Does Collaboration Improve Female Representation in Academic Fields?

By Soohyung Lee and Benjamin A. Malin*

* Lee: Department of Economics, Sogang University, 35 Baekbeom-ro, Mapo-gu, Seoul, 04107 South Korea (soohlee@sogang.ac.kr). Malin: Federal Reserve Bank of Minneapolis, 90 Hennepin Avenue, Minneapolis, MN 55401 (e-mail: Benjamin.malin@mpls.frb.org). This research has been supported by the Basic Science Research Program of the National Research Foundation of Korea (NRF), which is funded by the Ministry of Science, ICT & Future Planning (Grant 2017R1A2B1009094). We thank George Krivorotov, Lei Ma, Sara Brandel, and Mark Fiegener for providing research assistance and comments. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

Over the past half century, gender gaps have narrowed along many dimensions, including college enrollment, labor force participation, and earnings (Goldin, 2014). However, sizable gender gaps remain elsewhere; an example that has received considerable attention is women's underrepresentation in science, technology, engineering, and mathematics (STEM). Studies have investigated various factors affecting female participation in academic fields, such as discrimination, differential preferences, math intensity, and ability beliefs (e.g., Ceci et al., 2014, and Leslie et al., 2015). In this paper, we explore a new factor — collaboration — to explain variation in female representation across fields and over time.

Our focus on collaboration stems from research suggesting that the norms of and approaches to collaboration differ across genders. In particular, men and women exhibit, on average, different preferences for teamwork (Kuhn and Villeval, 2015), which, in academic settings, manifests itself in measurable differences between genders in the number of collaborators and network size (Bozeman and Gaughan, 2011). These differences suggest that, as the extent of collaboration in academia changes over time, the forces influencing the gender gap may also change. However, as far as we know, no study directly examines the relationship between collaboration and female representation.

Collaboration, as measured by coauthorship, has increased markedly in recent decades, with the mean number of authors per article published in peer-reviewed scientific journals increasing from 2.4 in 1975 to 5.4 in 2014. That said, it is not obvious whether the effect of increased collaboration should be an increase or a decrease in female representation, as there are mechanisms that work in both directions.[1] This is the question we empirically examine in this paper.

[1] For example, some studies (e.g., Kuhn and Villeval, 2015; Niederle and Vesterlund, 2011) find that women are more likely than men to prefer working in teams or less competitive environments, which suggests that the increase in collaboration may raise female

We collect data on the female share of new Ph.D. recipients in the United States and the average number of authors per article published in each year and academic field since 1975. We find that, after controlling for field and year fixed effects, one additional author on the average published paper is associated with an increase of 2.5 percentage points in the female share. This estimate suggests that the variation in collaboration across fields in 2014 can account for 25 percent of the variation in female share that year, and the increased collaboration from 1975 to 2014 can account for 31 percent of the increased female share. Our findings are robust to analyzing fields within STEM and non-STEM fields separately and to using alternative collaboration measures. Moreover, we also explore the relationship between collaboration and racial minority shares, finding a consistent, though less robust, relationship. We discuss implications below.

Finally, our findings are relevant for the important policy debate about how to improve female representation in academic fields — and STEM fields in particular. Policy makers may consider incentives, such as funding opportunities, to promote collaboration in academia, which could reduce the gender gap.

representation. On the other hand, increased collaboration could negatively affect females if they are less likely to be included on teams (see, e.g., Sheltzer and Smith, 2014).

I. Data and Methodology

We compile a dataset with measures of female representation and collaboration across the entire academic spectrum, including both STEM and non-STEM fields, from 1975 to 2014. Using annual data from the National Science Foundation's (NSF) Survey of Doctorate Recipients, we measure female representation as the share of women among new U.S. doctorate recipients. We measure collaboration by the mean number of authors per article published in a given year and academic field, using data from Thomson Reuters' Web of Science (WoS). The resulting longitudinal dataset covers 30 academic fields for every fifth year from 1975 to 2010, as well as 2014.[2]

We use the following panel regression model to measure the relationship between the degree of collaboration in a field and its female representation:

(1) $$Y_{ft} = \propto + \beta Collaboration_{ft} + \delta_f + \theta_t + \varepsilon_{ft},$$

where Y_{ft} is the female share of Ph.D. recipients in field f and year t, $Collaboration_{ft}$ is the average number of authors per article, and ε_{ft} is an error term. Parameters δ_f and θ_t capture field and time

[2] See details in the Online Appendix.

fixed effects, respectively, while β measures the extent to which an additional coauthor per article can account for the variation in female representation across time and fields.

We note that using longitudinal data, rather than solely cross-sectional data as is typical in the literature, is advantageous. It enables us to account for the possibility that some fields innately attract more women than others and allows us to investigate important changes in female share over time. Indeed, a significant portion (between 30 and 50 percent) of the cross-sectional variation in female share in 2014 comes from differential growth rates in female share over the past 40 years.[3]

II. Graphical Examination and Results

We begin by plotting the underlying data. Figure 1 shows the relationship between the change in female share and the change in coauthorship for each academic field from 1975 to 2014.[4] We estimate the linear relationship between the change in female share and the change in the amount of coauthorship for all fields (solid line) and separately for STEM (dotted line) and non-STEM (dashed line) fields. In all cases, we find a positive relationship, statistically significant at the 5 percent level.

Panel A of Table 1 reports estimates of β for all fields and separately for STEM and non-STEM fields. For all fields (column 1), one additional author on the average published paper is associated with an increase of 2.5 percentage points in the female share. This positive relationship also holds when we estimate the regression model using STEM and non-STEM fields separately. Columns (2) and (3) show that one additional author on the average published paper is associated with an increase of 3.7 percentage points in the female share among STEM fields and 6.0 percentage points among non-STEM fields, respectively.[5]

To gauge the quantitative importance of collaboration, we calculate the share of variation in female representation that may be accounted for by collaboration. In 2014, the difference between the maximum and minimum field female shares was 57.2 percentage points, while the corresponding difference for coauthorship was 5.7. The

[3] Table S1 in the Online Appendix reports the necessary data for this variance decomposition: namely, measures of collaboration and female share by academic field in both 1975 and 2014.

[4] Because three of our fields are missing some information in 1975, Figure 1 plots data for 27 fields. The three omitted fields include (1) computer and information sciences, (2) political science, and (3) electrical, electronics, and communications engineering.

[5] The average of these two effects is not the same as the result in column (1) because the regression specification corresponding to column (1) assumes that the time effects are the same across both STEM and non-STEM fields. That is, if our model allowed for differing time effects for STEM and non-STEM fields, the estimated coefficient would be the same as the (weighted) average of the coefficients in columns (2) and (3).

variation in collaboration can thus account for about 25 percent of the variation in female share (5.7 * 2.51 / 57.2). A similar calculation made over time suggests that increased collaboration can account for 31 percent of the increased female share. In sum, not only are coauthorship and female share positively correlated, but the variation in coauthorship may account for a sizable portion of the variation in female share across academic fields.

III. Discussion

Our finding of a positive relationship between coauthorship and female share raises additional questions that are likely to be of interest to practitioners who wish to increase female representation in their fields. For example, what are the mechanisms through which these two variables are linked? And are some types of collaboration more effective at promoting female representation than others? Although our data do not allow for direct answers to these questions, we conduct some additional exercises that provide hints.

We first explore whether female representation is more closely related to collaboration *within* or *across* institutions. Working in the same institution facilitates regular face-to-face interactions, which may be particularly important for senior researchers collaborating with junior researchers. From our WoS data, we calculate two additional collaboration measures: the average number of institutions and average number of authors per institution.[6] The former captures *across*-institution collaboration, and the latter captures the *within* variety. Panel B of Table 1 shows that an increase in the number of institutions is associated with a higher female share, whereas Panel C indicates that an increase in the number of authors per institution show no correlation.[7] Our finding thus suggests that face-to-face interaction and vertical collaboration, as measured by within-institution coauthorship, may not be particularly important for female representation.

We next expand our analysis to consider how collaboration interacts with the share of minority groups other than women. To the extent that the relationship between collaboration and female share is driven by a greater demand for diversity in more collaborative fields, we would expect to find similar patterns between collaboration and racial/ethnic minorities. Because information on race/ethnicity is only available for Ph.D.

[6] See the Online Appendix for detailed information on how to construct the sample and variables.

[7] Although the coefficient estimates in Panel C of Table 1 are much higher than those in Panel B, the number of institutions varies much less than the number of authors. Thus, the two measures account for similar amounts of the variation in female share

recipients who are U.S. citizens or permanent residents, we first redo our primary empirical analysis on this sample.[8] As shown in Panel A of Table 2, the relationship between coauthorship and female representation is qualitatively unchanged. Panels B and C show that collaboration is positively associated with the share of blacks and Asian Americans, although the statistical relationship is not as strong as the one we find for female share. For the estimates that are statistically significant, we calculate the fraction of the variation in representation that can be attributed to collaboration. In 2014, collaboration accounts for 26 percent of the variation in Black share across all fields, 76 percent of Black share variation across non-STEM fields, and 65 percent of Asian share variation across STEM fields.

In contrast, Hispanic share (Panel D) is not well accounted for by collaboration. We do not have sufficient data to identify the factors that drive these differences across minority groups. It could be that, to the extent increased collaboration increases demand for diversity, this increased demand does not uniformly apply to all minorities. Alternatively, some minority groups may find collaborative environments more appealing than do others.

IV. Conclusion

Using panel data for 30 academic fields from 1975 to 2014, we find that academic fields in which coauthorship has expanded more rapidly over the past 40 years have also experienced faster growth in female representation. One plausible mechanism through which collaboration may increase female share is that women are more likely to collaborate than their male counterparts (Bozeman and Gaughan, 2011), so as coauthorship becomes the norm, women engage in collaborative projects relatively more than do men. To the extent that such interactions contribute to success in academia, as suggested by McDowell et al. (2006) and Blau et al. (2010), this effect leads to higher female retention in collaborative fields. At any rate, future research is needed to more fully explore the mechanisms underlying the observed relationship.

Our findings suggest that those policies/initiatives that focus on increasing interactions between minorities (with respect to gender and race) and others in their fields may be beneficial. Policies of this type include public funding that supports mentoring for female scientists and provides

[8]The Online Appendix describes the construction of this sample.

opportunities for collaborative efforts (e.g., the support by the Committee on the Status of Women in the Economics Profession [CSWEP] in economics), as well as private, grassroots initiatives (e.g., Leanin.org) that seek to create safe and collaborative environments in which women may thrive.

REFERENCES

Blau, Francine D., Janet M. Currie, Rachel T. A. Croson, and Donna K. Ginther. 2010. "Can mentoring help female assistant professors? Interim results from a randomized trial," *American Economic Review: Papers & Proceedings* 100, 348–352.

Bozeman, Barry, and Monica Gaughan. 2011. "How do men and women differ in research collaborations? An analysis of the collaborative motives and strategies of academic researchers," *Research Policy* 40, 1393–1402.

Ceci, Stephen J., Donna K. Ginther, Shulamit Kahn, and Wendy M. Williams. 2014. "Women in academic science: A changing landscape," *Psychological Science in the Public Interest* 15, 75–141.

Goldin, Claudia. 2014. "A grand gender convergence: Its last chapter," *American Economic Review* 104, 1091–1119.

Kuhn, Peter, and Marie Claire Villeval. 2015. "Are women more attracted to co-operation than men?" *Economic Journal* 125, 115–140.

Leslie, Sarah-Jane, Andrei Cimpian, Meredith Meyer, and Edward Freeland. 2015. "Expectations of brilliance underlie gender distributions across academic disciplines," *Science* 347, 262–265.

McDowell, John M., Larry D. Singell Jr., and Mark Stater. 2006. "Two to tango? Gender differences in the decisions to publish and coauthor," *Economic Inquiry* 44, 153–168.

Niederle, Muriel, and Lise Vesterlund. 2011. "Gender and competition," *Annual Review of Economics* 3, 601–630.

Sheltzer, Jason M., and Joan C. Smith. 2014. "Elite male faculty in the life sciences employ fewer women," *Proceedings of the National Academy of Sciences* 111, 10107–10112.

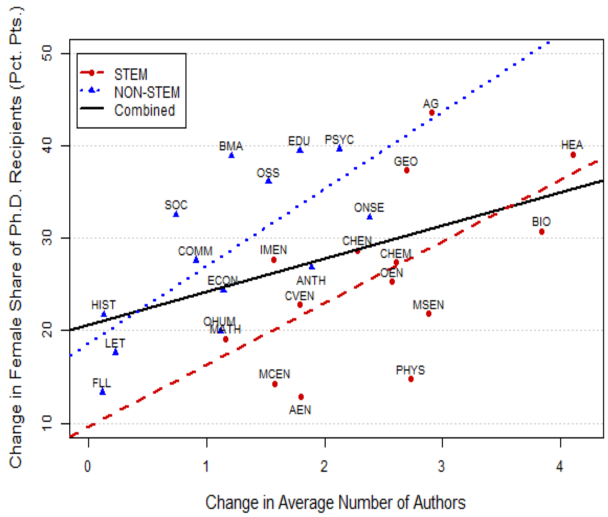

FIGURE 1. CHANGE IN COAUTHORSHIP AND FEMALE SHARE OF PHD RECIPIENTS FROM 1975 TO 2014

Note: This figure shows the relationship between the change in coauthorship and that in the female share of Ph.D. recipients from 1975 to 2014. The lines represent OLS estimates of the relationship across academic fields using all fields (solid line), STEM fields (red-dashed line), and non-STEM fields (blue-dotted line), respectively. The regression coefficient (standard error) for the various samples is 3.5 (1.6) for all, 6.6 (2.6) for STEM, and 8.3 (2.5) for non-STEM fields.

TABLE 1— FEMALE SHARE AND COLLABORATION

	All	STEM	Non-STEM
	(1)	(2)	(3)
Panel A. Mean authors	2.508***	3.656***	5.972***
	[0.811]	[1.285]	[1.330]
Panel B. Mean institutions	11.060***	16.210***	16.640***
	[2.449]	[4.758]	[3.038]
Panel C. Mean authors per institution	-0.184	0.999*	-1.925**
	[0.495]	[0.543]	[0.744]
Number of observations	266	141	125

Notes: Additional control variables include year dummies (total of 8), NSF-field specific dummies (total of 29), and a constant. Heteroskedasticity robust standard errors are reported in brackets.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

TABLE 2— FEMALE/MINORITY SHARE AND COLLABORATION

	All	STEM	Non-STEM
	(1)	(2)	(3)
Panel A. Female share	3.234***	3.497**	7.151***
	[0.888]	[1.443]	[1.220]
Panel B. Black share	0.651**	0.309	1.832***
	[0.255]	[0.517]	[0.490]
Panel C. Asian American share	-0.133	2.158**	0.0853
	[0.402]	[0.878]	[0.359]
Panel D. Hispanic share	-0.207	-0.182	0.0748
	[0.213]	[0.340]	[0.534]
Number of observations	266	141	125

Notes: This analysis used a sample that includes only Ph.D. recipients who are U.S. citizens or permanent residents. Collaboration is measured by mean number of authors. Additional control variables include year dummies (total of 8), NSF-field specific dummies (total of 29), and a constant. Heteroskedasticity robust standard errors are reported in brackets.

*** Significant at the 1 percent level. ** Significant at the 5 percent level. * Significant at the 10 percent level.

Online Appendix for

Does Collaboration Improve Female Representation in Academic Fields?

Authors: Soohyung Lee[1†], Benjamin A. Malin[2*].

Affiliations:

[1] Department of Economics, Sogang University, 35 Baekbeum-ro, Mapo-gu, Seoul 04107, South Korea; IZA; and Maryland Population Research Center and University of Maryland, College Park.

[2] Federal Reserve Bank of Minneapolis, 90 Hennepin Avenue, Minneapolis, MN 55401.

† Correspondence to: soohlee@sogang.ac.kr.

* These authors contributed equally to this work.

This file includes:

Materials and Methods
Female share and collaboration in the full-sample and restricted sample
Table: S1-S2

Materials and Methods

Doctorate Recipients by Field of Study: The annual number of doctorate recipients by field of study, gender, ethnicity/race, and U.S. citizenship status comes from the "Doctorate Recipients from U.S. Universities" reports, which are published annually by the National Science Foundation (NSF).[1] Because information by field and gender starts in 1966, our sample period ranges from 1966 to 2014. Information on ethnicity and race is available only from 1973 onward. Mark Fiegener, a project officer in the NSF's National Center for Science and Engineering Statistics, provided us with consistent data over this sample period for 34 academic disciplines.[2]

For each year and academic discipline, we calculate several variables. "Female share" is the number of female doctorates divided by the sum of female and male doctorates. We omit doctorates whose gender is missing. "Female share (US)" is the female share among doctorates with U.S. citizenship or permanent residency. We also calculate ethnic/racial shares as the number of doctorates of a given ethnicity/race (black, Hispanic, or Asian Americans) divided by doctorates whose ethnicity/race is reported. These shares are calculated from doctorates who are U.S. citizens or permanent residents.

Collaboration and Field of Study: We commissioned Thomson Reuters to construct a panel data set for us from their Web of Science (WoS) data. Specifically, for each of 251 WoS academic categories and every fifth year from 1970 to 2005 and annually from 2008 to 2014, Thomson Reuters provided the following variables: mean (median) authors per document, mean (median) institutional affiliations per document, mean (median) authors per institution per document, and the number of documents.

We placed a few restrictions on the documents that were included in the sample. First, we restricted documents to those classified as "original research," which includes journal articles and book chapters from journal editions. Thus, for example, meeting abstracts, editorial material, book reviews, reviews, proceedings papers, and corrections were excluded. Second, as a minimal quality requirement, we restricted our sample to documents that have received at least one citation. We refer to the resulting set of documents as our "full sample."

Additional sample restrictions were made to facilitate the analysis of institutional affiliation. These restrictions differ before and after 2008 because authors are not directly

[1] See National Science Foundation, "S&E Doctorate Awards," https://www.nsf.gov/statistics/doctorates/.
[2] Data from the NSF's Survey of Earned Doctorates (SED) is available online, but it is not suitable for our purposes. In particular, the WebCASPAR system (https://ncsesdata.nsf.gov/webcaspar/) does not provide data on gender, ethnicity, race, or citizenship after 2006. The SED Tabulation Engine (https://ncses.norc.org/NSFTabEngine/#WELCOME) performs such tabulations, but currently only from 2006 to 2012. The NSF provided us with the full range of annual SED data (1966–2014), using the classification system of the published tables. Note that the published tables differ somewhat from the online (WebCASPAR) data; specifically, 300+ subfields are aggregated into 47 "Detailed Disciplines" in the WebCASPAR data but into 34 "Major Fields" in the published tables.

linked to their institutions in the raw WoS data prior to 2008; rather, all institutions affiliated with any of a document's authors are listed together. For pre-2008 publications, our sample is simply restricted to papers that have at least one institutional affiliation. For post-2008 publications, we required documents to have at least one affiliation *and* also have link data (to match the author with the institution). We refer to the set of documents that meet these criteria as our "restricted sample." Thomson Reuters constructed all variables — that is, those constructed using the full sample and those that require institutional information — for this sample.

Finally, we construct the institutional affiliation of authors differently before and after 2008. Because authors are not linked to their institution before 2008, the number of institutions for a document is simply the minimum of the number of authors and listed institutions. After 2008, however, we can use linked data to deal with cases in which authors have multiple affiliations.[3] Because we are interested in whether collaboration takes place within or across institutions, we do not want to randomly assign one of the author's institutions to be their primary institution. Instead, we use the following algorithm to identify coauthors' shared institutions, which will produce a lower bound on the number of institutions per document. Namely, for each document:

1. Identify authors with only one affiliation.
    a. List those institutions and set the institution count to the number of institutions on that list.
    b. Remove any authors, including multiple-institution authors, who have affiliations on that list.
2. Identify the most frequently occurring affiliation for the remaining authors.
    a. Increment the institution count.
    b. Remove all authors associated with that institution.
3. Repeat previous step until institutions appear only once.
4. Add the number of remaining *authors* (*not institutions*) to the institution count.

Combining the Data: We combine the data on doctorate recipients and coauthorship to produce a panel data set, which covers 30 academic fields for every fifth year from 1975 to 2010, as well 2014, the latest year available. We drop 1970 data because the number of journal categories is significantly lower than in other years (229 vs. 248–251), and even though we have data for every year after 2008, we continue to use data from every fifth year for consistency with the earlier part of the sample. Our data set includes only 30 NSF academic fields (rather than 34) because we merge all education-related fields (i.e., education administration, education research, teacher education,

[3] By "multiple affiliations," we mean multiple appointments — for example, UCLA and University of Minnesota — and not just parent institutions — for example, the University of California System and UCLA. We deal with the latter case by using only the institution that is at the lowest rung, so to speak.

teaching fields, and other education). We do this because the education-related WoS journal categories do not allow for a clean mapping into separate NSF education fields.

To aggregate the 251 WoS journal categories into the 30 NSF major fields, we make use of NSF detailed subfields, which are listed at http://www.nsf.gov/statistics/2016/nsf16300/data/tab16.pdf and assign each journal category to an NSF subfield based on our best judgment. With over 300 NSF subfields, journal categories and subfields are at a similar level of disaggregation, making the mapping straightforward for most categories. The most frequent reason a mapping was less than clear-cut was that the journal category could have been classified as either biological/biomedical sciences or health sciences. Alternative mappings of these categories do not significantly affect our results. We also omit five WoS categories for which the mapping was unclear: Crystallography, Energy & Fuels, Microscopy, Nanoscience & Nanotechnology, and Spectroscopy.

We also drop a few year-*by*-WoS-category observations that display an exceptionally large number of authors because we are concerned that these outliers could mask the relationship between average collaboration and the composition of doctorates. For example, the mean number of authors per document in "Physics, Particles & Fields" jumped from less than 15 in 2010 (and all previous years) to 47 in 2014. This is an extreme outlier: across WoS categories and years, the mean and standard deviation of authors per document are 3.43 and 2.67, respectively. We drop observations that had a mean number of authors per document more than +/− 5 standard deviations from the cross-sectional mean (i.e., greater than 17 authors). This removes three observations out of 2,259. All three are in 2014: "Astronomy & Astrophysics," "Physics, Nuclear," and "Physics, Particles & Fields."

For each of our 30 academic fields and each year, we construct the weighted average of our collaboration statistics (mean authors per article, etc.) across journal categories assigned to the field, where the weight is based on the number of documents in the WoS category. Specifically, the weight for a WoS category is its number of documents relative to the number of documents for all WoS categories in the NSF field.

Finally, for analysis involving our primary measure of collaboration (i.e., number of authors), we use our full sample, so as to avoid differences in sample selection pre- and post-2008.[4] However, when we analyze the number of institutions and number of authors per institution, we must use our restricted sample. This requires care in how we compare

[4] Another reason to focus on the full sample is that WoS policy for assigning institutions appears to have changed in 1998. Documents have a "reprint/corresponding" address (until recently, just one per document) and also "researcher" addresses. The latter are the full address lists from the publication, while the former can manifest itself differently in the full text of the article. It appears that, in 1998, publications without "researcher" addresses but with a "reprint" address started to have the latter assigned to the former. This led to a big increase in our restricted sample and could cause a spurious drop in authors per document.

results across time; specifically, it is one reason we include time-fixed effects in our regressions.

Female share and collaboration in the full-sample and restricted sample

In Table 1 of the main text, we examine the relationship between female share and alternative measures of collaboration. As explained in the previous section, the variables of “mean institution” and “mean authors per institution” (Panels B and C of Table 1, respectively) are constructed using a restricted sample of journal articles, while the results for "mean authors" (Panel A) use our full sample. We show here that the use of different samples is not drive our results. To do so, we estimate the regression model reported in Panel A but use the restricted sample rather than the full sample. The results are reported in Table S2. Panel A repeats our baseline result (Panel A of the main text's Table 1), while Panel B reports the new estimate based on the restricted sample. We see that restricting the sample does not significantly alter the results.

Table S1: Female Share and Collaboration by NSF Categories

NSF Category	STEM	Female share (%)		Mean authors	
		1975	**2014**	**1975**	**2014**
	(1)	(2)	(3)	(4)	(5)
Agricultural sciences.; natural resources	Yes	4.72	48.31	2.22	5.13
Biological, biomedical sciences	Yes	23.05	53.86	2.52	6.36
Health sciences	Yes	30.95	70.02	2.77	6.88
Chemistry	Yes	10.92	38.30	2.65	5.26
Computer and information sciences	Yes	—	19.72	1.76	3.64
Geosciences	Yes	3.95	41.30	1.97	4.66
Mathematics	Yes	9.50	28.59	1.33	2.49
Physics and astronomy	Yes	5.38	20.11	2.48	5.29
Anthropology	No	35.75	62.62	1.41	3.30
Economics	No	9.61	34.01	1.31	2.45
Political science	No	—	44.13	1.21	1.85
Psychology	No	31.73	71.41	1.93	4.06
Sociology	No	30.88	63.42	1.36	2.10
Other social sciences	No	20.53	56.65	1.43	2.95
Aerospace, aeronautical, and astronautical engineering	Yes	1.42	14.25	1.88	3.68
Chemical engineering	Yes	1.08	29.77	2.16	4.44
Civil engineering	Yes	1.03	23.82	1.78	3.57
Electrical, electronics, and communication engineering	Yes	—	16.77	2.16	4.01
Industrial and manufacturing engineering	Yes	2.17	29.87	1.88	3.45
Materials science engineering	Yes	3.79	25.60	2.25	5.13
Mechanical engineering	Yes	0.62	14.87	1.84	3.42
Other engineering	Yes	2.06	27.35	2.36	4.93
Education	No	25.30	64.81	1.54	3.33
Foreign languages and literature	No	49.88	63.20	1.03	1.14
History	No	22.32	44.03	1.35	1.48
Letters	No	39.93	57.57	1.02	1.25
Other humanities	No	25.63	45.56	1.08	2.19
Business management and administration	No	3.56	42.49	1.51	2.72
Communication	No	30.30	57.94	1.40	2.31
Non-S&E fields not elsewhere classified	No	25.55	57.82	2.07	4.46

Notes: "—" indicates that there were less than 10 Ph.D. recipients in the particular field-year pair.

Table S2: Female Share and Collaboration

	All (1)	STEM (2)	Non-STEM (3)
Panel A. Mean authors	2.51***	3.66***	5.97***
	[0.81]	[1.29]	[1.33]
Panel B. Mean authors (restricted)	2.17***	2.90**	4.34***
	[0.72]	[1.20]	[1.02]

Notes: Additional control variables include year dummies (total of 8), NSF-field-specific dummies (total of 29), and a constant. Heteroskedasticity robust standard errors are reported in brackets. Significance: * 10 percent, ** 5 percent, *** 1 percent.