Dissertation authors and their advisors. Does gender matter?

Elena Chechik
European University at St.Petersburg
Center for Institutional Analysis of Science & Education
echechik@eu.spb.ru

September 22, 2021

Abstract

We investigate mentoring relationships and observe that men work with men, and women with women, more often than expected if dissertation authors and advisors assort randomly. We show significant bias to same-gender collaborations in most research areas and detect a correlation with the gender structure of the research area.

Keywords: gender; homophily; mentorship

1 Introduction

Homophily – the tendency for individuals to form relationships with those similar to themselves – occurs to varying degrees for many aspects of social life (race, gender, age, education, etc.) (McPherson, Smith-Lovin, and Cook 2001).

In academia, gender homophily is considered homophily in the co-authorship (Ghiasi, Larivière, and Sugimoto 2015), research teams (Campbell et al. 2013), and mentorship. Studies of mentorship have looked for associations with publications productivity (Gaule and Piacentini 2018), educational satisfaction (Seeber and Horta 2021), and "leaking" from academia (Shaw and Stanton 2012). Most works focused on a single research field, e.g., economics (Hilmer and Hilmer 2007) or chemistry (Gaule and Piacentini 2018), and the few studies that encompassed several fields did not report the variation of homophily throughout them (Sheltzer and Smith 2014). Schwartz, Liénard, and David (2021) presents a wide range of fields and shows a weak positive correlation between homophily and the proportion of female mentors or trainees.

We investigate mentorship between dissertation authors (trainees) and their mentors in Russia. We test whether the probability of working the same genders exceeded the proportion expected if mentors matched to trainees randomly. We have considered the problem of the Wahlund effect following Holman and Morandin (2019) to avoid overvaluing homophily within heterogeneous areas. To do this, we separately provide estimates for general research fields and their subfields and for Ph.D. and doctoral dissertations.

We detect gender homophily almost in all general research fields in Russia. Moreover, homophily in narrow research fields has positively correlated with the share of female mentors. Our estimation of gender homophily for mentoring relationships is the first on this type of data for Russia. Considering Russian specifics allows us to discuss some potential causes of gender homophily in various sciences and its consequences.

2 Data

We analyzed data from the Higher Attestation Commission website (Cyrillic: ВАК, Latin: VAK). The database integrate information on Ph.D. dissertations (Cyrillic: кандидатская диссертация) and on the next degree level attestation thesis – doctoral dissertations (Cyrillic: докторская диссертация). For more about data collection and availability by years, see Guba, Sokolov, and Sokolova (2020).

The VAK database explicitly indicated a mentorship relationship. We inferred trainee and mentor genders from the ending of patronym (Cyrillic: отчество). Gender inference was available for 97.2% of trainees and 87,7% of mentors. We excluded individuals who did not have patronym at all or have patronym with a high probability of association with both genders. Also, we excluded dissertations with more than one mentor (8,1%).

It is possible to reduce sample attrition by improving gender detection by the algorithm Panchenko and Teterin (2014), which estimates the probability that a typical first and second name user identifies as male or female based on social media data. Also, we can include multiple mentor cases in the future analysis by analogy with the authorship position typically used in the co-authorship studies (Holman and Morandin 2019).

The resulting dataset included 45608 dissertations, where 32972 is Ph.D. (between 2012 and 2016), and 12636 is doctoral (between 2008 and 2016). It is a multidisciplinary database with a wide range of research areas which, according to the VAK standard, consist of general fields (Table 1).

There is not a significant bias by general research fields in the dataset after excluding part of VAC database observations through comparing with Library of Dissertation Abstracts by the Book Chamber (Cyrillic: Летописи авторефератов диссертации Книжной палаты) (Guba, Sokolov, and Sokolova 2020). Comprehensive data on the authors of the dissertations and their mentors by gender are not available from the Book Chamber. However, the gender composition of our dataset correlates with aggregated data from the

Table 1: Dissertations by general research fields

	PhD (N=32972)	Doctoral (N=12636)	Overall (N=45608)
Agriculture	1030 (3.1%)	604 (4.8%)	1634 (3.6%)
Art Studies	292 (0.9%)	94 (0.7%)	386 (0.8%)
Biology	2139 (6.5%)	915 (7.2%)	3054 (6.7%)
Chemistry	1599 (4.8%)	314 (2.5%)	1913 (4.2%)
Culturology	233 (0.7%)	174 (1.4%)	407 (0.9%)
Earth Sciences	1430 (4.3%)	335 (2.7%)	1765 (3.9%)
Economics	4078 (12.4%)	1945 (15.4%)	6023 (13.2%)
Education	1861 (5.6%)	740 (5.9%)	2601 (5.7%)
History	947 (2.9%)	540 (4.3%)	1487 (3.3%)
Law	1387 (4.2%)	553 (4.4%)	1940 (4.3%)
Medical Sciences	3613 (11.0%)	2242 (17.7%)	5855 (12.8%)
Philology	1816 (5.5%)	739 (5.8%)	2555 (5.6%)
Philosophy	515 (1.6%)	366 (2.9%)	881 (1.9%)
Physics & Math.	2909 (8.8%)	614 (4.9%)	3523 (7.7%)
Political Science	413 (1.3%)	183 (1.4%)	596 (1.3%)
Psychology	634 (1.9%)	181 (1.4%)	815 (1.8%)
Sociology	468 (1.4%)	220 (1.7%)	688 (1.5%)
Technical science	7608 (23.1%)	1877 (14.9%)	9485 (20.8%)

Table 2: Dissertations by mentor‘s and trainee’s gender

	PhD (N=32972)	Doctoral (N=12636)	Overall (N=45608)
Mentor			
female	10812 (32.8%)	2884 (22.8%)	13696 (30.0%)
male	22160 (67.2%)	9752 (77.2%)	31912 (70.0%)
Trainee			
female	16212 (49.2%)	5905 (46.7%)	22117 (48.5%)
male	16760 (50.8%)	6731 (53.3%)	23491 (51.5%)

Russian Statistical Office.

We test whether mentors collaborate with same-gender trainees while controlling for the Wahlund effect (Holman and Morandin 2019). The Wahlund effect will tend to inflate the frequency of same-gender mentorship if the data is composed of two or more disconnected subsets with different trainees and mentors’ gender ratios. We show significant differences in gender composition across general research fields. Therefore, we split these general research fields into narrow research disciplines as far as the data allow us. For example, Physics, Medicine, and Technical Sciences have a quite reasonable standard split into narrow research disciplines. However, for Economics, it looks like even textual analysis of dissertation titles does not help split huge massive into narrow research fields (supplementary data accessed at link).

Another way to control for the Wahlund effect is to look at doctoral dissertations and Ph.D. separately. It will allow us to separate subsets of people at different levels of an academic career. The gender structure of mentorship relations by types of dissertations (Table 2) shows that an analysis of Ph.D. and doctoral dissertations separately can help to obtain more accurate estimates of gender homophily.

For the Ph.D. subset, trainees are distributed by gender evenly – 49.2% of them are female, and 50.8% are male. The doctoral subset already has shifted toward a higher proportion of males, but this is still a gap of 6.6%, which does not look impressive compared to the gaps we see among mentors. Among the Ph.D. subset's mentors, 32.8% are female, while females are only 22.8% through doctoral dissertation mentors.

3 Methods

To quantify gender homophily in mentoring relationships, we calculated the degree to which same-gender mentoring relationships exceeded the proportion expected if trainees matched to mentors randomly.

Following Schwartz, Liénard, and David (2021), we measured homophily as the degree to which same-gender mentoring relationships exceeded the proportion expected if trainees matched to mentors randomly. Homophily was first calculated separately for men and women:

$$\text{homphily}_F = Pr(\text{trainee}_F|\text{mentor}_F) - Pr(\text{trainee}_F) \tag{1}$$

$$\text{homphily}_M = Pr(\text{trainee}_M|\text{mentor}_M) - Pr(\text{trainee}_M) \tag{2}$$

Overall homophily was then computed as their sum, weighted by the total number of training relationships with mentors in each group:

$$\text{homphily}_{Total} = Pr(\text{mentor}_F) * \text{homphily}_F + Pr(\text{mentor}_M) * \text{homphily}_M \tag{3}$$

Positive values indicate that students and mentors of the same gender tend to work together, while negative values indicate that students and mentors of different gender tend to work together. A value of 0 means that students of any gender have an equal chance of training with mentors of any gender.

For comparisons of homophily across narrow research areas, only fields with at least 50 dissertations were included (319/385 fields with any women mentors). Data from research areas with less than 50 dissertations were included in Supplementary data, Table S1 (accessed at link).

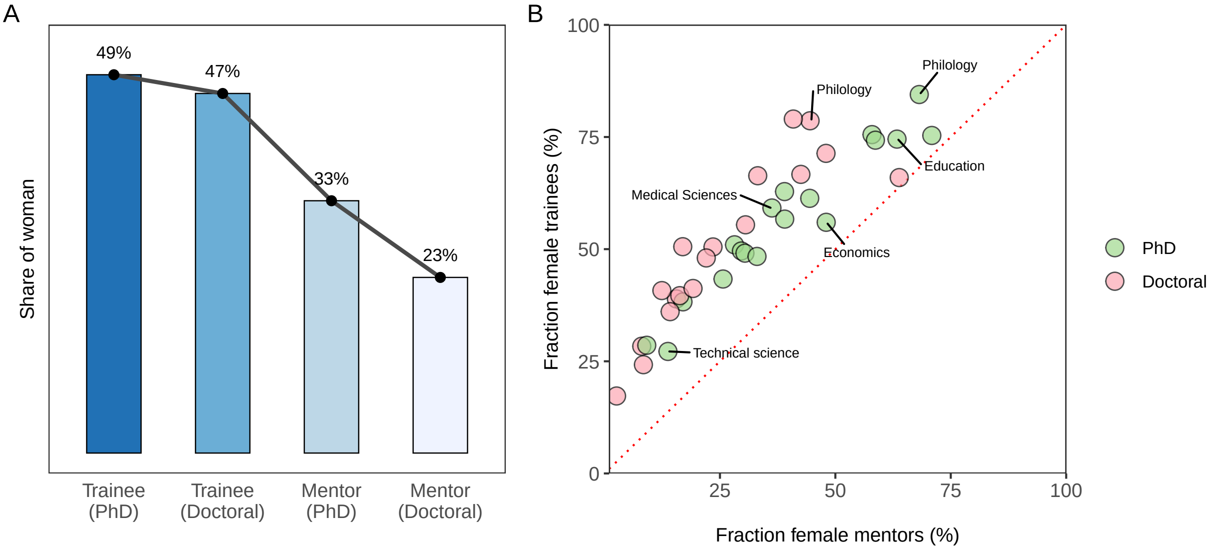

Figure 1: (A) Share of women according on their role in the dissertation production. (B) Share of female mentors less than female trainees across major fields

4 Results

4.1 Academic mentorship framework

Previous investigations show that the research areas are not homogeneous by gender - there are more male areas like STEM and more female areas like Education (Holman, Stuart-Fox, and Hauser 2018). Possible explanations vary from the academic environment (Moss-Racusin et al. 2012) to personal characteristics and behavioral patterns (Bleidorn et al. 2016).

In Russia, the proportion of women among trainees has varied by general fields – from 17% in Physics to 84% in Philology. However, the fraction of female mentors is always smaller than the fraction of female trainees, even in “female” areas like Education and Philology (Fig 1B). This result is expected and consistent with a wide range of investigations into the attrition of women across the academic career track (sometimes known as the “leaky pipeline”) (Stack 2004).

Fig 1A shows that the proportion of women is lower at progressively later stages of the academic career track, from Ph.D. trainee to Doctoral trainee to Ph.D. mentor to Doctoral mentor (supplementary data to provide the pipe accessed at link). Although this result indicates the population of academic mentors remains skewed towards men in all research areas, it does not indicate whether female trainees are continuing to mentorship positions at the same rate as male trainees. In addition, it does not indicate whether structural gender biases that affect women as mentors indirectly affect the retention of their trainees in academia. Fig 1A-B present just a static snapshot of the moment and tell us nothing about processes.

Every point in Fig 1B presents one of the general research areas, and if we expand

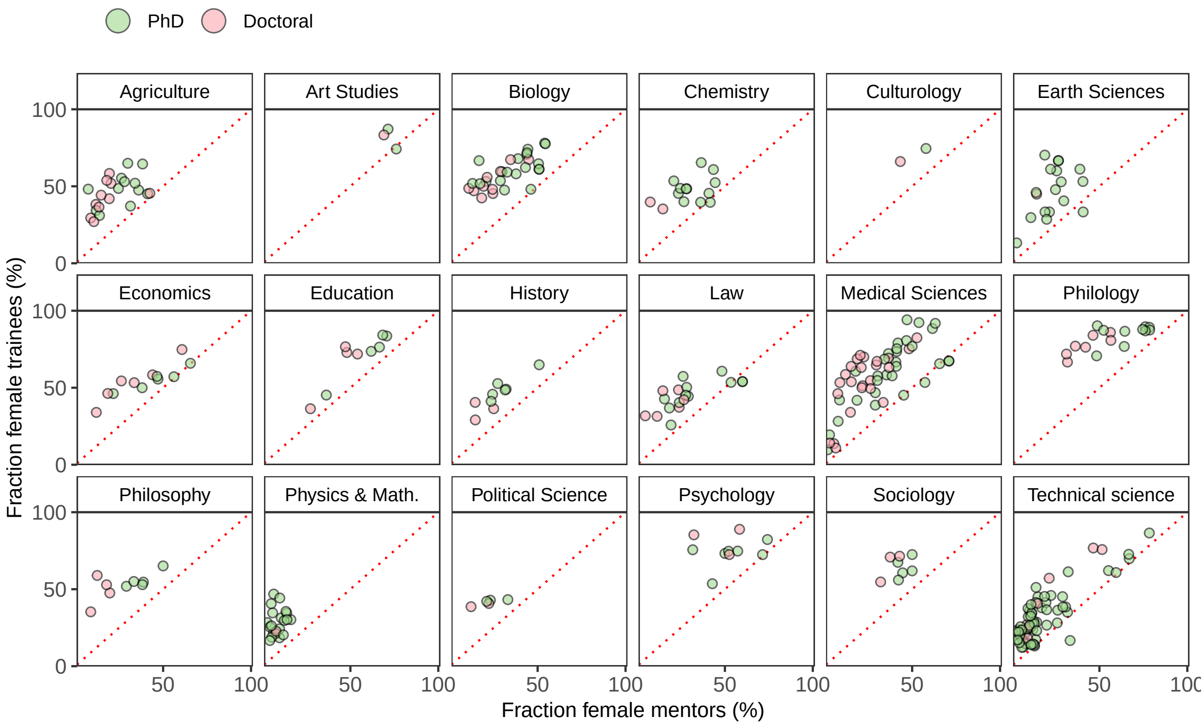

Figure 2: Share of female mentors less than female trainees across narrow research areas

it to the narrow research areas, we receive Fig 2. For example, for the general research area of Art Studies, we get two narrow research areas for Ph.D. dissertations and one for doctoral (list of narrow research areas accessed at link). Fig 2 confirms for us the persistence of the tendency that fraction female mentors are smaller than fraction female trainees. We find different dispersion of narrow research areas into a particular general research area. Narrow research areas from Physics and Math.'s are tightly concentrated in one space, while narrow areas from Medicine have extended distribution along the axes. Fig 2 demonstrate the importance of analyzing homophily not only in general areas but necessarily considering their internal structure and distributional characteristics in narrow areas (Holman and Morandin 2019).

Another point is that the proportion of women is lower at progressively later stages of the academic career track, and the depth of this decline varies by field (Fig 3). Culturology and Art Studies have wide confidence intervals for each stage. They do not allow us to speak with confidence about the continuation of the general trend towards a decrease in the proportion of women. Art Studies, in general, is close to balance - the proportion of women on all stages is more stable than other general fields. Psychology and Sociology show an increase in the proportion of female doctoral dissertation trainees compared to the Ph.D. trainees. However, when we look at the confidence interval of bars, we can check the proximity of these shares, which is also detected in other fields, such as Economics.

Thus, in Russia, several characteristics do not replicate the patterns of mentoring relationships evident in data from other countries. However, the key trends seem to be

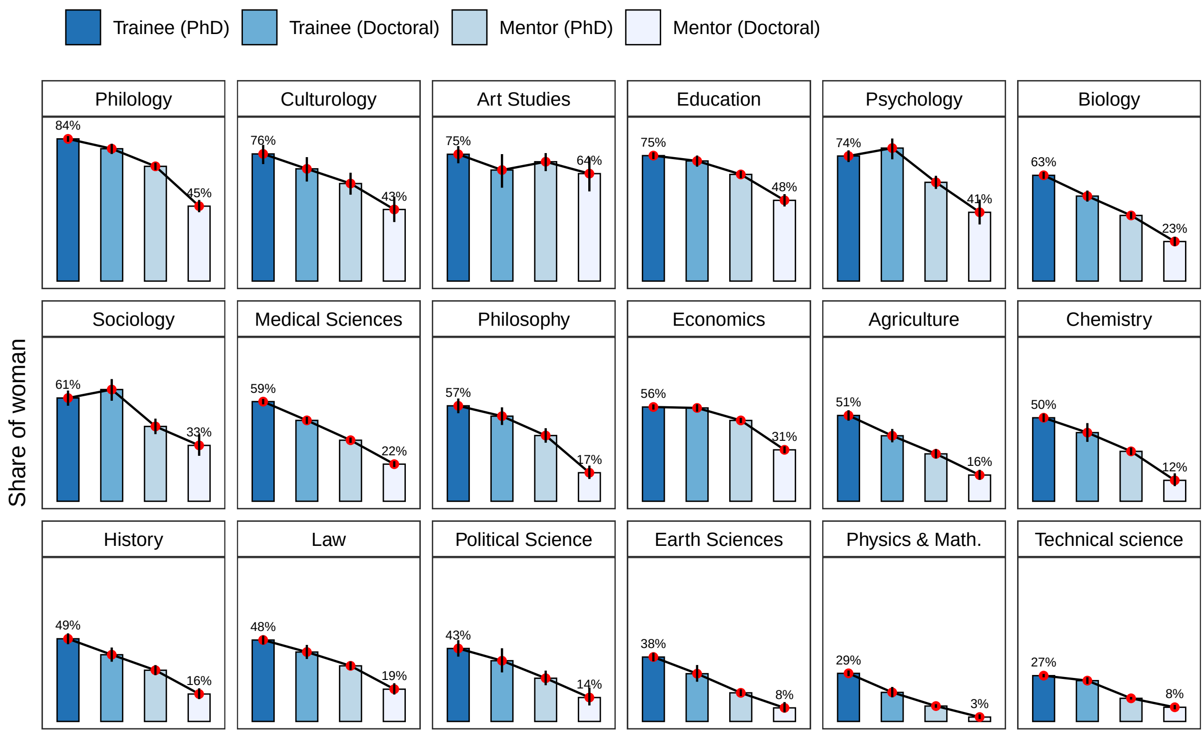

Figure 3: Share of women according on their role in the dissertation production – is significantly lower at progressively later stages of the academic career track almost in all major fields

Table 3: Dissertations by mentor's and traine's gender

	PhD (N=32972)	Doctoral (N=12636)	Overall (N=45608)
female mentor + male thainee	3302 (10.0%)	927 (7.3%)	4229 (9.3%)
male mentor + female thainee	8702 (26.4%)	3948 (31.2%)	12650 (27.7%)
both female	7510 (22.8%)	1957 (15.5%)	9467 (20.8%)
both male	13458 (40.8%)	5804 (45.9%)	19262 (42.2%)

universal and persistent regardless of data origin (Schwartz, Liénard, and David 2021). Focusing on the noticeable gaps in the proportion of women among trainees and mentors for Ph.D. and doctoral dissertations and analyzing two levels of research areas split - general and narrow allow us to make a more accurate assessment of homophily.

4.2 Gender homophily in mentorship

A summary of the types of collocations we observe in our data was presented in Table 3.

The minority of collaborations is the «female advisor + male trainee» type: for Ph.D. dissertations is 10%, for doctoral – 7.3%. The most common type of collaboration is the «both male» type – 40.8% for Ph.D. and 45.9% for doctoral dissertations. In addition, we see that total female mentors primarily supervise female trainees (20.8% of all collaborations for both types of dissertations). In comparison, the "female advisor

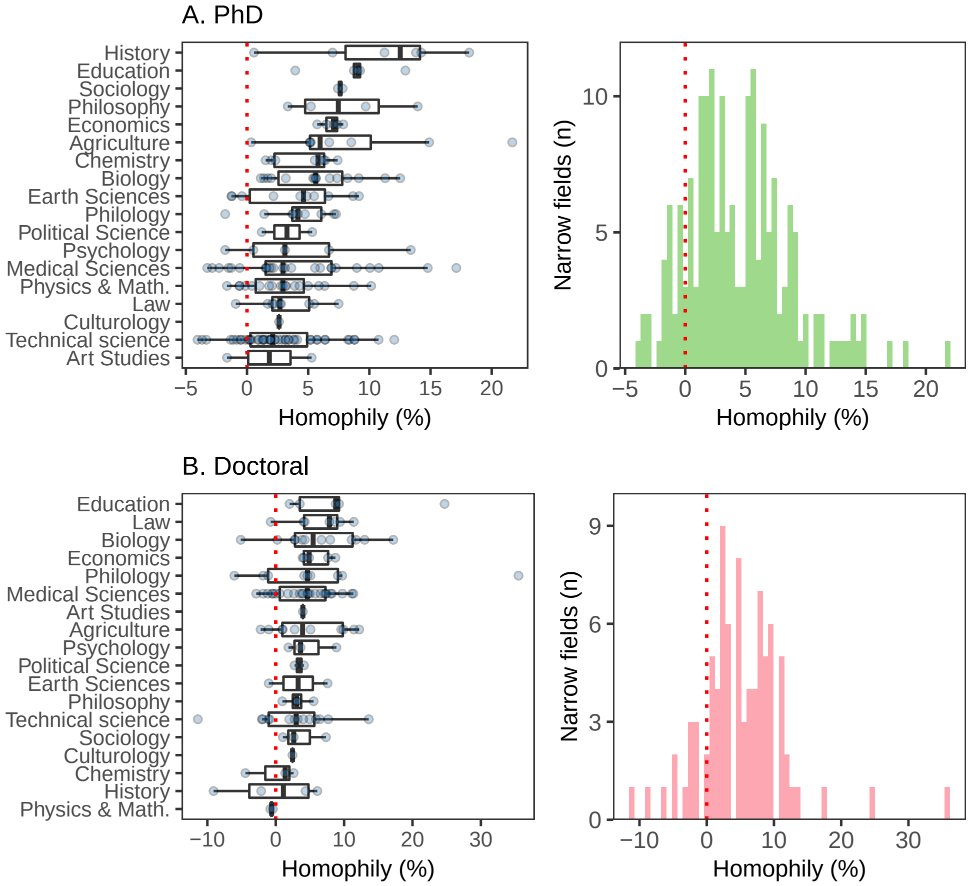

Figure 4: Distribution of narrow research areas within general research areas by homophily

+ male trainee" combination is only 9.3%.

We quantified gender homophily for narrow research areas following Holman and Morandin (2019) (see Methods). Homophily is positive when same-gender mentors and trainees work together more often than expected, negative when opposite-gender mentors and trainees work together more often than expected (heterophily), and equal to zero when mentors and trainees assort randomly for gender.

Fig 4A shows the degree of homophily for each narrow research area sorted by general research areas to which they belong. The boxplot median values are positive for all general areas. Nevertheless, we can also see particular points on the left from the dotted red line– these are narrow research areas characterized by negative homophily.

While homophily was detected for most general areas, no significant correlation with the gender structure of this area was found (Fig 6A). However, if we consider the level of narrow areas, the positive correlation strengthens and becomes significant. Although some estimates of homophily at this level may be imprecise due to small sample sizes.

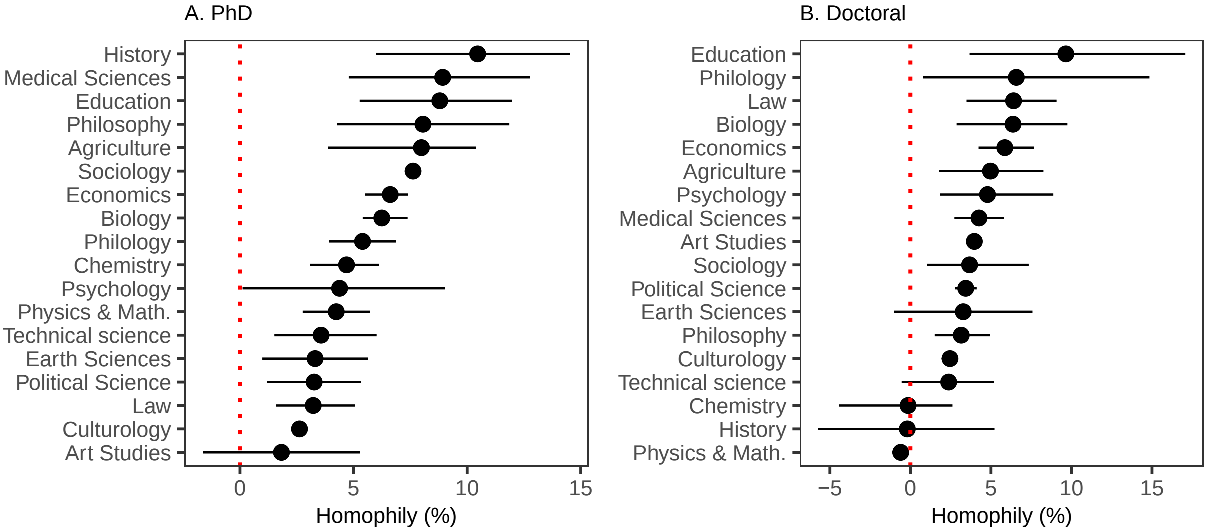

Figure 5: Homophily across general research areas based on narrow research areas

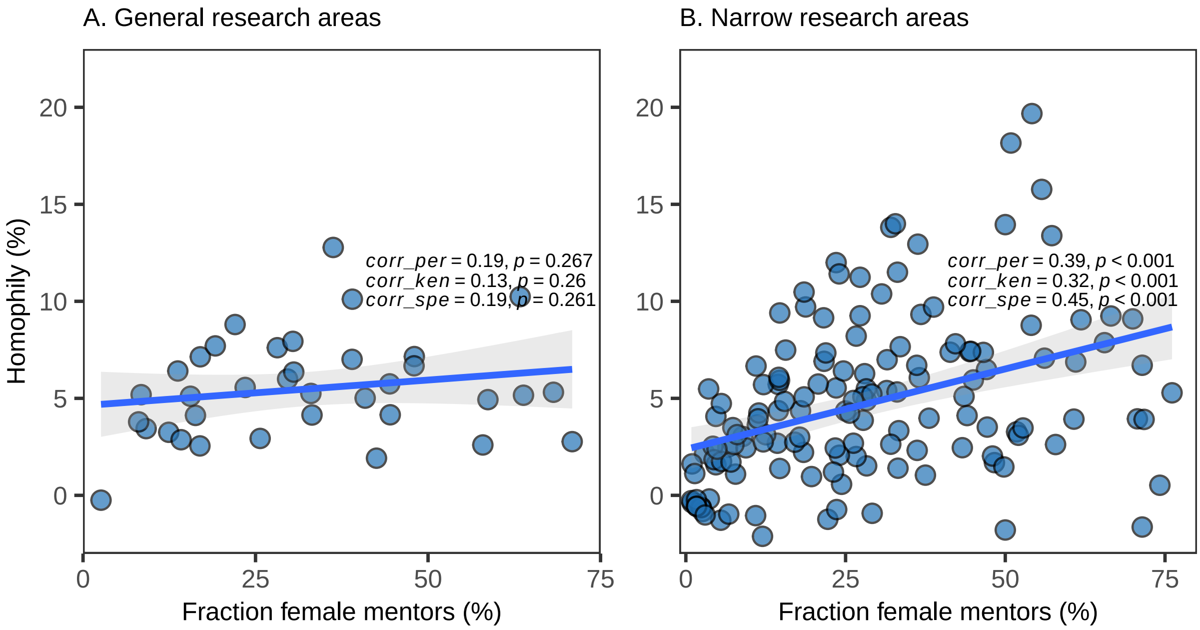

Figure 6: Homophily and share of female mentors across narrow and general research areas

Table 4: Homophily across general research areas

	PhD				Doctoral			
	Diss. (n)	Female trainees (%)	Female mentors (%)	Homoph. (%)	Diss. (n)	Female trainees (%)	Female mentors (%)	Homoph. (%)
Agriculture	1030	51	28	8	604	39	16	5
Art Studies	292	75	71	3	94	66	64	5
Biology	2139	63	39	7	915	50	23	6
Chemistry	1599	50	30	6	314	41	12	3
Culturology	233	76	58	3	174	67	43	2
Earth Sciences	1430	38	17	7	335	28	8	4
Economics	4078	56	48	7	1945	55	31	6
Education	1861	75	63	10	740	71	48	7
History	947	49	30	8	540	40	16	4
Law	1387	48	33	5	553	41	19	8
Medical Sciences	3613	59	36	13	2242	48	22	9
Philology	1816	84	68	5	739	79	45	4
Philosophy	515	57	39	10	366	51	17	3
Physics & Math.	2909	29	9	3	614	17	3	0
Political Science	413	43	26	3	183	36	14	3
Psychology	634	74	59	5	181	79	41	5
Sociology	468	61	44	6	220	66	33	4
Technical science	7608	27	14	6	1877	24	8	5

References

Bleidorn, Wiebke, Ruben C Arslan, Jaap JA Denissen, Peter J Rentfrow, Jochen E Gebauer, Jeff Potter, and Samuel D Gosling. 2016. "Age and Gender Differences in Self-Esteem—a Cross-Cultural Window." *Journal of Personality and Social Psychology* 111 (3): 396.

Campbell, Lesley G, Siya Mehtani, Mary E Dozier, and Janice Rinehart. 2013. "Gender-Heterogeneous Working Groups Produce Higher Quality Science." *PloS One* 8 (10): e79147.

Gaule, Patrick, and Mario Piacentini. 2018. "An Advisor Like Me? Advisor Gender and Post-Graduate Careers in Science." *Research Policy* 47 (4): 805–13.

Ghiasi, Gita, Vincent Larivière, and Cassidy R Sugimoto. 2015. "On the Compliance of Women Engineers with a Gendered Scientific System." *PloS One* 10 (12): e0145931.

Guba, Katerina, Mihail Sokolov, and Nadezhda Sokolova. 2020. "[RUS]dinamika Dissertacionnoj Industrii v Rossii: 2005-2015 Gg. Izmenil Li Novyj Institucional'nyj Trafaret Akademicheskoe Povedenie?" *Ekonomicheskaya Sociologiya* 21 (3).

Hilmer, Christiana, and Michael Hilmer. 2007. "Women Helping Women, Men Helping Women? Same-Gender Mentoring, Initial Job Placements, and Early Career Publishing Success for Economics PhDs." *American Economic Review* 97 (2): 422–26.

Holman, Luke, and Claire Morandin. 2019. "Researchers Collaborate with Same-Gendered Colleagues More Often Than Expected Across the Life Sciences." *PloS One* 14 (4): e0216128.

Holman, Luke, Devi Stuart-Fox, and Cindy E Hauser. 2018. "The Gender Gap in Science: How Long Until Women Are Equally Represented?" *PLoS Biology* 16 (4): e2004956.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27 (1): 415–44.

Moss-Racusin, Corinne A, John F Dovidio, Victoria L Brescoll, Mark J Graham, and Jo Handelsman. 2012. "Science Faculty's Subtle Gender Biases Favor Male Students." *Proceedings of the National Academy of Sciences* 109 (41): 16474–79.

Panchenko, Alexander, and Andrey Teterin. 2014. "Detecting Gender by Full Name: Experiments with the Russian Language." In *International Conference on Analysis of Images, Social Networks and Texts*, 169–82. Springer.

Schwartz, Leah P, Jean Liénard, and Stephen V David. 2021. "Impact of Gender on the Formation and Outcome of Mentoring Relationships in Academic Research." *arXiv Preprint arXiv:2104.07780*.

Seeber, Marco, and Hugo Horta. 2021. "No Road Is Long with Good Company. What Factors Affect Ph. D. Student's Satisfaction with Their Supervisor?" *Higher Education Evaluation and Development.*

Shaw, Allison K, and Daniel E Stanton. 2012. "Leaks in the Pipeline: Separating Demographic Inertia from Ongoing Gender Differences in Academia." *Proceedings of the Royal Society B: Biological Sciences* 279 (1743): 3736–41.

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 (28): 10107–12.

Stack, Steven. 2004. "Gender, Children and Research Productivity." *Research in Higher Education* 45 (8): 891–920.