Check for updates

The MIT Press

RESEARCH ARTICLE

Predicting the impact of *American Economic Review* articles by author characteristics

Tolga Yuret (iD)

Department of Economics, Istanbul Technical University, Istanbul, Turkey

ABSTRACT

Authors who publish in *American Economic Review* (*AER*) have career paths confined to a few prestigious institutions, and they mostly have exceptional past publication performance. In this paper, I show that authors who are educated and work in the top 10 institutions and have better past publication performance receive more citations for their current *AER* publications. Authors who have published in the top economic theory journals receive fewer citations even after controlling for the subfield of their *AER* article. The gender of the authors, years of post-PhD experience, and the location of the affiliated institution do not have any significant effect on the citation performance. An opportunistic editor can exploit the factors that are related to citation performance to substantially improve the citation performance of the journal. Such opportunistic behavior increases the overrepresentation of authors with certain characteristics. For example, an opportunistic editor who uses the predicted citation performance of articles to select a quarter of the articles increases the ratio of authors who works at the top 10 institutions from 30.8% to 52.0%.

1. INTRODUCTION

Most authors who publish in top economics journals have similar academic backgrounds. They are educated in and work at a few elite institutions. There are many reasons why having experience at a prestigious institution may help authors to publish in top journals. Most coauthors are either current or past coworkers, so having productive coworkers may facilitate publishing in top journals (Yuret, 2020). It is also more likely that an author from an elite institution has ties with the journal editor, which is an important factor for an article to be accepted (Colussi, 2018).

Another possible reason why most authors who publish in top economics journals work at a few prestigious institutions is that the journals may improve their impact by accepting more articles from authors who work at elite institutions. Likewise, the journals can improve their impact by accepting articles that are written by authors with other characteristics that indicate a higher citation potential.

In this paper, I have two aims. First, I want to identify the author characteristics that are strongly correlated with the citation performance of *American Economic Review* (*AER*) articles. In particular, I want to see whether the prevalent author characteristics, such as being affiliated with one of the top 10 institutions, are strongly and positively related to the citation performance. Second, I want to see whether a hypothetical opportunistic editor can substantially improve the citation performance of *AER* by using the predicted citation performance. In

this way, I want to see whether the opportunistic behavior increases the concentration of certain author characteristics.

I use a General Linear Model with Logarithm to analyze the effect of author characteristics on the number of citations that *AER* articles receive. The continent of the institution where the authors work and their past publication productivity are examples of the author characteristics that I include. In addition to the author characteristics, I include a few article-related factors that are easily observed and have a clear impact on citation performance, such as the month of the issue that the article is published in.

Next, I perform simulations where two types of opportunistic editors try to optimize the average citation performance of the journal. The opportunistic editor with perfect foresight has perfect information regarding the future citation performance of the articles and selects better performing articles. The opportunistic editor who uses the regression results computes the predicted citation performance of the articles to select the articles that have better expected performance.

In the simulations, I try to see whether large improvements can be attained by opportunistic editors. Moreover, I try to see whether the lack of perfect foresight benefits authors who have characteristics that are already overrepresented. In this way, I provide an additional perspective to interpret the regression results.

2. RELATED LITERATURE

There is a large literature on the factors that affect the number of citations that an article receives. Tahamtan, Afshar, and Ahamdzadeh (2016) present a comprehensive literature review by classifying the research on this topic into three categories. The first category includes paper-related factors, such as characteristics of titles and choice of topic. The second category includes journal-related factors, such as the Journal Impact Factor and the language of the journal. The third category includes author-related factors, such as the author's reputation and gender.

I focus on the effect of author characteristics on citation performance, so this paper mainly belongs to the third category. I only study the citation performance of a single journal so there are no journal-related factors in my analysis. Although my focus is on the author characteristics, I include a few paper-related factors that are easily observed and have a clear effect on the citation performance of an article.

A paper-related factor that I analyze is the month in which an article is issued. Articles that are published in the early months of the year have more time to be recognized, so they receive more citations than articles published in later months of the year. The effect is demonstrated for the citation performance in the first 3 years after publication by a study that analyzes all journals indexed in the Web of Science (Donner, 2018). In contrast, another study that focuses on information science articles claims that the month of the issue is not important because of the improved access in the digital age (Xie, Gong et al., 2019).

The publication year of the articles is another paper-related factor that I analyze. Because I use the same 2-year citation window for all the articles, the number of years in which articles receive citations does not change. However, the number of citations that journals receive in a given year increases over the years (Petersen, Pan et al., 2019). For this reason, the publication year is an important determinant for citation performance, even for a specific 2-year window.

Subdiscipline is another important paper-related factor that affects the number of citations that an article receives. Empirical economics articles typically receive more citations than theoretical economics articles (Johnston, Piatti, & Torgler, 2013). In contrast, empirical papers receive fewer citations than theoretical papers in psychology (Buela-Casal, Zych et al., 2009). There is evidence that subdiscipline categories other than the empirical-theoretical divide are also important for the number of citations that an article receives in economics (Medoff, 2003), and mathematics (Smolinsky & Lercher, 2012).

I include the subdiscipline of an article both as a paper-related factor and as an author-related factor. For the paper-related approach, I judge whether the papers are theoretical or empirical. For the author-related approach, I use a dummy variable that indicates whether authors have ever published in the top three theoretical economics journals.

Trimble and Ceja (2013) find that authors from the United States and Europe receive more citations in astrophysics. Of course, the geography and quality of institutions are related, as most prestigious institutions are in either North America or Europe. Chan, Guillot et al. (2015) find that authors from high-ranking institutions get more citations for their *Econometrica* and *AER* publications, and Amara, Landry, and Halilem (2015) find that authors from high-ranking Canadian business schools receive more citations than authors from low-ranking Canadian business schools. However, authors from top institutions do not necessarily receive more citations in all academic fields. For example, Haslam, Ban et al. (2008) conclude that psychologists from high-ranking institutions do not get more citations than psychologists from low-ranking institutions.

It is natural to expect that the past citation performance of authors is a good predictor of their current citation performance. A group of studies support this relation for economists (Medoff, 2003), physicists (Wang, Fan et al. 2019), and various fields of science (Onodera & Yoshikane, 2015). It is less clear why authors who published many papers in the past receive more citations. Card and DellaVigna (2020) find that more productive economists receive more citations. Lindahl (2018) claims that productive mathematicians also receive more citations, but the effect is clearer for those who publish at top journals. Hurley, Ogier, and Torvik (2013) also find that productivity affects citation performance; however, the effect is not significant when other factors such as authors' past citation performance are also considered.

There is no consensus on whether the gender of authors is an important factor for citation performance. Nunkoo, Hall et al. (2019) find that men receive significantly more citations than women in the tourism field, whereas Hengel and Moon (2020) find that women receive significantly more citations than men in economics. Nielsen (2017) could not find any significant gender effects for management journals. Thelwall (2020) analyzes citation performance in 27 academic fields and shows that whether or not women have a higher citation performance depends on the academic field.

Another inconclusive issue is the effect of the number of authors on the number of citations. Kosteas (2018) finds that the number of authors is a significant factor that affects citations in economics. Levitt (2015) also finds that the number of authors and the number of citations are positively related in economics, but the effect is only valid for articles with fewer than three authors. Bornmann, Schier et al. (2012) do not find any significant relationship between the number of authors and citations of an article for chemistry journals.

The closest to this study is Hamermesh (2018), which analyzes most of the author characteristics included in this paper and uses data from *AER* and four other top economics journals. The study finds that there is no significant difference between men and women in citation performance, but it is found that more senior authors, authors with better past publication

performance, authors from top institutions, and articles with more coauthors receive more citations. However, the study considers each factor affecting citation performance in isolation. In contrast, I run a regression including all the factors and use its results to perform simulations to show the effect of author characteristics on the citation performance from an additional perspective.

This paper is also related to opportunistic editorial practices. Martin (2016) summarizes the common editorial misconduct that tries to enhance the citation performance of journals. The editors force authors to cite papers from their journals; they form journal cartels, where journal A cites journal B, and journal B cites journal A; and they create an online queue of accepted papers, so they can choose the papers that are developing a better citation performance to publish.

There are some opportunistic editorial practices to boost citation performance that cannot be classified as unethical. For example, it is known that positive and strong results are more easily published. Franco, Malhotra, and Simonovits (2014) analyze all the projects in a prestigious sponsored program in social sciences, and they show that if the results are not significant and positive, the results are not likely to be written up, and the papers are not likely to be published if written. The null results are not published, possibly because of their low expected citation performance. The fact that stronger results get more citations is called *citation bias* (Fanelli, Costas, & Ioannidis, 2017).

Editors may choose not to publish articles in certain subfields of economics because of their citation performance. For example, there are few replication experiments (Andrews & Kasy, 2019) in the top economics journals. It is also known that heterodox papers are not published in the top economics journals (Earl & Peng, 2012). This may be due to the fact that heterodox articles do not receive many citations from mainstream articles (Lee, 2012).

Editorial decisions are not made by the editor alone: The editorial team is important. Card and DellaVigna (2020) find that editors closely follow referee decisions for economics journals. They also note that the referee recommendations correlate strongly with the citation performance of the papers. Naturally, the editors have some discretion. For example, they follow the recommendations of the referees who are more productive more closely, although referees who are less productive have equally good performance in predicting the citation performance.

3. DATA

In this paper, I focus on the citation performance of *AER* articles because there is considerable evidence that *AER* is the top economics journal. A survey among a large group of economists concludes that respondents see *AER* as the top economics journal (Axarloglou & Theoharakis, 2003). Many studies that rank economics journals by citation analysis also rank *AER* as the top economics journal (Kalaitzidakis, Mamuneas, & Stengos, 2011).

I collected information about *AER* articles and their authors. I used data from Web of Science, Econ-Lit, and biographies obtained from the internet. The description and the source of the data, and the interpretation of their summary statistics are given as follows.

3.1. Article-Related Factors (Rows 1–5 in Table 1)

I collected the article-related factors of all *AER* articles that were published between 2008 and 2017 from the Web of Science. The citation performance of articles in *AER* is considerably higher than that of other types of publications in *AER*. Therefore, I only considered regular

Table 1. Summary statistics

Row No.	Variables	No. of observations	Mean	Standard deviation
Article-level variables:				
1	Two-year impact	1,052	12.0	13.5
2	Number of authors	1,052	2.33	0.94
3	Year published	1,052	2012.7	2.81
4	Month issued	1,052	7.12	3.46
5	Dummy for theoretical articles	1,052	0.248	0.432
Author-level variables:				
6	Women	2,453	0.156	0.363
	Dummy variables for the continent of the author's affiliated institution			
7	North America	2,453	0.709	0.455
8	Europe	2,453	0.250	0.433
9	Rest of the World	2,453	0.041	0.199
	Dummy variables for the Quacquarelli Symonds rank of the author's affiliated institution			
10	Top 10	2,453	0.280	0.433
11	Ranked between 11 and 50	2,453	0.300	0.459
12	Ranked below 50	2,453	0.419	0.493
	Dummy variables for the years of post-PhD experience			
13	Below 10 years	2,453	0.465	0.499
14	Between 11 and 25 years	2,453	0.379	0.485
15	More than 25 years	2,453	0.119	0.323
16	PhD Top 10	2,453	0.479	0.500
17	No PhD	2,453	0.037	0.189
18	Average citations at top five journals	2,453	5.72	7.46
19	No publications at top five journals	2,453	0.383	0.486
20	Average publications in the last 4 years	2,453	1.17	1.11
21	Published in the top three theoretical journals	2,453	0.271	0.444

articles and excluded other types of publications, such as proceedings, editorials, corrections, and comments. There are 1,052 *AER* articles published in this period.

The first article-related factor is "2-year impact." This variable corresponds to the number of citations that an *AER* article that was published in year T receives in years $T + 1$ and $T + 2$.

I could not get a larger than 2-year citation window for all articles because I collected citation data in July and August of 2020 and included articles from 2017. I did not want to use different time windows for *AER* articles published in different years so as not to complicate the analysis. Moreover, a 2-year window is also used for the Journal Impact Factor, which is a standard journal quality metric. Lastly, there is clear evidence that short-term citation performance is a very strong predictor for long-term citation performance in economics (Kosteas, 2018); therefore the conclusions of this analysis may not be restricted only to short-term impact.

The average citation performance per *AER* article is 12.0 citations for a 2-year window, as seen in Table 1 (Row 1). This number may seem small for a top field journal. However, the average number of citations that economics journals receive is smaller than that of many scientific fields. According to Journal Citation Report 2017, a median economics journal has a 1.11 impact factor, whereas a median physics (multidisciplinary) journal has a 1.65 impact factor and a median chemistry (multidisciplinary) journal has a 2.20 impact factor.

The average number of authors per article is 2.33, as seen in Table 1 (Row 2). Hamermesh (2018) reports that the average number of authors in the top five economics journals is 2.01 for publications in 2007 and 2008. The average number of authors in *AER* articles is also comparable to that of the typical publication in economics. For example, Yuret (2015) reports that the average number of authors for 2012 publications by all faculty members in US economics departments is 2.10.

The last article-related variable is a dummy variable that indicates that the article is a theoretical article (Row 5). This information is collected by skimming through the *AER* articles. Around a quarter of the *AER* articles in my sample are theoretical.

3.2. Author's Gender (Row 6)

I obtained gender information from pictures in the authors' internet biographies. This is not an ideal method, as errors can be made during the process. I independently collected gender information for randomly selected 245 (10% of the sample) authors without using picture information. Most biographies do not contain direct gender information, so I mostly obtained gender information from gender pronouns. I could deduce the gender of 224 (91% of the random sample). The gender information collected through pictures did not contain any errors for these 224 authors. As I could obtain pictures of 100% of the authors, I used the gender obtained from the pictures in the analysis.

I see from Table 1 (Row 6) that 15.6% of the authors are women. The low female ratio in economics is not specific to the top journals. According to Hamermesh (2018), the ratio of females that have higher than median post-PhD experience is as low as 9% in the top 30 economics departments.

3.3. Continent and Ranking of the Author's Affiliated Institution (Rows 7–12 in Table 1)

I found the affiliated institution of the authors by analyzing the address information given by the Web of Science record of the article. A simple internet search revealed the continent of the affiliated institution. The rank of the institutions is taken from the 2020 Quacquarelli Symonds (QS) economics rankings[1].

---

[1] Obtained from https://www.topuniversities.com/university-rankings/university-subject-rankings/2020 /economics-econometrics.

Table 1 shows that 70.7% of all authors work in North America (Row 7), 25.0% of them work in Europe (Row 8), and only 4.1% of the authors work in the rest of the world (Row 9) at the time of the publication. The concentration of authors in North America and Europe is related to the number of top institutions in these continents. Seven out of the top 10 institutions are in North America, and only 12 institutions out of the top 50 are outside Europe and North America according to QS 2020 economics rankings.

Table 1 shows that 28.0% of the authors work at the top 10 institutions (Row 10), 30.0% of the authors work at the institutions ranked 11 to 50 (Row 11), and 41.9% of the authors work at the institutions that are ranked below 50 (Row 12). The high concentration of authors affiliated with a few institutions is not unique to *AER*. Wu (2007) analyzes *AER* and two other journals that are among the top five economics journals. The study finds that 34.3% of the pages in *AER*, 47.2% of the pages in *Journal of Political Economics*, and 57.3% of the pages in *Quarterly Journal of Economics* are written by authors who are affiliated with the top 10 economics departments for years 2000 to 2003.

I collected a sample of 100 articles that are written by 203 authors from five journals that are ranked between 26 and 30 according to Kalaitzidakis et al. (2011) to see whether the statistics from lower ranked journals are comparable to those of *AER*[2]. In this sample, 53.2% of the authors are from North America, 37.0% of the authors are from Europe, and 9.9% of the authors are from the rest of the world. In other words, *AER* has more author affiliations from North America, and fewer author affiliations from Europe and the rest of the world than that of the lower ranked journals. Moreover, only 9.9% of the authors are from the top 10 institutions, 13.8% of the authors are from institutions ranked between 11 to 50, and 76.4% of the authors are from institutions ranked below 50 in the lower ranked journals. Therefore, authors are much less concentrated in the top institutions in the lower ranked journals compared to *AER*.

3.4. Author's PhD Information (Rows 13–17 in Table 1)

Yuret (2020) studied the social network of authors of the top five economics journals for the same period as this study. The study includes PhD information that was obtained from internet biographies. Therefore, I obtained the PhD information from that study's data set.

Table 1 shows that the authors are relatively young economists: 46.5% of the authors obtained their PhDs less than 10 years ago (Row 13). If I exclude 3.7% of authors who did not get a PhD (Row 17), the median post-PhD experience is 10. In my sample of lower ranked journals, the median post-PhD experience is 11. Hamermesh (2018) finds that the median post-PhD experience is 17 for the faculty members in the top 30 economics departments. Therefore, the post-PhD experience of *AER* authors is similar to that of the authors of lower ranked journals but smaller than that of the faculty members in the top economics departments.

A total of 47.9% of the authors obtained their PhDs from the top 10 institutions, as can be seen from Table 1 (Row 16). This level of high concentration is not seen in the lower ranked journals. In my sample of 203 authors from lower ranked journals, the ratio of authors who obtained their PhDs from the top 10 institutions is 25.1%.

---

[2] Twenty articles from each of the following journals are selected randomly: *Journal of Economic Growth*, *Journal of Human Resources*, *Journal of Economic Dynamics & Control*, *Journal of Economic Behavior & Organization*, and *Journal of Business & Economic Statistics*.

3.5. Author's Citation Performance in Top 5 Publications (Rows 18–19 in Table 1)

I collected the authors' citation performance for their past publications at the top five economics journals from Web of Science[3]. If an author publishes an *AER* article in year *T*, I find the citation performance of all of that author's articles at the top five economics journals published from 1980 to $T - 1$. I also used a 2-year window for citations received by past publications. The reason why I restrict past author citation performance to the top five economics journals is twofold. First, I could not use an automated method to pick up citations for a 2-year window, so handpicking citations from more journals was time consuming. Second, I wanted to consider articles in journals that are of comparable quality to *AER*. This is because an author may receive a considerably higher citation performance for publications in the top journals than for those in more modest journals.

Table 1 shows that authors received on average 5.72 citations for their publications in the top five economics journals in a 2-year window (Row 18). This variable treats the average citation performance of 38.3% of authors (Row 19) who have never published in the top five economics journals as zero. If I exclude them, then the average citation performance increases to 9.27. This is still smaller than 12.0, which is the 2-year impact of the current *AER* publications (Row 1). This may be because the number of citations that articles receive increases in general through years and past publications are naturally older than current *AER* publications.

3.6. Author's Past Publication Information (Rows 20–21 in Table 1)

Econ-Lit covers more years and includes more economics journals than Web of Science. Moreover, the index includes full names of authors for more years, so name confusion can be kept at a minimum. As mentioned above, the citation performance of authors in the top five economics journals is obtained from the Web of Science. This is because Econ-Lit does not have citation information.

I had Econ-Lit publication records of the authors in my sample because I have already collected and cleaned the data for Yuret (2020). From these records, I obtained the publications in the last 4 years. For example, for an *AER* article that is published in the year 2012, I gathered all of the authors' publications from years 2008 to 2011. Then, I divide the number of the total publications by four to get a yearly average. I also deduced whether an author has published any articles in the top three theoretical journals from the Econ-Lit publication records[4]. For example, for an *AER* article that is published in the year 2012, I gathered the all of the authors' publications from 1980 to 2011 and see whether there are any articles in the top three theoretical journals.

Table 1 shows that *AER* authors publish an average of 1.17 articles, which may seem to be rather small (Row 20). However, this low performance is typical in economics. Yuret (2016) shows that an average faculty member in the top 10 economics departments publishes 1.94 articles, compared with 1.24 articles in economics departments that are ranked between 11 and 25. The average performance decreases to 0.73 for faculty members in economics departments that are ranked between 151 and 200. The study also reports that the publication performance for chemists is at least five times more than that of economists.

---

[3] Most economics journal rankings agree on the top five economics journals (Card & DellaVigna, 2013). The list of these journals is (in alphabetical order) *AER, Econometrica, Journal of Political Economy, Quarterly Journal of Economics,* and *Review of Economic Studies*.

[4] I took the top three theoretical economics journals from the rankings in Kalaitzidakis et al. (2011). These journals are *Journal of Economic Theory, Games and Economic Behavior,* and *Economic Theory*.

Table 2. Empirical vs. theoretical *AER* articles by authors who have published at least one article in the top three theoretical journals

| | | | *AER* Article (Row 5) | |
		Empirical	Theoretical	Total
Author	No articles in top three theoretical journals	1,569 (87.7 %)	220 (12.3 %)	1,789 (100 %)
	At least one article in top three theoretical journals (Row 21)	351 (52.9 %)	313 (47.1 %)	664 (100 %)
	Total	1,920 (78.2 %)	533 (21.8 %)	2,453 (100 %)

I see from Table 1 that 27.1% of the authors had at least one article in the top three theoretical journals (Row 21). I have already incorporated a dummy variable for the theoretical articles (Row 5). However, authors do not publish in a single subfield. Many theorists also take a role in empirical projects. Table 2 demonstrates that there is no perfect correlation between the subfield of the *AER* article and whether the author has published any papers in the top three theoretical journals in the past. More than half of the authors who have published in the top three theoretical journals publish an empirical *AER* paper. Some 12.3% of the authors who have not yet published in the top three theoretical papers publish a theoretical *AER* article.

4. REGRESSION ANALYSIS

Deschacht and Engels (2014) list the regression methods in the literature that analyze the citation performance. Thelwall and Wilson (2014) conclude that the best performing regression method to analyze citation performance is the General Linear Model with Logarithm because the citation data is skewed. Therefore, I use the General Linear Model with Logarithm for my regression analysis.

There are 1,052 articles, written by 2,453 authors. If three authors write a paper, all article-related variables would take the same value for these three observations. Therefore, these observations are not independently and identically distributed (iid). For this reason, I clustered standard errors at the article level. There are authors who publish more than one *AER* article during the 10 years that I included. In fact, there are 1,845 distinct authors who write these 1,052 articles. However, most author-level variables do not take the same value for the same author. Post-PhD experience, affiliated institution, and publication records may be different for the same author who publishes in different years. Nevertheless, the observations with the same author would not be iid. Therefore, I also report the standard errors that are clustered at the author level.

I give my regression results in Table 3. The coefficients for article-related factors are all in the expected direction and significant at 1% as seen in the first four rows of Table 3. Articles with more coauthors, published in a later year, and published in an early month in the year get more citations. The theoretical articles get fewer citations than empirical articles. As discussed in Section 2, these results are largely consistent with the literature.

Also as discussed in Section 2, there is no consensus on citation performance by gender. I see that there is no significant gender effect on citation performance in my regression (Row 5). The continent of the affiliated institution does not seem to matter for citation performance (Rows 6 and 7)[5]. In contrast, the ranking of the institution matters. Authors from the

---

[5] When I run the regression without other author-related factors, the North American affiliation is positive and significant at 1%, but the result does not hold as more author-related controls are added.

Table 3. Regression results (General Linear Model with Logarithm) Dependent Variable: ln(Two-year impact + 1)

Row #	Variables	Coefficients	Standard deviation (article-level cluster)	Standard deviation (author-level cluster)	Significance
1	Number of authors	0.126	0.024	0.015	***
2	Year published	0.023	0.009	0.006	***
3	Month issued	−0.056	0.006	0.005	***
4	Dummy variable for theoretical articles	−0.484	0.062	0.044	***
5	Female	−0.068	0.043	0.041	
	Dummy variables for the continent of the author's affiliated institution				
6	North America	0.021	0.050	0.039	
7	Rest of the world	−0.112	0.080	0.070	
	Dummy variables for the QS rank of the author's affiliated institution				
8	Top 10	0.148	0.043	0.040	***
9	Ranked below 50	−0.059	0.043	0.038	
	Dummy variables for the years of post-PhD experience				
10	Below 10 years	0.032	0.040	0.038	
11	More than 25 years	−0.010	0.055	0.055	
12	PhD top 10	0.148	0.037	0.040	***
13	No PhD	−0.064	0.091	0.090	
14	Average citations at top five journals	0.018	0.003	0.003	***
15	No publications at top five journals	0.144	0.053	0.049	***
16	Average publications in the last 4 years	0.061	0.018	0.010	***
17	Published in the top three theoretical journals	−0.168	0.046	0.042	***
18	Constant	−44.1	17.4	11.4	***
19	R_squared	0.248			
20	No. of observations	2453			

Notes: (i) Base categories are excluded to avoid dummy trap. Base categories are: "Dummy variable for the continent of the affiliated institution: Europe," "Dummy variable for the rank of the affiliated institution: Ranked between 11 and 50," and "Post-PhD experience between 11 and 25 years"; (ii) *Significant at 10%, **Significant at 5%, ***Significant at 1%; (iii) Standard errors are clustered at the article-level and author-level. The significance levels are the same for both clusters.

top 10 institutions receive more citations (Row 8). However, there is no significant citation performance effect for authors who work at institutions ranked below 50 (Row 9).

There is no significant effect for the years of post-PhD experience (Rows 10 and 11). Authors who obtained their PhD from the top 10 institutions receive significantly more citations (Row 12). This is even though I have other author-related controls such as the ranking of their affiliated institution. There are no significant effects of not having a PhD (Row 13).

All four publication performance variables are significant at 1%, as can be seen in Table 3 (Rows 14–17). Average citations that have been received by authors' earlier top-five journal publications are positively related to the citation performance of their current *AER* article (Row 14). It is interesting to note that authors who have not published in the top five journals previously also have a higher citation performance on their current *AER* publication (Row 15). Therefore, an author who does not have any experience publishing in the top five journals has a higher expected citation performance than an author who has previously published in the top five journals but attains low citation performance for these publications.

Table 3 shows that authors who have published more articles in the past 4 years receive more citations (Row 16). This is interesting because there are no quality adjustments for the number of publications. Yet, authors who have more publications receive more citations. Although the reason for this relation is not obvious, similar results have been reported in previous studies (Card & DellaVigna, 2020).

Authors who have published at least one article in the top three theoretical journals receive fewer citations than authors who do not have any such publications, as can be seen from Table 3 (Row 17). This is interesting because the subfield of the *AER* article is controlled for (Row 4). Therefore, a theorist has significantly less citation performance even after controlling for the subfield of the article.

The fitness level of the regression can be improved by considering additional article-related characteristics. Moreover, the fitness level may be low due to the fact that I am unable to include some unobserved author characteristics. For example, some authors may have better writing skills, and this might improve their citation performance. However, the focus of the paper is the observable author-related factors, and I wonder why authors with certain characteristics are more likely to publish in *AER*. For example, I note that there is a high concentration of authors affiliated with the top 10 institutions. My result, that their citation performance is better, may serve as an explanation, despite the fact that the regression coefficients may be biased due to observed and unobserved variables.

5. SIMULATION

In this section, I present a simulation to show the extent to which *AER*'s average citation performance can be improved if an opportunistic editor is biased towards authors who are expected to receive more citations. It is obvious from the regression analysis that an opportunistic editor can improve the citation performance by accepting proportionally more articles from authors who are educated in and work at higher ranked institutions and have better past publication performance. The simulation in this section shows exactly how much the citation performance can be improved by an opportunistic editor. Moreover, the simulation shows the degree to which authors with favorable backgrounds become more concentrated by an opportunistic policy. Therefore, the main aim of the simulation exercise is to show the effects already presented in the regression analysis more clearly.

I should note that the simulation in this section is a purely hypothetical exercise and I do not claim that editors actually use policies to inflate their Journal Impact Factor. If they use opportunistic policies to maximize the immediate citation potential of the journal, then it would be detrimental to the prestige of the journal in the long term. It is totally fine if the editors aim to accept better articles and these articles happen to have higher citation performance, because the number of citations is a good proxy for quality. However, if the editors directly aim for the citation potential, researchers would not submit to a journal that does not fairly judge the quality of their article. Nevertheless, the simulation exercise shows that there is

Table 4. An example for the simulation exercise

Article	Two-year impact	Predicted citation performance
A	10	11
B	8	7
C	6	8
D	4	2

a temptation for an opportunistic editorial policy because the immediate citation improvement by such a policy is substantial.

I consider two types of opportunistic editors. First, I consider the opportunistic editor who has perfect foresight, so that the editor knows exactly the number of citations that an article will receive before the article is accepted. Second, I consider the opportunistic editor who has limited information. I reran the General Linear Model with Logarithm presented earlier in the regression analysis section by restricting the sample for articles published from 2008 to 2014. Then, I use the coefficients to compute the predicted citation performance for articles published from 2015 to 2017. Therefore, the opportunistic editor uses the information by using the first 7 years of my sample to form expectations for the last 3 years of my sample. Lastly, the opportunistic editor simply takes the average of the predicted performance of authors to find the predicted citation performance of the article[6]. Both types of editors select half and a quarter of articles within each of the 3 years.

Table 4 gives a simple example of just four articles that were published in a single year to demonstrate how the simulation works. I see from the second column of the table that the total 2-year citation performance of four articles is 28, so the average citation performance is 7. Let's assume that the opportunistic editors select half of the articles. Then the opportunistic editor who has perfect foresight would select articles A and B so that the average citation performance increases to 9. The opportunistic editor who can only use the predicted performances of the articles would pick A and C, so the average citation performance of the articles increases to 8. In the simulation exercise, I also want to see what would happen if the opportunistic editor selects a quarter of the articles. Then, opportunistic editors select only one article. In this case, both types of opportunistic editor would select article A and have an average citation performance of 10.

Figure 1 shows the average 2-year citation performance of both types of opportunistic editor. The first bar is the average citation performance of all articles that are published between 2015 and 2017, and this value is used as a benchmark. The second and third bars give the average citation performance when half and a quarter of the articles are selected by an opportunistic editor who has perfect foresight, respectively. The average citation performance of the articles increases from 13.7 to 22.3 when half of the articles are selected and to 32.0 when a quarter of the articles are selected. In other words, an opportunistic editor with perfect foresight can improve the citation performance by 63.0% and 133.8% when selecting half or a quarter of the articles, respectively.

---

[6] There are more sophisticated methods to account for the expected performance of an article from the expected performance of its authors (see Ahmadpoor & Jones, 2019). However, I prefer using the average to keep the simulation exercise simple.
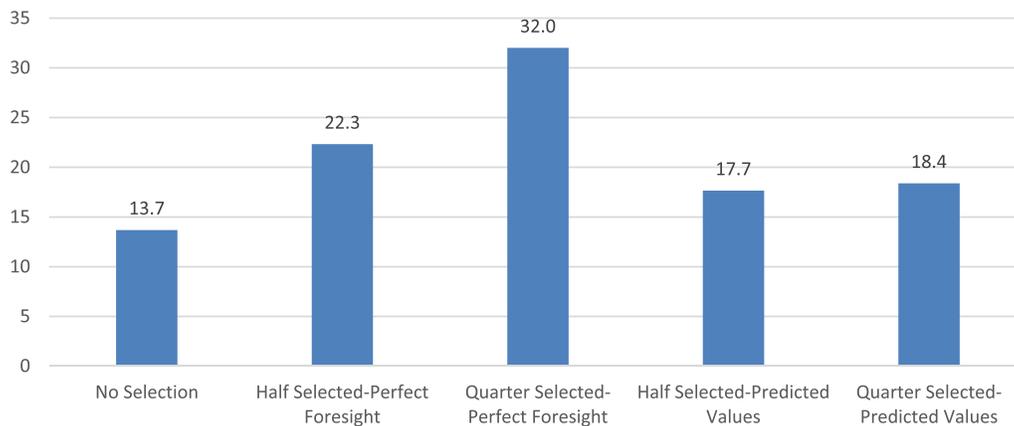
Figure 1. Average 2-year impact of *AER* articles: no selection, half and a quarter of articles selected under perfect foresight, and half and a quarter of articles selected by using predicted values.

The fourth and fifth bars of Figure 1 give the average citation performance of the articles when half and a quarter of the articles are selected by the opportunistic editor who relies on predicted values from regression results, respectively. The improvement is lower than the perfect foresight case, but it is still large. The average 2-year impacts are 17.7 and 18.4 when half and a quarter of the articles are selected, respectively. Compared to the benchmark value of 13.7, this corresponds to an increase of 28.9% and 34.2% when half and a quarter of the articles are selected, respectively.

Figure 2 shows that the ratio of authors who work at the top 10 institutions increases from 30.8% to 39.0% and 38.5% when half and a quarter of the articles are selected by an opportunistic editor with perfect foresight, respectively. The opportunistic editor increases the ratio of the authors who work in the top 10 institutions even more when the predicted values are used. As I discussed at the end of the regression analysis section, there may be missing observed and unobserved variables in the regression. Therefore, the opportunistic editor has to rely more on the author characteristics that I include, and this increases the representation of authors with certain characteristics.

Figure 3 shows the ratio of authors who work at institutions in North America. A total of 69.7% of the authors work at institutions in North America in all *AER* publications between 2015 and 2017. This benchmark value shows a high level of concentration. However, the
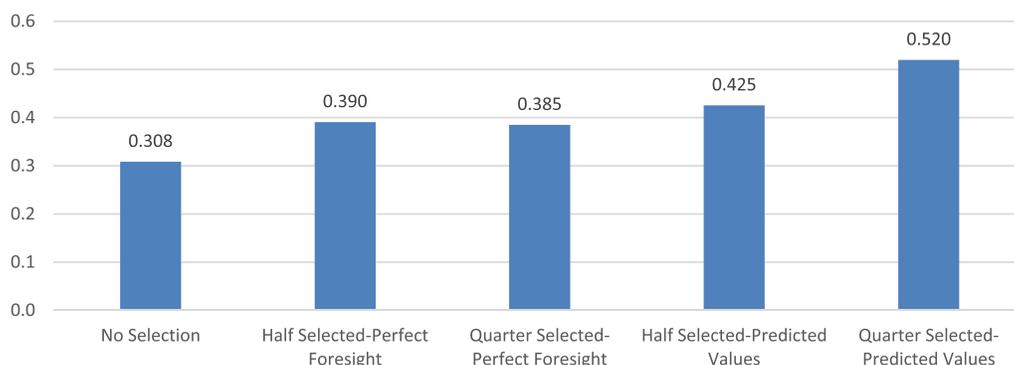

Figure 2. Ratio of authors who work at top 10 institutions: no selection, half and a quarter of articles selected under perfect information, and half and a quarter of articles selected by using predicted values.
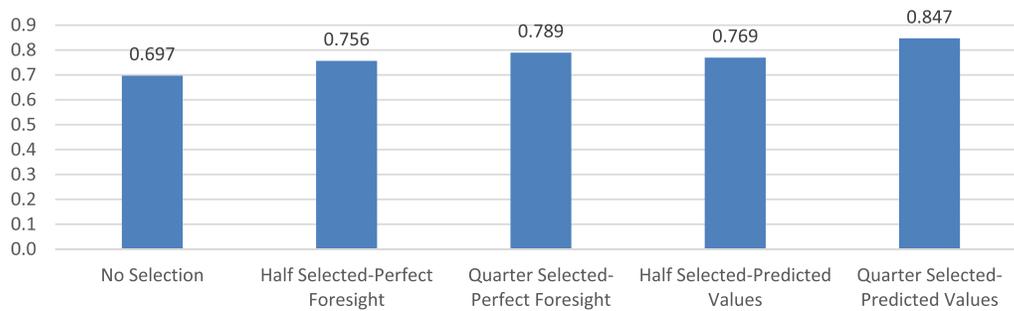
Figure 3. Ratio of authors who work in North America: no selection, half and a quarter of articles selected under perfect information, and half and a quarter of articles selected by using predicted values.

authors are even more concentrated when half and a quarter of the articles are selected. The ratio of authors who work in North American institutions increases to 84.7% when a quarter of authors are selected by using predicted values.

In Figure 4, I see that authors who have published in the top three theoretical journals are represented less when half and a quarter of articles are selected to improve *AER*'s citation performance. The ratio of authors who have published at least one article in the top three theoretical journals decreases from 26.8 to 17.0% when half of the articles are selected, and decreases to 12.6% when a quarter of articles are selected by the opportunistic editor with perfect foresight. The ratio of authors who have theoretical publications becomes even smaller when the opportunistic editor relies on predicted values. Therefore, a policy that solely maximizes citation performance would exclude many authors who had theoretical publications in their publication records.

Figure 5 shows the ratio of female authors who publish in *AER*. Just 16.6% of authors are female between 2015 and 2017; that is, female authors are highly underrepresented. The ratio of female authors increases slightly when half and a quarter of the publications are selected, under both perfect insight and predicted performance. Therefore, an opportunistic editorial policy to improve citation performance of *AER* articles would increase the representation of women in *AER* publications.

There are obvious limitations to the simulation results. I rely on accepted papers rather than all submitted papers. The editor may learn about the citation performance of the rejected
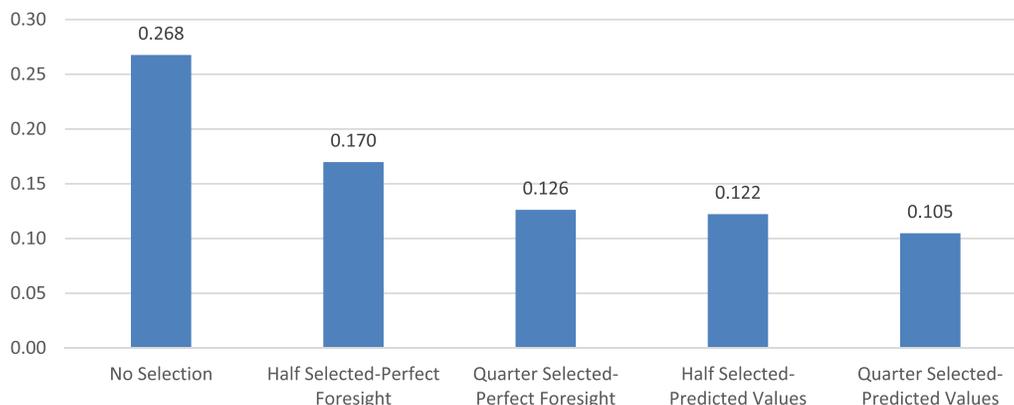

Figure 4. Ratio of authors who have published at least one paper in the top three theoretical journals: no selection, half, and a quarter of articles selected under perfect foresight, and half and a quarter of articles selected by using predicted values.
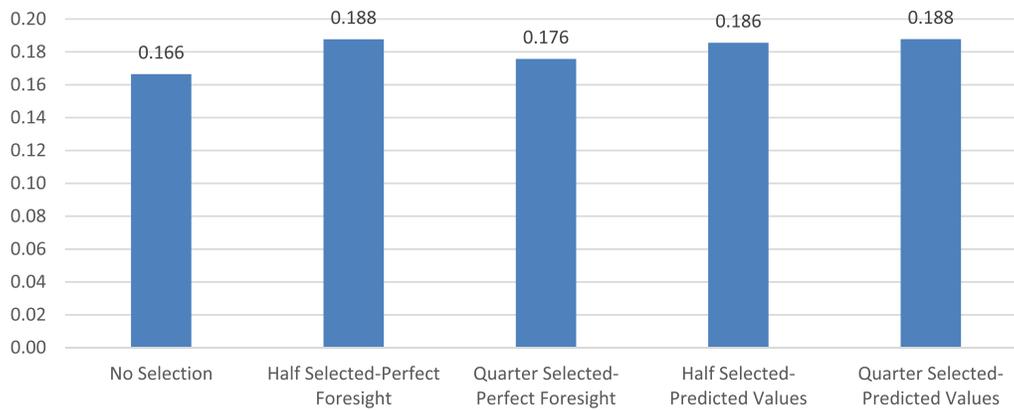
Figure 5. Ratio of female authors: no selection, half and a quarter of articles selected under perfect foresight, and half and a quarter of articles selected by using predicted values.

papers and will have a better idea about the citation potential of any given article that I suggest. It is also the case that editors observe some variables that are unobservable to us. For example, the editor may see whether the subject matter of the article is popular. In this case, the opportunistic editor's selection may be different and citation performance improvements may be less substantial than I suggest.

6. CONCLUSION

I have demonstrated that there is a high concentration of authors who have excellent academic backgrounds among the authors of *AER*. Then, I predicted the citation performance by mainly using author characteristics. I show that authors who have stronger academic backgrounds receive more citations. Next, I performed simulations by asking a hypothetical question: What would happen if a subset of the accepted papers is selected by an opportunistic editor. I considered two types of opportunistic editors: an opportunistic editor who has perfect foresight can improve the average citation performance of *AER* by 133.8%, and an opportunistic editor who uses predicted values from my regressions can improve the average citation performance of *AER* by 34.2% when a quarter of the current articles are selected.

The citation performance is used as the standard proxy for the quality of an article. Therefore, I cannot claim that the authors with certain characteristics are overrepresented because the journal inflates its Impact Factor. Because of the positive correlation between the quality of an article and its citation performance, authors who are expected to receive more citations may be preferred for the quality of their articles. Nevertheless, my analysis shows that there is a potential that the average citation performance of *AER* can be improved substantially if it is desired by an opportunistic editor. Another problem that is created by the opportunistic editor is that the journal would rely on the contribution of a smaller group of researchers, which might be a handicap. For example, authors who publish in theoretical journals are less likely to publish at *AER* if the journal relies on citation performance. This would limit the intellectual diversity of the journal.

An honest editor may want to maximize the average quality of articles in a journal. Suppose for a moment that the number of citations is a perfect measure for quality, and the quality of the article and its citations are perfectly correlated. In this case, the problem that the honest editor faces would be the same as the opportunistic editor in my simulation exercise. For example, the honest editor with perfect foresight would accept more articles by the authors who are

affiliated with the top 10 institutions, because these authors write higher quality articles. Moreover, the honest editor who has limited information would accept even more authors who are affiliated with the top 10 institutions, because being affiliated with these institutions serve as a good proxy for the quality of their articles. In other words, the honest editor needs to discriminate against authors with certain characteristics to maximize the expected average quality of the articles in the journal.

AER considers itself a general-interest economics journal that is among the most scholarly journals in economics[7]. I show that the author characteristics are concentrated in *AER*. It is a good feature of such a high-quality journal to give a fair chance to authors. Blind reviewing would solve the problem, but it is getting more difficult to hide the identity of authors. Another policy would be to give statistics about the authors in papers that are accepted and rejected. For example, if theorists' articles are disproportionately rejected, the journal may investigate whether this is the result of a fair procedure.

COMPETING INTERESTS

The author has no competing interests.

FUNDING INFORMATION

This research did not receive any funding.

DATA AVAILABILITY

I am not able to share the data of this study because the data include the citation performance of individual articles that are extracted from the Web of Science.

REFERENCES

Ahmadpoor, M., & Jones, B. F. (2019). Decoding team and individual impact in science and invention. *Proceedings of the National Academy of Sciences*, *116*(28), 13885–13890. https://doi.org/10.1073/pnas.1812341116, PubMed: 31235568

Amara, N., Landry, R., & Halilem, N. (2015). What can university administrators do to increase the publication and citation scores of their faculty members? *Scientometrics*, *103*(2), 489–530. https://doi.org/10.1007/s11192-015-1537-2

Andrews, I., & Kasy, M. (2019). Identification of and correction for publication bias. *American Economic Review*, *109*(8), 2766–2794. https://doi.org/10.1257/aer.20180310

Axarloglou, K., & Theoharakis, V. (2003). Diversity in economics: An analysis of journal quality perceptions. *Journal of the European Economic Association*, *1*(6), 1402–1423. https://doi.org/10.1162/154247603322752584

Bornmann, L., Schier, H., Marx, W., & Daniel, H. (2012). What factors determine citation counts of publications in chemistry besides their quality? *Journal of Informetrics*, *6*(1), 11–18. https://doi.org/10.1016/j.joi.2011.08.004

Buela-Casal, G., Zych, I., Medina, A., Viedma Del Jesus, M. I., Lozano, S., & Torres, G. (2009). Analysis of the influence of the two types of the journal articles; theoretical and empirical on the impact factor of a journal. *Scientometrics*, *80*(1), 265–282. https://doi.org/10.1007/s11192-008-1715-6

Card, D., & DellaVigna, S. (2013). Nine facts about top journals in economics. *Journal of Economic Literature*, *51*(1), 144–161. https://doi.org/10.1257/jel.51.1.144

Card, D., & DellaVigna, S. (2020). What do editors maximize? Evidence from four economics journals. *Review of Economics and Statistics*, *102*(1), 195–217. https://doi.org/10.1162/rest_a_00839

Chan, H., Guillot, M., Page, L., & Torgler, B. (2015). The inner quality of an article: Will time tell? *Scientometrics*, *104*(1), 19–41. https://doi.org/10.1007/s11192-015-1581-y

Colussi, T. (2018). Social ties in academia: A friend is a treasure. *Review of Economics and Statistics*, *100*(1), 45–50. https://doi.org/10.1162/REST_a_00666

Deschacht, N., & Engels, T. C. E. (2014). Limited dependent variable models and probabilistic prediction in informetrics. In: Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring scholarly impact: Methods and practice* (pp. 193–214). Springer. https://doi.org/10.1007/978-3-319-10377-8_9

Donner, P. (2018). Effect of publication month on citation impact. *Journal of Informetrics*, *12*(1), 330–343. https://doi.org/10.1016/j.joi.2018.01.012

---

7 https://www.aeaweb.org/journals/aer

Earl, P. E., & Peng, T. (2012) Brands of economics and the Trojan horse of pluralism. *Review of Political Economy*, *24*(3), 451–467. https://doi.org/10.1080/09538259.2012.701927

Fanelli, D., Costas, R., & Ioannadis, P. A. (2017). Meta-assessment of bias in science. *Proceedings of the National Academy of Sciences*, *114*(14), 3714–3719. https://doi.org/10.1073/pnas .1618569114, PubMed: 28320937

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502–1504. https://doi.org/10.1126/science .1255484, PubMed: 25170047

Hamermesh, D. S. (2018). Citations in economics: Measurement, uses and impacts. *Journal of Economic Literature*, *56*(1), 115–156. https://doi.org/10.1257/jel.20161326

Haslam, N., Ban, L., Kaufman, L., Loughan, S., Peters, K., … Wilson, S. (2008). What makes an article influential? Predicting impact in social and personality psychology. *Scientometrics*, *76*(1), 169–185. https://doi.org/10.1007/s11192-007-1892-8

Hengel, E., & Moon, E. (2020). Gender and quality at top economics journals. *Mimeo*.

Hurley, L. A., Ogier, A. L., & Torvik, V. I. (2013). Deconstructing the collaborative impact: Article and author characteristics that influence citation count. *Proceedings of the American Society for Information Science and Technology*, *50*(1), 1–10. https://doi .org/10.1002/meet.14505001070

Johnston, D. W., Piatti, M., & Torgler, B. (2013). Citation success over time: Theory or empirics? *Scientometrics*, *95*(3), 1023–1029. https://doi.org/10.1007/s11192-012-0910-7

Kalaitzidakis, P., Mamuneas, T. P., & Stengos, T. (2011). An updated ranking of academic journals in economics. *Canadian Journal of Economics*, *44*(4), 1525–1537. https://doi.org/10.1111/j.1540 -5982.2011.01683.x

Kosteas, V. D. (2018). Predicting long-run citation counts for articles in top economics journals. *Scientometrics*, *115*(3), 1395–1412. https://doi.org/10.1007/s11192-018-2703-0

Lee, F. S. (2012). Heterodox economics and its critics. *Review of Political Economy*, *24*(2), 337–351. https://doi.org/10.1080 /09538259.2012.664360

Levitt, J. M. (2015). What is the optimal number of researchers for social science research? *Scientometrics*, *102*(1), 213–225. https:// doi.org/10.1007/s11192-014-1441-1

Lindahl, J. (2018). Predicting research excellence at the individual level: The importance of publication rate, top journal publications, and top 10% publications in the case of early career mathematicians. *Journal of Informetrics*, *12*(2), 518–533. https://doi .org/10.1016/j.joi.2018.04.002

Martin, B. R. (2016). Editors' JIF-boosting stratagems—Which are appropriate and which not? *Research Policy*, *45*(1), 1–7. https:// doi.org/10.1016/j.respol.2015.09.001

Medoff, M. H. (2003). Collaboration and the quality of economics research. *Labor Economics*, *10*(5), 597–608. https://doi.org/10 .1016/S0927-5371(03)00072-1

Nielsen, M. W. (2017). Gender and citation impact in management research. *Journal of Informetrics*, *11*(4), 1213–1228. https://doi .org/10.1016/j.joi.2017.09.005

Nunkoo, R., Hall, C. M., Rughoobur-Seetah, S., & Teeroovengadum, V. (2019). Citation practices in tourism research: Toward a gender conscientious engagement. *Annals of Tourism Research*, *79*, 102755. https://doi.org/10.1016/j.annals.2019.102755

Onodera, N., & Yoshikane, F. (2015). Factors affecting citation rates of research articles. *Journal of the Association for Information Science and Technology*, *66*(4), 739–764. https://doi.org/10 .1002/asi.23209

Petersen, A. M., Pan, R. K., Pammolli, F., & Fortunato, S. (2019). Methods to account for citation inflation in research evaluation. *Research Policy*, *48*(7), 1855–1865. https://doi.org/10.1016/j .respol.2019.04.009

Smolinsky, L., & Lercher, A. (2012). Citation rates in mathematics: A study of variation by subdiscipline. *Scientometrics*, *91*(3), 911–924. https://doi.org/10.1007/s11192-012-0647-3

Tahamtan, I., Afshar, A. S., & Ahamdzadeh, K. (2016). Factors affecting number of citations: a comprehensive review of the literature. *Scientometrics*, *107*(3), 1195–1225. https://doi.org/10 .1007/s11192-016-1889-2

Thelwall, M., & Wilson, P. (2014). Regression for citation data: An evaluation of different methods. *Journal of Informetrics*, *8*(4), 963–971. https://doi.org/10.1016/j.joi.2014.09.011

Thelwall, M. (2020). Gender differences in citation impact for 27 fields and six English-speaking countries 1996–2014. *Quantitative Science Studies*, *1*(2), 599–617. https://doi.org/10.1162/qss_a_00038

Trimble, V., & Ceja, J. A. (2013). Are American astrophysics papers accepted more quickly than others? Part II: correlations with citation rates, subdisciplines, and author numbers. *Scientometrics*, *95*(1), 45–54. https://doi.org/10.1007/s11192-012-0793-7

Wang, F., Fan, Y., Zeng, A., & Di, Z. (2019). Can we predict ESI highly cited publications? *Scientometrics*, *118*(1), 109–125. https://doi.org/10.1007/s11192-018-2965-6

Wu, S. (2007). Recent publishing trends at the *AER, JPE* and *QJE*. *Applied Economics Letters*, *14*(1), 59–63. https://doi.org/10 .1080/13504850500425212

Xie, J., Gong, K., Li, J., Ke, Q., Kang, H., & Cheng, Y. (2019). A probe into 66 factors which are possibly associated with the number of citations an article received. *Scientometrics*, *119*(3), 1429–1454. https://doi.org/10.1007/s11192-019-03094-z

Yuret, T. (2015). Interfield comparison of academic output by using department level data. *Scientometrics*, *105*(3), 1653–1664. https://doi.org/10.1007/s11192-015-1621-7

Yuret, T. (2016). An analysis of the foreign-educated elite academics in the United States. *Journal of Informetrics*, *11*(2), 358–370. https://doi.org/10.1016/j.joi.2017.02.008

Yuret, T. (2020). Co-worker network: How closely are researchers who published in the top five economics journals related? *Scientometrics*, *124*(3), 2301–2317. https://doi.org/10.1007/s11192 -020-03589-0