Gender differences in citations at top economics journals
Even more evidence that women are held to higher standards in peer review*

Erin Hengel†

January 2019

This paper examines gender differences in citations for articles published in top-five economics journals. On average, male-authored papers are cited more than female-authored papers. Yet this finding is driven by a small number of highly cited papers, most of which were written by men—many of whom are Nobel prize winners—over 30 years ago. After controlling for time, author prominence and including either fixed effects for 91 superstar economists or all Nobel prize winners, I find female-authored papers are actually cited more. Moreover, when skewness in the distribution of citations is adjusted using the inverse hyperbolic sine function, female-authored papers are always cited more. I additionally estimate the marginal impact of co-authoring with more women for male authors, only. Using a fixed effects framework, I find that men earn 13 log points more citations when they increase the share of female co-authors on a paper by 50 percent. I conclude by roughly applying a theoretical framework developed in Hengel (2018) to identify the cause behind higher citations in female-authored papers. The results demonstrate suggestive evidence that women's higher citations are driven by factors outside their control.

* See Hengel (2018) for further information on data sources, citations and references. All acknowledgements given in that paper apply here as well.

† University of Liverpool, Department of Economics; email: ehengel@liverpool.ac.uk.

1 Introduction

Women are underrepresented in top economics journals. The average ratio of female authors per paper is 12.5 percent at the *American Economic Review* (*AER*), 10 percent at the *Quarterly Journal of Economics* (*QJE*), 7.7 percent at the *Review of Economic Studies* (*Restud*), 6.3 percent at the *Journal of Political Economy* (*JPE*) and 4.2 percent at *Econometrica*. Only a small share of papers are majority female-authored: 6.8 percent (*AER*), 4.9 percent (*QJE*), 3.8 percent (*JPE*), 3.4 percent (*Restud*) and 1.7 percent (*Econometrica*).

This paper investigates peer review's contribution to women's underrepresentation in economics journals. It builds on earlier research suggesting women are held to a higher bar in the peer review process: in Hengel (2018) I found that higher writing standards in peer review cause female economists to write at least seven percent more clearly than they otherwise would and this plausibly contributes to their noticeably longer review. In this paper, I investigate whether higher standards in peer review also lead to more citations for female-authored work.

I find that the ratio of female authors on a paper is negatively correlated with citation counts—but this is entirely driven by a very small subset of articles with extraordinarily high citations. These papers were predominately authored by men in the 1980s. Many of the authors are Nobel prize winners; many of the papers are why they won. After including fixed effects for the top 1 percent of authors in order to account for this and controlling for time, I find female-authored papers generate about 10–15 more citations than male-authored papers. When skewness in the distribution of citations is adjusted using the inverse hyperbolic sine function, female-authored papers are always cited more.

I additionally analyse citations at the author-level in order to tease out the marginal impact of co-authoring with more women for male authors, only. Using a fixed effects framework, I show that men who go from co-authoring a paper with no women to co-authoring a paper with 50 percent women will experience an increase in citations of about 13 log points.

This paper concludes with a rough application of the logical framework developed in Hengel (2018). Although the evidence is non-causal, the resulting analysis suggests factors beyond women's control are at least partially responsible for women's higher citations.

This paper joins a number of papers investigating gender differences in citations. Most evidence suggests there isn't any (for an overview of past research, see Ceci et al., 2014). Yet studies specific to economics are relatively sparse and somewhat contradictory. Bransch and Kvasnicka (2017), Hamermesh (2018), Laband (1987), and Smart and Waldfogel (1996) find no significant difference. Others do. Ginther and Kahn (2004) find that women have more citations per publication when measured by top–10 publications, journal publications and total publications; Card et al. (2018) find that female-authored papers at four general interest journals receive about 25 percent more citations after controlling for referee recommendations. Grossbard et al. (2018) find female-authored papers published in two field journals are cited more. Meanwhile, an earlier study by Ferber (1988) found that male-authored papers were cited more than female-authored papers.

The remainder of the paper proceeds in the following order. In the next section, I describe the data used throughout the paper. In Section 3, I give an overview of women's representation in top economics journals and discuss the distribution of citations in those journals. Analyses and results are presented in Section 4 and followed by conclusions (Section 5).

2 Data

The data include 10,954 articles published between 1950–2015 in the *American Economic Review* (*AER*), *Econometrica*, the *Journal of Political Economy* (*JPE*), the *Quarterly Journal of Economics*

Table 1: Article count, by journal and decade

Decade	*AER*	*ECA*	*JPE*	*QJE*	*Restud*	Total
1950–59		120				120
1960–69		344	184			528
1970–79		660	633	1	228	1,522
1980–89	180	648	562	401	490	2,281
1990–99	476	443	478	409	383	2,189
2000–09	695	519	408	413	430	2,465
2010–15	732	382	181	251	303	1,849
Total	2,083	3,116	2,446	1,475	1,834	10,954

Notes. Included is every article published with an English abstract between January 1950 and December 2015 in *AER*, *Econometrica*, *JPE* and *QJE* and every article published after 1,976 at *Restud*. Papers published in the May issue of *AER* (*Papers & Proceedings*) are excluded. Final row and column display total article counts by journal and decade, respectively. Data were originally collected for Hengel (2018) in order to analyse gender differences abstract readability (*AER*, *Econometrica*, *JPE* and *QJE*) and submit-accept times (*Econometrica* and *Restud*).

(*QJE*) and the *Review of Economic Studies* (*Restud*). Articles from the *Papers & Proceedings* issue of the *AER* (*P&P*), errata and corrigenda are excluded. Table 1 breaks down coverage by journal and decade.

Data were originally collected for Hengel (2018) in order to analyse gender differences in abstract readability at *AER*, *Econometrica*, *JPE* and *QJE* and submit-accept times at *Econometrica* and *Restud*. The largest sample is from *Econometrica* which regularly published abstracts with articles before 1950; *JPE* followed in the 1970s; *AER* and *QJE* in the 1980s. *Restud*'s sample starts in 1976, the first year it printed submission and acceptance dates in published papers.

Each author was manually assigned a gender either by me or a research assistant based on (i) obviously gendered given names (*e.g.*, "James" or "Brenda"); (ii) photographs on personal or faculty websites; (iii) personal pronouns used in text written about the individual; (iv) by contacting an author himself or associates and institutions connected to him. We identified the gender of all 7,574 distinct authors in the data.

Citation data were obtained from Web of Science, a comprehensive database of all social science research published since 1900 (Web of Science, 2018). Counts correspond to the number of published papers in the Web of Science database that cite a given article, including self-citations to later work.

To match articles in my dataset with those in Web of Science, I first searched for an exact match using the journal name, article title, volume, issue and page numbers. When exact matches were not found—due either to errors in the Web of Science database or my own—I manually searched for a match. All 10,954 articles were eventually matched with a Web of Science record.[1] Data for *AER*, *Econometrica*, *JPE* and *QJE* were collected in early January 2018; citations for *Restud* were collected in late October 2018.

Analysing gender citation gaps raises several issues: (i) younger articles have had less time to accrue citations; (ii) older articles are disproportionately male; (iii) mean citation counts may be distorted by a small number of superstar economists; (iv) superstar economists are *also* disproportionately male (see Section 3.2 for further discussion). (i) and (iii) skew the distribution of citations; (ii) and (iv) distort gender differences if they correlate with uncontrolled for (or unob-

[1] In an earlier version of this paper, I noted that Web of Science had no entry for Ryosuke Hotaka's article "Some Basic Problems on Excess Demand Functions" (*Econometrica*, 1971). They have subsequently updated their database to include it.

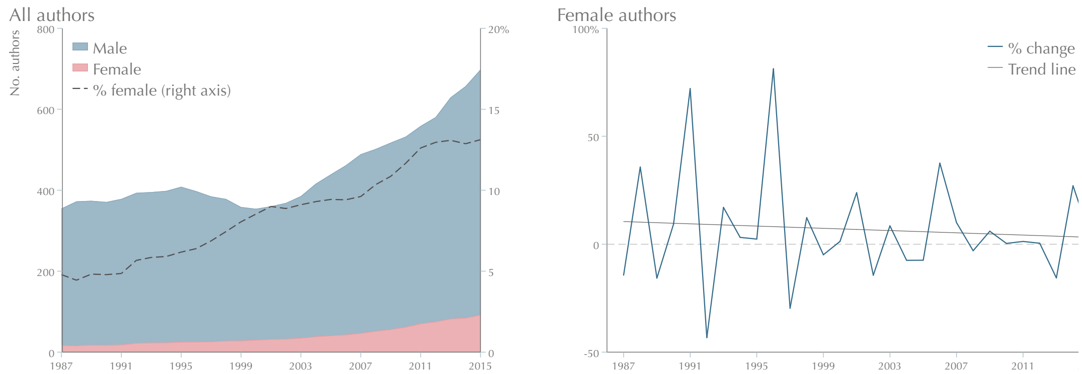

Figure 1: Authors published in a top-five journal

Notes. Figure on the left displays the stacked five-year moving averages of the total number of female (pink) and male (blue) authors published in a top-five journal between 1987–2015; the percentage of female authors is shown in grey (right axis). Graph on the right is the five-year moving average percentage change in the share of female authors (blue); the gray line is the trend over time.

served) factors that generate citations—*e.g.*, winning the Nobel prize—or produce self-reinforcing loops that create citations all on their own.

I principally account for (i)–(iv) by controlling for journal-year fixed effects and transforming raw counts with the inverse hyperbolic sine function (asinh).[2] I additionally construct a set of "superstar" fixed effects in order to illustrate how (iii) and (iv) combine to distort the gender citation gap at the average.

I define "superstar" to include any author who satisfies one or more of the following criteria: (1) 17 or more top-five publications; (2) 10 or more top-five publications, one of which is cited at least 2,500 times; (3) 5 or more top-five publications, one of which is cited at least 5,000 times.

(1) defines superstar according to quantity, alone. It is based on the assumption that most economists are familiar with the work of anyone publishing 17 or more top-five articles, whatever their quality. It is satisfied by one percent of authors in the database. Seventeen also corresponds to one plus the lifetime number of publications of the most prolific female economist.[3] (2) and (3) account for famous economists who are less prolific—*e.g.*, Paul Krugman—operate in fields with slower production functions—*e.g.*, industrial organisation—or publish extensively in other disciplines—*e.g.*, Daniel Kahneman. General results and conclusions do not change by making marginal adjustments to any criteria or swapping superstar effects with fixed effects only for authors winning the Nobel Prize.

About 1.2 percent of authors in the database satisfy conditions (1), (2) or (3). On average, each has published 21 times in a top-five journal; his highest cited paper is cited 1,797 times. Almost a third either won the Nobel prize, the John Bates Clark medal or both. Every single one is male. Appendix A lists them all.

Other control variables used in the analysis include journal fixed effects, year fixed effects, tertiary *JEL* category fixed effects, number of co-authors (N) and a control for author seniority (max. t_5) equal to the total number of top-five articles for the most prolific co-author at the time an article was published. In a subset of regressions, I additionally control for how prominent an author is at the time citation counts were collected (max. T_5). It is equal to the total number of top-five articles for the most prolific co-author as of the end of 2015.

[2] For a discussion of the function's properties with respect to citations, see Card and DellaVigna (2017).

[3] The female economist with the largest publication count is Esther Duflo. As on December 2015, she had 16 top-five publications.

Figure 2: Single-sex papers

Notes. First three figures display five-year moving averages of the number of male- (blue) and female-authored (pink) single-sex papers that are solo-authored, co-authored by two people and co-authored by three or more people, respectively. Grey line is the percentage of female-authored papers among each graph's respective sample. Bottom right-hand graph shows the stacked five-year moving averages of the total number of female (pink) and male (blue) authors on single-sex papers. Grey dashed line shows the percentage who are female.

3 Empirical setting

3.1 Where are the women?

The share of female authors published in a top-five journal was 7–8 percent higher in 2015 than it was in 1986 (Figure 1, left panel). Nevertheless, growth in that share has recently slowed (right panel); the percentage of women appears to be converging to around 15 percent. In this section, I investigate the source of the original growth and reasons for its slowdown.

First, I find no evidence that journals are publishing more exclusively female-authored papers nor that single-sex papers by women are co-authored by more people. The first three graphs in Figure 2 plot the number of single-sex articles. Top-five journals publish about as many solo female-authored papers today as they did in the late 1980s (Figure 2, top left)—seven in 1986, nine in 1992 and 11 in 2015—but the number of solo male-authored papers has declined since the 1990s—in 1986, 125 were published, that fell to 99 in 1992 and 45 by 2015. As a result, the proportion of solo-authored papers by women has increased from 5 percent in 1986 to 20 percent in 2015 (Figure 2 top left grey line).

Nevertheless, men have made up for their declining share of solo-authored papers by publishing more papers co-authored only with other men. The upper right hand and lower left hand graphs in Figure 2 show absolute numbers of articles by two and three or more (respectively) authors of the same gender. In 1987, top-five journals collectively published 79 articles co-authored by two men and zero articles co-authored by two women. In 2015, the corresponding figures were

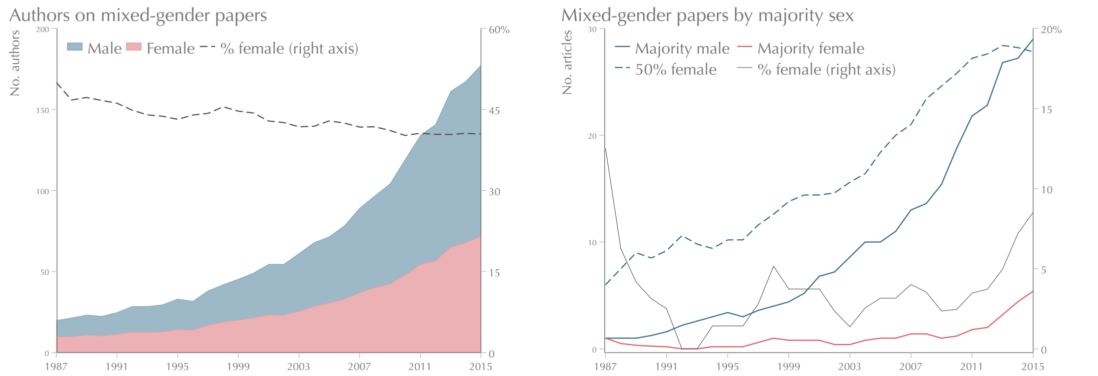

Figure 3: Mixed-gender papers (I)

Notes. Left-hand figure displays the total number of male (blue) and female (pink) authors of mixed-gender papers; percentage female shown with the dashed grey line (right axis). Right-hand figure displays the total number of mixed-gender papers with a strict majority of male (blue line) and female authors (pink line); papers with an equal number of each gender shown with the dashed blue line. The percentage of papers with a strict majority of female authors across all mixed-gender papers are shown with the solid grey line (right axis). All graphs shown as five year moving averages.

103 and one. Meanwhile, journals have sharply increased the number of single-sex articles they publish by three or more men—99 were published in 2014 versus 15 in 1986[4]—only ***six*** however, have ever been published by women.

The lower right-hand graph in Figure 2 plots the total number of male (blue) and female (pink) authors publishing single-sex papers in top-five journals. The male decline in solo-authored papers is more than offset by the substantial increase in exclusively male co-authored papers. As a result, the proportion of female authors among all authors of single-sex papers has hardly changed since in the late 1980s. Since single-sex papers make up the vast majority of published articles and just one top-five co-authored publication satisfies tenure requirements at most research active institutions, the dearth of co-authored papers entirely by women probably contributes to their lower tenure rates.[5]

Second, I do not find that women make up a greater share of authors on mixed-gender papers. Journals are publishing more papers with at least one female author, but the number of male authors on these papers has increased slightly faster than the number of female authors. As a result, the share of women on mixed gender papers has declined since the late 1980s (Figure 3, left-hand-graph).

Moreover, majority female-authored papers are almost as rare today as they were 30 years ago. Figure 3's right-hand graph shows aggregate numbers of mix-gender papers that are authored by a male majority vs. a female majority. Before 2000, few papers with a strict majority of female authors (pink) were published in top-five journals. Since then, they publish about three a year.[6] Because the number of mixed-gendered papers with a majority of *male* authors (blue) and equal number of male and female authors (dashed line) have also risen, however, there has been little or no growth in majority female-authored papers as a share of all mixed-gender papers (grey line).

Third, top-five journals publish more mixed-gender papers. This is the only reason why the share of female authors has grown since the late 1980s. The left-hand graph of Figure 4 plots the total number of articles published each year that are exclusively male-authored (dark blue), exclusively female-authored (pink) and co-authored by members of both sexes (light blue). The

[4]71 and 74 were published in 2012 and 2013, respectively, but only 65 were published in 2015.

[5]Single-sex papers made up two-thirds of all articles—and their authors three quarters of all authors—published in top-five journals in 2015.

[6]That increased to almost eight a year (on average) starting in 2012.

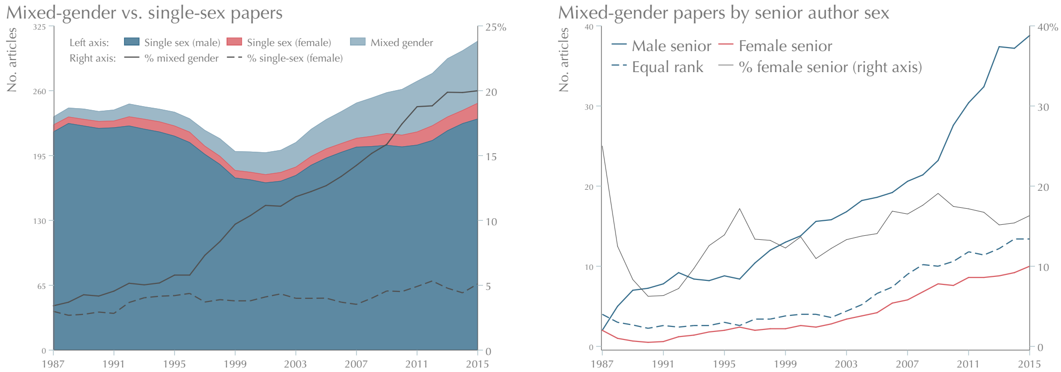

Figure 4: Mixed-gender papers (II)

Notes. Left-hand figure plots stacked the total number of 100 percent male-authored (dark blue), 100 percent female-authored (pink) and co-authored papers (light blue) published between 1986–2015 in a top-five journal; percentage co-authored and 100 percent female-authored shown with the grey and dashed lines, respectively. Right-hand figure displays the total number of mixed-gender papers with a male (blue line) and female (pink line) senior author. Papers with two or more senior authors of the opposite gender are represented by the dashed blue line. The percentage of mixed-gender papers with a senior female author shown in grey (right axis). All graphs shown as five year moving averages.

percentage of mixed-gender papers closely tracks the rising share of female authors from Figure 1.

Furthermore, the fraction of mixed-gender papers with senior female authors or authors of equal rank have not increased with time. Figure 4's right-hand graph displays the numbers of mixed gender articles with male (blue) and female (pink) senior authors. Papers with male senior authors have steadily increased since the late 1980s. Growth in papers with a senior female author (pink line) or male and female co-authors of equal rank (dashed blue line) is more modest. Senior female-authored papers make up about 14–16 percent of all mixed-gender papers; this figure remained relatively constant over the past three decades (grey line).

3.2 Male superstardom

> "In certain kinds of economic activity there is concentration of output among a few individuals, marked skewness in the associated distributions of income and very large rewards at the top." (Rosen, 1981, p. 845).

Citation counts may be highly skewed by a small number of superstar academics. Superstar academics probably merit more citations, but superstardom itself may also generate more citations conditional on an article's quality. This complicates gender analysis of citation counts if superstars in economics are also disproportionately male. In this section, I investigate the extent to which this statement is true

Seventy-five percent of articles published in top economics journals are cited by 115 or fewer papers; 50 percent are cited by 44 or fewer; 25 percent by 16 or fewer; 10 percent by (at most) six other papers. One percent are not cited at all. These figures mask a long right tail. Ten percent of articles in a top-four journal are cited by at least 250 other papers; five percent by more than 400; one percent are cited over 1,000 times.

Moreover, the distribution of papers in this latter category is itself rightward-skewed. The highest cited paper—Daniel Kahneman and Amos Tversky's paper "Prospect Theory: An Analysis of Decision under Risk" (*Econometrica*, 1979)—is cited 14,575 times. Halbert White's "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity" (*Econometrica*, 1980) is second highest (9,473 citations). In third is James Heckman's "Sample Selection Bias as a Specification Error" (*Econometrica*, 1979) with 8,513 citations. Conditional on having more than 1,000 cites, however, the median citation count is only 1,565.

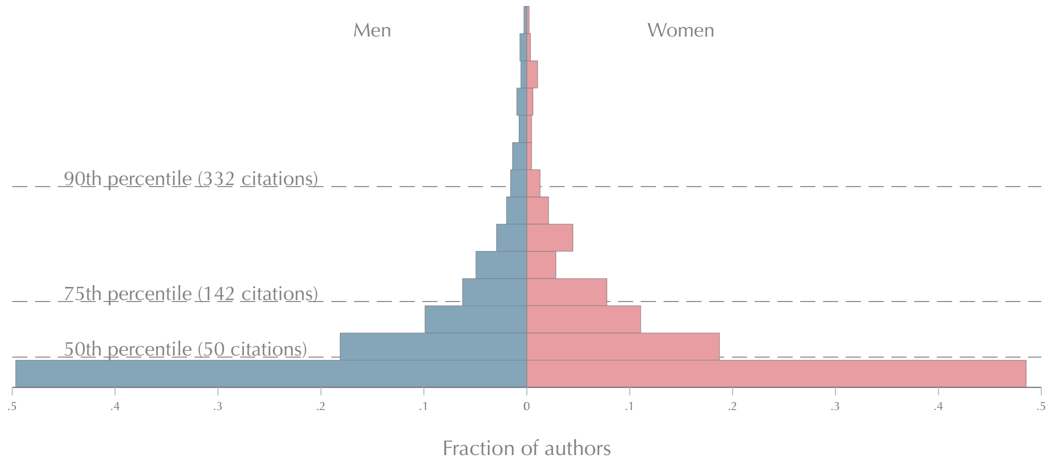

FIGURE 5: Histogram of authors' maximum lifetime citation counts

Notes. Graph displays the maximum citation count acrosss an author's lifetime number of top-five publications between 1950–2015 for all authors below the 95th percentile. Blue represents male authors; pink represents female authors.

Highly cited papers (top one percent) share two things with poorly cited papers (bottom one percent): they were written the 1980s, predominantly by men. Papers cited more than 1,000 times were published, on average, 32 years ago; even the youngest is thirteen years old.[7] Their average ratio of female authors is three percent. Similarly, papers that have *never* been cited are, on average, 36 years old. The average ratio of female authors is also three percent.

Conclusions are similar when data are analysed from the perspective of authors. To disaggregate the article-level data, I replicated each observation by its number of co-authors and uniquely assigned each individual co-author to exactly one duplicate observation. This created a panel dataset following authors' publication histories.

When evaluated by their highest cited paper, 59 authors—roughly one percent of all authors—are cited more than 2,000 times. Corresponding to the article-level analysis, this group is both relatively old and largely male. On average, their first top-four paper was published in 1982 (even the youngest in this group published a top four paper first in 2003). Ninety-five percent are male; almost a quarter won a Nobel prize.[8] Again, age and gender are factors the top one percent of cited authors share with the bottom one percent. Authors never cited at all are 96 percent male and were published, on average, in 1984.

When evaluated by their highest cited paper, male economists are cited more than female economists. By this measure, women are cited, on average 124 times; men 170 times. Yet the previous observation that senior male economists dominate the top and bottom one percent of article and author citation counts suggests that higher average citations for men is skewed rightward by a very small number of senior men and the fact that citation counts are censored from below at zero.

Indeed, when per-author citations are determined by their minimum, mean or median citation count, women are always cited at least as frequently as men both at the mean and median across all authors: on average, women's mean cited paper is cited 87 times, men's 84; women's median cited paper is cited 82 times, men's 71. When restricted to the lower 90th percentile of authors, women's highest cited papers are cited more than men's, both at the median and the mean.

[7] Lawrence Christiano, Martin Eichenbaum and Charles Evans's "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy" (*Journal of Political Economy*, 2005).

[8] Included in this count are Fischer Black and Amos Tversky, both of whom died before the the prize was awarded to co-authors for joint work

Figure 5 emphasises this point. It is a histogram of authors' median citation count for authors in the lower 95th percentile of median citation counts. Blue represents male authors; pink represents female authors. Across the vast majority of the distribution, women are cited at least as frequently as men.

4 Empirical analysis

Table 2 displays results from regressing citation counts (panel one) and the inverse hyperbolic sine (asinh) of citation counts (panels two and three) on a paper's share of female authors.[9] Panels one and two are estimated on article-level data using OLS; results in panel three are from a fixed effects regression on author-level data (male authors, only).

Column (1) is the baseline specification and controls for author seniority—*i.e.*, the number of top-five publications of the most prolific co-author at the time an article was published (max. t_5)—the number of co-authors (N) and journal fixed effects.

Additional controls and sample restrictions are progressively added in columns (2)–(8). First, publishing more papers increases one's prominence and prestige in the field, which in turn may influence citations to one's work. Column (3) accounts for this by controlling for max. T_5—*i.e.*, the maximum number of lifetime publications across all authors on a paper.

Columns (4)–(5) explicitly control for distortions introduced by a small number of disproportionately male "superstar" economists. Column (4) adds fixed effects for papers by the 91 authors listed in Appendix A.[10] Column (5) replaces them with fixed effects only for authors who win the Nobel prize.

Columns (6) and (7) repeat analyses from columns (4) and (5) (respectively) on the sample of articles published after 1990. I do this for three reasons. First, it generates a more accurate estimation of gender's contemporaneous effect. Second, the ratio of female authors per paper was very low before 1990. The combination of women's poor representation in economics before that and the positive yet non-linear relationship between the age of an article and citation count mean year fixed effects may not fully correct for selection issues related to the age of an article.[11] Third, *JEL* codes—which I control for in column (8)—were significantly revised in 1990; comparable codes are not available for periods pre- and post-reform.

4.1 Article-level analysis

Raw citation counts. As discussed in Section 2, the gender citation gap may be biased by the following two issues: (i) older articles have had more time to accrue citations and are disproportionately male; (ii) mean citation counts may be distorted by a small number of superstar economists who are *also* disproportionately male. The first panel of Table 2 regresses female ratio of raw citation counts in order to illustrate the impact of both.

According to to the baseline result (column (1)), papers by women are cited, on average, 28 fewer times than papers by men. This gap is substantially smaller, however, once year and journal-year fixed effects are added to account for selection issues related to the age of an article (column (2)). Column (3) includes controls for author prominence (max. T_5). This reverses the direction of the gender citation gap—female-authored papers are cited about two more times than papers by men who are equally prominent.

[9] See Hengel (2018) for a discussion and justification for using the female ratio to represent the "gender" of an article. All results are robust to the alternative proxies of article gender discussed in Hengel (2018, Appendix J) (available on request).

[10] See Section 2 for the criteria used to generate this list and further discussion.

[11] Specifically, older articles are more likely to have higher citation counts, but citation counts are increasing in year only for a very small number of articles generally authored by superstar men.

Table 2: Gender differences in citation counts

	1950–2015					1990–2015		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Article-level analysis (OLS)								
Dependant variable: no. citations								
Female ratio	-27.504^{***}	-5.780	1.962	9.386	6.095	14.643^{**}	10.730	10.005
	(6.958)	(6.700)	(6.631)	(6.368)	(6.532)	(7.279)	(7.430)	(8.036)
Max. t_5	1.530^{***}	1.929^{***}	-6.715^{***}	-6.648^{***}	-6.983^{***}	-8.045^{***}	-7.403^{***}	-8.035^{***}
	(0.390)	(0.345)	(1.092)	(1.045)	(1.045)	(1.115)	(1.164)	(1.045)
N	5.379	18.493^{***}	17.086^{***}	8.643^{**}	9.817^{***}	16.344^{***}	17.907^{***}	15.155^{***}
	(3.566)	(3.989)	(3.986)	(3.792)	(3.667)	(3.201)	(3.584)	(3.422)
Max. T_5			6.762^{***}	8.343^{***}	6.758^{***}	8.727^{***}	7.226^{***}	9.309^{***}
			(0.896)	(0.817)	(0.784)	(1.002)	(1.068)	(1.005)
Article-level analysis (OLS)								
Dependant variable: no. citations (asinh)								
Female ratio	0.077	0.163^{***}	0.223^{***}	0.274^{***}	0.234^{***}	0.269^{***}	0.244^{***}	0.143^{***}
	(0.059)	(0.051)	(0.050)	(0.050)	(0.050)	(0.052)	(0.052)	(0.054)
Max. t_5	0.027^{***}	0.020^{***}	-0.047^{***}	-0.046^{***}	-0.049^{***}	-0.051^{***}	-0.048^{***}	-0.049^{***}
	(0.003)	(0.002)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)
N	0.106^{***}	0.187^{***}	0.176^{***}	0.142^{***}	0.161^{***}	0.160^{***}	0.170^{***}	0.145^{***}
	(0.018)	(0.016)	(0.015)	(0.016)	(0.016)	(0.017)	(0.016)	(0.018)
Max. T_5			0.053^{***}	0.086^{***}	0.053^{***}	0.071^{***}	0.051^{***}	0.075^{***}
			(0.003)	(0.004)	(0.003)	(0.005)	(0.003)	(0.005)
Author-level analysis (fixed effects; male authors only)								
Dependant variable: no. citations (asinh)								
Female ratio	-0.300^{***}	0.208^{**}	0.247^{***}	0.266^{***}	0.241^{***}	0.244^{***}	0.231^{**}	0.219^{**}
	(0.109)	(0.087)	(0.088)	(0.089)	(0.088)	(0.090)	(0.090)	(0.096)
Max. t_5	-0.031^{***}	0.004	-0.012^{***}	-0.012^{***}	-0.013^{***}	-0.023^{***}	-0.020^{***}	-0.024^{***}
	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
N	0.061^{***}	0.165^{***}	0.148^{***}	0.128^{***}	0.144^{***}	0.137^{***}	0.148^{***}	0.145^{***}
	(0.020)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.018)
Max. T_5			0.021^{***}	0.037^{***}	0.021^{***}	0.043^{***}	0.025^{***}	0.048^{***}
			(0.003)	(0.006)	(0.003)	(0.007)	(0.004)	(0.007)
Journal effects	✓	✓	✓	✓	✓	✓	✓	✓
Year effects		✓	✓	✓	✓	✓	✓	✓
Journal×Year effects		✓	✓	✓	✓	✓	✓	✓
Superstar effects				✓		✓		✓
Nobel effects					✓		✓	
JEL (tertiary) effects								✓

Notes. 10,954 articles in panels one and two, columns (1)–(5); 6,287 articles in panels one and two, columns (6)–(8). 13,934 observations (3,025 distinct authors) in the third panel, columns (1)–(5); 9,037 observations (2,335 distinct authors) in the third panel, columns (6)–(8). Figures represent the coefficient on female ratio from an OLS (panels one and two) and fixed effects (panel three) regression. Dependent variables are raw citation count (panel one) and the inverse hyperbolic sine of it (panels two and three). Panel three results estimated on the sample of male authors, only. Robust standard errors in parentheses. ***, ** and * statistically significant at 1%, 5% and 10%, respectively.

The coefficient on female ratio jumps to nine and six once "superstar" and Nobel prize effects are included in columns (4) and (5), respectively. Neither figure is statistically significant at traditional levels when estimated on the entire sample. Both are larger, however, when estimated on the sample of articles published between 1990–2015 (columns (6) and (7)). The former is statistically significant; it suggests contemporary female-authored papers are cited about 15 more times than similar papers by men.

Column (8) adds tertiary *JEL* fixed effects to the equation estimated in column (6). The coefficient on female ratio is somewhat smaller; its standard error larger.[12] It suggests female-authored papers are cited about 10 more times than male-authored papers published on the same topic, although the difference is no longer statistically significant.

Inverse hyperbolic sine of citation counts. The second panel in Table 2 uses the inverse hyperbolic sine function of citations as the dependent variable. This functional transformation should reduce distributional skew from superstar economists and the age of an article, both of which may correlate with an article's share of female authors.

Results reinforce conclusions derived from the analysis using raw counts: without year controls, male- and female-authored papers are not statistically different from one another. After year controls are included, the gender citation gap favours women and is highly significant. Results in columns (3)–(7) suggest papers by women receive 22–27 log points more citations than than similar papers by men.[13] The coefficient on female ratio falls about six log points after including tertiary *JEL* effects but remains highly significant.

Including controls for author prominence and superstar or Nobel fixed effects hardly alters the coefficients and standard errors on female ratio. This suggests that selection effects related to the age of an article, superstar economists and the fact that both correlate with author gender are adequately handled by combining the asinh functional form transformation with journal-year fixed effects.

Coefficients on max. t_5, N and max. T_5 are roughly equivalent in size and magnitude to those from panel one. Papers with more co-authors (N) and by more prominent authors (max. T_5) receive more citations. After controlling for the latter, citations and author seniority (max t_5) are negatively related—*i.e.*, authors' earlier papers are cited more than their later papers, conditional on prominence.

4.2 Author-level analysis (male authors only)

I next analyse citations at the author-level in order to tease out the marginal impact of co-authoring with more women for male authors, only. To disaggregate the data, each article is duplicated N_j times, where N_j is article j's number of co-authors; observation $j_k \in \{1, \ldots, N_j\}$ is assigned article j's kth author. This creates a panel data set following author i over the $t \in \{1, \ldots, T_i\}$ articles he publishes in a top-five journal, where $t = 1$ for his first, $t = 2$ his second, *etc.* and T_i is the total number of top-five articles he has published in the data. I then estimate Equation (1) within a fixed effects framework:

$$\text{No. citations (asinh)}_{it} = \alpha_i + \beta_1 \text{ female ratio}_{it} + \boldsymbol{\theta}\, \mathbf{X}_{it} + \varepsilon_{it}, \tag{1}$$

where $\mathbf{X}_{it}$ is a vector of observable controls, (max. t_5, N, max. T_5 and fixed effects for journal, year, journal and year interactions, superstars, Nobel prize winners and tertiary *JEL* categories),

[12] Results and conclusions are very similar if Nobel fixed effects or both Nobel and superstar fixed effects are used instead or if tertiary *JEL* codes are replaced with primary or secondary codes. (Results not shown but available on request.)

[13] $\text{asinh}(x)$ closely mirrors the slope of $\log(x)$ for $x > 1$. Hence, the coefficient on female ratio represents log point effects. See Card and DellaVigna (2017) for further discussion.

α_i are author-specific fixed effects and ε_{it} is an idiosyncratic error. I eliminate α_i by demeaning the other variables.

β_1 represents the marginal change in citations male economists would receive if they went from co-authoring with no women to co-authoring a paper exclusively by women.[14] According to the baseline regression, men are cited less when they co-author with more women. After controlling for year fixed effects, however, the coefficient reverses sign and is highly significant: men are cited *more* when they co-author with more women.

Coefficients on all co-variates are roughly equivalent across columns (3)–(8). They are also very similar to their corresponding marginal effects from panel two.

4.3 Experience

Table 2 suggests that, after accounting for the year in which an article was published and including fixed effects for a very small number of male superstar economists, female-authored papers are cited more than male-authored papers. Moreover, the author-level analysis suggests male authors are cited more when they co-author with more women.

Explanations as to *why* can be broadly categorised into two mutually exclusive groups: (i) factors outside authors' control; (ii) factors within authors' control. In Hengel (2018), I develop a model to distinguish between (i) and (ii). Although highly specific to a different measure of a paper's quality—it's readability—the framework's general intuition is applicable to any situation where people are repeatedly judged on and respond to feedback about some quantifiable component of their output.

The logic applied to citations is as follows. Assume preferences are fixed over time. Then authors increase their citation counts today relative to yesterday only if they believe doing so reduces rejection rates. Although a variety of factors, such as author sensitivity and/or poor information, may lead to errors in initial beliefs, authors correct such mistakes with experience. Hence, citations that decline with experience mean errors in beliefs are behind initial counts. If they remain constant, then preferences probably are. But authors probably only *increase* citation counts if they actually reduce rejection rates. Moreover, if citation counts from female authors in the latter category exceed those of equivalent male authors *despite* identical acceptance rates, then higher citations in female-authored papers are due to circumstances beyond women's control—circumstances male authors do not endure.

Based on its logic, higher citations in female-authored papers are the result of external factors beyond their control when the following three conditions are satisfied:[15]

1. Experienced women's papers are cited more than papers by equivalent men.

2. Women's citations increase with experience.

3. Female-authored papers are accepted no more often than equivalent male-authored papers.

Causal identification requires that each condition holds for the same author in two different situations—before and after gaining experience and when compared to an equivalent, experienced author of the opposite gender. Hengel (2018) uses a matching identification strategy to account

[14] This interpretation is slightly awkward since a man's own contribution to a paper he co-authors means its share of female authors can never be one. Because the regression model implicitly assumes a linear relationship between asinh of citations and female ratio, however, the marginal change between co-authoring with no women and co-authoring with a share of x women, $x \in (0, 1)$ is $\beta_1 \times x$. Moreover, $\beta_1 \times x$ roughly corresponds to the coefficients obtained if female ratio is replaced with dummy variables equal to 0 for exclusively male-authored papers and 1 for x percent female-authored papers. (Regression results available on request.)

[15] See Hengel (2018) for a more explicit and complete model.

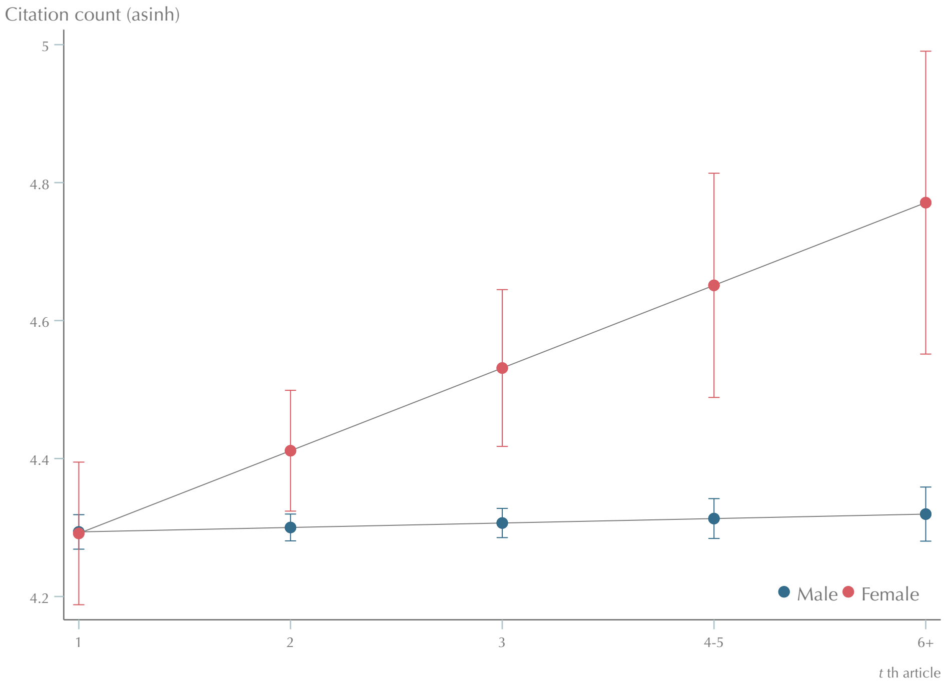

Figure 6: Citations for authors' tth publication

Notes. Estimated mean citation counts for authors' first, second, third, 4th–5th and sixth and up publications in the data. Pink represents women co-authoring only with other women; blue are men co-authoring only with other men. Vertical lines represent 90 percent confidence intervals from standard errors clustered on author.

for this. Here, I conduct a more limited investigation. It focuses on testing whether conditions 1 and 2 hold, on average.[16] To conduct this analysis, I estimated Equation (2) using a random effects framework:

$$\begin{aligned} \text{No. citations (asinh)}_{it} = {} & \alpha_i + \beta_1 \,\text{female ratio}_{it} + \beta_2 \,\text{female ratio}_{it} \times \text{male}_i \\ & + \beta_3 \, t + \beta_4 \, t \times \text{male}_i + \beta_5 \,\text{female ratio}_{it} \times t \\ & + \beta_6 \,\text{female ratio}_{it} \times t \times \text{male}_i + \boldsymbol{\theta}\, \mathbf{X}_{it} + \varepsilon_{it}, \end{aligned} \tag{2}$$

where t represents the number of articles an author has previously published in a top-five journal—*i.e.*, $t = 1$ for an author's first publication, $t = 2$ for his second, *etc.*— female ratio$_{it}$ is the share of female authors on author i's tth paper and male$_i$ is his gender (1 for male; 0 for female).[17] $\mathbf{X}_{it}$ is a vector of observable controls which includes author seniority, year and journal fixed effects as well as their interaction effects and fixed effects for superstar economists. α_i are author-specific effects and ε_{it} is an idiosyncratic error. α_i are assumed to be uncorrelated with other independent variables.

Figure 6 displays the marginal effect of t for men and women. Male effects estimated for male authors co-authoring with no females ($\beta_3 + \beta_4$); female effects for female authors co-authoring with no males ($\beta_3 + \beta_5$).

Citations remain relatively constant as men publish more papers—all are very close to the mean effect and none are statistically different from it (0.005, standard error 0.008). For a woman,

[16] The conclusions derived are non-causal. Nevertheless, they provide correlation evidence of whether authors' behaviour is consistent with internal factors or external constraints.

[17] Given the small number of observations above $t = 3$—particularly for female authors—publications four and five are grouped together, as are publications 6 and up.

however, every additional paper is cited more than her last. The marginal change in citations with respect to t is about 0.11 log points (standard error 0.04).

Figure 6 provides suggestive evidence that conditions (1) and (2) hold. Assuming female authored papers are accepted no more often than male-authored papers (condition (3))—and extensive study elsewhere suggest they do not (see, *e.g.*, Ceci et al., 2014)—then this suggests that external factors are behind female authors' higher citation rates.

5 Conclusion

The ratio of female authors on a paper is negatively correlated with citation rates. As I show, however, this effect is largely driven by a very small set of highly cited papers written almost exclusively by senior male economists—many of whom are Nobel prize winners. Once these individuals and publication year are properly controlled for, female-authored papers are cited more. A rough application of the logical framework in Hengel (2018) provides suggestive (non-causal) evidence that factors beyond women's control are at least partially responsible for women's higher citations during this period.

The evidence in this paper is correlational. It does, however, suggest that the median female economist must produce more provocative, exciting and *citable* research to publish in top economics journals. Higher standards applied to any dimension impose a quantity/quality tradeoff. Thus, women's higher quality papers probably result in fewer papers. Failing to account for this could reinforce lower promotion rates and wages among female economists.

References

Bransch, F. and M. Kvasnicka (2017). "Male Gatekeepers Gender Bias in the Publishing Process?" IZA Discussion Paper Series, No. 11089.

Card, D. and S. DellaVigna (2017). "What do Editors Maximize? Evidence from Four Leading Economics Journals". NBER Working Paper Series, No. 23282.

Card, D. et al. (2018). "Are Referees and Editors in Economics Gender Neutral?" Mimeo.

Ceci, S. J. et al. (2014). "Women in Academic Science: A Changing Landscape". *Psychological Science in the Public Interest* 15 (3), pp. 75–141.

Ferber, M. A. (1988). "Citations and Networking". *Gender and Society* 2 (1), pp. 82–89.

Ginther, D. K. and S. Kahn (2004). "Women in Economics: Moving Up or Falling Off the Academic Career Ladder?" *Journal of Economic Perspectives* 18 (3), pp. 193–214.

Grossbard, S., T. Yilmazer, and L. Zhang (2018). "The Gender Gap in Citations of Economics Articles: Lessons from Economics of the Household". Mimeo.

Hamermesh, D. S. (2018). *Citations in Economics: Measurement, Uses and Impacts*. Forthcoming 1. Cambridge, MA: National Bureau of Economic Research. arXiv: arXiv:1011.1669v3.

Hengel, E. (2018). "Publishing while Female: Are Women Held to Higher Standards? Evidence From Peer Review." Mimeo.

Laband, D. N. (1987). "A Qualitative Test of Journal Discrimination against Women". *Eastern Economic Journal* 13 (2), pp. 149–153.

Rosen, S. (1981). "The Economics of Superstars". *American Economic Review* 71 (5), pp. 845–858.

Smart, S. and J. Waldfogel (1996). *A Citation-Based Test for Discrimination at Economics and Finance Journals*. Tech. rep. Cambridge, MA: National Bureau of Economic Research, pp. 197–205.

Web of Science (2018). *Web of Science: Summary of Coverage*.

Appendices

A Male superstars

Table A.1: List of superstar economists

Author name	T_5	Max. cites	Nobel	JBC
Abel, Andrew B.	17	255		
Acemoglu, Daron	54	2,481		2005
Aghion, Philippe	23	1,826		
Alesina, Alberto	26	985		
Andrews, Donald W. K.	32	1,419		
Arellano, Manuel	7	7,407		
Banerjee, Abhijit V.	18	1,445		
Barro, Robert J.	26	2,852		
Becker, Gary S.	23	3,719	1992	1967
Bénabou, Roland	18	629		
Bernheim, B. Douglas	26	428		
Besley, Timothy J.	23	514		
Blackorby, Charles	18	124		
Blanchard, Olivier J.	17	1,155		
Blundell, Richard W.	25	681		
Bolton, Patrick	21	569		
Browning, Martin J.	19	300		
Caballero, Ricardo J.	20	241		
Campbell, John Y.	19	1,052		
Caplin, Andrew S.	18	216		
Card, David E.	25	583		1995
Chiappori, Pierre-André	21	430		
Cooper, Russell	17	553		
Crawford, Vincent P.	24	1,111		
Deaton, Angus S.	19	504	2015	
Diamond, Peter A.	20	549	2010	
Dixit, Avinash K.	18	705		
Engle, Robert F.	12	7,972	2003	
Epstein, Larry G.	26	958		
Farhi, Emmanuel	17	228		
Fehr, Ernst	17	2,840		
Feldstein, Martin S.	26	240		1977
Fisher, Franklin M.	21	97		1973
Fudenberg, Drew	28	908		
Gale, Douglas	19	706		
Granger, Clive W. J.	5	7,972	2003	
Green, Jerry R.	17	335		
Grossman, Gene M.	33	1,344		
Grossman, Sanford J.	13	2,546		1987
Gruber, Jonathan	17	349		
Gul, Faruk	17	321		
Hamilton, James D.	11	2,591		

Table A.1 (continued)

Author name	T_5	Max. cites	Nobel	JBC
Hansen, Lars Peter	13	3,791	2013	
Hart, Oliver D.	26	2,546	2016	
Hausman, Jerry A.	14	4,808		1985
Heckman, James J.	32	8,513	2000	1983
Helpman, Elhanan	37	1,239		
Jackson, Matthew O.	26	262		
Jovanovic, Boyan	30	1,316		
Kahneman, Daniel	7	14,575	2002	
Kehoe, Patrick J.	20	484		
Kremer, Michael	21	488		
Krugman, Paul R.	11	2,599	2008	1991
Laffont, Jean-Jacques	32	456		
Laroque, Guy	21	413		
Levine, David K.	19	296		
Levitt, Steven D.	19	341		2003
List, John A.	23	351		
Mankiw, N. Gregory	16	2,666		
Maskin, Eric S.	30	908	2007	
Milgrom, Paul R.	23	1,352		
Murphy, Kevin M.	18	1,245		1997
Newey, Whitney K.	23	1,026		
Pakes, Ariel	18	1,086		
Palfrey, Thomas R.	19	284		
Persson, Torsten	18	864		
Phillips, Peter C. B.	42	1,189		
Plott, Charles R.	20	254		
Postlewaite, Andrew	20	238		
Ray, Debraj	23	402		
Robinson, Peter M.	17	769		
Romer, David H.	16	2,666		
Rosen, Sherwin	11	3,133		
Rosenzweig, Mark R.	28	400		
Roth, Alvin E.	27	551	2012	
Rubinstein, Ariel	20	1,915		
Saez, Emmanuel	19	673		2009
Samuelson, Larry	20	578		
Sargent, Thomas J.	23	739	2011	
Scheinkman, José Alexandre	17	1,450		
Shleifer, Andrei	33	4,556		1999
Stein, Jeremy C.	18	810		
Stiglitz, Joseph E.	29	472	2001	1979
Tirole, Jean	59	791	2014	
Tversky, Amos	8	14,575		
Vishny, Robert Ward	11	4,556		
Weil, David N.	11	2,666		
Weitzman, Martin L.	19	326		
White, Halbert	16	9,473		
Wolpin, Kenneth I.	20	323		
Wright, Randall	19	403		
Zame, William R.	17	104		

Notes. Table lists the names, number of top-five publicatins, number of citations for their highest cited paper and the year in which they won the Nobel prize for economists.