

Publish or Perish: Selective Attrition as a Unifying Explanation for Patterns in Innovation
over the Career

Huifeng Yu, Gerald Marschke, Matthew B. Ross, Joseph Staudt, Bruce Weinberg

Online Appendix

Appendix A: Model

Researchers work (as researchers) a maximum of two periods. In each period, t , a new generation of young researchers begin their research career. The young researchers conduct research in period t and at t 's end either continue on as an old researcher in $t+1$, or attrit to a non-research job. At the end of period $t+1$, all researchers retire. Because new researchers are “born” in each period, in each period young researchers overlap with old researchers. Research output is comprised of research papers, with each young researcher writing a single paper with one old researcher. The number of papers each old researcher writes is the ratio of young researchers to old researchers.

Each researcher has research ability α , which is initially unobserved by old researchers and the funding agency, but is revealed at the end of first period of her career. Young researchers have either high-ability with ability α_H or low-ability with ability α_L , where $\alpha_H > \alpha_L$. The ability type of the young is unobserved but the fraction of H -types among the young is known and in each generation is p .

One way to have selective attrition in this model is to assume that science requires external funding and thus a funding agency that divides a budget in each period between young and old researchers. If an individual receives 1 unit of funding, she works as a researcher. Otherwise she takes a non-researcher job. The agency's budget, expressed in people who can be supported, in each period t is N_t . Because each researcher requires 1 unit of funding, in period t , there are N_{yt} young researchers and N_{ot} old researchers, with N_{yt} and N_{ot} chosen by the funding agency such that $N_{ot} + N_{yt} \leq N_t$. The funding agency only offers high-ability researchers funding for a second period. The young researchers brought in are useful to the scientific enterprise because they pair

with old researchers to write papers, but also, they (potentially) become next period's old researchers who lead research projects. Define the stock of potential old researchers at the beginning of t as $K_t = pN_{yt-1}$ and thus $N_{ot} = K_t - x_t$, where x_t ($0 \leq x_t \leq K_t$) is the number of high-ability researchers entering at the end of their first period who are not funded in their second period.

Each old scientist has a unit of effort to allocate to her papers. Thus if there are N_{yt} young researchers and N_{ot} old researchers, there are N_{yt} papers (one for each young researcher) with each old researcher writing $\frac{N_{yt}}{N_{ot}}$ papers and devoting $\varepsilon_t = \frac{N_{ot}}{N_{yt}}$ units of effort per paper, which determines the quality of articles. The scientific contribution embodied by a paper is assumed linear in the scientist's effectiveness adjusting for effort: $\delta f(\varepsilon_t)\alpha_H + \alpha_L$ and $\delta f(\varepsilon_t)\alpha_H + \alpha_H$ for papers with low- and high-ability young scientists, respectively, where $f(\varepsilon_t)$ is assumed rising in ε_t but at a decreasing rate. Assuming for example $f'(0) = \infty$ guarantees the funding agency supports at least some old researchers.

Over a researcher's career, the quantity of research increases. A stable N_{yt} implies that in each t $N_{ot} < N_{yt}$. Thus, while each young scientist writes one article, each older scientist supervises several. Note the number of articles published by a (two-period) scientist increases over her career from 1 to $\frac{N_{yt}}{N_{ot}}$.

Over a researcher's career, the quality of research falls. This is by construction. Because she works only with high-ability researchers as a young researcher and a mix of abilities as an old

researcher, her average research quality must fall. Research quality falls by $(\alpha_H - \alpha_L)(1 - p)$.¹ The quality decline over the career is greater the higher her chances of matching to low-ability workers when older and the greater the ability difference high to low-ability workers. Note that this result holds even if there is no decline in ability, $\delta = 1$.

Over a researcher's career, her effectiveness on any paper falls, if $\delta f(\varepsilon_t) < 1$. With $\delta \in [0,1)$ her effectiveness on papers falls from α_H when young to $\delta f(\varepsilon_t)\alpha_H$ when older, declining both because she is spread more thinly in her supervisory role and because of the age-related decline in her skills, though she may still be contributing more to the paper than her young co-author. If her human capital accumulation exceeds depreciation ($\delta > 1$), her effectiveness on papers would still fall if she were sufficiently burdened with young researchers to supervise.

Change in effectiveness and ability of cohort with age. Mean cohort ability rises with age and this is by construction (because low-ability researchers are never funded for a second period). The mean effectiveness of the cohort rises with age if $(1 - \delta)$ is small relative to the number of low ability researchers and their ability disadvantage. That is,

$$\delta \alpha_H > [\alpha_L(1 - p) + \alpha_H p] \text{ iff } (1 - \delta)\alpha_H < (1 - p)[\alpha_H - \alpha_L].$$

The cohort's mean effectiveness rises of course when $\delta > 1$.

¹ $(\alpha_H - \alpha_L)(1 - p) = [\delta f(\tau_t)\alpha_H + \alpha_H] - \{[\delta f(\tau_t)\alpha_H + \alpha_L](1 - p) + [\delta f(\tau_t)\alpha_H + \alpha_H]p\}$.

Appendix B: Sample Construction and Description

Appendix Table B.1: Sample Creation

Sample	Obs.
MEDLINE	22,376,714
SCIE-MEDLINE Intersection	15,085,762
1980-2009	9,897,775
Research articles	7,198,087
Multi-author articles	6,691,006
Articles without truncated author lists	6,648,200
First author or last author experience not missing	6,399,621
First author or last author aged less than 40	6,186,312
Last author not aged 0	5,665,869
Team size less than or equal to 13	5,613,189

Appendix Table B.2: Variable Description

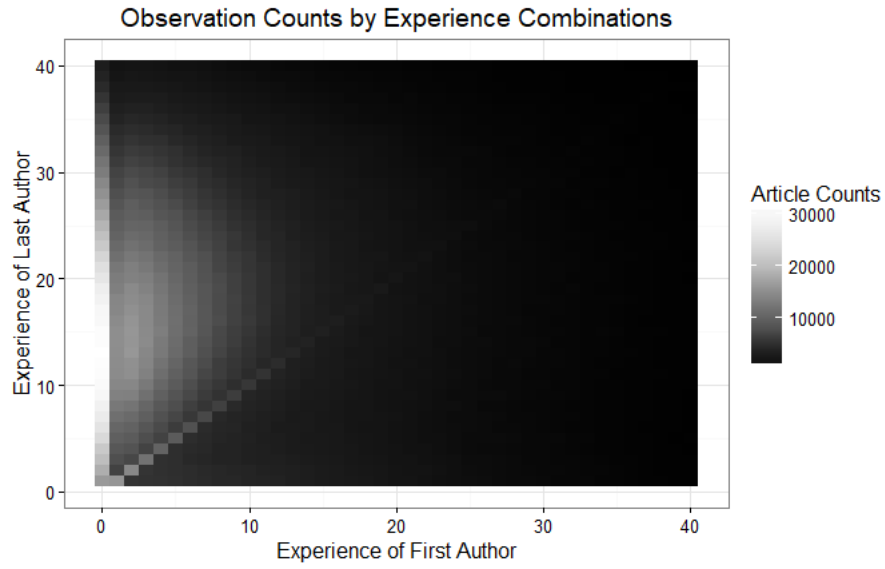
Variables	Description	Data Source
Utilization of Work		
Number of Citations (High Impact)	Total citations received by an article (forward citations) within MEDLINE until May 20, 2014	WoS
Disciplinary Diversity of Citations (Broad Impact)	One minus Herfindahl index of the field diversity of forward citations	WoS/MEDLINE
Article Inputs		
Mean Age of References (Distance from Frontier)	Difference between the article's publication year and the publication year of an article it cites, averaged over all articles it cites.	WoS
Disciplinary Diversity of References (Multi-Disciplinary)	One minus Herfindahl index of the field diversity of references.	WoS/MEDLINE
Number of New Concepts Used (Builds on new ideas)	The number of recent, top 0.1 percent concepts that an article's title or abstract uses (the vintage year of a recent concept is within 5 years of the article's publication year)	MEDLINE
Number of Old Concepts Used	The number of older top 0.1 percent concepts that an article's title or abstract uses (the vintage year of an older concept is more than 5 years before the article's publication year)	MEDLINE
Mean Age of Concepts Used	Mean age of all concepts used by an article	MEDLINE
Number of References	Total references cited by an article	WoS
Mean Citation of Recent References (log)	Log of mean of forward citations of recent references (references published within 5 years of the citing article's publication year)	WoS
Mean Citation of Old References (log)	Log of mean of forward citations of older references (references published more than 5 years before the citing article's publication year)	WoS
Diff. in Citations to Recent and Old References (log)	Difference in log of mean forward citations of recent references and log of mean forward citations of older references	WoS
Team Characteristics		
Number of Authors	Number of authors on the article	MEDLINE
Size of Team's Network	Count of distinct coauthors of all authors	Author-ity
Number of Authors' Cities	Number of distinct cities among author affiliation addresses	WoS
Mean Experience of Middle Author	Mean experience of all authors except for first and last author, if all middle authors are disambiguated	Author-ity
First Author Publication Count	First author's publication count over the 5-year period when this article is published	Author-ity
Last Author Publication count	Last author's publication count over the 5-year period when this article is published	Author-ity
Individual Productivity		
Best Paper	Articles with the highest citations	WoS, Author-ity

Mean of Best N Paper	Mean citations of N articles with highest citations, N is defined as number of articles published during the first 5-year bin of authors' career	WoS, Author-ity
Experience		
1st/Last 0	Experience of first/last author in article's publication year is 0 years	Author-ity
1st/Last a to b	Experience of first/last author in article's publication year is between a and b years	Author-ity

Appendix Table B.3: Correlations of variables

	Number of Citations (log)	Disciplinary Diversity of Citations	Number of Authors	Number of Authors' Cities	Size of Team's Network	Number of References	Mean Age of References	Disciplinary Diversity of References	Number of New Concepts Used	Number of Old Concepts Used	Mean Age of Concepts Used
Number of Citations (log)	1										
Disciplinary Diversity of Citations	0.467	1									
Number of Authors	0.145	0.0985	1								
Number of Authors' Cities	0.0953	0.0398	0.371	1							
Size of Team's Network	0.171	0.105	0.529	0.305	1						
Number of References	0.277	0.170	0.000173	0.0441	0.00405	1					
Mean Age of References	-0.181	-0.140	-0.112	-0.0277	-0.133	-0.0199	1				
Disciplinary Diversity of References	0.141	0.342	0.108	0.0220	0.108	0.211	-0.168	1			
Number of New Concepts Used	0.220	0.0982	0.0849	0.0265	0.110	0.0774	-0.178	0.0785	1		
Number of Old Concepts Used	0.255	0.269	0.219	0.0396	0.238	0.259	-0.250	0.324	0.241	1	
Mean Age of Concepts Used	-0.177	-0.123	-0.140	-0.0238	-0.159	-0.0627	0.256	-0.128	-0.303	-0.403	1

Note: All variables are transformed as the residual from a regression of the variable on time effects. The correlation matrix is produced from the subset of data for which no variables are missing. The observation count is 4,352,171. Forward citation counts are in (natural) log form.



Appendix Figure B.1: Article Counts by Experience Combinations

Note. The figure plots the number of articles written by people at each combination of first and last author experience.

Appendix C: Field Construction

Appendix C.1: Aggregating MeSH Terms to Construct Fields

We use the Medical Subject Headings (MeSH) that tag most articles in MEDLINE to characterize the fields to which each article belongs. There are 27,149 raw terms in the 2014 MeSH vocabulary and they vary widely in their descriptive detail. For instance, some articles are tagged with general terms such as “Body Regions” and some are tagged with more detailed terms such as “Peritoneal Stomata”. Thus, in order to construct comparable fields, we aggregate all MeSH terms to a similar level of descriptive detail.

To understand our aggregation method, first note that MeSH terms have a hierarchical structure. At the top of the hierarchy (first-level terms) are 16 very general terms such as “Anatomy”, “Organisms”, and “Diseases”. Each of these 16 first-level terms are identified by a unique capital letter. For instance, “Anatomy” is identified by the letter “A”, “Organisms” is identified by “B”, and so on. Beneath each of these first-level MeSH terms is a group of second-level MeSH terms. For instance, “Body Regions” is a second-level MeSH term beneath the top-level term “Anatomy”. Each second-level MeSH term is identified by the capital letter of the first-level MeSH term it is beneath and by two numbers. For instance, “Body Regions” is identified by “A01”. Beneath each second-level MeSH term is a group of third-level MeSH terms identified by the capital letter of the first-level term it is beneath, the two numbers of the second-level term it is beneath, and three subsequent numbers. For instance, “Anatomic Landmarks” is a third-level MeSH term under “Body Regions” and is identified as “A01.111”. This structure continues to depths of up to 12 levels.

Aggregating MeSH terms (that is, classifying lower level MeSH terms as a part of higher level MeSH terms) is not straightforward because some MeSH terms are beneath more than one higher level MeSH term and some articles can be tagged with multiple MeSH terms.

Consider the MeSH term “Asthma” This term has four separate identifiers: “C08.127.108”, “C08.381.495.108”, “C08.674.095”, and “C20.543.480.680.095”. Thus, “Asthma” falls under the first level MeSH term “Diseases” (identified by “C”). It also falls under the second-level terms “Respiratory Tract Diseases” (“C08”) and “Immune System Diseases” (“C20”). The problem arises because MEDLINE records only contain the MeSH terms themselves, not their identifiers. For instance, if a MEDLINE record is tagged with the MeSH term “Asthma”, we would not know if it was the “Asthma” that is beneath “Respiratory Tract Diseases” or “Immune System Diseases”. If we wanted to aggregate all MeSH terms to the second level, we would have to find a way to split “Asthma” between “Respiratory Tract Diseases” and “Immune System Diseases”. We opt for the straightforward method of assigning half to each higher-level term. If we wanted to aggregate all MeSH terms to the fourth level, “Asthma” would fall under the fourth-level terms

- Lung Diseases, Obstructive [C08.381.495]
- Hypersensitivity, Immediate [C20.543.480]
- Asthma [C08.127.108]
- Asthma [C08.674.095]

In this case, we assign a quarter of the raw term “Asthma” to each of these four fourth-level terms. Thus, overall, $1/4$ will be assigned to “Lung Diseases, Obstructive”, $1/4$ to “Hypersensitivity, Immediate”, and $1/4+1/4=1/2$ to “Asthma” itself.

Our last step is to apportion each article indexed in MEDLINE to the newly aggregated MeSH terms. Again, examples are the most illustrative. Suppose we come across an article that is tagged by “Asthma” and we want to aggregate to the fourth level. We know from above that $1/4$ will be assigned to “Lung Diseases, Obstructive”, $1/4$ to “Hypersensitivity, Immediate”, and $1/4+1/4=1/2$ to “Asthma” itself.

However, suppose (as is usually the case) that this article is also tagged with other MeSH terms. Specifically, suppose that the article is also tagged with the terms “Neck” (identified by A01.598) and “Health Information Exchange” (identified by L01.700.253, L01.399.500.500, L01.313.500.500, and E05.318.308.940.968.625.500.500). Figure C.1 illustrates this example. By the process discussed above, $1/4$ of “Health Information Exchange” will be assigned to each of the four fourth-level MeSH terms: “Health Information Exchange” itself (L01.700.253), “Health Information Management” (L01.399.500), “Medical Informatics” (L01.313.500), and “Data Collection” (E05.318.308). Since the lowest level of aggregation for “Neck” is the third level, we cannot assign it to a fourth-level term. We deal with this by simply eliminating “Neck”. It is too highly aggregated for our purposes.

We assume that each of the original remaining MeSH terms, “Asthma” and “Health Information Exchange” receive equal weight in characterizing the article. Under this assumption, the article will be apportioned to each fourth level MeSH term as follows:

- $1/2 * 1/4 = 1/8$ to “Lung Diseases, Obstructive”
- $1/2 * 1/4 = 1/8$ to “Hypersensitivity, Immediate”
- $1/2 * 1/4 = 1/8$ to “Asthma”

- $1/2 * 1/4 = 1/8$ to “Asthma”
- $1/2 * 1/4 = 1/8$ to “Health Information Exchange”
- $1/2 * 1/4 = 1/8$ to “Health Information Management”
- $1/2 * 1/4 = 1/8$ to “Medical Informatics”
- $1/2 * 1/4 = 1/8$ to “Data Collection”

Obviously, $1/8 + 1/8 + 1/8 + 1/8 + 1/8 + 1/8 + 1/8 + 1/8 = 1$. Thus, the article that was originally tagged by the three MeSH terms “Asthma”, “Neck” and “Health Information Exchange” is now apportioned between eight different fourth-level MeSH terms. In general, we apportion each MEDLINE article across aggregated MeSH terms in two stages. First, we equally apportion the original MeSH terms across the higher-level MeSH terms of which they are a part (e.g. apportion “Asthma” equally across “Lung Diseases, Obstructive”, “Hypersensitivity, Immediate”, “Asthma”, and “Asthma”). Second, we weight this apportionment by the inverse of the number of original MeSH terms of the proper level that tag the article (e.g. our hypothetical article was tagged by three original MeSH terms, but only two at the proper level of aggregation, and so we weight by $1/2$)

Appendix C.2: Detailed Description of Usage of MeSH Aggregation in Diversity Measures

Construction

Given the way we aggregate MeSH terms to define fields, one article may have multiple fields. Thus, we incorporate weights from the MeSH aggregation to construct our measures of disciplinary diversity of citations and disciplinary diversity of references. Here are the definitions of these two diversity measures.

Let $\underline{\mathcal{F}}$ denote the set of fields (indexed by \underline{f}) that article i 's citations belong to, $\mathbb{Z}_{i\underline{f}}$ denote the set of articles that cite article i (indexed by k) in field \underline{f} , and $w(\underline{f}|k)$ denote the weight of article k assigned to field \underline{f} . The weighted number of articles from field \underline{f} that cites article i is $\sum_{\forall k \in \mathbb{Z}_{i\underline{f}}} w(\underline{f}|k)$. The share of citations received by article i from field \underline{f} is then

$$s_{i\underline{f}} = \frac{\sum_{\forall k \in \mathbb{Z}_{i\underline{f}}} w(\underline{f}|k)}{\sum_{\forall \underline{f} \in \underline{\mathcal{F}}} \sum_{\forall k \in \mathbb{Z}_{i\underline{f}}} w(\underline{f}|k)}$$

and the disciplinary diversity of citations is

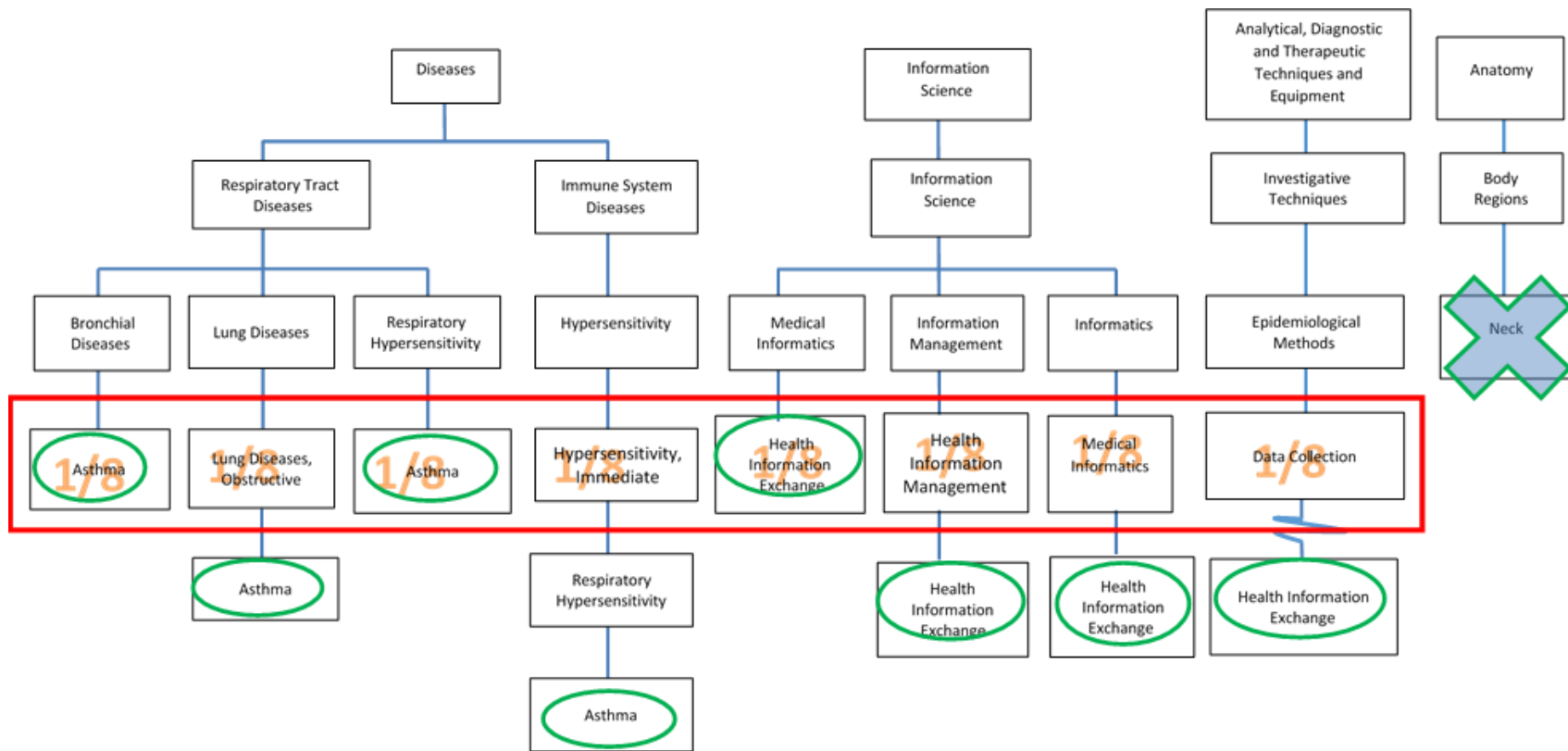
$$\text{Disciplinary Diversity of Citations}_i = \left(1 - \sum_{\forall \underline{f} \in \underline{\mathcal{F}}} s_{i\underline{f}}^2\right) \times 1000.$$

Let $\overline{\mathcal{F}}$ denote the set of fields (indexed by \overline{f}) that article i 's references belong to, $\mathbb{R}_{i\overline{f}}$ denote the set of articles that article i references (indexed by k) in field \overline{f} , and $w(\overline{f}|k)$ denote the weight of article k assigned to field \overline{f} . The weighted number of articles from field \overline{f} that article i cites is $\sum_{\forall k \in \mathbb{R}_{i\overline{f}}} w(\overline{f}|k)$. The share of reference article i cites from field \overline{f} is then

$$s_{i\bar{f}} = \frac{\sum_{\forall k \in \mathbb{R}_{i\bar{f}}} w(\bar{f}|k)}{\sum_{\forall \bar{f} \in \bar{\mathcal{F}}} \sum_{\forall k \in \mathbb{R}_{i\bar{f}}} w(\bar{f}|k)}$$

and the disciplinary diversity of references is

$$\text{Disciplinary Diversity of References}_i = (1 - \sum_{\bar{f} \in \bar{\mathcal{F}}} s_{i\bar{f}}^2) \times 1000.$$



Appendix Figure C.1: Process used to aggregate MeSH terms in order to construct fields.
 Note: Example shows MeSH terms Asthma and Neck.

Appendix D: Extracting and Processing Text

We first extract the title and abstract from each article. We then index all words, word pairs and word triplets (1-, 2-, and 3-grams). We then process the n-grams contained in these titles and abstracts by performing the following operations:

1. Convert all text to lower-case.
2. Eliminate 2- and 3-grams with words that cross the following characters: ,?!;:)(\}\{][--.
3. Eliminate all remaining characters that are not alphanumeric.
4. Eliminate all n-grams that contain words appearing in the stopword list provided by the NLM at this address: https://mbr.nlm.nih.gov/Download/2009/WordCounts/wrd_stop
5. Eliminate all n-grams that contain the following character sequences: web, www, http, pubmed, MEDLINE.
6. Eliminate all n-grams that contain more than two adjacent numbers.
7. Eliminate all n-grams that have a length of less than three characters.
8. Keep all 1-grams with character length 3-29, 2-grams with character length 7-59, and 3-grams with character length 11-89.
9. Stem each word from each n-gram using the module *Lingua::Stem* from the Comprehensive Perl Archive Network (CPAN).

Index all the processed n-grams from each title and abstract into 746 tab-delimited text files corresponding to the 746 MEDLINE XML files.

Appendix E: Additional Results

Appendix Table E.1: P-values of One-sided T-tests of Experience Coefficients Compared to Experience Peak

Dependent Variable	Exper. Group	0	1-5	6-10	11-15	16-20	21-25	26-30	31-40
<i>Panel A: First Author without Author-Position FE</i>									
Number of Citations		0.000	0.076	Peak	0.001	0.002	0.002	0.040	0.012
Disciplinary Diversity of Citations		0.000	Peak	0.000	0.000	0.000	0.000	0.000	0.000
<i>Panel B: First Author with Author-Position FE</i>									
Number of Citations		Peak	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Disciplinary Diversity of Citations		Peak	0.000	0.000	0.000	0.000	0.000	0.000	0.000
<i>Panel C: Last Author without Author-Position FE</i>									
Number of Citations			0.000	0.000	Peak	0.000	0.000	0.000	0.000
Disciplinary Diversity of Citations			0.000	0.000	0.001	0.005	0.000	Peak	0.000
<i>Panel D: Last Author with Author-Position FE</i>									
Number of Citations			Peak	0.000	0.000	0.000	0.000	0.000	0.000
Disciplinary Diversity of Citations			Peak	0.000	0.000	0.000	0.000	0.000	0.000

Note: Each row shows the peak of the variable indicated for each specification and the P-value of a one-sided T-test that each group differs from the peak. Panels A and C are based on the regression results from Table 2 column (2), Panels B and D are based on the regression results from Table 2 column (4).

Appendix Table E.2: Experience and Article Quality (Self-Citations Excluded)

Least Square				
Dependent Variable:	Number of Citations		Disciplinary Diversity of Citations	
Mean (Std. Dev.)	18.59 (60.47)		939.47 (41.96)	
	(1)	(2)	(3)	(4)
1st 0	0.078 (0.843)	11.509*** (1.301)	1.312*** (0.503)	4.410*** (0.871)
1st 1-5	1.905** (0.902)	9.666*** (1.267)	2.864*** (0.650)	3.579*** (0.796)
1st 6-10	2.427** (0.993)	5.624*** (0.836)	2.220*** (0.555)	2.326*** (0.522)
1st 11-15	1.380** (0.686)	3.990*** (0.679)	1.709*** (0.427)	1.797*** (0.396)
1st 16-20	0.701 (0.493)	3.410*** (0.788)	0.990*** (0.226)	1.268*** (0.295)
1st 21-25	-0.186 (0.202)	1.771*** (0.508)	0.424** (0.168)	0.497*** (0.190)
1st 26-30	0.252 (0.174)	2.174*** (0.732)	0.708*** (0.171)	0.719** (0.346)
Last 1-5	1.457*** (0.340)	9.076*** (1.295)	-1.869*** (0.277)	3.263*** (0.504)
Last 6-10	2.276*** (0.353)	6.947*** (0.927)	-0.709*** (0.139)	1.971*** (0.312)
Last 11-15	2.726*** (0.335)	4.775*** (0.691)	0.696*** (0.123)	1.604*** (0.220)
Last 16-20	2.167*** (0.242)	2.776*** (0.445)	0.799*** (0.105)	0.938*** (0.180)
Last 21-25	1.099*** (0.170)	1.169*** (0.286)	0.413*** (0.081)	0.475*** (0.112)
Last 26-30	1.091*** (0.117)	1.173*** (0.284)	0.911*** (0.076)	0.473*** (0.143)
Number of Author	0.662*** (0.047)	2.056*** (0.127)	0.440*** (0.046)	0.820*** (0.029)
Number of Authors' Cities	0.883*** (0.202)	0.274 (0.171)	0.102 (0.209)	-0.076* (0.040)
Size of Team's Network	0.017*** (0.002)	-0.003*** (0.001)	0.017*** (0.001)	0.007*** (0.001)
Author-Position FE		Yes		Yes
Observations	4345648	3226694	3865922	2851524
R ²	0.001	0.400	0.007	0.482

Note: Standard errors in parentheses are clustered by field. The dependent variables are residuals from first stage regressions that control for year and field fixed effects. The omitted groups are 31-40 years of experience for both 1st and last authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table E.3: Experience and Article Quality (Controlling for Middle Author Experience)

Dependent Variable:	Number of Citations		Disciplinary Diversity of Citations	
Mean (Std. Dev.)	29.49 (77.27)		946.12 (33.34)	
	(1)	(2)	(3)	(4)
1st 0	-0.071 (1.391)	18.380*** (1.516)	1.420*** (0.394)	4.861*** (0.581)
1st 1-5	2.725* (1.472)	14.575*** (1.413)	2.490*** (0.530)	3.846*** (0.487)
1st 6-10	3.015** (1.532)	7.706*** (1.059)	1.630*** (0.383)	2.496*** (0.291)
1st 11-15	1.573 (1.007)	5.152*** (0.652)	1.249*** (0.276)	1.849*** (0.204)
1st 16-20	0.548 (0.537)	3.844*** (0.501)	0.798*** (0.143)	1.234*** (0.155)
1st 21-25	-0.546* (0.331)	1.846*** (0.412)	0.311** (0.136)	0.564*** (0.207)
1st 26-30	0.792*** (0.275)	3.764*** (0.703)	0.711*** (0.190)	0.757** (0.322)
Last 1-5	0.791** (0.352)	13.921*** (1.935)	-1.816*** (0.189)	3.558*** (0.501)
Last 6-10	2.059*** (0.384)	9.258*** (1.309)	-0.591*** (0.105)	2.141*** (0.371)
Last 11-15	3.297*** (0.386)	6.253*** (0.870)	0.539*** (0.077)	1.629*** (0.211)
Last 16-20	2.931*** (0.312)	3.935*** (0.604)	0.636*** (0.070)	1.043*** (0.168)
Last 21-25	1.479*** (0.215)	1.680*** (0.351)	0.335*** (0.046)	0.451*** (0.084)
Last 26-30	1.908*** (0.146)	2.139*** (0.334)	0.818*** (0.056)	0.738*** (0.108)
Number of Author	1.134*** (0.119)	3.349*** (0.229)	0.457*** (0.034)	0.688*** (0.015)
Size of Team's Network	0.036*** (0.003)	0.006*** (0.001)	0.012*** (0.001)	0.005*** (0.001)
Number of Authors' Cities	1.980*** (0.266)	0.692*** (0.245)	0.152 (0.160)	-0.056** (0.024)
Mean Age of Middle Authors	-0.258*** (0.036)	-0.000 (0.012)	-0.023* (0.014)	-0.008 (0.005)
Author-Position FE		Yes		Yes
Observations	3803052	2744246	3710602	2676320
R ²	0.010	0.484	0.009	0.518

Note: Standard errors in parentheses are clustered by field. The dependent variables are residuals from first stage regressions that control for year and field fixed effects. The omitted groups are 31-40 years of experience for both 1st and last authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table E.4: Experience and Article Quality (Controlling for Author's Publication Count)

Dependent Variable:	Number of Citations		Disciplinary Diversity of Citations	
Mean	29.49		946.12	
(Std. Dev.)	(77.27)		(33.34)	
	(1)	(2)	(3)	(4)
1st 0	6.968*** (0.715)	19.920*** (1.934)	2.656*** (0.439)	3.330*** (0.544)
1st 1-5	8.918*** (0.883)	16.373*** (1.906)	3.917*** (0.548)	2.543*** (0.451)
1st 6-10	5.643*** (1.466)	9.100*** (1.332)	2.679*** (0.432)	1.748*** (0.241)
1st 11-15	3.012*** (1.018)	6.453*** (1.015)	1.835*** (0.347)	1.302*** (0.166)
1st 16-20	1.272* (0.673)	5.063*** (1.095)	1.030*** (0.189)	0.856*** (0.149)
1st 21-25	-0.151 (0.338)	2.454*** (0.585)	0.414*** (0.136)	0.181 (0.197)
1st 26-30	3.579*** (0.471)	4.393*** (0.886)	0.732*** (0.257)	0.537* (0.296)
Last 1-5	2.930*** (0.482)	13.804*** (1.858)	-3.150*** (0.311)	2.128*** (0.460)
Last 6-10	3.002*** (0.455)	9.768*** (1.351)	-1.000*** (0.167)	1.411*** (0.311)
Last 11-15	4.414*** (0.520)	6.706*** (1.079)	0.562*** (0.107)	1.050*** (0.190)
Last 16-20	3.508*** (0.383)	3.990*** (0.680)	0.684*** (0.102)	0.672*** (0.147)
Last 21-25	1.707*** (0.256)	1.739*** (0.379)	0.319*** (0.061)	0.262*** (0.080)
Last 26-30	4.134*** (0.394)	2.054*** (0.342)	1.079*** (0.125)	0.407*** (0.109)
Number of Author	1.184*** (0.116)	3.410*** (0.216)	0.695*** (0.128)	0.719*** (0.029)
Size of Team's Network	0.023*** (0.003)	0.007*** (0.001)	0.013*** (0.001)	0.006*** (0.001)
Number of Authors' City	1.910*** (0.323)	0.633** (0.273)	-0.095 (0.286)	-0.101** (0.041)
First Author Publication Count	0.182*** (0.018)	0.043*** (0.009)	-0.003 (0.005)	0.001 (0.002)
Last Author Publication Count	0.088*** (0.011)	-0.009* (0.005)	0.012*** (0.003)	-0.002** (0.001)
Author Position FE		Yes		Yes
Observations	4,372,875	3,248,324	4,259,127	3,163,030
R ²	0.012	0.434	0.015	0.561

Note: Standard errors in parentheses are clustered by field. The dependent variables are residuals from first stage regressions that control for year and field fixed effects. The omitted groups are 31-40 years of experience for both 1st and last authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table E.5: Experience and Article Quality (Long Career Authors)

Dependent Variable:	Number of Citations		Disciplinary Diversity of Citations	
Mean (Std. Dev.)	32.75 (85.35)		947.81 (30.77)	
	(1)	(2)	(3)	(4)
1st 0	0.466 (1.513)	19.378*** (2.357)	1.911*** (0.430)	5.448*** (0.602)
1st 1-5	3.903** (1.569)	15.388*** (2.345)	3.089*** (0.546)	4.311*** (0.521)
1st 6-10	4.720*** (1.720)	8.827*** (1.631)	2.164*** (0.408)	2.679*** (0.292)
1st 11-15	2.369** (1.178)	6.220*** (1.244)	1.568*** (0.277)	1.943*** (0.197)
1st 16-20	1.186 (0.823)	5.263*** (1.345)	0.958*** (0.141)	1.360*** (0.173)
1st 21-25	-0.494 (0.403)	2.454*** (0.643)	0.301** (0.125)	0.474** (0.212)
1st 26-30	0.498 (0.314)	3.488*** (1.062)	0.773*** (0.222)	0.927*** (0.347)
Last 1-5	11.218*** (0.852)	13.736*** (2.124)	2.245*** (0.393)	3.794*** (0.393)
Last 6-10	11.511*** (0.983)	10.366*** (1.477)	2.530*** (0.174)	2.817*** (0.326)
Last 11-15	8.402*** (0.567)	7.420*** (1.015)	1.901*** (0.095)	1.837*** (0.207)
Last 16-20	4.170*** (0.320)	4.461*** (0.616)	0.972*** (0.067)	1.133*** (0.152)
Last 21-25	1.586*** (0.225)	1.925*** (0.413)	0.325*** (0.051)	0.519*** (0.085)
Last 26-30	2.013*** (0.173)	2.452*** (0.378)	0.846*** (0.065)	0.871*** (0.115)
Number of Author	1.235*** (0.117)	3.616*** (0.236)	0.457*** (0.040)	0.734*** (0.018)
Number of Authors' Cities	1.991*** (0.327)	0.759*** (0.273)	0.152 (0.146)	-0.091*** (0.027)
Size of Team's Network	0.041*** (0.004)	0.008*** (0.002)	0.011*** (0.001)	0.005*** (0.001)
Author-Position FE		Yes		Yes
Observations	3089901	2427015	3021991	2371107
R ²	0.009	0.417	0.008	0.496

Note: The sample is restricted to people articles on which the last author has or will accumulate at least 20 years of publication experience by 2009. No restrictions are made on first author career length. Standard errors in parentheses are clustered by field. The dependent variables are residuals from first stage regressions that control for year and field fixed effects. The omitted groups are 31-40 years of experience for both 1st and last authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table E.6: Test for the Significance of Differences of Fields from Overall Peaks.

Variable		Position	Fields with Discrepancies	Fields with Significant Discrepancies	Share of Fields with Discrepancies that are Significant
Number of Citations	Without Fixed Effect	First	143	47	32.87%
		Last	126	31	24.60%
	With Fixed Effect	First	14	0	0.00%
		Last	36	5	13.89%
Disciplinary Diversity of Citations	Without Fixed Effect	First	86	25	29.07%
		Last	176	66	37.50%
	With Fixed Effect	First	54	10	18.52%
		Last	54	2	3.70%

Note: The table focuses on fields where the peak for the field differ from the peak for the full sample for that specification. Fields with Discrepancies is the count of subfields where the peak age group for the field is different from the full sample results. Fields with Significant Discrepancies is the count of subfields where a one-sided t-test of the peak for the subfield is significantly different from the full sample results at the 5% confidence level. As with Table 3, the full sample comprises 289 fields with at least 57,905 articles (the 99th percentile of the distribution of subfield size).

Appendix Table E.7: Experience and Article Inputs (References) (Self-Citations Excluded)

Dependent Variable:	Mean Age of References		Disciplinary Diversity of References	
Mean	8.46		946.64	
(Std. Dev.)	(4.26)		(39.66)	
	(1)	(2)	(3)	(4)
1st 0	-0.693*** (0.077)	-1.298*** (0.086)	2.432*** (0.496)	2.721*** (0.535)
1st 1-5	-0.910*** (0.097)	-1.092*** (0.089)	4.226*** (0.568)	3.325*** (0.527)
1st 6-10	-0.596*** (0.081)	-0.609*** (0.058)	2.960*** (0.453)	2.454*** (0.420)
1st 11-15	-0.457*** (0.056)	-0.442*** (0.053)	2.484*** (0.324)	1.779*** (0.408)
1st 16-20	-0.307*** (0.026)	-0.326*** (0.035)	1.736*** (0.207)	1.185*** (0.315)
1st 21-25	-0.162*** (0.015)	-0.167*** (0.022)	0.874*** (0.195)	0.486** (0.219)
1st 26-30	-0.228*** (0.024)	-0.270*** (0.040)	1.177*** (0.184)	0.562*** (0.200)
Last 1-5	-0.804*** (0.067)	-1.079*** (0.071)	-3.906*** (0.430)	0.575 (0.573)
Last 6-10	-0.659*** (0.054)	-0.667*** (0.059)	-1.255*** (0.171)	1.244*** (0.268)
Last 11-15	-0.560*** (0.038)	-0.456*** (0.044)	0.369*** (0.093)	1.132*** (0.206)
Last 16-20	-0.394*** (0.027)	-0.283*** (0.032)	0.552*** (0.071)	0.711*** (0.148)
Last 21-25	-0.208*** (0.015)	-0.132*** (0.021)	0.351*** (0.056)	0.378*** (0.086)
Last 26-30	-0.281*** (0.016)	-0.203*** (0.017)	0.598*** (0.075)	0.359*** (0.121)
Number of Author	-0.084*** (0.015)	-0.063*** (0.006)	0.774*** (0.057)	0.586*** (0.037)
Number of Authors' Cities	0.100*** (0.019)	0.074*** (0.007)	-0.677*** (0.223)	-0.159*** (0.034)
Size of Team's Network	-0.003*** (0.000)	-0.000*** (0.000)	0.008*** (0.001)	0.001 (0.001)
Author-Position FE		Yes		Yes
Observations	4414977	3276714	4400544	3269003
R ²	0.021	0.591	0.009	0.538

Note: Standard errors in parentheses are clustered by field. The dependent variables are residuals from first stage regressions that control for year and field fixed effects. The omitted groups are 31-40 years of experience for both 1st and last authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table E.8: Experience and Article Quality (Controlling for Article Inputs)

Dependent Variable	Number of Citations		Disciplinary Diversity of Citations	
Mean	29.49		946.12	
Standard Deviation	(77.27)		(33.34)	
	(1)	(2)	(3)	(4)
First Author Experience				
0	19.146*** (1.933)	20.274*** (1.656)	5.448*** (0.585)	8.083*** (0.501)
1-5	15.536*** (1.894)	16.341*** (1.647)	4.290*** (0.482)	6.402*** (0.433)
6-10	8.872*** (1.317)	9.601*** (1.172)	2.729*** (0.260)	4.122*** (0.217)
11-15	6.317*** (0.983)	7.084*** (0.925)	2.005*** (0.177)	3.089*** (0.164)
16-20	5.043*** (1.053)	5.651*** (1.013)	1.371*** (0.152)	2.119*** (0.142)
21-25	2.532*** (0.561)	2.862*** (0.564)	0.463** (0.192)	0.849*** (0.178)
26-30	3.777*** (0.864)	3.959*** (0.862)	0.925*** (0.308)	1.301*** (0.297)
Last Author Experience				
1-5	14.149*** (1.883)	14.535*** (1.584)	3.709*** (0.489)	5.829*** (0.419)
6-10	9.803*** (1.333)	10.029*** (1.111)	2.336*** (0.342)	3.595*** (0.289)
11-15	6.727*** (1.026)	6.994*** (0.877)	1.718*** (0.205)	2.629*** (0.173)
16-20	4.012*** (0.655)	4.269*** (0.555)	1.102*** (0.158)	1.716*** (0.134)
21-25	1.781*** (0.385)	1.909*** (0.333)	0.498*** (0.089)	0.806*** (0.078)
26-30	2.255*** (0.341)	2.374*** (0.343)	0.797*** (0.112)	1.153*** (0.103)
Number of Authors	3.389*** (0.220)	2.964*** (0.237)	0.744*** (0.018)	0.527*** (0.022)
Number of Authors' Cities	0.678*** (0.263)	0.666** (0.269)	-0.080*** (0.027)	-0.066** (0.033)
Size of Team's Network	0.007*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.004*** (0.001)
Number of References		0.363*** (0.013)		0.109*** (0.004)
Mean Age of References		-0.199*** (0.029)		-0.085*** (0.012)
Disciplinary Diversity of References		0.014 (0.009)		0.132*** (0.011)
Number of New Concepts Used		6.309*** (0.183)		0.301*** (0.048)
Number of Old Concepts Used		0.074 (0.053)		0.136*** (0.016)
Mean Age of Concepts Used		-0.038*** (0.010)		-0.015*** (0.004)
Author-Position FE	Yes	Yes	Yes	Yes
Observations	3,248,324	3,248,324	3,163,030	3,163,030
R ²	0.428	0.433	0.514	0.526

Note: Standard errors in parentheses are clustered by field. The dependent variables are residuals from first stage regressions that control for year and field fixed effects. The omitted groups are 31-40 years of experience for both 1st and last authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table E.9: Experience and Quantity (not Prorated for Team Size)

Dependent Variable	All Publication		Publication as First Author		Publication as Middle Author		Publication as Last Author	
Mean	4.897		1.121		2.496		1.068	
(Std. Dev.)	(6.938)		(1.940)		(4.007)		(3.121)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1-5	-2.624*** (0.036)	-0.431*** (0.035)	0.258*** (0.006)	0.860*** (0.008)	-0.895*** (0.019)	-0.175*** (0.019)	-1.725*** (0.018)	-1.022*** (0.018)
6-10	-1.990*** (0.035)	-0.000 (0.035)	0.250*** (0.006)	0.739*** (0.008)	-0.603*** (0.020)	-0.031 (0.019)	-1.441*** (0.018)	-0.688*** (0.018)
11-15	-1.312*** (0.035)	0.561*** (0.035)	0.152*** (0.006)	0.576*** (0.008)	-0.341*** (0.020)	0.137*** (0.019)	-0.977*** (0.018)	-0.183*** (0.018)
16-20	-0.748*** (0.035)	0.996*** (0.035)	0.076*** (0.007)	0.444*** (0.008)	-0.159*** (0.020)	0.253*** (0.019)	-0.558*** (0.018)	0.237*** (0.019)
21-25	-0.299*** (0.035)	1.227*** (0.035)	0.031*** (0.006)	0.332*** (0.008)	-0.040** (0.020)	0.300*** (0.019)	-0.218*** (0.018)	0.518*** (0.019)
26-30	0.016 (0.033)	1.198*** (0.034)	0.003 (0.006)	0.227*** (0.008)	0.028 (0.019)	0.279*** (0.019)	0.031* (0.017)	0.623*** (0.019)
31-35	0.123*** (0.028)	0.815*** (0.030)	-0.010* (0.006)	0.119*** (0.007)	0.041** (0.017)	0.185*** (0.017)	0.119*** (0.014)	0.470*** (0.016)
Author Fixed Effect	Yes		Yes		Yes		Yes	
Observations	8123499	7247263	8123499	7247263	8123499	7247263	8123499	7247263
R ²	0.018	0.586	0.002	0.507	0.007	0.532	0.038	0.511

Note: Standard errors in parentheses are clustered by individual. The dependent variables are residuals from first stage regressions that control for year and field fixed effects, where the field assigned is the field with the highest weight across an author's first publication. This table covers individuals who published at least one research article during the sample period and tracks their career until age 40 or their last publication. Years between the first and last publication are included with 0 publications. The omitted group are people with 36-40 years of experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table E.10: Experience and Quantity (Prorated for Team Size)

Dependent Variable	All Publication		Publication as First Author		Publication as Middle Author		Publication as Last Author	
Mean (Std. Dev.)	1.291 (2.003)		0.328 (0.612)		0.436 (0.653)		0.316 (0.872)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1-5	-0.754*** (0.009)	-0.040*** (0.010)	0.054*** (0.002)	0.236*** (0.003)	-0.120*** (0.003)	0.025*** (0.003)	-0.426*** (0.004)	-0.206*** (0.005)
6-10	-0.565*** (0.009)	0.123*** (0.010)	0.050*** (0.002)	0.200*** (0.003)	-0.074*** (0.003)	0.048*** (0.003)	-0.345*** (0.004)	-0.106*** (0.005)
11-15	-0.372*** (0.009)	0.298*** (0.010)	0.024*** (0.002)	0.155*** (0.003)	-0.034*** (0.003)	0.073*** (0.003)	-0.215*** (0.004)	0.039*** (0.005)
16-20	-0.216*** (0.009)	0.415*** (0.010)	0.005*** (0.002)	0.120*** (0.003)	-0.010*** (0.003)	0.084*** (0.003)	-0.104*** (0.004)	0.149*** (0.005)
21-25	-0.094*** (0.009)	0.461*** (0.010)	-0.004* (0.002)	0.092*** (0.003)	0.004 (0.003)	0.082*** (0.003)	-0.022*** (0.004)	0.210*** (0.005)
26-30	-0.011 (0.008)	0.419*** (0.010)	-0.006*** (0.002)	0.065*** (0.003)	0.012*** (0.003)	0.069*** (0.003)	0.030*** (0.004)	0.216*** (0.005)
31-35	0.018*** (0.007)	0.268*** (0.008)	-0.006*** (0.002)	0.035*** (0.002)	0.012*** (0.002)	0.043*** (0.003)	0.041*** (0.004)	0.149*** (0.004)
Author Fixed Effect	Yes		Yes		Yes		Yes	
Observations	8123499	7247263	8123499	7247263	8123499	7247263	8123499	7247263
R^2	0.018	0.579	0.002	0.484	0.006	0.514	0.035	0.519

Note: Standard errors in parentheses are clustered by individual. The dependent variables are residuals from first stage regressions that control for year and field fixed effects, where the field assigned is the field with the highest weight across an author's first publication. This table covers individuals who published at least one research article during the sample period and tracks their career until age 40 or their last publication. Years between the first and last publication are included with 0 publications. Articles are prorated for by the inverse of the number of authors (i.e. if an article has 5 authors, the publication counts as .2 publications for all authors). The omitted group are people with 36-40 years of experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table E.11: Experience and Quantity (for People who are Ever First or Last Authors)

Sample:	Ever First Author		Ever Last Author	
Mean	6.036		6.654	
Std. Dev.	(8.074)		(8.887)	
	(1)	(2)	(3)	(4)
1st 1-5	-3.332*** (0.050)	-1.048*** (0.050)	-2.067*** (0.045)	0.362*** (0.045)
1st 6-10	-2.475*** (0.050)	-0.251*** (0.050)	-0.985*** (0.046)	1.492*** (0.045)
1st 11-15	-1.556*** (0.050)	0.566*** (0.050)	-0.287*** (0.046)	2.161*** (0.045)
1st 16-20	-0.808*** (0.050)	1.172*** (0.050)	0.160*** (0.045)	2.458*** (0.045)
1st 21-25	-0.253*** (0.050)	1.488*** (0.050)	0.454*** (0.045)	2.464*** (0.045)
1st 26-30	0.125*** (0.048)	1.479*** (0.048)	0.538*** (0.043)	2.133*** (0.043)
1st 31-35	0.218*** (0.040)	1.006*** (0.041)	0.443*** (0.036)	1.386*** (0.037)
Individual FE	Yes		Yes	
Observations	5,028,753	4,675,460	3,985,695	3,871,142
R ²	0.023	0.579	0.011	0.570

Note: Standard errors in parentheses are clustered by individual. The dependent variables are residuals from first stage regressions that control for year and field fixed effects, where the field assigned is the field with the highest weight across an author's first publications. This table covers individuals who published at least one research article as first author (columns (1) and (2)) or last authors (columns (3) and (4)) during the sample period, and tracks their career until age 40 or their last publication. Years between the first and last publication are included with 0 publications. The omitted group are people with 36-40 years of experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table E.12: Experience and Article Quality (Best Papers)

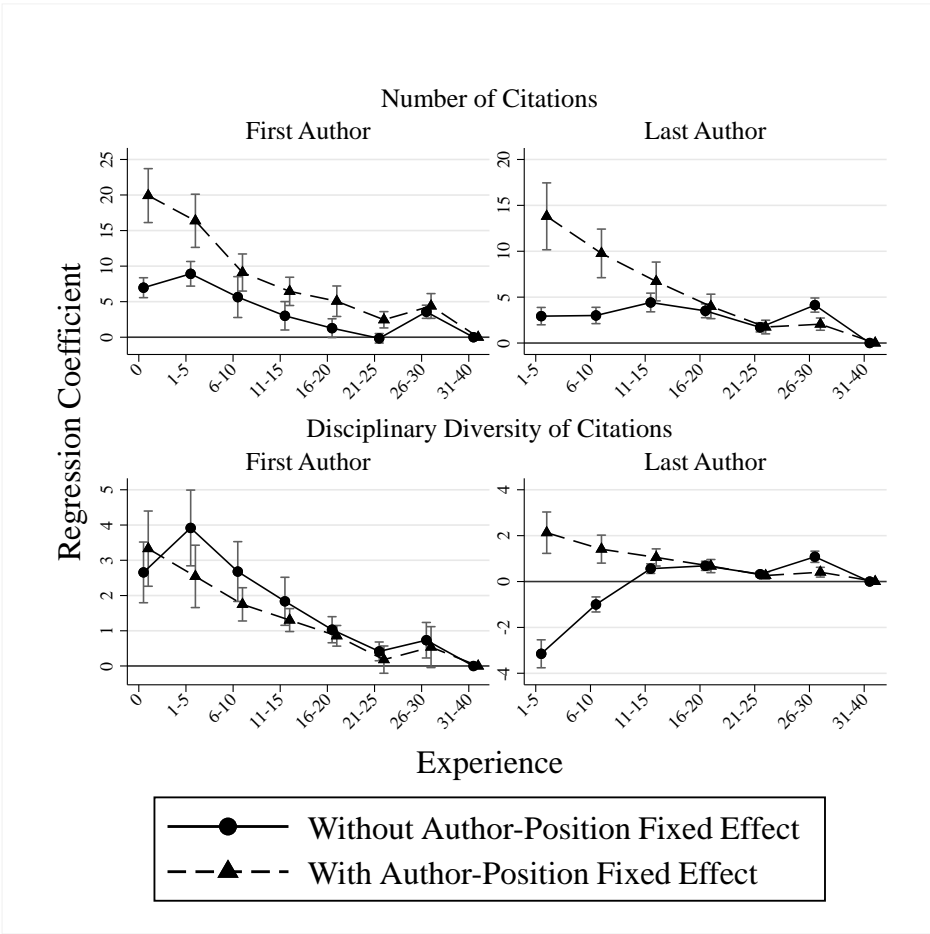
Dependent Variable	Best Paper		2nd Best Paper		3rd Best Paper		Mean of Best “Initial N” Papers	
Mean (Std. Dev.)	62.09 (199.05)		39.24 (76.18)		31.79 (55.06)		35.92 (87.10)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1-5	-19.313*** (0.666)	10.204*** (1.107)	-12.556*** (0.384)	6.231*** (0.428)	-10.546*** (0.308)	4.258*** (0.343)	-26.952*** (0.774)	6.508*** (0.796)
6-10	-10.361*** (0.667)	11.058*** (0.975)	-5.739*** (0.389)	8.831*** (0.434)	-4.836*** (0.310)	7.423*** (0.344)	-11.936*** (0.776)	13.874*** (0.795)
11-15	-4.634*** (0.683)	11.271*** (0.952)	-2.352*** (0.390)	9.187*** (0.434)	-1.886*** (0.312)	8.396*** (0.347)	-6.053*** (0.782)	13.341*** (0.805)
16-20	-0.464 (0.697)	12.028*** (0.922)	0.126 (0.393)	9.685*** (0.439)	0.266 (0.314)	9.050*** (0.353)	-1.933** (0.784)	13.687*** (0.815)
21-25	1.905*** (0.700)	11.589*** (0.900)	1.434*** (0.392)	9.036*** (0.439)	1.377*** (0.312)	8.499*** (0.353)	0.409 (0.786)	13.104*** (0.828)
26-30	2.868*** (0.715)	9.563*** (0.886)	1.846*** (0.385)	7.318*** (0.433)	1.556*** (0.303)	6.803*** (0.347)	0.715 (0.775)	10.409*** (0.814)
31-35	1.423** (0.694)	5.160*** (0.814)	1.001*** (0.359)	4.125*** (0.410)	0.825*** (0.273)	3.884*** (0.318)	0.305 (0.742)	6.562*** (0.789)
Author Fixed Effect	Yes		Yes		Yes		Yes	
Observations	4766781	3671139	3324004	2516110	2520042	1900612	3044821	2288511
R ²	0.002	0.335	0.005	0.507	0.007	0.556	0.015	0.436

Note: Standard errors in parentheses are clustered by individual. The dependent variables are residuals from first stage regressions that control for year and field fixed effects, where the field assigned is the field with the highest weight across an author’s first publication. This table covers individuals who published at least one, two, or three research articles during the sample period, and tracks their career until age 40 or their last publication. In the last pair of columns, we calculate the number of papers the researcher publishes in the first 5 years of the career (the “initial N”) and calculate the mean citations among the most cited papers in each period up to that “initial N.” The omitted group are people with 36-40 years of experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

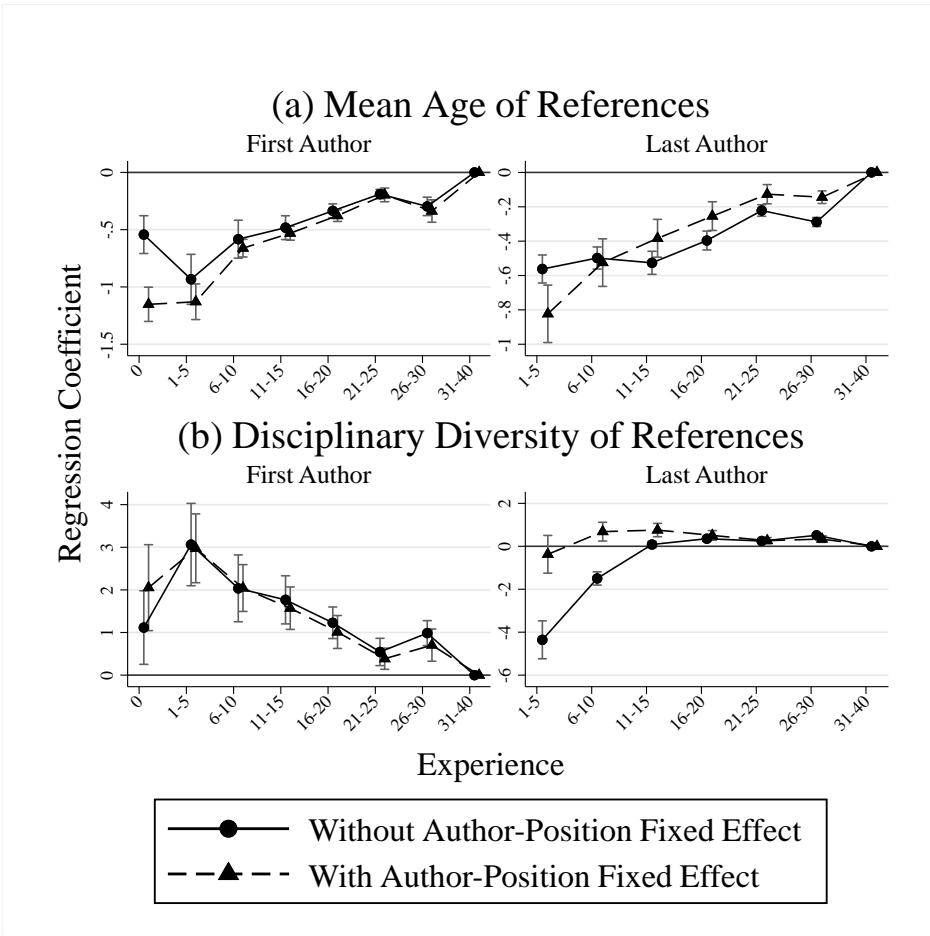
Appendix Table E.13: Experience and Article Quality (Best Papers, Prorated for Team Size)

Dependent Variable	Best Paper		2nd Best Paper		3rd Best Paper		Mean of Best N Paper	
Mean	13.86		8.59		6.89		7.82	
(Std. Dev.)	(98.73)		(19.96)		(13.54)		(26.88)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1-5	-5.468*** (0.207)	3.889*** (0.792)	-3.157*** (0.089)	2.473*** (0.123)	-2.519*** (0.069)	1.842*** (0.092)	-6.226*** (0.226)	10.408*** (0.300)
6-10	-3.104*** (0.184)	4.847*** (0.603)	-1.461*** (0.090)	3.441*** (0.126)	-1.112*** (0.069)	2.836*** (0.094)	-2.618*** (0.224)	11.788*** (0.276)
11-15	-1.614*** (0.194)	5.094*** (0.553)	-0.670*** (0.091)	3.534*** (0.128)	-0.446*** (0.069)	3.035*** (0.095)	-1.080*** (0.226)	10.893*** (0.276)
16-20	-0.525*** (0.191)	5.155*** (0.500)	-0.069 (0.091)	3.526*** (0.127)	0.027 (0.069)	3.046*** (0.096)	-0.249 (0.220)	9.663*** (0.265)
21-25	0.196 (0.198)	4.841*** (0.455)	0.212** (0.090)	3.130*** (0.124)	0.249*** (0.068)	2.725*** (0.094)	0.184 (0.226)	8.142*** (0.265)
26-30	0.801*** (0.257)	4.176*** (0.445)	0.309*** (0.087)	2.442*** (0.118)	0.307*** (0.065)	2.133*** (0.090)	0.194 (0.228)	6.036*** (0.258)
31-35	0.682*** (0.211)	2.517*** (0.327)	0.224** (0.089)	1.444*** (0.116)	0.172*** (0.059)	1.217*** (0.079)	0.197 (0.215)	3.620*** (0.247)
Author Fixed Effect	Yes		Yes		Yes		Yes	
Observations	4766781	3671139	3324004	2516110	2520042	1900612	3044821	2288511
R ²	0.000	0.183	0.005	0.479	0.007	0.549	0.008	0.327

Note: Standard errors in parentheses are clustered by individual. The dependent variables are residuals from first stage regressions that control for year and field fixed effects, where the field assigned is the field with the highest weight across an author's first publication. This table covers individuals who published at least one, two, or three research articles during the sample period, and tracks their career until age 40 or their last publication. In the last pair of columns, we calculate the number of papers the researcher publishes in the first 5 years of the career (the "initial N") and calculate the mean citations among the most cited papers in each period up to that "initial N." Prorated estimates are prorated by the inverse of the number of authors (i.e. if an article has 5 authors, the publication counts as .2 publications for all authors). The omitted group are people with 36-40 years of experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

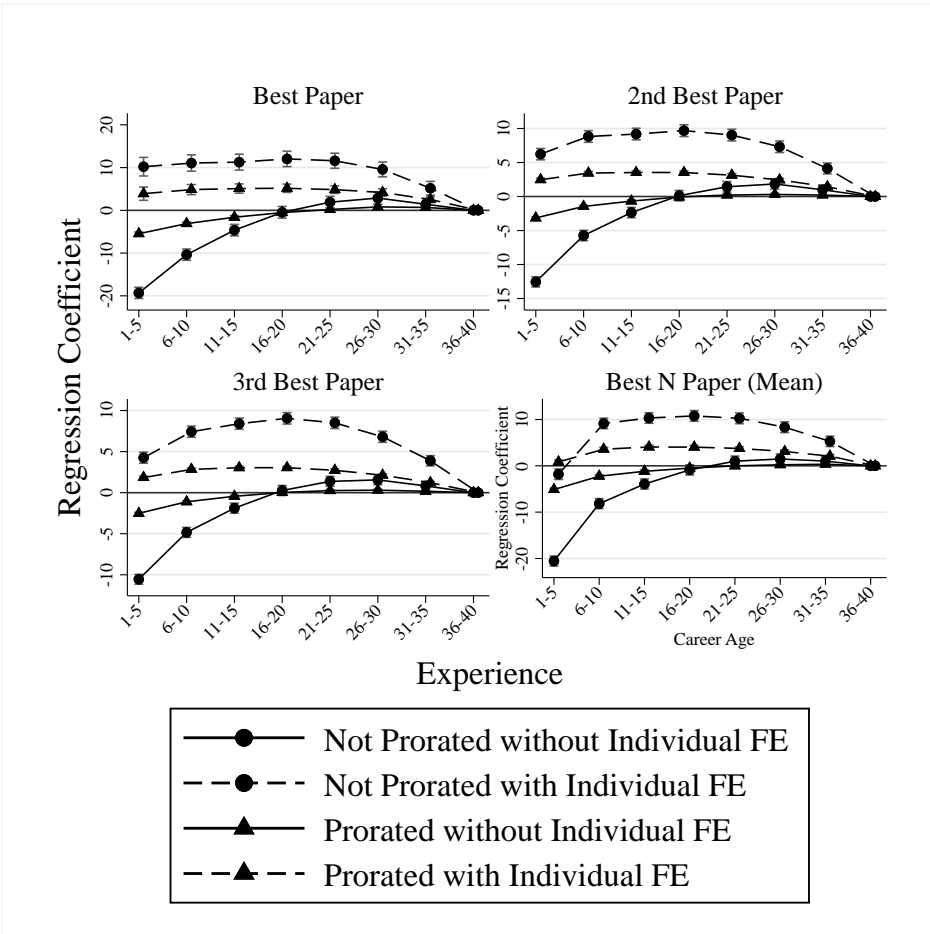


Appendix Figure E.1: Experience and Article Quality (Controlling for Author’s Publication Count)
 Note. Estimates from Appendix Table E.4.



Appendix Figure E.2: Experience and Article Inputs (References)

Note. Estimates from columns (1) to (4) of Table 5



Appendix Figure E.3: Experience and Article Quality (Best Papers)

Note. Estimates from Appendix Tables E.12 and E.13.

Appendix F: Robustness Check for Approach to Address Multicollinearity

Appendix Table F.1: Experience and Article Quality, Alternative Approach to Address Multicollinearity

Dependent Variable	Number of Citations		Disciplinary Diversity of Citations	
	Mean	Standard Deviation	Mean	Standard Deviation
	29.49	(77.27)	946.12	(33.34)
	(1)	(2)	(1)	(2)
1st 0	0.003 (1.335)	19.289*** (1.750)	1.585*** (0.449)	5.526*** (0.438)
1st 1-5	2.660* (1.393)	14.745*** (1.715)	2.561*** (0.562)	3.670*** (0.365)
1st 6-10	2.781* (1.439)	7.202*** (1.137)	1.377*** (0.350)	1.477*** (0.276)
1st 11-15	1.494 (0.955)	4.682*** (0.887)	1.159*** (0.238)	0.801*** (0.241)
1st 16-20	0.758 (0.623)	4.124*** (0.997)	0.848*** (0.105)	0.697*** (0.223)
1st 21-25	-0.247 (0.333)	2.190*** (0.585)	0.366*** (0.110)	0.194 (0.229)
1st 26-30	0.866** (0.248)	3.692*** (0.872)	0.884*** (0.205)	0.878*** (0.259)
Last 1-5	0.905** (0.435)	16.380*** (1.806)	-2.219*** (0.175)	5.311*** (0.487)
Last 6-10	1.546*** (0.404)	10.664*** (1.251)	-1.341*** (0.117)	2.962*** (0.296)
Last 11-15	2.857** (0.400)	7.043*** (0.939)	-0.013 (0.077)	1.911*** (0.163)
Last 16-20	2.570*** (0.299)	4.165*** (0.601)	0.335*** (0.064)	1.171*** (0.122)
Last 21-25	1.427*** (0.222)	1.927*** (0.364)	0.220*** (0.051)	0.569*** (0.079)
Last 26-30	2.286** (0.186)	2.789*** (0.340)	1.062*** (0.073)	1.156*** (0.112)
Number of Author	1.142*** (0.096)	3.474*** (0.207)	0.605*** (0.039)	0.810*** (0.016)
Size of Team's Network	0.028*** (0.003)	-0.001 (0.001)	0.007*** (0.001)	-0.000 (0.001)
Number of Authors' City	1.772*** (0.298)	0.722*** (0.264)	0.067 (0.177)	-0.043 (0.028)
Author-Position FE		Yes		Yes
Observations	4372875	3248324	4259127	3163030
R ²	0.007	0.428	0.008	0.515

Note: Standard errors in parentheses are clustered by field. The dependent variables are residuals from first stage regressions that control for linear time trend and field fixed effects. The omitted groups are 31-40 years of experience for both 1st and last authors.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table F.2: Experience and Article Inputs, Alternative Approach to Address Multicollinearity

Dependent Variable:	Number of New Concepts		Number of Old Concepts		Mean Age of Concepts		Average Age of		Disciplinary Diversity of	
	Used		Used		Used		References		References	
Mean (Std. Dev.)	0.33 (0.97)		16.67 (12.06)		53.78 (14.11)		8.83 (5.84)		948.36 (35.23)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1st 0	0.074*** (0.014)	0.177*** (0.026)	1.441*** (0.335)	2.089*** (0.329)	-1.594*** (0.214)	-2.134*** (0.310)	-0.542*** (0.084)	-1.159*** (0.069)	1.150*** (0.444)	2.108*** (0.510)
1st 1-5	0.098*** (0.015)	0.129*** (0.023)	2.244*** (0.408)	2.164*** (0.274)	-2.148*** (0.252)	-1.751*** (0.251)	-0.928*** (0.111)	-1.110*** (0.072)	3.105*** (0.497)	3.087*** (0.403)
1st 6-10	0.063*** (0.009)	0.059*** (0.013)	1.702*** (0.284)	1.318*** (0.194)	-1.524*** (0.151)	-0.959*** (0.187)	-0.566*** (0.083)	-0.618*** (0.036)	2.068*** (0.402)	2.163*** (0.277)
1st 11-15	0.040*** (0.006)	0.024*** (0.009)	1.448*** (0.182)	1.019*** (0.149)	-1.226*** (0.097)	-0.567*** (0.159)	-0.472*** (0.051)	-0.488*** (0.030)	1.788*** (0.290)	1.658*** (0.252)
1st 16-20	0.019*** (0.004)	0.008 (0.007)	0.947*** (0.077)	0.747*** (0.117)	-0.787*** (0.048)	-0.359*** (0.119)	-0.332*** (0.030)	-0.354*** (0.024)	1.228*** (0.192)	1.060*** (0.195)
1st 21-25	0.006* (0.003)	0.004 (0.005)	0.450*** (0.050)	0.395*** (0.080)	-0.402*** (0.047)	-0.211** (0.083)	-0.188*** (0.021)	-0.188*** (0.031)	0.541*** (0.164)	0.411*** (0.127)
1st 26-30	0.011* (0.006)	0.015** (0.007)	0.548*** (0.098)	0.656*** (0.067)	-0.500*** (0.098)	-0.391*** (0.111)	-0.301*** (0.041)	-0.336*** (0.048)	1.002*** (0.151)	0.735*** (0.193)
Last 1-5	0.095*** (0.012)	0.250*** (0.038)	-0.674*** (0.159)	0.815*** (0.226)	-0.999*** (0.158)	-1.844*** (0.121)	-0.549*** (0.042)	-0.879*** (0.084)	-4.331*** (0.452)	-0.473 (0.458)
Last 6-10	0.077*** (0.008)	0.163*** (0.025)	-0.271*** (0.093)	0.522*** (0.129)	-0.797*** (0.097)	-1.118*** (0.083)	-0.470*** (0.032)	-0.545*** (0.068)	-1.466*** (0.158)	0.628*** (0.228)
Last 11-15	0.067*** (0.007)	0.105*** (0.018)	0.480*** (0.045)	0.471*** (0.091)	-0.816*** (0.042)	-0.784*** (0.071)	-0.505*** (0.033)	-0.388*** (0.054)	0.110 (0.079)	0.749*** (0.162)
Last 16-20	0.045*** (0.006)	0.059*** (0.011)	0.526*** (0.044)	0.319*** (0.065)	-0.611*** (0.026)	-0.482*** (0.041)	-0.384*** (0.027)	-0.255*** (0.041)	0.371*** (0.060)	0.513*** (0.117)
Last 21-25	0.025*** (0.003)	0.031*** (0.006)	0.331*** (0.027)	0.156*** (0.038)	-0.323*** (0.026)	-0.230*** (0.031)	-0.217*** (0.017)	-0.128*** (0.028)	0.256*** (0.047)	0.264*** (0.072)
Last 26-30	0.031*** (0.003)	0.036*** (0.006)	0.571*** (0.038)	0.264*** (0.068)	-0.612*** (0.030)	-0.438*** (0.032)	-0.296*** (0.014)	-0.157*** (0.018)	0.532*** (0.053)	0.342*** (0.079)
Number of Author	0.017*** (0.001)	0.023*** (0.001)	0.626*** (0.042)	0.554*** (0.017)	-0.402*** (0.036)	-0.233*** (0.015)	-0.127*** (0.015)	-0.074*** (0.006)	0.742*** (0.054)	0.556*** (0.037)
Number of Authors' City	-0.015*** (0.004)	-0.007*** (0.002)	-0.586*** (0.104)	-0.109*** (0.016)	0.454*** (0.077)	0.113*** (0.030)	0.120*** (0.019)	0.040*** (0.005)	-0.445** (0.212)	-0.054 (0.049)
Size of Team's Network	0.001*** (0.000)	0.000*** (0.000)	0.011*** (0.001)	0.003*** (0.001)	-0.010*** (0.001)	-0.001*** (0.000)	-0.004*** (0.000)	-0.001*** (0.000)	0.009*** (0.001)	0.002*** (0.001)
Author Position FE		Yes		Yes		Yes		Yes		Yes
Observations	4488900	3328811	4488900	3328811	4423970	3279249	4402841	3263662	4430478	3293024

R^2	0.011	0.470	0.052	0.582	0.020	0.529	0.016	0.533	0.012	0.583
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Note: Standard errors in parentheses are clustered by field. The dependent variables are residuals from first stage regressions that control for linear time trend and field fixed effects. The omitted groups are 31-40 years of experience for both 1st and last authors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table F.3: Experience and Quantity (not Prorated for Team Size), Alternative Approach to Address Multicollinearity

Dependent Variable	All Publication		Publication as First Author		Publication as Middle Author		Publication as Last Author	
Mean	4.897		1.121		2.496		1.068	
(Std. Dev.)	(6.938)		(1.940)		(4.007)		(3.121)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
1-5	-2.706*** (0.036)	-0.487*** (0.036)	0.248*** (0.006)	0.864*** (0.008)	-0.961*** (0.020)	-0.268*** (0.019)	-1.745*** (0.018)	-1.047*** (0.018)
6-10	-2.075*** (0.036)	-0.196*** (0.036)	0.240*** (0.006)	0.723*** (0.009)	-0.667*** (0.020)	-0.217*** (0.019)	-1.459*** (0.018)	-0.739*** (0.018)
11-15	-1.374*** (0.036)	0.412*** (0.036)	0.146*** (0.006)	0.570*** (0.009)	-0.385*** (0.020)	-0.018 (0.019)	-0.988*** (0.018)	-0.225*** (0.019)
16-20	-0.791*** (0.035)	0.903*** (0.036)	0.071*** (0.007)	0.445*** (0.008)	-0.187*** (0.020)	0.139*** (0.019)	-0.563*** (0.018)	0.209*** (0.019)
21-25	-0.329*** (0.035)	1.176*** (0.036)	0.028*** (0.006)	0.336*** (0.008)	-0.057*** (0.020)	0.222*** (0.019)	-0.220*** (0.018)	0.499*** (0.019)
26-30	-0.008 (0.033)	1.174*** (0.035)	-0.001 (0.006)	0.230*** (0.008)	0.019 (0.019)	0.233*** (0.019)	0.029* (0.017)	0.611*** (0.019)
31-35	0.100*** (0.028)	0.795*** (0.030)	-0.016*** (0.006)	0.116*** (0.007)	0.035** (0.017)	0.163*** (0.017)	0.116*** (0.015)	0.461*** (0.016)
Author Fixed Effect	Yes		Yes		Yes		Yes	
Observations	8123499	7247263	8123499	7247263	8123499	7247263	8123499	7247263
R ²	0.018	0.577	0.002	0.505	0.008	0.523	0.039	0.509

Note: Standard errors in parentheses are clustered by individual. The dependent variables are residuals from first stage regressions that control for linear time trend and field fixed effects, where the field assigned is the field with the highest weight across an author's first publication. This table covers individuals who published at least one research article during the sample period, and tracks their career until age 40 or their last publication. Years between the first and last publication are included with 0 publications. Articles are weighted by the inverse of the number of authors (i.e. if an article has 5 authors, the publication counts as .2 publications for all authors). The omitted group are people with 36-40 years of experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table: F.4: Experience and Quantity (Prorated for Team Size), Alternative Approach to Address Multicollinearity

Dependent Variable	All Publication		Publication as First Author		Publication as Middle Author		Publication as Last Author	
Mean	1.291		0.328		0.436		0.316	
Std. Dev	(2.003)		(0.612)		(0.653)		(0.872)	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
1-5	-0.754*** (0.009)	0.011 (0.010)	0.055*** (0.002)	0.247*** (0.003)	-0.132*** (0.003)	0.003 (0.003)	-0.429*** (0.004)	-0.204*** (0.005)
6-10	-0.570*** (0.009)	0.149*** (0.010)	0.050*** (0.002)	0.209*** (0.003)	-0.085*** (0.003)	0.012*** (0.003)	-0.347*** (0.004)	-0.109*** (0.005)
11-15	-0.380*** (0.009)	0.331*** (0.010)	0.024*** (0.002)	0.166*** (0.003)	-0.041*** (0.003)	0.042*** (0.003)	-0.216*** (0.004)	0.039*** (0.005)
16-20	-0.226*** (0.009)	0.453*** (0.010)	0.004** (0.002)	0.131*** (0.003)	-0.014*** (0.003)	0.060*** (0.003)	-0.105*** (0.004)	0.151*** (0.005)
21-25	-0.105*** (0.009)	0.497*** (0.010)	-0.005** (0.002)	0.101*** (0.003)	0.002 (0.003)	0.065*** (0.003)	-0.023*** (0.004)	0.213*** (0.005)
26-30	-0.023*** (0.008)	0.446*** (0.010)	-0.008*** (0.002)	0.071*** (0.003)	0.011*** (0.003)	0.058*** (0.003)	0.029*** (0.004)	0.217*** (0.005)
31-35	0.006 (0.007)	0.278*** (0.008)	-0.009*** (0.002)	0.036*** (0.002)	0.011*** (0.003)	0.038*** (0.003)	0.039*** (0.004)	0.149*** (0.004)
Author Fixed Effect	Yes		Yes		Yes		Yes	
Observations	8123499	7247263	8123499	7247263	8123499	7247263	8123499	7247263
R ²	0.017	0.577	0.002	0.485	0.006	0.506	0.035	0.518

Note: Standard errors in parentheses are clustered by individual. The dependent variables are residuals from first stage regressions that control for linear time trend and field fixed effects, where the field assigned is the field with the highest weight across an author's first publication. This table covers individuals who published at least one research article during the sample period, and tracks their career until age 40 or their last publication. Years between the first and last publication are included with 0 publications. Articles are weighted by the inverse of the number of authors (i.e. if an article has 5 authors, the publication counts as .2 publications for all authors). The omitted group are people with 36-40 years of experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table F.5: Experience and Article Quality (Best Papers), Alternative Approach to Address Multicollinearity

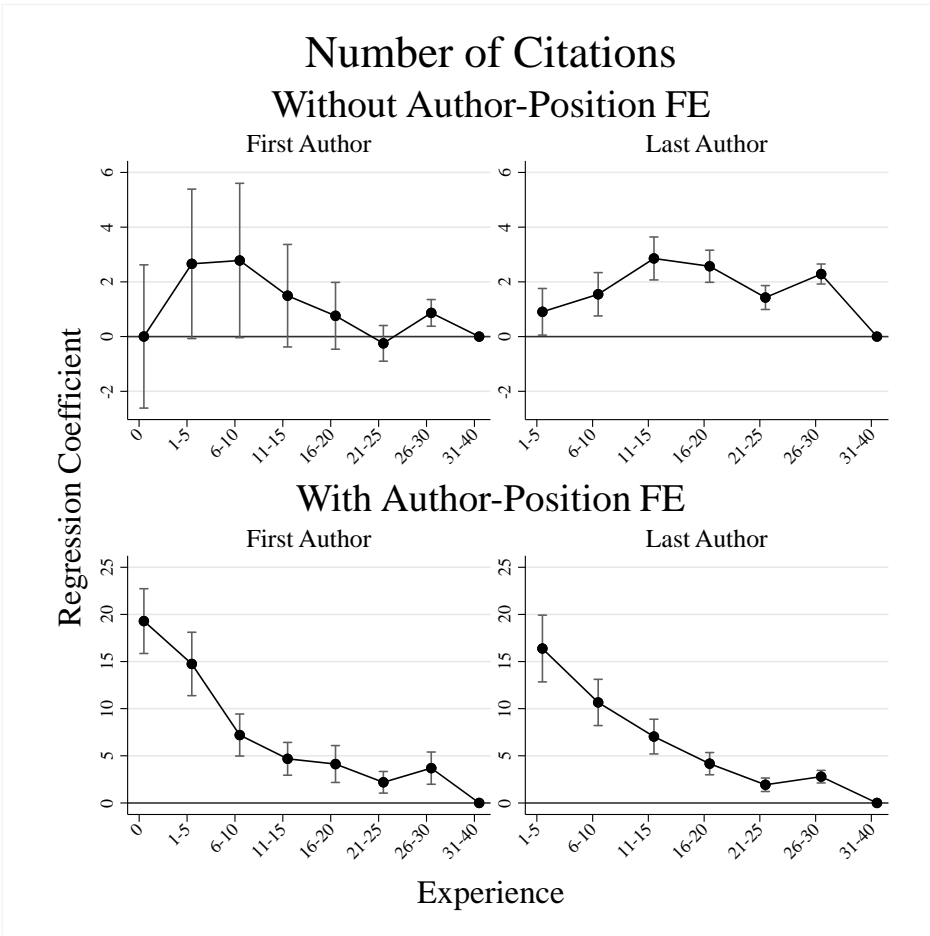
Dependent Variable	Best Paper		2nd Best Paper		3rd Best Paper		Mean of Best N Paper	
Mean (Std. Dev.)	62.09 (199.05)		39.24 (76.18)		31.79 (55.06)		35.92 (87.10)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1-5	-20.556*** (0.669)	6.883*** (1.112)	-12.880*** (0.386)	5.684*** (0.434)	-10.580*** (0.310)	4.673*** (0.349)	-20.718*** (0.554)	-4.365*** (0.581)
6-10	-12.725*** (0.669)	3.168*** (0.981)	-6.884*** (0.391)	5.549*** (0.440)	-5.522*** (0.313)	5.796*** (0.351)	-9.042*** (0.557)	4.001*** (0.580)
11-15	-6.287*** (0.686)	3.157*** (0.957)	-3.100*** (0.393)	5.570*** (0.441)	-2.302*** (0.314)	6.405*** (0.354)	-4.368*** (0.562)	5.203*** (0.587)
16-20	-1.286* (0.700)	5.789*** (0.929)	-0.108 (0.396)	6.961*** (0.446)	0.245 (0.317)	7.617*** (0.360)	-0.751 (0.566)	7.012*** (0.596)
21-25	1.706** (0.704)	7.632*** (0.907)	1.532*** (0.395)	7.434*** (0.447)	1.575*** (0.315)	7.762*** (0.361)	1.642*** (0.566)	8.105*** (0.600)
26-30	2.959*** (0.719)	7.490*** (0.894)	2.110*** (0.388)	6.672*** (0.442)	1.865*** (0.307)	6.650*** (0.356)	2.233*** (0.560)	7.369*** (0.595)
31-35	1.480** (0.697)	4.332*** (0.821)	1.161*** (0.363)	3.987*** (0.418)	1.003*** (0.277)	3.970*** (0.326)	1.442*** (0.535)	4.908*** (0.576)
Author Fixed Effect		Yes		Yes		Yes		Yes
Observations	4766781	3671139	3324004	2516110	2520042	1900612	4032350	3099151
R ²	0.002	0.333	0.005	0.501	0.007	0.547	0.010	0.401

Note: Standard errors in parentheses are clustered by individual. The dependent variables are residuals from first stage regressions that control for linear time trend and field fixed effects, where the field assigned is the field with the highest weight across an author's first publication. This table covers individuals who published at least one research article during the sample period, and tracks their career until age 40 or their last publication. The omitted group are people with 36-40 years of experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Table F.6: Experience and Article Quality (Best Papers, Prorated for Team Size), Alternative Approach to Address Multicollinearity

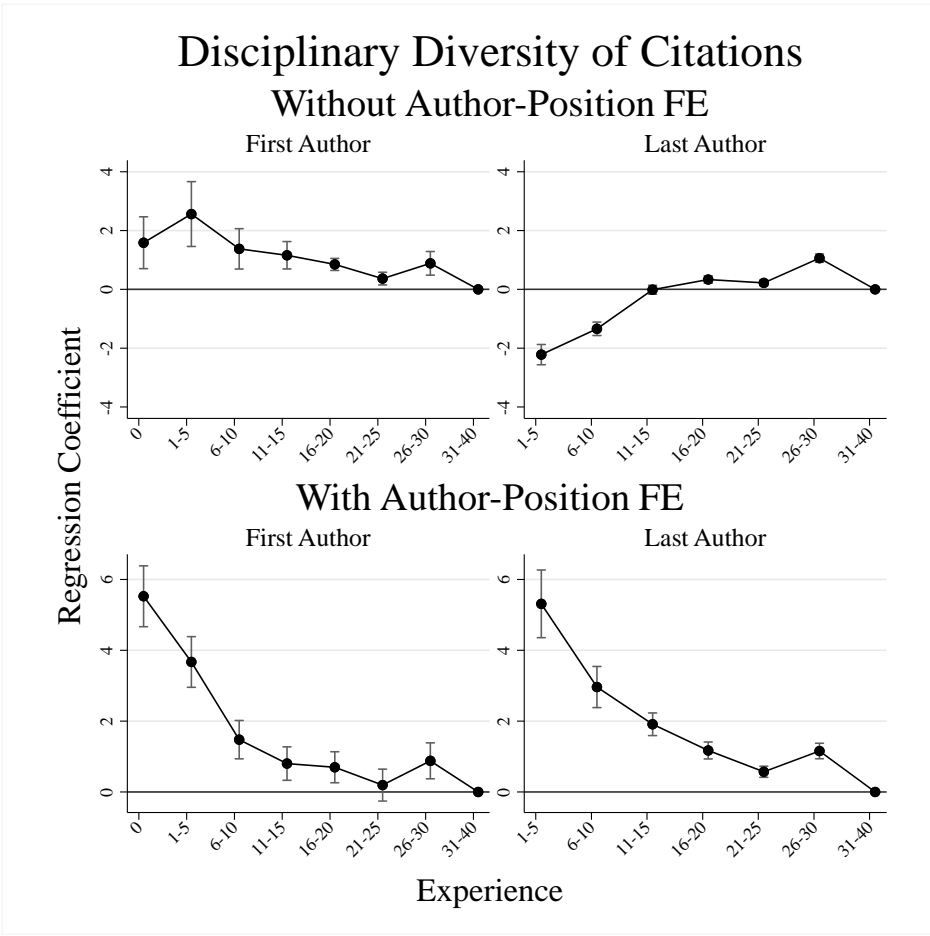
Dependent Variable	Best Paper		2nd Best Paper		3rd Best Paper		Mean of Best N Paper	
Mean (Std. Dev.)	13.86 (98.73)		8.59 (19.96)		6.89 (13.54)		7.82 (26.88)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1-5	-5.615*** (0.207)	3.785*** (0.792)	-3.217*** (0.089)	2.439*** (0.124)	-2.540*** (0.069)	1.908*** (0.093)	-5.115*** (0.146)	0.852*** (0.215)
6-10	-3.374*** (0.185)	4.137*** (0.604)	-1.629*** (0.090)	3.003*** (0.127)	-1.222*** (0.069)	2.594*** (0.095)	-2.238*** (0.145)	3.591*** (0.199)
11-15	-1.821*** (0.194)	4.295*** (0.553)	-0.785*** (0.091)	3.029*** (0.129)	-0.515*** (0.070)	2.728*** (0.095)	-1.166*** (0.147)	4.062*** (0.199)
16-20	-0.635*** (0.191)	4.577*** (0.500)	-0.111 (0.091)	3.146*** (0.128)	0.019 (0.070)	2.821*** (0.097)	-0.499*** (0.140)	4.043*** (0.187)
21-25	0.176 (0.198)	4.573*** (0.455)	0.221** (0.091)	2.924*** (0.125)	0.278*** (0.069)	2.616*** (0.095)	-0.009 (0.146)	3.756*** (0.186)
26-30	0.777*** (0.257)	4.092*** (0.445)	0.333*** (0.088)	2.372*** (0.119)	0.345*** (0.066)	2.112*** (0.091)	0.236 (0.150)	3.099*** (0.181)
31-35	0.635*** (0.212)	2.490*** (0.328)	0.225** (0.089)	1.428*** (0.117)	0.186*** (0.059)	1.226*** (0.080)	0.359*** (0.139)	2.093*** (0.171)
Author Fixed Effect	Yes		Yes		Yes		Yes	
Observations	4766781	3671139	3324004	2516110	2520042	1900612	4032350	3099151
R ²	0.000	0.183	0.005	0.479	0.007	0.549	0.006	0.303

Note: Standard errors in parentheses are clustered by individual. The dependent variables are residuals from first stage regressions that control for year and field fixed effects, where the field assigned is the field with the highest weight across an author's first publication. This table covers individuals who published at least one research article during the sample period and tracks their career until age 40 or their last publication. Prorated estimates are prorated by the inverse of the number of authors (i.e. if an article has 5 authors, the publication counts as .2 publications for all authors). The omitted group are people with 36-40 years of experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



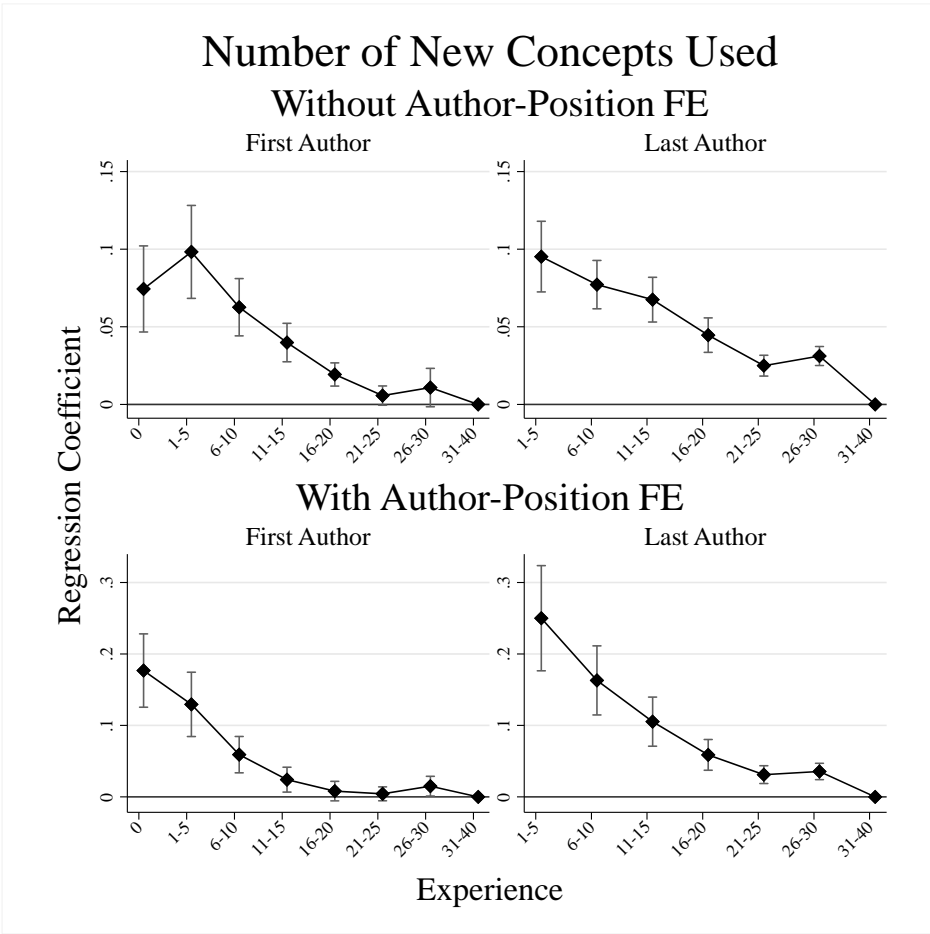
Appendix Figure F.1: Experience and the Number of Citations, Alternative Approach to Address Multicollinearity

Note. The figure plots the coefficients from columns (1) and (2) of Appendix Table F.1.



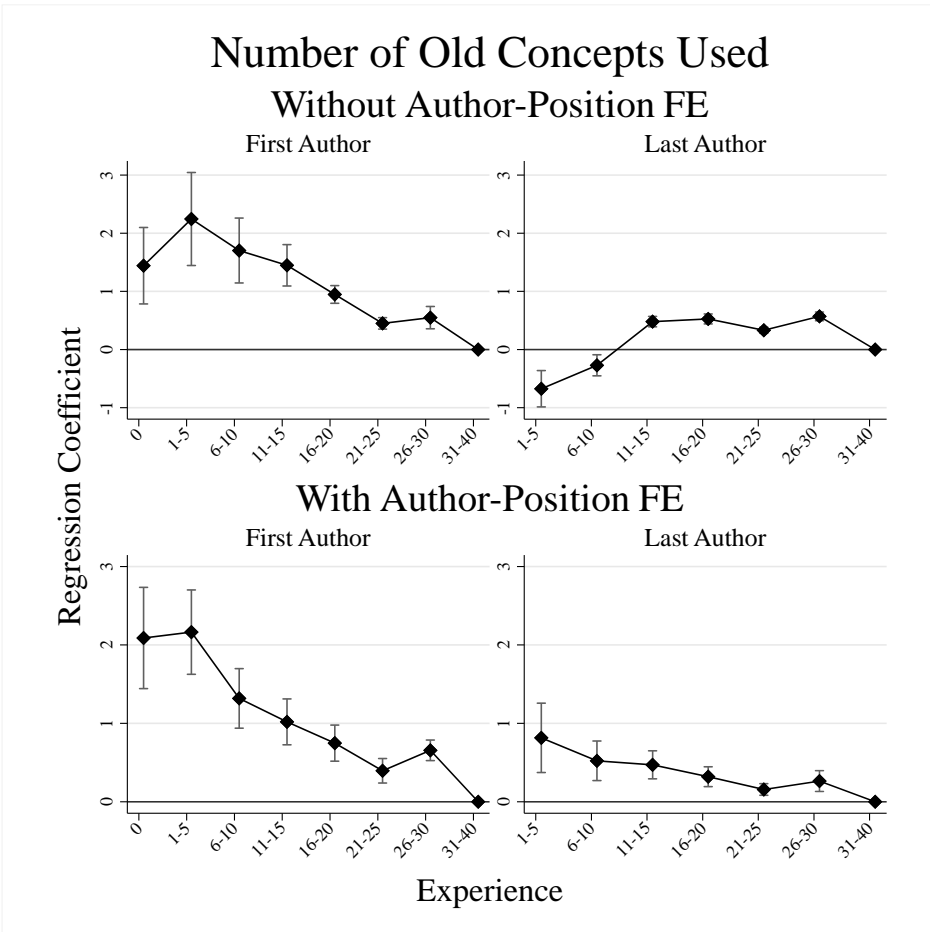
Appendix Figure F.2: Experience and the Disciplinary of Citations, Alternative Approach to Address Multicollinearity

Note. The figure plots the coefficients from columns (3) and (4) of Appendix Table F.1.



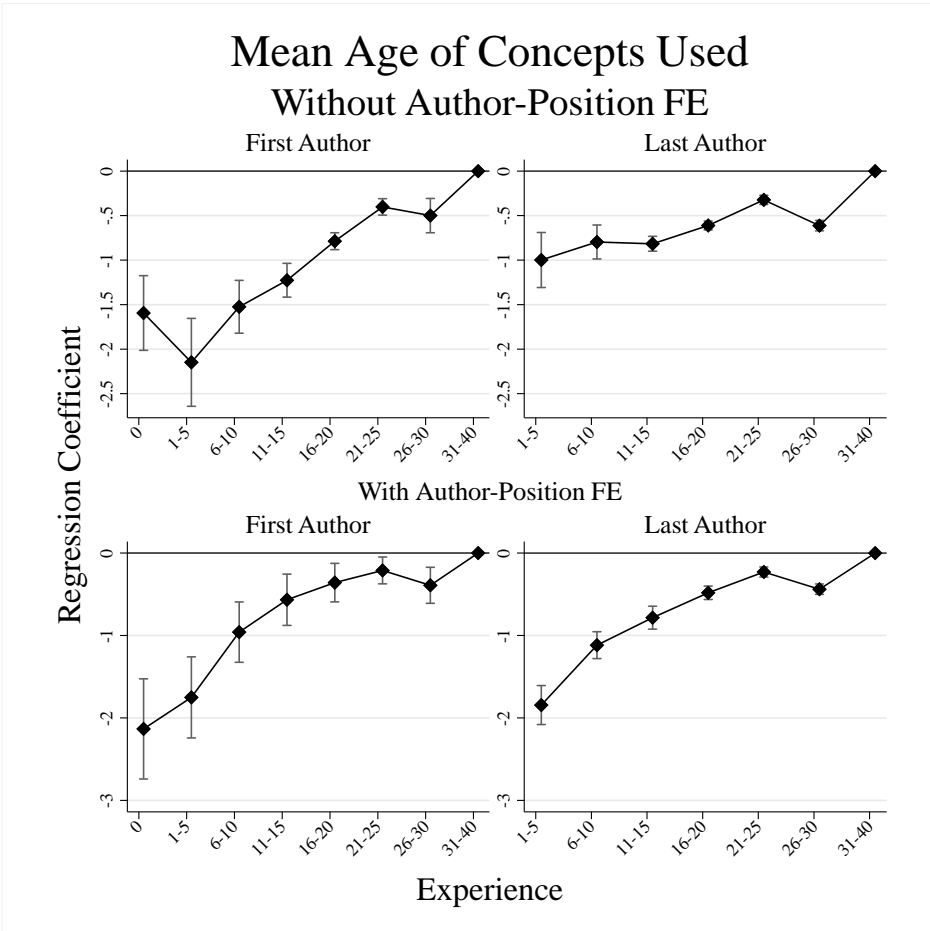
Appendix Figure F.3: Coefficients for Number of New Concepts Used, Alternative Approach to Address Multicollinearity

Note. The figure plots the coefficients from columns (1) and (2) of Appendix Table F.2.



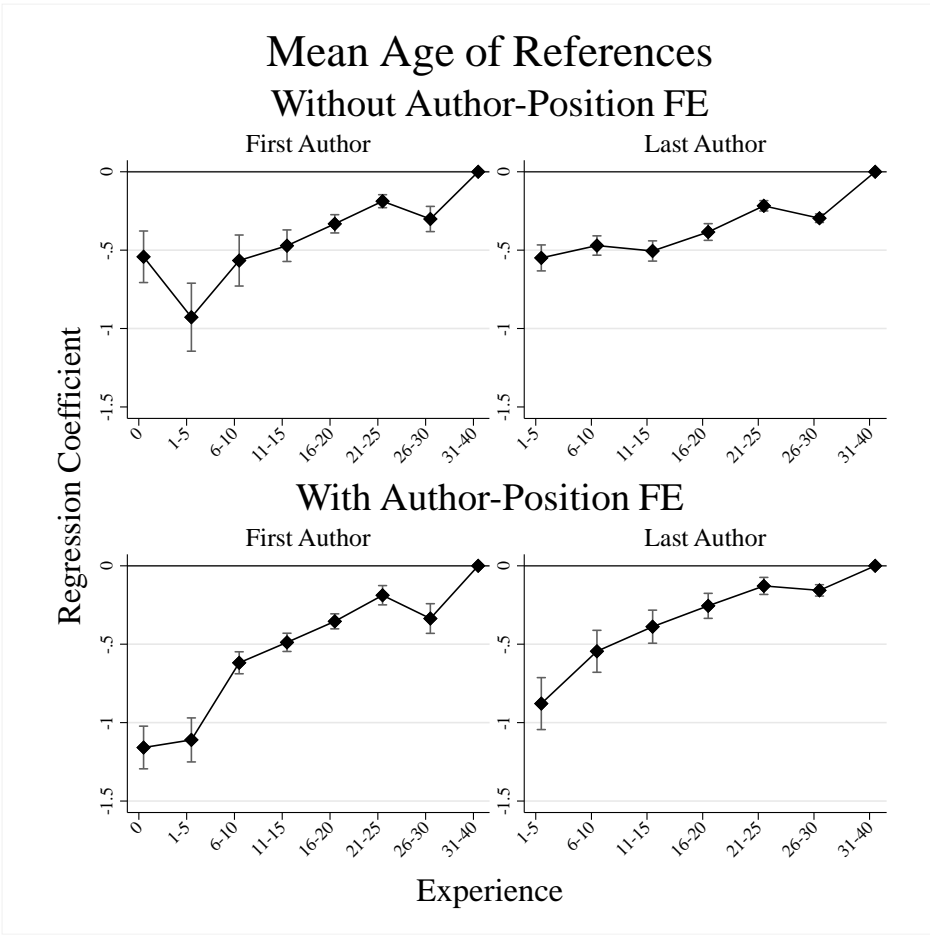
Appendix Figure F.4: Experience and the Number of Old Concepts Used, Alternative Approach to Address Multicollinearity

Note. The figure plots the coefficients from columns (3) and (4) of Appendix Table F.2



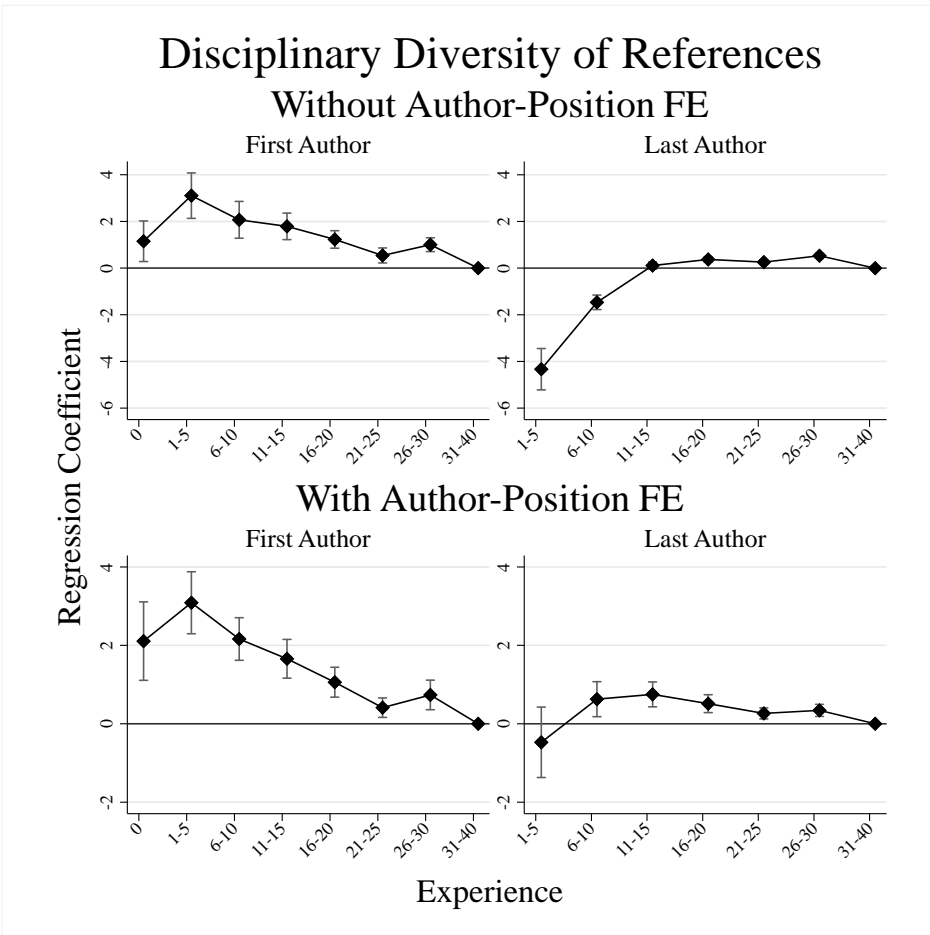
Appendix Figure F.5: Experience and the Mean Age of Concepts Used, Alternative Approach to Address Multicollinearity

Note. The figure plots the coefficients from columns (5) and (6) of Appendix Table F.2.



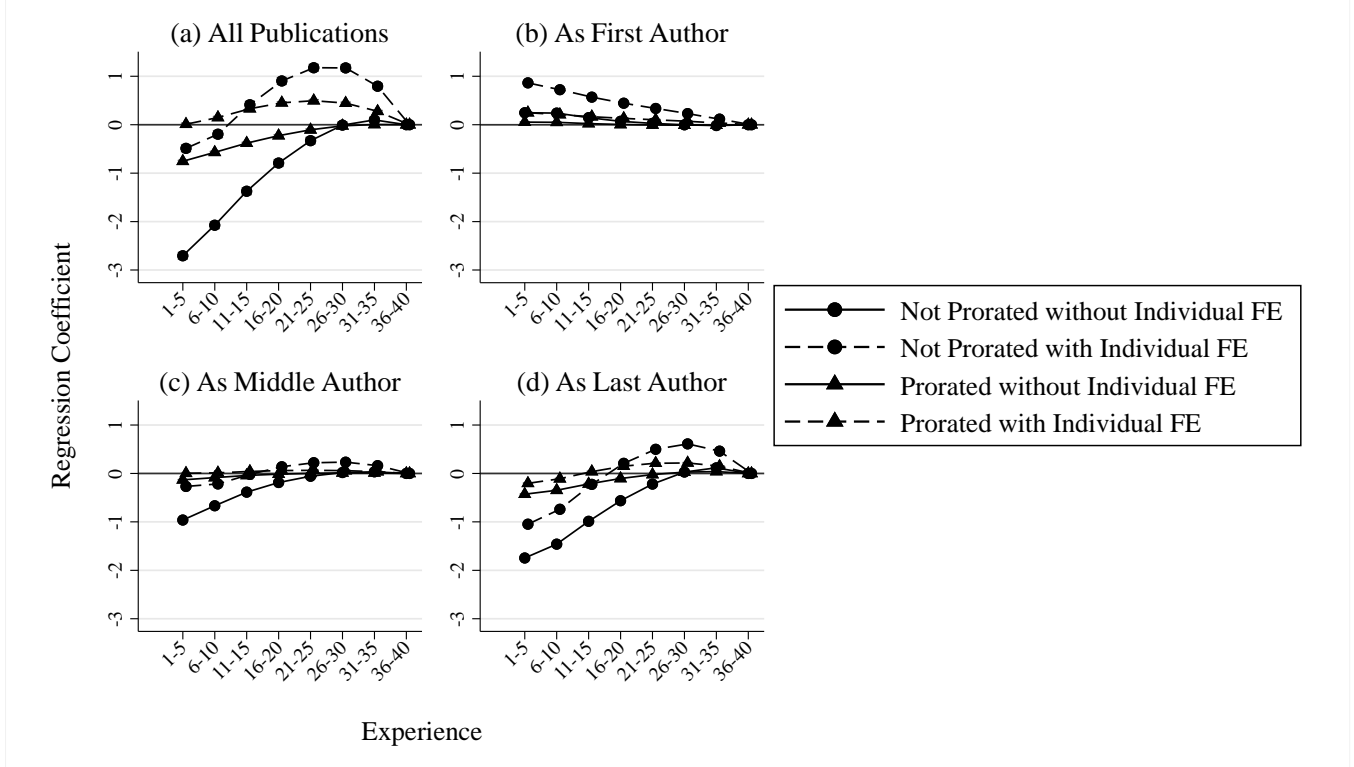
Appendix Figure F.6: Experience and the Mean Age of References, Alternative Approach to Address Multicollinearity

Note. The figure plots the coefficients from columns (7) and (8) of Appendix Table F.2.



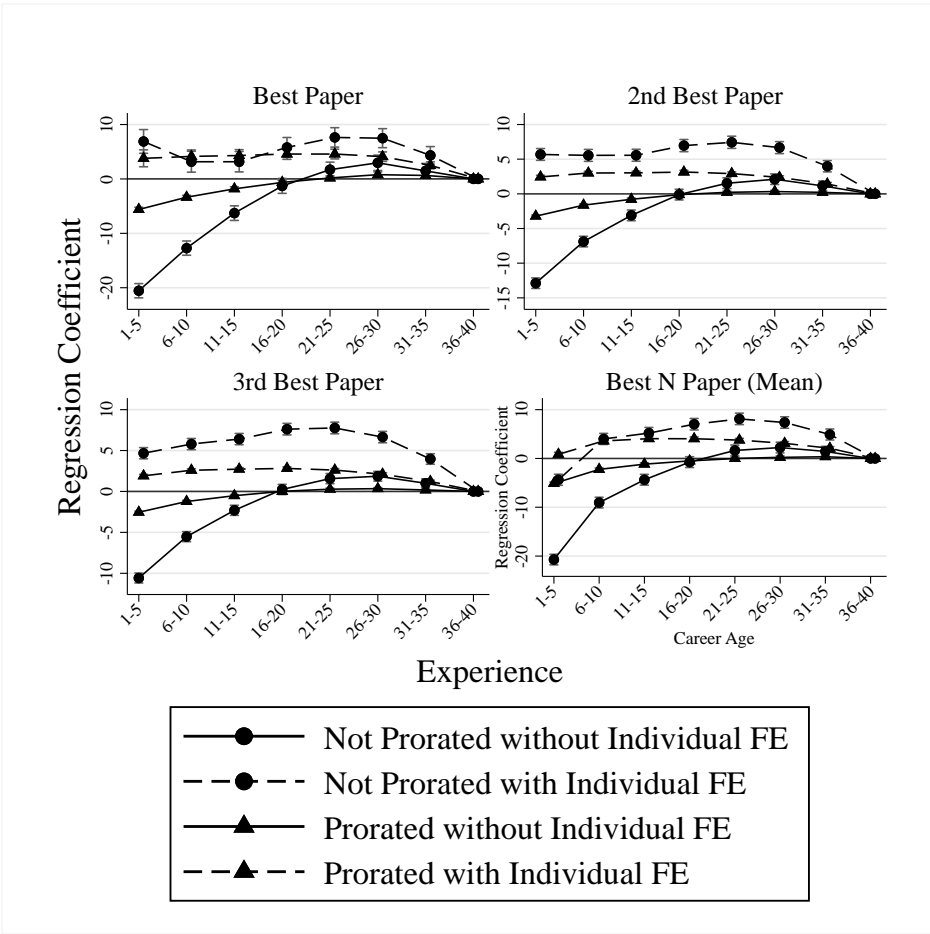
Appendix Figure F.7: Experience and the Disciplinary Diversity of References, Alternative Approach to Address Multicollinearity

Note. The figure plots the coefficients from columns (9) and (10) of Appendix Table F.2.



Appendix Figure F.8: Experience and Quantity, Alternative Approach to Address Multicollinearity

Note. Based on estimates in Appendix Table F.3 and Appendix Table F.4. In the weighted estimates, articles are prorated by the inverse of the number of authors (i.e. if an article has 5 authors, the publication counts as .2 publications for all authors).



Appendix Figure F.9: Experience and Article Quality (Best Paper), Alternative Approach to Address Multicollinearity

Note. Based on estimates in Appendix Table F.5 and F.6.