The Evolving Landscape of Political Science: Two Decades of Scholarship in a Growing Discipline

Guy Grossman* William Dinneen† Carolina Torreblanca‡

March 5, 2025

Abstract

This study examines publication trends in political science over the past two decades (2003–2023), analyzing over 140,000 articles from 174 peer-reviewed journals. Using bibliometric methods and text-as-data innovations, the study investigates key aspects of scholarly output, including research volume, author productivity, topical focus, methodological approaches, and research design choices. We find that political science is a growing discipline primarily driven by an increasing number of contributing authors rather than individual productivity gains. The study documents a shift toward quantitative methods and the rise and decline of various research designs. Additionally, it explores the relationship between research specialization, topical novelty, and scholarly impact, revealing that novelty and focus in research are not associated with placement in top outlets but, conditional on publication, topically-focused and novel research is often better cited. The findings provide a comprehensive overview of the evolving landscape of political science scholarship, offering insights into future research avenues.

*David M. Knott Professor of Global Politics, Department of Political Science and PDRI-DevLab, University of Pennsylvania. Email: ggros@upenn.edu.

†PDRI-DevLab Predoctoral Fellow and Data Scientist, University of Pennsylvania. Email: wdinneen@sas.upenn.edu

‡PDRI-DevLab Postdoctoral Fellow, Department of Political Science, University of Pennsylvania. Email: catba@sas.upenn.edu.

Introduction

Scholarly publications largely form the foundation of a discipline's body of knowledge. By analyzing them, we gain valuable insights into how knowledge is generated and accumulated, who contributes to its creation, the methodologies employed in its production, and its substantive breadth. This paper explores how Political Science (PS) has changed over the past two decades, as reflected in journal articles.[1] To do so, we analyze over 140,000 papers published in 174 journals classified as "political science" outlets by Clarivate's Social Science Index.[2] We discuss our data and how we assembled it in Section 1.

Building on a growing Bibliometric body of work and recent innovations in text-as-data methods, our analysis of publication trends in the discipline centers on three key issues. First, we examine the discipline's growing volume of research (Section 2) and test some of its determinants, including collaboration patterns and researcher productivity (Section 3). Second, we explore what political scientists study (topical focus) and how they study those topics by classifying papers by their methodological and research design choices (Section 4). Third, we explore trends in research specialization, which we measure via papers' topical focus and topical novelty, and test the extent to which research specialization is rewarded in political science (Section 5). We conclude with a short discussion of how our study may pave the way for future research (Section 6).

[1]Our dataset does not include monographs and chapters in edited volumes because systematic information on these document types is not available in digital form.

[2]Clairvate classified 318 journals as Political Science (PS). We excluded journals with an impact factor of less than 1, journals that do not use peer review, and non-English journals. We include journals that are cross-listed as PS and International Relations (e.g., *World Politics*) but exclude journals classified exclusively as IR, such as *American Journal of International Law*. See SI Table B2 for a list of included and excluded journals.

1 Data

In this section, we describe the data we use to explore publishing trends in political science over two decades, from 2003 to 2023. We define "the discipline of political science" as the set of papers published in political science journals, irrespective of the author's status, disciplinary background, and institution. Figure 1 shows our data collection process, which we briefly describe below.[3]

Our starting point is the list of 188 journals with an impact factor of at least 1, which Clarivate classifies as political science.[4] We then search for each of these journals by name in Scopus, a comprehensive bibliographic database for academic research managed by Elsevier. We exclude three journals not indexed by Scopus, three that are not peer-reviewed, and eight that are not published in English, leaving us with 174 peer-reviewed, English-language political science journals. This number of journals is a marked improvement on previous reviews of the discipline, which generally only uses a limited sample of journals: for example, Fisher et al. (1998) base their trends analysis on three journals; Wæver (1998) uses seven journals, Kristensen (2012) uses 59 journals, Metz and Jäckle (2017) use 96 journals and Carammia (2022) bases their analysis on 100 journals.

Identifying these 174 journals on Scopus allowed us to download the journal's meta-data, including journal metrics (e.g., yearly citations) and, most importantly, a complete index of papers published in each journal. We collected paper-level data on 129,751 articles

[3] In SI Section B, we discuss the trade-off associated with defining the relevant corpus using either political science journals or, instead, authors (political scientists) and why we ultimately chose the former over the latter approach.

[4] Clarivate is an analytics company that provides tools for scientific research and academic performance evaluation via its *Web of Science* platform. We set a minimal impact factor to ensure we do not include very low-quality papers not representative of the discipline.

published in our sample of 174 journals between 2003 and 2023. These include information on each article's authors, title, abstract, publication date, and DOI link. We then gather the metadata for the 85,654 unique authors of these articles, including information on each author's yearly number of publications, annual citation count, and affiliation country. For these 85,654 authors, we also downloaded the basic metadata of an additional 807,288 papers published in non-political science journals, which we use in supplementary analyses. We enrich Scopus's metadata with measures of authors' gender, which we predict using the genderize.io package,[5] and summary measures of publication success such as h-index.[6]

To reliably classify trends in the topics studied and methods used in political science, we downloaded the full text of PS articles. Specifically, we successfully scraped the full paper text of 111,560 articles. Based on the first 1,000 words of each article, we classify paper topics using Structure Topic Modeling (STM). We classify the methods used by articles using a combination of Supervised Machine Learning and ChatGPT. We provide additional information on these classification exercises in Section 4 below.

2 Volume

We begin by analyzing trends in the volume of political science articles. Studying the volume of articles over time provides valuable insights into the discipline's evolution, the field's growth, and the process through which knowledge is produced. Mapping trends in volume further sheds light on the political science academic community's productivity

[5]The *genderize.io* package predicts binary sex based on the frequency of first names (and country when available). We were able to assign sex for about 97% of authors based on their name and country of origin with a mean posterior probability of 96.8%.

[6]For a more comprehensive discussion of Scopus journal, paper, and author metadata, see SI Section B.

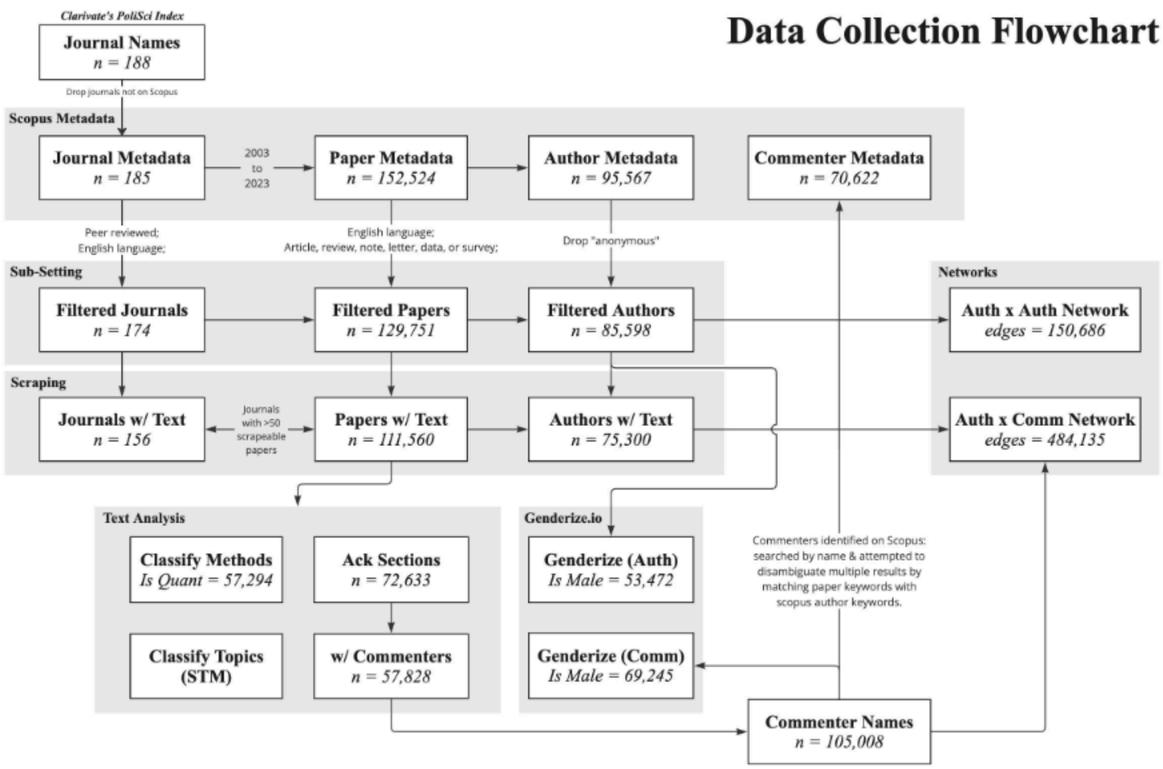

Figure 1: This Figure shows the data path from the initial list of journal names to the fully enriched data set. This article does not use data we retrieve from the acknowledgment sections, as those data are the basis of a companion paper.

expectations over time. It also lays out the foundation for studying factors affecting a scholar's production function (e.g., team size and composition and the importance of the author's resources proxied by factors such as institutional affiliation and seniority). We use our journal- and article-level datasets to explore trends in volume and productivity in political science in the past 21 years.

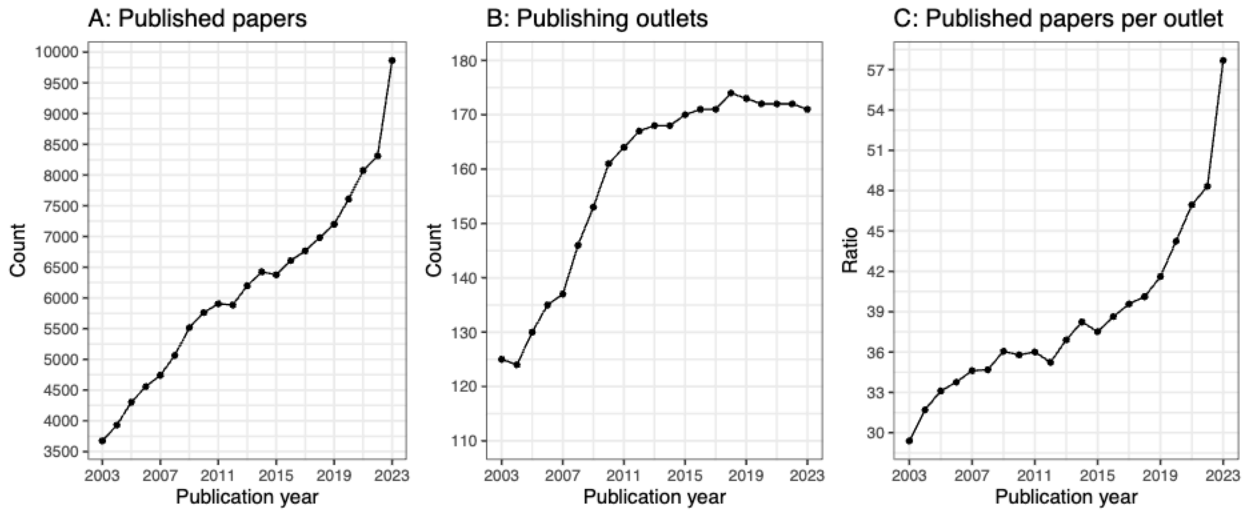

Figure 2: The left panel shows the count of papers published in all identified political science outlets each year. The middle panel plots the count of political science outlets publishing each year. The right panel plots and the ratio of published papers over publishing outlets.

Mimicking trends observed in many other academic disciplines (Wang and Barabási 2021, ch. 15), political science is a growing discipline. From about 3,500 articles published in political science journals in 2003 to almost 10,000 in 2023 alone (Figure 2, left panel). This dramatic growth comes with a growing number of outlets — from 125 political science journals in 2003 to over 170 publishing outlets in 2023 (Figure 2, mid-panel), but also with more papers per outlet — from about 30 articles a year per journal in 2003 to 58 articles a year per journal (Figure 2, right panel).[7]

[7]The growth in the number of published papers per journal is similar if we limit the sample to papers published in journals that existed in 2003. See Figure C1 in the Appendix.

2.1 Trends in co-authorship

The sheer growth in the number of publications is expected to influence the size of research teams: as the volume of new knowledge grows exponentially, the time a scientist can dedicate to absorbing new knowledge remains finite. This leads to specialization, and thus a need to put together bigger teams to tackle questions that span more than one topic.

Indeed, we find that co-authored research represents an increasingly large share of the output produced by political science, similar to a trend that has been documented in other disciplines (Wang and Barabási 2021, ch. 8). Figure 3 shows that the proportion of yearly published papers that were solo-authored (in red) has declined over time. In contrast, the proportion of co-authored articles, the green line (2 authors) and the blue line (3 +), has increased consistently over the period. Since 2021, most of the articles published in political science outlets have been co-authored (put differently, starting in 2021, the median number of authors increased from 1 to 2).[8]

Collaboration has not only become more frequent, but teams are also growing larger. Figure 3 shows that the proportion of papers published by teams of three or more members is increasing at a higher rate than two-person teams, narrowing the gap between the blue and green lines. Similarly, the right panel of Figure 4 also shows the trend toward team size expansion: while in 2003, the mean number of authors per paper published in PS journals (red line) was 1.4 in 2023, the mean number of authors had jumped to 2.

[8]Focusing on a smaller set of core journals, Metz and Jäckle (2017, p. 158) identify 2013 as the first year in which the majority of PS published papers were co-authored.

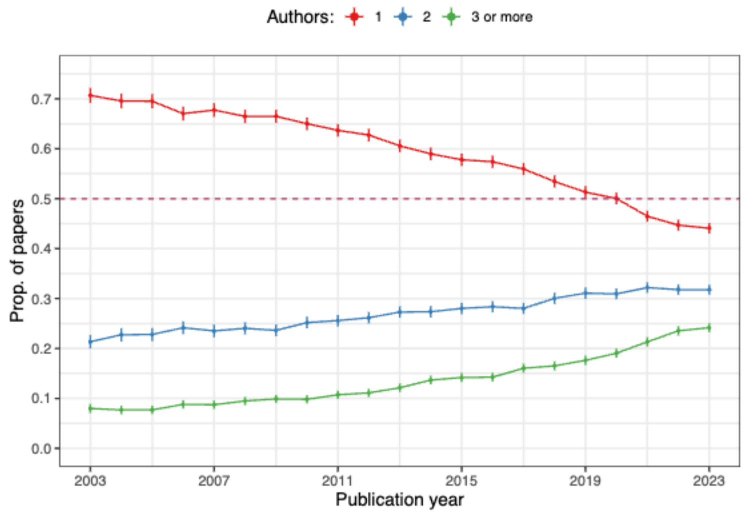

Figure 3: The Figure shows the proportion of yearly published papers in Political Science authors that had one author (red) two authors (blue) and three or more authors (green).

2.2 What explains the increase in volume?

The volume of political science increased almost three-fold between 2003 and 2023. What accounts for this rapid growth? A natural place to start is by assessing whether productivity changes can explain the volume increase. The left panel of Figure 4 shows the total count of unique authors (in red) and published papers (in blue) each year throughout the study period. The number of unique authors *and* the number of papers has increased drastically since 2003. While in 2003 only 4,358 unique authors published a paper in a political science journal, by 2023, that number had increased more than three-fold to 15,354.

Notably, the number of unique authors has increased *faster* than the number of unique papers, as seen in the widening gap between the blue and red lines in the left panel. Consequently, as the right panel shows (in gray), the yearly ratio of unique papers over unique authors, or papers per capita, is decreasing. On average, there were 0.84 papers per capita in 2003. By 2023, there were only 0.64 papers per capita. The decrease in the number of papers per capita results from the increasing popularity of co-authorship we documented above seen in Figure 3. Put simply, the time it takes to write a paper with (say) two other co-authors is larger than a third of the time it takes to write a single-authored paper.[9]

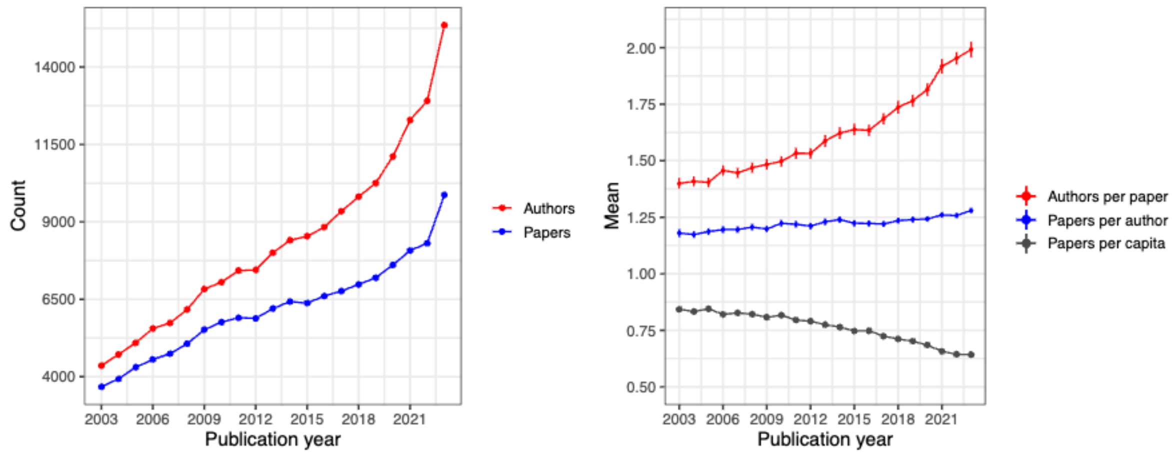

Figure 4: The left panel shows trends in the count of unique authors (in red) and published papers (in blue) in the study period. The right panel shows the average number of authors per published paper (in red) and the average number of published papers by each unique author in our sample (in blue) and the overall number of papers per unique authors (in gray). To calculate the average number of authors per paper, we computed the mean number of authors per published paper per year. To calculate the average number of papers per author, we first computed the total number of papers published by each individual author each year. Then we computed the mean of that number for each year.

While co-authorship might depress the number of papers per capita, individual researchers' productivity may still increase over time due to growing publication pressure, technological innovations (e.g., computing power, LLMs), and lower costs of conducting some forms of research. Figure 4 shows (in red) that while the number of authors per paper increased from 1.3 in 2003 to 2 in 2023 on average (54% increase), the average number of papers published by each unique author in our dataset (in blue) has increased from 1.18 papers

[9]Recent technological innovations, such as Overleaf, Zoom, Microsoft Teams, Open Science Framework and Slack, may make collaborative work more efficient over time.

in 2003 to 1.28 papers in 2023 (8.5% increase). Thus, author-level productivity, marginalized over co-authored and single-authored work, increased by only 8.5%. Back-of-the-envelope calculations suggest that even in the absence of this increase in productivity if the number of 2023 authors each produced 1.18 papers per year (the 2003 average), and assuming that the modal 2023 paper has two authors – 2023 would have seen more than 9,000 papers published. Consequently, the pace of the modest increase in productivity accounts for only a small part of the PS discipline's corpus expansion we documented above.

Slight increases in political scientists' productivity are insufficient to explain the rapid growth of the discipline (5% annualized over 21 years). Instead, the expansion of the discipline might be explained by the growth of political science-writing researchers: More political science is a product of more political scientists. The left panel of Figure 5 plots the count of new authors publishing in political science each year from 2003 to 2014. The colors of the bars distinguish whether researchers will publish again in political science within the next 10 years (green), whether they will publish again but not in PS journals within the next 10 years (blue), or whether they will only publish once within that period (red). The right panel shows the proportion of each group by year.

As shown in Figure 5, every year (post-2003), there are more than a thousand new Political Science authors, but a little under two-thirds of them will never publish again in political science outlets within 10 years. Of those, we note the reversal in the ratio of 'single hitters' (red) compared to those who have published only once in political science outlets but published again elsewhere (in blue). That the latter group is mostly comprised researchers working in related fields like Economics, Psychology, Communication, and Sociology suggests the growing connection and relevance of political science to other disciplines. In sum, our analysis suggests that more authors, instead of more productive authors, best explain the bulk of the increase in political science volume. This conclusion is consistent with findings from other disciplines (Wang and Barabási 2021, p. 173).

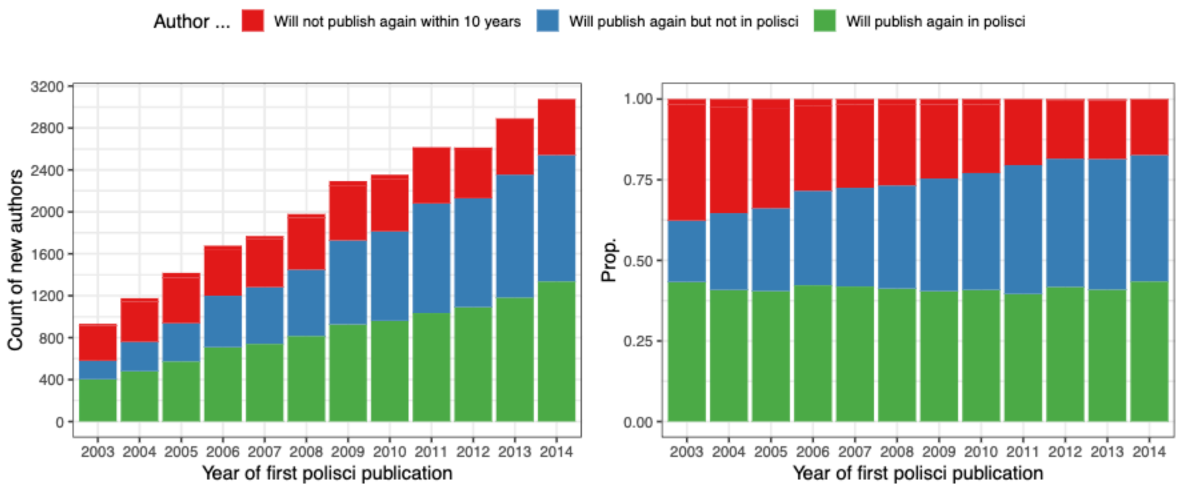

Figure 5: The left panel of the Figure shows the count of new authors, defined as authors who published their first in-sample political science publication each year, who will not publish again in any discipline within ten years (red), will publish again within 10 years but not in a political science journal (in blue) and those who will publish again within that same interval in an in-sample political science journal (in green). The right panel shows the corresponding yearly proportions of each of the three groups.

3 A deeper dive into researcher productivity

Above we concluded that changes in productivity are unlikely to explain why political science output has increased in the past two decades. However, we have also shown that around 60% of debutant political science authors will not publish in the discipline again. Could the productivity of individual political scientists be increasing but be obfuscated by this growing number of "single hitters," i.e., researchers who only publish once in the discipline? Alternatively, could it be that younger cohorts of political scientists increasingly publish outside of the discipline, for example, in high-impact multidisciplinary general interest journals such as Science, Nature, and PNAS, or journals of other related disciplines, such as Economics, Sociology, Psychology, and Communication? While the latter phenomenon, if true, cannot explain the increase in the volume of published political science, it would have important implications for political science as a discipline.

To test whether political scientists have become more productive, within or outside political science, once we remove "single hitters" first, we classify "political scientists" as individuals who published at least five articles in any journal over ten years *and* who published at least half of their papers after ten years in a political science journal. Next, we compare two cohorts of political scientists, thus defined, over ten years: those who published their first article, not necessarily in political science, in 2004 and those who published it in 2014. With this sample, we estimate the mean number of published papers per author by fitting the following model:

$$\text{Publications}_{ic} = \sum_{y=1}^{10} \beta_y \mathbb{I}[\text{Year since first publication}_i = y] + \epsilon_{ic} \tag{1}$$

Here, the β_y's are estimators of the mean number of yearly published papers by authors from cohort c each year y after their first publication, and ϵ_i are robust errors clustered at the author-level. Figure 6 shows the results for two outcomes of interest. The left panel shows the estimated mean number of yearly published papers in political science outlets for political science researchers, as defined above. The right panel shows the mean number of yearly published papers in all outlets, regardless of the field, for the same sample of researchers. Figure 6 underscores two interesting dynamics. First, there does not seem to be much difference in productivity between the 2004 (red) and 2014 (blue) cohorts when it comes to publishing political science journals, at least during the first seven years since the authors' first publication. After year one and until year seven, both cohorts publish at similar (increasing) rates. After year seven, however, the younger cohort publishes at a slightly higher rate, around .95 papers per year, while the older cohort publishes around .65 papers per year (Figure 6, left panel).

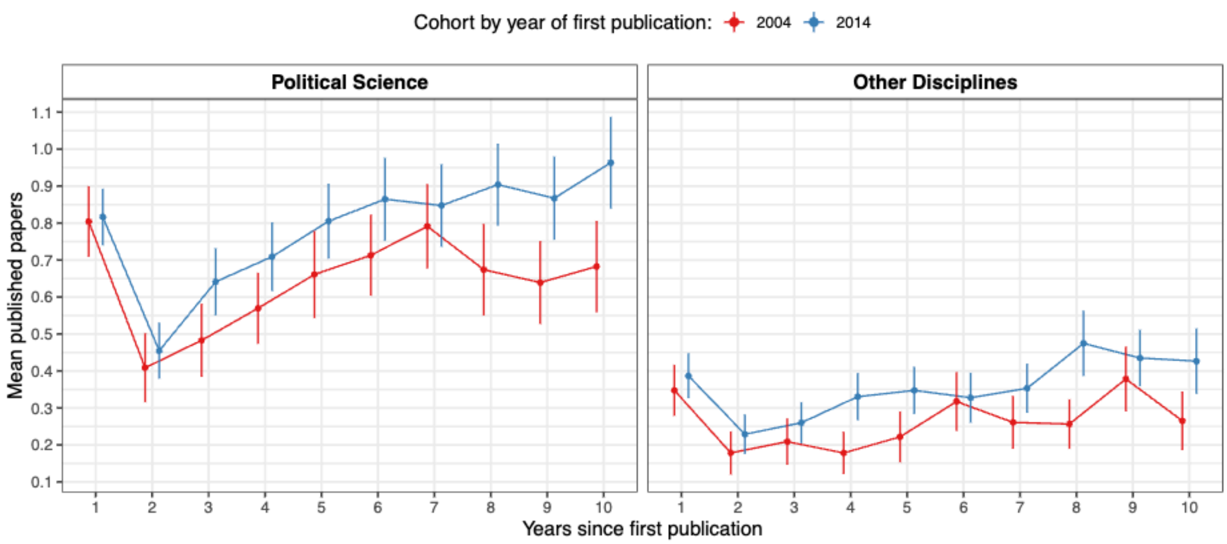

Figure 6: The left panel shows the estimated number of papers published in political science outlets during the first ten years of researchers' careers, for researchers who first published a paper in 2004 (red) or 2014 (blue), and published five papers or more, at least half of which were published in political science outlets. The right panel shows the estimated number of papers published in the first ten years of PS researchers' careers in non-PS outlets for PS researchers as defined above.

Second, the right panel shows similar productivity trends for publications in other fields. Although the 2014 cohort slightly outproduces their 2004 cohort in the mean number of publications in other fields throughout, the gap is small and inconsistent. By the 10th year since their first publication, the 2014 cohort was publishing .45 papers in non-political science outlets per year, on average, while the 2004 cohort was publishing only .25 papers.

In sum, the growth in the volume of published work in political science journals is largely due to an increase in the number of unique authors publishing in PS outlets. At the same time, we witness a slight increase in the productivity of younger cohorts of political scientists, who publish at slightly higher rates both in political science outlets and other disciplines, especially as their careers advance.

4 Content and Methods

A discipline is characterized by the issues it studies and the methods used to examine those issues. Interestingly, to the best of our knowledge, none of the earlier reviews of political science trends have reported trends in content and methods. This section fills in this gap.

4.1 Topics

We extract the topics of political science studies by applying structural topic modeling to our corpus of full-text political science articles ($n = 111,560$). First, we follow the method proposed by Gerlach, Shi and Amaral (2019) to pre-process the text and remove uninformative words and stopwords, followed by the manual removal of high-frequent terms that are not associated with any specific topic (such as *article*, *publication*, *publish*, *work*, and *political*). Second, we use the STM *R* package (Roberts, Stewart and Tingley 2019), estimating the model with year and journal as prevalence covariates, and experimented with the number of topics. We found that 30 topics were optimal in that it allowed each topic to be unique and recognizable without redundancies (which appeared when we increased the number of topics) and without omissions (which occurred when we reduced the number of topics). Third, we manually labeled topics based on the words with the highest probability.

We note that instead of labeling each article with a unique topic, STM outputs a posterior proportion, θ_{it}, that paper i is allocated to topic t, for each of the thirty topics. On average, however, only 2-3 topics result in a posterior proportion higher than 10%.[10] We further note our analysis below focuses on *between-topic* frequency, yet the substantive interest might change over time within a given topic. For example, twenty years ago, the political communication literature focused extensively on media bias, but in recent years, the field has concentrated on misinformation. Similarly, 20 years ago, the "Democracy and Autocracy" topic focused on hybrid regimes, and 10 years ago, it strongly emphasized autocratic resilience. However, in recent years, scholars writing on this topic have shifted attention to questions surrounding democratic erosion. Such within-topic refocus is not captured in our data. Bracketing these caveats, in Figure 7, we plot the mean topic proportion and 95% confidence intervals for all papers published in the top-20 political science outlets each year (blue) and all other outlets (red).[11] Several trends are worth noting.

First, some topics have become more prevalent in recent years (in particular, Political Communication, Indigenous politics, Social Movements, and Race and Immigration). Other topics, like Terrorism, War, the US Presidency, Lobbying, and Normative Theory, have lost popularity. Last, some topics seem ever-present, like Democracy and Autocracy, Identity politics, Federalism and Decentralization, and Democratization.

Second, there is a difference between the intensity with which topics are covered in the top 20 PS outlets and the rest of the discipline's journals. Some topics are consistently over-represented in Top-20 outlets compared to the broader discipline: for example, Civil Wars and Intergroup Conflict, Public Opinion, Quantitative Methods, and Fiscal politics. Other topics — European politics, post-Soviet politics, Critical Theory, and Terrorism —

[10] Overall, 74.1% of the papers have at most three topics with a proportion higher than 10%. See Tables E4 and E5 in the Appendix for details.

[11] We identify the top 20 outlets by calculating each journal's SJR impact factor, averaged across the entire period, and selecting the 20 journals with the highest mean impact. Table B1 in the Appendix lists the 20 journals in the group.

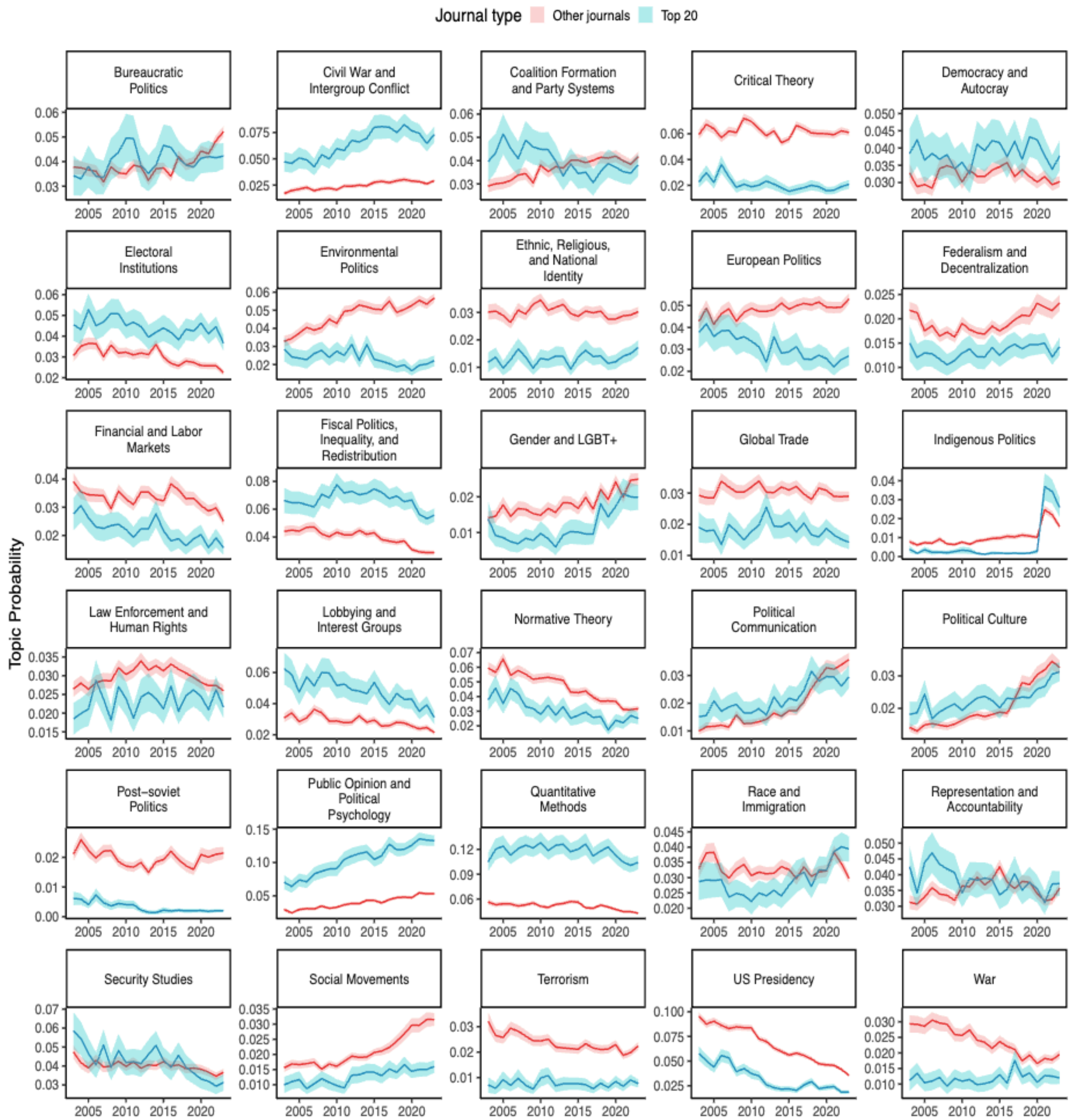

Figure 7: Figure shows the mean topic proportion and 95% confidence intervals for all papers published in the top 20 political science outlets each year (blue) and all other outlets (red). Topics are estimated using Structure Topic Modeling and were labeled manually.

are consistently under-represented in Top-20 journals compared to the broader discipline.

Third, trends in topic popularity tend to track in the same direction in the top-20 journals and the rest of the discipline (e.g., the decline in research on the US Presidency and Financial and Labor markets or the increase in Public Opinion scholarship and research applying a political culture lens). However, this is not always the case. For example, Environmental politics has grown tremendously in the past two decades, but Environmental politics scholars have struggled to make inroads into political science's most prestigious journals.

4.2 Methods

We are interested not only in what political science studies but also in how. We thus classified each paper by its method. After cleaning the full text and removing stop-words, we leverage a sample of papers manually coded by Teele and Thelen (2017) that overlaps with our sample to train our classification model. We begin with three labels: quantitative, qualitative/normative, and formal. In stage one, we train a TF-IDF + Logistic Regression model to identify key features (i.e., words), then proceed to a second stage using only the top 50 features to reduce noise.[12] Next, to distinguish between qualitative/normative papers, we utilize ChatGPT along with a custom prompt. [13] Measuring these results against Teele and Thelen and additional papers from the TRIP Journal Article Database (*TRIP Journal Article Database Release (Version 3.3).* 2020; Maliniak and Tierney 2018), we achieve an F1

[12]We achieve a mean macro F1 score of 82% (+/- 0.06) across a cross-fold validation test. For example, here are some representative words and their corresponding classification coefficients for each category: Quantitative — equilibrium (-3.9); variables (2.9); results (2.8). Formal — equilibrium (5.6); model (2.5); game (1.9). Qualitative/Normative — model (-2.7); variable (-2.2); justice (1.8). A full classification report and a list of coefficients can be seen in the appendix under Tables F6 and F7.

[13]The prompt text and performance is available in the Appendix under Prompt 1 and Table F8.

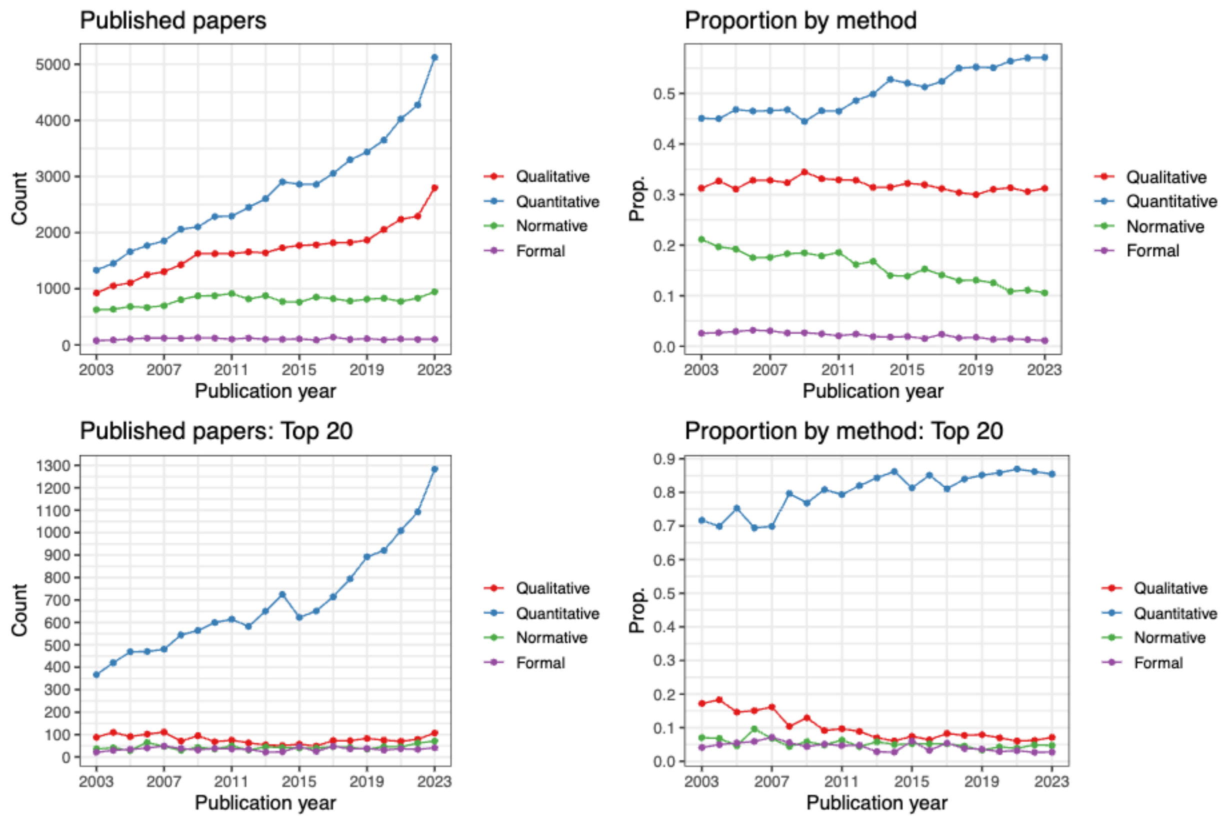

Figure 8: The top-left panel shows the number of papers published in all political science outlets each year, classified as qualitative (in red), quantitative (in blue), normative (in blue), and formal (in purple). The top-right panel shows the proportion of all published papers by method over all published papers that year. The bottom row shows the same analyses but only papers published in the top 20 PS outlets per their SJR ranking.

score of 88%. With the model and prompt together, we classified the papers in our sample into four categories: quantitative, qualitative, normative, and formal.

We highlight four notable trends. First, the share of political science papers that rely on quantitative analysis has increased substantially over the past two decades. In 2003, only 45% of papers published in political science used quantitative methods; by 2023, the share had grown to 57% of published papers. This trend is accentuated in Top-20 outlets, where papers using quantitative methods rose from 70% in 2003 to 85% in 2023. Second, while the number of qualitative papers has increased, the proportion of political science papers that rely exclusively on qualitative methods has remained hovering around 30% throughout the study period. However, the share of exclusively qualitative papers has decreased significantly over the past two decades (from 15% in 2003 to 7% in 2023) in the top journals in the discipline.[14] Third, while the yearly output of normative theory papers has constantly hovered around 1,000 throughout the period, the share of normative papers has decreased substantially (from 17% of papers in 2003 to 10% in 2023). Last, the share of papers published in political science journals that use formal theory is small (around 2%) and unchanging whether we focus on the entire corpus of journals or when zooming into the Top 20 journals.

4.3 Research Design Trends

We conclude this section with a deeper dive into trends in research design. The last two decades witnessed a methodological shift in political science research. In particular, the "credibility revolution" refocused quantitative research towards designed-centered approaches (Blair, Coppock and Humphreys 2023). Qualitative scholars, too, have increasingly centered causal processes in their investigations by adopting and refining methods such as process tracing (Collier 2011) and counterfactual reasoning within comparative case studies frameworks (Mahoney 2004). This section uses our massive data to give a data-driven overview of trends in an important facet of PS as a discipline: its research design.

We primarily focus on quantitative causal inference since our classification approach, described below, performed significantly better for quantitative than qualitative methods. This choice is, in part, due to the fact that qualitative scholars do not always use consis-

[14]Our classification scheme classifies mixed methods papers as quantitative. As such, the share of papers using qualitative methods, such as archival research, in-person interviews, focus group discussions and participatory observation, is higher than 30%.

tent labeling for the same method.[15] In addition, quantitative causal inference methods are significantly more likely to be accompanied by auxiliary terms — usually referencing identification assumptions — that help reduce false negative coding.[16] We thus chose to report trends in qualitative methods only in the online Appendix with some caveats.

To classify the research designs used in each paper, we use a two-stage procedure. Our process begins with a rule-based keyword detection system, followed by GPT-based refinement to to enhance precision. First, we construct a dictionary of methodological keywords, distinguishing between "main" terms (e.g., "instrumental variable" or "Difference-in-differences") and "sub" terms reflecting implementation details (e.g., "exclusion restriction" or "parallel trends"). If at least one "main" *and* at least one "sub" keyword appear in a paper, we tentatively mark that paper as using that research design and extract a 500-character context window around each keyword match for later validation with GPT.[17] Finally, to avoid matches based on incidental mentions of keywords, we use GPT alongside a custom prompt,

[15]Consider, for example, process tracing, which our text analysis suggests has risen significantly in popularity starting around 2009. Process tracing was labeled as a method in political science in around 2008 and later systematized in landmark studies (e.g., Collier 2011) and textbooks (e.g., Beach and Pedersen 2016). Of course, scholars were using the method without calling it that before 2009, mostly in the guise of single- or comparative case study designs. Our classification approach cannot separate between the growing popularity of process tracing versus more consistent labeling of the method.

[16]For example, instrumental variable (IV) methods will (almost) always be accompanied by the terms "first stage" and "exclusion restriction." Similarly, regression discontinuity designs (RDD) will appear along terms such as "bandwidth." "forcing variable," and "threshold." Such consistency in auxiliary terms is not always found in qualitative methods.

[17]To address overlapping classifications between field experiments and survey experiments, which share important keywords (e.g. "random assignment"), we implement conflict rules that prioritize survey experiment specific keywords (e.g. "conjoint experiment", or "attention check"). Papers matching both field and survey experiment keywords are assigned to the latter category based on methodological precedence. The performance of this approached, tested against manual RA coding, can be seen in Appendix Table F10.

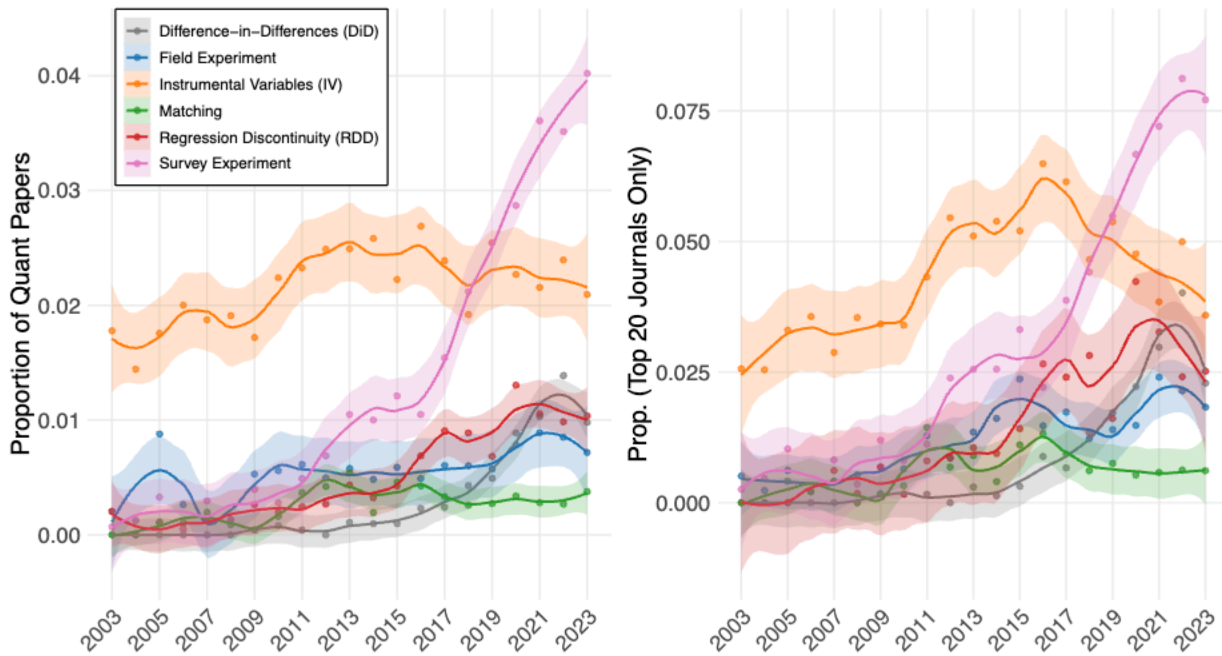

Figure 9: Figure shows the proportion of quantitative papers coded as using each research design over time. The left panel shows this value for all quantitative papers in our sample. The right panel shows results for only the top 20 journals in our sample.

providing it with the 500-character context windows to evaluate whether keywords indicate actual method use. Overall, 79% of the keyword-based labels—which we report as the final labels—were confirmed by GPT. [18]

We report our findings in Figure 9 for quantitative papers published in any political science journal in the study period (left panel) and the top 20 outlets (right panel). We highlight three key findings. First, there has been a general increase in the adoption of a host of credible research designs starting around 2011-12, with trends being more pronounced in the top 20 journals. Notwithstanding this general trend, some research design methods, such as matching and instrumental variable approaches, go "out of fashion" as the PS community becomes more aware of these methods' limitations (more on this below).

[18]When we tried a similar approach on qualitative methods, only 45% of keyword-based labels were confirmed by GPT.

Second, we wish to focus our readers not only on trends but also on levels: Except survey experiments, which have truly taken off since 2011 and appear in about 8% of published quantitative papers in the top 20 outlets by 2023, causal inference designs are still rather rare. Difference-in-differences, two-way fixed effects and event study designs (all under DiD), Regression Discontinuity Designs (RDD), and matching estimators each appear in about 1% of quantitative PS articles. Field experiments — arguably due to their high costs — are used in less than 0.5% of published quantitative papers in our sample. The relative rarity of credible causal inference designs speaks to the persistence of "selection on observables" regressions, which, if we assume make up the rest of our quantitative sample, are still used in about 75-80% of quantitative PS papers.

Third, our data allows us to identify the rise (and fall) of specific research designs. Consider trends in instrumental variables (IVs) and survey experiments. Between 2003 and 2015, IV estimators were both the most popular and the fastest-growing causal inference method in quantitative political science: in 2003, 2.6% of articles published in top 20 outlets used an IV estimator; in 2015 that share was close to 7%.[19] Yet, the use of IVs has plummeted in the past decade —almost returning to 2003 levels— as political scientists became increasingly aware of the rarity in which the method's core identification assumptions hold.[20]

In contrast to IVs, survey experiments have consistently increased in popularity and

[19]The rise in the use of IVs post-2003 likely owes to the broad impact of several seminal papers using IVs, such as Acemoglu, Johnson and Robinson (2001) and Miguel, Satyanath and Sergenti (2004).

[20]Influential critiques include Sovey and Green (2011) that focused on exclusion restriction violations, and Lal et al. (2024) that demonstrated how researchers often overestimate the strength of their instruments due to non-i.i.d. error structures. IVs, however, remain a relatively popular research design (in 2023, there were still twice as many IV studies than RDDs), in part because of the growing popularity of shift-share (Bartik) instruments in trade and especially migration studies (e.g., Dipoppa 2024).

account now for about 8% of quantitative papers published in the top 20 PS outlets. This explosion in popularity is likely due to a confluence of factors beyond the credibility revolution itself. These include the entry of new survey firms that provided relatively inexpensive access to online opt-in samples (such as Bovitz, Lucid, Dynata), which dramatically reduced the cost of conducting survey experiments,[21] and the broad appeal of several methodological and substantive seminal studies. We note in particular, Hainmueller, Hopkins and Yamamoto (2014), which introduced Conjoint experiments to PS, Coppock (2019), which showed that survey experiments often generalize to other populations, and Tomz and Weeks (2013), which popularized the use of survey experiments among (a subset of) IR scholars. Finally, the growing competition over a shrinking PS tenure track market,[22] (likely) incentivizes scholars at the PhD and postdoc stages to run survey experiments, which often results in faster time-to-publication.

5 Novelty and Focus in Political Science Research

Political science research has expanded significantly, with rising collaboration and shifting trends in topic popularity. However, more research does not necessarily mean better, more insightful, or more impactful work. This section complements section 4.1 by examining the content of political science research through two key dimensions: topical novelty and topical focus. Following Heiberger, Galvez and McFarland (2021), we assess *topical focus* as a function of whether papers concentrate on a few topics or span many, and *topical novelty*

[21]Relatedly, we note the importance of several influential studies, which gave PS scholars the green light to use inexpensive M-Turkers as experimental subjects without, arguably, compromising quality (Berinsky, Huber and Lenz 2012), nor generalizability (Coppock 2019).

[22]Data assembled by McGrath and Diaz (2023) based on APSA jobs posting suggest a 27% decline in the number of tenure track PS jobs between 2012 and 2022.

as a function of the extent to which papers focus on frequently paired topics or introduce rare combinations.

There are good reasons to examine topical novelty when taking stock of an academic discipline. Knowledge advances both through piecemeal, cumulative studies, and via breakthrough. Novel approaches commonly drive scientific breakthroughs. Novel research contributes to the evolution of a field by introducing fresh perspectives, challenging established theories, and addressing contemporary problems with innovative solutions. Novelty, however, is also risky, and may take longer to be recognized by peers (Wang, Veugelers and Stephan 2017). Meanwhile, the importance of topical focus for "better science" is less clear-cut. A narrow topical focus allows for a deep, rigorous investigation of a specific topic. From an individual researcher's perspective, topic specialization enables scholars to assert expertise and (may) increases their credibility and influence within the academic community. While highly specialized studies contribute deeply to a (niche) area, they may have limited applicability beyond their specific domain. Indeed, broader studies can bridge sub-fields, facilitate innovation, and adapt faster to emerging trends in their field. Understanding the trends and differences in novelty and topical focus over a lengthy period sheds light on the nature of extant political science research.

First, we explain how we construct the measures of paper-level topical focus and novelty and provide basic summary statistics. Next, we explore whether collaboration with other researchers systematically relates to topical focus and novelty. Last, we study how these two paper-level characteristics correlate with paper-level success, measured by paper journal placement and the number of citations (standardized within year).

5.1 Measuring Novelty and Focus

To construct paper-level measures of topical focus and novelty, we adapt and refine measures proposed by Heiberger, Galvez and McFarland (2021). For both indices, we use the results from the STM model, detailed in Section 4. Recall that the STM model assumes that each paper is a mixture of different topics. The STM outputs the posterior proportion of paper i allocated to topic t, θ_{it}, that paper i includes topic t, for each of the 30 topics described in Section 4. We define topical focus as a Herfindahl index of topic focus for each of the papers in our sample. It is computed by adding the squared topic proportion θ_{it} of paper i including topic t for all 30 topics; $Focus = \sum_{t=1}^{T} \theta_{it}^2$. The index ranges from 0 to 1, and its interpretation is straightforward; the closer the index is to one, the more topically focused a paper is. The closer it reaches zero, the more topically diverse the paper is.

Our novelty index measures how rare or common a topic combination in a given paper is relative to the topics in political science papers published during the preceding three years. It is constructed by first assigning to each paper the two highest proportion topics, as estimated by the STM model.[23] Next, we compute the proportion of all papers written over the preceding three years with either topic as one of their two most probable. Using this proportion, we can calculate the *expected* number of papers written on each two-topic combination throughout the period. The intuition is straightforward: when a topic is written about frequently, we should expect it to co-occur with other frequent topics more often by chance alone. Last, we calculate the actual observed share of papers that combine each dyad of topics and take the ratio of observed over expected papers with that specific

[23]We exclude 7,498 papers for which the second most probable topic results in an estimated posterior probability lower than .1, as we consider such papers single-topic focused. See Section G.1 in the Appendix for details on the index construction and the sample. Figure E2 in the Appendix shows topic-level trends for each paper's second most likely topic.

topical combination r. We then assign a novelty score $N = 1 - r$ to each paper in our sample. A zero value in the novelty score indicates we have as many papers on that combination of topics as expected. A score close to one, the theoretical maximum, indicates a paper that combines topics in a perfectly novel way. Conversely, papers with negative scores combine topics often used together during the past three years more than expected and are hence not novel combinations. Table 1 provides basic summary statistics for both outcomes.

Index	Mean (SD)	Min, Max	Median	N
Topical Focus	0.267 (0.134)	0.068, 0.997	0.230	111,560
Topical Novelty	0.081 (0.626)	-4.11, 0.985	0.232	98,000

Table 1: The table shows summary statistics for the topical focus and novelty indices. We calculate topical novelty using a rolling three-year count of all articles, for which the STM estimates a posterior probability of the second most likely topic being more than 1. Consequently, we calculate the novelty index for a subset of papers published after 2004, while the focus index is calculated for the entire sample of papers included in the STM.

5.2 Collaboration, Topical Diversity, and Novelty

Does team size and composition have a systematic relationship with political science's topical novelty and focus? The answer has important implications for the increasingly collaborative discipline. However, whether collaboration should foster or depress novel or topically diverse research is *a priori* theoretically unclear. On the one hand, research conducted in teams can rely on members' distinct substantive and methodological expertise that translates into more topical diversity or a larger thematic scope when conducting research. Additionally, approaching a research topic from diverse backgrounds could result in more novel research. On the other hand, collaboration could be more frequent between researchers with similar research interests, substantive and methodological approaches. Large teams can also be riskaverse, prioritizing well-trodden topics. Such collaboration will unlikely lead to more novel or diverse research relative to single-authored work.

The relationship between publication success, team size, and composition has been studied extensively in the natural sciences (Freeman and Huang 2014). However, it has received insufficient attention in political science.[24] Extant evidence points to a positive return to diversity in scientific teams (see, for example, AlShebli, Rahwan and Woon (2018), Freeman and Huang (2014), and Powell (2018)). According to Wang and Barabási (2021), diversity within a scientific team promotes the team's effectiveness by enhancing productivity, resulting in works with higher impact or both (p. 114).

If collaboration facilitates more diverse research, we expect to observe that, on average, co-authored work would be more novel and less topically focused than single-authored work. We test this expectation and report our findings in Figure 10. The left panel shows the mean yearly topical focus index for co-authored (in red) political science research and single-authored (in blue) work. Recall that higher index values represent more topically focused research, while smaller values represent more topically diverse papers. Overall, political science papers have become slightly *more* topically focused in recent years, especially starting in 2013. Further, co-authored papers are slightly more topically focused than single-authored work. While an average paper in 2003 had a diversity score of .25, by 2023, the average score for single-authored papers was .275 and .285 for co-authored work. However, the magnitude of the differences is small: a 10% increase for single-authored papers and a 14% increase for co-authored work.

Conversely, the right panel shows that topical novelty in political science has consistently increased in the past decades, especially for co-authored work. While research

[24]Though see Teele and Thelen (2017) for an important discussion on gender diversity in political science teams.

published by teams in 2005 was as novel as research published by solo authors, by 2023, the novelty index for teams was, on average, 42% larger. These findings are consistent with results from Heiberger, Galvez and McFarland (2021), who find that dissertation topic novelty, but not topical focus, are positively associated with the likelihood graduating sociology PhD s become advisors themselves.

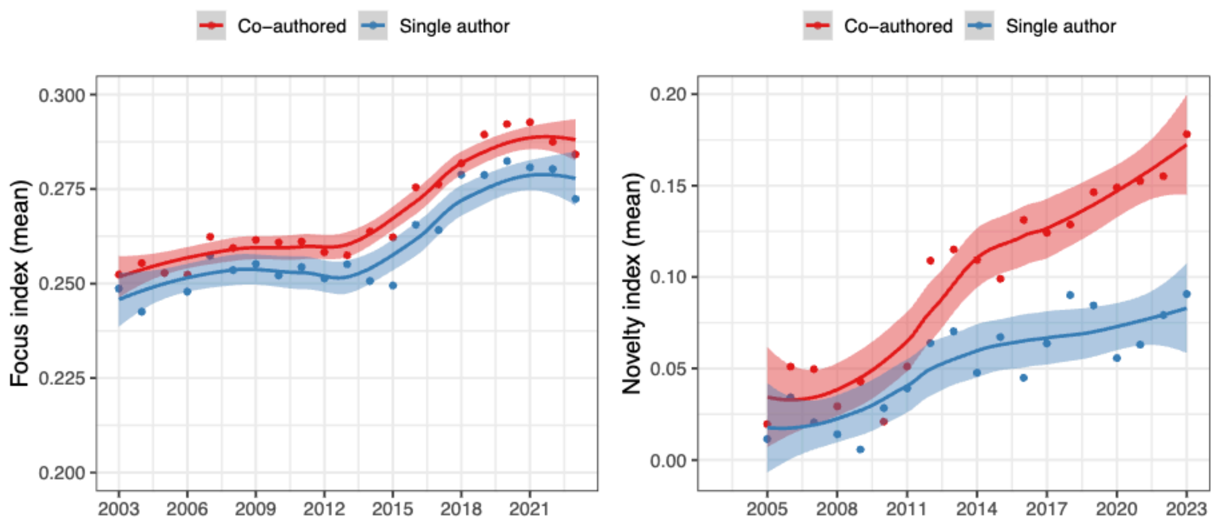

Figure 10: The left panel shows the yearly mean diversity index for co-authored (in red) and single-authored (in blue) papers. The right panel shows the yearly mean novelty index for co-authored (in red) and single-authored (in blue) papers. Gray bands mark the 95% confidence intervals.

The analyses so far have underscored trends in topical novelty and focus as well as systematic associations between these measures and formal collaboration in political science research. However, research teams and research projects form endogenously, making it tenuous to ascertain from such raw associations whether team composition and size cause changes in topical novelty or focus or whether other factors shape the topical breadth of political science research and simultaneously influence the size and composition of teams. To assuage some concerns about the interpretability of the results, given the endogenous process through which our data is produced, we leverage the richness of our data to compare how topical focus and novelty relate to co-authorship *after parsing out all time-invariant author-level characteristics.* To do so, we construct a dataset where the paper-author is the unit of observation and compare how topical novelty and focus are related to co-authorship by comparing solo-authored and co-authored papers written by the same authors. We thus estimate the following model:

$$\text{Index}_{ia} = \beta_1 \text{CoAuthored}_{ia} + \gamma_a + \varepsilon_{ia} \tag{2}$$

where Index_{ia} is the standardized index of either topical novelty or focus, for paper i, published by author a, CoAuthored_{ia} is either a binary variable that takes the value of 1 if paper i written by author a was co-authored and zero otherwise, or the log number of co-authors in paper i, including author a. γ_a are author-fixed effects, and ϵ_{ia} are robust standard errors, double-clustered at the paper and author levels. For comparability, we subset the sample and retained only papers for which we could estimate both indices.

We report our estimates in Table 2. When looking at the extensive margin with the binary measure and the intensive margins with the continuous measure, we find that collaboration results in more novel published research after accounting for all-time invariant author-level characteristics. When we include author-fixed-effects, a co-authored paper is, on average, 3.8% of a standard deviation more novel than a single-authored work. Alternatively, going from a single-authored paper to a paper with 2 co-authors, increases the novelty index by an average of $(log(2) - log(1)) \times .030 \approx 0.02$, or 2% of a standard deviation. Conversely, after accounting for individual characteristics, co-authorship has no statistically significant association with paper focus and the association is very precisely estimated at zero.

	Novelty (1)	Novelty (2)	Focus (3)	Focus (4)
Co-author	0.038***		0.000	
	(0.008)		(0.001)	
log(Authors)		0.030***		−0.001
		(0.009)		(0.001)
Author FE	✓	✓	✓	✓
Num.Obs.	166 387	166 387	166 387	166 387
R2	0.528	0.528	0.560	0.560

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: The table reports the estimated difference in the standardized novelty index and the focus index for co-authored papers relative to single-authored papers and the expected marginal change in both indices when the count of authors in a paper increases by one unit in the logarithmic scale (or around 2.7 more authors). Dependent variables are standardized. The unit of observation is the author-paper. Robust standard errors, double clustered at the author and paper, are reported in parentheses.

In the above analyses, co-authorship teams are a blunt category that captures greater heterogeneity in authors' research interests, and methodological approaches relative to single-authored work. To the degree that co-authorship teams result in more diverse backgrounds, teamwork is associated with topically more novel work. However, we know that *team composition* varies greatly: some co-authorship teams are made up of researchers with similar backgrounds, while other teams might include researchers of different genders, seniority, and substantive focus. In Figure 11, we explore one dimension of diversity: gender. Specifically, following Yang et al. (2022), we probe whether diversity in the gender of team members is related to novelty and topical focus by examining how these indices differ with the gender composition of co-authorship teams. As a benchmark, we explore how gender in single-authored papers relates to the same measures.

In Figure 11 (left panel), we subset the sample to co-authored papers, analyzing only papers where we could construct both topical focus and topical diversity indices. Both indices are standardized for comparison ease. The first striking result is that the gender composition of teams is associated with no significant difference in the topical focus of published papers, consistent with results from Table 2. However, when it comes to novelty, all-female teams publish the most topically novel papers, followed by mixed-gender teams and all-male teams. Comparing with the right panel, we can see that overall, single authors, regardless of their gender, publish papers that are no more nor less topically focused than what is produced by teams. However, solo male researchers publish the least topically novel work.

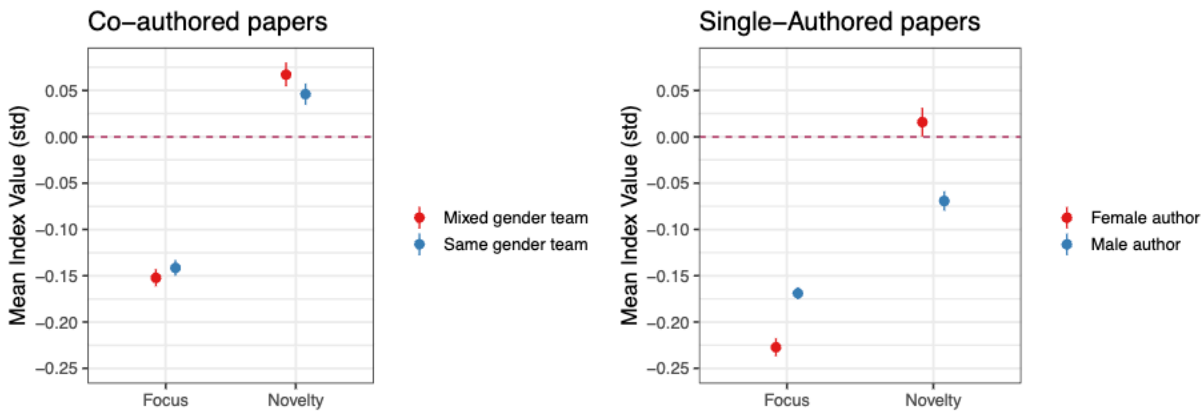

Figure 11: The left panel plots the mean diversity and novelty indices and 95% confidence intervals for co-authored papers written by all-female teams (in red), all-male teams (in blue), and teams of mixed gender (in green) between 2005 and 20023.

Results from Figure 11 are in tension with recent findings showing that gender-diverse teams produce more novel research (Yang et al. 2022). Indeed, in our data too, mixed-gender teams produce more novel research than single-gender teams: when pooling (as Yang et al. (2022) do) all-female teams with all-male teams, the average novelty score is lower than that of mixed-gender teams. This result, however, obfuscates the fact that all-female teams produces, on average, the most novel research in the discipline whereas all-male teams produce, on average, the least novel output.[25]

[25]Figure G.3 plots the proportion of papers written by teams, according to the teams'

5.3 Novelty, Diversity, and Publication Success

Results thus far suggest co-authorship generally leads to more novel research but has no statistical association with papers' topical focus. In this section, we explore the consequences of topical novelty and focus as they pertain to publication-level success.

Studies examining the association between the topical novelty of research and its success appear to be mixed. On the one hand, previous work has found that newness in research is rewarded with recognition in the scientific community (Heiberger, Galvez and McFarland 2021; Antons, Joshi and Salge 2019). Other studies found that producing novel research might be "high risk /high reward" whereby highly novel papers exhibit a larger variance in success, at least as measured by citations (Wang, Veugelers and Stephan 2017). There is also a temporal dimension to novelty success: novel papers are more likely to be among the top cited in the long run, *but* less likely to be cited in the short run.

The consequences of topical focus for research success are also a-priory unclear. On the one hand, there are benefits from a paper having a topical specialization; focusing on one topic can allow researchers to engage with it more thoroughly, produce deeper insights, or become a touchstone for other scholars contending with the same topic. On the other hand, topically diverse papers could be more impactful because their contributions are broader, catching the attention of more scholars, from diverse fields. Related scholarship has failed to detect a consistent association between topical focus and downstream metrics of *researcher-level* success (Heiberger, Galvez and McFarland 2021). However, its consequences for *paper-level* success have received insufficient attention, especially in the social sciences.

---

gender composition, that cover each topic. 13.1% written by all female teams cover Gender and LGBT+. Conversely, 1.6% of all-male teams write papers on Gender and LGBT+ topics. Novelty in all-female teams' output is thus partly a result of their focus on gender, a seldom explored topic for all-male and mixed teams.

We focus on two core matrices for paper-level success: citation counts and a journal's reputation proxied by its impact factor. Citation counts are widely regarded as a measure of publication success because they reflect how much a research work influences the academic community.[26] When a paper is cited, it signifies that other researchers have found its content valuable enough to use as a foundation, reference, or justification for their work. This makes citation counts a good proxy for the paper's impact: higher citation counts often indicate greater recognition, contribution, relevance, and utility within a given field.

The number of citations of a publication depends on the time frame within which citations are counted. While a more extended period might be more accurate, it also excludes mechanically recent publications. Studies have found that two to three years is sufficient to obtain robust citation impact indicators at the paper level (Wang 2013). Raw citation counts, however, have several limitations. They vary, for example, significantly across fields (American Politics versus Political Theory), publication type (e.g., articles versus book reviews), and time (early versus recent periods). Put simply, specific fields, publication types, and later periods produce more publications and (hence mechanically) more citations than others (Waltman 2016). We address these limitations by normalizing citation counts by the year a publication appeared and document type.

We (partially) account for field diversity in citation by using journal placement. Specifically, we compare citation count within journals and account for secular trends in citations because the more time that elapses the more chance there is to get cited, by standardizing citations within journal-year. The latter measure tracks whether a specific paper garnered more citations than other papers published in the same year and outlet, and allows

[26]See Waltman (2016) for an insightful review of the literature on citation impact indicators, including their pros and cons.

us to control for journal-level differences in popularity.[27] Because both indices have different scales, we transform them to standard deviations by standardizing them as well. In so doing, we are able to compare the magnitude of differences. We estimate the following model:

$$\text{Success}_{ia} = \beta_1 \text{Index}_{ia} + \gamma_a + \varepsilon_{ia} \tag{3}$$

where success_{ia} is a measure of paper-level success, either the standardized count of citations or a binary variable that takes the value of one if paper i is published in a top-20 journal and zero otherwise. Index_{ia} measures the standardized novelty or topical focus of paper i, published by author a; γ_a are author fixed effects and ε_{ia} are robust standard errors clustered at the author-paper level.

We report results in Table 3. When comparing to papers published in the same year and outlet, a one standard deviation increase in the novelty index is associated with a 1.1% of a standard deviation increase in the number of citations after accounting for all time-invariant author-level characteristics, the preferred specification. Regarding topical diversity, a one standard deviation increase in the focus index is associated with an increase in the number of citations of 5.9% of a standard deviation. However, after accounting for time-invariant author characteristics with fixed effects, topical focus and topical novelty are both statistically unrelated to the probability that a paper is published in a Top 20 PS outlet.

Overall, the results suggest that while novel and focused papers tend to perform better

[27] In Table G11 in the Appendix, we replicate all of the results with a measure of citations standardized within year. Results for focus are consistent in magnitude. However, results for novelty are estimated to be precisely zero. These results suggest that while, relative to other papers published that year and in the same outlet, more novel papers get more citations after accounting for authors' quality, it is not the case that relative to other papers published that year overall, more novel papers get more citations, after accounting for authors' quality.

	Std. Citations (1)	Top 20 (2)	Std. Citations (3)	Top 20 (4)
Novelty (std)	0.011*	0.002		
	(0.005)	(0.002)		
Focus (std)			0.059***	0.002
			(0.010)	(0.003)
Author FE	✓	✓	✓	✓
Num.Obs.	166 387	166 387	166 387	166 387
R2	0.460	0.542	0.461	0.542

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Table shows the estimated change in paper citations (standardized within journal-year) when a paper's novelty or focus index and the estimated change in the probability that paper i is published in a Top 20 outlet (as per their SJR Impact ranking, see Table B1 in the Appendix for details.). Robust standard errors, clustered at the author and paper level, reported in parentheses.

in terms of citations than other contemporaneous papers published in the same outlets, they are not systematically more likely to be published in high-impact outlets. This aligns with research indicating that novelty is a risky strategy, occasionally yielding significant rewards (Wang, Veugelers and Stephan 2017). Novel papers may struggle to be published in prestigious outlets, but once published, they have the potential to accumulate citations as their value becomes recognized. We further examine this possibility in Figure G.2 in the Appendix. We show that more novel papers perform the same as less novel papers for the first 5-10 years after publication. After which more novel papers slowly grow to become better cited. Similarly, focused research may be less likely to be published in a high-impact journal, potentially finding a place in lower-impact but more specialized journals where they eventually receive attention from the right audience and become more frequently used by scholars relative to comparable papers published in the same year and outlet.

6 Discussion

Using a large corpus of over 140,000 papers published in political science journals and employing machine learning (in particular, text-as-data) innovations, we provide an overview of core trends in political science output in the past two decades (2003-2023). Our analysis focused on three key issues: (a) volume and researcher productivity, (b) topical and methodological focus and research design choices, and (c) research specialization, measured via topical focus and topical novelty.

As for volume, we show that political science is a steadily growing discipline: with annualized growth of slightly over 5%, the number of PS papers doubles every 13-14 years. We further show that the increase in volume is mainly because more researchers are publishing in PS outlets and not individual researchers' productivity gains. Nonetheless, we also find that younger cohorts are slightly more productive than older cohorts, at least when measured after 10 years since a researcher's first publication.

Regarding topical focus, applying STM to the entire corpus of PS papers, we show which topics became more and less popular over time and that popularity in the discipline, writ large, is not always reflected in PS's top 20 journals (with Environmental Politics, in particular, being a case in point). Classifying the method used in each paper in our corpus, we quantify the growing move toward quantitative research at the expense of normative work and studies that rely exclusively on qualitative methods. We also document trends in research design, demonstrating, in particular, the dramatic increase in the use of survey experiments, the rise and fall of certain methods (e.g., IVs), and the stubborn staying power of "selection on observables" study designs, which still account for a large share of quantitative papers.[28]

[28]Not all residual category studies necessarily use 'selection on obsevrables,' though many undoubtedly do.

The latter finding suggests that while the credibility revolution has undoubtedly impacted the discipline (mainly as reflected in top-20 publications), causal inference research designs are still far from the norm in the discipline.[29]

Finally, we explored trends in topical diversity and novelty. Consistent with growing specialization observed in other disciplines, political science papers are becoming more topically focused, and topical specialization is rewarded by being cited at higher rates than more topically diverse papers published in the same outlet. By contrast, attesting to the high reward/high-risk nature of novel research, we also show that in political science, as in science more generally (Wang, Veugelers and Stephan 2017), topical novelty is not positively correlated with top-20 outlet placement, and is only weakly rewarded, on average, with citation counts. We also documented that topical novelty has a gendered component: compared to all female teams, which, on average, produce more novel research, all male research teams are more likely to focus on incremental research.

Our study makes several notable contributions. First, we contribute to a nascent literature taking stock of publication trends in political science as a discipline. Earlier work has focused on identifying a more limited set of trends, such as the number of authors per publication (e.g., Fisher et al. 1998), the gender composition of research teams (Teele and Thelen 2017), and the content of published articles (e.g., Mas-Verdu et al. 2021; Saraceno 2020). While important, these contributions have notable limitations (e.g. selection on the DV, elite universities bias) in part due to their reliance on a limited set of selective, unrepresentative journals.[30] Our focus on methodological and research design choices, as

[29]We do not maintain that all PS quantitative studies necessarily need to use a causal inference methods: For example, there is much value in high quality descriptive work, or work that synthesize findings across studies (e.g., meta-analyses and reviews).

[30]For example, Teele and Thelen (2017)'s analysis is based on 10 to-ranked PS journals, while the analysis of Mas-Verdu et al. (2021) and Saraceno (2020) is based on tracking

well as on topical diversity and novelty complement a small political science literature that uses bibliometric analysis to explore collaboration patterns (Leifeld et al. 2017; Metz and Jäckle 2017), and knowledge production trends (Carammia 2022; Kaiser, Tóth and Demeter 2023).[31]

Second, we introduce a unique, large dataset of the metadata of papers published in political science journals over 21 years, which we plan to make publicly available. Most importantly, in addition to the data already assembled by Scopus (e.g., the number of authors and journal placement), we classify each paper by its method, research design, novelty score, and topical focus. We suspect these classifications will allow other researchers to answer (many) additional questions we have not explored herein. Consider our measure of the degree of novelty in PS publications over two decades. This measure could allow PS scholars to assess areas of rapid development, distinguish between incremental advancements and transformative research (as well as their determinants and rewards), highlight limits to external validity, and thereby open new avenues of research.

Our study helps understand the potentially unequal underlying processes fueling the growth of political science. According to the American Political Science Association's latest report on eJob Postings[32] the number of advertised jobs has declined since the 2010-2011 academic year: from 1,215 to 1,121 in 2023-2024. However, in Section 3, we show that there

---

trends over decades in a single journal (*European Political Science*, and *Journal of Politics*, respectively).

[31] Carammia (2022) studies the contribution of European political science scholarly communities, finding a tendency towards increasing diversity in the geographic basis of our discipline's scientific production. Kaiser, Tóth and Demeter (2023) uses Scopus journal index to study the characteristics of publishing house ownership, open access trends, and trends in the country of residence of authors publishing in political science for the past two decades.

[32] Available at: https://apsanet.org/wp-content/uploads/2025/02/2023-2024-eJobs-Report.pdf

are more political scientists than ever. While this excess supply of talent might lead to better selection and higher quality research, it comes at the expense of job uncertainty for younger cohorts relative to their senior fellows.

Our study also sheds more light on the nature of collaboration in PS. We documented that collaboration might be good for science (co-authored work is, on average, more novel) but not necessarily for the individual scientist (novel research is not necessarily published by more impactful journals). While our analysis of reward to topical diversity and novelty was at the *paper level*, future research can explore instead reward outcomes at the *career-level* (as in Heiberger, Galvez and McFarland (2021)), which will allow quantifying the career implications of various collaboration choices. Future work should also take more seriously the networked aspect of PS collaborations, applying network science tools to analyze the bibliometric data (at the journal and paper level) we have assembled.

Acknowledgment

We wish to extend a special thank Irene Tang for help in setting the data pipeline and Shuning Ge who took the lead in estimating the Structural Topic Models (STM). We also thank Noa Rubinstein, Micaela Montero Johnson, Ria Rege, Clarice Wang, Erika Kishino, and Arihant Tripathi for excellent research assistance.

References

Acemoglu, Daron, Simon Johnson and James A Robinson. 2001. "The colonial origins of comparative development: An empirical investigation." *American economic review* 91(5):1369–1401.

AlShebli, Bedoor K, Talal Rahwan and Wei Lee Woon. 2018. "The preeminence of ethnic diversity in scientific collaboration." *Nature communications* 9(1):5163.

Antons, David, Amol M. Joshi and Torsten Oliver Salge. 2019. "Content, Contribution, and Knowledge Consumption: Uncovering Hidden Topic Structure and Rhetorical Signals in Scientific Texts." *Journal of Management* 45(7):3035–3076.
URL: *https://doi.org/10.1177/0149206318774619*

Beach, Derek and Rasmus Brun Pedersen. 2016. *Causal case study methods: Foundations and guidelines for comparing, matching, and tracing.* University of Michigan Press.

Berinsky, Adam J, Gregory A Huber and Gabriel S Lenz. 2012. "Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk." *Political analysis* 20(3):351–368.

Blair, Graeme, Alexander Coppock and Macartan Humphreys. 2023. *Research design in the social sciences: declaration, diagnosis, and redesign.* Princeton University Press.

Blei, David M, Andrew Y Ng and Michael I Jordan. 2003. "Latent dirichlet allocation." *Journal of machine Learning research* 3(Jan):993–1022.

Blei, David M and John D Lafferty. 2007. "A correlated topic model of science.".

Carammia, Marcello. 2022. "A bibliometric analysis of the internationalisation of political science in Europe." *European Political Science* 21(4):564–595.

Collier, David. 2011. "Understanding process tracing." *PS: political science & politics* 44(4):823–830.

Coppock, Alexander. 2019. "Generalizing from survey experiments conducted on Mechanical Turk: A replication approach." *Political Science Research and Methods* 7(3):613–628.

Dipoppa, Gemma. 2024. "How Criminal Organizations Expand to Strong States: Local Agreements and Migrant Exploitation in Northern Italy." *Journal of Politics* .

Fisher, Bonnie S, Craig T Cobane, Thomas M Vander Ven and Francis T Cullen. 1998. "How many authors does it take to publish an article? Trends and patterns in political science." *PS: Political Science & Politics* 31(4):847–856.

Freeman, Richard B and Wei Huang. 2014. "Collaboration: Strength in diversity." *Nature* 513(7518):305–305.

Gerlach, Martin, Hanyu Shi and Luís A Nunes Amaral. 2019. "A universal information theoretic approach to the identification of stopwords." *Nature Machine Intelligence* 1(12):606–612.

Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments." *Political analysis* 22(1):1–30.

Heiberger, Raphael H., Sebastian Munoz-Najar Galvez and Daniel A. McFarland. 2021. "Facets of Specialization and Its Relation to Career Success: An Analysis of U.S. Sociology,

1980 to 2015." *American Sociological Review* 86(6):1164–1192.
URL: *https://doi.org/10.1177/00031224211056267*

Kaiser, Tamás, Tamás Tóth and Marton Demeter. 2023. "Publishing Trends in Political Science: How Publishing Houses, Geographical Positions, and International Collaboration Shapes Academic Knowledge Production." *Publishing Research Quarterly* 39(3):201–218.

Kristensen, Peter M. 2012. "Dividing discipline: structures of communication in international relations." *International Studies Review* 14(1):32–50.

Lal, Apoorva, Mackenzie Lockhart, Yiqing Xu and Ziwen Zu. 2024. "How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice Based on 67 Replicated Studies." *Political Analysis* 32(4):521–540.

Leifeld, Philip, Sandra Wankmüller, Valentin TZ Berger, Karin Ingold and Christiane Steiner. 2017. "Collaboration patterns in the German political science co-authorship network." *PloS one* 12(4):e0174671.

Mahoney, James. 2004. "Comparative-historical methodology." *Annu. Rev. Sociol.* 30(1):81–101.

Maliniak, Daniel, Susan Peterson-Ryan Powers and Michael J. Tierney. 2018. "Is International Relations a Global Discipline? Hegemony, Insularity, and Diversity in the Field." *Security Studies* 23(3).

Mas-Verdu, Francisco, Jose-Maria Garcia-Alvarez-Coque, Paula Andrea Nieto-Aleman and Norat Roig-Tierno. 2021. "A systematic mapping review of European Political Science." *European Political Science* 20(1):85.

McGrath, Erin and Ana Diaz. 2023. "2021-2022 eJobs Report: the Political Science Job Market.".

Metz, Thomas and Sebastian Jäckle. 2017. "Patterns of publishing in political science journals: an overview of our profession using bibliographic data and a co-authorship network." *PS: Political Science & Politics* 50(1):157–165.

Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004. "Economic shocks and civil conflict: An instrumental variables approach." *Journal of political Economy* 112(4):725–753.

Powell, Kendall. 2018. "These labs are remarkably diverse–here's why they're winning at science." *Nature* 558(7708):19–23.

Roberts, Margaret E, Brandon M Stewart and Dustin Tingley. 2019. "Stm: An R package for structural topic models." *Journal of statistical software* 91:1–40.

Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Edoardo M Airoldi et al. 2013. The structural topic model and applied social science. In *Advances in neural information processing systems workshop on topic models: computation, application, and evaluation*. Vol. 4 Harrahs and Harveys, Lake Tahoe pp. 1–20.

Saraceno, Joseph. 2020. "Disparities in a flagship political science journal? Analyzing publication patterns in the journal of politics, 1939–2019." *The Journal of Politics* 82(4):e45–e55.

Sovey, Allison J and Donald P Green. 2011. "Instrumental variables estimation in political science: A readers' guide." *American Journal of Political Science* 55(1):188–200.

Teele, Dawn Langan and Kathleen Thelen. 2017. "Gender in the journals: Publication patterns in political science." *PS: Political Science & Politics* 50(2):433–447.

Tomz, Michael R and Jessica LP Weeks. 2013. "Public opinion and the democratic peace." *American political science review* 107(4):849–865.

TRIP Journal Article Database Release (Version 3.3). 2020. https://trip.wm.edu/dashboard/journal-articles.

Wæver, Ole. 1998. "The sociology of a not so international discipline: American and European developments in international relations." *International organization* 52(4):687–727.

Waltman, Ludo. 2016. "A review of the literature on citation impact indicators." *Journal of informetrics* 10(2):365–391.

Wang, Dashun and Albert-László Barabási. 2021. *The science of science*. Cambridge University Press.

Wang, Jian. 2013. "Citation time window choice for research impact evaluation." *Scientometrics* 94(3):851–872.

Wang, Jian, Reinhilde Veugelers and Paula Stephan. 2017. "Bias against novelty in science: A cautionary tale for users of bibliometric indicators." *Research Policy* 46(8):1416–1436.

Yang, Yang, Tanya Y Tian, Teresa K Woodruff, Benjamin F Jones and Brian Uzzi. 2022. "Gender-diverse teams produce more novel and higher-impact scientific ideas." *Proceedings of the National Academy of Sciences* 119(36):e2200841119.

Online Appendix "Trends in Political Science: 20-years"

(not intended for publication)

A Glossary

SJR (SCImago Journal Rank) is a metric used in Scopus to assess the impact and prestige of academic journals. A key feature of SJR is that it is a Prestige-Weighted Metric: it measures the scientific influence of journals by considering both the number of citations they receive and the prestige of the journals where those citations come from. Unlike raw citation counts, SJR assigns higher value to citations from more influential journals. Specifically, SJR is the number of citations in a specific year (e.g., 2023) to articles published in the previous three years (e.g., 2020–2022), considering the prestige of the citing journals.

B Data Collection

In general, there are two approaches to systemically map an academic discipline, each with its pros and cons. Choosing between them entails a trade-off.

One approach is to first map the scholars who make the discipline; for example, by assembling the names of all standing faculty in political science departments (nationally or globally, depending on the study's objective). This is the approach taken by Leifeld et al. (2017) when analyzing trends in political science scholarship in Germany. One advantage of this approach is that it allows tracing all publications of political science faculty whether these are published in political science outlets, or elsewhere (e.g., in general interest science journals or journals associated with related fields. Another advantage is that it allows enumerating faculty's publications beyond peer reviewed journals, including monographs and books chapters. In essence, this approach assumes that political science's discipline's output is simply what political scientists, hired by academic departments, publish.

While using scholars as the start point has its benefits, it also comes with a major drawback: selecting on success and survival. This is especially problematic for studies that wish to (also) explain publication success. As always, selecting on the dependent variable introduces bias. Using academics as a start point has another drawback: it omits publications written by those who are not hired by political science departments, including academics from related disciplines, non-academics, graduate students who ended up not taking academic jobs, etc. In short, using political scientists as a start point is a problem if we assume that political science's output as a discipline is what gets published in political science outlets, irrespective of the home institution of the author. As Wæver (p. 697 1998) argued "Journals are the most direct measure of the discipline itself."

These problems are solved if the starting point is political science journals (irrespective of the author's status and institution). This approach does not suffer from selection on survival and success, and it does not arbitrarily omit publications in political science journals just because their author is (currently) not hired by a political science department. Of course, using political science journals as a starting points also has drawback: we underestimate the productivity of scholars who published in non-PS outlets. Since our goal is to account for trends in political science as a discipline (and not trends of political scientists), we chose journals as our starting point, and complemented information about authors' publications outside PS journals using Scopus's metadata. Below we describe how we moved from the names of political science journals to assemble four attribute datasets that capture a 21-year period (2003-2023): (a) journal-level dataset, (b) article-level dataset, (c) author-level dataset, (d) commenter-level dataset. In addition, we assemble two relational dataset: (e) network of authors, and (f) network of authors and commenters.

Journal-level data

Our starting point is the full list of all 188 political science journals as classified by Clarivate — an analytics company that provides tools for scientific research and academic performance evaluation via its Web of Science platform. This number of journals is a marked improvement on past reviews of the discipline that generally only uses a sample of journals: for example, Fisher et al. (1998) base their trends analysis on three journals; Wæver (1998) uses seven journals, Kristensen (2012) uses 59 journals, Metz and Jäckle (2017) use 96 journals and Carammia (2022) bases their analysis on 100 journals.

We first drop 3 journals that are not on Scopus, 3 journals that are not peer-review, and 9 journals that do not publish in English (e.g., Historia Y Politica), leaving us with a set of 174 peer-review, English language, political science journals. We then used journal names to match each journal listed by Clarivate to the journal's metadata as measured by Scopus, a comprehensive bibliographic database for academic research managed by Elsevier. Our journal-level data includes information such as yearly publication count, yearly citation count, and most importantly, for our analysis of publication success, Scoups's metrics that allow assessing the journal's relative performances, such as CiteScore, SNIP (Source Normalized Impact per Paper), and SJR (SCImago Journal Rank).[33] In addition, for each journal we calculate additional variables such as the number of unique authors (including by sex), and the type of articles it publishes, by methods and topics.

[33]See SI Section A for a glossary of measures and concepts used throughout this manuscript.

Table B1: Top 20 Journals

Title	Coverage Start	N Paps	w/ Full Text	Mean SJR	2023 SJR
American Political Science Review	1906	1269	1268	6.50	5.07
International Organization	1947	594	594	5.44	4.93
American Jrnl of Political Science	1982	1341	1341	6.17	4.63
Political Analysis	1989	641	641	5.12	4.61
Comparative Political Studies	1968	1267	1267	3.13	3.49
Political Communication	1980	661	660	2.09	3.35
European Jrnl of Political Research	1973	1220	1220	2.38	3.33
British Jrnl of Political Science	1971	1093	1093	3.12	3.32
World Politics	1948	373	366	3.67	3.02
Jrnl of Public Admin. Research & Theory	1991	799	791	3.48	2.98
Jrnl of Politics	1939	2013	2011	3.44	2.79
Political Behavior	1979	862	862	2.42	2.69
Political Science Research & Methods	2018	356	356	2.34	2.43
Quarterly Jrnl of Political Science	2006	248	246	3.08	2.03
Jrnl of Conflict Resolution	1957	1182	1175	3.02	1.86
Jrnl of Peace Research	1964	1158	1151	2.69	1.74
Public Opinion Quarterly	1937	844	835	2.08	1.64
International Security	1984	580	0	3.37	1.58
International Studies Quarterly	1978	1273	1271	2.34	1.50
European Union Politics	2000	593	593	2.28	1.38

Table B2: Journal Sample Composition

Journal Title	**ISSN**	**Journal Title**	**ISSN**	**Journal Title**	**ISSN**
		In Sample (With Text)			
International Organization	0020-8183	Political Communication	1058-4609	American Political Science Review	0003-0554
Contemporary Security Policy	1352-3260	Environmental Politics	0964-4016	Political Analysis	1047-1987
European Journal of Political Research	0304-4130	British Journal of Political Science	0007-1234	Comparative Political Studies	0010-4140
World Politics	0043-8871	Policy and Internet	nan	International Journal of Press/politics	1940-1612
Global Environmental Politics	1526-3800	Political Psychology	0162-895X	Journal of Chinese Political Science	1080-6954
Review of International Political Economy	0969-2290	Journal of Public Administration Research and Theory	1053-1858	West European Politics	0140-2382
American Journal of Political Science	0092-5853	Journal of European Public Policy	1350-1763	New Political Economy	1356-3467
Political Geography	0962-6298	Review of International Organizations	1559-7431	Political Behavior	0190-9320
Political Science Research and Methods	2049-8470	Policy Studies Journal	0190-292X	Perspectives on Politics	1537-5927
Socio-economic Review	1475-1461	Journal of Peace Research	0022-3433	Public Administration	0033-3298
Canadian Journal of Political Science	0008-4239	Public Opinion Quarterly	0033-362X	International Environmental Agreements: Politics, Law and Economics	1567-9764
Politics and Gender	1743-923X	International Studies Review	1521-9488	International Theory	1752-9719

Continued on next page

Table B2: Journal Sample Composition

Journal Title	ISSN	Journal Title	ISSN	Journal Title	ISSN
South European Society and Politics	1360-8746	Ps - Political Science and Politics	1049-0965	East European Politics	2159-9165
Democratization	1351-0347	European Political Science Review	1755-7739	Journal of Democracy	1045-5736
Journal of Politics	0022-3816	Journal of Conflict Resolution	0022-0027	Political Studies	0032-3217
Social Movement Studies	1474-2837	Regulation and Governance	1748-5983	Journal of European Integration	0703-6337
Geopolitics	1465-0045	Governance	0952-1895	African Affairs	0001-9909
Government and Opposition	0017-257X	Annals of the American Academy of Political and Social Science	0002-7162	Research and Politics	nan
Party Politics	1354-0688	Studies in Comparative International Development	0039-3606	International Studies Quarterly	0020-8833
Journal of Information Technology and Politics	1933-1681	Territory, Politics, Governance	2162-2671	Journal of Current Southeast Asian Affairs	1868-1034
International Political Sociology	1749-5679	Terrorism and Political Violence	0954-6553	Journal of Public Policy	0143-814X
Politics and Society	0032-3292	European Union Politics	1465-1165	Electoral Studies	0261-3794
Comparative Politics	0010-4159	Nations and Nationalism	1354-5078	Philosophy and Public Affairs	0048-3915
Politics and Governance	nan	International Political Science Review	0192-5121	Post-soviet Affairs	1060-586X
Journal of Common Market Studies	0021-9886	Political Research Quarterly	1065-9129	Swiss Political Science Review	1424-7755
Review of Policy Research	1541-132X	Cooperation and Conflict	0010-8367	European Political Science	1680-4333
Studies in Conflict and Terrorism	1057-610X	European Journal of Political Economy	0176-2680	Political Quarterly	0032-3179
Latin American Politics and Society	1531-426X	Political Studies Review	1478-9299	Global Policy	1758-5880
Journal of Strategic Studies	0140-2390	Social Science Quarterly	0038-4941	Local Government Studies	0300-3930
Journal of Human Rights	1475-4835	International Journal of Conflict and Violence	1864-1385	New Left Review	0028-6060
Politics	0263-3957	Contemporary Political Theory	1470-8914	Publius: the Journal of Federalism	0048-5950
Business and Politics	1469-3569	International Journal of Public Opinion Research	0954-2892	Journal of Elections, Public Opinion and Parties	1745-7289
British Journal of Politics and International Relations	1369-1481	Citizenship Studies	1362-1025	Journal of Political Philosophy	0963-8016
Meditteranean Politics	1362-9395	Journal of International Relations and Development	1408-6980	Quarterly Journal of Political Science	1554-0626
Comparative European Politics	1472-4790	European Security	0966-2839	The International Journal of Transitional Justice	1752-7716
International Feminist Journal of Politics	1461-6742	Problems of Post-communism	1075-8216	Public Choice	0048-5829
Politics and Religion	nan	American Politics Research	1532-673X	Legislative Studies Quarterly	0362-9805
Europe-asia Studies	0966-8136	German Politics	0964-4008	International Politics	1384-5748
Armed Forces and Society	0095-327X	Political Theory	0090-5917	Review of African Political Economy	0305-6244
Journal of Contemporary European Studies	1478-2804	Scandinavian Political Studies	0080-6757	Ethics and International Affairs	0892-6794
Current History	0011-3530	International Affairs	0020-5850	Political Science	0032-3187

Continued on next page

Table B2: Journal Sample Composition

Journal Title	ISSN	Journal Title	ISSN	Journal Title	ISSN
Intelligence and National Security	0268-4527	Nationalities Papers	0090-5992	Acta Politica	0001-6810
Parliamentary Affairs	0031-2290	Survival	0039-6338	Australian Journal of Political Science	1036-1146
Revista Brasileira De Politica Internacional	0034-7329	Journal of Women, Politics and Policy	1554-477X	British Politics	1746-918X
Scottish Journal of Political Economy	0036-9292	Polity	0032-3497	Japanese Journal of Political Science	1468-1099
Journal of Theoretical Politics	0951-6298	Economics and Politics	0954-1985	Human Rights Quarterly	0275-0392
Contemporary Southeast Asia	0129-797X	Politics, Philosophy and Economics	1470-594X	Communist and Post-communist Studies	0967-067X
Australian Journal of Politics and History	0004-9522	Presidential Studies Quarterly	0360-4918	Critical Review	0891-3811
Political Science Quarterly	0032-3195	Studies in American Political Development	0898-588X	Forum (germany)	1540-8884
East European Politics and Societies	0888-3254	Latin American Perspectives	0094-582X	Irish Political Studies	0790-7184
Ethics and Global Politics	1654-4951	Historical Materialism	1465-4466	European History Quarterly	0265-6914
Politicka Ekonomie	0032-3233	Journal of Policy History	0898-0306	Telos	0090-6514
In Sample (Without Text)					
Annual Review of Political Science	1094-2939	Policy and Society	1449-4035	Policy and Politics	0305-5736
Journal of Chinese Governance	2381-2346	Contemporary Politics	1356-9775	Journal of Political Ideologies	1356-9317
International Security	0162-2889	Critical Policy Studies	1946-0171	State Politics and Policy Quarterly	1532-4400
Cambridge Review of International Affairs	0955-7571	Revista De Ciencia Politica	0716-1417	Peacebuilding	2164-7259
Politische Vierteljahresschrift	0032-3470	Monthly Review	0027-0520	Journal of Australian Political Economy	0156-5826
Politikon	0258-9346	Nation	0027-8378	Austrian Journal of Political Science	nan
Journal of Cold War Studies	1520-3972	Dissent	0012-3846	Politica Y Gobierno	1405-1060
Historia Y Politica	1575-0361	Lex Localis	1581-5374	Revista De Estudios Politicos	0048-7694
Independent Review	1086-1653	Internasjonal Politikk	0020-577X	Osteuropa	0030-6428
Politix	0295-2319	Romanian Journal of Political Science	1582-456X		
Out of Sample					
Earth System Governance	2589-8116	Chinese Political Science Review	2365-4244	Journal of Experimental Political Science	2052-2630
European Policy Analysis	nan	Journal of Genocide Research	1462-3528	Journal of Political Power	2158-379X
European Journal of Politics and Gender	2515-1088	Frontiers in Political Science	nan	Contemporary Italian Politics	2324-8823
Research & Politics	nan	World	nan	Journal of Political Marketing	1537-7857
Global Discourse	2326-9995	Political Research Exchange	nan	Politics Groups and Identities	2156-5503

Continued on next page

Table B2: Journal Sample Composition

Journal Title	ISSN	Journal Title	ISSN	Journal Title	ISSN
Critical Studies on Security	2162-4887	Italian Political Science Review-rivista Italiana Di Scienza Politica	0048-8402	Regional and Federal Studies	1359-7566
African Security	1939-2206	Review of Economics and Political Science	2356-9980	Peace Economics Peace Science and Public Policy	1079-2457
Global Public Policy and Governance	2730-6291	Critical Studies on Terrorism	1753-9153	Journal of Politics in Latin America	1866-802X
Global Social Policy	1468-0181	State Crime	2046-6056	Business and Politics	nan
Politics & Policy	1555-5623	Studies in Social Justice	1911-4788	Interest Groups & Advocacy	2047-7414
Zeitschrift Fur Vergleichende Politikwissenschaft	1865-2646	Election Law Journal	1533-1296	Behavioral Sciences of Terrorism and Political Aggression	1943-4472
European Journal of Political Theory	1474-8851	Civil Wars	1369-8249	Ethnopolitics	1744-9057
Public Administration and Policy-an Asia-pacific Journal	1727-2645	Constellations-an International Journal of Critical and Democratic Theory	1351-0487	Journal of International Political Theory	1755-0882
Journal of Comparative Politics	1338-1385	East Asian Policy	1793-9305	Capital and Class	0309-8168
International Journal of Politics Culture and Society	0891-4486	Asian Journal of Comparative Politics	2057-8911	Asian Politics & Policy	1943-0779
Forum-a Journal of Applied Research in Contemporary Politics	2194-6183	New Perspectives	2336-825X	Moral Philosophy and Politics	2194-5616
Journal of Human Rights Practice	1757-9619	Partecipazione E Conflitto	1972-7623	Journal of Political Science Education	1551-2169
Critical Review of International Social and Political Philosophy	1369-8230	National Identities	1460-8944	Democratic Theory-an Interdisciplinary Journal	2332-8894
Global Constitutionalism	2045-3817	Global Responsibility to Protect	1875-9858	Issues & Studies	1013-2511
Colombia Internacional	0121-5612	Rethinking Marxism-a Journal of Economics Culture & Society	0893-5696	Revista Espanola De Ciencia Politica-recp	1575-6548
Latin American Policy	2041-7365	Democracy & Security	1741-9166	Teoria Y Realidad Constitucional	1139-5583
Politics Religion & Ideology	2156-7689	Polis-politicheskiye Issledovaniya	1026-9487	Journal of Civil Society	1744-8689
Global Change Peace & Security	1478-1158	Nordic Journal of Human Rights	1891-8131	Russian Politics	2451-8913
Rusi Journal	0307-1847	Socialist Studies	1918-2821	Revue D Economie Politique	0373-2630
Taltech Journal of European Studies	2674-4600	Asian Journal of Political Science	0218-5377	French Politics	1476-3419
Commonwealth & Comparative Politics	1466-2043	Politologicky Casopis-czech Journal of Political Science	1211-3247	Insight Turkey	1302-177X
Politica Y Sociedad	1130-8001	Caucasus Survey	2376-1199	Journal of Public Finance and Public Choice	2515-6918
Scottish Affairs	0966-0356	New Political Science	0739-3148	Intersections-east European Journal of Society and Politics	nan
China Quarterly of International Strategic Studies	2377-7400	Desafios	0124-4035	International Critical Thought	2159-8282
Geopoliticas-revista De Estudios Sobre Espacio Y Poder	2172-3958	Politicke Vedy	1335-2741	Politique Europeenne	1623-6297

Continued on next page

Table B2: Journal Sample Composition

Journal Title	ISSN	Journal Title	ISSN	Journal Title	ISSN
Populism	2588-8064	Security and Human Rights	1874-7337	Otoritas-jurnal Ilmu Pemerintahan	2088-3706
Studia Europejskie-studies in European Affairs	1428-149X	Studies in Indian Politics	2321-0230	Obrana a Strategie-defence & Strategy	1214-6463
Strategic Review for Southern Africa	1013-1108	Urvio-revista Latinoamericana De Estudios De Seguridad	1390-3691	American Political Thought	2161-1580
Politeia-journal of Political Theory Political Philosophy and Sociology of Politics	2078-5089	Australasian Parliamentary Review	1447-9125	Idp-internet Law and Politics	1699-8154
Revista Brasileira De Estudos Politicos	0034-7191	Revista Internacional De Pensamiento Politico	1885-589X	International Journal of Cyber Warfare and Terrorism	1947-3435
America Latina Hoy-revista De Ciencias Sociales	1130-2887	Conflict Studies Quarterly	2285-7605	Revista De Investigaciones Politicas Y Sociologicas	1577-239X
Ciencia Politica	1909-230X	Reflexion Politica	0124-0781	Izquierdas	0718-5049
Canadian Political Science Review	1911-4125	New Proposals-journal of Marxism and Interdisciplinary Inquiry	1715-6718	Politicka Misao-croatian Political Science Review	0032-3241
Analele Universitatii Bucuresti-stiinte Politice	1582-2486	European Journal of Transformation Studies	2298-0997	Siyasal-journal of Political Sciences	nan
Icelandic Review of Politics & Administration	1670-6803	African Journal on Conflict Resolution	1562-6997	Revista Del Clad Reforma Y Democracia	1315-2378
Revista Estudios Socio-juridicos	0124-0579	Tocqueville Review	0730-479X	Scienza & Politica-per Una Storia Delle Dottrine	1590-4946
Analecta Politica	2027-7458	Temas Y Debates	1666-0714	Turkish Policy Quarterly	1303-5754
Anacronismo E Irrupcion	2250-4982	Cimexus	1870-6479	Ciudad Paz-ando	2011-5253
Pensamiento Al Margen	2386-6098	Revista Estudos Politicos	2177-2851	Sravnitelnaya Politika-comparative Politics	2221-3279
Revista Andina De Estudios Politicos	2221-4135	Politica & Societa	2240-7901	Revista Mexicana De Analisis Politico Y Administracion Publica	2007-4425
Storia Del Pensiero Politico	2279-9818	Totalitarismus Und Demokratie	1612-9008	Laboratoire Italien-politique Et Societe	1627-9204

Paper-level data

Using Scopus's metadata, we further extract information on all 154,738 articles published in our sample. Scopus's metadata includes information on the article's authors, title, abstract, publication date, and DOI link. While informative, Scopus's article metadata is limited; for example, it does not tell much about the article's topic of inquiry, nor the method used or the identity of those commenting on papers along the way. We, therefore, supplemented Scopus's metadata by first downloading all articles that we were able to.[34] Specifically, we were able to successfully scrape the paper text of 111,854 articles. Armed with the full text of the articles, we classified each paper by topic using structure topic modeling (STM) and

[34]We were unable to scrape certain articles for two main reasons: first, we had limited institutional access to the articles, or second, the websites themselves were un-scrapable due to technical reasons.

by method using a combination of Supervised Machine Learning and ChatGPT.

Author-level data

We identify 95,567 unique authors who wrote the 154,738 articles published in the complete list of political science journals between 2003 and 2023 (85,654 unique authors in the filtered list of journals). Using Scopus's metadata, extracting information on these authors is relatively straightforward. Scopus metadata includes information on each author's yearly number of documents (i.e., publications), yearly citation count, and affiliation country. We supplement Scopus with measures of authors' sex, constructed using the genderize.io package,[35] as well as summary measures of publication success such as h-index and Euclid scores.

Commenter-level data

Given the importance of informal forms of collaboration, a key innovation of this study is assembling systematic information on those commenting on journal articles. Neither Scopus nor Clarivate collects this information. We construct our original dataset of commenters in three steps. First, we extracted the acknowledgment section, when such a section exists, from each of the 111,854 journal articles we scraped for a total of 77,101 acknowledgment sections. In the next step, we use GPT to extract the names of commenters for 63,522 articles with an acknowledgment section and names of commenters (1,579 articles had an acknowledgment section but did not thank anyone as a commenter). This stage left us with the names of 105,008 unique commenters. In the third step, we matched the commenters' names, as extracted directly from the acknowledgment section, back to Scopus's metadata using a fuzzy match algorithm. Notably, only 66,967 commenters (or 64% of named commenters) have a Scopus ID. In such cases, we constructed a dataset with similar metadata information on authors (e.g., number of yearly publications and citations, h-index, and Euclid scores.) This information does not exist for the 36% of commenters we could not match to a recognizable scholar.

C Volume

In Figure 2, we show that the rate of published papers per outlet is consistently increasing. This could be the result of (a) existing journals publishing more paper per issue or (b) new journals publishing more paper per issue from their inception. In Figure C1, we replicate the right panel of Figure 2 but subletting the data only to journals which were already

[35]The *genderize.io* package predicts binary gender based on the frequency of first names (and country when available) in a labeled dataset of over one billion public social media profiles. We were able to assign gender for about 97% of authors based on their name and country of origin with a mean posterior probability of 96.8%.

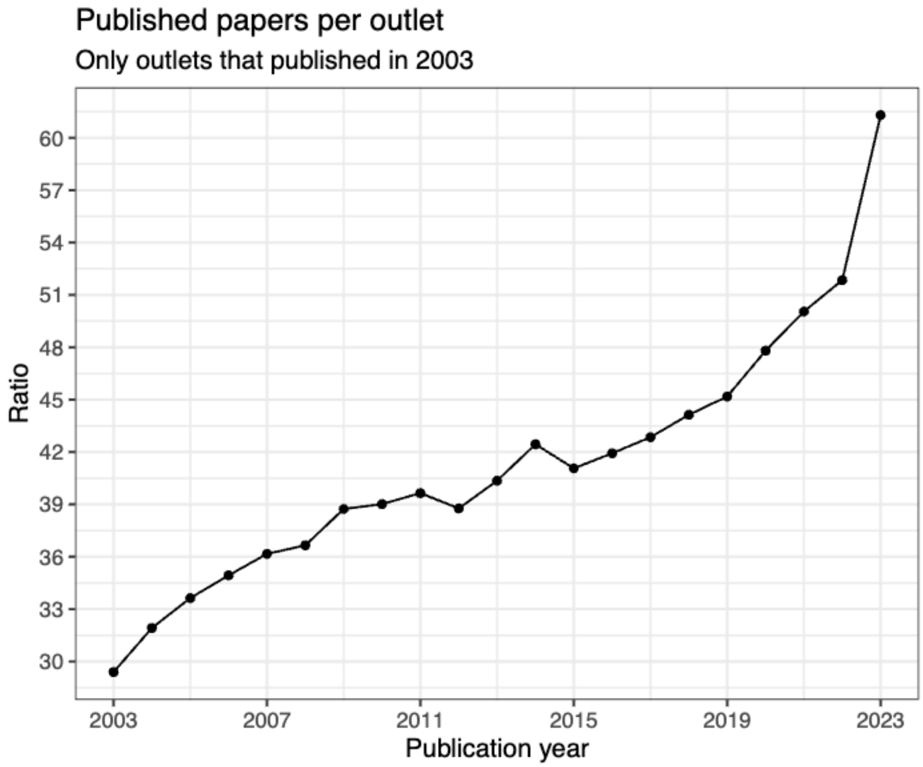

Figure C1: Figure shows the ratio of published papers over publishing outlets per year, for papers published in outlets that were already existent in 2003.

publishing in 2003. Even after reducing the sample, we can see the same trend in the number of published papers per outlet, suggesting that existing journals became more voluminous.

D Productivity

E Topics (STM)

To discover and extract thematic and semantic structures embedded in each paper, we fit Structural Topic Model (STM) (Roberts et al. 2013) to the first 1000 words of each paper. Topic modeling approaches such as Latent Dirichlet Allocation (LDA) (Blei, Ng and Jordan 2003) and the Correlated Topic Model (CTM) (Blei and Lafferty 2007) has been widely used in the field of social sciences and political sciences. Topic models is a generative model that defines the data generating process to be the following: each document (paper) first draws topics from a document-topic distribution, then conditional on the chosen topics, each word is again draw from a topic-word distribution to construct a full document. The input of a topic model is a document-term matrix where each document is represented by its unique compositing words and their frequency. The model uses variational inference to estimate parameters upon convergence. Final outputs include a list of identified topics, the posterior proportion of each document allocated to each topic, and the posterior probability distribution of words associated with respective topic.

Compared to these traditional topic modeling methods, we prefer STM because it allows researchers to incorporate document attributes or metadata into the topic modeling. In this case, either topical prevalence or topical content, or both, can be modeled as a function of the document-specific covariates. Topical prevalence impact the document-topic distribution whereas topical content refers to topic-word distribution. Our STM was estimated with publication year and journal placement as prevalence covariates, allowing the proportion of topics allocated to paper to vary across years and journals.

Since STM is an unsupervised learning process, the number of topics needs to be specified by researchers. We selected the model with 30 topics after carefully investigating topic interpretations with specifying the number of topics being 10, 20, 25, 30, 40, or 50. We present the first 10 highest frequency words for each topic, along with our manually created names in Table E3.

Table E3: Names & Top 10 Words Per STM Topic

Topic Name	**Top 10 Words (lemmatized)**
1 - Electoral Institutions	*elect, vote, voter, candid, elector, parti, campaign, turnout, incumb, system, presidenti, democrat, seat, district, poll, result, effect, ballot, win, support*
2 - Quantitative Methods	*data, model, use, measur, analysi, differ, variabl, estim, effect, method, compar, number, empir, result, two, approach, time, test, set, case*
3 - Ethnic, Religious, and National Identity	*nation, ident, ethnic, religi, citizenship, cultur, religion, state, minor, group, polit, communiti, muslim, languag, islam, memori, church, nationalist, peopl, christian*
4 - Federalism and Decentralization	*local, region, state, govern, nation, feder, level, polit, territori, citi, system, urban, central, municip, parti, decentr, subnat, area, rural, differ*
5 - Democracy and Autocracy	*democraci, regim, polit, democrat, institut, state, authoritarian, power, countri, elit, reform, rule, transit, govern, latin, leader, opposit, econom, system, chang*
6 - Political Communication	*media, news, polit, communic, polar, inform, internet, digit, onlin, public, coverag, social, content, mass, technolog, platform, audienc, televis, frame, discuss*
7 - Financial and Labor Markets	*econom, market, economi, capit, bank, industri, labor, growth, busi, sector, financial, product, financi, invest, crisi, rate, worker, privat, labour, employ*

Table E3: Names & Top 10 Words Per STM Topic

Topic Name	**Top 10 Words (lemmatized)**
8 - Political Culture	*social, group, differ, polit, trust, individu, attitud, support, cultur, result, factor, countri, prefer, econom, level, valu, analysi, toward, import, determin*
9 - Post-Soviet Politics	*russia, russian, soviet, ukrain, communist, ukrainian, eastern, war, putin, europ, region, moscow, union, countri, republ, west, foreign, western, polici, postcommunist*
10 - Law, Human Rights, and Courts	*right, law, human, state, legal, court, intern, justic, constitut, protect, rule, case, enforc, norm, polic, crime, govern, crimin, violat, judici*
11 - Gender and LGBTQI+	*women, gender, men, femal, feminist, sexual, polit, male, equal, represent, quota, differ, sex, marriag, role, publish, gap, experi, intersect, masculin*
12 - Terrorism	*terror, intellig, terrorist, attack, australian, australia, group, ireland, secur, cite, northern, british, threat, radic, govern, irish, activ, oper, term, suicid*
13 - US Presidency	*american, presid, time, polit, year, mani, govern, public, press, peopl, nation, histori, end, made, publish, day, unit, like, british, earli*
14 - Critical Theory	*social, polit, concept, theori, practic, cultur, relat, discours, world, idea, way, histor, histori, critic, approach, form, develop, narrat, mean, modern*
15 - Environmental Politics	*polici, chang, govern, environment, institut, climat, actor, develop, process, global, approach, focus, intern, problem, energi, framework, system, environ, state, policymak*
16 - Coalition Formation and Party Systems	*parti, polit, govern, coalit, system, parliamentari, posit, ideolog, parliament, elector, populist, polici, minist, leader, right, support, issu, left, chang, elect*
17 - Social Movements	*movement, organ, social, network, protest, mobil, activ, group, societi, civil, activist, action, collect, polit, stakehold, relationship, interest, structur, transnat, sport*
18 - Indigenous Politics	*indigen, land, des, les, canadian, que, canada, food, los, las, politiqu, mexico, latin, del, sur, dan, agricultur, con, est, para*

Table E3: Names & Top 10 Words Per STM Topic

Topic Name	**Top 10 Words (lemmatized)**
19 - Bureaucratic Politics	*public, govern, servic, agenc, administr, manag, organ, inform, health, perform, provid, student, program, privat, respons, bureaucrat, sector, organiz, develop, implement*
20 - Lobbying and Interest Groups	*legisl, polici, presid, group, member, interest, state, congress, committe, execut, power, senat, decis, hous, legislatur, major, court, presidenti, prefer, polit*
21 - Global Trade	*china, countri, trade, develop, global, econom, world, africa, chines, south, aid, intern, region, african, nation, asia, foreign, oil, japan, govern*
22 - Public Opinion and Political Psychology	*survey, public, polit, opinion, effect, respond, attitud, support, inform, individu, behavior, experi, respons, like, peopl, percept, affect, question, evalu, find*
23 - Security Studies	*intern, state, secur, power, foreign, polici, relat, nuclear, war, unit, global, domest, world, threat, strateg, nation, cooper, order, strategi, interest*
24 - European Politics	*european, polici, union, member, integr, state, nation, europ, countri, institut, commiss, govern, crisi, germani, process, council, actor, negoti, level, polit*
25 - Normative Theory	*moral, theori, liber, reason, human, argument, individu, claim, peopl, polit, principl, right, valu, argu, way, view, ethic, good, justic, equal*
26 - Race and Immigration Politics	*educ, immigr, social, state, racial, american, black, famili, school, welfar, popul, white, children, migrat, polici, health, migrant, increas, percent, age*
27 - War	*militari, war, forc, iraq, oper, arm, armi, civilian, israel, arab, unit, afghanistan, soldier, secur, iran, defens, support, isra, state, pakistan*
28 - Civil War and Intergroup Conflict	*iolenc, conflict, war, peac, group, conflict, civil, state, violent, arm, rebel, intern, like, actor, govern, polit, effect, increas, support, victim*
29 - Fiscal Politics, Inequality, and Redistribution	*govern, cost, polici, econom, tax, effect, model, public, countri, corrupt, spend, incom, increas, good, literatur, prefer, result, incent, polit, level*
30 - Representation and Accountability	*polit, citizen, democraci, particip, democrat, public, institut, repres, process, govern, interest, legitimaci, represent, engag, system, peopl, account, civic, decis, deliber*

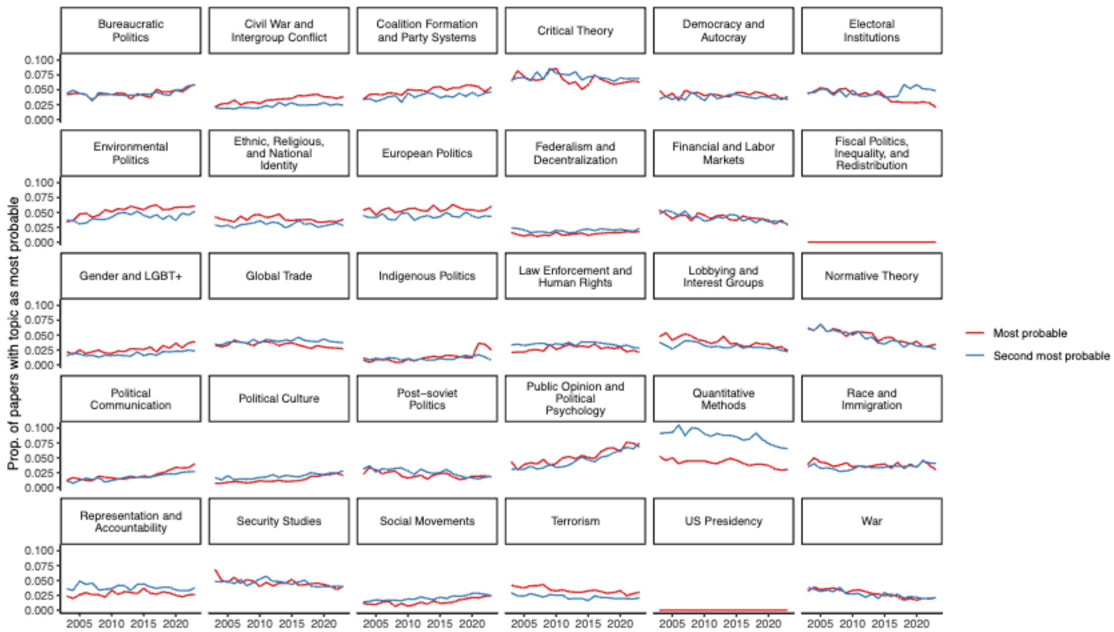

Figure E2: Figure shows the proportion of yearly papers for which the STM model estimates each topic as the most probable (in red) and for which the STM model estimates each topic as the second most probable (in blue).

Topic rank	Mean Prob.	SD Prob.	Min Prob.	Max Prob.	Median Prob.
1	0.392	0.164	0.0108	0.999	0.362
2	0.199	0.069	0.0004	0.491	0.196
3	0.120	0.048	0.0002	0.306	0.119
4	0.0772	0.035	0.0001	0.222	0.076
5	0.0520	0.027	0.0001	0.172	0.050

Table E4: Descriptive statistics of posterior topic proportion, for the five most likely topics per paper.

Count of Topics with $P(\theta_t) \geq .1$	Number of Papers	Prop. of Papers	Cumulative Prop. of Papers
0	1	0.00	0.00
1	7,497	6.72	6.72
2	30,220	27.09	33.81
3	44,925	40.27	74.08
4	23,789	21.32	95.40
5	4,794	4.30	99.70
6	332	0.30	99.99
7	2	0.00	100.00

Table E5: Count and proportion of papers in the sample, according to the number of topics for which the STM estimates a posterior probability higher than .1.

F Method Classification

Overall Methods

To classify papers with available full-texts into methods: quant, qual/normative, formal, we use the following procedure.

- First, we clean and prepare the raw text of the paper by: a) removing any preamble before the abstract and the bibliography, b) text cleaning including removing white space, removing lines with fewer than 5 words, and removing lines with less than 50% text characters, c) and removing stopwords from the text using The Natural Language Toolkit (NLTK) Python library's preset list of English stopwords as well as a custom set of stopwords which are listed in Table **??**.

- Second, we prepare a sample of papers manually coded by Teele and Thelen (2017) that overlaps with our own sample to use in training our classification model. Teele and Thelen classified 1,694 papers in our full-text sample into six categories: statistical (1,138), experiment (145), qualitative (140), formal theory (118), political theory (105), and conceptual (48), which we consolidated into our three categories: quant (1283), qual/normative (293), formal (118).

- Third, we use this corpus of text to train two stage classifier using the Scikit-learn Python ML library. In the first stage, we vectorize the text with TfidfVectorizer, then use a OneVsRestClassifier alongside a LogisticRegression to fit one classifier per class. Initially, we limit the vocabulary of the model to 450 features, including both unigrams (single words) and bigrams (two consecutive words). In stage two, we take the 70 most significant features from the initial classifiers and retrain a model with this limited vocabulary to reduce noise in the text features (a list of the top 40 features, along with their respective coefficients can be seen in Table F7). All parameters were optimized using GridSearchCV, and the full configuraion can be seen in Table **??**.

- Fourth, after evaluating performance of the model with a 10 fold cross validation, (see results in Table F6) we train a final classifier on the entire labeled corpus and use it to classify our entire sample.

Although we initially classified our sample into three categories, quant, qual/normative, and formal, we wanted to distinguish between qualitative and normative papers. After achieving poor results on this task using a Logistic Regression, we resorted to the use of OpenAI's GPT-4o-Mini large language model alongside a custom prompt (see Prompt 1) to classify these papers. Specifically, we fed GPT the title, abstract, and first 2,000 characters of each paper previously classified as "qual/normative". Ultimately, this left us with the papers in our sample classified into four final categories: quantitative, qualitative, normative, and formal.

To evaluate the performance of GPT, we again used the manually coded sample from Teele and Thelen (2017), but added to this labeled data by a) taking papers from the journals *International Theory*, *Journal of Political Philosophy*, *Philosophy and Public Affairs*, *Contemporary Political Theory*, *Political Theory*, *Journal of Political Philosophy* on the assumption that they were normative, and b) adding additional papers labeled as qualitative from the TRIP Journal Article Database (*TRIP Journal Article Database Release (Version 3.3)*. 2020; Maliniak and Tierney 2018). In total, the resulted in a test set with 2,149 papers labeled as normative, and 422 papers labeled as qualitative. The results of the classification are shown in Table F8.

Category	Precision	Recall	F1-score	Support
Quantitative	0.93	0.96	0.95	1283
Formal	0.88	0.65	0.75	118
Qual/Theory	0.85	0.81	0.83	293
Micro avg	0.92	0.92	0.92	1694
Macro avg	0.88	0.81	0.84	1694
Weighted avg	0.91	0.92	0.91	1694
Samples avg	0.92	0.92	0.92	1694

Table F6: LogReg Method Classification Report

Quant		Formal		Qual/Norm	
equilibrium	-3.89	equilibrium	5.57	model	-2.66
variables	2.91	model	2.53	results	-2.56
results	2.86	game	1.93	effect	-2.42
variable	2.83	bargaining	1.79	variable	-2.30
data	2.57	cost	1.78	variables	-2.19
effect	2.38	data	-1.52	justice	1.77
respondents	2.11	political	-1.41	social	1.76
likely	2.09	probability	1.30	interests	1.73
case	-1.98	variables	-1.28	data	-1.73
table	1.94	rule	1.09	new	1.67
attitudes	1.77	variable	-1.05	causal	1.63
effects	1.73	respondents	-1.02	respondents	-1.62
sample	1.72	war	1.01	good	1.59
good	-1.64	significant	-1.00	likely	-1.58
new	-1.62	subjects	-0.97	attitudes	-1.55
example	-1.59	democratic	-0.96	did	1.54
social	-1.57	table	-0.94	probability	-1.49
measure	1.51	effects	-0.93	theory	1.42
causal	-1.50	case	0.93	equilibrium	-1.42
cost	-1.44	type	0.90	table	-1.41
using	1.43	public	-0.90	sample	-1.40
subjects	1.39	network	0.88	people	1.39
strategy	-1.37	results	-0.88	way	1.38
justice	-1.36	target	0.87	human	1.38
estimates	1.36	control	-0.85	effects	-1.35
ndings	1.31	treatment	-0.84	like	1.32
statistically	1.28	beliefs	0.82	liberal	1.28
interests	-1.27	strategy	0.82	state	1.25
way	-1.26	lower	0.81	society	1.22
signicant	1.26	people	-0.80	power	1.22
significant	1.24	equation	0.78	market	1.20
theory	-1.23	result	0.76	organizations	1.19
impact	1.22	elections	-0.76	example	1.19
rule	-1.19	investment	0.76	reform	1.19
levels	1.16	women	-0.76	view	1.17
make	-1.15	partisan	-0.75	law	1.16
panel	1.14	terrorist	0.74	scholars	1.16
year	1.14	using	-0.74	crisis	1.16
problem	-1.13	estimates	-0.74	problem	1.15
private	-1.12	median	0.73	case	1.15

Table F7: Top 40 words by classification coefficient for each method category

Prompt 1: Qualitative v. Normative GPT Prompt

```
"""
<!--begin excerpt-->
Title: {TITLE}
Abstract: {ABSTRACT}
FIRST_2000_CHARS: {FIRST_2000_CHARS}
<!--end excerpt-->
Your task is to analyze the above excerpts from a political science paper. After reading them carefully, you should construct a block of json metadata for the paper according to the following guidelines:
- **paperType**: Determine if this is a "normative paper or a "qualitative" (empirical) paper.
  - **normative** indicates the paper is primarily focused on theoretical frameworks, moral or ethical arguments, prescriptive claims, or conceptual innovations without reliance on empirical data or case studies.
  - **qualitative** indicates an empirical (data-driven) study, which may include interviews, case studies, comparative analysis, content analysis, historical data, or meta-reviews of existing literature.

### Guidelines for Classification
1. **Normative (pure theory / moral arguments)**
   - Papers that primarily develop or critique theories, concepts, or frameworks in a philosophical or moral sense.
   - They often involve ethical, moral, legal, or cultural judgments about right and wrong, good and bad, or appropriate and inappropriate.
   - They may pose 'what ought to be' questions, emphasize values or justice, or propose normative principles, prescribing what should be done or believed.
   - They usually rely on logic, reasoned argumentation, and conceptual analysis rather than data collection, interviews, or observational findings.
2. **Qualitative (empirical)**
   - Research designed to make descriptive or explanatory inferences based on empirical information about the world, descriptive or explanatory (connecting causes and effects) in nature
   - Papers that draw on real-world data, case analyses, interviews, focus groups, participant observation, archival research, ethnography, historical narratives, content and thematic analysis, or other forms of empirical research
   - They may investigate political phenomena, test theories using specific evidence, or compare policy outcomes in different contexts.
   - Even if the paper references theories, the focus is on evidence-based conclusions, case studies, or empirical findings. Meta-reviews also fall under this category if they synthesize existing empirical studies.

### Special Considerations
- A paper can reference normative elements (e.g., an ethical framework) while
```

```
still featuring extensive empirical analysis. If the primary emphasis is on data or real-world evidence, classify as **qualitative**.
- A paper can reference empirical examples as background, but if the central argument is a moral or theoretical proposition with minimal systematic evidence, classify as **normative**.
- Focus on the paper's overall purpose, tone, and evidence usage to make your final classification.
Output your response as valid json in the following structure (with very short sentence justification stored in a **justification** variable if needed if needed in nuanced cases):
"""
```

Class	Precision	Recall	F1-Score	Support
isQualitative	0.74	0.87	0.80	422
isNormative	0.97	0.94	0.96	2149
Accuracy		0.93		2571
Macro Avg	0.85	0.90	0.88	2571
Weighted Avg	0.93	0.93	0.93	2571

Table F8: Qualitative v. Normative Classification Report

Paper Submethods/Research Designs

We were also interested in classified quantitative and qualitative papers into a more fine grained measure of popular submethods or research designs (e.g. Diff-In-Diff, RDD, Ethnography, etc.). Since these submethods, when used, use very specific language, we used a relatively simple keyword approach to categorize papers. We categorized each paper into the research designs listed in Table F9, utilizing the corresponding sets of main and sub keywords.

For quantitative designs, we searched the full text of each paper for any occurrence of the corresponding main and sub keywords, only labeling the paper as that design if *both* the main and sub keywords were mentioned. The rational behind this approach is based on the assumption that although papers might use the main keywords without actually using that design (for example, a research might cite a paper which uses a "survey experiment" without implementing one themselves), they are less likely to use both the main and sub keywords (for example, the aforementioned researcher, though they cite a "survey experiment," are much less likely to also mention "randomized question order" or "balance test" unless they are using the design themselves). For qualitative designs, we only used a single set of keywords under the assumption that the language in qualitative research tends to be more varied and context-dependent than in quantitative studies, making it less amenable to a strict main-and-sub keyword structure.

Finally, to ensure that we weren't merely coding papers that mentioned the keywords without actually using the research design in question, we leveraged GPT to filter out false positive results. Taking a 500-character context window around each keyword match, we feed those excerpts, along with the paper abstract, to GPT using Prompt 2.

Table F9: Quantitative and Qualitative Research Design Keywords

Design	**Main Keywords**	**Sub Keywords**
	Quantitative Methods	
Difference-In-Difference	*difference-in-difference, DiD approach, two-way fixed effects, DiD design*	*parallel trends, common trends assumption, parallel trend holds, time x treatment interaction, test for parallel trends, counterfactual trends*
Instrumental Variables	*instrumental variable*	*2SLS, two-stage least squares, exclusion restriction, endogenous, endogeneity, instrument validity, overidentification, instrument, first stage, second stage, monotonicity*
Field Experiment	*RCT, randomized controlled trial, randomized control trial, field experiment*	*balance test, treatment group, random assignment, compliance, compilers, CACE, ATE, SATE, ITT, spillover effect, noncompliance, attrition, random walk, enumeration, enumerators, endline survey*
Regression Discontinuity Design	*RDD, regression discontinuity, RD design*	*running variable, threshold, bandwidth, forcing variable*
Event Study	*event study*	*event window, pre-event period, post-event period, dynamic treatment effects, relative time indicator, placebo, parallel trends*
Synthetic Control	*synthetic control, synthetic group*	*donor pool, Abadie, weights, weighting, reweighted, donor, Athey*
Survey Experiment	*survey experiment, conjoint experiment, list experiment, conjoint survey experiment*	*embedded experiment, randomized question order, choice-based conjoint, paired conjoint design, attribute-based conjoint, fully randomized conjoint, balance test, SATE, attention check, satisficers, Qualtrics, question wording*

Design	Main Keywords	Sub Keywords
Matching	*propensity score matching, matching estimator, nearest neighbor matching, coarsened exact matching, ATE matching, genetic matching*	*covariate balance, common support, Mahalanobis distance*
	Qualitative Methods	
Process Tracing	*process tracing, within-case causal mechanism, causal process observations, mechanistic evidence, sequential analysis*	—
Qualitative Comparative Analysis	*qualitative comparative analysis, QCA, csQCA, fsQCA, set-theoretic methods, configurational analysis, causal chain*	—
Critical Discourse Analysis	*critical discourse analysis, CDA, Foucauldian discourse, language-power relations, Fairclough approach, discourse-historical method*	—

Design	Main Keywords	Sub Keywords
Ethnography	*political ethnography, ethnography, field immersion, participant observation, thick description, contextual fieldwork, narrative field notes, in-depth field engagement, emic perspective*	—
Participatory Action Research	*participatory action research, community-based research, co-creation approach, collective inquiry, emancipatory methodology, action-oriented research, PAR cycle*	—

Class	Precision	Recall	F1-score	Support
Field Experiment	0.66	0.95	0.78	40
Survey Experiment	0.89	0.65	0.75	79
Neither	0.74	0.78	0.76	80
Accuracy	—	—	0.76	199
Macro Avg	0.76	0.79	0.76	199
Weighted Avg	0.78	0.76	0.76	199

Table F10: A comparison of field & survey experiments hand coded by undergraduate research assistants to test overlap between field and survey experiments.

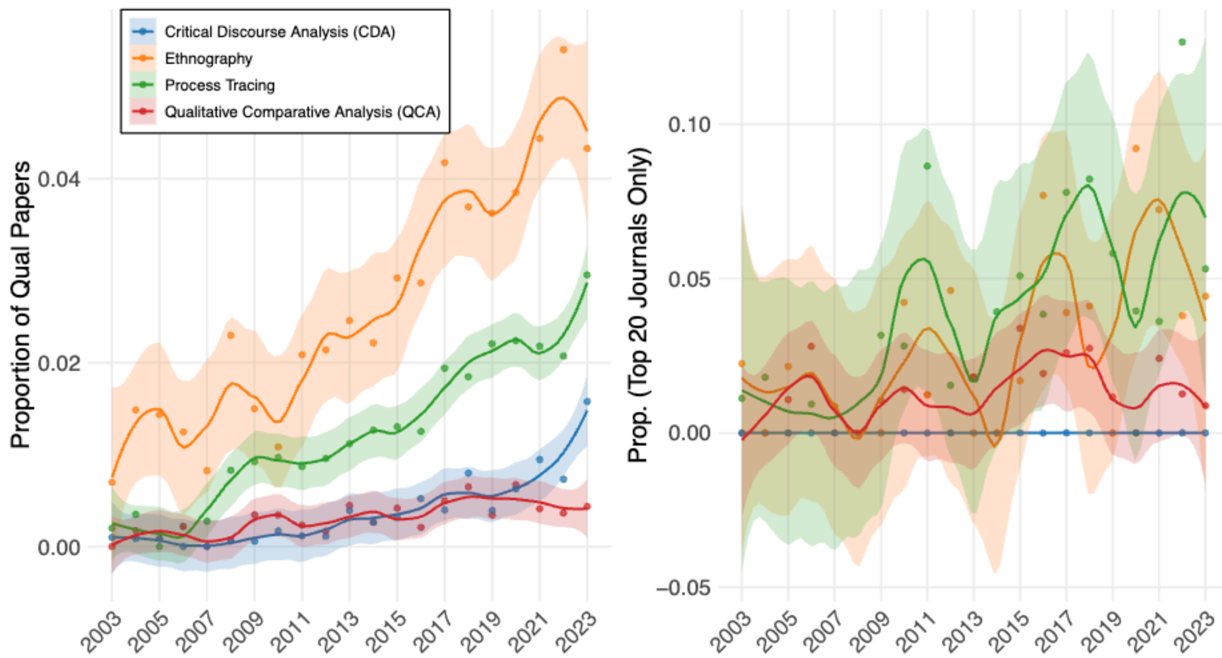

Figure F3: Figure shows the proportion of qualitative papers coded as using each research design over time. The left panel shows this value for all qualitative papers in our sample. The right panel shows results for only the top 20 journals in our sample.

Prompt 2: Prompt For Checking Submeths True Positives

```
<!--begin excerpt-->
Predicted Method: {PREDICTED_METHOD}, {PREDICTED_METHOD_DESCRIPTION}
================
Paper Abstract:
{ABSTRACT}
================
Excerpts & Keywords:
{EXCERPTS_1000_CHARS}

<!--end excerpt-->
Your goal is to analyze the excerpts above and determine whether the paper under discussion genuinely applies the predicted method ({PREDICTED_METHOD}) in its research design, or whether it merely references the method in passing (for example, citing other papers that used it or explaining why the authors themselves did not adopt it). The decision hinges on explicit or implied language indicating the actual use of the method for data collection or analysis.

Guidelines for Determination
Method is Actually Used
- The paper explicitly states it employs this method in collecting or analyzing data.
- The excerpt points to direct application: e.g., we conducted an RDD study using or the analysis follows a Diff-in-Diff approach
- Clear mention of how data or evidence is gathered or processed with the specified method.
Method is Not Used (Only Referenced or Rejected)
- The excerpt indicates the authors are describing or critiquing how others have used the method, without applying it themselves.
- The authors mention the method as a possibility but ultimately report not implementing it.
- Discussion focuses on theoretical explanations of the method or historical references rather than application.

Output the response as valid JSON. Incorporate a short justification in a justification field only if needed to clarify the decision. The structure should be:

'methodConfirmed': 'Used' or 'Not_Used', 'justification': '... short note if needed ...'
```

G Focus & Novelty

G.1 Index Construction

In this section, we explain how the topical focus and topical novelty indices are constructed using the STM output.

Focus Index

To measure the degree to which a paper is topically focused, we use the STM estimated proportion $\theta_{i,t}$ of each topic t in a given paper i. We then square these proportions, $\theta_{i,t}^2$, and sum the squared values for each paper. The resulting focus index is given by:

$$\text{Focus Index}_i = \sum_t \theta_{i,t}^2$$

Higher values indicate a greater concentration on fewer topics.

Novelty Index

To assess how novel a paper's topic combination is, we first identify the two most prevalent topics in each paper according to the STM output. Next, we define three-year rolling windows and count the total number of papers published on each topic within each rolling window. Using these counts, we determine the total number of papers published on each topic within the window.

To estimate how many papers would be expected to contain the identified topic combination by chance, we use the following formula:

$$E(T_a, T_b) = \frac{N_a \times N_b}{N_{\text{total}}}$$

where N_a is the number of papers on Topic a in the window, N_b is the number of papers on Topic b in the window, N_{total} is the total number of papers published in the window, and $E(T_a, T_b)$ represents the expected number of papers that contain both topics by chance.

For example, if 10 papers were published on Topic 1, 20 on Topic 2, and 100 papers in total within a three-year window, the expected number of papers containing both topics by chance is:

$$E(T_1, T_2) = \frac{10 \times 20}{100} = 2$$

The ratio of the actual number of observed papers with the topic combination to the expected number is computed as:

$$R(T_a, T_b) = \frac{O(T_a, T_b)}{E(T_a, T_b)}$$

where $O(T_a, T_b)$ is the observed number of papers with the topic combination. Finally, the novelty index is calculated as:

$$\text{Novelty Index} = 1 - R(T_a, T_b)$$

A higher novelty score indicates that a paper's topic combination is less common than expected by chance, relative to the last three years of publications.

G.2 Focus, Novelty, and Paper-Level Performance

In Table 3 we show that topical focus and topical novelty are systematically related to higher citations, relative to other papers published in the same year and the same journal. We further show there is not systematic relationship between novelty and the probability that a paper is published in a Top 20 outlet. In the following Table (Table G11), we estimate the association between focus, novelty, and a paper's citations. However, we compare paper-level citations with those of other papers published in the same year, regardless of the outlet where they were published (by standardizing the citation count within year). Results for topical focus are consistent in magnitude and positive, suggesting focus is positively associated with paper-level performance overall. However, results for novelty are estimated to be precisely zero. These results suggest that while, relative to other papers published that year and in the same outlet, more novel papers get more citations after accounting for authors' quality, it is not the case that relative to other papers published that year overall, more novel papers get more citations, after accounting for authors' quality.

	Citations (year-std)	Top 20	Citations (year-std)	Top 20
Novelty	0.000	0.002		
	(0.006)	(0.002)		
Focus			0.059***	0.002
			(0.010)	(0.003)
Author FE	✓	✓	✓	✓
Num.Obs.	166 387	166 387	166 387	166 387
R2	0.419	0.542	0.461	0.542

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table G11: Table shows the estimated change in paper citations (standardized within year) when a paper's novelty or focus index and the estimated change in the probability that paper i is published in a Top 20 outlet (as per their SJR Impact ranking, see Table B1 in the Appendix for details.). Robust standard errors, clustered at the author and paper level, reported in parentheses.

G.3 Team Composition and Novelty

In Figure 11 we show that all female teams produce the most novel research. In this section we probe deeper into this result. In Figure G.3 we plot the proportion of papers published by teams, according to their gender composition, that have each topic as one of their two most common topics. Topics are arranged from largest difference between all-male and all-female teams to smallest. As can be seen, all female teams publish most often on Critical Theory, followed by Gender and LGBT+. In contrast, all male teams publish most frequently in Critical Theory, Quantitative Methods, and Security Studies. Mixed gender teams most often write papers about Public Opinion and Political Psychology.

In the main paper, we find that novelty is positively associated with paper-level success, measured as the standardized number of citations within a journal-year, but not when citations are standardized within years. We argue, based on extant literature, that novelty could potentially be a "high-risk, high-reward" strategy that pays off in the long run rather than the short run.

In plot G.3, we examine whether novelty is, in fact, correlated with long-term paper-level success. The left panel shows that, regardless of a paper's novelty, the mean percentile rank in citations for highly novel papers (in purple) and non-novel papers (in red) starts off quite similar within the first few years of publication. However, by the 10th year, papers ranked in the lowest quantile of novelty plateau and actually decline in ranking. Conversely, highly novel papers become more highly cited over time.

The same pattern is evident in the right panel. For the first 10 years after publication, novelty appears to be orthogonal to raw citation count. However, after the 10th year, the

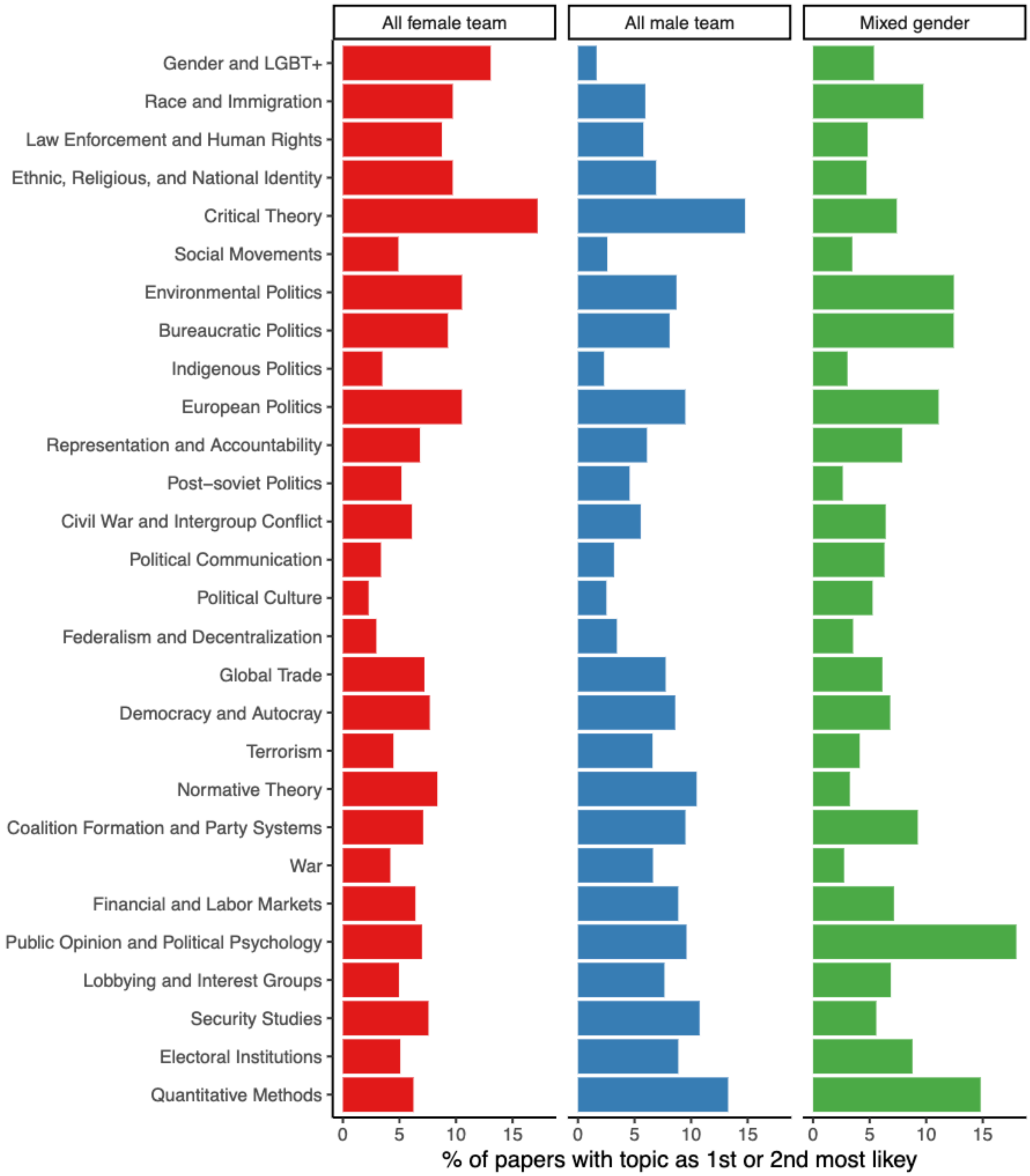

Figure G4: Figure shows the percentage of papers written by all female teams (in red), all male teams(in blue), and mixed gender teams (in green) with each topic as the 1st or 2nd most common

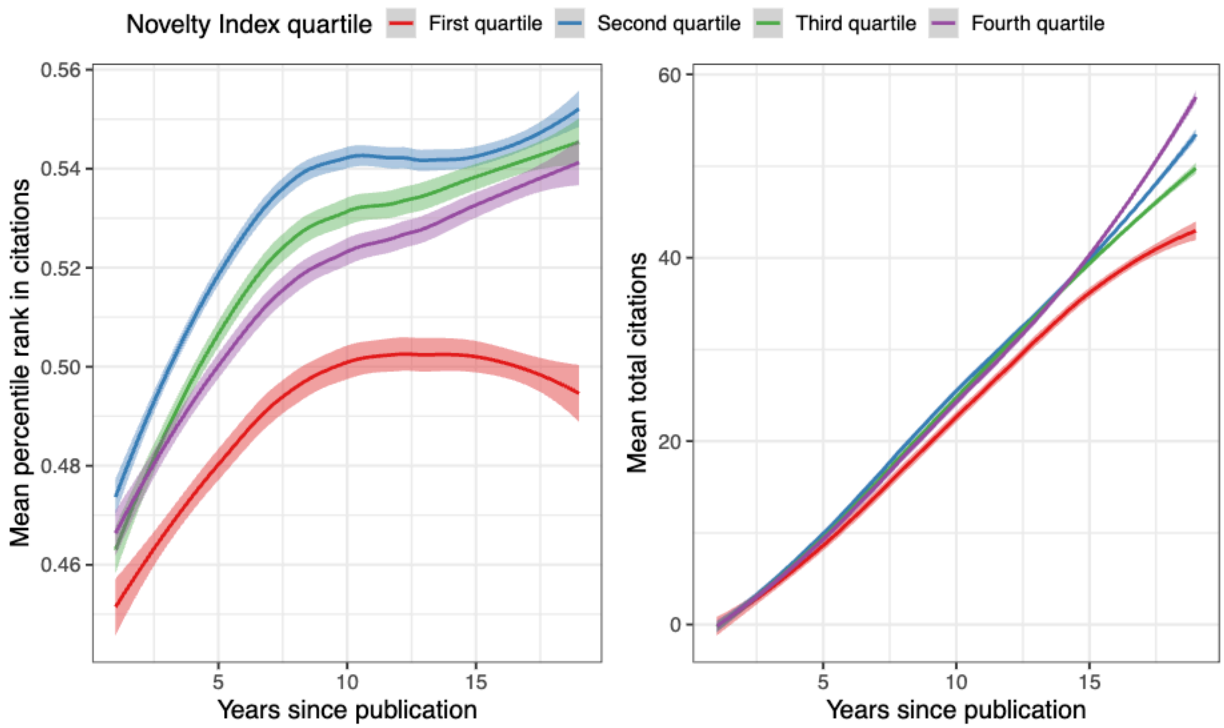

Figure G5: The left panel shows the mean percentile rank in citations of papers in the first (red), second (blue), third (green), and fourth (purple) quartile of novelty as years since publication increases. The right panel shows the mean raw count of total citations for papers in each of these two groups as years since publication increases.

most novel group of papers begins to be slightly more cited, while the least novel group plateaus. By the 15th year, a noticeable gap emerges, where novelty is positively associated with citation count.