(0.0105)
Observations	13,136	13,136	13,136	13,136	13,136	13,136
Predictive success rate (from logit)	0.5940	0.5887	0.6052	0.6092	0.6135	0.6404

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.3: Demographics: Employment and Marital Status

	Should the UK leave the EU				
	Ecological fallacy				
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)
Current activity = Self-employed	-0.0489***		-0.0473**		-0.0263
	(0.0184)		(0.0183)		(0.0195)
Current activity = In paid employment	-0.0773***		-0.0775***		-0.0336***
	(0.0112)		(0.0112)		(0.0125)
Current activity = Unemployed	0.0199		0.0161		0.0896***
	(0.0266)		(0.0269)		(0.0269)
Self-employment rate (2015)		-0.0026	-0.0043		-0.0052
		(0.0082)	(0.0081)		(0.0079)
Employment rate (2015)		0.0027	0.0049		0.0027
		(0.0113)	(0.0113)		(0.0111)
Unemployment rate (2015)		0.0156	0.0157		0.0188*
		(0.0114)	(0.0114)		(0.0112)
Current legal marital status = Single				-0.1714***	-0.0775***
				(0.0149)	(0.0166)
Current legal marital status = Married or civil partner				-0.0890***	-0.0679***
				(0.0136)	(0.0139)
Age = 30 or younger					-0.1213***
					(0.0160)
Age = 60 or older					0.0639***
					(0.0136)
Constant	0.4653***	0.4223***	0.4654***	0.5201***	0.5006***
	(0.0107)	(0.0087)	(0.0108)	(0.0137)	(0.0175)
Observations	13,136	13,136	13,136	13,136	13,136
Predictive success rate (from logit)	0.5839	0.5839	0.5843	0.5937	0.5921

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.4: Health

VARIABLES	Should the UK leave the EU (1)	(2)	(3)	(4)	(5)	(6)
Visits GP in 12m = None	-0.0271**	-0.0113				
	(0.0107)	(0.0106)				
Visits GP in 12m = Over 10	0.0892***	0.0847***				
	(0.0203)	(0.0199)				
Age = 60 or older		0.1213***		0.1200***		0.1226***
		(0.0113)		(0.0114)		(0.0114)
Share of suspected cancer patient treated within 62 Days (2015)		-0.0213**		-0.0211**		-0.0214**
		(0.0088)		(0.0087)		(0.0088)
Out-patient in 12m = None			-0.0331***	-0.0166*		
			(0.0085)	(0.0086)		
Out-patient in 12m = Over 10			0.0791***	0.0717***		
			(0.0277)	(0.0271)		
In-patient in 12m = Yes					0.0475***	0.0353**
					(0.0150)	(0.0150)
Constant	0.4226***	0.3826***	0.4379***	0.3926***	0.4182***	0.3820***
	(0.0091)	(0.0096)	(0.0097)	(0.0103)	(0.0088)	(0.0091)
Observations	13,136	13,136	13,136	13,136	13,136	13,136
Predictive success rate (from logit)	0.5818	0.5974	0.5880	0.5910	0.5839	0.5944

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.5: Housing

	Should the UK leave the EU					
	Ecological fallacy					
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)	(6)
Value of property: home owners				-0.0428***	-0.0422***	-0.0408***
				(0.0111)	(0.0114)	(0.0112)
Housing tenure = Owned (outright + mortgage)	0.1595**		0.1269*	0.1141*	0.1143*	0.1303*
	(0.0760)		(0.0694)	(0.0626)	(0.0682)	(0.0712)
Housing tenure = Rented from employer or private	-0.0109		0.0100		-0.0020	0.0206
	(0.1831)		(0.1681)		(0.1675)	(0.1688)
Housing tenure = Local authority rent	0.2391		0.1811			0.2285
	(0.1923)		(0.1880)			(0.1836)
Age = 60 or older				0.1311***	0.1317***	0.1318***
				(0.0151)	(0.0151)	(0.0151)
Owned (outright + mortgage) share (2001)		0.0799***	0.0783***	0.0372***	0.0201	0.0469**
		(0.0183)	(0.0184)	(0.0069)	(0.0180)	(0.0214)
Owned (outright + mortgage) share growth (2001-2011)				-0.0184**	-0.0391*	-0.0613***
				(0.0075)	(0.0214)	(0.0232)
Private rented share (2001)		0.0132	0.0132		-0.0139	0.0013
		(0.0116)	(0.0116)		(0.0119)	(0.0134)
Privated rented share growth (2001-2011)					-0.0170	-0.0396*
					(0.0209)	(0.0228)
Council rented share (2001)		0.0360**	0.0350**			0.0380**
		(0.0174)	(0.0174)			(0.0189)
Council rented share growth (2001-2011)						-0.0115
						(0.0093)
Constant	0.2609***	0.4187***	0.2928***	0.2611***	0.2607***	0.2444***
	(0.0757)	(0.0076)	(0.0691)	(0.0622)	(0.0678)	(0.0708)
Observations	6,425	6,425	6,425	6,425	6,425	6,425
Predictive success rate (from logit)	0.5772	0.5729	0.5729	0.6026	0.6002	0.6050

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Two-way clustered standard errors at the local authority and household levels are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.6: Current Employment: All Individuals

	Should the UK leave the EU			
	Ecological fallacy			
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)
No work last week & doesn't have paid job	0.0979***		0.0970***	0.0939***
	(0.0101)		(0.0101)	(0.0098)
Unemployment rate (2015)		0.0144*	0.0127	0.0171**
		(0.0083)	(0.0082)	(0.0074)
Manufacturing employment share (2001)				0.0347***
				(0.0079)
Manufacturing employment share change (2001-2011)				0.0042
				(0.0081)
Construction employment share (2001)				0.0086
				(0.0105)
Construction employment share change (2001-2011)				0.0088
				(0.0098)
Retail employment share (2001)				0.0361***
				(0.0078)
Retail employment share change (2001-2011)				-0.0237***
				(0.0068)
Finance employment share (2001)				0.0023
				(0.0078)
Finance employment share change (2001-2011)				-0.0131
				(0.0085)
Constant	0.3852***	0.4223***	0.3856***	0.3866***
	(0.0097)	(0.0087)	(0.0098)	(0.0084)
Observations	13,136	13,136	13,136	13,136
Predictive success rate (from logit)	0.5839	0.5839	0.5797	0.5942

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.7: Current Employment: Individuals With Paid Jobs

	Should the UK leave the EU					
	Ecological fallacy					
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)	(6)
Current job = Permanent				0.0924***	0.0580***	0.0504***
				(0.0177)	(0.0176)	(0.0174)
Current job sector = Manufacturing	0.1395***		0.1159***	0.1352***	0.0731***	0.0597***
	(0.0191)		(0.0187)	(0.0191)	(0.0191)	(0.0186)
Current job sector = Construction	0.1704***		0.1604***	0.1659***	0.0858***	0.0811***
	(0.0268)		(0.0266)	(0.0269)	(0.0260)	(0.0257)
Current job sector = Wholesale & retail	0.0877***		0.0747***	0.0854***	0.0350**	0.0284*
	(0.0177)		(0.0176)	(0.0177)	(0.0172)	(0.0172)
Current job sector = Finance	-0.0449		-0.0415	-0.0484*	-0.0437	-0.0426
	(0.0277)		(0.0274)	(0.0278)	(0.0275)	(0.0272)
Current job = Self-employed	0.0267		0.0301*	0.0380**	0.0136	0.0153
	(0.0177)		(0.0173)	(0.0175)	(0.0168)	(0.0166)
Sex = Male					0.0328***	0.0329***
					(0.0099)	(0.0100)
Age = 30 or younger					-0.1338***	-0.1306***
					(0.0142)	(0.0138)
Age = 60 or older					0.0184	0.0200
					(0.0189)	(0.0185)
Highest qualification = Degree					-0.2714***	-0.2563***
					(0.0124)	(0.0127)
Highest qualification = Other higher degree					-0.1020***	-0.0979***
					(0.0185)	(0.0188)
Manufacturing employment share (2001)		0.0470***	0.0437***			0.0332***
		(0.0092)	(0.0091)			(0.0091)
Manufacturing employment share change (2001-2011)						0.0005
						(0.0082)
Construction employment share (2001)		0.0250**	0.0236**			0.0038
		(0.0109)	(0.0107)			(0.0098)
Construction employment share change (2001-2011)						-0.0001
						(0.0093)
Retail employment share (2001)		0.0386***	0.0377***			0.0260***
		(0.0088)	(0.0087)			(0.0083)
Retail employment share change (2001-2011)						-0.0143*
						(0.0077)
Finance employment share (2001)		0.0086	0.0106			0.0134
		(0.0092)	(0.0093)			(0.0090)
Finance employment share change (2001-2011)						-0.0178**
						(0.0090)
Self-employment rate (2015)		0.0149*	0.0144			0.0127
		(0.0089)	(0.0088)			(0.0089)
Unemployment rate (2015)						0.0161**
						(0.0078)
Constant	0.3550***	0.3874***	0.3583***	0.2709***	0.4354***	0.4375***
	(0.0103)	(0.0084)	(0.0095)	(0.0195)	(0.0210)	(0.0200)
Observations	8,434	8,434	8,434	8,434	8,434	8,434
Predictive success rate (from logit)	0.6111	0.6204	0.6132	0.6118	0.6370	0.6560

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.8: Unearned Income and State Benefits

	Should the UK leave the EU		
VARIABLES	(1)	(2)	(3)
Receives core benefits	0.0357***		
	(0.0124)		
Receives pensions	0.1177***		
	(0.0115)		
Receives disability benefits	0.1119***		
	(0.0167)		
Receives other benefits or credits	0.0867***		
	(0.0131)		
Receives other sources of income	-0.0496***		
	(0.0152)		
Core benefits:			
Income Support		0.2002***	
		(0.0354)	
Job Seeker's Allowance		0.0467	
		(0.0404)	
Child Benefit		0.0006	
		(0.0131)	
Universal Credit		0.0357	
		(0.0472)	
Other sources of income:			
Education Grant other than a Student Loan or Tuition Fee Loan			-0.2329***
			(0.0286)
Trade Union or Friendly Society Payment			0.1427
			(0.1883)
Maintenance or Alimony			0.0636*
			(0.0361)
Payments from a family member not living with you			-0.1063***
			(0.0402)
Rent from Boarders or Lodgers (not family members) living here with you			-0.0823
			(0.0646)
Rent from any other property even if that only covers that property's mortg			-0.0317
			(0.0201)
Or any other regular payment			-0.0812**
			(0.0381)
Constant	0.3615***	0.4182***	0.4287***
	(0.0102)	(0.0090)	(0.0089)
Observations	13,136	13,136	13,136
Predictive success rate (from logit)	0.5977	0.5875	0.5841

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.9: Life Satisfaction

	Should the UK leave the EU				
	Ecological fallacy				
VARIABLES	(1) Individual	(2) Aggregate	(3) Both	(4)	(5)
Dissatisfied with health				0.0549***	0.0533***
				(0.0116)	(0.0116)
Dissatisfied with income				0.0643***	0.0652***
				(0.0121)	(0.0121)
Dissatisfied with amount of leisure time				-0.0625***	-0.0629***
				(0.0111)	(0.0111)
Dissatisfied with life overall	0.0252**		0.0262**	-0.0101	-0.0085
	(0.0125)		(0.0125)	(0.0156)	(0.0156)
CV life satisfaction APS well-being data (2015)		0.0263***	0.0265***		0.0263***
		(0.0073)	(0.0073)		(0.0073)
Constant	0.4187***	0.4223***	0.4186***	0.4107***	0.4108***
	(0.0089)	(0.0085)	(0.0087)	(0.0095)	(0.0093)
Observations	13,136	13,136	13,136	13,136	13,136
Predictive success rate (from logit)	0.5839	0.5848	0.5866	0.5804	0.5850

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Table C.10: Nationality and Ethnicity

	Should the UK leave the EU			
VARIABLES	(1)	(2)	(3)	(4)
Born in UK	0.1239***		0.0706***	0.0581***
	(0.0160)		(0.0184)	(0.0182)
Ethnic group = Mixed		-0.2298***	-0.2154***	-0.2020***
		(0.0274)	(0.0271)	(0.0280)
Ethnic group = Asian		-0.1186***	-0.0817***	-0.0823***
		(0.0228)	(0.0252)	(0.0266)
Ethnic group = Black		-0.1864***	-0.1497***	-0.1301***
		(0.0210)	(0.0222)	(0.0281)
Ethnic group = Other ethnic group		-0.0687	-0.0274	-0.0252
		(0.0549)	(0.0561)	(0.0594)
EU migrant resident share (2001)				-0.0382***
				(0.0112)
Non-EU migrant resident share (2001)				0.0150
				(0.0170)
EU migrant resident growth (2001-2011)				0.0083
				(0.0139)
Non-EU migrant resident growth (2001-2011)				-0.0136
				(0.0128)
Constant	0.3102***	0.4379***	0.3706***	0.3812***
	(0.0151)	(0.0092)	(0.0194)	(0.0195)
Observations	13,136	13,136	13,136	13,136
Predictive success rate (from logit)	0.5839	0.5839	0.5839	0.5850

Notes: The table reports results from linear probability regressions (OLS). Non-dummy variables are standardized. Authority-level clustered standard errors are presented in parentheses, asterisks indicate *** p<0.01, ** p<0.05, * p<0.1.

Bibliography

Agranov, M., J. K. Goeree, J. Romero, and L. Yariv (2018). What makes voters turn out: The effects of polls and beliefs. *J. Eur. Econ. Assoc. 16*(3), 825–856.

Alabrese, E., S. O. Becker, T. Fetzer, and D. Novy (2019). Who voted for brexit? individual and regional data combined. *European Journal of Political Economy 56*, 132–150.

Ali, B. S. N. and C. Lin (2013). Why People Vote: Ethical Motives and Social Incentives. *Am. Econ. J. Microeconomics 5*(2), 73–98.

Allcott, H. and M. Gentzkow (2017). Social media and fake news in the 2016 election. *J. Econ. Perspect. 31*(2), 211–236.

Althaus, S. L., P. F. Nardulli, and D. R. Shaw (2002). Candidate Appearances in Presidential Elections, 1972-2000. *Polit. Commun. 19*(1), 49–72.

Arzheimer, K. (2009). Contextual factors and the extreme right vote in western europe, 1980–2002. *American Journal of Political Science 53*(2), 259–275.

Ashcroft, L. (2016). How the united kingdom voted on thursday… and why. *Lord Ashcroft Polls 24*, 1–14.

Ashworth, S. and J. D. Clinton (2006). Does Advertising Exposure Affect Turnout? *Quart. J. Polit. Sci. 1*(1), 3–23.

Azoulay, P., A. Bonatti, and J. L. Krieger (2017). The career effects of scandal: Evidence from scientific retractions. *Res. Policy 46*(9), 1552–1569.

Azoulay, P., J. L. Furman, Krieger, and F. Murray (2015, dec). Retractions. *Rev. Econ. Stat. 97*(5), 1118–1136.

Banducci, S. and C. Hanretty (2014, nov). Comparative determinants of horse-race coverage. *Eur. Polit. Sci. Rev. 6*(04), 621–640.

Bar-Ilan, J. and G. Halevi (2017). Post retraction citations in context: a case study. *Scientometrics 113*(1), 547–565.

Barone, G., A. D'Ignazio, G. De Blasio, and P. Naticchioni (2016). Mr. rossi, mr. hu and politics. the role of immigration in shaping natives' voting behavior. *Journal of Public Economics 136*, 1–13.

Barzel, Y. and E. Silberberg (1973). Is the Act of Voting Rational? *Public Choice 16*(Fall 1973), 51–58.

Becker, S. O., T. Fetzer, et al. (2016). Does migration cause extreme voting? *Center for Competitive Advantage in the Global Economy and The Economic & Social Research Council*, 1–54.

Becker, S. O., T. Fetzer, and D. Novy (2017). Who voted for brexit? a comprehensive district-level analysis. *Economic Policy 32*(92), 601–650.

Bennion, E. A. (2005). Caught in the ground wars: Mobilizing voters during a competitive Congressional campaign. *Ann. Am. Acad. Pol. Soc. Sci. 601*, 123–141.

Blais, A. (2006). What affects voter turnout? *Annu. Rev. Polit. Sci. 9*, 111–125.

Bordignon, F. (2020). Self-correction of science: a comparative study of negative citations and post-publication peer review. *Scientometrics 124*(2), 1225–1239.

Brainard, J. (2018, oct). Rethinking retractions. *Science 362*(6413), 390–393.

Bratton, M. and M. S. Kimenyi (2008). Voting in kenya: Putting ethnicity in perspective. *Journal of Eastern African Studies 2*(2), 272–289.

Brennan, G. and J. Buchanan (1984). Voter Choice: Evaluating Political Alternatives. *Am. Behav. Sci. 28*(2), 185–201.

Buck, N. and S. McFall (2011). Understanding society: design overview. *Longitudinal and Life Course Studies 3*(1), 5–17.

Burgoon, B. (2013). Inequality and anti-globalization backlash by political parties. *European Union Politics 14*(3), 408–435.

Burn-Murdoch, J. (2016). Brexit: Voter turnout by age. *Financial Times 24*, 16.

Bursztyn, L., D. Cantoni, P. Funk, F. Schönenberger, and N. Yuchtman (2020). Identifying the Effect of Election Closeness on Voter Turnout: Evidence from Swiss Referenda. (grant 172739).

Buscha, F., A. Maurel, L. Page, and S. Speckesser (2012). The Effect of Employment while in High School on Educational Attainment: A Conditional Difference-in-Differences Approach. *Oxf. Bull. Econ. Stat. 74*(3), 380–396.

Cancela, J. and B. Geys (2016). Explaining voter turnout: A meta-analysis of national and subnational elections. *Elect. Stud. 42*, 264–275.

Card, D. and S. Dellavigna (2020). What do editors maximize? Evidence from four economics journals. *Rev. Econ. Stat. 102*(1), 195–217.

Card, D., S. Dellavigna, P. Funk, and N. Iriberri (2020). Are Referees and Editors in Economics Gender Neutral? *Q. J. Econ. 135*(1), 269–327.

Chicco, D. and G. Jurman (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics 21*(1), 1–13.

Cipullo, D. and A. Reslow (2019). Biased Forecasts to Affect Voting Decisions? The Brexit Case.

Clarke, H. D., M. Goodwin, and P. Whiteley (2017). Why britain voted for brexit: An individual-level analysis of the 2016 referendum vote. *Parliamentary Affairs 70*(3), 439–464.

Clarke, S. and M. Whittaker (2016). *The importance of place: Explaining the characteristics underpinning the Brexit vote across different parts of the UK.* Resolution Foundation.

Coate, S., M. Conlin, and A. Moro (2008). The performance of pivotal-voter models in small-scale elections: Evidence from Texas liquor referenda. *J. Public Econ. 92*(3-4), 582–596.

Colantone, I. and P. Stanig (2018). Global competition and brexit. *American political science review 112*(2), 201–218.

Correia, S., P. Guimarães, and T. Z. Zylkin (2020). Fast Poisson estimation with high-dimensional fixed effects. *Stata J. 20*(1), 95–115.

Cox, A., R. Craig, and D. Tourish (2018). Retraction statements and research malpractice in economics. *Res. Policy 47*(5), 924–935.

Cox, G. W. (1999). Electoral Rules and the Calculus of Mobilization. *Legis. Stud. Q. 24*(3), 387.

Cox, G. W. and M. C. Munger (1989). Closeness, Expenditures, and Turnout in the 1982 U.S. House Elections. *Am. Polit. Sci. Rev. 83*(1), 217–231.

Cox, G. W., F. M. Rosenbluth, M. F. Thies, and M. F. Thies (1998). Mobilization, Social Networks, and Turnout: Evidence from Japan. *World Polit. 50*(3), 447–474.

Creighton, M. J. and J. Amaney (2017, July). Masked intolerance toward immigrants in the uk before and after the vote to leave the eu. *Understanding Society Scientific Conference. University of Essex. Unpublished conference paper*.

Dale, A. and A. Strauss (2009). Don't forget to Vote: Text Message as a Mobilization. *Am. J. Pol. Sci. 53*(4), 787–804.

Darvas, Z. (2016). High inequality and poverty helped trigger the brexit protest vote. *London: London School of Economics and Political Science.*.

De Vries, C. and H. Solaz (2017, July). Brexit and the left behind? disentangling the effects of assets, benefits, income and skills on voting behaviour in the 2016 brexit referendum. *EU Referendum Project Symposium. University of Essex. Unpublished conference paper*.

DellaVigna, S. and E. Kaplan (2007). The Fox News Effect: Media Bias and Voting Author. *Q. J. Econ. 122*(3), 1187–1234.

Dellavigna, S., J. A. List, U. Malmendier, and G. Rao (2017). Voting to tell others. *Rev. Econ. Stud. 84*(1), 143–181.

Dippel, C., R. Gold, and S. Heblich (2015). Globalization and its (dis-) content: Trade shocks and voting behavior.

Doebler, S., G. C. S. S. and B. Hayes (2017, July). What is the role of regional economic decline, personal economic struggle and political disengagement as drivers of support for brexit? a longitudinal analysis using understanding society (ukhls) data. *EU Referendum Project Symposium. University of Essex. Unpublished conference paper*.

Dorn, D., G. Hanson, K. Majlesi, et al. (2016). Importing political polarization? the electoral consequences of rising trade exposure. *NBER Working Paper 22637*.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *J. Polit. Econ. 65*(2), 135–150.

Duffy, J. and M. Tavits (2008). Beliefs and Voting Decisions: A Test of the Pivotal Voter Model. *Am. J. Pol. Sci. 52*(3), 603–618.

Dumas-Mallet, E., A. Garenne, T. Boraud, and F. Gonon (2020). Does newspapers coverage influence the citations count of scientific publications? An analysis of biomedical studies. *Scientometrics 123*(1), 413–427.

Enikolopov, B. R., M. Petrova, and E. Zhuravskaya (2011). Media and Political Persuasion: Evidence from Russia. *Am. Econ. Rev. 101*(7), 3253–3285.

Enos, R. D. and A. Fowler (2014). Pivotality and Turnout: Evidence from a Field Experiment in the Aftermath of a Tied Election. *Polit. Sci. Res. Methods 2*(02), 309–319.

Fanelli, D. (2013, jun). Any publicity is better than none: newspaper coverage increases citations, in the UK more than in Italy. *Scientometrics 95*(3), 1167–1177.

Farber, H. S. (2010). Rational Choice and Voter Turnout: Evidence from Union Representation Elections.

Feddersen, T. and A. Sardoni (2006). A Theory of Participation in Elections. *Am. Econ. Rev. 96*(4), 1271–1282.

Ferree, K. E., G. B. Powell, and E. Scheiner (2014). Context, electoral rules, and party systems. *Annual Review of Political Science 17*, 421–439.

Fetzer, T. (2018). Did austerity cause brexit? *CAGE Working Paper 381/2018*.

Franklin, M. N., C. van der Eijk, D. Evans, M. Fotos, W. Hirczy de Mino, M. Marsh, and B. Wessels (2004, apr). *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies since 1945.* Cambridge University Press.

Fujiwara, T., K. Meng, and T. Vogl (2016). Habit formation in voting: Evidence from rainy elections. *Am. Econ. J. Appl. Econ. 8*(4), 160–188.

Funk, P. (2010). Social incentives and voter turnout: Evidence from the Swiss mail ballot system. *J. Eur. Econ. Assoc. 8*(5), 1077–1103.

Furman, J. L., K. Jensen, and F. Murray (2012). Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. *Res. Policy 41*(2), 276–290.

Gareth, J., W. Daniela, H. Trevor, and T. Robert (2013). *An introduction to statistical learning: with applications in R.* Spinger, New York.

Gentzkow, B. M., J. M. Shapiro, and M. Sinkinson (2011). The Effect of Newspaper Entry and Exit on Electoral Politics. *Am. Econ. Rev. 101*(7), 2980–3018.

Gentzkow, M. (2006). Television and Voter Turnout. *Q. J. Econ. 121*(3), 931–972.

Gentzkow, M. and J. Shapiro (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica 78*(1), 35–71.

Gerber, A., M. Hoffman, J. Morgan, and C. Raymond (2020, jul). One in a Million: Field Experiments on Perceived Closeness of the Election and Voter Turnout. *Am. Econ. J. Appl. Econ. 12*(3), 287–325.

Gerber, A. S. and D. P. Green (2000). The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment. *Am. Polit. Sci. Rev. 94*(3), 653–663.

Gerber, A. S., J. Gruber, and D. M. Hungerman (2015). Does Church Attendance Cause People to Vote? Using Blue Laws' Repeal to Estimate the Effect of Religiosity on Voter Turnout. *Br. J. Polit. Sci. 46*, 481–500.

Gimpel, J. G., K. M. Kaufmann, and S. Pearson-merkowitz (2007). Battleground States versus Blackout States: The Behavioral Implications of Modern Presidential Campaigns. *J. Polit. 69*(3), 786–797.

Goncalves, D., J. Libgober, and J. Willis (2021). Learning versus Unlearning: An Experiment on Retractions.

Gourieroux, A. C., A. Monfort, and A. Trognon (1984). Pseudo Maximum Likelihood Methods: Theory. *Econometrica 52*(3), 681–700.

Großer, J. and A. Schram (2010). Public opinion polls, voter turnout, and welfare: An experimental study. *Am. J. Pol. Sci. 54*(3), 700–717.

Heckman, J., H. Ichimura, J. Smith, and P. Todd (1998, sep). Characterizing Selection Bias Using Experimental Data. *Econometrica 66*(5), 1017.

Heckman, J. J., H. Ichimura, and P. E. Todd (1997, oct). Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Rev. Econ. Stud. 64*(4), 605–654.

Herrera, H., M. Morelli, and T. Palfrey (2014). Turnout and power sharing. *Econ. J. 124*(574).

Hesselmann, F., V. Graf, M. Schmidt, and M. Reinhart (2017). The visibility of scientific misconduct: A review of the literature on retracted journal articles. *Curr. Sociol. Rev. 65*(6), 814–845.

Hobolt, S. B. (2016). The brexit vote: a divided nation, a divided continent. *Journal of European Public Policy 23*(9), 1259–1277.

Hochschild, A. R. (2018). *Strangers in their own land: Anger and mourning on the American right.* The New Press, New York.

Hoffman, M., G. León, and M. Lombardi (2017). Compulsory voting, turnout, and government spending: Evidence from Austria. *J. Public Econ. 145*, 103–115.

Ivanova, A., M. S. Schäfer, I. Schlichting, and A. Schmidt (2013). Is There a Medialization of Climate Science? Results From a Survey of German Climate Scientists. *Sci. Commun. 35*(5), 626–653.

Jin, G. Z., B. Jones, S. F. Lu, and B. Uzzi (2019). The reverse matthew effect: Consequences of retraction in scientific teams. *Rev. Econ. Stat. 101*(3), 492–506.

Knies, G. (2016). Understanding society - the uk household longitudinal study: Waves 1-6, (2009-2015) user guide. *University of Essex, Colchester.*.

Langella, M. and A. Manning (2016). Who voted leave: the characteristics of individuals mattered, but so did those of local areas. *British Politics and Policy at LSE*.

Lazer, D. M. J., M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, and J. L. Zittrain (2018, mar). The science of fake news. *Science 359*(6380), 1094–1096.

León, G. (2017). Turnout, political preferences and information: Experimental evidence from Peru. *J. Dev. Econ. 127*(December 2016), 56–71.

Levine, D. K. and T. R. Palfrey (2007). The Paradox of Voter Participation? A Laboratory Study. *Am. Polit. Sci. Rev. 101*(1), 143–158.

Lewandowsky, S., U. K. Ecker, C. M. Seifert, N. Schwarz, and J. Cook (2012). Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychol. Sci. Public Interes. Suppl. 13*(3), 106–131.

Liberini, F., A. J. Oswald, E. Proto, and M. Redoano (2017). Was brexit caused by the unhappy and the old?

Lu, S. F., G. Z. Jin, B. Uzzi, and B. Jones (2013). The Retraction Penalty: Evidence from the Web of Science. *Sci. Rep. 3*, 1–5.

Madestam, A., D. Shoag, S. Veuger, and D. Yanagizawa-Drott (2013). Do political protests matter? Evidence from the tea party movement. *Q. J. Econ.*, 1633–1685.

Martin, N. and M. Sobolewska (2017, July). Ethnic minority support for leaving the european union. *EU Referendum Project Symposium. University of Essex. Unpublished conference paper.*.

Matsusaka, J. G. (1993). Election closeness and voter turnout: Evidence from California ballot propositions. *Public Choice 76*(4), 313–334.

Matsuyama, W., H. Mitsuyama, M. Watanabe, K.-i. Oonakahara, I. Higashimoto, M. Osame, and K. Arimura (2005, dec). RETRACTED: Effects of Omega-3 Polyunsaturated Fatty Acids on Inflammatory Markers in COPD. *Chest 128*(6), 3817–3827.

Mehra, M. R., S. S. Desai, F. Ruschitzka, and A. N. Patel (2020). RETRACTED: Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. *Lancet*.

Mongeon, P. and V. Larivière (2016, mar). Costly collaborations: The impact of scientific fraud on co-authors' careers. *J. Assoc. Inf. Sci. Technol. 67*(3), 535–542.

Morton, R. B., D. Muller, L. Page, and B. Torgler (2015). Exit polls, turnout, and bandwagon voting: Evidence from a natural experiment. *Eur. Econ. Rev. 77*, 65–81.

Mullainathan, S. and J. Spiess (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives 31*(2), 87–106.

Ortoleva, B. P. and E. Snowberg (2015). Overconfidence in Political Behavior. *Am. Econ. Rev. 105*(2), 504–535.

Palfrey, T. R. and H. Rosenthal (1985, mar). Voter Participation and Strategic Uncertainty. *Am. Polit. Sci. Rev. 79*(1), 62–78.

Peng, H., D. M. Romero, and E.-Á. Horvát (2022, jun). Dynamics of cross-platform attention to retracted papers. *Proc. Natl. Acad. Sci. 119*(25).

Phillips, D. P., E. J. Kanter, B. Bednarczyk, and P. L. Tastad (1991, oct). Importance of the Lay Press in the Transmission of Medical Knowledge to the Scientific Community. *N. Engl. J. Med. 325*(16), 1180–1183.

Piller, C. (2021, jan). Disgraced COVID-19 studies are still routinely cited. *Science 371*(6527), 331–332.

Pollock, G. (2017, July). Melancholy politics: Brexit as a consequence of weakening mainstream party attachment in the united kingdom. *Understanding Society Scientific Conference. University of Essex. Unpublished conference paper.*.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program 14*(3), 130–137.

Prat, A. and D. Strömberg (2013). The Political Economy of Mass Media. In D. Acemoglu, M. Arellano, and E. Dekel (Eds.), *Adv. Econ. Econom. Tenth World Congr.*, pp. 135–187. Cambridge: Cambridge University Press.

Riker, W. H. and P. C. Ordeshook (1968). A Theory of the Calculus of Voting. *Am. Polit. Sci. Rev. 62*(1), 25–42.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review 15*(3), 351–357.

Rosati, D. (2021, apr). Citations are not opinions: a corpus linguistics approach to understanding how citations are made.

Sampson, T. (2017). Brexit: the economics of international disintegration. *Journal of Economic perspectives 31*(4), 163–184.

Sarathchandra, D. and A. M. McCright (2017). The Effects of Media Coverage of Scientific Retractions on Risk Perceptions. *SAGE Open 7*(2).

Schneider, J., D. Ye, A. M. Hill, and A. S. Whitehorn (2020). *Continued post-retraction citation of a fraudulent clinical trial report, 11 years after it was retracted for falsifying data*, Volume 125.

Schuessler, A. A. (2000). Expressive Voting. *Ration. Soc. 12*(1), 87–119.

Selb, P. (2009). A deeper look at the proportionality-turnout nexus. *Comp. Polit. Stud. 42*(4), 527–548.

Serghiou, S., R. M. Marton, and J. P. Ioannidis (2021). Media and social media attention to retracted articles according to Altmetric. *PLoS One 16*(5 May 2021), 1–19.

Serra-Garcia, M. and U. Gneezy (2021, may). Nonreplicable publications are cited more than replicable ones. *Sci. Adv. 7*(21), 1–23.

Shachar, B. R. O. N. and B. Nalebuff (1999). Theory and Evidence on Political Participation. *Am. Econ. Rev. 89*(3), 525–547.

Smith, J. and P. Todd (2005, mar). Does matching overcome LaLonde's critique of nonexperimental estimators? *J. Econom. 125*(1-2), 305–353.

Strömberg, D. (2004). Radio's impact on public spending. *Q. J. Econ. 119*(1), 189–221.

Sugawara, Y., T. Tanimoto, S. Miyagawa, M. Murakami, A. Tsuya, A. Tanaka, M. Kami, and H. Narimatsu (2017). Scientific misconduct and social media: Role of twitter in the stimulus triggered acquisition of pluripotency cells scandal. *J. Med. Internet Res. 19*(2), 1–10.

Sumner, P., S. Vivian-Griffiths, J. Boivin, A. Williams, L. Bott, R. Adams, C. A. Venetis, L. Whelan, B. Hughes, and C. D. Chambers (2016). Exaggerations and caveats in press releases and health-related science news. *PLoS One 11*(12), 1–15.

Sumner, P., S. Vivian-Griffiths, J. Boivin, A. Williams, C. A. Venetis, A. Davies, J. Ogden, L. Whelan, B. Hughes, B. Dalton, F. Boy, and C. D. Chambers (2014). The association between exaggeration in health related science news and academic press releases: Retrospective observational study. *BMJ 349*(December), 1–8.

Tahamtan, I., A. Safipour Afshar, and K. Ahamdzadeh (2016). Factors affecting number of citations: a comprehensive review of the literature. *Scientometrics 107*(3), 1195–1225.

Teixeira da Silva, J. A. and H. Bornemann-Cimenti (2017). Why do some retracted papers continue to be cited? *Scientometrics 110*(1), 365–370.

The Economist (2021). Zombie research haunts academic literature long after its supposed demise.

Tolbert, C. J. and R. E. Hero (1996). Race/ethnicity and direct democracy: An analysis of california's illegal immigration initiative. *The journal of Politics 58*(3), 806–818.

Weingart, P. (1998, dec). Science and the media. *Res. Policy 27*(8), 869–879.

Whitaker, R. and P. Lynch (2011). Explaining support for the uk independence party at the 2009 european parliament elections. *Journal of Elections, Public Opinion & Parties 21*(3), 359–379.

Whitely, W. P. (1994). The scientific community's response to evidence of fraudulent publication. The Robert Slutsky case. *JAMA J. Am. Med. Assoc. 272*(2), 170–173.

Wohlrabe, K. and C. Bürgi (2021). Do working papers increase journal citations? Evidence from the top 5 journals in economics. *Appl. Econ. Lett. 28*(17), 1531–1535.

Woo, S. J. and J. P. Walsh (2021). On the Shoulders of Fallen Giants What do references to retracted research tell us about citation behaviors ?

Worcester, R. M. (1980). Pollsters, the Press, and Political Polling in Britain. *Public Opin. Q. 44*(4), 548–566.

Yin, Y., Y. Dong, K. Wang, D. Wang, and B. F. Jones (2022, jul). Public use and public funding of science. *Nat. Hum. Behav.*.

Ziegler, L. (2021). What is the Media Impact of Research in Economics ?