# Over a century of economics research collaboration

Willy Chen[*] Xiao Qiao[†] Hanzhe Zhang[‡]

March 21, 2025[§]

**Abstract**

We document patterns of research collaboration among academic economists using 140 years of bibliographic information of publications and working papers, beginning in 1886. We document an increasing number of authors per paper and an increasing fraction of multi-authored papers. Collaboration across institutions first decreased and then increased in the past 25 years, while collaboration across economists of different experience levels remained stable over time. We employ a simple framework to study the benefits and costs of collaboration. We find a benefit from collaboration: The most-cited papers are increasingly written by larger research teams. Using COVID as a natural experiment, we find a polarizing effect from the increased costs of collaboration: While some researchers retreated and worked more by themselves, others took advantage of the remote work tools and teamed up with more people.

**Keywords:** collaboration, economics, COVID, working papers

**JEL:** A1

[*]Department of Economics, Michigan State University, willyc@msu.edu.

[†]Department of Data Science, City University of Hong Kong, xiaoqiao@cityu.edu.hk.

[‡]Department of Economics, Michigan State University, hanzhe@msu.edu.

[§]We thank Valentin Bolotnyy, Steven Haider, Alec Kirkley, Wendy Stock, Jeff Wooldridge, and Mo Xiao, as well as seminar participants at Amazon, American Economic Association annual meeting, Michigan State University, and Midwest Economics Association annual meeting for helpful suggestions. We thank financial support from Amazon and National Science Foundation.

# 1 Introduction

The production of a research paper has undergone tremendous transformation in past decades. Most notably, the number of authors on research papers has steadily increased over time; *inter alia*, Guimera et al. (2005) and Wuchty, Jones, and Uzzi (2007) document increased collaboration in many fields across both the natural sciences and the social sciences. Despite ample evidence of increased collaboration, there is limited agreement among researchers on the underlying drivers of this shift.[1] Explanations of collaboration patterns mostly remain descriptive and suggestive, and few attempts have been made to formally test hypotheses about the sources of rising teamwork among academic economists.

This paper sets out to accomplish two goals. First, we aim to document more comprehensive—more longitudinal and more granular—facts about research collaboration to establish a wider and deeper understanding of this phenomenon. Whereas existing papers tend to focus on recent decades when studying collaboration patterns, our investigation expands this analysis to the beginning of modern economics research in the nineteenth century: For example, *Quarterly Journal of Economics* was founded in 1886 and *Journal Political Economy* was founded in 1891. We also explore collaboration patterns among economists of different experience levels and across academic institutions to shed light on additional dimensions of research cooperation. Second, we examine potential drivers of collaboration patterns. We seek to understand the choice an economist faces: working alone or collaborating with others. Two main features arise in this framework: the cost associated with the production of a paper, and the quality of the paper which we refer to as returns. We study how variation in cost and return can produce changes in collaboration patterns.

A key challenge standing in the way of our research questions is a lack of quality data. Extant work often relies on data compromised by misclassification, inaccuracies, and omissions. In some cases, overly inclusive classification of economics papers likely results in a dataset only loosely related to economics research.[2] In other cases, author affiliation can be stale or is simply missing from the database. To overcome this issue, we meticulously put together a database of economics research papers and author information. We start from 64 journals representative of economics research, including general-interest and field-specific journals. We compile raw bibliometric data from published papers in these journals via OpenAlex, a large bibliographic catalog of research papers. In addition, we obtain working papers from

[1] Increases in team size have been attributed to, among others, improvements in communication technology (Jones, Wuchty, and Uzzi, 2008; Xie, 2014), increased specialization of research areas (Hunter and Leahey, 2008), and rising complexity of modern problems that require interdisciplinary solutions (Milojević, 2014; Falk-Krzesinski et al., 2011).

[2] For example, Jones (2021) counts articles from over 3,000 journals as economics-focused.

the Social Science Research Network (SSRN) via OpenAlex, and those from the National Bureau of Economic Research (NBER) directly on its website. This comprehensive database provides a consistent definition of economics papers and academic economists, enabling us to address our research questions in a consistent environment. Our assembly of data consists of not only publications and working papers, but also author affiliation information over time. Because a researcher may move from one institution to another, accurately capturing changing affiliations is especially important in our study of inter-institutional collaboration.

Our empirical findings reveal that increased collaboration first began in the 1950s. We find a secular trend toward more collaboration arriving in waves: a shift from primarily single-authored papers in the 1950s and 1960s to two-authored papers in the 1970s, three-authored papers in the 1990s, and to four-authored papers in the 2010s. The proportion of multi-authored papers, as well as the average number of authors among these papers, rises over time. A more granular view reveals that inter-institutional collaboration initially decreases, then increases in the most recent 25 years to reflect greater cooperation across institutions.[3] Conditional on the number of authors on papers, experience assortativity—whether researchers tend to work with those with comparable levels of experience—remains unchanged for much of our sample. These results extend and expand existing findings to a broader set of papers (working and published) going further back in time, as well as reveal new dimensions to collaboration.

Next, we study the drivers of these patterns. We employ a simple random utility model with two main components, cost and return. The cost term encompasses all factors contributing to the overall expense of producing a paper, including time, opportunity, and financial costs. Return captures the impact of the paper, for which we use the trailing five-year citation count (compared to other papers in the same year) as the empirical proxy.

We consider the COVID pandemic as a natural experiment with an immediately large impact on the cost but a limited effect on other aspects of paper production. While there may be changes to returns or to idiosyncratic preferences to collaborate, we assume they are small relative to changes in costs. We find positive deviations from trend in the shares of single-authored papers as well as papers with four or more authors during COVID, indicating that the pandemic had a polarizing effect. While some researchers retreated and worked more by themselves, others took advantage of the novel productivity tools and normalized remote work culture to team up with more people. In terms of inter-institutional collaboration, we find little evidence of changes during COVID compared to the prior period.

[3] Jones, Wuchty, and Uzzi (2008) find that research is increasingly done across universities. Using a longer sample, we demonstrate that there was an earlier decrease in inter-institutional collaboration before a recent increase.

While COVID serves as a plausible cost shifter, it does not provide an informative event for returns. We define a homerun paper as a publication whose trailing five-year citation count is in the top 10% among all economics papers published in the same year. Regressions of homerun probabilities on time fixed effects, controlling for other paper characteristics, show that returns to multi-authored papers change through time. While single-authored papers are the most likely to be highly cited in the 1950s, two-authored papers are the most likely in the 1960s and 1970s. Three-authored papers are the most likely to be homerun in the 1990s and 2000s, only to be overtaken by four-plus-authored papers in the 2010s. Evidently, returns to collaborative work have increased over time.

We investigate whether returns may be a driving factor in research collaboration. Our hypothesis is that economists respond to higher returns to multi-authored papers by writing more papers in teams. We test this hypothesis through a statistical model with the fraction of $N$-authored papers as the dependent variable and lagged shares of homerun papers as the independent variable. Following the seemingly unrelated regression (SUR) methodology of Tomz, Tucker, and Wittenberg (2002), we find that researchers tend to assemble into larger teams of four or more co-authors following higher returns, while they do not respond much to increased returns to two-authored or three-authored papers.

Our parsimonious model provides structure to our empirical analysis. Minimal structure allows us to proceed from a description of collaboration to a potential explanation, en route to testing more hypotheses. Our expanded and more granular empirical documentation of research collaboration patterns also encourages additional theoretical explanations. Stylized facts related to the cross-sectional variation in experience levels and institutional affiliation must be accounted for in newly proposed theories of collaboration.

After a literature review, the rest of the paper is organized as follows. Section 2 describes the construction of our unique database of author and paper information. Section 3 documents patterns of collaboration among economics researchers. Section 4 introduces a simple framework to understand collaboration patterns and conducts empirical tests of possible channels. Section 5 concludes.

## 1.1 Literature Review

Our paper contributes to three strands of literature: (i) documenting facts in research collaboration, (ii) empirically testing the mechanisms of research collaboration, and (iii) causally identifying the effect of COVID on productivity and collaboration.

Several papers document a rise in research collaboration across scientific fields. Wuchty, Jones, and Uzzi (2007) illustrate an increase in the number of authors on a research paper

for both the natural sciences and the social sciences. Uzzi et al. (2013) compare the research outputs of research teams and individuals and find that teams generate more creative ideas. As research fields become increasingly specialized, researchers tend to train longer and receive their doctoral degrees later (Jones, 2009, 2010). These studies tend to measure collaboration intensity by the number of coauthors, offering a broad view of trends across disciplines, with less focus placed on the composition of research teams. Compared to previous work, we analyze collaboration trends using the most comprehensive database of economics researchers spanning 140 years, and we document facts not only about the number of collaborators, but also about the makeup of teams. Our more comprehensive data allows us to account for variables such as author seniority, institutional rankings, and location of affiliation at a finer level of granularity compared to extant work limited by data quality and availability.

In contrast to ample descriptive studies of research collaboration, there is less of a focus on empirical tests of formal hypotheses. Milojević (2014) asserts that there is no model that describes the size of a research team and proposes one to capture its cross-sectional distribution. With a focus on the physical sciences and patents, Ahmadpoor and Jones (2019) find that teams tend to assemble among authors with similar citation levels, providing corroborating evidence for positive assortative matching. In a comprehensive survey, Liu et al. (2023) classify empirical work into two categories: (1) to discover and estimate empirical regularities and (2) to identify the underlying mechanism. The authors discuss significantly more papers in the former category than the latter. Our paper adds to each of these two categories, with an emphasis on testing the implications of a simple framework.

A number of papers examine the effect of COVID on workplace behavior (Butler and Jaffe, 2021; Bayhan et al., 2022; Ford et al., 2021; Yang et al., 2022; Dong et al., 2023; Jain et al., 2024). Survey-based research highlights a differential impact of the pandemic on genders, suggesting heightened challenges for women, especially those with caregiving responsibilities (King and Frederickson, 2021; Liu et al., 2022; Sinatra et al., 2023). Heo et al. (2022) collect more than 2,000 surveys from 100 countries and identify delays in STEM research reliant on labs and fieldwork, a disproportionate burden on female scientists, and reduced opportunities for informal collaboration due to the shift to virtual conferences. Whereas existing work tends to focus on COVID itself as the event of interest, we treat it as a laboratory to examine how costs can change collaboration patterns, using the onset of the pandemic as a test of our framework on collaboration choices.

# 2 Data

We gather both working papers and publications. We obtain paper-level information from the bibliometric database OpenAlex, which includes publications and working papers posted on the Social Science Research Network (SSRN). SSRN working papers are available from its inception in 1994 to 2023. Working papers posted on the National Bureau of Economic Research (NBER) from the beginning of the working papers in 1973 to 2023 are also collected. For publications, we consider papers published in 64 journals commonly considered to be the most prominent economics research outlets including general-interest journals (e.g., *Journal of Political Economy*, *Economic Journal*, and *Economic Inquiry*) and field journals (e.g., *Journal of Development Economics*, *Games and Economic Behavior*, and *Social Choice and Welfare*). Table 2 lists the 64 journals by field.[4] Publications are available since 1886, when the *Quarterly Journal of Economics* was founded. We will refer to all publications among the 64 journals as "EC64", and the top five general interest journals (*American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*) as "top five." We refer to anyone who has published an EC64 paper as an "academic economist."

We record the year of publication or posting of working papers in repositories.[5] Since the dating is precise only up to the year, we define the pre-COVID period as all years through 2019 and the post-COVID period as 2021 and beyond. To ensure our focus on economics research and valid comparison before and after COVID, we exclude papers primarily about the COVID pandemic by filtering out any titles that include "COVID," "coronavirus," or "sars-cov-2." We also distinguish between papers in which all authors share a single institutional affiliation—indicating intra-institutional collaboration—and those with multiple affiliations, which indicate inter-institutional collaboration. In our analysis, a paper's "major affiliation" is defined as the most frequently occurring affiliation among its authors.

## 2.1 Author Affiliation Records

Author affiliation through time is necessary for studying collaboration patterns across institutions. Existing datasets typically do not contain a time series of affiliation information for each author. Furthermore, available affiliation from bibliographic catalogs may be incorrect or stale. This problem is worse for working papers, since repositories such as SSRN

[4]The list of journals comes from the School of Economics at Shanghai University of Finance and Economics and is consistent with the top journals in ranking (Kalaitzidakis, Mamuneas, and Stengos, 2003; Heckman and Moktan, 2020; Ham, Wright, and Ye, 2021), and is more comprehensive compared to lists maintained by other institutions (e.g., Cabral, 2020; Tilburg University, 2025).

[5]The date of many papers appears as January 1st in OpenAlex.

often only retain the latest affiliation information and overwrite past positions even for works posted while the author was in those positions.[6] To overcome this issue, we construct author-affiliation records using published papers, which contain the correct affiliation on the date of publication. By chaining together multiple publications, we are able to construct a panel data of author affiliations. Our approach follows the framework proposed by Lin et al. (2023) for completing institutional data from the Microsoft Academic Graph (MAG), a large database containing scientific publication records, citation relationships, and author information.

MAG stopped its updates in December 2021. The non-profit organization OurResearch proposed an open access bibliographic database to continue MAG called OpenAlex. As of March 2024, OpenAlex includes metadata for over 200 million papers and books, 13 million authors, and metadata for over 100,000 institutions. Given its comprehensive coverage, OpenAlex has been used by universities to gauge the progress of their research teams.

To construct author-affiliation records, we obtain all articles indexed in OpenAlex that are published in the 64 economics journals we consider. We pull relevant paper information from OpenAlex to construct records of each economist's affiliation and fill in any gaps using the steps described below. To process author-affiliation data with missing records, we implement the following procedure.

1. **Identify active years**. For each author-affiliation pair, determine the range of years with existing records. Specifically, note the first recorded year as `minyear` and the last as `maxyear`.

2. **Forward fill missing years**: For each affiliation, fill in any missing years between `minyear` and `maxyear` using forward filling. This technique replaces missing values with the last observed non-missing value, ensuring continuity in the data.

3. **Retain most recent forward-filled entry**: For years that contain only filled records, retain the most recent forward-filled entry. This approach ensures that the data reflects the latest known affiliation information.

By implementing these steps, we can achieve a consistent and accurate representation of affiliations over time, effectively addressing gaps due to missing records. Our methodology of affiliation construction relies on authors having active publication histories. For authors who

[6]For example, suppose that author A was affiliated with institution I from 2005 to 2010 and institution J from 2011 to 2020. If author A posted a working paper on SSRN in 2008 (while in institution I), but the information was not scraped until 2018, then it is more than likely that OpenAlex would record that this paper was published with author A from institution J instead of institution I. We exclude the affiliation records for NBER and IZA, due to the inter-institutional nature of these organizations.

do not have publications beyond 2018, we forward fill their 2018 affiliation through 2023. For any authors whose affiliation is still missing after this step, we include affiliation information from working papers posted in 2018 or later. Our approach yields a 98% complete panel of author affiliation records.[7]

We check the quality of our construction affiliations. We manually collect MIT tenure-track faculty's education and career histories to check against our generated data. For the years that an affiliation was present, our data construction yields an 85% accuracy rate. In addition, we check our dataset against the education and career histories of 200 award-winning economists collected in Freeman et al. (2024), and the results show an 87% accuracy rate.

## 2.2 Paper-Year Records

Along with author affiliation records, our dataset includes paper and author information. We construct a dataset of paper-year records including title, publication year, journal/repository, as well as authors and their affiliations. We obtain available information from OpenAlex, incorporate author affiliation, and construct additional variables. For a paper $i$ published in year $t$ in journal/repository $j$, we form the following outcome variables of interest:

- ($Num_{ijt}$) Number of authors on the paper
- Inter-institutional collaboration variables
  - ($Intra_{ijt}$) Intra-institutional collaboration dummy: An indicator variable that equals 1 if all authors share a common institution.
  - ($pct_maj_{ijt}$): The fraction of authors from the major institution, the most common affiliation for paper $i$.
- Experience variables
  - ($pct_jun_{ijt}$) Fraction of junior economists: the percent of authors that are economists whose first EC64 paper was published zero to nine years ago.
  - ($pct_sen_{ijt}$) Fraction of senior economists: the percent of authors that are economists whose first EC64 paper was published 10 or more years ago.

[7] The only authors missing an affiliation from this construction are those who do not have an affiliation listed on their working paper *and* have no publication or another working paper that listed an affiliation in a given year.

Working papers from NBER need to be linked to author profiles from OpenAlex in order to populate author affiliation records that we construct from EC64 papers. In lieu of unique identifiers that we can map between the two databases, we adopt a fuzzy matching algorithm that seeks to identify similar elements from different data sets while allowing for some small degree of imperfect matching. We fuzzy match author names from NBER working papers with those from our author affiliation records, retaining pairs of names with a Jaro-Winkler distance with $p = 0.1$ smaller than 0.05. When applied to 31,721 working papers from 1973 to 2023, 24,563 papers (77.4%) have fully matched author information, 6,591 papers (20.8%) have partial matches, and 567 papers (1.8%) have no matched author information.

Table 1 presents the summary statistics of the four sets of papers. Although working paper repositories are a more recent occurrence, with the NBER paper series starting in 1973 and SSRN in 1994, the number of papers on these platforms approaches those of published papers in the top five journals and EC64, respectively. The average number of authors per paper is somewhat higher for working papers than for publications, possibly reflecting the more recent nature of repositories. There is a great deal of collaboration across institutions in all types of papers and venues. In particular, 92.6% of NBER working papers are written by teams across institutions. There is also a large share of younger economists in published papers: More than 60% of authors in multi-authored papers have fewer than 10 years of experience publishing papers.

# 3 Stylized Facts in Economics Research Collaboration

In this section, we document stylized facts about research collaboration among economics researchers. As the COVID pandemic began in 2020, which had a significant impact on collaboration patterns, we focus on data through 2019 to study secular trends over time, and we reserve 2021 onward for analysis specific to COVID.

## 3.1 Number of Authors

Economics as a field has shown significant growth since 1886, when *Quarterly Journal of Economics* was first founded. Figure 1 illustrate that the number of economists and the number of journals (among the EC64 that we track) have increased substantially. While the number of journals hovered below 10 prior to the 1950s, many more journals were introduced between 1960 and 2000, resulting in almost our complete set of 64 journals by the early 2000s. The number of economists also increased more rapidly in the post-WWII era. In Figure 1, we show the trend of four types of economists defined as follows:

1. Assumed Active: For each economist, identify their first and last EC64 publication years. All the intervening years are assumed to be active. The number of active economists is the total count in that year.

2. Publishing: For each year, count the number of unique economists who published in EC64.

3. New: Economists who published in EC64 for the first time in that year.

4. Cumulative: For each economist, record only their first EC64 publication year. Then, create a cumulative count of economists for each year, ensuring that once an economist is added, they remain included in subsequent years.

Figure 2 plots the number of published and working papers over time. While the number of papers published in the top five general-interest journals rose steadily over time, the rate of increase of publications among all 64 journals is markedly higher, indicating a faster expansion in top-field and other general-interest journals. NBER working papers grew at a similar rate compared to EC64 publications. Since its founding in 1994, SSRN has been hosting an increasingly large number of papers, with a rate of growth that dwarfs publications or NBER working papers. This dramatic increase stabilized around 10,000 papers per year in 2012, until a spike in 2020.

Figure 3 highlights the evolution of the number of authors on a research paper over time. Single-authored papers dominated all papers prior to 1950. At the same time, it was exceedingly rare to have papers with three or more authors. The number of single-authored papers continued to rise until 1990, from which point we have seen a steady decline for more than 30 years. The popularity of multi-authored papers appear to take on waves: Beginning in the 1950s, there was a noticeable increase in the number of two-authored papers. There was a clear increase in the number of three-authored papers in the 1980s, and papers with four or more authors became more common in the 2000s.

Figure A1 illustrates similar patterns when we focus on papers published in the top five economics journals. Multi-authored papers became more common at approximately the same times as those of EC64 papers. Publications in top five journals exhibit a more extreme shift toward a greater number of collaborators, as shown by the sharp decline in single-authored papers from 400 per year to fewer than 100 per year from the 1970s to the 2010s.[8]

Breaking up these growths by the number of authors in the paper, Figure 6 shows that much of this growth may be attributed to the growing number of papers with three or more

[8]Until 2017, the May issue of the AER each year includes papers presented at the American Economic Association's annual meeting that January, increasing the overall paper count for top five journals. Figure A4 presents the result without the AER, which shows the same patterns over time with fewer papers.

authors. Figure 7 presents the same breakdown by the share of papers with $N$ authors in each set of papers. Trends in top five papers may be difficult to see in Figure 6 because of the volume of papers, but it is clear in Figure 7 that the share of papers with three or more authors has increased in all samples.

We analyze changes in the composition of research papers through a linear trend specification:

$$100 \cdot \mathbb{1}\{Num_{ist} = n\} = \beta_0 + \beta_{1,n} t + \alpha_s + \varepsilon_{ist} \tag{1}$$

where $\mathbb{1}\{Num_{ist} = n\}$ is an indicator variable for whether the number of authors on paper $i$ posted at source $s$ (journal or repository name) at time $t$ is $n = 1, 2, 3, 4+$.[9] $\alpha_s$ captures differences in the number of authors across distinct outlets.

Table 4 provides the results for Equation (1). Column (1) shows that the share of single-authored papers decreases by an average of 1% for each year that passes. Column (2) indicates that two-authored papers also tend to decrease over time, falling at an average of 0.55% per year. Columns (3) and (4) show increasing shares of three- and four-plus-authored articles, which grow at an average of 0.88% and 0.67% per year, respectively. NBER, SSRN, and top-five dummy variables suggest that working papers and the most selective published papers tend to be more collaborative—i.e., more authors—compared to the universe of papers published in EC64 journals. To see if there is any discrepancy across the four sets of papers, we repeat our analysis by estimating Equation (1) separately for each set of papers in Table B1. While there exists some variation in point estimates across datasets, we continue to observe the salient pattern that *research production has shifted from an individual exercise to a team effort.*

## 3.2 Collaboration across Institutions

As research teams become larger over time, do their compositions change? The first dimension we examine relates to whether researchers mainly work with colleagues at their home institutions who are in close physical proximity, or also with those from other institutions. On the one hand, to the extent that the increase in the number of authors on articles is primarily a result of better collaboration tools and lower communication hurdles, we may expect to see more frequent cooperation across institutions over time. On the other hand, if rising collaboration is more of a product of an increasingly competitive publication environment, we do not necessarily expect to see changes in teamwork across institutions.

[9] We group papers with four or more authors into one category: Until 2010, less than 10% of papers have four or more authors.

We study two variables related to inter-institutional collaboration. First, we calculate the share of inter-institutional papers, i.e., those that contain at least two distinct affiliations, among all multi-authored papers. Second, we compute the share of authors in the major institution, defined as the most common affiliation on a paper. Figure 8 plots the time series of the fraction of inter-institutional papers. Focusing on the period since 1950, we observe that the fraction of papers pertaining to cross-institutional cooperation declined from 1950 to 2000 from about 90% of all multi-authored papers to 60-70%. This negative trend reversed in the most recent decades, rising to 75% by 2023. In comparison, Jones, Wuchty, and Uzzi (2008) examine collaboration from 1975 to 2005 and find that scientific research is increasingly conducted across universities. The difference in findings may be attributable to our expanded sample period and the particular focus on economics research rather than across fields.

Working papers from NBER and SSRN tend to contain more inter-institutional collaboration compared to published papers. The fraction of inter-institutional papers hovers around 80% for SSRN, and 90-95% for NBER. While the share of inter-institutional NBER papers slightly increases over time, the share for SSRN papers remains mostly stable.

Figure 9 shows that, among inter-institutional papers, the fraction of authors from the major institution on a paper decreased over time. This trend is the most salient for publications from EC64 journals and top-five journals. Working papers also show decreased concentration of authors from the same major institution, albeit to a smaller degree than publications. This pattern suggests that growing collaboration may not simply be driven by old colleagues who moved to new institutions, but rather by collaborators at multiple institutions.

We quantify the magnitude of inter-institutional collaboration by estimating trend models for the fraction of inter-institutional papers and the share of authors from the major institution:

$$100 \cdot Inter_{ist} = \beta_0 + \beta_1 t + \alpha_s + \varepsilon_{ist} \tag{2}$$

$$pct_maj_{ist} = \beta_0 + \beta_1 t + \alpha_s + \varepsilon_{ist} \tag{3}$$

where $Inter_{ist}$ is an indicator variable that is 1 if all authors of the paper have a common institution and $pct_maj_{ist}$ is the share of authors from the major institution on a paper containing at least two distinct affiliations. The right-hand variables are defined the same way as those in Equation (1).

Table 5 presents the results estimated from 2001 to 2019. The left-hand variable in Panel A is the fraction of inter-institutional papers. Column (1) indicates a significantly positive

trend of 0.18% per year. Working papers have statistically and quantitatively higher fractions on average when compared to published papers in EC64 journals, and the top-five journals a somewhat lower fraction on average. Columns (2) and (3) separately examine the trends for working papers and publications, which reveal that the increase in inter-institutional collaboration is much more evident for published papers. The trend coefficient is six times larger for publications (an average of 0.42% increase per year) than for working papers (0.07% per year). SSRN papers have a lower fraction of inter-institutional papers compared to NBER papers.

The left-hand variable in Panel B of Table 5 is the share of authors in the major institution. Column (1) shows a significantly negative trend in this period of -0.12% per year. Working papers tend to have a lower share of authors from the same major institution, with NBER working papers having the lowest average (9.10% lower compared to EC64 papers). Papers in the top five journals also tend to have a lower share compared to the broader EC64 group. Columns (2) and (3) show that the trend is similar for working papers when compared to publications, with comparable point estimates of -0.10% and -0.14%, respectively. Across the board, it appears that papers are becoming less dominated by team members from one institution.

## 3.3 Collaboration across Experience Levels

Another dimension to the composition of a research team relates to the experience level of authors. We categorize each author into three categories: (i) *junior*: author whose first EC64 publication was nine or fewer years prior to the publication; (ii) *senior*: author whose first EC64 publication was 10 or more years prior to the publication; and (iii) *non-economist*: author who has never published a paper in EC64. Using these categories, we evaluate whether researchers tend to work with others with similar or different experience levels. A partnership between two researchers of similar experience levels may indicate a more horizontal relationship, such as a combination of distinct expertise, while teamwork across experience levels may suggest a more vertical one, such as a mentor-mentee relationship or a team with more specialized roles.

Figure 10 presents the shares of two-authored publications among EC64 journals by collaboration type: two junior co-authors, one junior and one senior, and two seniors. From 1950 onward, the share of papers written by two junior economists initially rose into the 1970s, but then declined. The share of papers by one junior and one senior economist has been consistently around 40%, and the fraction of papers produced by two senior collaborators has somewhat increased. Overall, the split among the three categories has remained

relatively stable over time.

Figure 11 plots the share of collaboration types among SSRN working papers. In addition to the three types we can define for publications, we can further construct two categories involving authors who have never published in EC64 journals (non-economists). This figure exhibits similar trends as Figure 10: A stable share of junior-senior and senior-senior collaboration, and a decreasing share of junior-junior papers. We also observe a growing proportion of working papers posted by an economist and a non-economist, suggesting an increase in interdisciplinary collaboration.[10] Figure 12 expands the analysis to all publications among EC64 journals. We continue to observe relatively stable shares of entirely senior, entirely junior, and mixed papers over time. In particular, there appears to be a limited rise in papers written by authors with seniority levels.

We estimate trend models for collaboration across experience levels:

$$100 \cdot \mathbb{1}\{pct_jun = 100\}_{ist} = \beta_0 + \beta_1 t + \alpha_s + \varepsilon_{ist} \tag{4}$$

$$100 \cdot \mathbb{1}\{pct_sen = 100\}_{ist} = \beta_0 + \beta_1 t + \alpha_s + \varepsilon_{ist} \tag{5}$$

where $\mathbb{1}\{pct_jun = 100\}_{ist}$ is an indicator variable for a paper written entirely by junior economists, and $\mathbb{1}\{pct_sen = 100\}_{ist}$ entirely by senior economists. The right-hand variables follow the same notation as that of Equation (1).

Table 7 shows that there is a decline in papers written by all junior collaborators, while there is a much smaller change for papers written by all senior collaborators. Point estimates in Column (1) demonstrate that among SSRN working papers, those with exclusively junior authors tend to decline by 0.65 percentage point per year, whereas those with exclusively senior authors decline at a smaller 0.13 percentage point each year. We observe a similar pattern for NBER working papers and publications on EC64 journals or top-five journals.

# 4 Drivers of Collaboration

So far we have documented patterns of economics research collaboration. The natural follow-up question relates to potential drivers of such trends. What factors cause changes in the number of authors or in the composition of a research team? Our investigation seeks to address these questions next.

Consider a random utility model (McFadden, 1974) in which the economist's utility from a project is comprised of returns, costs, and taste. Based on these three factors, the economist

[10] This observation could also be due to the expansion of the economics profession and more collaboration between PhD students (who have yet to publish in the 64 journals) and their supervisors.

decides to either work alone and produce a single-authored paper or collaborate with other economists and produce a multi-authored paper. The return of a single-authored paper is denoted by $r_1$ and a multi-authored paper $r_m$. Let the per-author cost of production of each type of paper be $c_1$ and $c_m$, respectively. Individual researchers may also have idiosyncratic tastes for working by themselves or with others, captured by $\varepsilon_1$ and $\varepsilon_m$. The decision for an author $j$ can then be summarized as follows. She would prefer to write a multi-authored paper over a single-authored paper if

$$r_{m,j} - c_{m,j} + \varepsilon_{m,j} > r_{1,j} - c_{1,j} + \varepsilon_{1,j}. \tag{6}$$

This model provides two directions of exploration: cost and return. The cost term encompasses all factors contributing to the production of a paper including time cost, opportunity cost, and financial cost. The return term seeks to capture the quality and impact of a paper. All else equal, researchers presumably would like to write high-quality papers that are well-regarded by the profession and leave an impact. Similar to Jones (2021), we use the trailing five-year citation count as the empirical proxy for the return of a paper. Because the number of researchers changes over time, two equally influential papers written at different points in time may receive different numbers of citations. To provide a fair comparison across time, we compare citations with other papers published in the same year, and define highly influential papers as those that rank in the top decile of their respective cohorts.[11]

We use Equation (6) as a guide to understand the empirical patterns in collaboration. Suppose returns were held constant, then any changes in collaboration must arise due to variations in the relative cost of working alone versus with co-authors. If instead costs were held constant, then collaboration patterns must be entirely attributable to returns. In reality, neither returns nor costs can be held constant, while the relative degree of their variation may differ. On the one hand, if there are drastic and differential shocks to the cost of producing each type of paper while the returns are relatively constant, then we can study the relationship between these shocks and the share of multi-authored papers. On the other hand, we can use publication records in a prolonged period absent of any shocks to costs to study the effect of returns on the number of co-authors on a paper. We explore these directions in the following sections.

[11] The return to a paper may be defined in other ways. For example, Heckman and Moktan (2020) define relative returns as additional hazard rate for successful tenure.

## 4.1 COVID as a Natural Experiment

The onset of the COVID pandemic swiftly impacted daily life. In-person events quickly dwindled and in many cases were halted completely. The result is an immediate change to costs associated with doing research, increasing the difficulty of face-to-face communication and collaboration while making it easier to work together virtually. Because few people anticipated the pandemic (for most academic researchers in economics), the return to a paper likely did not shift as quickly as costs did, and the impact that a particular paper imparts on the profession likely persisted. Furthermore, since COVID was a rare event, economists lacked prior experience in dealing with a massive pandemic, and it is unlikely that the profession quickly adjusted the standard with which papers are viewed. Therefore, we treat COVID as a natural experiment which, in relative terms, primarily changed costs to collaboration.

Assuming that the relative returns of multi-authored papers and the distribution of individual preferences for collaborative work remained relatively constant before and after COVID, we can attribute changes to shares of multi-authored papers in the post-COVID era to changes in costs $c_s$ and $c_m$.[12] We provide an overview of some of the challenges and opportunities of COVID, which form the basis for our hypotheses regarding changes to research collaboration.

### 4.1.1 The Impact of COVID

The COVID pandemic potentially had an immediate impact on the following channels. We separately list those that may have raised the cost of collaboration and those that served to lower costs.

The following channels likely increased the cost of collaboration.

**Disruption of Research Activities.** Field experiments came to a halt, and in-person access to labs and offices was suspended. There was also increased mental drain related to the pandemic, which may have reduced motivation and research productivity.

**Strain on Funding and Resources.** The pandemic led to financial strain for many funding agencies and research institutions. Some projects, especially those not directly related to COVID, experienced budget cuts or delays in funding.

[12] It is also possible that changes in the composition of papers is due to changes in taste. For example, Heckman and Moktan (2020) assert that single-authored papers provide higher weight than multi-authored papers in tenure decisions. This is outside of our scope that focuses on returns and costs of paper production.

**Cancellation of Conferences and Meetings.** The cancellation of academic conferences, workshops, research visits, and meetings limited the opportunities for researchers to interact face-to-face, which are often crucial for establishing and maintaining collaborative relationships.

The following channels may have reduced the cost of collaboration.

**An Aggregate Shift Toward Virtual Collaboration.** The pandemic accelerated the adoption of virtual collaboration tools, allowing researchers to continue to cooperate through online platforms. Video conferences, cloud-based collaboration tools, and virtual events became more commonplace and more culturally accepted, making collaboration easier and eliminating travel time. Virtual collaboration also reduced the cost of inter-institutional collaboration relative to working with an in-person colleague.

**Increased Open Science and Data Sharing.** There was a notable shift towards open science, with researchers sharing data and findings more freely and more rapidly than before the pandemic. Pre-print servers saw a surge in usage, facilitating faster dissemination of research results and enabling collaborative efforts to build on emerging knowledge without the delays associated with traditional publishing.

COVID may also have long-term impact on the cost of cooperation. Determining the long-term impact is more difficult compared to the short-term, since other factors can confound the initial impact of the pandemic. The mass adaptation of new collaborative tools serves to decrease the cost of collaboration, and their prevalence persists after the pandemic subsided. Researchers retain different outside options during the work-from-home (WFH) era compared to afterward. For example, suppose that COVID limited the capacity of researchers to work on single-authored papers due to increased mental load of handling a health crisis. This would cause an immediate negative shock to the feasibility of single-authored papers and encourage more researchers to work in teams. As the effect of COVID diminished, those with a strong preference for individual work would revert to solo work once that option became feasible.

Combining the perspectives above, we evaluate the net effect of COVID on collaboration patterns to understand how costs can influence teamwork.

#### 4.1.2 The Effect on the Number of Authors

Empirically, COVID had a significant impact on collaboration. Figure 2 illustrates that the economics profession experienced a positive shock in both the number of working papers and publications in 2020. In 2021, the number of working papers on SSRN and NBER

experienced negative growth, but not the publications in top five and EC64. This delay in the growth of working papers to publications is expected, given the widely heterogeneous submission to the publication timeline in economic journals, as documented in Hadavand, Hamermesh, and Wilson (2024).

Because the fraction of multi-authored papers contains a time trend, observed changes during COVID may simply be attributed to a secular shift over time. As such, we must control for the time trend when evaluating the effect of COVID and estimate:

$$100 \cdot \mathbb{1}\{Num_{ist} = n\} = \beta_0 + \sum_{c=2020}^{2023} WP_i \cdot \mathbb{1}\{t = c\} + \sum_{c=2020}^{2023} \mathbb{1}\{t = c\} + \beta_1 t + WP_i + \varepsilon_{ist}, \tag{7}$$

where $WP_i$ is an indicator variable that is 1 if paper $i$ is a working paper, and 0 otherwise.

Table 8 reports the results. The yearly trend variable shows that single-authored and two-authored papers are on the decline, whereas three- and four-authored papers are on the rise. During the COVID pandemic, there appears to be some decrease in three-authored publications and increase in four-authored publications from 2020 through 2023. What is more striking is the magnitude and persistence of change in working papers. Specifically, Column (1) shows that the share of single-authored working papers experienced positive and growing deviations from the negative trend from 2020 to 2023, while the share of single-authored publications was largely unchanged. In contrast, Column (2) shows that the share of two-authored working papers experienced negative and growing deviations relative to their trend from 2020 to 2023, while the share of two-authored publications was virtually unchanged. Column (3) shows that the share of all three-authored papers, whether published or working, experienced negative and growing deviations from the positive trend from 2020 to 2023. Column (4) shows that all papers with four or more authors experienced positive and growing deviations from the positive trend. In summary, our estimates suggest that, after COVID, the share of two- and three-authored working papers decreased, while the share of single-authored papers and papers with four or more authors increased.

### 4.1.3 The Effect on Inter-Institutional Collaboration

Following the shift to work-from-home, economists can be encouraged to initiate collaboration that may have otherwise been unattractive due to great physical distances. At the same time, the halt of in-person events can discourage researchers from forming new collaborations, especially outside of their home institutions. What is the net effect on cooperation across institutions? We focus on the share of papers that are inter-institutional and the

share of authors who are in the major institution as our outcome variables of interest.

Given that the estimated trends of institutional assortativity varies substantially by sample in Table 6, we estimate deviations from the historical trend separately by sample as well. Similar to our preceding analysis, we estimate the COVID period deviations for our outcome variables by adding indicator variables of post-COVID years, revising eq. (2) and eq. (3) to:

$$100 \cdot \mathbb{1}\{Inter_{ist}\} = \alpha_s + \sum_{c=2020}^{2023} \mathbb{1}\{t = c\} + \beta_1 t + \varepsilon_{ist} \tag{8}$$

$$pct_m aj_{ist} = \alpha_s + \sum_{c=2020}^{2023} \mathbb{1}\{t = c\} + \beta_1 t + \varepsilon_{ist} \tag{9}$$

Table 9 presents the results for inter-institutional collaboration. Columns (1) and (4) show that SSRN working papers and papers published in the top five journals do not exhibit any significant deviation from their pre-COVID trends. Columns (2) and (3) suggest some decreases in teamwork across institutions for NBER working papers and publications on EC64 journals.

Table 10 shows the results for the fraction of authors in the major institution in multi-authored papers. SSRN and top five papers show almost no change during COVID compared to the prior period, whereas working papers on NBER and published papers on EC64 journals display some decline, albeit not always consistent across the COVID years. Taken together, the evidence from Table 9 and Table 10 suggest that there is limited evidence of a clear change in inter-institutional collaboration during COVID. In other words, researchers did not significantly alter the composition of their co-authors, even if they collaborated more.

## 4.2 Changes in Returns to Collaboration

We capture the return to a research paper through its impact on the profession, proxied by the number of citation it receives. Jones (2021) hones in on citation counts of a paper and documents that multi-authored economics papers, relative to single-authored papers, have a larger likelihood of becoming highly cited. We follow Jones (2021) to focus on highly-cited papers among teams of different sizes, with the difference that we focus on trailing five-year citations rather than cumulative citations. Our measure favors papers whose impact is recognized relatively quickly, a desirable trait in the competitive landscape of academia, where promotions and prestige are often tied to the recent impact of one's work. Let $i$ denote the paper, $j$ the journal, $t$ the year, and $n$ the number of authors. Let the number of $n-$author papers in journal $j$ in year $t$ be $M_{njt}$. We construct the following variables linked to the impact of papers:

- $HR_{nijt}$ (Homerun) : A binary variable that is 1 if paper $i$ is a homerun paper, defined as a publication whose trailing five-year citation is among the top 10% of all EC64 papers published in the same year.
- $RNI_t = \left( \frac{\sum_i HR_{Nijt} \Big/ \sum_j M_{Njt}}{\sum_i HR_{1ijt} \Big/ \sum_j M_{1jt}} \right)$ (Relative $N$-Author Impact): The ratio between the share of $N$-authored papers that are homerun and the share of single-authored papers that are homerun. This is a comparison between the homerun conversion rate of $N$-authored papers to that of single-authored papers.
- $RNR_t = \left( \frac{\sum_i HR_{Nijt} \Big/ \sum_{i,n} HR_{nijt}}{\sum_i HR_{1ijt} \Big/ \sum_{i,n} HR_{nijt}} \right) = \left( \frac{\sum_i HR_{Nijt}}{\sum_i HR_{1ijt}} \right)$ (Relative $N$-Author Returns): The ratio between the share of all homerun papers that are $N$-authored and the share of homerun papers that are single-authored. This is a comparison between the contribution to all homerun papers from $N$-authored collaboration and individual work.

We study the time variation and driving factors of the above variables to understand differential returns from writing single-authored and multi-authored papers.

### 4.2.1 Predicting a High Impact, Homerun Paper

We estimate the relative returns of having multiple authors on a paper using the specification in Equation (10) below.

$$100 \cdot HR_{nijt} = \alpha + \sum_N \beta_N \mathbb{1}\{Num_{ist} = N\} + \gamma X_{ijt} + \kappa_t + \phi_j^J + \phi_j^F + \varepsilon_{nijt} \tag{10}$$

where $\beta_N$'s are coefficients on $N$-author fixed effects, $\kappa_t$ capture year fixed effects, $\phi_j^J$ are fixed effects for the journal, and $\phi_j^F$ are fixed effects for the field. $X$ is a vector of binary variables including whether a paper is inter-institutional, has senior author, has an author from one of the top 10 schools, has a senior author from top 10 schools, has an author from the top 11-30 schools, has a senior author from top 11-30 schools, has an author from US institution, and has at least one international collaborator.

Table 3 reports the results from Equation (10). Column (1) controls for paper characteristics including inter-institutional collaboration, author experience and institutional rank, whether one author is affiliated with an US institution, and whether the collaboration is international. The strongest predictor for whether a paper will be highly cited is having an

author from a top 10 institution, followed by having an author from a top 11 to 30 institution. Other characteristics have relatively small effects on the probability of becoming a homerun paper, with authors from U.S. institutions, the number of authors, and international collaboration also containing some explanatory power. Controlling for characteristics, articles with multiple authors are more likely to become homerun papers. Two- and three-authored papers are 2.36% and 2.94% more likely, compared to single-authored papers, to highly cited, while four-plus-authored papers do not show a significant difference. Column (2) include the same control variables as Column (1), plus year fixed effects. Four-plus-authored papers now show a statistically significant increase in the likelihood of becoming a homerun paper compared to single-authored papers. Column (3) further adds fixed effects for fields (see Table 2 for the categorization of fields and journals), and Column (4) includes journal fixed effects. We observe similar estimates in these columns, with similar economic magnitudes across two-, three-, and four-plus-authored articles when controlling for latent differences across journals. Table B2 repeats the analysis, excluding publications in the top five journals. While some point estimates change, overall patterns remain unchanged. We also re-estimate the model without the interaction term between the experience level of an author and her institutional ranking, and change the indicator variable for author institutional ranking to whether the majority of authors on a paper belong to a certain ranking. Table B3 shows that the results also remain largely unchanged.

An alternative measure to the impact of a paper is its percentile among the trailing five-year citations of all papers published in the same year. Whereas $HR_{nijt}$ is a binary variable that turns on if a paper is in the top 10% of citations, this alternative measure is a continuous variable with larger values indicating more impact. All right-hand variables remain identical to those in Equation (10). Table B4 exhibit similar estimates as those in Table 3: More impactful papers are associated with authors from top 10 institutions, top 11 to 30 institutions, multiple authors, and U.S. authors. Interestingly, whether a paper includes international collaborators appears to be a stronger driver of its citation percentile compared to the likelihood of becoming a homerun paper.

### 4.2.2 Time-Varying Returns to Multi-Authored Papers

The preceding results show that research teams can have different probabilities of creating impactful work. We further investigate whether these probabilities change over time, and if teams of a certain size consistently outperform other sizes. We re-estimate Equation (10) to obtain the likelihood of becoming a homerun paper by two-, three-, and four-plus-authored papers over 10-year intervals. Figure 5 illustrates the likelihood of becoming a homerun paper for different-sized teams, compared to a benchmark of single-authored works. For

example, a two-authored paper is 2.5% less likely to become highly cited when compared to a single-authored paper in the 1950s, but 2% more likely in the 1960s.

There appears to be several waves of change in the type of research that is most likely to become a homerun paper. In the 1950s, single-authored papers were the most likely to be highly cited, followed by two- and three-authored papers. Four-plus-authored papers are a full 5% less likely to be highly cited compared to single-authored papers. In the 1960s and 1970s, two-authored papers overtook single-authored papers to become the most likely to be highly cited. Three-authored papers also became more likely to be highly cited compared to single-authored papers in the 1970s. By the 1980s, two- and three-authored papers were both 3-4% more likely to be highly cited compared to single-authored papers, while four-plus-authored papers continued to see less impact. Since the 1960s, the likelihood of multiple authors producing impactful papers has consistently exceeded that of single-authored works. In particular, three-authored papers became the most likely to be highly cited in the 1990s and 2000s, while four-plus-authored papers became the most likely in the 2010s (5% higher compared to single-authored papers).

The patterns in Figure 5 reveal that there have been increasing returns to working in larger teams. Over time, larger teams became more effective at producing high-impact research. The timing of each successive wave of the highest-cited papers appears to coincide with the change in the shares of each type of paper (Figure 7). We plot the fraction of $N$-authored papers along with the fraction of homerun papers written by $N$-authored teams in Figure 13. The former can be viewed as the "popularity" of specific-sized teams, while the latter as the "success." For teams of all sizes, popularity and size track each other closely. Single-authored papers show a steady decrease in both popularity and success over time. Two-authored papers have seen a resurgence in popularity and success since the 1960s, compared to relatively steady shares from 1900 to 1960. Three- and four-plus-authored papers experienced both a surge in popularity and success in the 1980s and 2000s, respectively.

### 4.2.3 Do Returns Drive Popularity?

Figure 13 suggests a close relationship between the success of a multi-authored paper and how common it is. On the one hand, economics researchers may respond to increasing returns to larger teams by writing more multi-authored papers. On the other hand, stronger researchers working more often in teams can mean a greater fraction of highly influential multi-authored papers. In this section, we investigate these possibilities.

We explore whether economists collaborate more in response to increasing returns to

collaborative works. We would like to estimate the following model:

$$S_{N,t} = \alpha_N + \beta_{0,N} t + \beta_{1,N} HRS_{N,t-5} + \gamma_N S_{N,t-5} + e_{N,t} \tag{11}$$

where $S_{N,t} = \frac{M_{njt}}{\sum_n M_{njt}}$ is the share of $N$-author papers in year $t$. $HRS_{N,t-5}$ is the share of $N$-author papers that are in the top 10% of trailing five-year citations, in year $t-5$. The choice of a five-year lag aligns with our choice of a five-year trailing citation count.[13]

If we estimate Equation (11) for the sets of papers with different authors, their shares add to 100%, $\sum_{N=1}^{4+} S_{N,t} = 100$. This way, information present in one regression (for one set of papers) may be related to coefficients in another regression. We handle this issue by adopting the methodology of Tomz, Tucker, and Wittenberg (2002), who provide a simple way to estimate models under the seemingly unrelated regression framework with multiple categories of outcomes. We create the following variables:

- $LRNS_t = ln\left(\frac{S_{N,t}}{S_{1,t}}\right), \quad N = 2, 3, 4+$
- $LRNI_t = ln\left(RNI_{N,t}\right)$
- $LRNR_t = ln\left(RNR_{N,t}\right)$

Computing the relative share of papers (to the share of single-authored papers $S_{1,t}$ and taking logarithms breaks the constraint that the sum of all shares must be 100%. We then estimate the following system of equations using the SUR framework:

$$\begin{cases} \Delta LR2S_t = \alpha_2 + \beta_{0,2} t + \beta_{1,2} LR2I_{t-5} + \gamma_2 LR2S_{t-5} + e_t \\ \Delta LR3S_t = \alpha_3 + \beta_{0,3} t + \beta_{1,3} LR3I_{t-5} + \gamma_3 LR3S_{t-5} + e_t \\ \Delta LR4S_t = \alpha_4 + \beta_{0,4} t + \beta_{1,4} LR4I_{t-5} + \gamma_4 LR4S_{t-5} + e_t \end{cases} \tag{12}$$

where $\Delta LRNS_t = LRNS_t - LRNS_{t-1}$ is the first difference of log shares. We are interested in the coefficient associated with the relative $N$-Author Impact from five years ago, $LRNI_{t-5}$. A positive coefficient would indicate that researchers tend to assemble teams following their relative impact in the preceding period.

Table 11 provides the results. The top two panels, on two- and three-authored papers, show little support for the idea that economists form teams in response to larger team impact. In comparison, the bottom panel, for four-plus-authored papers, offers evidence that researchers assemble into larger teams in response to higher returns. Across the four alternative specifications in the columns, the coefficient on $LR4I_{t-5}$ remains consistently

[13] We estimate the same models with a six-year or a seven-year lag, and the results do not change.

positive and significant, with a similar economic magnitude. We also estimate the system from Equation (12) with relative $N$-author returns, $LRNR_t$, in place of $LRNI_t$. Table B5 shows similar results as those from Table 11. Whereas researchers respond to higher returns of the largest teams (four or more members), they do not appear to do so for smaller teams of two or three people.

# 5 Conclusion

This paper seeks to expand our current understanding of research collaboration along two dimensions. First, we document more granular collaboration patterns, including how economists of different experience levels work together, and offer more complete evidence on how collaboration across institutions has changed. Second, we study potential drivers of collaboration patterns through a simple framework that allows us to compare the benefit of a paper to its cost of production. Changes in benefits and costs can generate different collaboration patterns over time.

We construct a novel dataset combining working papers from NBER and SSRN, as well as published papers from OpenAlex, and provide detailed author-level information such as affiliation data across time. Consistent with existing work, we find a significant trend of an increasing number of co-authors on research papers over time. The share of multi-authored papers, whether as a fraction of working papers or publications, continues to rise over time. Teamwork across institutions has become more common, but only in the most recent 25 years. Researchers tend to work with others with similar experience and standing, such that the mix between senior and junior economists remains unchanged in recent decades.

Using COVID as a plausible exogenous shock to the cost of paper production, we find a polarizing effect on scientific scholarship. While some researchers tended to work more by themselves, others collaborated more broadly during the pandemic than ever before. Inter-institutional collaboration remained largely unchanged in this period. We also document increasing returns to collaborative work over time. While single-authored papers were the most likely to become highly cited in the 1950s, multi-authored papers became increasingly more likely to be highly cited in subsequent decades, with four-plus-authored papers finishing as the most influential in the 2010s. We test the hypothesis that the prevalence of multi-authored papers follows a rise in impact, and find that researchers tend to respond to higher returns of larger teams but show limited response to rising returns of medium-sized teams.

While our work takes a step toward understanding research collaboration through variation in returns and costs, our empirical investigation imposes minimal structure and mainly relies on reduced-form evidence. A more formal theoretical exploration of the mechanism

of research collaboration should account for the empirical patterns in returns and costs, and provide an explanation for their observed changes over time. We leave this intriguing direction to future work.

# Bibliography


Ahmadpoor, Mohammad and Benjamin F Jones. 2019. "Decoding team and individual impact in science and invention." *Proceedings of the National Academy of Sciences* 116 (28):13885–13890.

Bayhan, Hasan Gokberk, Sinem Mollaoglu, Hanzhe Zhang, and Kenneth Frank. 2022. "Project team collaborations during time of disruptions: Transaction costs, knowledge flows, and social network theory perspective." *Construction Research Congress* :1012–1023.

Butler, Jenna and Sonia Jaffe. 2021. "Challenges and gratitude: A diary study of software engineers working from home during covid-19 pandemic." In *2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*. IEEE, 362–363.

Cabral, Luís. 2020. "List of leading journals for economics faculty." Accessed: 2025-02-06.

Dong, Xu, Victoria R. Nelson, Hanzhe Zhang, Young Anna Argyris, Sinem Mollaoglu, Kenneth Frank, Arnav Jain, Xiao Qiao, Joseph Thekinen, and Haochen Liu. 2023. "Does the transition to online meetings exacerbate or alleviate gender and racial inequities in student participation in project teams? Theory-guided machine-learning analysis of online collaborative work during the pandemic." Mimeo.

Falk-Krzesinski, Holly J, Noshir Contractor, Stephen M Fiore, Kara L Hall, Cathleen Kane, Joann Keyton, Julie Thompson Klein, Bonnie Spring, Daniel Stokols, and William Trochim. 2011. "Mapping a research agenda for the science of team science." *Research Evaluation* 20 (2):145–158.

Ford, Denae, Margaret-Anne Storey, Thomas Zimmermann, Christian Bird, Sonia Jaffe, Chandra Maddila, Jenna L Butler, Brian Houck, and Nachiappan Nagappan. 2021. "A tale of two cities: Software developers working from home during the covid-19 pandemic." *ACM Transactions on Software Engineering and Methodology (TOSEM)* 31 (2):1–37.

Freeman, Richard B., Danxia Xie, Hanzhe Zhang, and Hanzhang Zhou. 2024. "High and rising concentration of award-winning economists." Mimeo.

Guimera, Roger, Brian Uzzi, Jarrett Spiro, and Luis A Nunes Amaral. 2005. "Team assembly mechanisms determine collaboration network structure and team performance." *Science* 308 (5722):697–702.

Hadavand, Aboozar, Daniel S. Hamermesh, and Wesley W. Wilson. 2024. "Publishing economics: How slow? Why slow? Is slow productive? How to fix slow?" *Journal of Economic Literature* 62 (1):269–293.

Ham, John C., Julian Wright, and Ziqiu Ye. 2021. "New rankings of economics journals:

documenting and explaining the rise of the new society journals." Mimeo.

Heckman, James J. and Sidharth Moktan. 2020. "Publishing and promotion in economics: The tyranny of the top five." *Journal of Economic Literature* 58 (2):419–470.

Heo, Seulkee, Alisha Yee Chan, Pedro Diaz Peralta, Lan Jin, Claudia Ribeiro Pereira Nunes, and Michelle L Bell. 2022. "Impacts of the COVID-19 pandemic on scientists' productivity in science, technology, engineering, mathematics (STEM), and medicine fields." *Humanities and Social Sciences Communications* 9 (1):1–11.

Hunter, Laura and Erin Leahey. 2008. "Collaborative research in sociology: Trends and contributing factors." *The American Sociologist* 39:290–306.

Jain, Arnav, Faizan Shafique, Sinem Mollaoglu, Xu Dong, Hanzhe Zhang, Shimeng Dai, Kenneth Frank, Dorothy Carter, Anna Young Argyris, Annick Anctil et al. 2024. "Iterative development of dynamic student project team interventions." In *Construction Research Congress 2024*. 109–118.

Jones, Benjamin F. 2009. "The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?" *The Review of Economic Studies* 76 (1):283–317.

———. 2010. "Age and great invention." *The Review of Economics and Statistics* 92 (1):1–14.

Jones, Benjamin F. 2021. "The rise of research teams: Benefits and costs in economics." *Journal of Economic Perspectives* 35 (2):191–216.

Jones, Benjamin F, Stefan Wuchty, and Brian Uzzi. 2008. "Multi-university research teams: Shifting impact, geography, and stratification in science." *science* 322 (5905):1259–1262.

Kalaitzidakis, Pantelis, Theofanis P. Mamuneas, and Thanasis Stengos. 2003. "Rankings of academic journals and institutions in economics." *Journal of the European Economic Association* 1 (6):1346–1366.

King, Molly M. and Megan E. Frederickson. 2021. "The pandemic penalty: The gendered effects of COVID-19 on scientific productivity." *Socius* 7:23780231211006977.

Lin, Zihang, Yian Yin, Lu Liu, and Dashun Wang. 2023. "SciSciNet: A large-scale open data lake for the science of science research." *Scientific Data* 10 (1):315.

Liu, Lu, Benjamin F. Jones, Brian Uzzi, and Dashun Wang. 2023. "Data, measurement and empirical methods in the science of science." *Nature Human Behaviour* 7 (7):1046–1058.

Liu, Meijun, Ning Zhang, Xiao Hu, Ajay Jaiswal, Jian Xu, Hong Chen, Ying Ding, and Yi Bu. 2022. "Further divided gender gaps in research productivity and collaboration during the COVID-19 pandemic: Evidence from coronavirus-related literature." *Journal of Informetrics* 16 (2):101295.

McFadden, Daniel. 1974. "The measurement of urban travel demand." *Journal of Public Economics* 3 (4):303–328.

Milojević, Staša. 2014. "Principles of scientific research team formation and evolution." *Proceedings of the National Academy of Sciences* 111 (11):3984–3989.

Sinatra, Roberta, Sandra Sousa, Yanmeng Xing, and M. Szell. 2023. "The Impact of COVID-19 on Scientific Productivity and Collaboration." In *Proceedings of the 9th International Conference on Computational Social Science (IC2S2)*. Copenhagen, Denmark, 1–4.

Tilburg University. 2025. "Economics and Business Schools Research Ranking Sandbox." URL https://rankings.tilburguniversity.edu/sandbox/. Accessed: 2025-02-06.

Tomz, Michael, Joshua A. Tucker, and Jason Wittenberg. 2002. "An easy and accurate regression model for multiparty electoral data." *Political Analysis* 10 (1):66–83.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Benjamin F. Jones. 2013. "Atypical combinations and scientific impact." *Science* 342 (6157):468–472.

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi. 2007. "The increasing dominance of teams in production of knowledge." *Science* 316 (5827):1036–1039.

Xie, Yu. 2014. ""Undemocracy": inequalities in science." *Science* 344 (6186):809–810.

Yang, Longqi, David Holtz, Sonia Jaffe, Siddharth Suri, Shilpi Sinha, Jeffrey Weston, Connor Joyce, Neha Shah, Kevin Sherman, Brent Hecht et al. 2022. "The effects of remote work on collaboration among information workers." *Nature Human Behaviour* 6 (1):43–54.

# 6 Figures

Figure 1: The Number of Economists and the Number of Journals in EC64

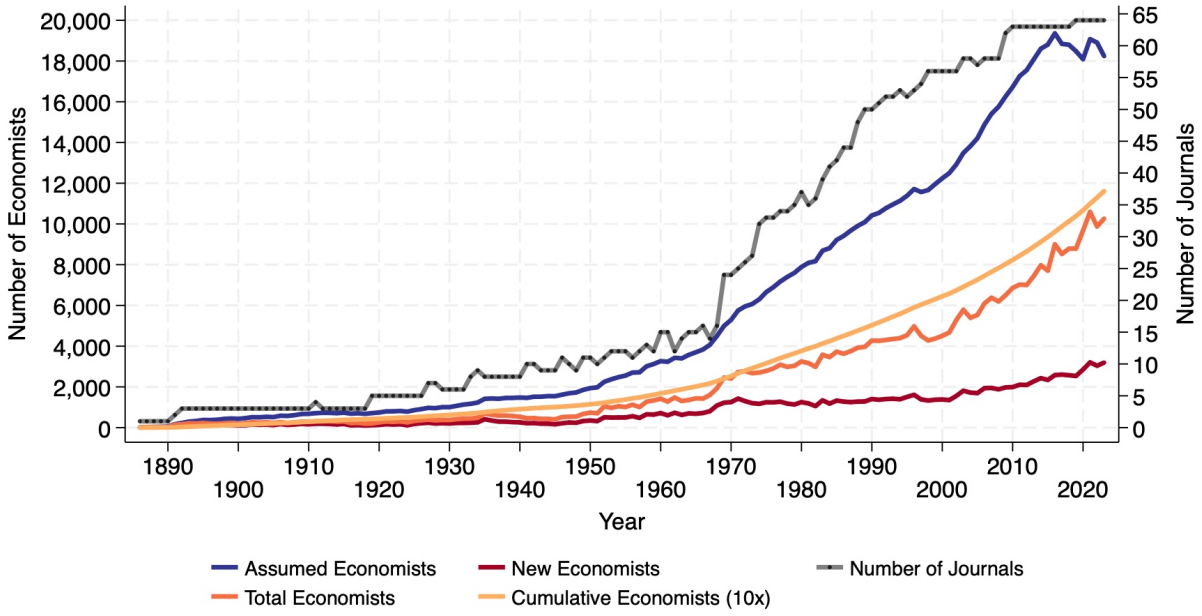



Figure 2: The Number of Economics Papers, 1886-2023

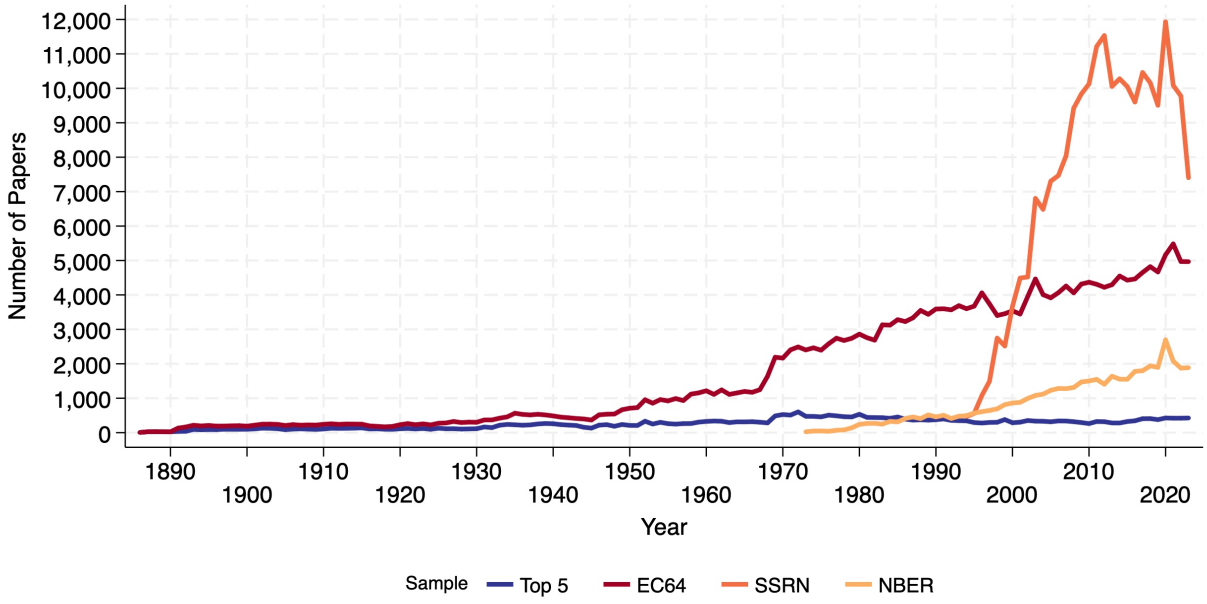


Figure 3: The Number of Authors on a Paper, EC64 Journals, 1886-2023

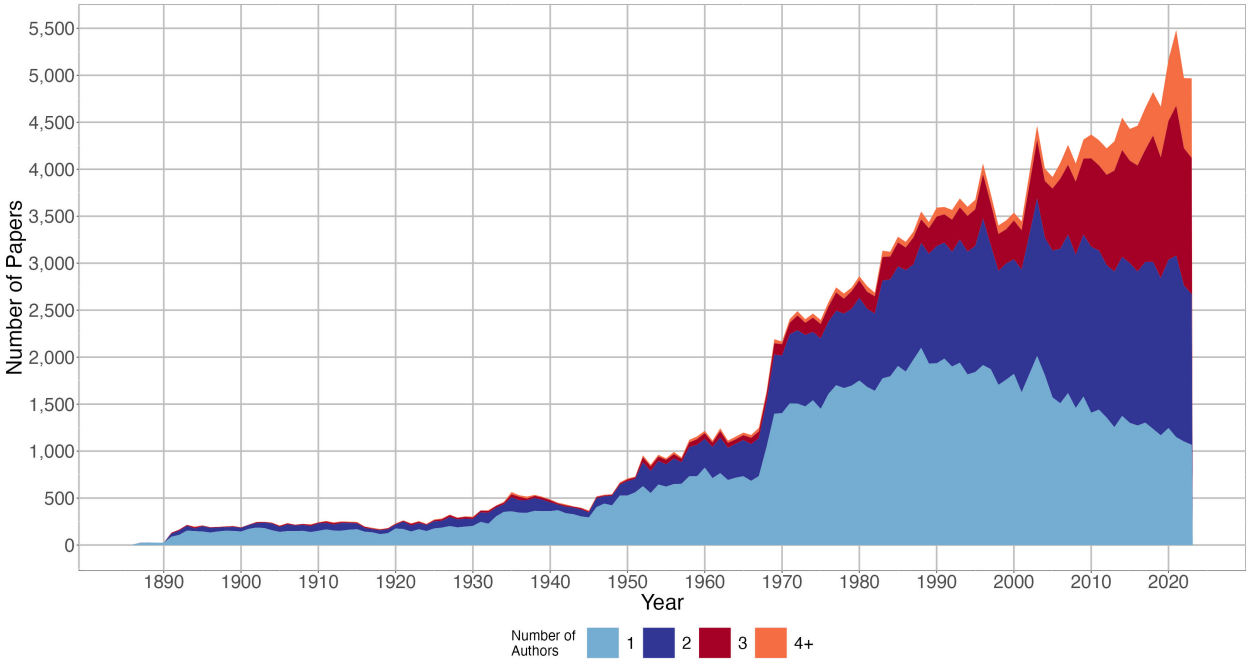
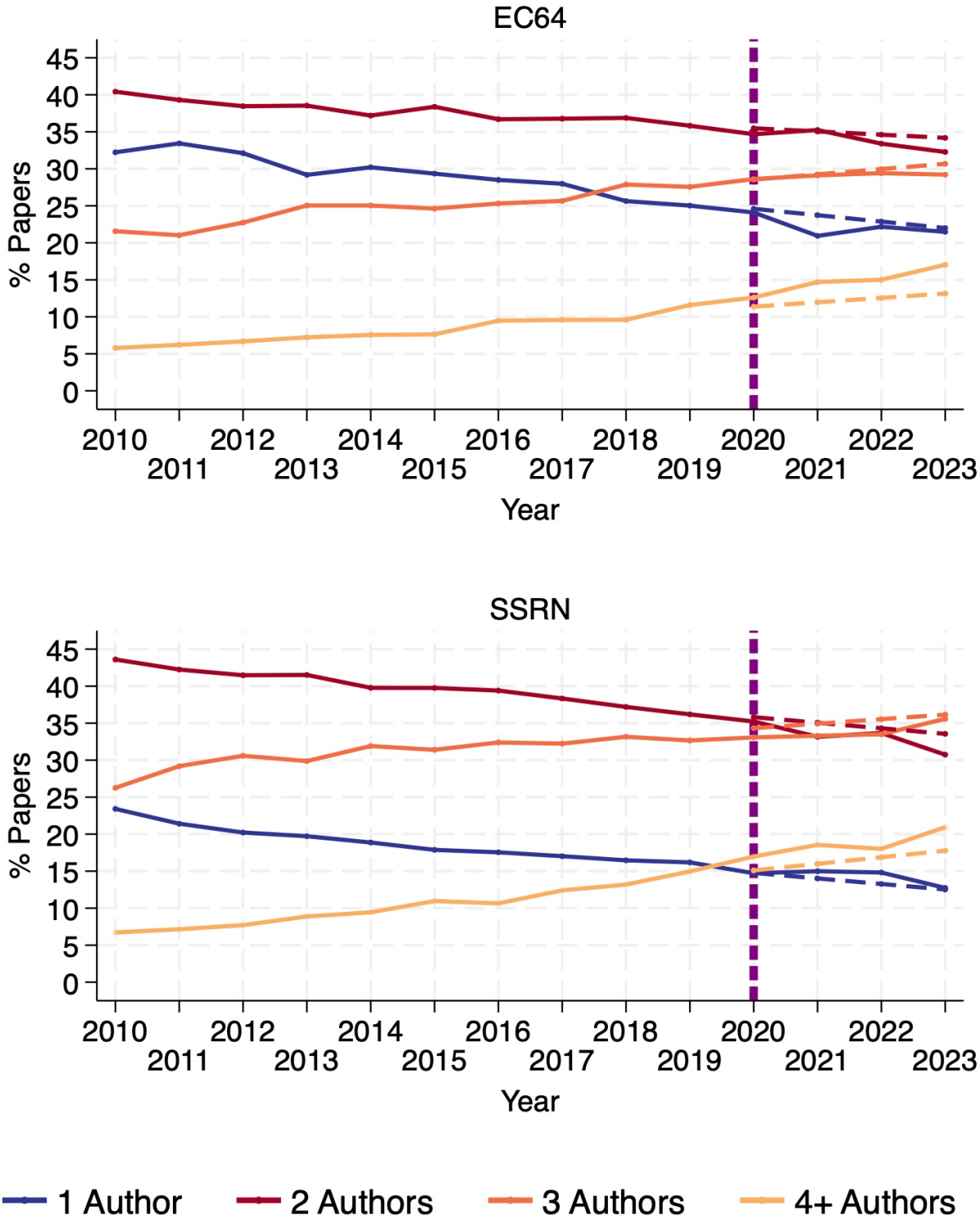
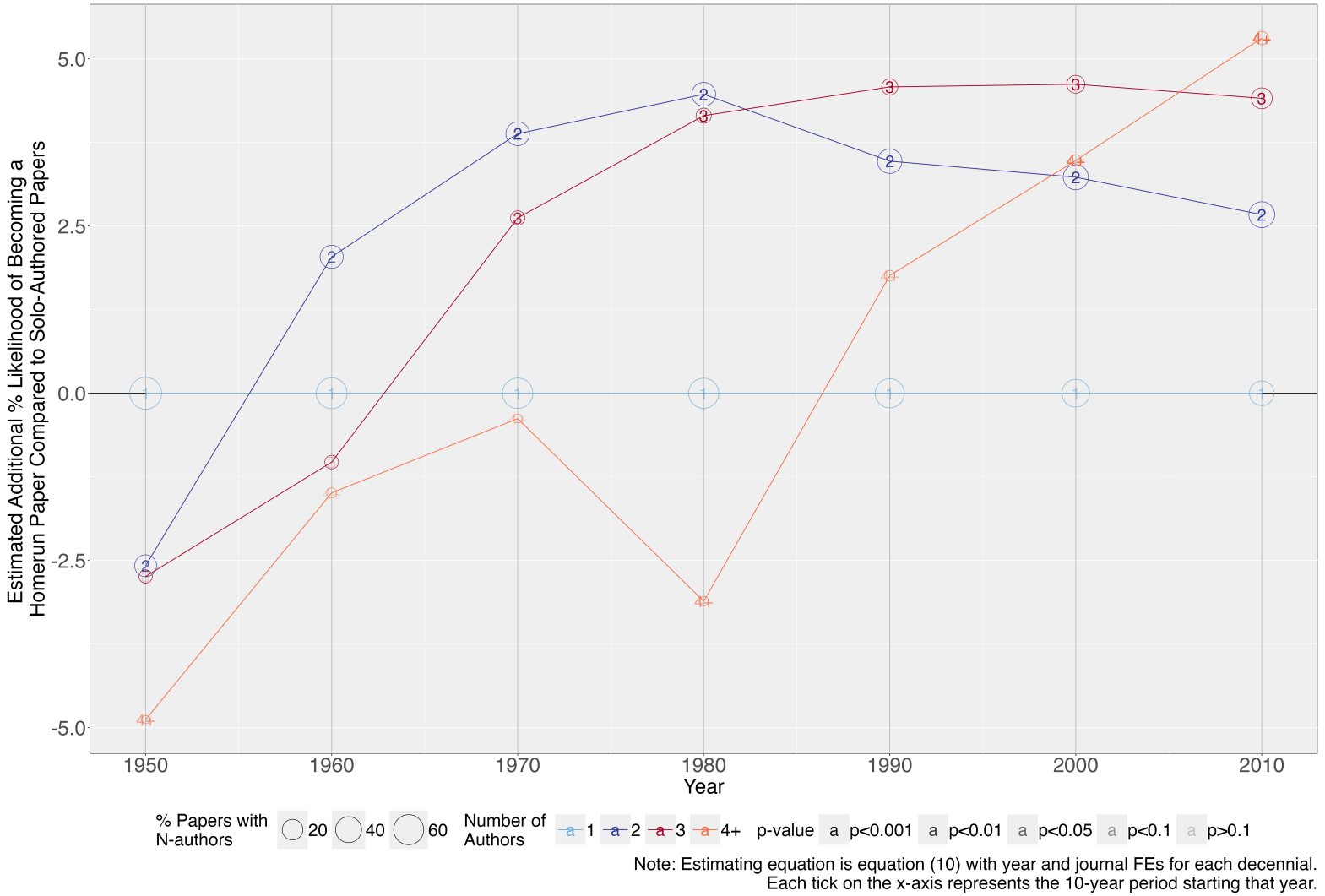


Figure 4: Deviation from Linear Trend, Number of Authors, 2010-2023




Note: Dashed lines are linear predictions.

Figure 5: Evolution of Estimated Returns to Number of Authors, 10-Year Periods




Note: Estimating equation is equation (10) with year and journal FEs for each decennial. Each tick on the x-axis represents the 10-year period starting that year.

Figure 6: Multi-Authored Papers Increased Over Time

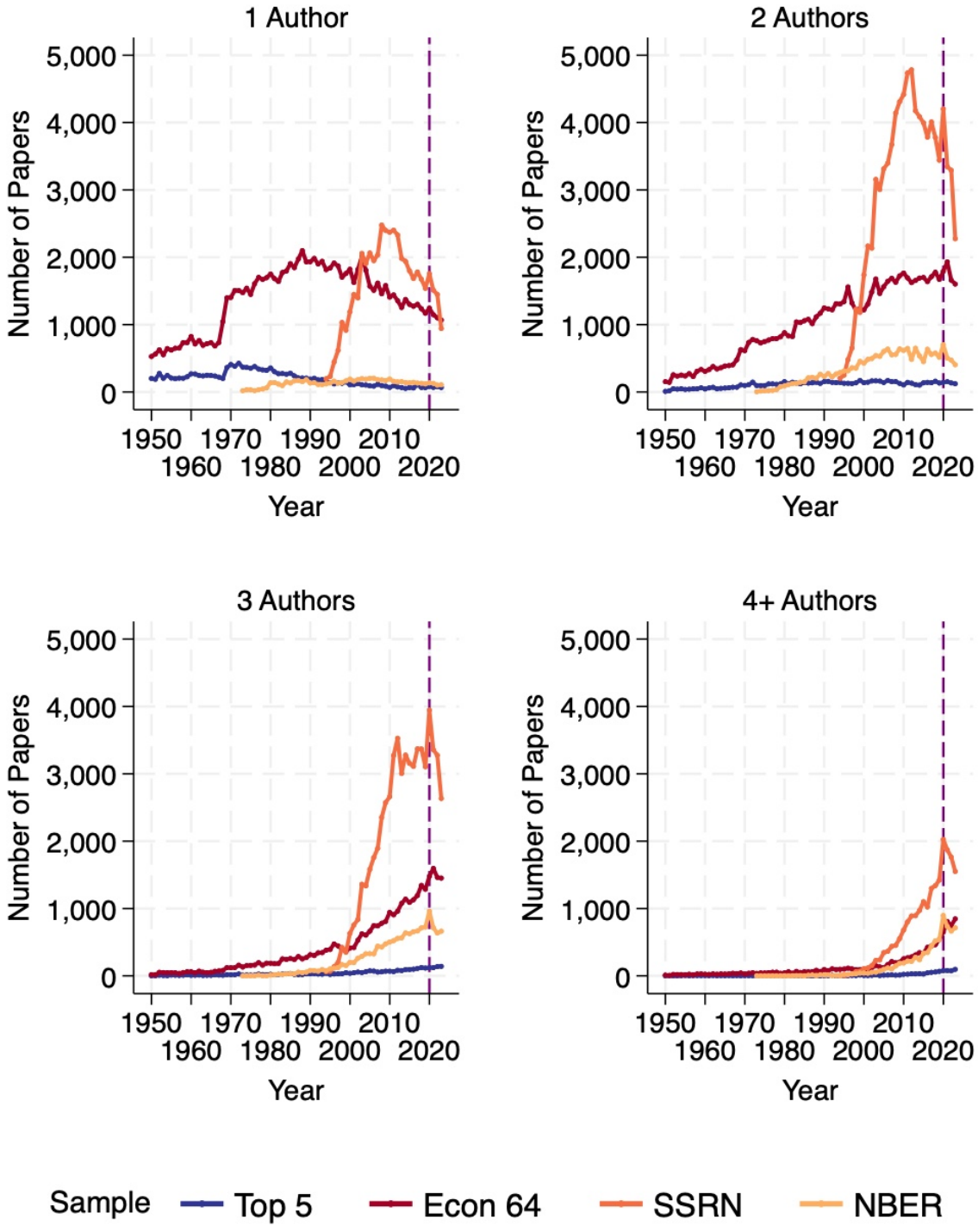


Figure 7: Number of Authors Increased Over Time

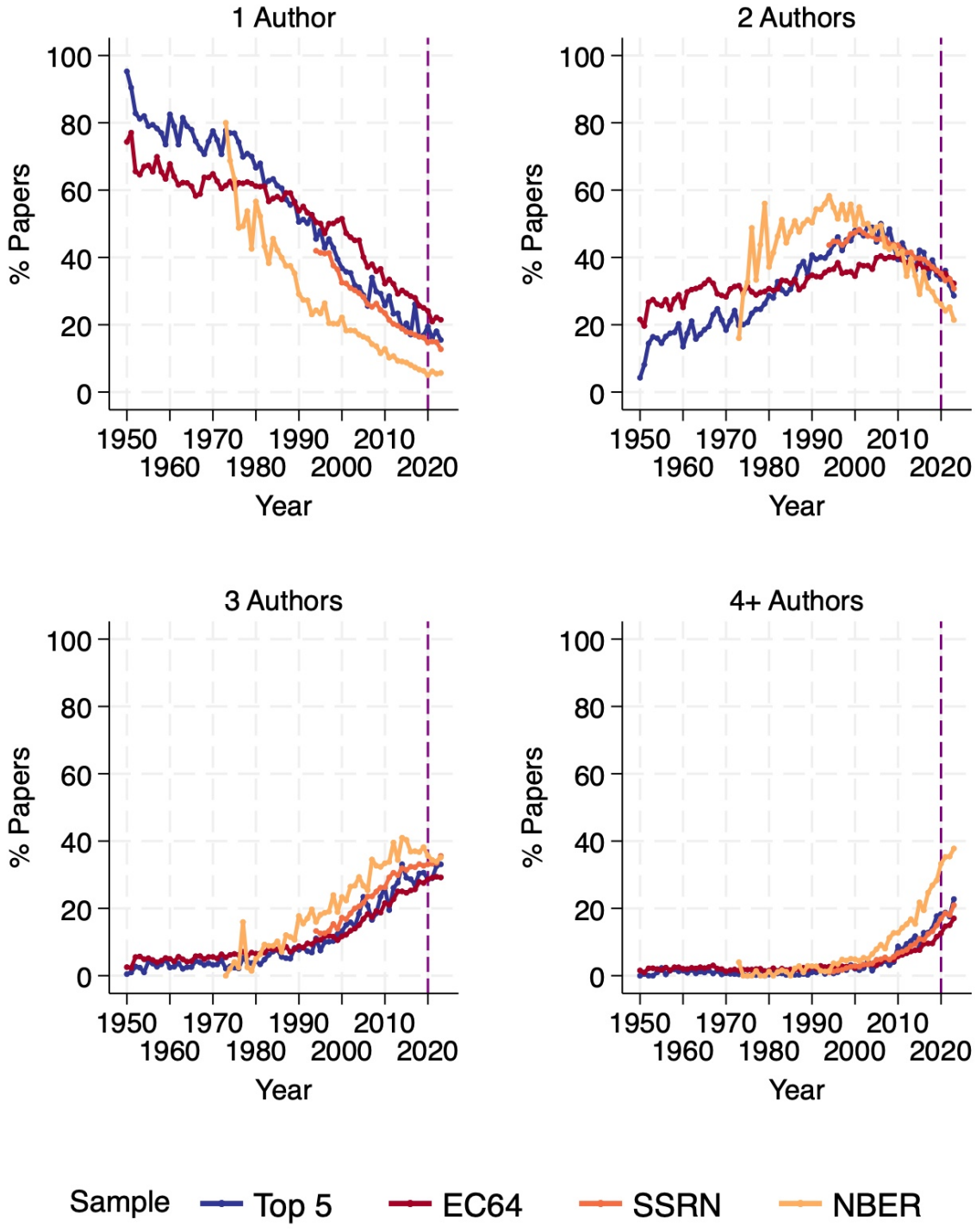


Figure 8: Trends in Inter-Institutional Collaboration

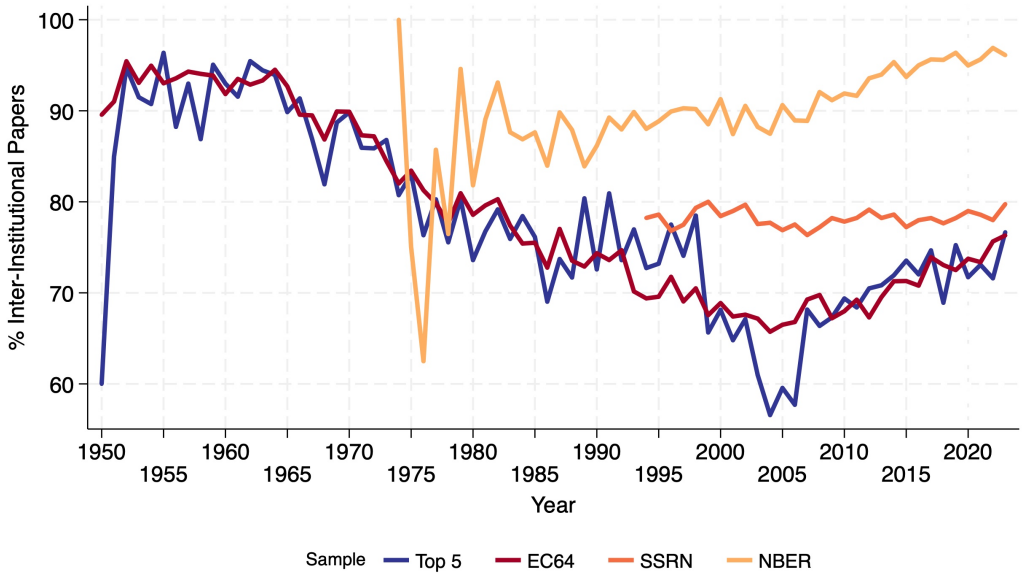



Figure 9: Decrease in Institutional Concentration

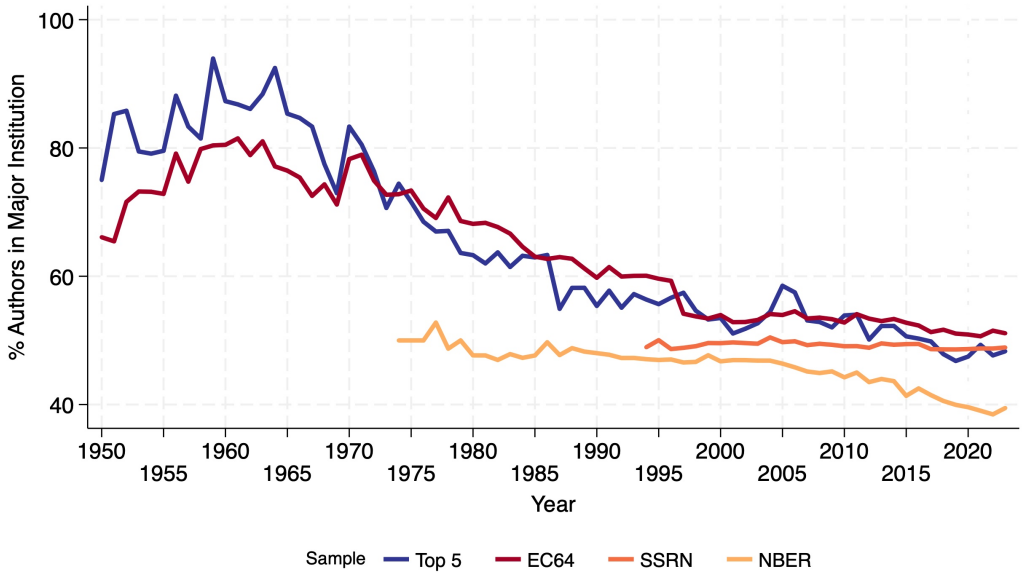


Figure 10: Stable Pattern of Experience Assortativity in 2-Authored EC64 Papers

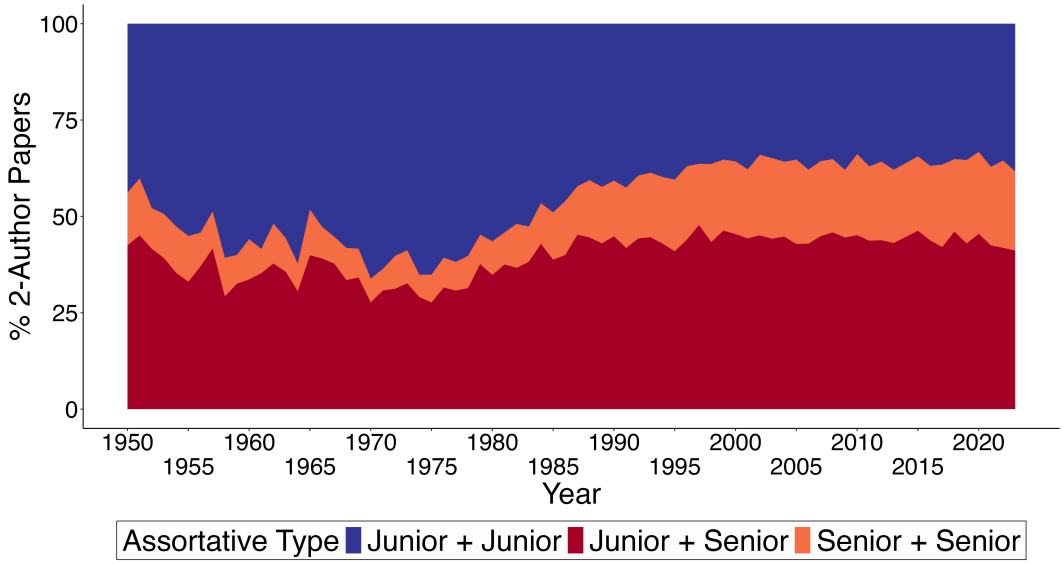



Figure 11: Stable Pattern of Experience Assortativity in 2-Authored SSRN Papers

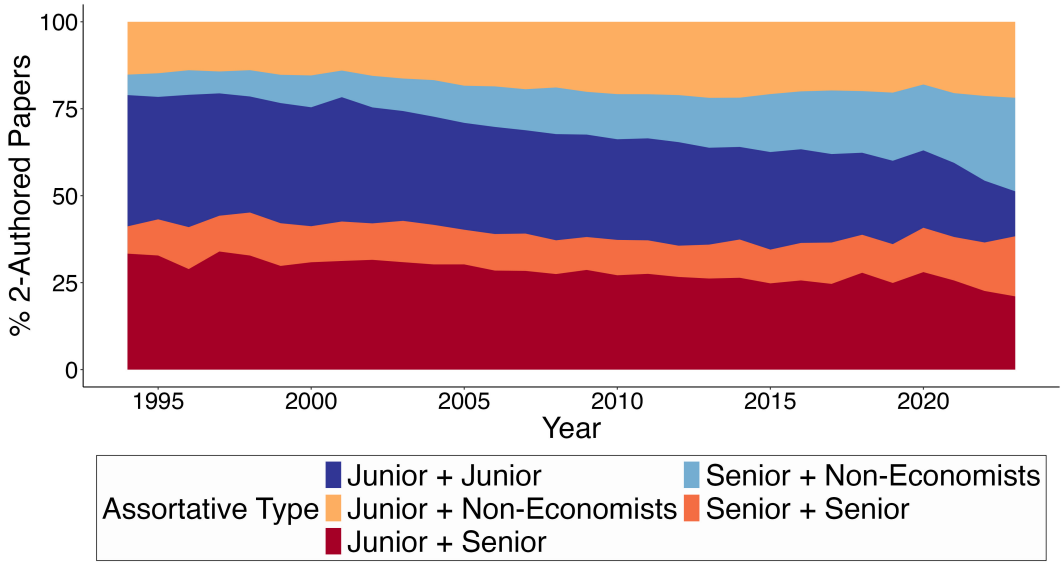


Figure 12: Experience Assortativity in EC64 Papers Conditional on Number of Authors

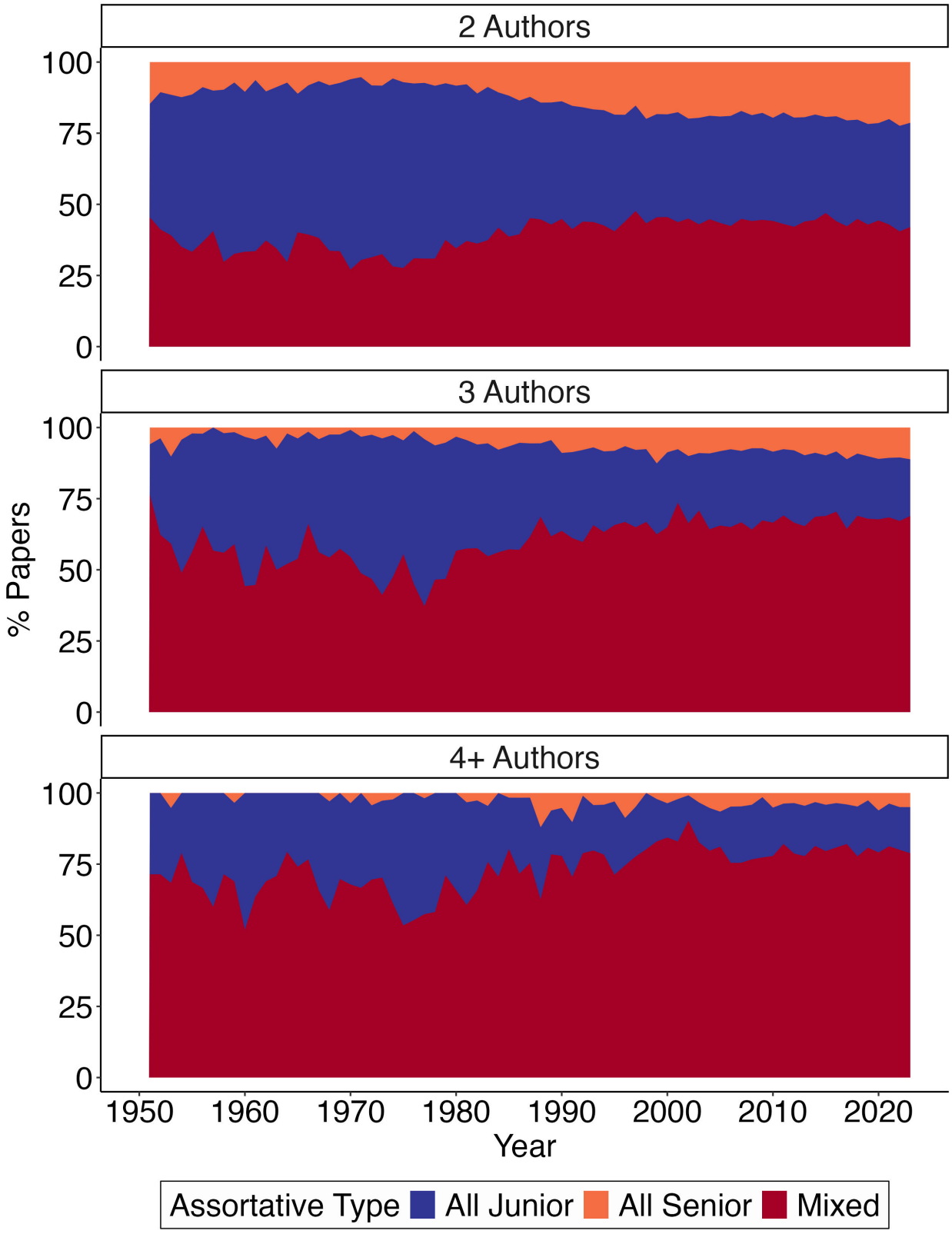
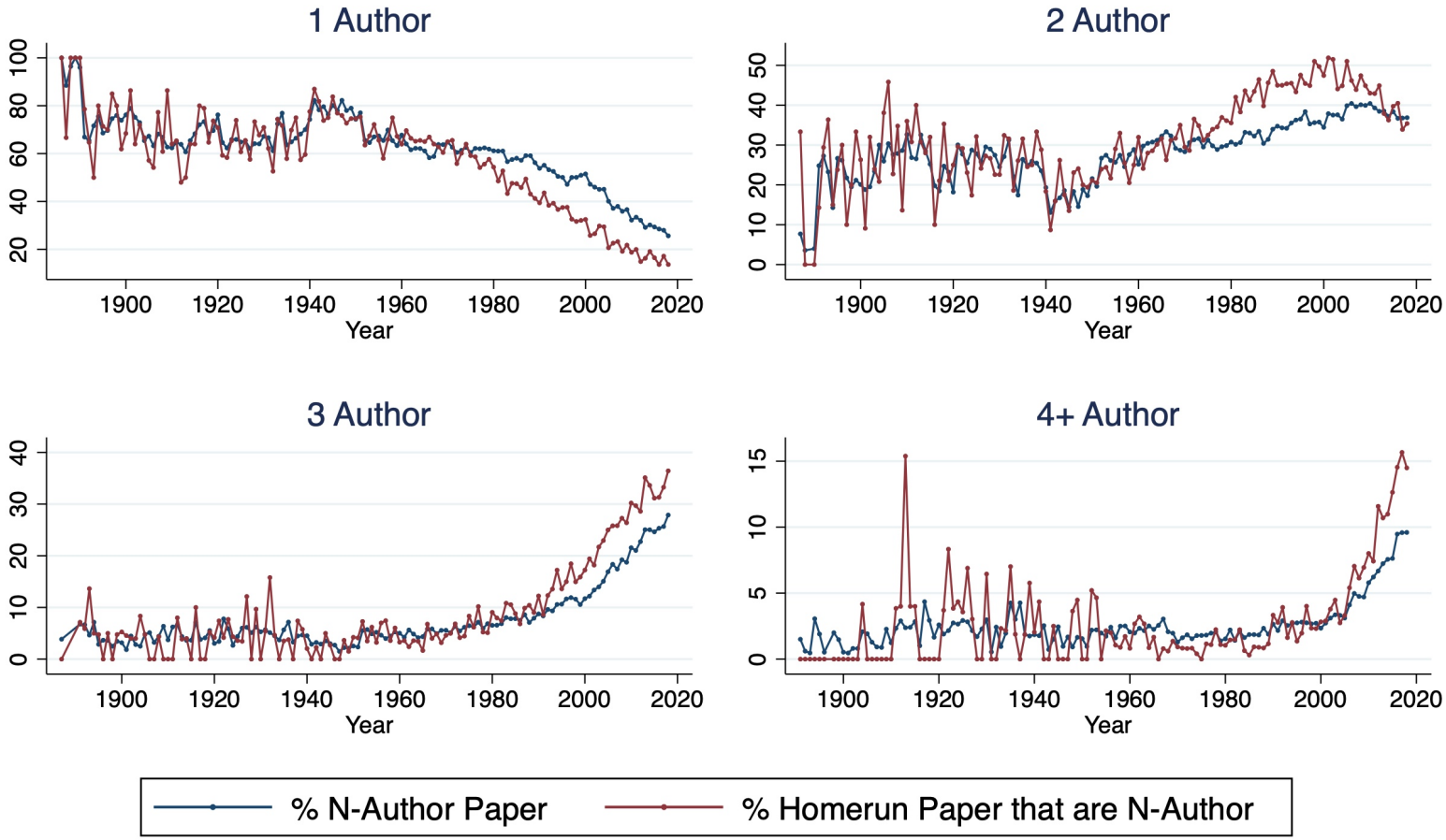


Figure 13: Evolution of the Fraction and Success of $N$-Author Papers



# 7 Tables

Table 1: Summary Statistics of Samples

| | SSRN | NBER | Top Five | EC64 |
|---|---|---|---|---|
| Coverage Years | 1994-2023 | 1973-2023 | 1886-2023 | 1886-2023 |
| Number of Papers | 220,446 | 31,154 | 38,074 | 242,158 |
| Number of Authors per Paper | 2.3 | 2.6 | 1.5 | 1.8 |
| % Inter-Institutional in Multi-Authored Papers | 78.2% | 92.6% | 74.5% | 74.8% |
| Avg % Junior in Multi-Authored Papers | 48.8% | 36.1% | 61.4% | 63.1% |

Table 3: Predictors of Homerun Papers, 1900-2018

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | % Likelihood of Becoming a Homerun Paper | | | |
| Number of Authors | | | | |
| 2 | 2.36*** | 2.92*** | 2.68*** | 3.81*** |
| | (0.16) | (0.17) | (0.16) | (0.17) |
| 3 | 2.94*** | 3.82*** | 3.28*** | 4.76*** |
| | (0.24) | (0.24) | (0.24) | (0.24) |
| 4+ | 0.52 | 1.28*** | 1.20*** | 3.49*** |
| | (0.38) | (0.38) | (0.37) | (0.37) |
| Inter-Institutional | 1.22*** | 0.49*** | -0.33** | -1.54*** |
| | (0.15) | (0.15) | (0.15) | (0.15) |
| Has Senior Author | -0.99*** | -0.72*** | -0.94*** | -0.71*** |
| | (0.15) | (0.15) | (0.15) | (0.15) |
| Has Top 10 Author | 15.93*** | 15.95*** | 13.81*** | 11.04*** |
| | (0.26) | (0.26) | (0.26) | (0.26) |
| Has Top 10 Senior | -0.72** | -0.82** | -1.00*** | -0.72** |
| | (0.35) | (0.35) | (0.35) | (0.34) |
| Has 11-30 Author | 7.80*** | 7.94*** | 6.80*** | 5.02*** |
| | (0.27) | (0.27) | (0.26) | (0.26) |
| Has 11-30 Senior | -0.42 | -0.47 | -0.37 | -0.42 |
| | (0.36) | (0.36) | (0.35) | (0.35) |
| US Institution | 0.67*** | 1.03*** | 2.64*** | 3.20*** |
| | (0.15) | (0.15) | (0.15) | (0.16) |
| International Collab. | -0.02 | 0.79*** | 0.81*** | 0.71*** |
| | (0.23) | (0.23) | (0.23) | (0.23) |
| Number of Papers | 213,947 | 213,947 | 213,947 | 213,947 |
| Year FE | | × | × | × |
| Field FE | | | × | |
| Journal FE | | | | × |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. A homerun paper is a paper that has a 5-year EC64 citation in the top 10 percentile among EC64 papers published in that year. Top 10 schools are Harvard, MIT, Yale, Stanford, Columbia, Princeton, UChicago, UPenn, Northwestern, and UCBerkeley. 11-30 schools are NYU, UCLA, UMich, Cal Tech, Cornell, UCSD, UW Madison, Duke, UMinn, Brown, CMU, BU, Johns Hopkins, UMD, UT Austin, UCDavis, Penn State, Rochester, UNC, UVA, Vanderbilt, WUSTL. The fields are categorized by journal according to Table 2.

Table 2: Journals by Field

| | |
|---|---|
| General Interests | American Economic Review |
| | Econometrica |
| | Journal of Political Economy |
| | Quarterly Journal of Economics |
| | Review of Economic Studies |
| | Journal of Economic Perspectives |
| | Journal of Economic Literature |
| | European Economic Review |
| | Economic Journal |
| | Journal of the European Economic Association |
| | Canadian Journal of Economics |
| | American Economic Review: Insights |
| | International Economic Review |
| Applied Microeconomics | American Economic Journal: Economic Policy |
| | Journal of Labor Economics |
| | Journal of Real Estate Finance and Economics |
| | Journal of Applied Econometrics |
| | The Review of Economics and Statistics |
| | American Economic Journal: Applied Economics |
| | The Journal of Economic Education |
| | Journal of Population Economics |
| | Journal of Environmental Economics and Management |
| | Journal of Development Economics |
| | Journal of Health Economics |
| | Journal of Human Resources |
| | Journal of Public Economics |
| | Journal of Urban Economics |
| | The Journal of Law and Economics |
| | Journal of International Economics |
| | American Journal of Agricultural Economics |
| Finance | The Journal of Finance |
| | Journal of Financial Economics |
| | Review of Financial Studies |
| Economic History | History of Political Economy |

| | |
|---|---|
| | The Economic History Review<br>The Journal of Economic History<br>Explorations in Economic History |
| Macroeconomics | Review of Economic Dynamics<br>Journal of Economic Dynamics and Control<br>Journal of Monetary Economics<br>Journal of Money, Credit and Banking<br>Macroeconomic Dynamics<br>American Economic Journal: Macroeconomics<br>Journal of Economic Growth |
| Econometrics | Quantitative Economics<br>Journal of Business & Economic Statistics<br>Econometric Theory<br>Journal of Econometrics |
| Microeconomic Theory | American Economic Journal: Microeconomics<br>Games and Economic Behavior<br>Journal of Regulatory Economics<br>Journal of Economics and Management Strategy<br>The Journal of Industrial Economics<br>The RAND Journal of Economics<br>Journal of Economic Behavior and Organization<br>Journal of Economic Theory<br>Social Choice and Welfare<br>Theoretical Economics<br>Experimental Economics<br>International Journal of Industrial Organization<br>Economic Theory<br>Journal of Mathematical Economics<br>Journal of Risk and Uncertainty<br>Journal of Comparative Economics |

Table 4: Estimated Yearly Trend of Number of Authors from 2001 to 2019

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | % 1 Author | % 2 Authors | % 3 Authors | % 4+ Authors |
| Yearly Trend | -1.00*** | -0.55*** | 0.88*** | 0.67*** |
| | (0.02) | (0.02) | (0.02) | (0.01) |
| NBER | -22.38*** | 1.40*** | 12.34*** | 8.65*** |
| | (0.25) | (0.34) | (0.32) | (0.23) |
| SSRN | -29.70*** | -5.70*** | 22.07*** | 13.32*** |
| | (0.33) | (0.37) | (0.32) | (0.20) |
| Top 5 | -8.63*** | 4.45*** | 2.69*** | 1.48*** |
| | (0.57) | (0.65) | (0.55) | (0.34) |
| Number of Papers | 282,156 | 282,156 | 282,156 | 282,156 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. We also estimated a version with a quadratic specification of trend; the results are not qualitatively different.

Table 5: Estimated of Yearly Trend of Institutional Assortativity in Collaboration from 2001 to 2019

| | (1) All | (2) Working Papers | (3) Publications |
|---|---|---|---|
| Panel A | | % Papers Inter-Institutional | |
| Yearly Trend | 0.18*** | 0.07*** | 0.42*** |
| | (0.02) | (0.02) | (0.04) |
| NBER | 23.03*** | | |
| | (0.30) | | |
| SSRN | 8.38*** | -14.64*** | |
| | (0.23) | (0.25) | |
| T5 | -1.54** | | -1.47** |
| | (0.71) | | (0.71) |
| Number of Papers | 201,968 | 144,183 | 57,785 |
| Panel B | | % Authors in Major Instiution | |
| Yearly Trend | -0.12*** | -0.10*** | -0.14*** |
| | (0.01) | (0.01) | (0.02) |
| NBER | -9.10*** | | |
| | (0.14) | | |
| SSRN | -3.62*** | 5.49*** | |
| | (0.10) | (0.11) | |
| T5 | -1.15*** | | -1.15*** |
| | (0.33) | | (0.33) |
| Number of Papers | 154,546 | 114,458 | 40,088 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models in panel A are estimated with only multi-authored papers. Models in panel B are further restricted to include only inter-institutional papers. We also estimated a version with a quadratic specification of trend; the results are not qualitatively different.

Table 6: Estimated Yearly Trend of Institutional Assortativity in Collaboration from 2001 to 2019

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Sample | SSRN | NBER | EC64 | T5 |
| Panel A | | % Papers Inter-Institutional | | |
| Yearly Trend | 0.02 | 0.49*** | 0.38*** | 0.77*** |
| | (0.02) | (0.04) | (0.04) | (0.12) |
| Number of Papers | 130,126 | 14,057 | 53,159 | 4,626 |
| Panel B | | % Authors in Major Institution | | |
| Yearly Trend | -0.06*** | -0.40*** | -0.13*** | -0.35*** |
| | (0.01) | (0.02) | (0.02) | (0.05) |
| Number of Papers | 101,443 | 13,015 | 36,946 | 3,142 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Models in panel A are estimated with only multi-authored papers. Models in panel B are further restricted to include only inter-institutional papers.

Table 7: Estimated Yearly Trend of Experience Assortativity in Multi-Authored Papers from 2001 to 2019

| | (1) SSRN | (2) NBER | (3) EC64 | (4) T5 |
|---|---|---|---|---|
| | % Papers with All Junior Authors | | | |
| Yearly Trend | -0.65*** | -0.39*** | -0.31*** | -0.51*** |
| | (0.02) | (0.04) | (0.04) | (0.11) |
| | % Papers with All Senior Authors | | | |
| Yearly Trend | -0.13*** | -0.16*** | -0.07** | -0.07 |
| | (0.01) | (0.04) | (0.03) | (0.10) |
| Number of Papers | 130,126 | 24,140 | 53,159 | 4,626 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models are estimated with only two- and three-authored papers.

Table 8: Estimated Deviations of Number of Authors from Linear Yearly Trend During COVID from 2001 to 2023

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | % 1 Author | % 2 Authors | % 3 Authors | % 4+ Authors |
| 2020 | -0.49 | 1.31* | -1.03 | 0.21 |
| | (0.61) | (0.68) | (0.64) | (0.47) |
| 2021 | -2.68*** | 2.59*** | -1.42** | 1.50*** |
| | (0.57) | (0.67) | (0.63) | (0.49) |
| 2022 | -0.37 | 1.10 | -1.76*** | 1.03** |
| | (0.61) | (0.69) | (0.67) | (0.51) |
| 2023 | -0.20 | 0.35 | -2.79*** | 2.64*** |
| | (0.61) | (0.69) | (0.67) | (0.54) |
| Working Paper × | | | | |
| 2020 | 2.13*** | -4.37*** | -2.12*** | 4.36*** |
| | (0.66) | (0.77) | (0.74) | (0.57) |
| 2021 | 5.74*** | -7.08*** | -2.62*** | 3.95*** |
| | (0.64) | (0.78) | (0.75) | (0.60) |
| 2022 | 4.16*** | -4.38*** | -3.02*** | 3.25*** |
| | (0.67) | (0.80) | (0.78) | (0.62) |
| 2023 | 6.14*** | -10.64*** | -5.74*** | 10.25*** |
| | (0.80) | (1.18) | (1.30) | (1.25) |
| Working Paper | -11.62*** | 4.31*** | 6.27*** | 1.05*** |
| | (0.19) | (0.21) | (0.18) | (0.11) |
| Yearly Trend | -1.00*** | -0.55*** | 0.88*** | 0.67*** |
| | (0.02) | (0.02) | (0.02) | (0.01) |
| Number of Papers | 344,763 | 344,763 | 344,763 | 344,763 |
| Sample Fixed Effects | × | × | × | × |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. SSRN sample is from 2004 to 2022, the rest of the samples are from 2001 to 2023.

Table 9: Estimated Deviations of Inter-Institutional Collaboration from Yearly Trend During COVID from 2001 to 2023

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Sample | SSRN | NBER | EC64 | T5 |
| 2020 | 0.87* | -1.94*** | 0.77 | -3.42 |
| | (0.46) | (0.68) | (0.80) | (2.75) |
| 2021 | 0.45 | -1.76** | 0.01 | -2.83 |
| | (0.51) | (0.74) | (0.79) | (2.74) |
| 2022 | -0.16 | -1.00 | 1.90** | -5.09* |
| | (0.53) | (0.70) | (0.82) | (2.86) |
| 2023 | 1.58*** | -2.28*** | 2.18*** | -0.79 |
| | (0.58) | (0.77) | (0.83) | (2.75) |
| Yearly Trend | 0.02 | 0.49*** | 0.38*** | 0.77*** |
| | (0.02) | (0.04) | (0.04) | (0.12) |
| Number of Papers | 163,663 | 18,664 | 69,181 | 6,027 |
| 2019 Mean | 78.22 | 96.37 | 72.51 | 75.24 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The models are estimated with only multi-authored papers. We also estimated a version with a quadratic specification of trend; the results are not qualitatively different.

Table 10: Estimated Deviations of % Authors in Major Institution Yearly Trend During COVID from 2001 to 2023

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Sample | SSRN | NBER | EC64 | T5 |
| 2020 | -0.07 | -0.74* | -0.87** | -1.12 |
| | (0.18) | (0.42) | (0.39) | (1.25) |
| 2021 | 0.08 | -0.90* | -1.01*** | 1.06 |
| | (0.19) | (0.47) | (0.38) | (1.25) |
| 2022 | 0.14 | -1.07** | -0.02 | -0.23 |
| | (0.20) | (0.49) | (0.40) | (1.26) |
| 2023 | 0.37 | 0.32 | -0.29 | 0.76 |
| | (0.23) | (0.53) | (0.42) | (1.37) |
| Yearly Trend | -0.06*** | -0.40*** | -0.13*** | -0.35*** |
| | (0.01) | (0.02) | (0.02) | (0.05) |
| Number of Papers | 127,865 | 17,429 | 48,919 | 4,169 |
| 2019 Mean | 48.59 | 39.95 | 51.06 | 46.80 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
The models are estimated with only inter-institutional multi-authored papers.
We also estimated a version with a quadratic specification of trend; the results are not qualitatively different.

Table 11: Estimated Correlations Between Returns and Popularity of Multi-Authored Papers, 1950-2018

| | | | | |
|---|---|---|---|---|
| | | | $\Delta LR2S_t$ | |
| $LR2I_{t-5}$ | -0.0068 | -0.0293 | -0.0615 | -0.0748 |
| | (0.0258) | (0.0364) | (0.0485) | (0.0525) |
| $LR2S_{t-5}$ | | 0.0416 | | -0.0780 |
| | | (0.0416) | | (0.0809) |
| Year | | | 0.0017 | 0.0034 |
| | | | (0.0010) | (0.0021) |
| | | | $\Delta LR3S_t$ | |
| $LR3I_{t-5}$ | 0.0043 | -0.0353 | -0.0634* | -0.0638* |
| | (0.0244) | (0.0341) | (0.0382) | (0.0387) |
| $LR3S_{t-5}$ | | 0.0499 | | -0.0583 |
| | | (0.0310) | | (0.0627) |
| Year | | | 0.0033** | 0.0056* |
| | | | (0.0014) | (0.0029) |
| | | | $\Delta LR4S_t$ | |
| $LR4I_{t-5}$ | 0.0690** | 0.0744* | 0.0666* | 0.0712* |
| | (0.0347) | (0.0409) | (0.0387) | (0.0410) |
| $LR4S_{t-5}$ | | 0.0077 | | -0.0517 |
| | | (0.0508) | | (0.0715) |
| Year | | | 0.0011 | 0.0024 |
| | | | (0.0016) | (0.0023) |
| Number of Years | 61 | 61 | 61 | 61 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. A homerun paper is a paper that has a 5-year EC64 citation in the top 10 percentile among EC64 papers published in that year.

# Appendix A Additional Figures

Figure A1: The Number of Authors on a Paper, Top 5 Journals, 1886-2023

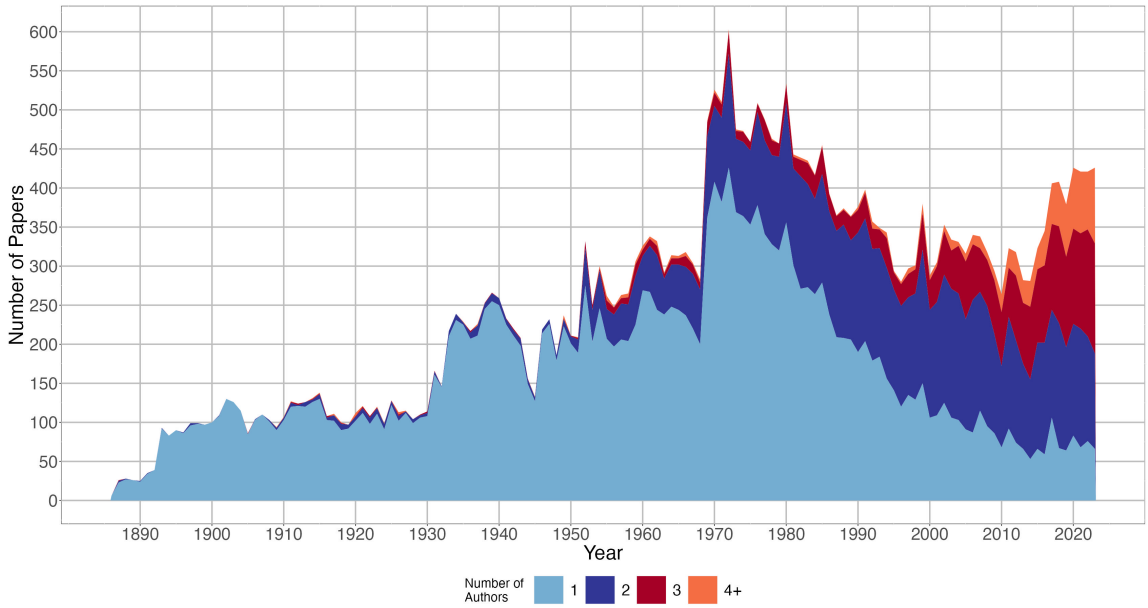



Figure A2: Overall Decrease of Experience Assortativity in EC64 Papers

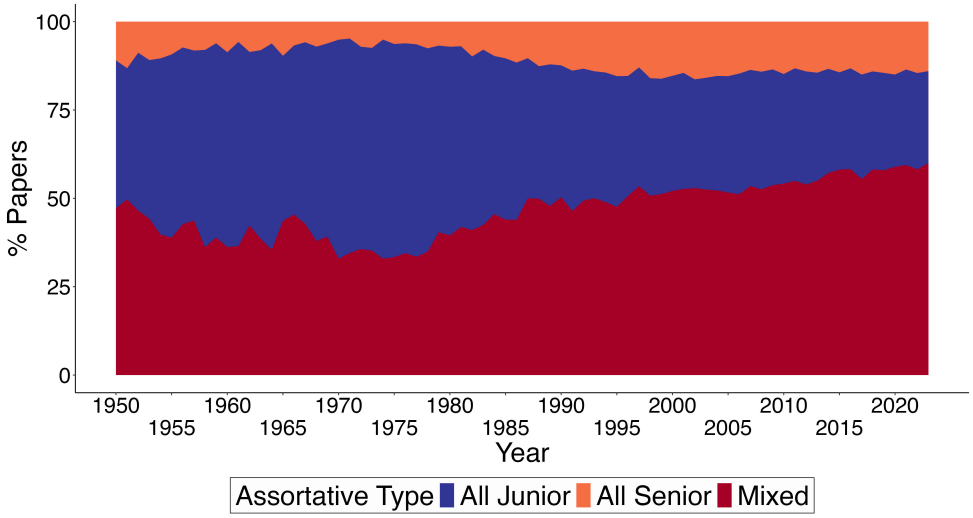


Figure A3: Evolution of Estimated Returns to Number of Authors, 5-Year Periods




Note: Estimating equation is equation (10) for each quinquennial.
Each tick on the x-axis represents the 5-year period starting that year.

Figure A4: The Number of Authors on a Paper, T5 Journals Excluding AER and ECMA, 1886-2023

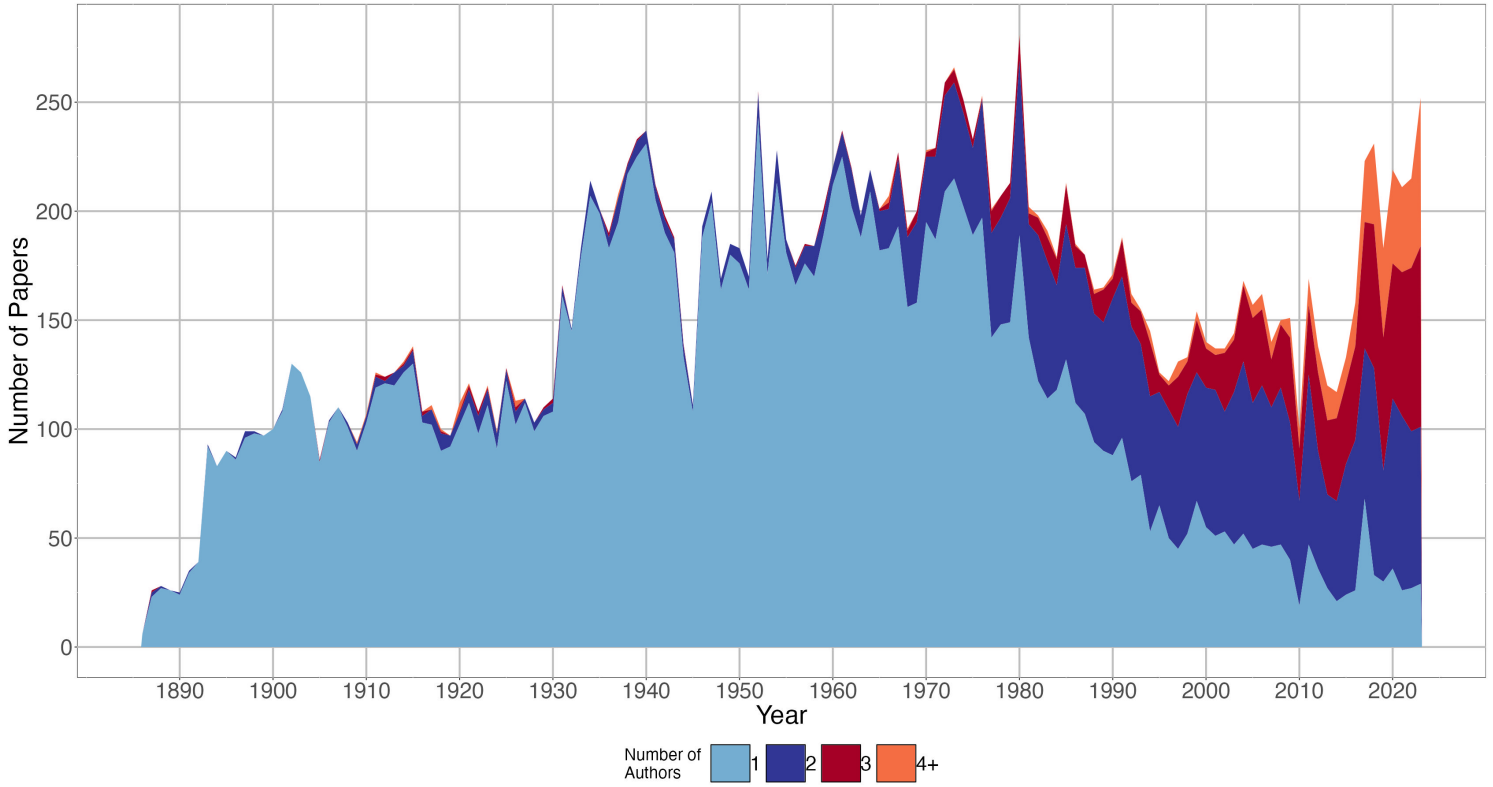
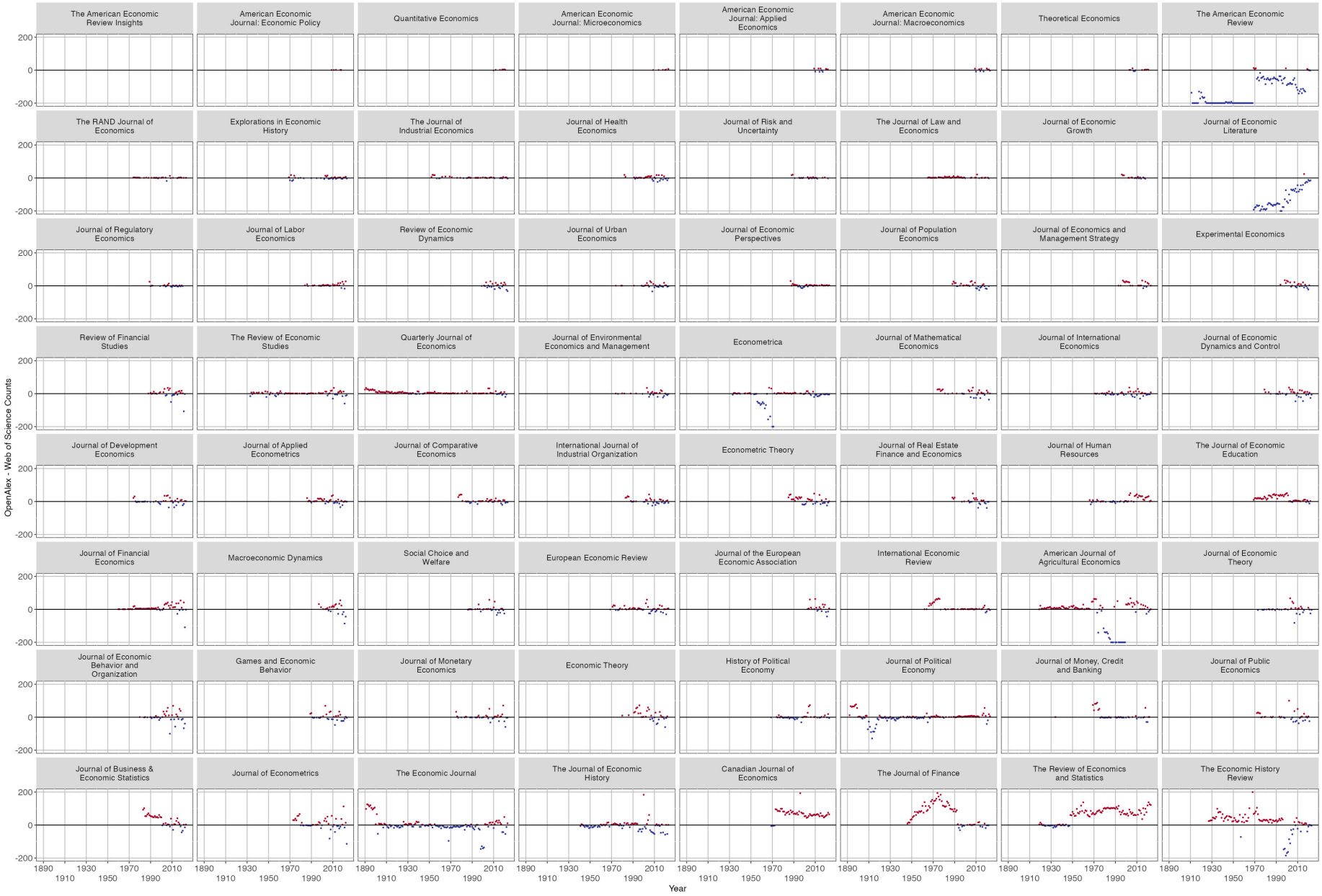


Figure A5: Differences Between OpenAlex and Web of Science Records by Journal



# Appendix B Additional Tables

Table B1: Estimated Linear Yearly Trend of Number of Authors from 2001 to 2019

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | % 1 Author | % 2 Authors | % 3 Authors | % 4+ Authors |
| Panel A: *SSRN* | | | | |
| Yearly Trend | -0.93*** | -0.64*** | 0.91*** | 0.66*** |
| | (0.02) | (0.02) | (0.02) | (0.01) |
| Number of Papers | 167,367 | 167,367 | 167,367 | 167,367 |
| Panel B: *NBER* | | | | |
| Yearly Trend | -0.73*** | -1.41*** | 0.80*** | 1.33*** |
| | (0.04) | (0.05) | (0.05) | (0.04) |
| Number of Papers | 27,261 | 27,261 | 27,261 | 27,261 |
| Panel C: *Top 5* | | | | |
| Yearly Trend | -0.97*** | -0.75*** | 0.91*** | 0.81*** |
| | (0.10) | (0.11) | (0.09) | (0.06) |
| Number of Papers | 6,258 | 6,258 | 6,258 | 6,258 |
| Panel D: *EC64* | | | | |
| Yearly Trend | -1.22*** | -0.11*** | 0.87*** | 0.46*** |
| | (0.03) | (0.03) | (0.03) | (0.02) |
| Number of Papers | 81,270 | 81,270 | 81,270 | 81,270 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. We also estimated a version with a quadratic specification of trend; the results are not qualitatively different.

Table B2: Predictors of Homerun Papers, Excluding T5 Papers , 1900-2018

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | % Likelihood of Becoming a Homerun Paper | | | |
| Number of Authors | | | | |
| 2 | 2.42*** | 2.81*** | 2.15*** | 2.46*** |
| | (0.16) | (0.16) | (0.16) | (0.16) |
| 3 | 2.74*** | 3.59*** | 2.41*** | 2.98*** |
| | (0.22) | (0.23) | (0.23) | (0.23) |
| 4+ | 0.58* | 1.27*** | 0.52 | 1.59*** |
| | (0.35) | (0.35) | (0.35) | (0.35) |
| Inter-Institutional | 0.69*** | -0.17 | -0.57*** | -1.24*** |
| | (0.15) | (0.15) | (0.15) | (0.15) |
| Has Senior Author | -0.74*** | -0.44*** | -0.10 | -0.10 |
| | (0.14) | (0.14) | (0.14) | (0.14) |
| Has Top 10 Author | 11.42*** | 11.47*** | 10.68*** | 8.96*** |
| | (0.26) | (0.26) | (0.26) | (0.26) |
| Has Top 10 Senior | 0.00 | -0.11 | -0.10 | -0.12 |
| | (0.35) | (0.35) | (0.35) | (0.35) |
| Has 11-30 Author | 5.99*** | 6.16*** | 5.21*** | 4.06*** |
| | (0.26) | (0.26) | (0.26) | (0.25) |
| Has 11-30 Senior | 0.05 | 0.03 | 0.29 | 0.29 |
| | (0.34) | (0.34) | (0.34) | (0.34) |
| US Institution | 1.08*** | 1.47*** | 1.85*** | 2.06*** |
| | (0.14) | (0.14) | (0.14) | (0.15) |
| International Collab. | -0.71*** | 0.50** | 0.69*** | 0.49** |
| | (0.21) | (0.22) | (0.22) | (0.22) |
| Number of Papers | 180,906 | 180,906 | 180,906 | 180,906 |
| Year FE | | × | × | × |
| Field FE | | | × | |
| Journal FE | | | | × |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. A homerun paper is a paper that has a 5-year EC64 citation in the top 10 percentile among EC64 papers published in that year. A senior author is an author whose first EC64 publication was published ten or more years ago. Top 10 schools are Harvard, MIT, Yale, Stanford, Columbia, Princeton, UChicago, UPenn, Northwestern, and UC Berkeley. 11-30 schools are NYU, UCLA, UMich, Cal Tech, Cornell, UCSD, UW Madison, Duke, UMinn, Brown, CMU, BU, Johns Hopkins, UMD, UT Austin, UCDavis, Penn State, Rochester, UNC, UVA, Vanderbilt, WUSTL. The fields are categorized by journal into General, Applied, Theory, Metrics, Macro, Finance, and History.

Table B3: Predictors of Homerun Papers, 1900-2018

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | % Likelihood of Becoming a Homerun Paper | | | |
| Number of Authors | | | | |
| 2 | 2.34*** | 2.90*** | 2.68*** | 3.83*** |
| | (0.16) | (0.17) | (0.16) | (0.17) |
| 3 | 4.24*** | 5.12*** | 4.41*** | 5.69*** |
| | (0.24) | (0.24) | (0.24) | (0.24) |
| 4+ | 2.52*** | 3.27*** | 2.95*** | 5.00*** |
| | (0.38) | (0.38) | (0.37) | (0.37) |
| Inter-Institutional | 1.46*** | 0.74*** | -0.11 | -1.38*** |
| | (0.15) | (0.15) | (0.15) | (0.15) |
| Has Senior Author | -1.18*** | -0.93*** | -1.17*** | -0.92*** |
| | (0.14) | (0.14) | (0.14) | (0.13) |
| Majority Author Top 10 | 15.89*** | 15.85*** | 13.58*** | 10.90*** |
| | (0.20) | (0.20) | (0.20) | (0.20) |
| Majority Author 11-30 | 7.92*** | 8.04*** | 6.92*** | 5.04*** |
| | (0.20) | (0.20) | (0.19) | (0.19) |
| US Institution | 0.78*** | 1.14*** | 2.74*** | 3.33*** |
| | (0.15) | (0.15) | (0.15) | (0.16) |
| International Collab. | 0.01 | 0.83*** | 0.83*** | 0.71*** |
| | (0.23) | (0.23) | (0.23) | (0.23) |
| Number of Papers | 213,947 | 213,947 | 213,947 | 213,947 |
| Year FE | | × | × | × |
| Field FE | | | × | |
| Journal FE | | | | × |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. A homerun paper is a paper that has a 5-year EC64 citation in the top 10 percentile among EC64 papers published in that year. Top 10 schools are Harvard, MIT, Yale, Stanford, Columbia, Princeton, UChicago, UPenn, Northwestern, and UCBerkeley. 11-30 schools are NYU, UCLA, UMich, Cal Tech, Cornell, UCSD, UW Madison, Duke, UMinn, Brown, CMU, BU, Johns Hopkins, UMD, UT Austin, UCDavis, Penn State, Rochester, UNC, UVA, Vanderbilt, WUSTL. The fields are categorized by journal into General, Applied, Theory, Metrics, Macro, Finance, and History.

Table B4: Predictors of Citation Percentile among Papers Published in the Same Year, 1900-2018

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Citation Percentile (Higher = More Citations) | | | |
| Number of Authors | | | | |
| 2 | 5.33*** | 6.26*** | 5.38*** | 6.06*** |
| | (0.15) | (0.16) | (0.15) | (0.16) |
| 3 | 5.60*** | 7.15*** | 5.76*** | 7.14*** |
| | (0.22) | (0.23) | (0.23) | (0.22) |
| 4+ | 2.11*** | 3.51*** | 2.70*** | 5.33*** |
| | (0.36) | (0.36) | (0.35) | (0.35) |
| Inter-Institutional | 1.38*** | 0.24 | -0.56*** | -1.99*** |
| | (0.14) | (0.15) | (0.14) | (0.14) |
| Has Senior Author | -2.42*** | -2.00*** | -1.70*** | -1.33*** |
| | (0.14) | (0.15) | (0.14) | (0.14) |
| Has Top 10 Author | 14.26*** | 14.21*** | 12.05*** | 8.44*** |
| | (0.25) | (0.25) | (0.24) | (0.24) |
| Has Top 10 Senior | -0.88*** | -1.02*** | -1.36*** | -1.00*** |
| | (0.33) | (0.33) | (0.33) | (0.32) |
| Has 11-30 Author | 9.13*** | 9.28*** | 7.90*** | 5.48*** |
| | (0.26) | (0.26) | (0.25) | (0.24) |
| Has 11-30 Senior | -0.67* | -0.70** | -0.57* | -0.60* |
| | (0.34) | (0.34) | (0.33) | (0.32) |
| US Institution | 1.65*** | 2.23*** | 3.59*** | 4.96*** |
| | (0.14) | (0.14) | (0.14) | (0.15) |
| International Collab. | 3.54*** | 5.12*** | 4.51*** | 3.85*** |
| | (0.22) | (0.22) | (0.22) | (0.21) |
| Number of Papers | 213,947 | 213,947 | 213,947 | 213,947 |
| Year FE | | × | × | × |
| Field FE | | | × | |
| Journal FE | | | | × |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. A homerun paper is a paper that has a 5-year EC64 citation in the top 10 percentile among EC64 papers published in that year. Top 10 schools are Harvard, MIT, Yale, Stanford, Columbia, Princeton, UChicago, UPenn, Northwestern, and UCBerkeley. 11-30 schools are NYU, UCLA, UMich, Cal Tech, Cornell, UCSD, UW Madison, Duke, UMinn, Brown, CMU, BU, Johns Hopkins, UMD, UT Austin, UCDavis, Penn State, Rochester, UNC, UVA, Vanderbilt, WUSTL. The fields are categorized by journal into General, Applied, Theory, Metrics, Macro, Finance, and History.

Table B5: Estimated Correlations Between Returns and Popularity of Multi-Authored Papers, 1950-2018

| | | | | |
|---|---|---|---|---|
| | | | $\Delta LR2S_t$ | |
| $LR2R_{t-5}$ | 0.0048 | -0.0293 | -0.0726 | -0.0748 |
| | (0.0151) | (0.0364) | (0.0496) | (0.0525) |
| $LR2S_{t-5}$ | | 0.0708 | | -0.0032 |
| | | (0.0716) | | (0.0793) |
| Year | | | 0.0032* | 0.0034 |
| | | | (0.0019) | (0.0021) |
| | | | $\Delta LR3S_t$ | |
| $LR3R_{t-5}$ | 0.0111 | -0.0353 | -0.0617* | -0.0638* |
| | (0.0126) | (0.0341) | (0.0342) | (0.0387) |
| $LR3S_{t-5}$ | | 0.0852 | | 0.0055 |
| | | (0.0599) | | (0.0702) |
| Year | | | 0.0056** | 0.0056* |
| | | | (0.0024) | (0.0029) |
| | | | $\Delta LR4S_t$ | |
| $LR4R_{t-5}$ | 0.0389* | 0.0744* | 0.0300 | 0.0712* |
| | (0.0222) | (0.0409) | (0.0295) | (0.0410) |
| $LR4S_{t-5}$ | | -0.0667 | | -0.1230 |
| | | (0.0799) | | (0.0928) |
| Year | | | 0.0010 | 0.0024 |
| | | | (0.0020) | (0.0023) |
| Number of Years | 61 | 61 | 61 | 61 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. A homerun paper is a paper that has a 5-year EC64 citation in the top 10 percentile among EC64 papers published in that year.

Table B6: Estimated Correlations Between Returns and Popularity of Multi-Authored Papers, 1950-2018

| | | | | |
|---|---|---|---|---|
| | $\Delta$ Log Ratio 2 Authors to 1 Author (t) | | | |
| $\Delta$LR2R(t-5) | -0.1033*** | -0.1165*** | -0.1035*** | -0.1176*** |
| | (0.0399) | (0.0403) | (0.0399) | (0.0403) |
| $\Delta$Log Ratio 2:1 (t-5) | | 0.1711* | | 0.1648 |
| | | (0.1010) | | (0.1016) |
| Year | | | 0.0004 | 0.0003 |
| | | | (0.0006) | (0.0006) |
| | $\Delta$ Log Ratio 3 Authors to 1 Author (t) | | | |
| $\Delta$LR3R (t-5) | -0.0334 | -0.0233 | -0.0393 | -0.0293 |
| | (0.0363) | (0.0362) | (0.0368) | (0.0364) |
| $\Delta$Log Ratio 3:1 (t-5) | | -0.0318 | | -0.0633 |
| | | (0.1090) | | (0.1122) |
| Year | | | 0.0012 | 0.0013 |
| | | | (0.0010) | (0.0010) |
| | $\Delta$ Log Ratio 4 Authors to 1 Author (t) | | | |
| $\Delta$R4R (t-5) | 0.0790** | 0.0643* | 0.0748** | 0.0623 |
| | (0.0374) | (0.0386) | (0.0373) | (0.0386) |
| $\Delta$Log Ratio 4:1 (t-5) | | 0.1737* | | 0.1495 |
| | | (0.1024) | | (0.1040) |
| Year | | | 0.0021 | 0.0018 |
| | | | (0.0015) | (0.0015) |
| Number of Years | 57 | 57 | 57 | 57 |

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. A homerun paper is a paper that has a 5-year EC64 citation in the top 10 percentile among EC64 papers published in that year.