Is the Bar Higher for Female Scholars? Evidence from Career Steps in Economics*

Niels Johannesen and Simon Muchardt

March 4, 2024

Abstract

Do gender disparities in academia reflect that female scholars are held to higher standards than males? We address this question by comparing the scientific merit of male and female academic economists who make the same career step. Across four domains – i.e. faculty positions, network affiliations, research grants and editor appointments – we find no evidence that standards are higher for females. By contrast, the average female has less citations and publications than the average male who makes the same career step. In most domains, this reflects a gender gap for "marginal" scholars, consistent with lower merit thresholds for females.

JEL codes: A11, I23, J16, J44, J62

Keywords: Gender differences, Discrimination, Unequal treatment, Gender gap, Academic labor markets

*Niels Johannesen is with the Centre for Business Taxation at Saïd Business School, Oxford University and with the Center for Economic Behavior and Inequality at the University of Copenhagen. Simon Muchardt is with the Université Paris-Dauphine and the University of Copenhagen. We are grateful for insightful comments from seminar participants at the University of Copenhagen. The activities of Center for Economics Behavior and Inequality are financed by the Danish National Research Foundation, Grant DNRF134. Declarations of interest: none. Correspondence: Niels Johannesen, Centre for Business Taxation, Saïd Business School, University of Oxford, Park End Street, Oxford OX1 1HP, United Kingdom.

1 Introduction

Across scientific fields, males tend to occupy most of the top positions in academia (e.g. European Commission, 2023). While females are more likely to complete university education, a "leaky pipeline" into doctoral programs and at each rung of the academic career ladder implies that professors are predominantly male. This pattern mirrors gender asymmetries in other parts of society. For instance, males continue to occupy the vast majority of the leading positions in business (e.g. Grant Thornton, 2022), politics (e.g. United Nations, 2023) and public administration (e.g. UNDP, 2021) despite long-run increases in female participation in professional and political life.

A potential explanation for the low share of women in top academic positions is that female scholars are held to higher standards than male scholars. Such unequal treatment would resonate with broader patterns of discrimination against women in the labor market (e.g. Neumark et al., 1996; Goldin and Rouse, 2000; Sin et al., 2022). It would also carry a somber message about the academic labor market being *unfair*, by putting female scholars at a disadvantage, and *inefficient* by misallocating female scholarly talent.

In this paper, we study whether female scholars in economics need more scientific merit than males to be selected for the same career steps.[1] Motivated by a simple conceptual framework, we develop an empirical methodology that draws on publicly available data and can be applied to different domains of academic careers. The basic premise is that scholars are selected for career steps based on their scientific merit but that selection committees may apply different threshold levels of scientific merit to males and females because of idiosyncratic gender preferences. Based on this premise, we can learn about systematic gender differences in the unobserved thresholds by comparing the scientific merit of males and females who successfully make the same career step. Intuitively, if merit thresholds are higher for females than for males, females who make a given career step would tend to have more scientific merit than the males who make the same career step.

The empirical analysis proceeds in two steps. First, we identify a large set of instances where male and female economists made precisely the same career step at the same time. Concretely, they were appointed to the same *faculty position* in the same economics department, they received the same type of research *grant* from the same research council, they became affiliates of the same academic *network* within the same program area or they were appointed *editor* at

[1] Recent work documents the low share of women in top positions in economics (Auriol et al., 2022) and finance (Sherman and Tookes, 2021) and lay out historical trends (Lundberg and Stearns, 2019).

the same journal. Second, we compare two measures of scientific merit across male and female economists who successfully made the same career step at the time: the number of journal *publications* adjusted for journal quality and the number of *citations* that their academic work has attracted.

Across all four domains and both measures of scientific merit, we find that female scholars, on average, have less scientific merit than male scholars at the time they make the same career step. For faculty positions, the estimated difference is around 0.1 log-points, which corresponds to a gap of around 10%. The difference widens to around 0.2 log-points for editor positions, 0.3-0.4 log-points for network affiliations and 0.4-0.6 log-points for research grants. These patterns are robust to alternative merit measures that discount co-authored work, control for academic age, and allow for forward-looking selection committees by including publications and citations three years after the career step. They are also qualitatively similar across subsamples when we separately consider different faculty positions (assistant, associate and full professor), different networks (NBER and CEPR), and grants from different research councils (France, United States, Germany and United Kingdom).

We complement these comparisons of means with richer distributional comparisons. Concretely, for each scholar making a career step, we define their "relative merit" as merit relative to the average of other scholars making the same career step at the same time. We then compare the male and female distributions of relative merit and measure the gender gap at different quantiles. Intuitively, comparing male and female scholars at the bottom of these distributions – "marginal" scholars whose merit is lower than others making the same career step at the same time – may be particularly informative about systematic gender differences in merit thresholds.

The distributional analysis adds important nuance to the main results. On the one hand, for faculty positions, the relative merit of males and females are highly similar through most of the distribution and the gender difference at the mean identified in the main analysis reflects a modest gap at the top. The fact that there is no gender difference at the bottom suggests that merit thresholds may be similar for males and females. On the other hand, for network affiliations and research grants, females have significantly less relative merit than males at the lowest quantiles and the gender gap is almost constant through the distribution. The salient gender difference for marginal scholars suggests that merit thresholds may be lower for females than for males.

In sum, we find no evidence that female scholars are held to higher scientific standards than male scholars in economics. In one domain, faculty appointments, our results suggest that the bar in terms of scientific merit is the same for males and females. In other domains such as network affiliations and research grants, the bar appears to be lower for females than for males. These conclusions come with two important caveats that reflect limitations inherent to the research design. First, selection for career steps generally reflect other qualities than scientific merit, e.g. skills in teaching in the context of faculty appointments. If female scholars possess these qualities to a larger extent than male scholars, we cannot rule out that the bar in terms of overall merit, including not just the quality of scientific work but also all other dimensions of merit, is higher for females. Second, our measures of scientific merit may themselves be gendered. If the same paper is harder to publish and less likely to be cited when the author is female, publication and citation scores do not adequately capture scientific merit (Card et al., 2020; Hengel, 2022). In this case, the estimated gender gaps in merit thresholds may reflect that selection is based on true scientific merit rather than biased publication and citation scores.

Our paper makes a methodological contribution to the literature that studies unequal treatment of men and women in academia and in labor markets more broadly. The most related papers focus on specific career steps, i.e. promotion to tenure (Sarsons et al., 2021) and selection of Fellows of the Econometric Society (Card et al, 2022) and Members of the National Academy of Science (Card et al, 2023). These studies compile data on both successful and unsuccessful candidates and estimate how the probability of success depends on gender conditional on scientific merit. A key disadvantage of this approach is that information about the set of unsuccessful candidates is often unavailable or is very costly to access.[2] While economics departments, research councils and scientific journals publish who is on the faculty, who received a grant and who is the editor, they generally do not reveal the names of candidates who were considered but not selected. By contrast, our approach relies only on information about successful candidates, which increases the feasible sample sizes and widens the range of possible applications considerably. Our approach is conceptually similar to recent studies of unequal treatment in academic publishing that make inference about gender differences in quality thresholds by comparing the ex post citations (Card et al, 2019) and readability scores (Hengel, 2022) of articles by female and male authors conditional on the (positive) outcome of the editorial process. We refine this approach by suggesting a distributional analysis that strengthens the case for identifying gender differences in the thresholds.

Beyond the methodological contribution, our paper also conveys new substantial insights about unequal treatment in economics, a mechanism often invoked to explain the low represen-

[2] Sarsons et al. (2021) is a good illustration of the tediousness of identifying unsuccessful candidates, in their setting scholars who were denied tenure.

tation of females in the profession. Our findings generally resonate with Card et al. (2022, 2023) who document that the bar for being admitted to the most prestigious societies in the sciences is currently lower for females than for males; however, the scope of our analysis is broader in two dimensions. First, we consider a range of outcomes that go beyond peer recognition and involve high stakes for the scholars involved, i.e. job promotions, research grants and so on. Second, our analysis covers a relatively broad group of economists at different levels of seniority and productivity and not just the extreme right tail of the merit distribution.

By suggesting that the low female representation in top positions in the economics profession does not simply reflect that female scholars are held to higher scientific standards than males when considered for career steps, our results suggest that more subtle mechanisms are at play. Some of these mechanisms have been investigated in the literature, e.g. gender stereotypes as reflected in comments on online platforms (Wu, 2018, 2020) and in reference letters (Eberhardt, Facchini and Rueda, 2023); gender differences in access to co-author networks (Ductor, Goyal and Prummer, 2023), in recognition for group work (Sarsons et al., 2021) and in the propensity to apply for promotions (Bosquet, Combes and García-Peñalosa, 2019); as well as gender biases in teaching evaluations (Boring, 2017; Mengel, Sauermann and Zölitz, 2019).

The paper proceeds by describing the data in Section 2, discussing the empirical design in Section 3, reporting the results in Section 4 and concluding in Section 5.

2 Data

We collect data for the analysis in two steps. First, we identify instances where economics scholars make career steps and use a standard algorithm to determine their gender. Second, we collect information about the publications and citations of these scholars at the time of the career step and use it to create measures of scientific merit. The resulting dataset serves to compare the scientific merit of male and female economists who make the same career step at the same time.

2.1 Career steps

We consider career steps within four distinct domains: (i) becoming an affiliate of a selective *network* for academic economists; (ii) receiving a scientific *grant* from a research council; (iii) being appointed *editor* of an economics journal; and (iv) being appointed to a *faculty position* at an economics department. We briefly discuss the data collection for each of the domains in turn. More details are available in the Online Appendix.

Networks

We consider affiliations of two highly regarded networks for academic economists: the National Bureau of Economic Research (NBER) in the United States and the Center for Economic Policy Research (CEPR) in Europe. The two networks are similar in selecting new affiliates in a competitive process that draws on nominations from existing members and in organizing activities within program areas such as labor studies, public economics and corporate finance.

On their websites, both networks maintain lists of current network affiliates and their program area affiliations. To obtain information about appointment years, which is key for our empirical design, we scrape archived versions of the network websites.[3] If one website version indicates that a scholar is a network affiliate and another version from around one year before indicates that the same scholar is not an affiliate, we infer that the scholar was appointed in the course of the year. We consider that scholars make the same career step if they become affiliates of the same network in the same program area in the same year. We identify 2,197 scholars making 629 distinct career steps in this domain over the period 2001-2022 (Table 1, Columns 1-3).

Grants

We consider grants from national research councils in the United States, United Kingdom, France and Germany. We focus on research grants where the recipient is an individual, rather than a network or a group, and where the key criterion is academic excellence. All of the research councils make lists of past grant awards available on their websites.

We scrape the websites to obtain, for each individual grant, the name of the principal investigator, the year of the award, and the type of grant. We consider that scholars make the same career step if they are awarded the same grant from the same research council in the same year. We identify 4,186 scholars making 271 distinct career steps in this domain over the period 1994-2022 (Table 1, Columns 1-3).

Editorships

We consider editor appointments at the 100 leading economics journals according to IDEAS/RePEc (2023).[4] We retain journals specializing in finance and econometrics, but disregard interdisciplinary journals and journals from adjacent fields where editors may publish primarily outside of economics journals.

[3] A similar approach has recently been adopted by Heckman and Moktan (2020) to collect data on job histories of academic economists.

[4] Among several alternatives, we consistently use the ranking of journals by their h-index. See discussion below.

Economics journals generally do not publish lists of past editors; however, they typically print the names of the current editors in the front matter of each issue. We hand-collect a dataset with the names of the editors at the top-100 journals. If one issue indicates that a scholar is an editor while another issue from around one year before indicates that the same scholar is not an editor, we infer that the scholar was appointed in the course of year. We consider that scholars make the same career step if they are appointed editors of the same journal in the same year. We identify 1,028 scholars making 496 distinct career steps in this domain over the period 2004-2022 (Table 1, Columns 1-3).

Faculty positions

We consider appointments to faculty positions at the 100 leading economics departments according to IDEAS/RePEc (2023).

Economics departments generally maintain lists of current faculty members on their websites including information about titles. We obtain information about appointment years by scraping archived versions of the websites. If one website version indicates that a scholar holds a given faculty position and another version from around one year before indicates that the same scholar holds another position or no position at all, we infer that the scholar was appointed in the course of year. We consider that scholars make the same career step if they are appointed to a faculty position with the same title in the same economics department in the same year. We identify 2,333 scholars making 1,206 distinct career steps in this domain over the period 2012-2022 (Table 1, Columns 1-3).

2.2 Scientific merit

For each of the scholars that make a career step in one of the four domains, we obtain detailed information about publications and citations by scraping Google Scholar. We use this information to construct two measures of the scholars' scientific merit at the time they make the career step: the cumulative number of citations that their work has attracted and the number of articles they have published in academic journals adjusted for the quality of the journals.[5]

To adjust the number of publications for journal quality, we express publications in AER-equivalents (Conroy et al., 1995). Concretely, we adopt the ranking of economics journals by their h-index from IDEAS/RePEc (2023) and assign to each journal in the top-500 an AER-equivalent, which is the ratio of the journal's own score and the score of the *American Economic Review*. With this procedure, the AER-equivalent of a paper published in *Econometrica* is

[5]Details are available in the Online Appendix

around 0.8, a paper in *Journal of Econometrics* around 0.5, a paper in *Journal of Human Resources* around 0.3, a paper in *Journal of Comparative Economics* around 0.2 and a paper in *Journal of International Development* around 0.1.[6]

The publication measure is not meaningful for non-economists whose publication outlets will typically not be among the top-500 economics journals. In practice, this does not represent a major challenge since most of the career steps we consider, e.g. appointments in economics departments and membership of economics networks, naturally limit the sample to academics who mainly publish in economics journals. Nevertheless, we address any remaining concern by excluding scholars from the estimation sample whose publication record indicate that they are not economists.[7]

Finally, we compute the two measures of scientific merit for all of the 50,000 economists registered at IDEAS/RePEc. This sample serves as a reference population of academic economists, against which we can compare the economists in our estimation datasets.

2.3 Descriptive analysis

Figure 1 provides context for the analysis by documenting key patterns in the data we have collected. Panel A shows that the female share among those who make a career step has exhibited a clearly increasing trend over the sample period in all four domains.[8] Panel B shows that, although citation and publication scores capture distinct dimensions of scientific merit, they are highly correlated and exhibit a relation that is strikingly close to log-linear over the entire distribution. Appealing to this property, we collapse the two outcomes into one in the final part of the analysis. Panels C-D show that our analysis primarily concerns economists from the upper half of the distribution of citation and publication scores. Intuitively, economists appointed to editor positions (green line) almost all come from the very top of the distribution, reflecting that editors are recruited from the ranks of senior academics with an excellent publication record. Economists appointed to faculty positions (brown line) and receiving research grants (red line) also generally have high citation and publication counts,

[6]We choose the ranking based on h-indexes because it produces a ranking that is consistent with the profession's strong priors about the top-5 journals in economics (i.e. *American Economic Review*, *Quarterly Journal of Economics*, *Journal of Political Economy*, *Econometrica* and *Review of Economic Studies*) and the top-3 journals in finance (i.e. *Journal of Finance*, *Journal of Financial Economics* and *Review of Financial Studies*). Indeed, the 7 highest ranked journals by this approach are the top-5 economics journals and 2 of the top-3 finance journals.

[7]Concretely, we exclude scholars if none of their five most cited papers is published in an outlet with a title containing words such economics, finance, econometrics. More details are available in the Online Appendix.

[8]A steep increase in the female share has been documented for related outcomes, e.g. publishing in top economics journals and election to the Econometric Society (Card et al., 2022).

but the distributions are much more dispersed, reflecting that both career steps occur at many levels of seniority, e.g. assistant vs full professorships and starting vs advanced grants. Finally, economists given network affiliations (blue line) exhibit a fairly tight distribution of citation and publication scores with a mode visibly below that of editors, reflecting that they tend to be accomplished, but relatively early-career economists.

3 Empirical design

3.1 Conceptual framework

The aim of the empirical analysis is to test if there is unequal treatment in the economics profession in the sense that female scholars need more (or less) scientific merit to make a career step than male scholars. We motivate the empirical specifications with a simple conceptual framework

Consider the selection of economists for a career advance among a larger set of candidates. Concretely, the advance could be an appointment to a faculty position or an editorial position, a research grant or an invitation to join a professional network. Each candidate has a level of scientific merit y. The selection committee has a preference for candidates with more merit, but may also have a gender preference. Hence, it chooses two merit thresholds, α_F and α_M, and selects female candidates with $y \geq \alpha_F$ and male candidates with $y \geq \alpha_M$. If $\alpha_M < \alpha_F$, we say that the bar is higher for females or that females are held to higher standards.

Empirically, we only observe the set of successful candidates and their merit. As we are unable to identify the unsuccessful candidates, we cannot directly estimate the effect of gender on success conditional on merit.[9] Given this constraint, we take two alternative approaches. First, we compare the mean level of merit $\overline{y}$ across the successful male and female candidates. If merit is distributed in the same way for males and females in the choice set, a higher mean for females implies a higher bar for females, i.e. $\overline{y}_M < \overline{y}_F \iff \alpha_M < \alpha_F$. Second, we compare the gender-specific distributions of merit for successful candidates. Intuitively, the candidates at the bottom of these distributions, for practical purposes say the 5th, 10th or 25th percentile, can be interpreted as "marginal" and their merit levels approximately identify the gender-specific merit thresholds.

Selection committees may consider other qualities than scientific merit. Concretely, this could be teaching skills for faculty appointments and project quality for research grants. Letting

[9] We cannot estimate $success_i = \alpha + \beta gender_i + f(merit_i) + \epsilon_i$ because we can only identify successful candidates and not the entire choice set.

ϵ denote such other qualities, selection would be based on $y + \epsilon$ rather than y. As ϵ is generally unobserved, we need the additional assumption that ϵ is uncorrelated with gender to make inference about unequal treatment.[10]

3.2 Empirical implementation

We implement this conceptual framework with the data on career steps, merit and gender described above. The main difference relative to the conceptual framework is that we want to pool observations from many different career steps in one regression while ensuring that the estimates are always identified from comparisons within sets of candidates who successfully made the same career step at the same time.

Denoting individuals by i and distinct career steps by c, we thus estimate the following equation separately for each of the four domains (i.e. faculty positions, network affiliations, research grants and editor positions):

$$y_{i,c} = \alpha_c + \beta^d \text{female}_{i,c} + \mu_{i,c} \tag{1}$$

where y is a measure of scientific merit and α_c represents a separate intercept for each career step c. The estimated β expresses the average merit of female economists who make a career step in domain d relative to the average merit of male economists who make the same career step in the same year.

It is useful to note that only when the set of economists making a career step includes at least one male and one female does the career step contribute to the identification of β. Table 1 provides information about the number of career steps and scholars contributing to identification within each domain (Columns 4-5).

To make comparisons at different positions of the merit distribution, we first estimate the following regression:

$$y_{i,c} = \alpha_c + \eta_{i,c} \tag{2}$$

and compute the residual $\overline{\eta}_{i,c}$ for each economist. The residual is a measure of "relative merit", i.e. an economist's merit relative to the average for economists who make the same career step. For instance, $\overline{\eta}_{i,c}$=0.1 means that the merit of economist i is 0.1 log-points higher (roughly 10%) than the average economist who made career step c. This metric is directly comparable across career steps. Hence, we construct the cumulative distribution of $\overline{\eta}_{i,c}$ for males and

[10] This assumption is analogous to the assumption that gender is uncorrelated with the error term in the equation $success_i = \alpha + \beta gender_i + f(merit_i) + \epsilon_i$ when the sample comprises successful and unsuccessful candidates.

females separately and make comparisons at specific quantiles. Systematic gender differences at the bottom of the distributions are suggestive of systematic merit differences between the "marginal" male and female candidates and, thus, of gender differences in the level of merit needed to make a career step.

4 Results

4.1 Main results

Figure 2 illustrates the main results. It shows the estimated coefficient on the female indicator across 8 separate regressions varying the measure of merit (i.e. publication and citation scores) and the domain (i.e. faculty positions, editor positions, network affiliations and research grants). In all regressions, the point estimate is negative implying that the average female economist who makes a career step has less scientific merit than the average male economist who makes the same career step.

While the estimates are consistently negative across all four domains, there is significant heterogeneity in the size of the merit gap. For faculty positions, the estimated difference is around 0.1 log-point for either merit measure, which corresponds to a merit gap of around 10%. The difference widens to around 0.2 log-points for editor positions, 0.3-0.4 log-points for network affiliations and 0.4-0.6 log-points for research grants. The gap tends to be slightly larger when merit is measured in terms of citations rather than publication scores.

Figure 3 illustrates the heterogeneity within each of the four domains. Panel A shows how the results for faculty positions vary with the type of position. The merit gap is consistently largest for appointments to associate professor, which typically coincides with the tenure decision. Panel B shows how the results for research grants vary across countries. The merit gap is smallest in France and quite similar in the United Kingdom, Germany and United States. Panel C shows the results for network affiliations for the two networks separately. The merit gap is larger for NBER than CEPR across both merit measures. Panel D reports results for editor appointments for the 20 highest-ranking journals and other journals separately. There is virtually no merit gap for the top journals, only in lower-ranking journals.

We also investigate potential heterogeneity over time by splitting the sample at the onset of the global #MeToo movement in 2017. As shown in Figure A.1 in the Online Appendix, our results do not provide clear evidence that this watershed moment for gender relations is the origin of the merit gap in the economics profession. The merit gap generally existed before

#MeToo and did not systematically become wider after.[11]

4.2 Robustness

We probe the robustness of the main results in three different ways and report the results in the Online Appendix.

First, the estimated merit gap could reflect gender differences in collaboration patterns if female economists work more alone or in smaller groups (Ductor, Goyal and Prummer, 2023). We re-estimate the model with alternative measures of scientific merit that weigh the contribution of each paper to the authors' citation and publication scores by the inverse of the number of authors. As shown in Figure A.2, the estimates remain qualitatively unchanged.

Second, decisions about career steps plausibly account for expected future publications and citations, which are relatively predictable over short horizons due to the time lag involved in the editorial process (e.g. a conditionally accepted paper is a strong predictor of a future publication) and the autocorrelation in annual citation counts (e.g. a high citation count this year is a strong predictor of a high citation count next year). To ensure that anticipation of future merit gains does not create a bias, we re-estimate the model with alternative merit measures that capture citation and publication scores 3 years after the career step. As shown in Figure A.3, the estimates are quite robust to this change.

Third, the merit gap may reflect that female economists are younger (in terms of years since PhD completion) than male economists making the same career step. This could happen if selection committees aim to improve the gender balance in the faculty, the network or the editorial group, and therefore choose to promote high-potential females earlier than high-potential males. This would introduce a merit gap at the time of the career step, even if merit measured over the entire career (in expectation) is equalized across genders. To explore this hypothesis, we re-estimate the model with an alternative outcome that measures each economist's scientific merit relative to other economists in the same PhD cohort, i.e. the within-cohort rank. As shown in Figure A.4, the merit gap remains highly significant for network appointments and research grants with this specification whereas it becomes insignificant for faculty positions and editor appointments. To the extent that the merit gap identified in the main analysis partly reflects that female scholars are appointed to editorial and faculty positions earlier than male scholars with the same long-run merit, this still implies an unequal treatment in favor of female

[11] By comparison, Card et al. (2022) find strong evidence of time trends in their study of selection to the Econometric Society, but over a much longer period: a penalty for females in the early years (1933-1979) and a premium in the late years (2010-2020).

scholars.

4.3 Distribution

Figure 4 plots the cumulative distribution of "relative scientific merit" separately for males (gray) and females (color) in each of the four domains. The results give rise to different conclusions across domains.

For faculty positions (Panel A), the male and female distributions are highly similar until the 80th percentile. Specifically, marginal candidates, i.e. those at say the 5th or the 10th percentile, do no exhibit any visible gender differences. Around the 80th percentile, the two distributions start diverging and at each position at the top, say at the 90th or 95th percentile, males have more scientific merit than females. This pattern suggests that the bar for selection is similar for male and female candidates and that the difference in average merit across successful male and female candidates reflects that the male distribution exhibits a thicker tale of candidates with extremely high levels of merit relative to the career step they are making.

For research grants and network affiliations (Panels B and C), the patterns are markedly different. The entire female distribution is shifted left relative to the male distribution, indicating that successful female candidates have less scientific merit than their male counterparts at every position of the distributions. Specifically, the difference is pronounced at the bottom, suggesting that the threshold for making a career step in these domains tends to be lower for female than for male candidates.

For editor positions (Panel D), the male and female distributions are similar at the very bottom; however, they diverge somewhere between the 5th and the 10th percentile and male candidates have more merit than female candidates at every position above that. Considering that the number of editors appointed at the same journal in the same year is generally low, many "marginal" candidates are above the 10th percentile.[12] Hence, the evidence suggests a lower threshold for females in this domain.

We complement the raw distributional comparisons with quantile regressions of relative scientific merit on gender. This yields estimates that are analogous to horizontal comparisons of the male and female distributions in Figure 4 and allow for statistical inference. The results are illustrated in Figure A5 in the Online Appendix. For faculty positions, the merit gap is statistically significant only at the top of the distribution (above 90th percentile); for research grants and network affiliations, it is significant at all quantiles; and for editor positions, it is

[12] The average number of economists in the "mixed-gender" career steps that contribute to identification is less than 4.

borderline significant at all quantiles except the very bottom (5th percentile) and the very top (95th percentile).

Finally, we conduct an alternative test of gender differences in the merit thresholds. For each unique combination of a career step and a year, we identify the successful candidate with the lowest level of scientific merit by each of the two merit measures. We then compare the fraction of females in these two sets of marginal candidates to the fraction of females in the overall set of successful candidates within each domain. The results are illustrated in Figure A6 in the Online Appendix. Across all four domains and both measures of merit, females constitute a larger share of the marginal candidates than overall.

5 Conclusion

Male scholars occupy most of the top positions in economics and more broadly in academia. A common view holds that this asymmetry reflects unequal treatment in the sense that female scholars need more scientific merit than males to be selected for the same career steps.

Our analysis of four types of career steps in economics does not support this view. We consistently find that the average female scholar has less scientific merit than the average male scholar who makes the same career step. In most domains, this reflects gender differences for marginal scholars, consistent with lower merit thresholds for females.

The results suggest that more subtle mechanisms are required to understand the low female representation in top academic positions.

References

[1] Auriol, E., Friebel, G., Weinberger, A., and Wilhelm, S. (2022). Underrepresentation of women in the economics profession more pronounced in the United States compared to heterogeneous Europe. *Proceedings of the National Academy of Sciences*, 119(16), e2118853119.

[2] Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, 145, 27-41.

[3] Bosquet, C., Combes, P. P., and García-Peñalosa, C. (2019). Gender and promotions: Evidence from academic economists in France. *Scandinavian Journal of Economics*, 121(3), 1020-1053.

[4] Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2020). Are referees and editors in economics gender neutral? *Quarterly Journal of Economics*, 135(1), 269-327.

[5] Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2022). Gender differences in peer recognition by economists. *Econometrica*, 90(5), 1937-1971.

[6] Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2023). Gender gaps at the academies. *Proceedings of the National Academy of Sciences*, 120(4), e2212421120.

[7] Conroy, M. E., Dusansky, R., Drukker, D., and Kildegaard, A. (1995). The productivity of economics departments in the US: Publications in the core journals. *Journal of Economic Literature*, 1966-1971.

[8] Ductor, L., Goyal, S., and Prummer, A. (2023). Gender and collaboration. *Review of Economics and Statistics* 105(6), 1366–1378.

[9] Eberhardt, M., Facchini, G., and Rueda, V. (2023). Gender differences in reference letters: Evidence from the economics job market. *Economic Journal*, 133(655), 2676-2708.

[10] European Commission (2023). 2023 report on gender equality in the EU. Available at `https://commission.europa.eu/system/files/2023-04/annual_report_GE_2023_web_EN.pdf`

[11] Goldin, C., and Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review*, 90(4), 715-741.

[12] Grant Thornton (2023). Women in Business 2023. Available at `https://www.grantthornton.com/content/dam/grantthornton/website/assets/content-page-files/advisory/pdfs/2023/women-in-business-2023.pdf`

[13] Heckman, J. J., and Moktan, S. (2020). Publishing and promotion in economics: The tyranny of the top five. *Journal of Economic Literature*, 58(2), 419-470.

[14] Hengel, E. (2022). Publishing while female: Are women held to higher standards? Evidence from peer review. *Economic Journal*, 132(648), 2951-2991.

[15] IDEAS/RePEc (2023). Economics and Finance Research. Available at `https://ideas.repec.org`.

[16] Lundberg, S., and Stearns, J. (2019). Women in economics: Stalled progress. *Journal of Economic Perspectives*, 33(1), 3-22.

[17] Mengel, F., Sauermann, J., and Zölitz, U. (2019). Gender bias in teaching evaluations. *Journal of the European Economic Association*, 17(2), 535-566.

[18] Neumark, D., Bank, R. J., and Van Nort, K. D. (1996). Sex discrimination in restaurant hiring: An audit study. *Quarterly Journal of Economics*, 111(3), 915-941.

[19] Sarsons, H., Gërxhani, K., Reuben, E., and Schram, A. (2021). Gender differences in recognition for group work. *Journal of Political economy*, 129(1), 101-147.

[20] Sherman, M. G., and Tookes, H. E. (2022). Female representation in the academic finance profession. *Journal of Finance*, 77(1), 317-365.

[21] Sin, I., Stillman, S., and Fabling, R. (2022). What Drives the Gender Wage Gap? Examining the Roles of Sorting, Productivity Differences, Bargaining, and Discrimination. *Review of Economics and Statistics*, 104(4), 636-651.

[22] United Nations (2023). Women in politics: 2023. Available at `https://www.unwomen.org/en/digital-library/publications/2023/03/women-in-politics-map-2023`

[23] UNDP (2021). Gender Equality in Public Administration. Available at `https://www.undp.org/publications/global-report-gender-equality-public-administration`

[24] Wu, A. H. (2018). Gendered language on the economics job market rumors forum. *American Economic Review* (Papers and Proceedings), 108, 175-179.

[25] Wu, A. H. (2020). Gender bias among professionals: an identity-based interpretation. *Review of Economics and Statistics*, 102(5), 867-880.

Table 1: Estimation samples. The table describes the samples of economists who make a career step in the four domains, i.e. are appointed to a faculty position at an economics department, receive a scientific research grant, become members of an academic network, and are appointed editor of an economics journal. The table shows the time period covered (Column 1), the total number of unique career steps (Column 2), the total number of economists making a career step (Column 3), the number of "mixed-gender" career steps, i.e. where the set of economists making the career step includes at least one male and one female (Column 4), and the number of economists making a "mixed-gender" career step (Column 5).

	(1)	(2)	(3)	(4)	(5)
		All career steps		**Mixed-gender career steps**	
	Period	# Distinct career steps	# Economics scholars	# Distinct career steps	# Economics scholars
Faculty	**2012 - 2023**	**1,206**	**2,333**	**288**	**998**
- Assistant professor	2012 - 2023	445	915	129	422
- Associate professor	2012 - 2023	351	666	78	269
- Full professor	2012 - 2023	410	752	81	307
Grants	**1994 - 2023**	**271**	**4,186**	**126**	**3,817**
- France	2005 - 2022	18	1,379	17	1,351
- Germany	1994 - 2023	170	669	46	382
- UK	2015 - 2020	6	236	6	236
- USA	1999 - 2023	77	1,902	57	1,848
Networks	**2001 - 2023**	**629**	**2,197**	**217**	**1,441**
- CEPR	2001 - 2022	402	1,220	137	745
- NBER	2001 - 2023	227	977	80	696
Editors	**2004 - 2023**	**496**	**1,028**	**111**	**424**
- Top20	2004 - 2023	168	372	41	156
- Below20	2004 - 2023	328	656	70	268

Figure 1: Descriptives. Panel A shows how the female share of the economists who make a career step in each of the four domains evolves over the sample period (3-year moving averages). Panel B shows the correlation between our two measures of scientific merit, i.e. citation counts and publication scores (both in logs), in a binned scatterplot. Panel C shows the distribution of citations for the full sample of academic economists registered at IDEAS/RePEc (black line) and each of the four samples of economists who make a career step (colored lines). Panel D shows the distribution of publication scores for the full sample of academic economists registered at IDEAS/RePEc (black line) and each of the four samples of economists who make a career step (colored lines).

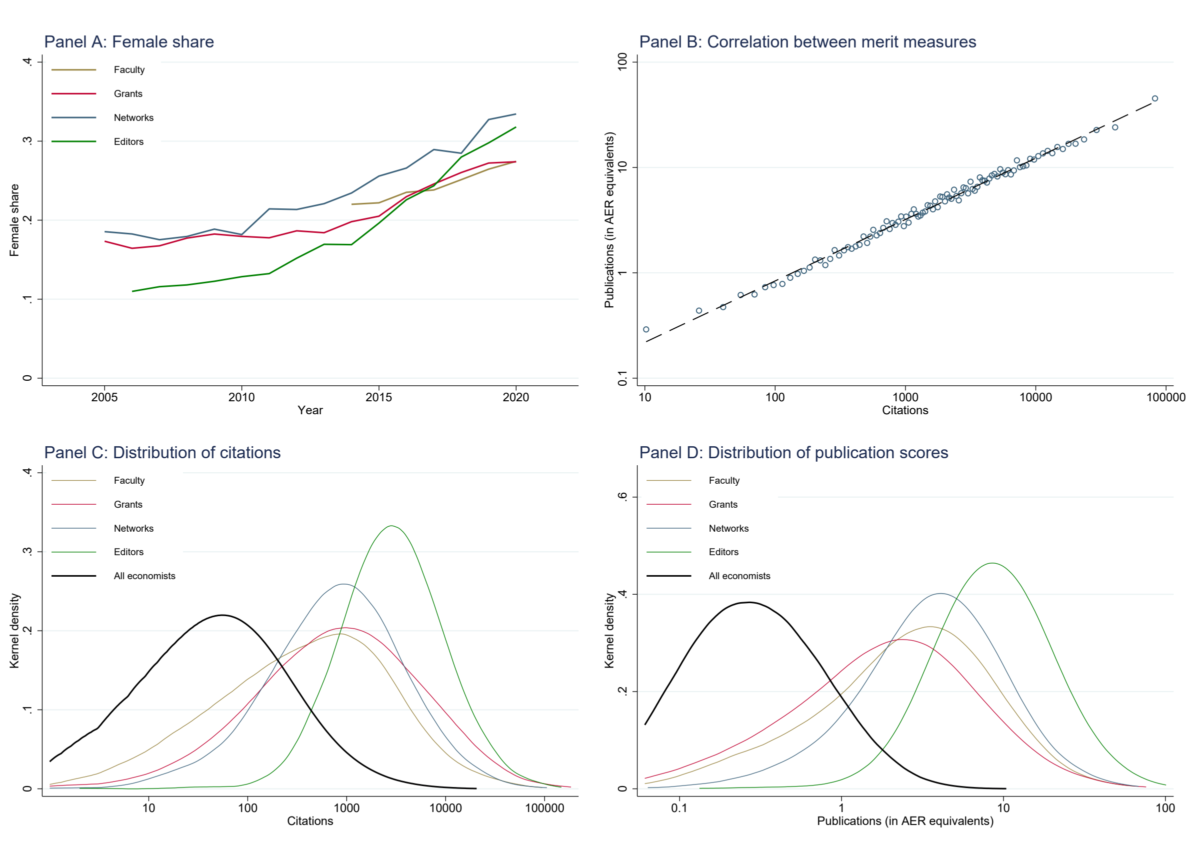


Figure 2: Main results. The figure illustrates the main regression results. Each bar represents the mean difference in publication scores (left side) and citation counts (right side) between female and male economists at the time they make the same career step in one of the four domains, i.e. are appointed to a faculty position at an economics department (brown bars), are appointed editor of an economics journal (green bars), become members of an academic network (blue bars), and receive a scientific research grant (red bars).

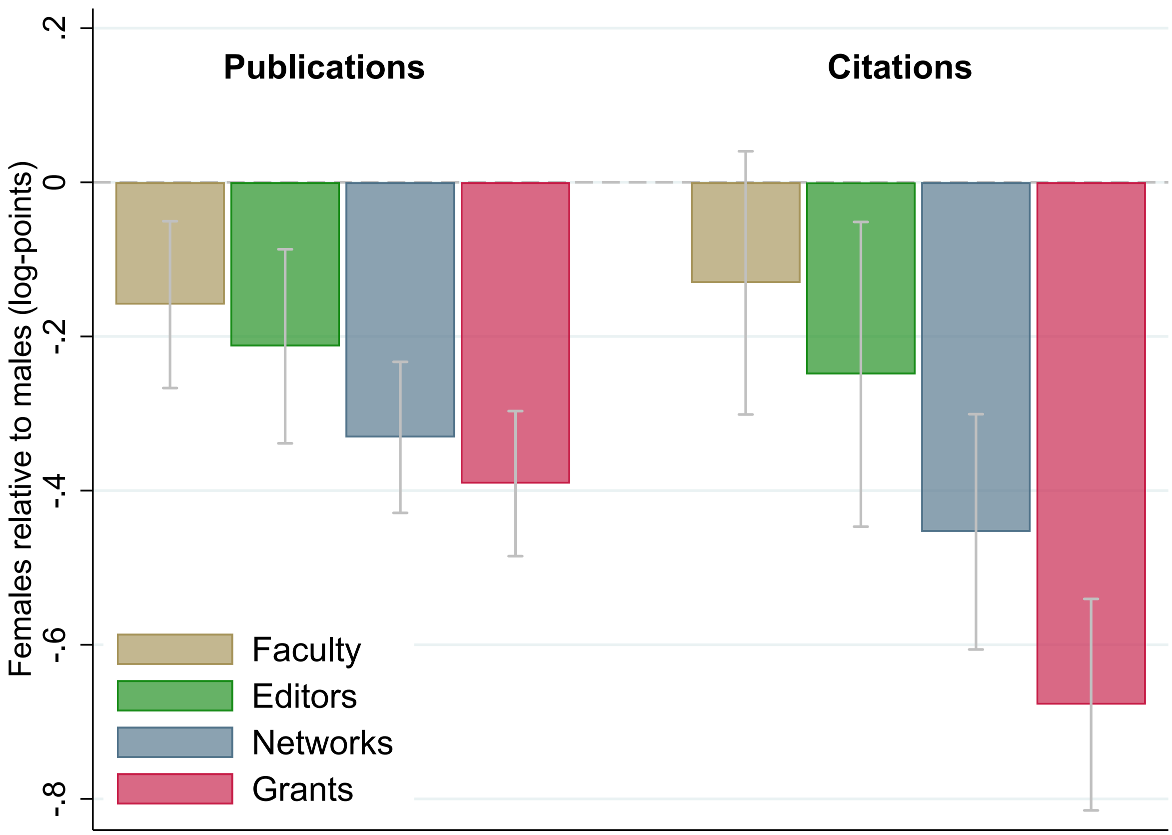


Figure 3: Heterogeneity. The figure shows how the main results vary across subsamples: Panel A shows the results for faculty positions for Assistant Professors, Associate Professors and Full Professors separately. Panel B shows the results for research grants for the French, U.S., German and U.K. research councils separately. Panel C shows the results for academic networks by CEPR and NBER separately. Panel D shows the results for editorships for Top-20 journals and other journals separately.

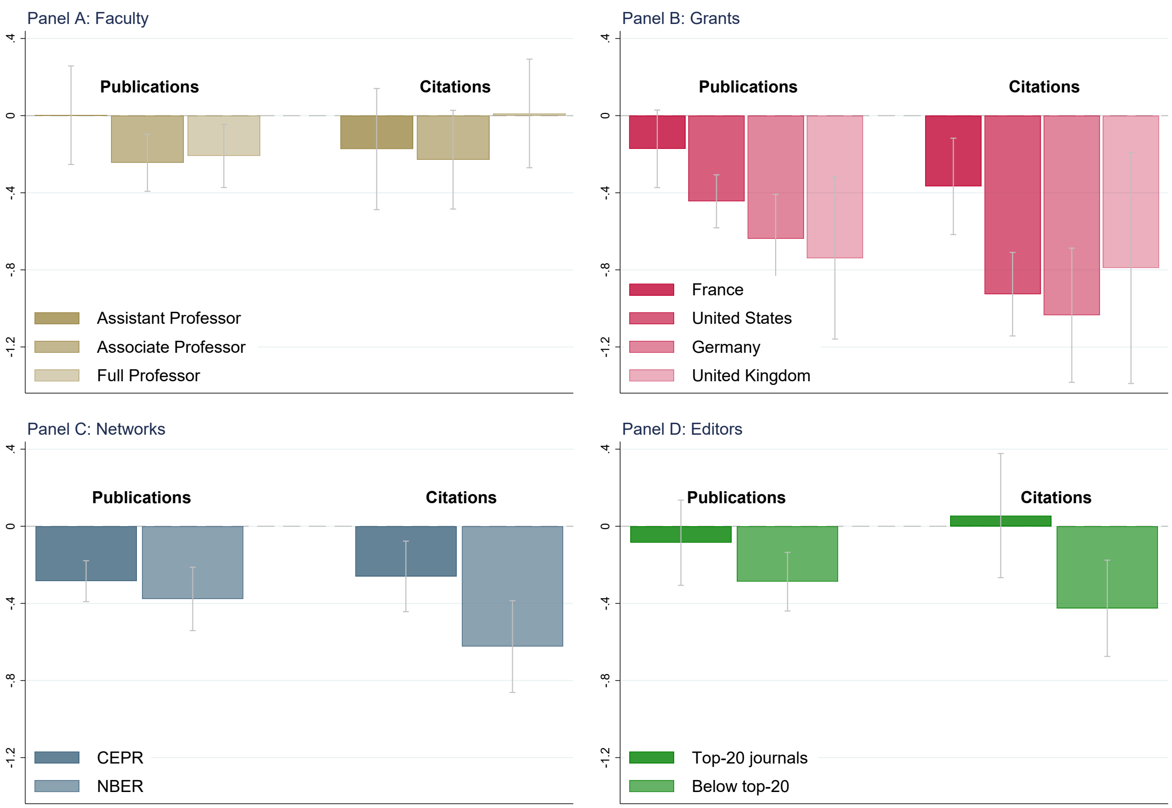
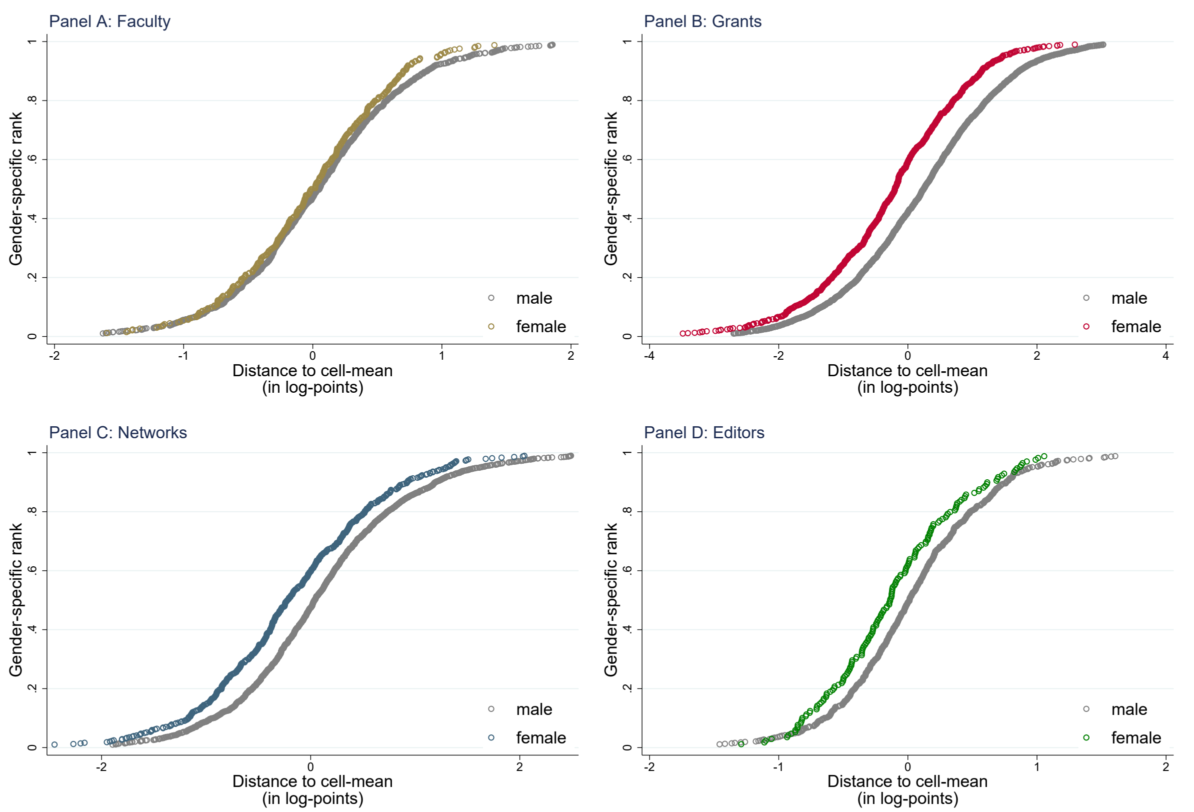


Figure 4: Distribution of relative scientific merit. The figure shows the cumulative distribution of scientific merit of male economists (gray lines) and female economists (colored lines) in our four samples measured *relative* to the average taken across other economists making the same career step in the same year. To construct the figures, we first regress each measure of scientific merit, i.e. citation counts and publication scores, on career step fixed effects. For each economist, we then take the average of their residuals in the citation and publication regressions to obtain a single measure of their relative scientific merit at the time they make a career step, i.e. merit relative to the other economists who make the same career step. Finally, we rank these relative merit measures, for male and female economists separately, and plot them against their ranks.



ONLINE APPENDIX

Appendix A: Additional results

Figure A1: Heterogeneity by time period. The figure shows estimates from the baseline regression where the sample period is split into two subperiods: the period before 2018 and the period from 2018 and onwards. The bars represent estimates of the mean difference in publication scores (left side) and citation counts (right side) between female and male economists at the time they make the same career step.

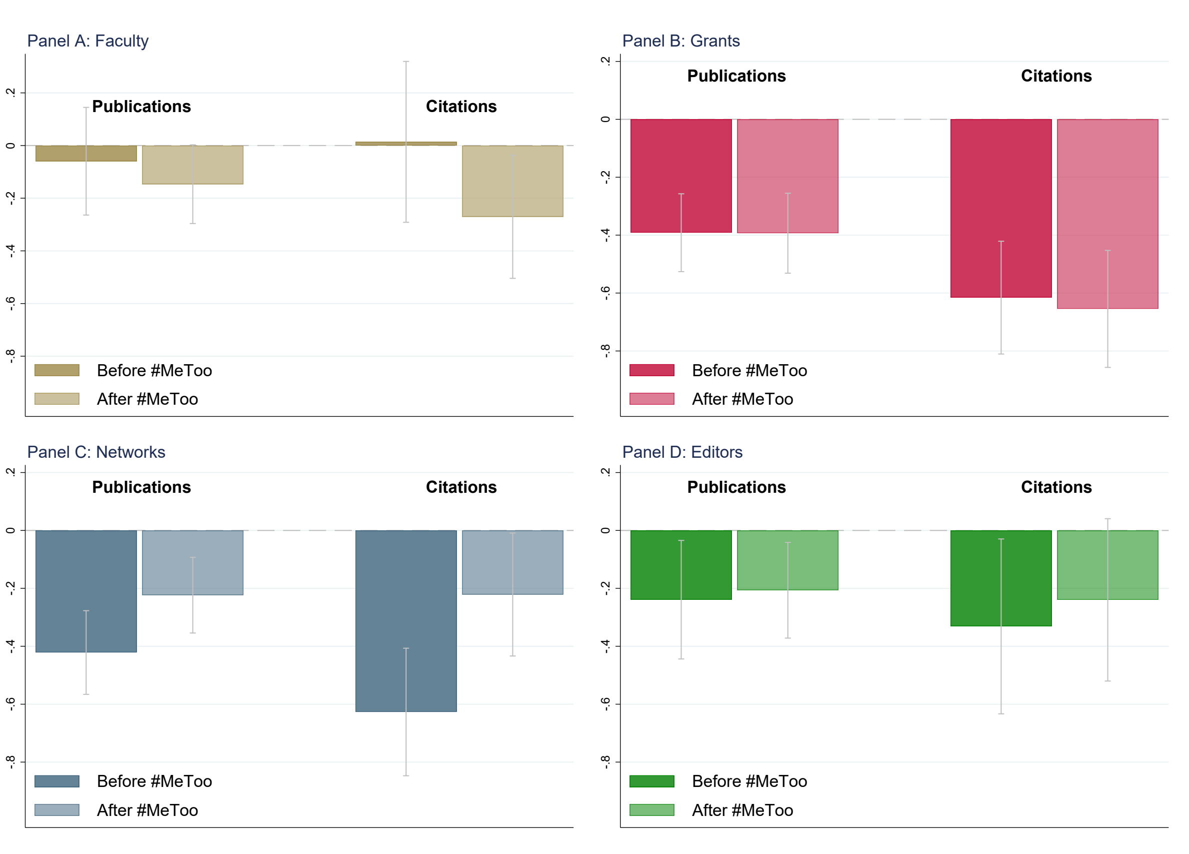


Figure A2: Robustness to discounting joint work. The figure shows estimates from the baseline regression where citation counts and publication scores are adjusted for the number of co-authors. The bars represent estimates of the mean difference in adjusted publication scores (left side) and adjusted citation counts (right side) between female and male economists at the time they make the same career step.

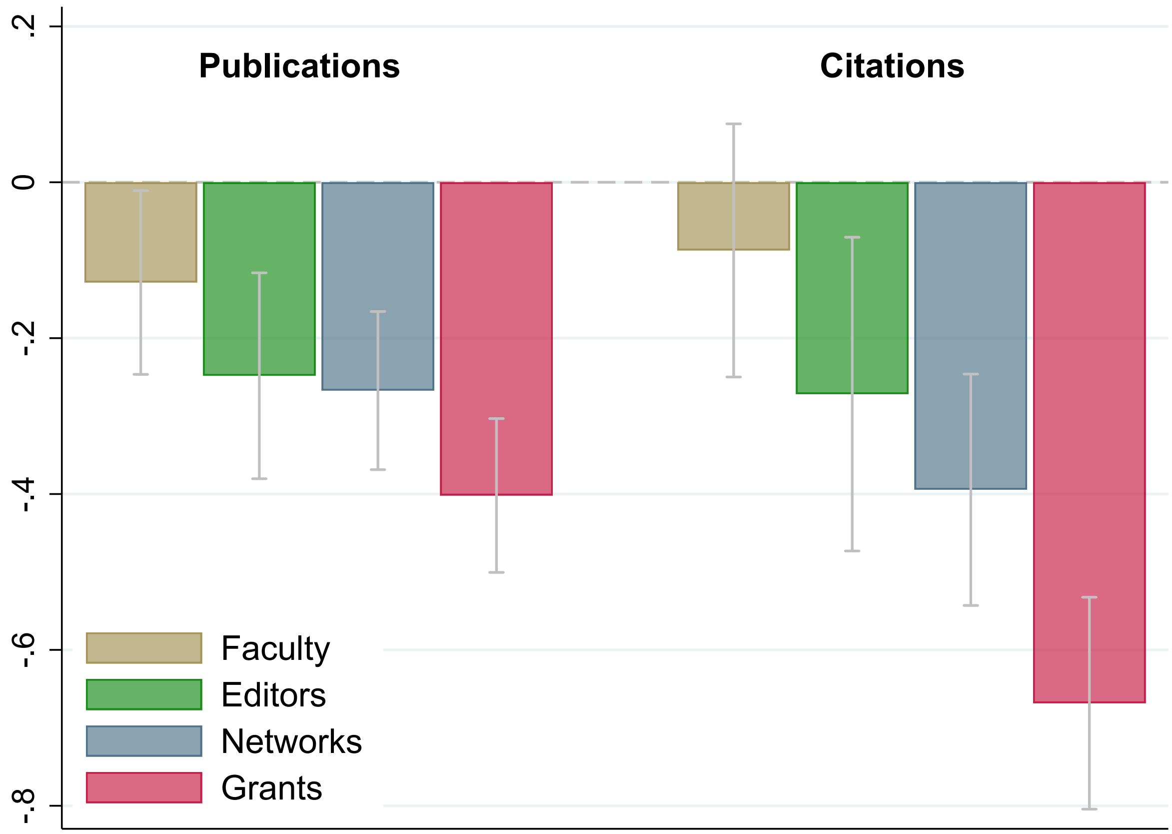


Figure A3: Robustness to using ex post measures of scientific merit. The figure shows estimates from the baseline regression where citation counts and publication scores are measured ex post. The bars represent estimates of the mean difference in publication scores (left side) and citation counts (right side) between female and male economists 3 years after they make the same career step.

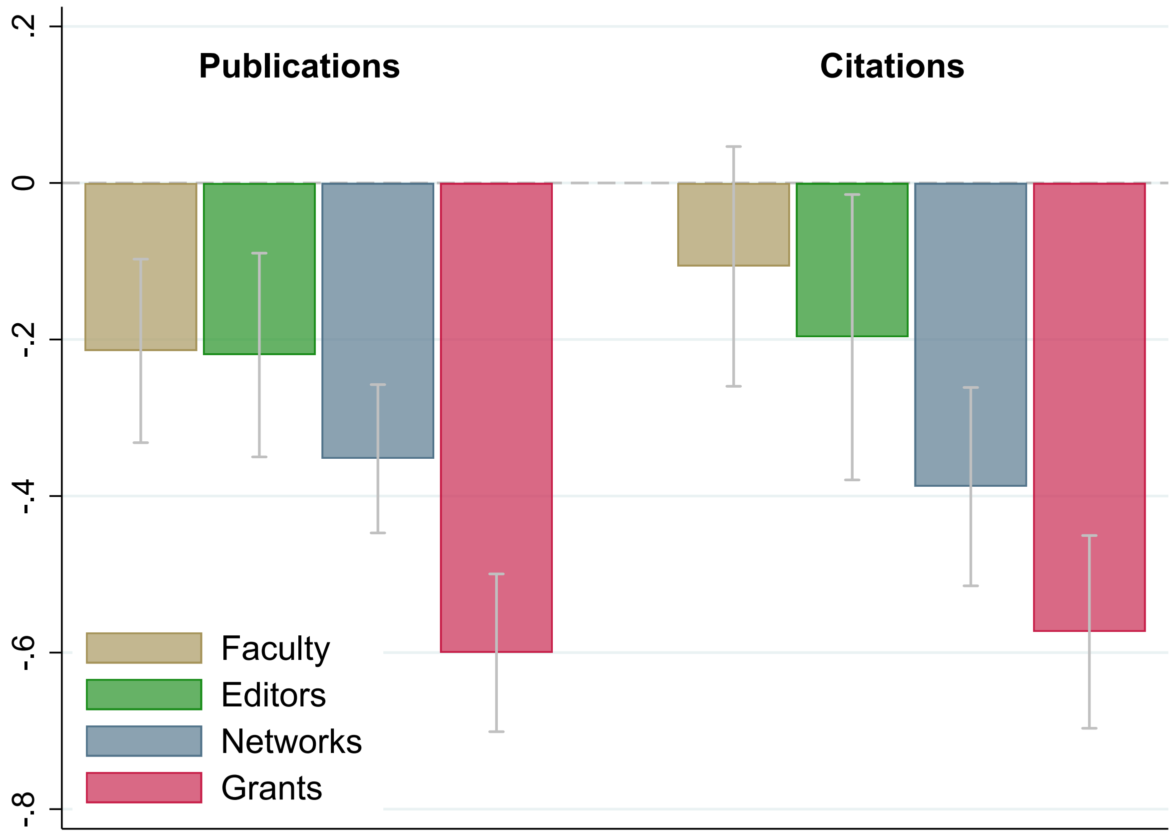


Figure A4: Robustness to using within-cohort ranks. The figure shows estimates from the baseline regression where scientific merit is measured as the rank in citation counts and in publication scores within the group of economists of the same PhD age registered in IDEAS/RePEc. The bars represent estimates of the mean difference in publication rank (left side) and citation rank (right side) between female and male economists at the time they make the same career step.

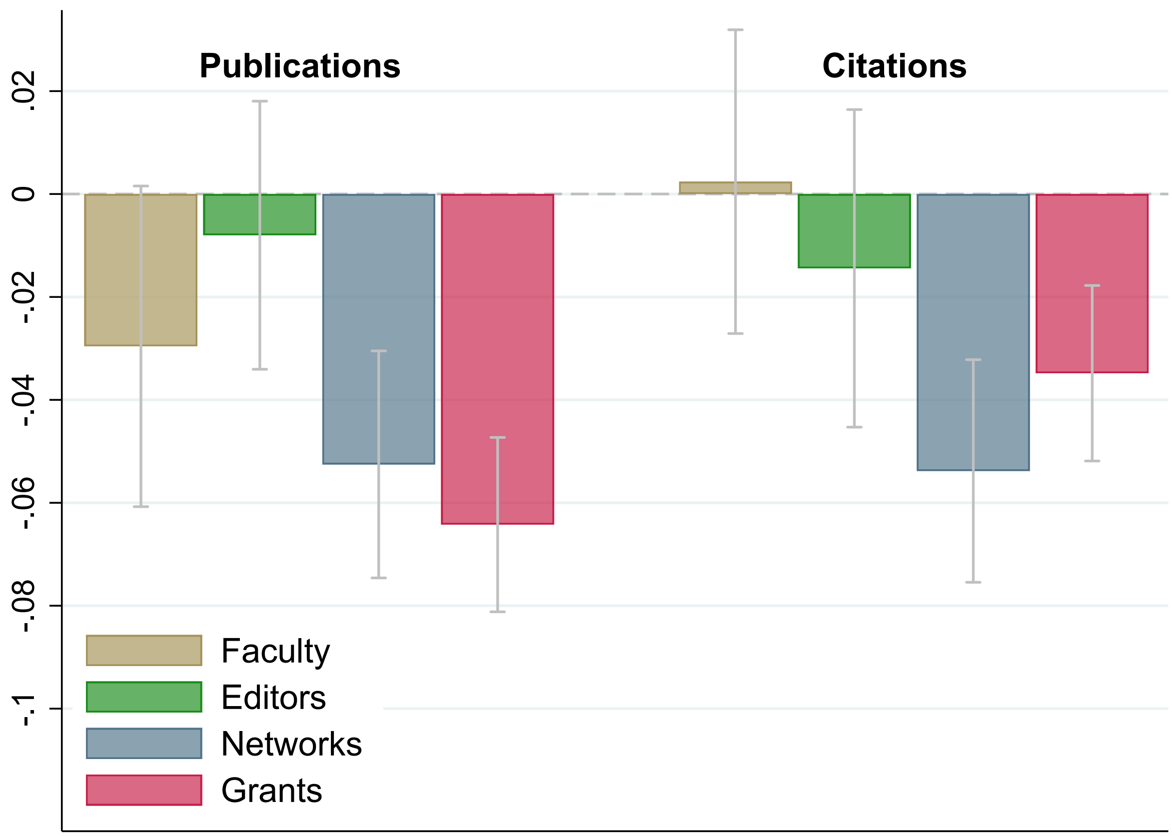


Figure A5: Quantile regression results. The figure shows the differences in scientific merit across male and female economists who make the same career step at various positions in the gender-specific distributions of relative merit. To construct the figures, we first regress each measure of scientific merit, i.e. citation counts and publication scores, on career step fixed effects. For each economist, we then take the average of their residuals in the citation and publication regressions to obtain a single measure of their relative scientific merit at the time they make a career step, i.e. merit relative to the other economists who make the same career step. Finally, we regress this relative merit measure on a female dummy in a series of quantile regressions.

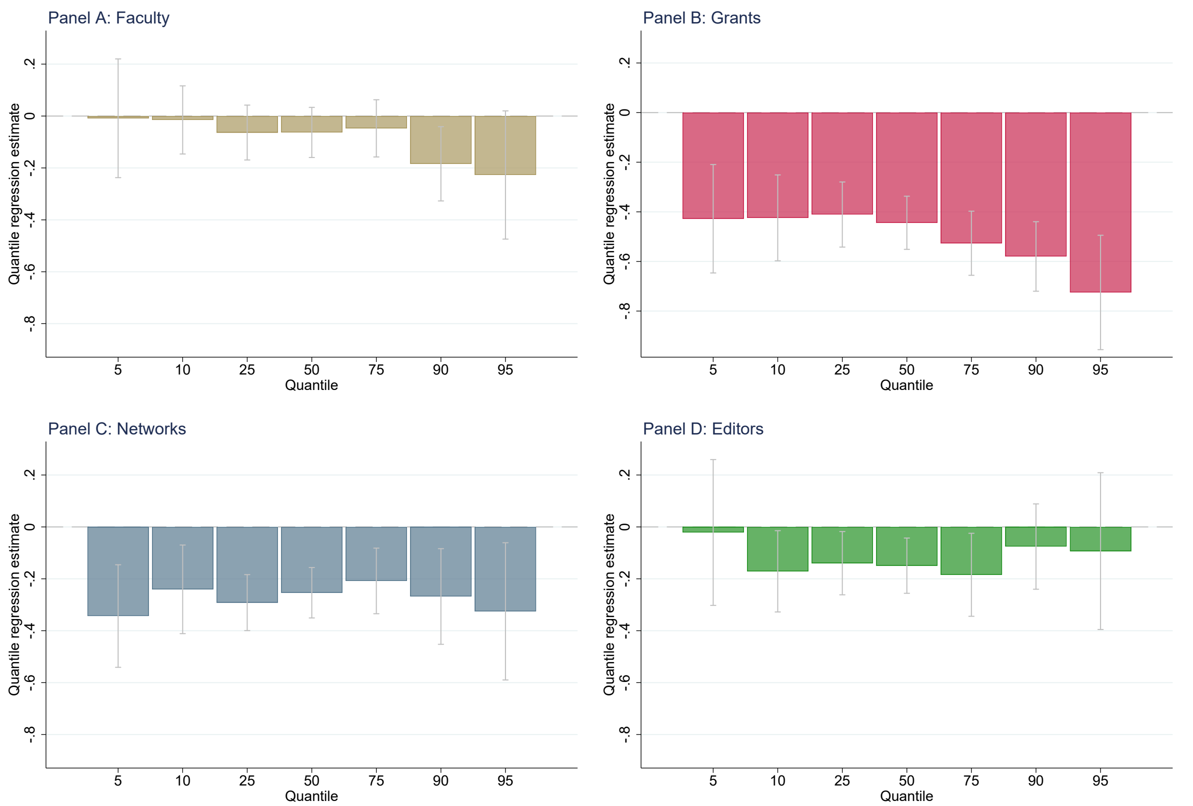
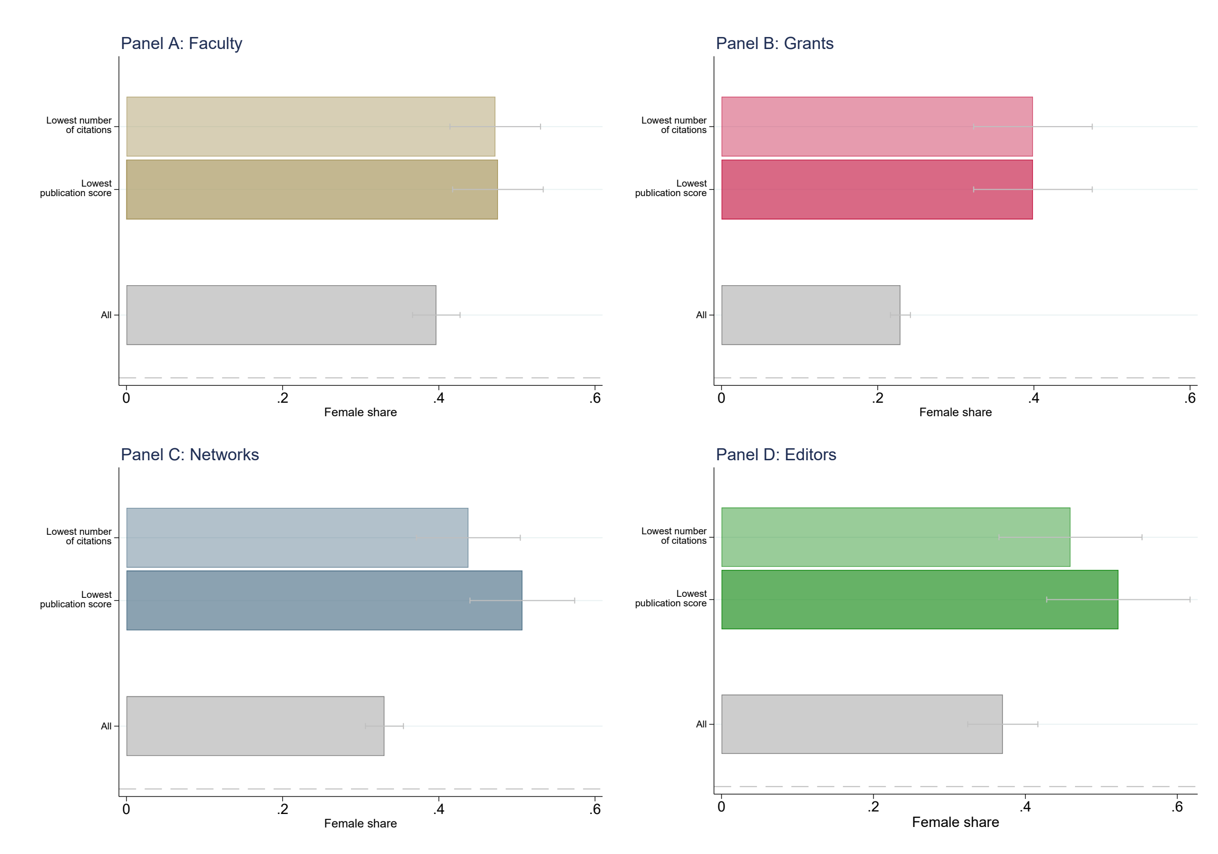


Figure A6: Female shares at the lowest level of relative merit vs. overall. The figure compares the female share overall among the economists who make a career step (gray bars) to the female share among the economists who have the lowest scientific merit among those making the same career step (colored bars). To construct the gray bars, we simply compute the share of females in each career step and take the average across all career steps in a given domain. To construct the colored bars, we identify the economist with the lowest publication score and citation count within each cell and compute the female share within each of these two groups of economists.



Appendix B: Data Construction

In this Appendix, we provide a detailed description of the steps involved in constructing the estimation dataset. Step 1 identifies career steps in each of the four domains: network affiliations, research grants, editor appointments and faculty positions. Step 2 identifies a reference population of academic economists. Step 3 identifies the Google Scholar pages of each of the scholars in our samples. Step 4 adds the gender of each of the scholars. Step 5 adds information about the publications and citations for each of the scholars. Step 6 computes AER equivalents of each paper published by the scholars. Step 7 creates the variables used in the empirical analysis.

We use the term *raw datasets* to refer to the datasets created in Step 1 (i.e. each row corresponds to a scholar who made a career step) and Step 2 (each row corresponds to a scholar in the reference population). We use the term *clean datasets* to refer to the datasets created in Steps 3-4 where information about the Google Scholar page and gender of each scholar is added. We use the term *full datasets* to refer to the much larger datasets created in Steps 5-6 by adding detailed information about publications and citations. We use the term *estimation datasets* to refer to smaller datasets created in Step 7 by collapsing the detailed information about publications and citations to the merit measures that are used in the empirical analysis.

All the source data comes from free and publicly available online sources. Most commonly, we collect the source data through web scraping. In these cases, we include the python programs that perform the scraping in the replication package and list them under the relevant step in the description below. Occasionally, the source data is directly available for download at the source. In these cases, we describe where the source files can be downloaded. One type of source data is hand-collected by manually copying information from non-scrapable text files.

All the scholar-level information in the replication package is de-identified. The only scholar-level dataset in the replication package is the estimation dataset where identifying information such as names and URLs of the Google Scholar pages is removed. The package does not include the source data nor any of the intermediate datasets with identifying information.

Step 1: Identify career steps

In Step 1, we identify career steps made by economics scholars in each of the four domains: network affiliations, research grants, editor appointments and faculty positions. The specific approach varies across domains depending on the available information; however, we always rely on free and publicly available online sources. A tool that we use repeatedly is the internet archive Wayback Machine, which contains archived versions of websites from different points in time (Available here: `https://archive.org/web/`).

Network affiliations

CEPR Appointments

We scrape historical versions of the CEPR website accessed through Wayback Machine.[13]

[13] `https://cepr.org/`

Specifically, for each year from 2000 to 2023, we scrape an archived version of the website to collect the name and program area (i.e. field) of every individual on the list of research fellows. We generally select website versions that are approximately one year apart. We were unable to find functional website versions for two years: 2002 and 2018.

We use this information to create a raw dataset of research fellow appointments at CEPR. We consider that a scholar is appointed in year t if her name appears on the list of research fellows in year t, and not in year $t-1$. Given that we were unable to find functional website versions 2002 and 2018, we cannot identify research fellows appointed in 2002-2003 and 2018-2019. We keep the following information in the raw dataset of CEPR appointments: name of scholar, CEPR appointment year and program area.

The replication package includes the python program that performs the web scraping and creates the raw dataset of CEPR appointments.

- *Program*: 0. MAKERAW_CEPR.py

NBER Appointments

We scrape historical versions of the NBER website accessed through Wayback Machine.[14] Specifically, for each year from 2000 to 2023, we scrape an archived version of the website to collect the name and program area (i.e. field) of every individual on the list of affiliated scholars. We generally select website versions that are approximately one year apart.

We use this information to create a raw dataset of affiliated scholar appointments at NBER. We consider that a scholar is appointed in year t if her name appears on the list of affiliated scholars in year t, and not in year $t-1$. We keep the following information in the raw dataset of NBER appointments: name of scholar, NBER appointment year and program area.

The replication package includes the python program that performs the web scraping and creates the raw dataset of NBER appointments.

- *Program*: 0. MAKERAW_NBER.py

Research grants

U.S. grants

The website of the National Science Foundation (NSF) contains files with comprehensive information about all the grants awarded by this agency since 2000.[15]. We download the files concerning grants awarded under the program "Economics".

We use this information to create a raw dataset of research grants awarded to scholars in economics in the United States. We consider that a scholar was awarded a grant in year t if she was the principal investigator of a project that was awarded an NSF grant in year t. We keep the following information in the raw dataset of U.S. research grants: name of scholar, grant year, agency, award instrument.

The replication package includes the python program that consolidates the information from the thousands of NSF source files (one file per grant) into the raw dataset (one row per grant).

- *Program*: 0. MAKERAW_GrantsUS.py

[14] https://www.nber.org/

[15] https://www.nsf.gov/awardsearch/download.jsp

U.K. grants

The website of the United Kingdom Research and Innovation (UKRI) contains files with comprehensive information about all the grants awarded by this agency over the period 2015-2021, including grants awarded to economists by the Economic and Social Research Council (ESRC).[16]. We download these files.

We use this information to create a raw dataset of research grants awarded to scholars in the social sciences in the United Kingdom, i.e. all the grants awarded by the ESRC. We consider that a scholar was awarded a grant in year t if she was the principal investigator of a project that was awarded a grant by the ESRC in year t. We keep the following information in the raw dataset of U.K. research grants: name of scholar, grant year, grant category. The challenge that the UKRI data does not distinguish directly between grants in economics and other social sciences is addressed in Step 3 below.

The replication package includes the python program that extracts the relevant information from the UKRI source files and creates the raw dataset of U.K. grant awards.

- *Program*: 0. MAKERAW_GrantsUK.py

French grants

The website of the Agence Nationale de la Recherche (ANR) contains files with information about all grants awarded by this agency since 2005.[17] We download these files.

We use this information to create a raw dataset of research grants awarded to academic scholars in France. We consider that a scholar was awarded a grant in year t if she was the principal investigator of a project that was awarded a grant by ANR in year t. We keep the following information in the raw dataset of French research grants: name of scholar and grant year. The challenge that the ANR data does distinguished directly between grants in economics and other sciences is addressed in Step 3 below.

The replication package includes the python program that extracts the relevant information from the ANR source files and creates the raw dataset of French grant awards.

- *Program*: 0. MAKERAW_GrantsFrance.py

German grants

The website of the Deutsche Forschungsgemeinschaft (DFG) contains a database with information about all grants awarded by this agency since 1994.[18] We scrape the database for information about grants awarded under the subject areas Economic Policy, Applied Economics and Economic theory.

We use this information to create a raw dataset of research grants awarded to academic scholars in Germany. We consider that a scholar was awarded a grant in year t if she was the principal investigator of a project that was awarded a grant by DFG in year t. We keep the following information in the raw dataset of German research grants: name of scholar, grant year and grant program.

[16] `https://www.ukri.org/what-we-do/what-we-have-funded/`

[17] `https://www.data.gouv.fr/fr/organizations/agence-nationale-de-la-recherche/`

[18] `https://gepris.dfg.de/gepris/OCTOPUS?task=showKatalog`

The replication package includes the python program that performs the web scraping and creates the raw dataset of German grant awards.

- *Program*: 0. MAKERAW_GrantsGermany

Editor appointments

Our starting point is the top-100 academic economics journals taken from the ranking by IDEAS/RePEc based on h-indexes.[19] From the gross list of 100 journals, we discard 33 journals that are either interdisciplinary (e.g. *World Development*) or belong to adjacent fields (e.g. *American Political Science Review*). We discard these journals because their editors often publish outside of economics, which renders the AER-equivalent publication scores a poor measure of academic merit. However, we retain journals specializing in finance (e.g. *Journal of Finance*) and econometrics (e.g. *Journal of Econometrics*). For each of the remaining 67 journals and for each year since 2004, we search for the names of the editors in the front matter of the journal. We always start from the first issue of the year and move to later issues if the editor names are not available in the first issue.

We use this hand-collected information to create a raw dataset of editor appointments. We consider that a scholar is appointed editor in year t if her name appears on the list of editors in year t, and not in year $t-1$. We keep the following information in the raw dataset of editor appointments: name of scholar, editor appointment year and journal ranking.

The replication package includes a spreadsheet with the list of the top-100 journals and the python program that extracts the relevant information from the hand-collected dataset and creates the raw dataset of editor appointments.

- *Source data*: SOURCE_Journals-Top100.xlsx
- *Program*: 0. MAKERAW_Editors

Faculty positions

Our starting point is the top-100 academic economics departments taken from the ranking by IDEAS/RePEc.[20] For each department and each year since 2012, we scrape the name and position of each academic scholar affiliated with the department from an archived version of the department website accessed through Wayback Machine. We generally select website versions that are approximately one year apart. From the gross list of 100 departments, we are unable to collect data for 18 departments as the historical versions of department websites are not functional. For the remaining 82 departments, we are able to obtain at least partial coverage. We keep information about positions at the level of assistant professors, associate professors and full professors (or similar such as lecturers at universities in the United Kingdom).

We use this information to create a raw dataset of appointments to new faculty positions. We consider that a scholar is appointed to a new position in a department in year t if her name appears with a given title on the department website in year t, and did not appear at all or appeared with a lower position in year $t-1$. We keep the following information in the

[19]https://ideas.repec.org/top/top.journals.hindex.html

[20]https://ideas.repec.org/top/top.econdept.html

raw dataset of faculty appointments: name of scholar, appointment year, position, university, ranking of the university.

The replication package includes a spreadsheet with the list of the top-100 departments as well as the python program that performs the web scraping and creates the raw dataset of appointments to new faculty positions.

- *Source data*: SOURCE_Departments-Top100.xlsx

- *Program*: 0. MAKERAW_Tenure

Step 2: Construct dataset of all IDEAS/RePEc economists

In Step 2, we initiate a dataset with information about all economists listed in the IDEAS/RePEc database (Available here: `https://ideas.repec.org/i/eall.html`). This will serve as a reference population of academic economists, to which we can compare the economists in our career step samples. In this step, we scrape the name of each economist listed in the IDEAS/RePEc database.

The replication package includes the python program that performs the web scraping.

- *Program*: 0. MAKERAW_Repec

Step 3: Identify Google Scholar pages

In Step 3, we match the scholars identified in Steps 1 and 2 to their Google Scholar pages. There are several challenges to this name-based matching: multiple scholars with the same name, middle names that are not used consistently and special characters with multiple representations in the standard English alphabet.

We use an automated process to search for each of the scholars' names in the Google Scholar search engine. We scrape the search results and the URLs of the Google Scholar pages. We retain only those search results where there is an exact match with the first and last name. Our automated process allows for the inversion of first and last names, the presence or absence of capital letters, as well as the presence or absence of specific characters (e.g., ß, ş, . . .) and of letters with accent (e.g., ü, à, . . .). We bypass the CAPTCHAs tests triggered by the high volume of requests made to Google Scholar by integrating a CAPTCHA solver extension, 2Captcha into Google Chrome when executing our scraping programs.[21].

For some scholars in our samples, there is a unique match to a Google Scholar page. This is likely to reflect a true match; however, we need to address the possibility that the scholar does not have a Google Scholar page and that the matched page belongs to another scholar with the precisely same name. For other scholars, there is no matching Google Scholar page. This is likely to reflect that the scholar does not have a Google Scholar page. For yet other scholars, there are multiple matching Google Scholar pages. This is likely to reflect that multiple scholars with precisely the same name have a Google Scholar page.

To address these challenges, we identify a set of matched Google Scholar pages, which are likely to be mismatches as they appear to belong to non-economists. For each of the matched Google Scholar pages, we examine whether the articles are published in what appears to be

[21] urlhttps://2captcha.com

economics or finance journals. We classify journals as economics or finance if the journal name includes either "conom" or "inanc", thus accommodating both "economics" and "economy" as well as "finance", and "financial" and allowing for capitalization of initial letters. We consider that a matched Google Scholar page belongs to a non-economist, and thus consistutes a mismatch, if none of the five most cited articles on the page are economics or finance journals according to this classification.

In cases with a unique match between a scholar in our sample and a Google Scholar page after discarding pages belonging to non-economists, we consider that the match is correct and retain the URL of the page. In cases with no match, we consider that the scholar is not a scholar of economics or does not have a Google Scholar page. Excluding non-economists at this stage is particularly important in the context of French and U.K. grants where we are not able to distinguish between grants to economists and other disciplines directly in the source data from the grant authorities. In cases where there are multiple matches, we consider ourselves unable to identify the correct Google Scholar page. Finally, for each sample, we create tables with two columns, the name of the scholar and the URL of the scholar's Google Scholar page.

The replication package includes a set of python programs, one program for each sample, that perform Step 3.

- *Program*: 1. MAKECLEAN_CEPR
- *Program*: 1. MAKECLEAN_NBER
- *Program*: 1. MAKECLEAN_Editors
- *Program*: 1. MAKECLEAN_GrantsFrance
- *Program*: 1. MAKECLEAN_GrantsGermany
- *Program*: 1. MAKECLEAN_GrantsUK
- *Program*: 1. MAKECLEAN_GrantsUS
- *Program*: 1. MAKECLEAN_Tenure
- *Program*: 1. MAKECLEAN_Repec

Step 4: Add Gender

In Step 4, we determine the gender of the scholars identified in Steps 1. We rely on the functionality of an existing program that determines gender based on forenames [22]. The program gives the following output for each name: male, female, mostly male, mostly female, or unknown. When the program gives output other than "male" or "female", we hand-collect information about the gender of the scholar. Finally, we add a gender column to the datasets created in Step 3.

This step is performed in the same set of programs that performed Step 3.

Step 5: Collect Google Scholar data on citations and publications

In Step 5, we scrape the Google Scholar pages identified under Step 3 for information about publications and citations. We bypass the CAPTCHAs tests triggered by the high volume of

[22] https://pypi.org/project/gender-guesser/

requests made to Google Scholar by integrating a CAPTCHA solver extension, 2Captcha into Google Chrome when executing our scraping programs.[23].

For the career-step scholars identified in Step 1, we extract information about the title and the co-authors of each paper, the journal of publication for each paper (if published), the year of publication for each paper, and the number of citations attracted by each paper in each year. Hence, we produce a full dataset with one row per scholar × paper × citation year.

For the reference population of economists identified in Step 2, we need less detailed information. We first extract information about the title of each paper, the journal of publication for each paper (if published), and the year of publication for each paper. We use that to produce a dataset with one row per scholar × paper. We then extract information about the overall number of citations attracted by each scholar in each year. We use that to produce a dataset with one row per scholar × citation year.

The replication package includes a set of python programs, one program for each sample, that perform Step 5.

- *Program*: 2. MAKEFULL_CEPR
- *Program*: 2. MAKEFULL_NBER
- *Program*: 2. MAKEFULL_Editors
- *Program*: 2. MAKEFULL_GrantsFrance
- *Program*: 2. MAKEFULL_GrantsGermany
- *Program*: 2. MAKEFULL_GrantsUK
- *Program*: 2. MAKEFULL_GrantsUS
- *Program*: 2. MAKEFULL_Tenure
- *Program*: 2. MAKEFULL_Repec

Step 6: Add AER-equivalent score

In Step 6, we create the AER-equivalent score for each journal. Our starting point is the top-500 academic economics journals taken from the ranking by IDEAS/RePEc based on h-indexes.[24]. We compute a journal's AER-equivalent score as the ratio between the journal's factor score and the factor score of the AER (the journal with the highest factor score). We then merge the AER-equivalent scores on the full datasets created in Step 5. The merging happens in the same set of programs that performed Step 5.

The replication package includes a spreadsheet with the titles of the top-500 journals along with information about their factor scores and the computed AER-equivalent scores.

- *Source data*: SOURCE_AERequivalents-Top500.xlsx

Step 7: Create estimation datasets

In Step 7, we create an estimation dataset for each sample of scholars making career steps. For a scholar making a career step in year t, we cumulate the AER-equivalent scores of the papers published in years $s \leq t$ to obtain a publication-based measure of merit in year t, and

[23] urlhttps://2captcha.com

[24] `https://ideas.repec.org/top/top.journals.hindex.html`

we cumulate the citations in years $s \leq t$ to obtain a citation-based measure of merit in year t. We construct alternative merit measures that allow for forward-looking selection in year t by cumulating publications and citations over years $s \leq t + 3$. We construct yet other merit measures that discount co-authored work by weighing citations and publications by the inverse of the number of authors. Finally, we construct rank measures of publication- and citation-based merit, which account for PhD age. In the reference population of academic economists, we compute the approximate PhD age of each scholar as the difference between year t and the year of the first publication (for the purposes of publication-based merit) and the year of the first citation (for the purposes of citation-based merit). We then compute the rank of each scholar within the set of individuals with the same PhD age and add it to the estimation dataset. In the small number of cases where a scholar in the estimation sample is not represented in the reference population of economist, we impute a rank from the rank of scholars with the same PhD age and a highly similar merit measure.

The replication package includes the stata program that imports the full datasets created in Steps 1-6 and creates four estimation datasets, one for each of the four domains where we observe career steps. It also includes de-identified versions of the four estimation datasets.

- *Program*: 3. MAKEESTIMATION_All
- *Estimation data*: ESTIMATION_AllEditors.dta
- *Estimation data*: ESTIMATION_AllGrants.dta
- *Estimation data*: ESTIMATION_AllNetworks.dta
- *Estimation data*: ESTIMATION_Tenure.dta