Networks in the market for researchers*

Flavio Hafner†

Christoph Hedtrich‡

December 9, 2024

Abstract

We study the role of collaboration networks in the U.S. market for early-career scientists. In 2014, about 21% of PhD graduates found their first job at a university where their advisor had a co-author—a threefold increase from 1990. Such connections more than double the probability of finding a job at a given university, even after controlling for unobserved factors at the level of PhD class–hiring university pair and for the topical fit between the PhD graduate's dissertation and research at the hiring university. We observe a sizable citation premium for graduates placed through the advisor's network. However, when comparing graduates hired at the same university, the productivity premium is fully explained by public information on the productivity of the PhD graduate and their advisor at the time of graduation. These results suggest that PhD advisors do not convey additional private information about match quality or the PhD graduate's productivity in their collaboration network. While the estimated role of network-based placements has doubled from 1990 to 2014, the citation premium for network hires has remained constant.

JEL Codes: J24, J63, M51

*We thank Oskar Nordström-Skans, Bas Hofstra and various seminar and conference participants for their valuable comments. We thank Mona Dapfer and Quan (Jack) Cheng Xie for their excellent research assistance. We acknowledge financial support by Jan Wallanders och Tom Hedelius Stiftelse/Tore Browaldhs Stiftelse through program grant P21-0134.

†Netherlands eScience Center, f.hafner@esciencecenter.nl

‡Uppsala University, christoph.hedtrich@nek.uu.se

The size of research teams and inter-university collaboration networks has grown substantially over the last decades (Fortunato et al., 2018; Wuchty, Jones and Uzzi, 2007; Jones, 2009; Freeman, Ganguli and Murciano-Goroff, 2014). Because of the importance of social connections in academic labor markets (Rose and Shekhar, 2023; Hadlock and Pierce, 2021), increasing inter-university collaboration can reshape how the labor market for early-career researchers works. However, existing evidence on this topic is limited, often focusing on specific fields or periods. We aim to fill this gap by providing systematic evidence on the role of collaboration networks in hiring PhD graduates at universities. Our analysis spans all research fields and covers the increase in collaboration from 1990 to 2014.

One reason for the reliance on networks in hiring is that social connections can solve an information problem (Rees, 1966): Because neither the firm nor the job candidate can observe all relevant aspects of the match, social connections can convey information that reduces search costs. Nevertheless, whether network hiring leads to better matches remains unclear. In an optimistic view, referrers promote candidates that are a good fit for the job, benefitting workers and firms by increasing their options and match productivity. In a pessimistic view, referrers may exploit their relationship with the employer to promote underperforming candidates, resulting in nepotism and inefficient matches. In the market for early-career researchers, PhD advisors are well-positioned to act as referrers, leveraging their private knowledge of graduates and their connections with hiring universities. Thus, whether network hiring through PhD advisors' co-author connections outperforms other hiring channels in selecting productive matches is theoretically and empirically unclear.

To address this, we examine whether network hires are more productive than their peers and assess whether connections convey additional private information beyond what is publicly available at the time of PhD graduation. Understanding the role of co-author connections in hiring practices provides insights into their impact on match quality and aggregate research productivity—an increasingly relevant issue as advisors' networks expand.

To investigate whether the rising number of co-author connections systematically influences PhD graduates' placement at universities and their productivity outcomes, we build a database on the research careers of PhD graduates. We track graduates starting from their dissertations (Proquest, 2023) and follow their affiliations and research output through publications in the Microsoft Academic Graph (Microsoft, 2021).

We estimate a multinomial logit model to show that advisors' collaboration networks predict where PhD graduates secure their first jobs. Specifically, having a connection to a university through the advisor's network doubles the probability of a graduate being hired by that university. These results are precise and robust, and they account for systematic hiring patterns at the PhD class $\times$ hiring university level as well as for the similarity between the graduate's research topics and those of the hiring universities.

The estimated effect of a single connection on placement remains stable, even as the number of connected universities and the overall influence of networks have grown. Between 1990 and 2014, the predicted impact of advisors' networks in placing graduates has doubled. Notably, systematic matching predicted by connections explains half of all placements at connected universities

To estimate the role of co-author connections in matching PhD graduates with universities, we build on the approach of Eliason et al. (2022). Their linear model and our multinomial logit model are very similar, except for the functional form assumption; an advantage of the latter is that it is consistent with matching models like Choo and Siow (2006).

Our estimates of the effect of advisors' co-author connections on matching have a causal interpretation, assuming that no unobserved factors simultaneously determine the advisor's network connections and the graduate job placement relative to their peers. These factors must be unobserved to analysts but known to market participants even without network connections. This assumption is consistent with those in many existing referral studies (Burks et al., 2015; Brown, Setren and Topa, 2016; Kramarz and Thesmar, 2013), with one exception being Rajkumar et al. (2022) who conduct an experiment on LinkedIn. In line with a causal interpretation, we show that placebo advisor connections

do not predict matches and that the estimated connection effect diminishes with time since the last collaboration.

After demonstrating that co-author connections predict at which universities PhD graduates continue their research careers, we then focus on the productivity of network hires. We find that graduates hired by a university within their advisor's co-author network produce more output after graduation than non-connected hires. However, theory implies that the screening benefits of network hiring do not necessarily increase productivity unconditionally, but rather conditional on public information at the time of hiring. We thus examine the extent to which such information can account for the productivity premium of network hires.

When comparing graduates within the same PhD class, we find that those placed at a university where their advisor has a co-author receive 27% more citations than those with a non-connected first placement. We then add controls for publicly available information about the graduate's productivity—the number of pre-graduation papers weighted by the citation impact of their venues—and for the advisor's productivity, measured by their citation decile within the field. Accounting for these factors reduces the productivity premium of network hires by half. The remaining productivity premium of network hires, after accounting for publicly available information, reflects both selection on graduate ability and the effects of placement at different universities.

To eliminate the role of the first placement itself, we then compare graduates with the same first post-PhD affiliation. Among PhD graduates hired at the same university, connected hires produce 11% more citations than those without a connection. Controls for public information explain the entire productivity premium of network hires, suggesting that connected hires are positively selected in terms of expected productivity compared to their colleagues at the same hiring university but that this selection does not reflect private information. The approximately zero premium of network hires holds in terms of several additional measures, including the number of papers they write, the number of co-authors at their first affiliation, and the probability of staying active in research. These results show that network hires outperform on average, but there is no evidence that

co-author connections reveal who performs better post-PhD compared to other hiring channels.

Our paper contributes to the literature on the role of networks in academic labor markets. A novel feature in our paper is the use of natural language processing to control for the topical fit between PhD graduates and hiring universities.

The most closely related study is by Rose and Shekhar (2023), who document the importance of the advisor's network centrality for placement. In contrast to their paper, we focus on individual connections between authors and examine the productivity of connected graduates. Further, they focus on the impact of the advisor's centrality on a PhD graduate, but here, the focus is on the performance of network hiring compared to other hiring channels. Our results are similar to Baruffaldi, Visentin and Conti (2016) from Science&Engineering PhD students—one career stage earlier than the one we study. In addition, there is a tendency for finance departments to repeatedly hire from the same PhD programs (Hadlock and Pierce, 2021). This suggests that the total extent of network hiring through social connections is likely larger than what we document for co-author connections. The focus on co-author connections, though, is motivated by the rise of teamwork in research.

Malmgren, Ottino and Nunes Amaral (2010) and Ma, Mukherjee and Uzzi (2020) highlight that mentors and PhD advisors specifically are predictive of success in academic careers. Our evidence is consistent with this, while we highlight the specific role co-author connections of PhD advisors play in the placement of early career researchers.

The academic environment has been shown to be important for research output (Azoulay, Liu and Stuart, 2017; Way et al., 2019). Similarly, Smeets, Warzynski and Coupé (2006) document the placement and subsequent performance of Economics PhD graduates based on which PhD program they come from. These results highlight the importance of considering the academic environment in the analysis of post-PhD outcomes, in order not to confound the role of network hiring with the general academic environment.

Terviö (2011) and Clauset, Arbesman and Larremore (2015) document systematic

hiring flows at the level of PhD granting × hiring university pair. Their results highlight the need to flexibly account for unobserved hiring determinants at the level of PhD class × hiring university pair to isolate the role played by individual co-author connections.

Our paper further contributes to the empirical literature on networks in the broader labor market (Dustmann et al., 2016; Brown, Setren and Topa, 2016). San (2022) estimates a matching model to study the role of parents' co-workers. He exploits the structure of the matching model to uncover worker productivity from observed wages and allocations. An advantage of our setting is that we can document the productivity of network hires directly. Additionally, our focus is not on the labor market outcomes of individual workers, but rather on the benefits of network hiring from the employers' perspective. We follow a similar approach to Burks et al. (2015) who document broadly similar patterns with productivity data in a general labor market setting. In our setting, we can observe pre- and post-hiring productivity measures that are specific to the job. This allows us to approximate productivity information about candidates available to hiring committees.

Finally, our results relate to theories about networks in the labor market: Our results indicate that the network does not reveal private information about match productivity (Dustmann et al., 2016; Simon and Warner, 1992; Galenianos, 2013). Further, even though we find that the productivity of the advisor can explain part of the productivity premium of network hires, our evidence is not in line with homophily-based models (Montgomery, 1991): There, network hiring enables firms to hire more productive workers on dimensions that are otherwise unobservable (Hensvik and Skans, 2016). However, in our setting, the advisor's productivity is publicly observable to all universities—independently of the network connection. This implies that all universities can use the advisor's productivity as a signal for the graduate's future output.

1 Theoretical Background

We review the main theoretical predictions of different mechanisms underlying network hiring[1]. We focus on two primary outcomes: First, whether a match is formed between graduate i and university j. Second, the productivity of connected hires compared to hires from other channels. For this purpose, consider the following decomposition of productivity into unobservable components:

$$y_{i,j} = a_i + b_j + f_{i,j}, \tag{1}$$

where a_i refers to the candidates' "ability", b_j the universities' "quality" and $f_{i,j}$ is a match-specific productivity component.

The simplest theory of network hiring is that information about the presence of vacancies and candidates is transmitted through the network without giving any direct information about any component of $y_{i,j}$ or other match-relevant payoffs. Thus, information transmitted through networks may simply reduce search frictions and directly increase the matching probability within the network.

In addition, social connections can help screen candidates. On one hand, connections may reveal private information about match-specific productivity $f_{i,j}$ (Simon and Warner, 1992; Galenianos, 2013; Dustmann et al., 2016). These theories assume that employers can screen connected workers with higher precision. Additional private information about connected hires implies higher productivity above and beyond what would be expected based on public information. In many situations, this also implies a higher probability of matching.

On the other hand, Montgomery (1991) considers the screening aspect when hiring workers not through referrals directly, but through connections alone. The key idea is that the firm knows who their best workers are, and if there is homophily in connections, the friends of these good employees are more productive on average. These productivity gains can stem from the candidate's unobserved ability a_i and the match-specific term

[1]We largely build upon Topa (2011) in this review, who also discuss predictions regarding wages in further detail.

$f_{i,j}$ and suggest a situation in which, conditional on public information, network hiring is prevalent, selective, and associated with higher average productivity.

In contrast, Cornell and Welch (1996) describe a situation of "screening discrimination". Their model allows for discrimination to occur even when the underlying quality distribution of candidates is known, and there is no match-specific productivity $f_{i,j}$ to learn about. The key assumption is that own-group candidates are easier to screen. This implies that from the perspective of a given employer, the posterior belief about own group candidates has fatter tails and thus implies that the candidate they believe to be most qualified is more likely to come from their own group. This happens even when all market participants know that the unconditional distribution of qualities is the same across groups. Applying their insights to network hiring suggests that hiring connected PhD graduates can be common if universities obtain additional information about candidates' quality a_i. The conditional posterior belief of universities about connected graduates has fatter tails despite the average network hire across all universities being of the same quality as non-network hires. This mechanism rationalizes a situation where network hiring is prevalent, but connected hires are not necessarily positively selected on realized productivity conditional on public information.

Theories of moral hazard apply when a hiring university can punish the sender of the referral—the PhD advisor—if the referred graduate performs poorly (Heath, 2018). In such a situation a hiring university could offer stronger incentives to network hires. This theory is most relevant when the sender of the referral is in close contact with the hiring party. In the current situation, the hiring university has limited control over the PhD advisor, and the PhD advisor has limited control over the PhD graduate after graduation. Therefore, we expect these mechanisms to have limited importance for the role of PhD advisors in the market for early career researchers.

Finally, consider the case where PhD advisors have private incentives to place their graduates, leading them to refer candidates who might otherwise struggle to secure an academic position. This implies a negative selection in terms of a_i among network hires. Beaman and Magruder (2012) provide evidence in line with this mechanism

from an experimental recruitment setting for casual jobs. Their findings suggest that the incentives to refer less qualified candidates complicate using referrals as an effective screening tool.

In sum, theories of network hiring generally imply a higher likelihood of a match within the network. Further, if social connections help screen candidates, then network hires should outperform other hires, conditional on public information. However, as PhD advisors have private incentives to place students, it is possible that they negatively select referrals, which can overturn the informational benefits of network hiring.

This paper's empirical analysis begins by estimating the role of PhD advisors' collaboration networks in PhD graduates' first placements. Motivated by the theoretical ambiguity surrounding network-based hiring, we then examine the productivity of network hires compared to their peers, focusing on whether public information explains the selection of network hires or whether a residual productivity premium of network hires remains.

2 Data

To track publication outcomes, affiliations, and co-author networks of PhD graduates and their advisors, we combine data on PhD dissertations with a comprehensive bibliographic database. The first data set is ProQuest Dissertations & Theses Global (PQDT)$^{\text{TM}}$, which provides information on PhD dissertations (Proquest, 2023). This data includes the names of PhD graduates and their advisors, the PhD-granting universities, and the titles and abstracts of the theses. The second data set is the Microsoft Academic Graph (MAG), which contains information on papers, authors, their affiliations, and citation links (Microsoft, 2021; Sinha et al., 2015; Wang et al., 2019).

Using the fuzzy matching algorithm `dedupe` (Gregg and Eder, 2022), we link PhD graduates as well as their advisors to their respective publication careers in MAG. The algorithm is trained with active learning: we label the potential links that the algorithm is least certain about. Links are identified when two records in the two data sets have

similar features. For graduates, this means a similar name, keywords, paper titles, as well as starting their publication career around the year of PhD graduation. For advisors, this means a similar name, affiliation name, as well as publishing activity around the student's graduation year.

We discuss the data processing and the quality of the data in more detail in Appendix A.1. First, both data sources are of high quality. The US Library of Congress uses ProQuest, and MAG's coverage of scientific works is comparable to data sources such as Scopus and Web of Science and has been used in previous studies (Huang et al., 2020). Second, while the active learning nature of our linking algorithm prevents us from calculating precision and recall that are representative of the linked sample, we show evidence of good precision and recall with two approaches. For advisors, who tend to be established researchers, we calculate a recall of about 0.75 across fields of study and years, assuming all advisors have a corresponding author entity in MAG. For chemistry graduates, who tend to publish during their PhD (Gaulé and Piacentini, 2013), we check whether the PhD-granting university is the same as the affiliation recorded on publications during the PhD. Using this information, which was not used during linking, we find a lower bound on the precision of linking of 0.78.

Our sample consists of PhD graduates from US universities who received their degrees between 1990 and 2014 and whom we linked to an author in MAG. Based on information in PQDT, we assign graduates to fields. We follow the MAG classification of scientific work into a hierarchical structure of fields. Level 0 groups all research into 19 fields, such as Physics, Biology, and Psychology. We further use the classification at level 1, which includes examples such as Thermodynamics, Econometrics, and Virology; we will refer to these as subfields throughout. There are 292 subfields in total.

We link universities in PQDT and affiliations in MAG to the Carnegie Classification of Institutions of Higher Education (American Council on Education, 2021). This allows us to identify institutions consistently across the two datasets. We consider only placements at institutions included on the Carnegie list, which we will refer to from now on as US universities.

We restrict the sample to graduates who (i) publish at least once within the first seven years after graduating and (ii) do so with an affiliation to a US university other than their PhD-granting university. Additionally, we limit our analysis to graduates for whom we have linked at least one advisor to MAG.

As hiring universities, we define for each graduate the set of universities that hire at least one graduate in the graduate's dissertation field within a five-year window of graduation. For example, for a biology graduate from Harvard in 1991, the set of hiring universities consists of all universities that hired at least one graduate in our sample of Biology PhD graduates from 1990 to 1994.

We define a graduate's first affiliation as the first university, other than their PhD-granting institution, listed as their affiliation on papers published between one and seven years after earning their PhD.

We construct each graduate's connections to universities based on their own and their advisor's co-authors. Co-authors are identified from all authors on papers published up to 20 years before the graduate's degree year. We include papers classified as "Journal," "Book," "Book Chapter," and "Conference." A university is considered connected through the advisor if the advisor has a co-author affiliated with that university—namely, if it is the co-author's most recent affiliation and if it was recorded within five years before the student's graduation. If there are multiple advisors listed on a dissertation, we record co-author connections of all advisors. Connections through the PhD graduate's own collaborations are defined analogously.

As outcomes for the post-PhD analysis, we summarize the count and 10-year forward citations of all papers in the first 7 years after graduating. We again include papers classified as "Journal," "Book," "Book Chapter," and "Conference." Further, we count the number of new co-authors at the first post-PhD affiliation, i.e., those the graduate did not co-author with up to the degree year. Researchers are assigned affiliations annually based on the most frequently recorded affiliation across their papers that year. The citations of graduate's first affiliations are calculated as the sum of 10-year forward citations for all papers published in the 10 years preceding graduation, at that university and in the

graduate's dissertation field.

We construct additional control variables based on PhD graduates' and their advisors' output during the PhD. First, we sum PhD graduates' publications, weighted by the average 10-year forward citations of all publications in the degree year of the publication venue. This measure is intended to capture the signal contained in the venue of the graduate's publication, as their own papers have had little time to accumulate citations. We create a dummy variable for graduates without publications, and among the remaining graduates, we use a dummy for the graduate's quintile in their field and cohort distribution (e.g., within biology in 1990–1994). Second, we construct the PhD advisor's decile in the citation distribution of advisors within the respective research field and within a five-year interval before graduation. For each advisor, we include only citations received within the 10 years up to and including the graduation year. For graduates with multiple advisors, we use the advisor listed first to construct control variables.

The alignment of the topics of graduates' dissertations and the research at hiring universities is likely an important determinant of hiring decisions and co-authorship. To account for topical alignment empirically, we compute topic vectors for graduates' dissertations and research output at hiring universities within the same field. Topics are based on the Fields of Study in the MAG hierarchy and are derived from the abstract text in dissertations and publications using Microsoft's language model (Wang et al., 2020, section 2.5), ensuring consistency in the topic definition across PQDT and MAG. We use the model's predicted topics and scores for each document to construct its topic vector. For the analysis, we use Fields of Study up to level 2 in the MAG field hierarchy, representing detailed topic areas[2]. We then calculate two cosine similarities: First, the similarity between a graduate's dissertation topic vector and the aggregated topic vector for their field at the hiring university. Second, the maximum similarity between the graduate's dissertation topic vector and individual researchers' vectors in that field affiliated with the hiring university. See Appendix A.7.2 for further details and Appendix

[2]There are 137,188 Fields of Study at level two, in addition to the 292 subfields (MAG level 1) and 19 fields (MAG level 0).

B.2.6 for descriptive statistics of the similarity measures.

Table 1: Summary statistics

Variable	Baseline	Advisor in MAG	Connected first affiliation	Not Connected first affiliation
Degree Year	2004	2004	2006	2004
Year First pub post PhD	2.91	2.88	2.57	2.96
Class Size	21.80	22.47	26.57	21.54
Link Score Student	0.97	0.98	0.97	0.98
Link Score Advisor		0.98	0.98	0.98
Connected Advisor		0.17	0.94	0
Connected Own Co-author		0.04	0.2	0
N Affiliation Connections				
through advisor		9.69	18.38	7.71
through co-author		2.03	3.86	1.62
Outcomes				
Cites 7y post PhD	364	363	578	314
Papers 7y post PhD	11.46	11.55	14.60	10.85
Co-authors first affil	3.99	4.02	6.64	3.42
First Affil 6y post PhD	0.59	0.59	0.49	0.61
Active 6y post PhD	0.89	0.89	0.87	0.89
Cites first affil	11,109	11,308	27,611	7,596
Observations				
N PhD Graduates	98,866	82,279	15,257	67,022
N PhD Graduates × hiring institutions	37,095,809	31,435,851	5,846,391	25,589,460

Notes: The baseline sample consists of PhD graduates in the Proquest Dissertations&Theses database from US universities included in the Carnegie Classification who graduated between 1990 and 2014 and who published at least once after their graduation with an affiliation that is not their PhD granting university. The column *Advisor in MAG* refers to PhD graduates for whom we have advisor information and have found a link to the Microsoft Academic Graph for the advisor. The rows *Link Score Student* and *Link Score Advisor* are the confidence scores about the identified links between entities. The scores are predicted values from the linking model of whether the records in PQDT and MAG refer to the same record. See Appendix A.1 for further information.

Table 1 shows summary statistics for the sample of linked PhD graduates. Each column shows the statistics for a different sample: the first column is for the baseline sample of PhD graduates; the second column is for those with a linked advisor—our analysis sample. Columns 3 and 4 split this sample into those with a network connection to

their first affiliation and those without. In Appendix A.7.1 we show the field composition of our sample. As our sample of graduates is not a random sample across fields, we will show the main results also separately by field.

The baseline sample of graduates linked to MAG consists of 98,866 graduates, for 82,279 of whom we have advisor information and could link the advisor to MAG. This is our analysis sample. In Appendix A.7 we show that our sample represents 7.4% of all PhD graduates in the PQDT dataset during the sampling period. The sample is not a representative sample of all US PhD graduates, but limited to those continuing an academic career in the US. This limits generalizability to graduates continuing their career in other sectors or countries.

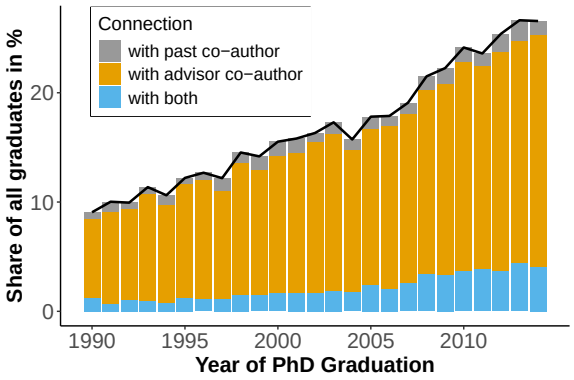
We group PhD graduates into PhD classes based on their PhD university $\times$ Field $\times$ 5 year window[3]. For example, we group PhD graduates from Stanford University in Computer Science in 1990-1994 into one PhD class and those graduating between 1995 and 1999 into a separate class. We only include graduates in our analysis sample in this class definition.

The table shows that the linking algorithm we use to match PhD advisors and graduates between PQDT and MAG is very certain about predicted links in all subsamples: On average, the identified links have a score of 97% or more, and this is the same for students and advisors independently of connection status.

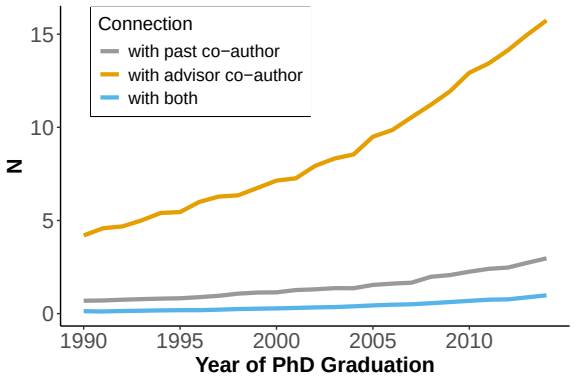
In the primary sample, 17% of graduates have a co-author connection to their first post-PhD affiliation through their advisor, and 4% through their own co-author network. This indicates that network connections can be important for matching PhD graduates to their first post-PhD affiliation, resonating with evidence reported by Brown (1965). In Figure 1a, we show the share of PhD graduates in our analysis sample by connection status to their first affiliation. The share of co-author connected placements has grown rapidly and is dominated by the PhD advisors' co-author connections. Figure 1b shows the average number of connected universities by connection type. On average, a graduate in 1990 was connected to 5 universities through their advisor, but a graduate in 2014 to

[3]Years are grouped into fixed windows, not rolling windows. That means we use year windows 1990-1994, 1995-1999, ... to group PhD graduates into classes.

15 universities. The number of connections to universities through PhD graduates' own co-authors have grown similarly but reached only 3 in 2014.



(a) Share of PhD graduates placed within collaboration network



(b) Number of connected affiliations

Figure 1: Co-author connections and placement of PhD graduates

We will estimate to what extent co-author connections predict placement, controlling for a variety of potential confounders, and to what extent the rising share of network hires is driven by the increasing number of connections or by changes in the importance of connections for matching.

Further, connected graduates are on average more productive than non-connected graduates; this holds in terms of the number of papers, number of citations, or year to first publication post-PhD. We will return to this pattern in the empirical section where we control for a range of confounders and discuss possible explanations.

3 Empirical Framework

We present the empirical framework used to investigate whether co-author connections across universities systematically relate to the placement of PhD graduates and their productivity.

Network connections and the first affiliation of PhD graduates To evaluate the role of network connections in shaping PhD graduates' transitions to their first post-PhD affiliation, we estimate a multinomial logit model. This model specifies the probability that a graduate i is matched with a hiring university j as:

$$P(d_{i,j} = 1) = \frac{e^{\alpha_{c(i),j} + \beta X_{i,j} + \gamma A_{i,j}}}{\sum_{j'} e^{\alpha_{c(i),j'} + \beta X_{i,j'} + \gamma A_{i,j'}}}. \tag{2}$$

In practice, we estimate the multinomial logit using a Poisson regression, which ensures computational feasibility and is equivalent, as discussed in Appendix C.1. Notably, this specification aligns with the equilibrium in matching models introduced by Choo and Siow (2006). The estimation of such models, including the possibility of using Poisson regression, is discussed in Galichon and Salanié (2024). Appendix C.2 further discusses this connection.

The variable $A_{i,j}$ is a binary indicator that equals one if graduate i has a network connection to university j. Our primary interest lies in the parameter γ, which quantifies how much a network connection increases the probability of a match.

To move toward a causal interpretation of this parameter, we control as finely as possible for factors that could influence both the probability of a match and the likelihood of having a network connection.

First, the fixed effects $\alpha_{c(i),j}$ account for factors that systematically impact hiring flows between pairs of hiring universities and graduating classes. The fixed effects $\alpha_{c(i),j}$ capture the possibility that universities closer to each other—either geographically or in terms of research interests—have more co-authorship connections and are more likely to hire graduates from each other. They also account for overall labor demand of each hiring university.

Second, we include controls $X_{i,j}$ that vary between each PhD graduate and hiring university. A key concern is specialization within fields: an advisor specializing in molecular biology is likely to collaborate more with departments specializing in molecular biology. The advisor's students are also likely to work on molecular biology. Thus, they are likely to be more attractive hires for other molecular biology-focused departments compared to a student from the same class who specializes in plant biology. We use topic similarity controls to account for such concerns, as introduced in Section 2. These controls capture the topical alignment between the graduate's research and the research conducted at hiring universities.

It is also important to discuss which confounders are not captured by our controls and which could invalidate a causal interpretation of the network parameter γ. Such confounding arises from factors that influence both the likelihood of having a connection and the graduate's matching outcome but are not accounted for by our controls. These factors must be relevant within a PhD-class and hiring university pair and must influence the matching outcome even in the absence of a connection. One example is ethnicity-based homophily in collaboration connections—both in the graduate-advisor and advisor-coauthor relationships—combined with an ethnicity-specific preference for locations. In such a scenario, graduates would be more likely to match with universities where their advisor has a co-author. This correlation would arise even if the advisor did not have a direct connection with the university.

Assessing changes over time The stark rise in the number of connections to universities highlights the possibility that the effect of a given connection changes over time. Mechanically, if the number of connections increases over time, the probability of a match through any single connection decreases at some point since graduates can only have one first placement university. However, the overall probability of a match within the network connections can increase over time. To test for changes in γ over time while avoiding confounding differences over time with changes in the composition of fields, we estimate the following multinomial logit model:

$$P(d_{i,j} = 1) = \frac{e^{\alpha_{c(i),j} + \beta X_{ij} + \Delta\gamma A_{i,j} \times (t(i) - 1990)}}{\sum_{j'} e^{\alpha_{c(i),j'} + \beta X_{ij'} + \Delta\gamma A_{i,j'} \times (t(i) - 1990)}} \tag{3}$$

We test whether $\Delta\gamma = 0$. In addition to the baseline controls, we also account for heterogeneity in network effects across fields and the number of connections. Specifically, we control for a field-specific quadratic function of the number of connections N_i: $\gamma^0_{F_i} A_{i,j} + \gamma^1_{F_i} A_{i,j} N_i + \gamma^2_{F_i} A_{i,j} N_i^2$, where F_i is the dissertation field of graduate i. The number of connections of the student to the potential hiring university are measured separately for the PhD graduate's own co-author connections and their advisor's co-author connections.

Post-PhD outcomes of network hires We compare the careers of PhD graduates whose first affiliation is connected through a co-author to those whose first affiliation is not connected. To analyze post-PhD outcomes y_i, we use the following Poisson regression framework:

$$y_i = \exp\left(\alpha_{c(i)} + \alpha_{j(i)} + \beta X_i + \gamma A_{i,j} + u_i\right), \tag{4}$$

where:

- $j(i)$ is the first affiliation post-PhD of graduate i
- $\alpha_{c(i)}$ is a fixed effect for the PhD class,
- $\alpha_{j(i)}$ is a fixed effect for the first affiliation,
- X_i includes individual controls measured up to the time of graduation,
- $A_{i,j(i)}$ indicates whether the graduate's first affiliation post-PhD had a network connection.

Our main interest lies in the parameter γ, which captures the difference in post-PhD outcomes between graduates placed through the network and those not placed through the network. By varying the fixed effects included when estimating equation (4), we analyze the role of the network from the perspectives of both graduates and hiring universities.

First, we compare graduates within the same class by including a class fixed effect, $\alpha_{c(i)}$. Second, we compare new hires at the same university by including a first-affiliation fixed effect, $\alpha_{j(i)}$. This specification closely relates to studies examining productivity outcomes within firms (Burks et al., 2015). Third, we combine these approaches by simultaneously including both types of fixed effects.

We also assess the role of the students' observable characteristics X_i. These variables include measures of the graduate's and their advisor's research output during the PhD, as described in Section 2. First, we control for the graduate's productivity during the PhD, measured as their publication output weighted by the average citations of the publication venue. This measure reflects the graduate's overall productivity during the PhD while

accounting for the limited time their work has had to accumulate citations. Second, we control for the advisor's research prominence in the graduate's research field.

The universities also observe similar information at the time of hiring. Our regressions, with and without these controls, will indicate whether networks provide additional information beyond these variables. This analysis closely relates to existing research on network hiring with skill measures (Burks et al., 2015; Pallais and Glassberg Sands, 2016). However, there are two key differences to these earlier studies. First, we observe an occupation-specific productivity measure instead of more general skills. Second, we observe this measure not only for graduates but also for advisors, who can play an important role in the referral process and for the graduates' expected productivity.

Assessing changes over time We investigate whether the gap in post-PhD outcomes between connected and non-connected graduates changed during our sampling period by expanding the framework in equation (4):

$$y_i = \exp\left(\alpha_{c(i)} + \alpha_{j(i)} + \beta X_i + \gamma_{F_i} A_{i,j(i)} + \delta A_{i,j(i)} \times (t(i) - 1990) + u_i\right), \tag{5}$$

We are primarily interested in the parameter δ, which captures a linear trend in the outcome gap between connected and non-connected graduates. Time $t(i)$ refers to the graduation year.

This specification addresses potential confounding of changes over time with composition change across fields in two ways. First, it accounts for variation in the gap between connected and non-connected graduates across fields, captured by the field-specific parameters γ_F. Second, the controls X_i include fixed effects at the level of the graduate's field $\times$ graduating year, capturing field-specific time trends in post-PhD research outcomes during our sampling period.

4 Results

We now present the main results of our analysis, examining the role of co-author connections in the hiring and career outcomes of PhD graduates. All regressions are estimated with the *fixest* package in `R` (Bergé, 2018).

First, we estimate the effect of co-author connections on the graduate's first academic placement. Second, we compare the post-PhD career outcomes of graduates hired through a network connection to those hired without one. In particular, we analyze whether network-based hiring conveys additional private information beyond publicly available signals at the time of PhD graduation. Throughout the analysis, while we focus on the advisor's connections, we also report results for the PhD graduates' own connections.

4.1 Network connections and the first placement of PhD Graduates

Table 2 shows the results from estimating equation (2). We estimate a multinomial logit model, where each observation corresponds to a pair of a PhD graduate and a potential hiring university and the outcome is whether the hiring university is the PhD graduates' first affiliation post-PhD.

The first column shows a coefficient of 0.728 for the advisor connection, indicating that a match between a graduate and a university is more than twice as likely if the advisor has a co-author at the university. A co-author connection directly linked to the PhD graduate yields an even larger estimated coefficient of 1.04. To inform about the overall importance of PhD advisors' and PhD graduates' co-author connections, one also needs to consider the number of connections. Before returning to this at the end of this section, we now tighten the regression models to account for potential confounders that drive co-author connections and placement.

In column 2, we add fixed effects for the class $\times$ (potential) hiring university pair. The coefficient on the advisor connection is barely altered to 0.747. This specification controls for any systematic matching determinants that vary at the level of the PhD

class × (potential) hiring university pair, for example, geographic distance or research specialization within fields at the university level (e.g. specialization into plant biology).

In columns 3 and 4, we include controls for the research productivity of the PhD student before graduation, the research productivity of the advisor, and the number of connected universities. Each of these is interacted with the decile of the (potential) hiring university's citations within the field's distribution. These controls account for systematic differences in demand for PhD graduates across universities with varying levels of research intensity. The estimated coefficients on both the advisor's and the PhD's connection are hardly changed.

Table 2: Multinomial Logit: Matching PhD Graduates to Universities

Panel A: Average effect of network connections on placement					
Dependent Variable:			Match formed		
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Advisor connection	0.728	0.747	0.743	0.744	0.684
	(0.025)	(0.025)	(0.025)	(0.025)	(0.023)
PhD connection	1.04	1.12	1.12	1.11	1.07
	(0.037)	(0.049)	(0.049)	(0.049)	(0.048)
Fixed-effects					
PhD Class×University ID		Yes	Yes	Yes	Yes
Advisor Citation Decile×University Citation Decile×Field			Yes	Yes	Yes
Pre Graduation Productivity×University Citation Decile×Field				Yes	Yes
Connections Decile×University Citation Decile×Field				Yes	Yes
Additional controls with varying slopes					
Max similarity to faculty members×Field					Yes
Avg similarity to faculty members×Field					Yes
Observations	650,527	650,527	650,527	650,527	650,527

Panel B: Assessing changes over time of the effect of network connections			
Dependent Variable:		Match formed	
Model:	(1)	(2)	(3)
Variables			
Advisor connection	0.655		
	(0.046)		
(Degree Year-1990) × Advisor connection	0.002	0.002	0.004
	(0.003)	(0.003)	(0.003)
PhD connection	1.28		
	(0.119)		
(Degree Year-1990) × PhD connection	-0.012	-0.012	-0.005
	(0.006)	(0.006)	(0.006)
Fixed-effects			
PhD Class×University ID	Yes	Yes	Yes
Advisor Citation Decile×University Citation Decile×Field	Yes	Yes	Yes
Pre Graduation Productivity×University Citation Decile×Field	Yes	Yes	Yes
Connections Decile×University Citation Decile×Field	Yes	Yes	Yes
Additional controls (with varying slopes)			
Max similarity to faculty members × Field	Yes	Yes	Yes
Avg similarity to faculty members × Field	Yes	Yes	Yes
Advisor connection × Field		Yes	Yes
PhD connection × Field		Yes	Yes
Advisor connection × N Advisor connection × Field			Yes
PhD connection× N PhD connection × Field			Yes
Advisor connection × N^2 Advisor connection × Field			Yes
PhD connection× N^2 PhD connection × Field			Yes
Observations	650,527	650,527	650,527

Notes: Unit of observation is a pair of PhD graduate and a hiring university. The sample is restricted to pairs of a PhD class (PhD university×Field×5 year window) and hiring university with variation in the connection status of PhD graduates and where at least one hire has occurred. See equation (2) for the multinomial logit specification for Panel A and equation (3) for Panel B. Clustered (PhD university×Field×5 year window + hiring university) standard-errors in parentheses

A potentially major confounder of the estimated effect of co-author connections is the research specialization of graduates. To address this concern, column 5 demonstrates that the measured effect of network connections on matching is not driven by the alignment

of the PhD graduate's specialization in research topics with the research specialization at potential hiring universities. We include controls for the cosine similarity between the topics of the graduate's dissertation and research output of potential hiring universities.

While the similarity measures predict which university a graduate matches with, they have limited impact on the estimated effect of network connections. The estimated coefficient for advisor connections decreases slightly to 0.684, and for PhD graduate connections to 1.07. Appendix B.2.6 presents the estimated coefficients on the similarity measures and summary statistics. The estimated coefficients, rescaled by one standard deviation of the similarity measures, are positive and of the same order of magnitude as the coefficients on connections. Thus, there is systematic matching on topic similarity, but controlling for it has limited impact on the estimated coefficients on connections.

Advisor connections play a significant role across all fields, as shown in Appendix B.2. Further, there is substantial variation in the point estimates across fields, which is unrelated to the average number of connections.

To validate the robustness of our findings we perform various checks. First, we assign PhD graduates a randomly chosen advisor from their PhD class. The placebo advisor's connections do not predict university placement, consistent with PhD advisors having a close relationship with their own students, but not other students in the same PhD class. Second, Appendix B.2.4 shows that our results are robust to changes in the level of detail of measured research topics. Third, Appendix B.2.3 demonstrates that the estimated effect of advisor connections diminishes with time since the last collaboration, consistent with connections weakening as temporal distance increases. Lastly, we drop the fixed effect at the level of PhD-class×hiring university and replace it with a fixed effect just for the hiring university in Appendix B.2.7. We do this to include most dyads in the matching regression. The results show that controlling for the fixed effect at the level of PhD-class×hiring university compared to just the hiring university has no impact on the point estimate of an advisor connection. However, in the larger sample, the point estimate for γ is 1.1 while it is 0.66 in the smaller sample. We report the main estimates with the more restrictive fixed effects, and smaller sample, as a conservative estimate for

the role of advisor connections in the placement of PhD graduates.

In Panel B of table 2, we show the estimated time trend in the advisor connection's effect on matching PhD graduates to their first university. The estimated connection effect is stable over time. In column (2), we add controls for differences in the effect of connections by field, and in column (3), we further allow the effect of connections to vary with the number of connections. The coefficient on the time trend in column (3) is modestly positive (0.004, s.e. 0.003) for advisor connections and modestly negative for PhD connections (-0.005, s.e. 0.006). Even when scaled to the difference between the first and last year, i.e., multiplying the estimated coefficients by 25, the point estimate implies a modest difference of 0.1 for advisor connections and -0.125 for PhD connections.

In summary, our findings so far indicate that network connections—both through the advisor's co-authors and the PhD graduate's own co-authors—significantly increase the likelihood of a graduate being placed at a university. This estimated effect of a connection to a university is stable over time despite a large increase in the number of connected universities. Now, we quantify to what extent the advisor and PhD connections increase the probability of being placed at any of the connected universities.

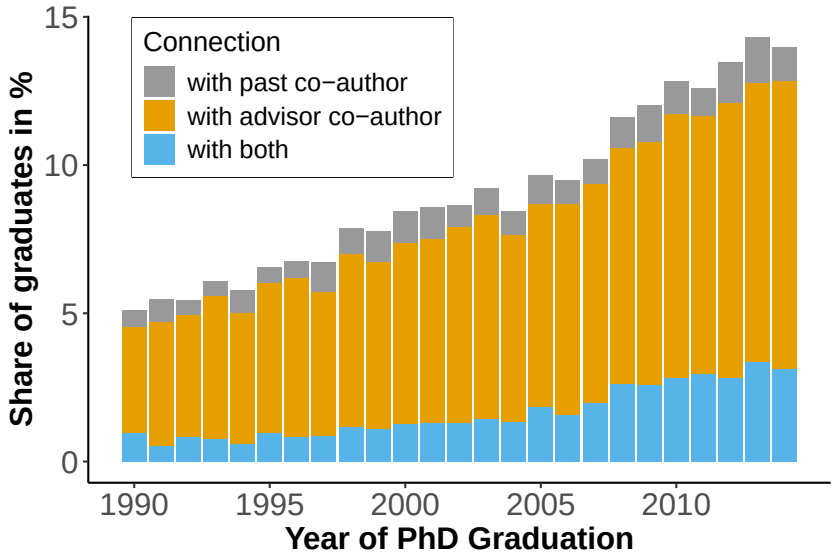


Figure 2: Estimated effect of co-author connections on placement of PhD graduates

Notes: We plot $\bar{P}_{j\in J_{A=1}} - \hat{P}^0_{j\in J_{A=1}}$ which is the share of connected hires minus the predicted share of connected hires in the absence of connections. See Appendix C.3 for the derivation. The values for $\bar{P}_{j\in J_{A=1}}$ are taken from Figure 1a and $\hat{P}^0_{j\in J_{A=1}} = \frac{e^{-\hat{\gamma}}\bar{P}_{j\in J_{A=1}}}{e^{-\hat{\gamma}}\bar{P}_{j\in J_{A=1}}+(1-\bar{P}_{j\in J_{A=1}})}$ is evaluated with the point estimates for γ shown in Table 10.

To assess the extent to which the share of network-based hiring in Figure 1a is systematic we subtract the share of hires at a connected university that would have been expected without a connection. Specifically, we use the estimated multinomial logit model to subtract the predicted share of hires at connected universities that would have occurred without a connection.

The predicted probability to match with a connected affiliation[4] in absence of connections is $\hat{P}^0_{i,j\in J_{A=1}} = \frac{e^{-\hat{\gamma}} P_{i,j\in J_{A=1}}}{e^{-\hat{\gamma}} P_{i,j\in J_{A=1}} + (1 - P_{i,j\in J_{A=1}})}$. See Appendix C.3 for the derivation. We evaluate the counterfactual $\hat{P}^0_{i,j\in J_{A=1}}$ at the sample average of the share of hires $\bar{P}_{j\in J_{A=1}}$ and at the estimated coefficients $\hat{\gamma}$ shown in Table 10. We plot the systematic, predicted share of connected hires $\bar{P}_{j\in J_{A=1}} - \hat{P}^0_{j\in J_{A=1}}$ in Figure 2. The yellow bars show the systematic role of PhD advisors' collaboration networks in placing PhD graduates, which has more than doubled from 1990 to 2014. The estimated increase in the probability of being placed at a university connected by an advisor's co-author reached over 10% in 2014. Most of the PhD's own co-author connections overlap with the advisors; their joint estimated role is shown in blue. PhD's connections without overlap with the advisor are shown in grey and are quantitatively almost irrelevant.

4.2 Post-PhD outcomes of network hires

We now study whether graduates placed through their advisor's co-author network have different career outcomes after graduation compared to graduates not placed through the advisor's network. We focus on outcomes measuring research productivity, such as citations and the number of papers published in the first 7 years after graduation. Furthermore, we complement these results with additional outcomes on (i) collaboration patterns, measured by the number of new co-authors at the first post-PhD affiliation, (ii) stability of the match, measured by an indicator for whether a graduate is still affiliated with the same university 6 years after their PhD, (iii) whether graduates produce any output more than 6 years after the PhD and (iv) the productivity of the hiring university measured by citations. We focus on the results regarding advisor connections but discuss

[4]The set of connected universities is $J_{A=1}$ and the probability to match with one of the universities $j \in J_{A=1}$ is denoted $P_{i,j\in J_{A=1}}$.

the difference between the estimates for PhD connections at the end of this section.

The results are shown in Table 3, separately for the different comparisons: Panel A shows the results from the point of view of the graduating class, Panel B from the point of view of the hiring university, and Panel C includes fixed effects for both. All results are from Poisson regressions, and so the coefficients measure differences in log points.

Comparing within the PhD class Starting with panel A, we find that graduates placed through advisor connections are more productive than their classmates not placed through connections. Column 1 shows that this gap is 23.8 log points for the number of citations.[5] This gap can reflect ability, access to productivity-enhancing employers, or a direct effect of the connection as such.

Columns 2 and 3 show that publicly observable predictors of productivity can partly explain this gap. Controlling for the expected citations of the graduate's work before graduation lowers the point estimate to 0.172; further controlling for the advisor's citations before graduation lowers the estimate to 0.121 (standard error 0.021). Thus, graduates from the same PhD class with an advisor connection to their first post-PhD university outperform, above and beyond, what is predictable based on their own and their advisor's pre-graduation research output. In the remaining columns of the table, we include these additional predictors as controls in the regressions.

Column 4 shows that a higher number of papers partly drives higher citations. The point estimate is 0.061 (standard error 0.01), indicating that about half the higher number of citations post-PhD is due to more papers. Thus, graduates placed through connections are more productive in terms of the quantity and quality of research.

Column 5 shows that graduates who are placed through connections collaborate more with their new colleagues. Despite increased collaboration, connected hires are more than 14% less likely to be affiliated with their first university 6 years after the PhD (column 6). This result is not driven by dropping out from research: Column 7 shows that the probability of producing any output 6 years after the PhD or later is the same. PhD graduates are likely to be in a postdoc position (Fiegener, 2010); thus, leaving the first

[5]The point estimate translates into $exp(0.238) - 1 \approx 27\%$ higher citations.

affiliation can be a positive sign as it may reflect a step up the career ladder to a faculty position.

Finally, in column 8, we show that graduates placed through connections are matched to substantially more productive universities, as measured by the citations of the university in a graduate's dissertation field.

To summarize, from the perspective of the PhD class, graduates placed through the network have substantially different careers compared to their peers. They outperform post-PhD in terms of both quantity and quality of research, are placed in more productive universities, collaborate more with their new colleagues, and are more likely to switch affiliations 6 years after the PhD. This holds even after controlling for pre-determined, publicly observable predictors of productivity. That is, among PhD graduates from a similar environment, there is substantial selection into network hiring, reflecting the ability of the graduate, the quality of first placement, and the quality of the match to their first post-PhD university.

To further highlight the selection into network hiring, even among graduates from similar environments, we now show whether the estimates reflect selection across advisors. We repeat the analysis of Table 3 Panel A but include advisor fixed effects. Figure 3 shows in grey the baseline estimate and in yellow and blue specifications with the advisor fixed effect. The relative productivity of connected and not connected graduates is similar when compared within a PhD class as it is within an advisor. Even for graduates with the same advisor, there is a large gap in productivity post-PhD, which highlights substantial selection into network hiring even conditional on a very similar PhD environment.

Comparing within the hiring university To provide evidence on whether network hiring outperforms as a hiring channel, we now compare graduates within a hiring university[6]. In Table 3, Panel B, we present results from the hiring university's perspective. Column 1 indicates that connected hires are approximately 11% more productive than other hires at the same university, suggesting that connected hires are positively selected based on productivity even when placed at the same university post-PhD.

[6]We interact the fixed effect for the hiring university with field and 5-year graduation window.

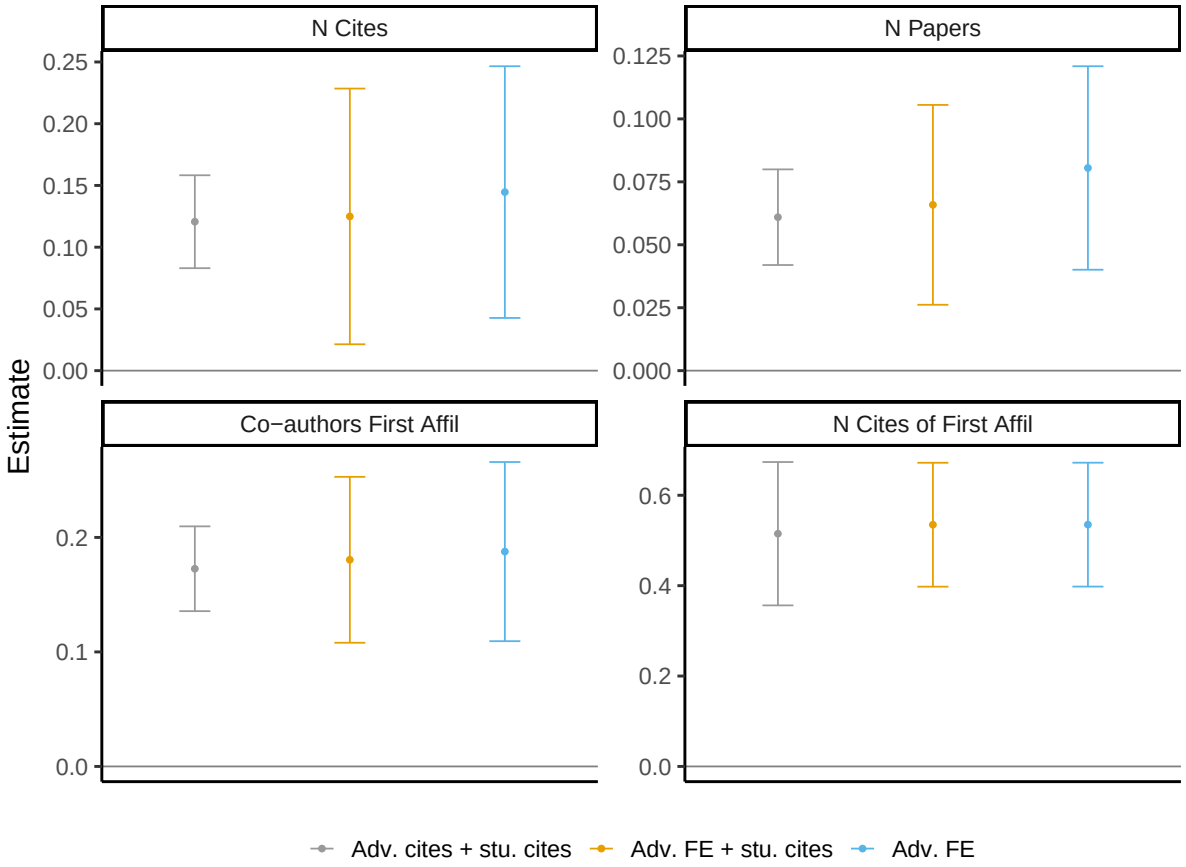


Figure 3: Poisson regression of post-PhD outcomes on advisor connection with advisor fixed effects

In columns 2 and 3, we progressively add the productivity predictors as in panel A, showing that publicly observable factors can explain the productivity premium of network hires. Including both the advisor's and the student's pre-graduation productivity (in column 3) makes connected and non-connected hires indistinguishable—the point estimate for γ drops to -0.013 (standard error 0.021). This means that connected hires are positively selected in terms of publicly observable productivity predictors at the time of PhD graduation but not in terms of private information revealed by the connection.

The remaining columns in panel B keep the same controls as column 3. Compared to other hires at the same university, they show that connected hires publish a similar number of papers and collaborate with a similar number of new co-authors but are more likely to leave the university despite continuing research at a similar rate.

The results about the role of publicly observable predictors of productivity in panel B are noteworthy for three reasons. First, they contrast with the findings from panel

A from the graduating class' point of view: Connected hires from the same PhD class substantially outperform their non-connected peers, but once the place of work is taken into account, this gap drops to 0. This implies that either the place of work has a direct effect on the productivity of graduates or that the hiring mechanism of these more productive places can uncover information about graduates that is unobserved to the analyst, also in the absence of a network connection. Second, the gap in column 1 resembles the findings in Burks et al. (2015) for trucking and high-tech. But, while they found that connected hires were similar along measures of general human capital (Burks et al., 2015, Table IV), we find that the gap can be explained by predictors of productivity specific to the job. Third, the results are not in line with theories that highlight the use of networks to reveal information about match quality: Such theories imply that connected hires should be more productive than non-connected hires even after controlling for any publicly observable predictors of productivity. We find no evidence for such effects.

In Appendix B.3.4 we present the within hiring university estimates for citations, corresponding to columns (1) and (3) of Panel B, separately by field.

Comparing within class and within hiring university Panel C in Table 3 shows results where we control for fixed effects for the hiring university and the PhD class. The results are similar to those in panel B: Connected hires outperform other hires at the same university, but this performance gap can be predicted by public information at the time of hiring. Point estimates are very similar when adding the fixed effect for the PhD class in addition to the hiring university. This implies that conditional on the selection done by the hiring university, there is limited additional information in which PhD class a student is from when comparing connected to not-connected hires.

PhD graduate's connections PhD graduates' own co-author connections to other universities also exhibit substantial positive selection in terms of post-PhD citations across Panels A, B, and C of Table 3. After controlling for productivity predictors in column (3), the gap in citations is small and positive but statistically insignificant in all panels. In contrast, the number of papers shows a consistently positive and significant gap, ranging

from 4.3 to 5.3 log points, even after applying controls. The differences between hires through PhD connections and advisor connections are further reflected in other outcomes, all of which display substantially more negative point estimates for PhD connections. PhD-connected hires have fewer *new* co-authors at the hiring university, partially due to pre-existing co-author relationships with researchers at the same university. They are also significantly less likely to remain at the hiring university. Additionally, they are more likely to cease publishing entirely. While the point estimates for PhD and advisor connections differ, our initial results on matching highlight that advisor connections play a much larger quantitative role.

Assessing Changes over time In the Appendix, in Table 14, we report the results from estimating equation (5), with the fixed effects for both class and field. We are unable to detect substantial changes in the difference in ex-post hiring outcomes between graduates hired through the network and graduates not hired through the network. Thus, despite the increasing importance of the collaboration network for PhD graduates' hiring, the selection into being hired through the network has remained stable during our sampling period.

Table 3: Poisson Regression: Post-PhD outcomes of connected vs. not connected hires

Panel A: Comparison of post-PhD outcomes with Class Fixed Effect

Dependent Variables:	N Cites PhD graduate			N papers	Co-authors First Affil	Same Affil PhD+6yrs	Any output PhD+6yrs	N Cites of First Affil
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Advisor connection	0.238	0.172	0.121	0.061	0.173	-0.142	-0.007	0.515
	(0.021)	(0.019)	(0.019)	(0.010)	(0.019)	(0.013)	(0.004)	(0.081)
PhD connection	0.232	0.027	0.045	0.053	-0.089	-0.505	-0.030	-0.041
	(0.030)	(0.030)	(0.029)	(0.015)	(0.032)	(0.032)	(0.006)	(0.035)
Fixed-effects								
PhD Class	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Graduation Productivity×Field		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Advisor Citation Decile×Field			Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Pseudo R^2	0.43	0.51	0.52	0.30	0.35	0.04	0.01	0.75
Observations	82,141	82,141	82,141	82,279	77,484	79,895	81,771	82,092

Panel B: Comparison of post-PhD outcomes with Destination Fixed Effect

Dependent Variables:	N Cites PhD graduate			N papers	Co-authors First Affil	Same Affil PhD+6yrs	Any output PhD+6yrs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
Advisor connection	0.117	0.059	-0.013	0.018	0.018	-0.067	-0.009
	(0.022)	(0.020)	(0.021)	(0.009)	(0.012)	(0.012)	(0.005)
PhD connection	0.208	0.014	0.039	0.048	-0.067	-0.465	-0.031
	(0.026)	(0.025)	(0.025)	(0.013)	(0.029)	(0.037)	(0.007)
Fixed-effects							
Hiring University Id×Field × 5 year window	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Graduation Productivity×Field		Yes	Yes	Yes	Yes	Yes	Yes
Advisor Citation Decile×Field			Yes	Yes	Yes	Yes	Yes
Fit statistics							
Pseudo R^2	0.51	0.57	0.58	0.37	0.38	0.06	0.01
Observations	81,660	81,660	81,660	82,279	67,755	75,833	80,463

Panel C: Comparison of post-PhD outcomes with Class and Destination Fixed Effect

Dependent Variables:	N Cites PhD graduate			N papers	Co-authors First Affil	Same Affil PhD+6yrs	Any output PhD+6yrs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
Advisor connection	0.080	0.043	-0.004	0.024	0.015	-0.068	-0.008
	(0.023)	(0.022)	(0.021)	(0.010)	(0.016)	(0.013)	(0.005)
PhD connection	0.218	0.026	0.039	0.049	-0.068	-0.483	-0.032
	(0.029)	(0.028)	(0.027)	(0.016)	(0.035)	(0.038)	(0.007)
Fixed-effects							
PhD Class	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hiring University Id×Field × 5 year window	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Graduation Productivity×Field		Yes	Yes	Yes	Yes	Yes	Yes
Advisor Citation Decile×Field			Yes	Yes	Yes	Yes	Yes
Fit statistics							
Pseudo R^2	0.63	0.68	0.68	0.43	0.45	0.09	0.02
Observations	81,614	81,614	81,614	82,279	66,551	74,276	80,151

Notes: The unit of observation is a PhD graduate. See equation (4) for the Poisson regression specification. Observations with zero outcomes that are perfectly predicted by fixed effects are dropped. Clustered (PhD university- Field - 5 year window + Hiring university) standard-errors in parentheses. N Cites PhD graduate measures citations received on articles published in the first 7 years post PhD graduation. N papers is the number of articles published in the same period. Co-authors First affil measures the number of new co-authors at the first post-PhD affiliation. Same Affil PhD+6yrs and Any Output PhD+6yrs indicate whether the PhD graduate, 6 years after the PhD graduation or later, is still affiliated with their first post-PhD affiliation and whether they publish any papers at that point.

Lastly, we provide additional results regarding the robustness of our main results in the Appendix. Our analysis of post-PhD outcomes and the interpretation of the results depends on whether factors that are known to market participants, determine both post-PhD outcomes and whether a graduate ends up a connected hire. We consider two candidates for such factors: (1) the size of the network, measured by the number of connected affiliation and (2) the topical alignment of the PhD graduate's dissertation with the hiring university. See Appendix B.3.6 and B.3.3 for the results. In both cases we find very similar results to those reported in Table 3.

5 Conclusion

We analyze the role of PhD advisors' collaboration networks for matching PhD graduates to universities. To do so, we build a novel database containing information about PhD graduates that allows us to track them and their advisors throughout their careers.

We find that co-author connections of advisors strongly predict at which university a PhD graduate finds their first job—having a connection doubles the probability of matching with a given university. Our estimates account for unobserved heterogeneity at the level of PhD class $\times$ hiring university pair and the similarity in research topics between the graduate's dissertation and her potential new colleagues at the hiring university. Overall, the importance of advisors' collaboration networks for matching PhD graduates to universities doubled between 1990 and 2014 due to the increasing size of the network.

We then show the gap between connected and non-connected hires on a range of post-PhD outcomes. When comparing graduates from the same PhD class, connected hires are more productive post-PhD. They receive over 12% more citations, even after controlling for public information about the graduate's and their advisor's research output during the PhD. Connected hires are still more productive when comparing network hires to their peers at the same hiring university. However, in this case, the productivity premium of network hires can be predicted by public information at the time of hiring.

Our results show that favoritism does not dominate network hiring, as network hires

are positively selected based on productivity. However, the lack of a productivity premium beyond public information suggests that the screening benefits of network hiring are limited in the PhD graduate market.

A Data appendix

The broad steps in our data pipeline are shown in Figure 4. The following subsections explain the details.

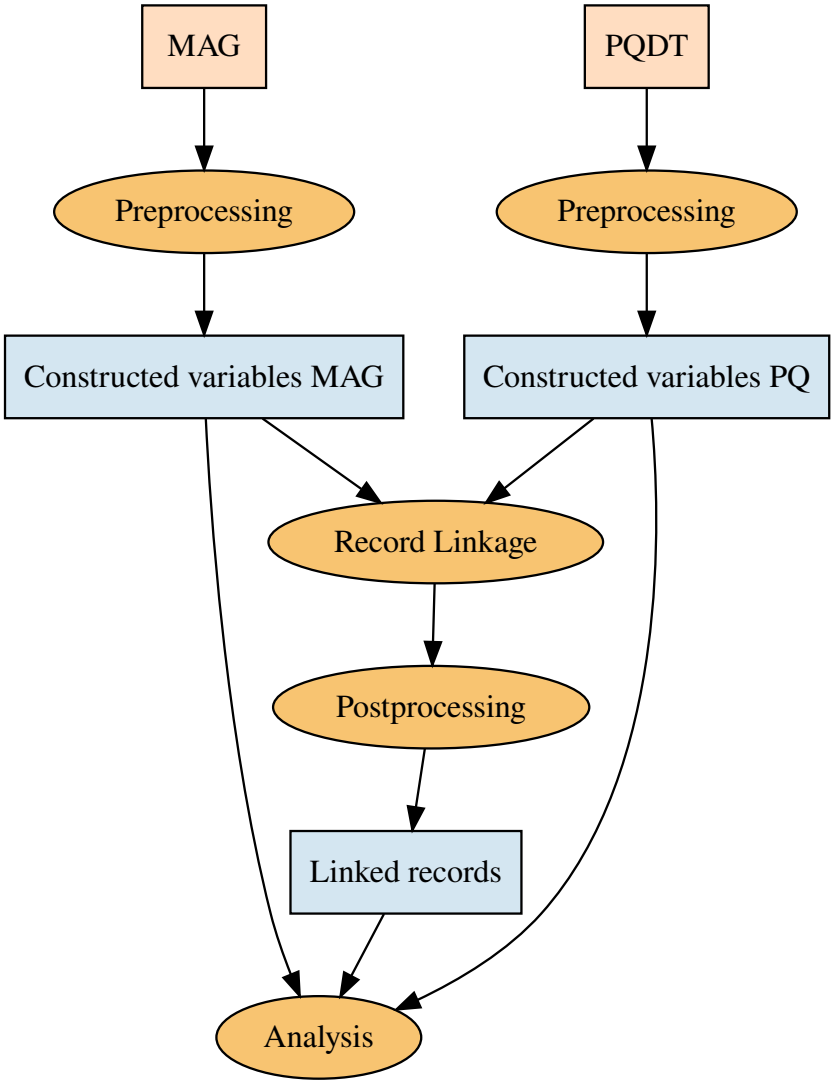


Figure 4: The end-to-end data pipeline.

The figure shows a schema of the data pipeline. Oval nodes are data handling steps; rectangular nodes are data (single or multiple relational tables). See text for details.

A.1 Data sources

A.1.1 Microsoft Academic Graph (MAG)

The data are collected from the internet by Microsoft's search engine Bing and from RSS Feeds of publishers. Six entities are extracted from the data: Author, Institution,[7] Paper, Field of Study, Venue (e.g., "American Economic Review"), and Event (e.g., a specific issue of the AER). For some entities (Author, Affiliation), the database reports both the original string and a disambiguated string together with an entity ID generated by MAG's algorithms, described in more detail by Sinha et al. (2015). The Affiliation entity is extracted from the metadata of published papers, and an author may have multiple affiliations for the same paper.

They use additional proprietary in-house and online data to create the Field of Study entity. There are five levels of Field of Study, and the labels represent the semantics of a paper. At the highest level, 0, the labels correspond to fields of study such as chemistry, biology, and mathematics. A paper is assigned several such Fields of Study from any level.

Compared to other databases often used in bibliometric analysis, such as Scopus or Web of Science, several studies find MAG to be comparable in coverage. Visser, van Eck and Waltman (2021) benchmark several databases (MAG, Crossref, Dimensions, and a restricted version of Web of Science) against Scopus for the years 2008 to 2017. They report that MAG has the highest coverage (81 percent). They also report that MAG covers many documents that are not in Scopus, and in a random sample of these, many of them are of a scientific nature. Martín-Martín et al. (2020) reports that Microsoft Academic has the second-highest coverage of citations among a set of services, including Google Scholar, Scopus, and Web of Science. Microsoft Academic finds 60% of all citations, with good coverage for many fields except some issues in Physics and some Humanities categories. Hug, Ochsner and Brändle (2017) find that a citation analysis with MAG gives the same results as with Scopus. Although they find some

[7] Institutions can be universities, firms, or international organizations. Since we focus on universities, we will refer to this entity as "university".

errors in metadata on papers, such as missing authors and wrong year of publication, they believe that MAG has the potential for "full-fledged bibliometric analysis". Hug and Brändle (2017) benchmark the quality and coverage of MAG with Scopus and Web of Science (WoS) for the research output of a specific university. They find high rank-rank correlations between the databases in citation counts and find that MAG has good coverage of venue entities. Harzing and Alakangas (2017) compare citation counts for a set of academics in different disciplines (Engineering, Social Sciences, Humanities, Life Sciences, Sciences) and find that MAG, behind Google Scholar, performs at least as good or better than WoS and Scopus across disciplines.

A.1.2 Proquest Dissertation & Theses (PQDT)

PQDT is a repository of PhD theses provided by Clarivate Analytics (Proquest, 2023). To our knowledge, it is the most complete repository of PhD theses in the United States. We use the metadata of this database: the name, graduating university, reported advisors of the student, the title, and the keywords of the dissertation.

A.2 Preprocessing

MAG Based on the MAG database, we build several intermediate tables for each author. We extract the Fields of Study they publish in and the start and end year of their publishing career. We extract the unique affiliation-year combinations of their published papers. From the first ten papers published within the first five years of an author's publication career, we also extract the unique keywords and the year-title combinations of the publications. The keywords are the MAG Fields of Study labels at level 1.

Because an author can publish papers with different affiliations in the same year, we use two types of universities: `main` and `all`. `main` refers to the one affiliation in a year at which the author publishes most of their papers. `all` refers to the list of all affiliations.

MAG is built from the raw data of each paper and thus requires author disambiguation. The disambiguation in MAG is conservative because high precision is traded off for lower recall. For instance, this creates some entities that are duplicates of other author

entities. These duplicates, however, tend to be associated with only one or a few papers and have shorter publication careers than the main entity for a researcher. For example, MAG will have multiple author entities named Jordi Gali. Still, one of them is clearly the main entity with basically all of his publications, and the duplicate author entities tend to have only one "publication" that does not provide sufficient information for author disambiguation, including online publications like short reports or blog posts.

For this reason, we follow Huang et al. (2020) and restrict the author sample to entities that publish at least two articles during their whole career and, on average, at most 20 papers per year. This is our main author sample used in the remaining analysis.

ProQuest We use the PQDT database (Proquest, 2023). Starting with the individual dissertation files, we extract tables for advisors, authors, tagged fields of study,[8] and predicted fields of study. The fields of study are predicted from the abstract of each thesis with the language model from MAG. We keep the ten most likely predicted fields as long as their score is above 0.4 (a heuristic threshold below which we found the predictions to be noisy). We export these tables from the Clarivate server and process them further.

We use the dissertation identifier from PQDT as an identifier for graduates. As advisors are not uniquely identified, we create identifiers called `relationship_id`. They uniquely identify the graduate and the advisor's position in the metadata. We uniquely identify universities based on their name. We correspond the fields reported in ProQuest to the MAG fields (lvl 0) utilizing the classification by Organisation for Economic Co-operation and Development (2007).

Crosswalking university entities In several steps of our analysis, it is important to distinguish universities from each other, but reported names differ between MAG and PQDT. We create a crosswalk of universities from the set of US universities on the Carnegie classification (American Council on Education, 2021): We correspond the university identifiers from MAG and from PQDT to the Carnegie list. We do so with the name and, where possible, the zip code and city of the university; for this task, we

[8]These tags are readily available but are different from the labels used in MAG.

use the table of zip codes from Missouri Census Data Center (2020). We link all R1 universities, most of R2 and R3 universities, and some additional institutions. Whenever necessary, we use the university name from the Carnegie classification, which we refer to as the crosswalked university name.

A.3 Record Linkage

Here, we present the approach to linking records across PQDT and MAG. We use the open-source library `dedupe` (Gregg and Eder, 2022) to train a record-linking algorithm. The record-linking algorithm is trained by labeling proposed links across two datasets. However, two random records across datasets are almost never a link. To avoid this problem, the algorithm blocks records into groups based on the similarity of the features used for linking. The user is then asked to label proposed links chosen from pairs of records blocked together and some randomly chosen pairs. The proposed pairs the user is supposed to label are chosen to be the ones that the algorithm is least certain about. This procedure allows the fuzzy linking of very large databases with limited cost in terms of labeling and computation.

We separately link PhD graduates and PhD advisors from PQDT to MAG. Further, we separately link within each field.

Using the following features, we search for one-to-one links between graduates in PQDT and authors in MAG:

- First name, last name and middle name.
- Dissertation keywords and the keywords in the publications at the start of the publishing career in MAG. The keywords are the fields of study entity from the MAG semantic language model. We use the fields at level 1 and aggregate fields at levels 2 to 5 to their most likely parent at level 1. We only consider fields where the algorithm is confident enough about the field (a score of at least 0.4).
- The year of graduation in PQDT and the year of first publication in MAG
- Thesis title in PQDT and the paper titles in MAG.

Because an author in MAG can be an advisor for multiple theses, we search for many-to-one links between advisors in PQDT and authors in MAG. We use the following features:

- First name, last name and middle name.
- The student's graduating year in PQDT and the start and end year of the advisor's publishing career.
- Dissertation university and MAG affiliation, both crosswalked to the Carnegie List, as described above (American Council on Education, 2021).

A.4 Training

Each of the two authors of this paper created one training sample per field of study. We provided 40 to 60 labeled pairs and followed the following protocol. To label a proposed link as a true link, the following needs to hold

1. For students
   (a) Similar names.
   (b) Overlap in paper titles or overlap in keywords.
   (c) Graduating year and year of first publication not more than 10 years apart.
2. For advisors
   (a) Similar names.
   (b) The student does not graduate before the year of the first publication of the advisor.
   (c) The advisor is at a university with a similar name as the student in a window around the student's graduation.[9] Proposed links that are more than 10 years apart are labelled as "no", even if the previous conditions are true.

[9]This often resulted in requiring exactly matching university names, although in some special cases (such as the University of California system) it did not.

We then label records for each of the research fields. Within the above boundaries, each author judges the similarity of records to actively label proposed links.

A.5 Postprocessing of the predicted links

In this step, we combine the predicted links from the two models into a single prediction. The step is based on a comparison as illustrated in table 4 for graduates and table 5. For each field, we compare the predicted links of the two models. Starting from the sample of graduates/advisors in ProQuest, we classify the predicted links into four groups: those where the predicted MAG entity is the same (column "Same entity"), where only one of the models found a link ("Only by 1", "Only by 2"), and where both models found a link but to different entities ("Different entity"). The table reports these fractions for each field of study and a weighted average across all fields. We will return to these numbers

in section A.6.

Table 4: Linking the graduates

	Fraction of links found				
Field	Same entity	Only by 1	Only by 2	Different entity	Number of links
Art	0.48	0.52	0.00	0.00	1717
Biology	0.67	0.29	0.03	0.01	51420
Business	0.67	0.15	0.13	0.05	12147
Chemistry	0.86	0.12	0.01	0.01	24490
Computer Science	0.88	0.04	0.07	0.02	19933
Economics	0.76	0.16	0.07	0.01	7680
Engineering	0.59	0.24	0.14	0.04	34597
Environmental Science	0.79	0.03	0.17	0.01	5091
Geography	0.66	0.29	0.04	0.01	4097
Geology	0.68	0.26	0.05	0.01	5260
History	0.88	0.07	0.05	0.01	5574
Materials Science	0.46	0.26	0.22	0.05	9771
Mathematics	0.63	0.32	0.04	0.01	12239
Medicine	0.44	0.22	0.31	0.03	8645
Philosophy	0.85	0.06	0.08	0.01	2721
Physics	0.62	0.04	0.31	0.03	7459
Political Science	0.85	0.08	0.06	0.01	7107
Psychology	0.89	0.06	0.04	0.01	33327
Sociology	0.71	0.05	0.23	0.01	4742
Total	0.69	0.19	0.10	0.02	258017

Note: The table summarizes the links found from ProQuest graduates to MAG authors. Graduates are defined as the authors of the dissertations in ProQuest. First, the columns headed by "Fraction of links found" compare the identified links across two different labelers as described in the text. The columns show the fraction of links found for two training sets constructed by two different labelers. "Same entity" are graduates for which the models trained on different training sets find the same MAG identifier. "Only by 1" and "Only by 2" are graduates for which only one of the two separately trained models found a link to MAG. "Different entity" are graduates for which both models find links to MAG but to different identifiers. Second, the last column reports the total number of links found for each field after postprocessing (see text for details). Third, the last row reports the total across fields. The fractions are weighted by the number of graduates in the respective fields in ProQuest 1990–2014.

Table 5: Linking the advisors

	Fraction of links found				
Field	Same entity	Only by 1	Only by 2	Different entity	Number of links
Art	0.88	0.00	0.11	0.01	10628
Biology	0.78	0.00	0.22	0.00	101852
Business	0.69	0.01	0.30	0.00	30963
Chemistry	0.93	0.00	0.02	0.05	48670
Computer Science	0.89	0.08	0.02	0.01	39618
Economics	0.96	0.00	0.02	0.02	27266
Engineering	0.81	0.00	0.19	0.00	99106
Environmental Science	0.73	0.23	0.03	0.01	11055
Geography	0.62	0.32	0.02	0.03	12229
Geology	0.84	0.00	0.14	0.01	11478
History	0.85	0.01	0.12	0.02	30228
Materials Science	0.74	0.01	0.13	0.12	23998
Mathematics	0.78	0.15	0.01	0.06	32938
Medicine	0.94	0.00	0.02	0.04	19264
Philosophy	0.93	0.00	0.06	0.00	12153
Physics	0.51	0.15	0.03	0.31	19825
Political Science	0.87	0.00	0.08	0.04	24309
Psychology	0.91	0.00	0.08	0.00	94144
Sociology	0.46	0.01	0.51	0.02	20914
Total	0.83	0.02	0.12	0.03	670638

Note: The table summarizes the links found from ProQuest advisors to MAG authors. An advisor is one relationship id as described in the text. First, the columns headed by "Fraction of links found" compare the identified links across two different labelers as described in the text. The columns show the fraction of links found for two training sets constructed by two different labelers. "Same entity" are advisors for which the models trained on the different training sets find the same MAG identifier. "Only by 1" and "Only by 2" are advisors for which only one of the two separately trained models found a link to MAG. "Different entity" are advisors for which both models find links to MAG but to different identifiers. Second, the last column reports the total number of links found for each field after postprocessing (see text for details). Third, the last row reports the total across fields. The fractions are weighted by the number of graduates in the respective fields in ProQuest 1990–2014.

We combine the links as follows. First, using the comparison previously explained, we only keep predicted links where the predictions from the two models agree. This means that, for both graduate and advisor entities, we accept links of three kinds:

1. The two models link the same MAG entity identifier to the ProQuest entity.

2. Only one of the models predicts a link, but the entity name in MAG is very similar to the entity name in ProQuest—a Jaro-Winkler similarity of 0.9 or more.

3. The two models assign a different MAG entity to the same ProQuest entity. Here, we only keep cases where two conditions hold: First, the names of the two entities in MAG are almost identical (a Jaro-Winkler similarity of 0.99 or more). Second, the entity of one of the predicted links publishes more than five times as many papers over the whole career as the entity of the other predicted link. This case covers duplicated entities in MAG, where a few papers are not merged to the main entity of an author.

Second, we deal with duplicates that arise from linking fields separately. For graduates, we require that both the MAG and the ProQuest entity are uniquely linked; in other words, if a graduate in ProQuest is linked to different MAG entities in the models for chemistry and biology, we drop them. This can happen if the dissertation field is not uniquely assigned to chemistry or biology. For advisors, we require that the ProQuest entity be uniquely linked to MAG, but not vice versa, since advisors can supervise multiple theses.

The last row in tables 4 and 5 report the number of links that result at the end of this postprocessing step. In total, we find around 258'000 one-to-one links for graduates and 670'000 many-to-one links (from many dissertations to one MAG entity) for advisors.

A.6 Empirical performance of the linking algorithm

We discuss the performance of the record linkage in terms of precision and recall and the role of human error in training the linking algorithm.

Precision and recall In prediction tasks where a representative sample of correctly labeled pairs exists, one can check the performance of the prediction by comparing predicted to true labels. Because of active learning, the labeled records are not randomly selected and, therefore, unsuitable for calculating recall and precision. Nevertheless, we suggest a lower bound for the precision in linking graduates and an approximation to recall for advisors.

First, we provide a check on the precision of our links for graduates in chemistry. Precision is defined as the number of true links divided by the total number of links made. For chemistry, we provide a test based on correctly linked graduates: First, because most graduates already publish during their dissertation (Gaulé and Piacentini, 2013), we can calculate the fraction of our linked chemistry graduates that satisfy this condition, conservatively discarding any other links as false positives. Second, because we do not link on the name of the PhD university, we can compare the name of the alma mater to the affiliation name of the linked author entity in MAG before graduation. We find that 18% of linked graduates do not publish during their dissertation. Further, 4% of those that publish do not publish at their alma mater. Therefore, at least a fraction of $(1 - 0.18) \times (1 - 0.04) = 0.78$ linked records was plausibly linked correctly. This performance is in the upper tercile of the methods discussed in Bailey et al. (2020, Table 1) for historical US census data.

Second, because advisors tend to be established researchers and likely publish regularly, we should find a high fraction of advisors in the MAG data. We find about 75 percent or more advisors in the MAG data in most fields and years.[10] This indicates that our linking strategy has a high recall for advisors, i.e., a high share of links found out all links that could be found. This also provides reassuring evidence that the affiliation information in MAG is accurate since it is required to identify co-author connections across universities.

[10]The exceptions are philosophy, where we find between 50 and 75 percent of advisors in each year, and business, where the fraction of advisors linked declines from 0.75 to 0.6 over the sample period.

Human error As the algorithm is trained through active labeling by humans, errors in labeling may propagate to a biased linked sample. By ensuring consistency between two individual labelers, the postprocessing of links reduces such concerns. Moreover, the linking statistics in Tables 4 and 5 show that the models make very similar predictions.[11]

First, the models agree for 69 percent of linked graduates and 83 percent of linked advisors. In most fields, the agreement is at least 60 percent. Second, the predicted MAG entity rarely differs. Across fields, this number is at most five percent for graduates. It is similar for advisors except for Materials Science and Physics. Third, the columns "Only 1" and "Only 2" also show that one of the two models is typically more conservative than the other. However, the postprocessing only keeps such links if the names of the linked entities are very similar.

A.7 Constructing the analysis sample

We look at graduates between 1990 and 2014 whom we can link to MAG and for whom we can link at least one advisor to MAG.

In the main text in Section 2, we describe the construction of the main variables used in the analysis. Here, we provide additional detail ommited from the main text.

The share of PhD graduates our sample represents is shown in Table 6. ProQuest records approximately 1.1 million dissertations from US universities between 1990 and 2014. Matching these dissertations to universities and colleges listed in the Carnegie Classification leaves us with 1.04 million theses. We link approximately 24% of these observations to MAG, where we restrict the MAG AuthorIds to those who have at least once a recorded affiliation in the US. We restrict our sample further to graduates whose advisors are also linked in MAG. Graduates must publish at least one paper within the time frame of one to seven years after their graduation.

Further, they need to have at least one recorded affiliation on the post-PhD publications. Lastly, that recorded affiliation needs to be on the Carnegie list and different from

[11]These similarities do not stem from overlap in the training sets: We verified in one field (graduates chemistry) that the training sets of the two labelers do not overlap (for one labeler, two out of 50 pairs are also in the training set of the other labeler).

the PhD awarding university. These restrictions leave us with 7.4% of the initial sample and 31.2% of the linked graduates.

Table 6: Sample selection

Sample Cut	N	Share	Share of linked graduates
Proquest Dissertations, US location, 1990-2014	1116087	100.0	
University in Carnegie list	1044436	93.6	
Linked to MAG AuthorId (with US affiliation)	263602	23.6	100.0
Advisor Linked to MAG AuthorId	212303	19.0	80.5
At least 1 paper post-PhD	194315	17.4	73.7
Post-PhD affiliation recorded	130323	11.7	49.4
Post-PhD affiliation in Carnegie List	96176	8.6	36.5
Post-PhD affiliation != PhD affiliation	82279	7.4	31.2

The sample is not a random representation of all PhD graduates but focuses on those in US academia with a publication record. This limits generalizability, especially to graduates entering industry, government, or non-academic sectors. While not fully representative, the sample is well-suited for analyzing academic trajectories and productivity.

A.7.1 Field composition of sample

In Figure 5 we show the field composition of our analysis sample of PhD graduates. The sample is not representative of the field distribution of all PhD graduates. Thus, to complement the analysis in the paper, we also show the main results split by field in Appendix B.2 and B.3.4.

A.7.2 Topical Fit

An innovation in our paper is a proxy for the topical fit between PhD graduates and universities, which we use as controls in our regressions. We compute these metrics from

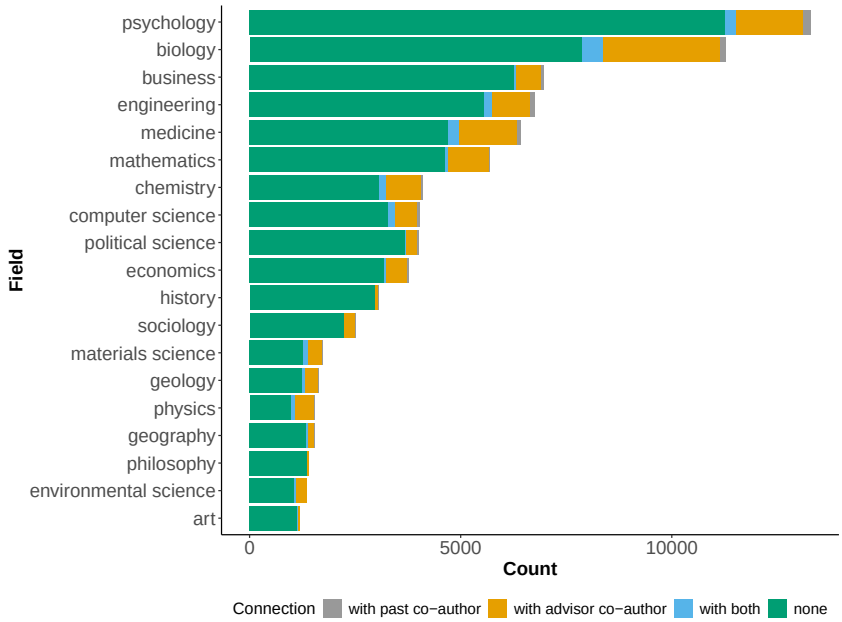


Figure 5: Field composition of PhD graduates in analysis sample

dissertation abstracts using the following methodology.

Our starting point is MAG's language model (Wang et al., 2020, section 2.5) that assigns Fields of Study to scientific text. The MAG database contains assigned Fields of Study for each paper, and their corresponding score. We apply MAG's language model to the dissertation abstract (or to the title if the abstract is missing), recording the ten highest-scoring Fields of Study for each and the corresponding score, as long as the score is above 0.4. This cutoff was chosen heuristically, as additional Fields of Study with low scores appear to be unrelated to the content of the abstracts.

We use these data to compare the research topics of graduates' dissertations and hiring universities' research output. As research topics, we use the Fields of Study of the MAG classification up to level 2 and show robustness regarding this choice in Appendix B.2.4.

For the documents in MAG, we use the same restriction on document types as before; that is, we include papers classified as "Journal," "Book," "Book Chapter," and "Conference." Microsoft has already assigned the topics for documents in MAG.

For hiring universities, we consider research output in the five years before the graduate's dissertation. We consider the set of authors in MAG sharing the given affiliation and

having publication careers in the same field (e.g., "economics"). We determine authors' affiliations based on the universities at which they publish most of their papers each year. For all department authors, we calculate the average score for each Field of Study across all papers published in the five years before the graduate's dissertation. We compute the hiring university × Field's overall topic vector as the sum of the authors' topic vectors.

Finally, we compute two cosine similarities for each pair of graduate and hiring university to proxy how well the graduate's research concepts align with the hiring university's. The first similarity compares the graduate's concept vector to the hiring university × Field's overall topic vector, capturing how well the graduate's research fits the general specialization. The second similarity is the maximum cosine similarity of the similarities between the graduate's topic vector and each individual author affiliated with the hiring university publishing in the same field[12]. A high score in this metric indicates that there is a researcher with very similar research interests to the PhD graduate in the hiring department.

The calculation of the cosine similarity in detail:

- $\vec{S_i}$ is the topic vector where each entry $n = 1, \ldots, N$ is the sum of the scores of topic n over the set of documents P_i by author i: $S_{i,n} = \sum_{p \in P_i} s_{n,p}$. Normalizing the topic vector does not affect the calculated similarity in the next steps.

- Similarity between two authors i, i':

$$cos\ similarity(i, i') = \frac{\sum_n (S_{i,n} * S_{i',n})}{\sqrt{\sum_n S_{i,n}^2}\sqrt{\sum_n S_{i',n}^2}}$$

- Max similarity between graduate i and department j. The set I_j of authors at department j to be considered are those,

  - who are affiliated with department j within the 5 years up to graduation of i,

  - whose main field is the same as of graduate i.

[12]To limit computational burden, we restrict attention to the 200 authors publishing the most papers in the given time period, affiliated with the hiring university and publishing in the same field. In practice, this is rarely a restriction.

– We restrict to documents in the last 5 years (up to graduation year of i).

- Max similarity between graduate i and department j is defined as

$$max\ cos\ similarity(i,j) = \max_{i' \in I_j} cos\ similarity(i,i').$$

- Average similarity between graduate i and department j is the cosine similarity between the topic vector for all researchers in I_j, defined as $\bar{S}_j = \sum_{i' \in I_j} \vec{S}_{i'}$.

$$avg\ cos\ similarity(i,j) = \frac{\sum_n (S_{i,n} * \bar{S}_{j,n})}{\sqrt{\sum_n S_{i,n}^2}\sqrt{\sum_n \bar{S}_{j,n}^2}}.$$

A.7.3 Quality of Affiliation Information

The average number of advisor connected affiliations per student increases from 5 to 15 over the sample period, as shown in Figure 1b. This rise, however, may partly reflect improvements in the coverage of affiliation information in MAG. To address this, Figure 6. Specifically, it shows the proportion of authors with at least one document in MAG in a given year who have available affiliation information. We plot the average of this share for each dissertation year, separately for advisors and graduates.

The figure reveals that affiliation coverage improves over time, increasing by nearly 20%. However, the number of connected affiliations rises by 300%. Even if the entire increase in affiliation coverage was attributed to measurement error, the observed growth in connected affiliations is still an order of magnitude larger. Thus, at most, a small share of the observed rise in the number of connected affiliations can be attributed to better coverage of affiliation information.

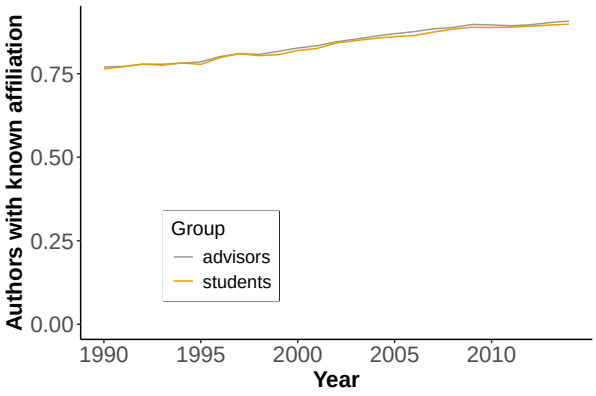


Figure 6: Share of (linked) Authors with known affiliation conditional on publishing in MAG by degree year in ProQuest Dissertations&Theses.

B Additional Results

B.1 Sample Selection

In this section, we provide additional evidence on the selection of observations that end up in our main analysis. Our main analysis examines post-PhD outcomes for graduates based on whether they join a connected affiliation. A potential concern is that graduates might be systematically selected into the sample based on their network characteristics. Specifically, graduates with certain network traits could be more likely to become connected hires due to factors that also influence post-PhD outcomes. If true, this would imply a correlation between being a connected hire and the unobserved error term u_i in the outcome equation (4).

While we cannot directly test this, we can investigate whether the likelihood of being included in the analysis is related to a key determinant of becoming a connected hire: the number of connected affiliations.

To do this, we estimate a logit model predicting inclusion in the analysis sample among all graduates that we could link to MAG. The predictors include variables constructed from research output up to the PhD graduation year, similar to those in our main post-PhD analysis. However, instead of expected citations based on publication venues, we use 10-year forward citations[13]. All variables are discretized into deciles within Field×5-year windows.

Table 7 shows the results. Columns 1 and 2 show the regression result regarding sample inclusion. Column 1 includes only productivity-related predictors, the number of papers and their citations of the graduate, and the citations of the advisor. Column 2 adds the size of the collaboration network, measured by the number of connected affiliations. The Pseudo R^2 changes minimally between these two specifications, suggesting that network size does not predict whether a graduate is included in our sample. This result is reassuring, as it indicates that sample selection is not driven by exposure to networks (conditional on PhD productivity).

[13]Here, we are not concerned that the information we use is not available in real-time for market participants; we want to document whether there is systematic selection into the sample

Next, we focus on the analysis sample to examine whether the number of connected affiliations predicts joining a connected affiliation. Columns 3 and 4 in Table 7 present these results, showing that network size is highly predictive of becoming a connected hire. This finding echoes our results on matching discussed in Section 4.1.

In summary, network size is closely related to the probability of a graduate joining a connected affiliation but is unrelated to whether they are included in our analysis sample. This reassures us that a key predictor of the variation in our main analysis—whether a graduate is a connected hire—is conditionally independent of whether a graduate is included in the sample.

Table 7: Logit Regression: In-Sample and Connected Post-PhD affiliation

Dependent Variables:	In sample		Connected advisor	
Model:	(1)	(2)	(3)	(4)
Fixed-effects				
Field×5 Year Window	Yes	Yes	Yes	Yes
Field×N Papers PhD Decile	Yes	Yes	Yes	Yes
Field×N Citations PhD Decile	Yes	Yes	Yes	Yes
Subfield	Yes	Yes	Yes	Yes
Field×Advisor Citations	Yes	Yes	Yes	Yes
N PhD Connections Decile×Field		Yes		Yes
N Advisor Connections Decile×Field		Yes		Yes
Fit statistics				
Pseudo R^2	0.126	0.129	0.143	0.178
Observations	212,254	212,254	80,992	76,728

Notes: The unit of observation is a PhD graduate. Columns (1) and (2) include all graduates linked to MAG and whose advisor also has been linked, independently of their post-PhD affiliation. In columns (3) and (4), we use the analysis sample, as in the main text.

B.2 Network connections and the first affiliation of PhD graduates

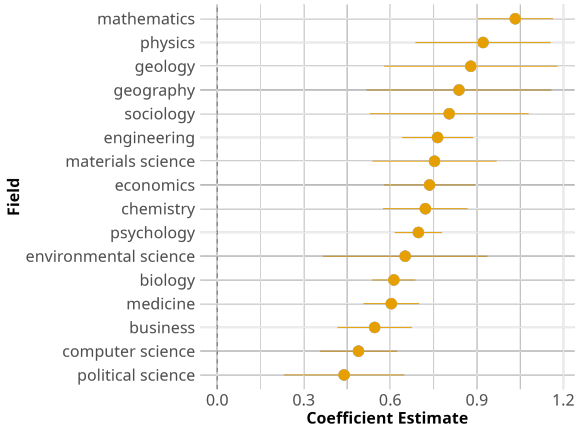
In this section, we present additional results on the role of the PhD advisor's co-author connections and the matching of PhD graduates to universities.

B.2.1 Heterogeneity by field

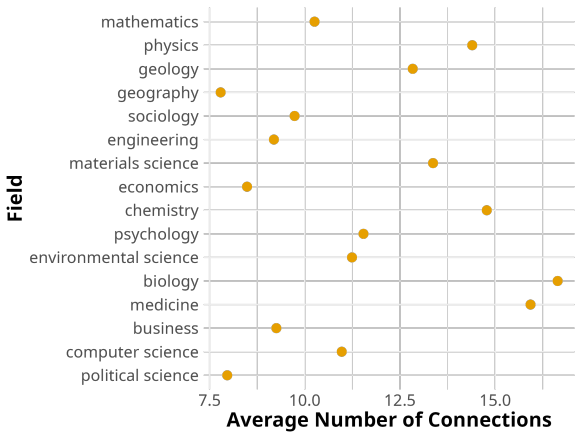
Figure 7a displays the estimated coefficients for advisor connections along with their 95% confidence intervals by field. Due to the limited sample size, we do not report results for History, Art, and Philosophy. The estimates correspond to the specification reported in the last column of the first panel of Table 2. Fields are ordered by the point estimate on "Advisor Connection", with Mathematics and Physics at the top, showing an estimated coefficient close to 1, decreasing to just above 0.4 for political science.

In Panel 7b, we present the average number of advisor-connected affiliations by field, which shows no systematic relationship with the estimated coefficients.

The results reveal heterogeneity in the estimated effect of advisor connections on placement across fields, but the estimated effect is significantly positive in all cases.



(a) Multinomial Logit: Advisor Connection Coefficient



(b) N Advisor Connected Universities

Figure 7: Matching PhD Graduates to Universities by Field: Advisor Connections

B.2.2 Placebo advisor analysis

Our main results highlight the role of the PhD advisor's co-author connections for the placement of PhD students. Here, we present a placebo analysis by repeating the matching analysis with a "placebo advisor." To do so, we randomly draw another advisor from the graduate's PhD class and repeat the estimation of equation (1) in the main text with the placebo advisor's connections. If the personal connection of the advisor with the student is important, then the estimated effect should be substantially lower for the placebo advisor. The placebo connections might pick up systematic matching based on shared characteristics that unobserved to us. Table 8 shows the results. Placebo connections predict a small decrease in the probability of matching, conditional on controls in the main specification in column (6). The absolute values are small compared to the main effect.

Table 8: Matching - Placebo connection

Dependent Variable:	Match formed				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Advisor Connection	0.728	0.744	0.740	0.742	0.681
	(0.024)	(0.025)	(0.025)	(0.025)	(0.023)
Placebo Advisor Connection	0.007	-0.045	-0.046	-0.046	-0.047
	(0.024)	(0.017)	(0.017)	(0.017)	(0.017)
PhD connection	1.04	1.12	1.12	1.11	1.07
	(0.038)	(0.049)	(0.049)	(0.049)	(0.048)
PhD Class×University ID		Yes	Yes	Yes	Yes
Advisor Citation Decile×University Citation Decile×Field			Yes	Yes	Yes
Pre Graduation Productivity×University Citation Decile×Field				Yes	Yes
Connections Decile×University Citation Decile×Field				Yes	Yes
Varying Slopes					
Max similarity to faculty members ×Field					Yes
Avg similarity to faculty members ×Field					Yes
Fit statistics					
Observations	650,527	650,527	650,527	650,527	650,527

Notes: Placebo connections are constructed by randomly assigning a PhD graduate another "placebo" advisor from her PhD class. The placebo advisor's connections are then used as an additional type of network in equation (2).

B.2.3 Connections - Decay with time since last collaboration

In our main results, we considered co-author connections of the advisor independently of the year of the collaboration. Here, we separate connections by the most recent year of collaboration between the advisor and the co-author at another institution. We expect connections further distant in the past to be weaker predictors of placement. Figure 8 shows the results by the difference between the year of graduation of the PhD graduate and the most recent collaboration year of the PhD advisor and their connection. Years since the last collaboration are binned into 3-year windows. The effect is strongest for connections from collaborations in the 3 years preceding the year of graduation of the focal PhD graduate and the point estimate decays almost linearly with time.

We see the year of most recent collaboration as a proxy for the strength of the connection between the advisor and the co-author at another institution, as it relates to how recent contact was. In line with that interpretation, the estimated effect of a co-author connection decays with time since the last collaboration.

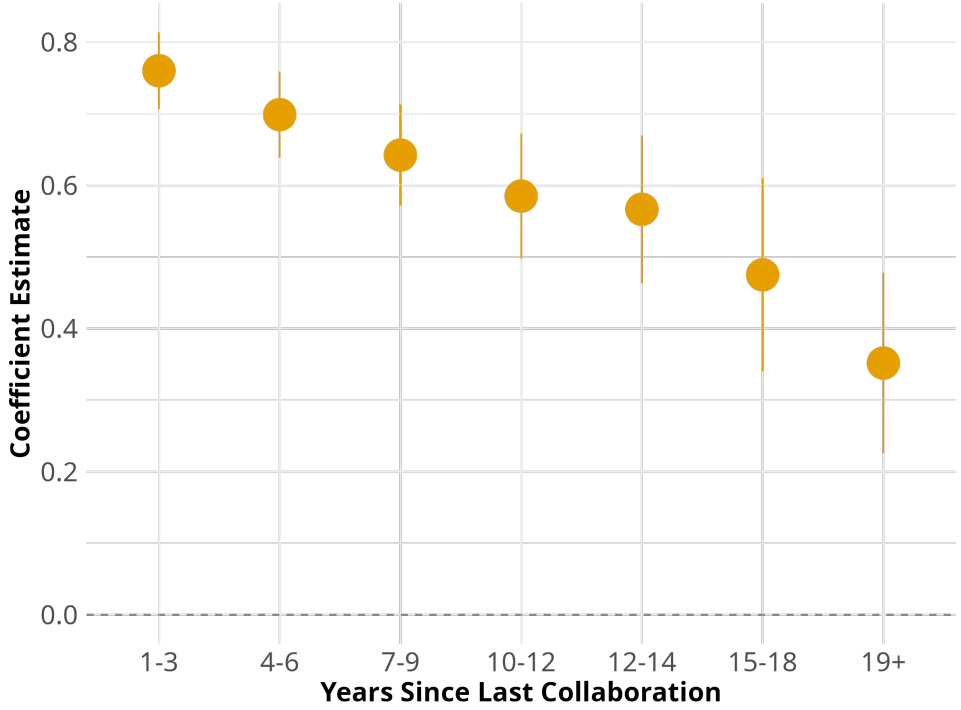


Figure 8: Matching: Advisor connection by most recent collaboration year

Estimates and associated 95 % confidence intervals of γ for the advisor's network by distance of most recent collaboration year to graduation year of focal PhD graduate.

B.2.4 Topic Similarity measurement

We use the MAG field hierarchy to assign detailed topics to researchers' output. In the main analysis, we consider all fields up to level two. Here, we compare results when we successively include all levels of the MAG field hierarchy.

Table 9: Matching - Topic similarity by maximum MAG field hierarchy level

Dependent Variable:	Match formed				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Advisor connection	0.718	0.684	0.668	0.663	0.662
	(0.024)	(0.023)	(0.023)	(0.022)	(0.022)
PhD connection	1.10	1.07	1.05	1.04	1.04
	(0.049)	(0.048)	(0.048)	(0.047)	(0.048)
Fixed-effects					
PhD Class×University ID	Yes	Yes	Yes	Yes	Yes
Advisor Citation Decile×University Citation Decile×Field	Yes	Yes	Yes	Yes	Yes
Pre Graduation Productivity×University Citation Decile×Field	Yes	Yes	Yes	Yes	Yes
Connections Decile×University Citation Decile×Field	Yes	Yes	Yes	Yes	Yes
Varying Slopes					
Max similarity to faculty members (Lvl 1) (Field)	Yes				
Avg similarity to faculty members (Lvl 1) (Field)	Yes				
Max similarity to faculty members (Lvl 2) (Field)		Yes			
Avg similarity to faculty members (Lvl 2) (Field)		Yes			
Max similarity to faculty members (Lvl 3) (Field)			Yes		
Avg similarity to faculty members (Lvl 3) (Field)			Yes		
Max similarity to faculty members (Lvl 4) (Field)				Yes	
Avg similarity to faculty members (Lvl 4) (Field)				Yes	
Max similarity to faculty members (Lvl 5) (Field)					Yes
Avg similarity to faculty members (Lvl 5) (Field)					Yes
Fit statistics					
Observations	650,527	650,527	650,527	650,527	650,527

When we increase the detail of the topic measurement of research by increasing the maximum level from one to five in the MAG field hierarchy. The estimated coefficient on connections changes slightly when adding level two but is almost constant afterward. Thus, we report the estimate with topics up to level two of MAG field hierarchy in the

main text.

B.2.5 Adding the interaction between PhD's and Advisor's connection

Here, we repeat the matching analysis presented in the main text but add the interaction between the PhD's and Advisor's connection. This effectively separates the connection status into three cases: only the advisor is connected, only the PhD graduate is connected, and lastly, both are connected. The results are presented in Table 10.

The coefficient on the advisor connection is for the first case, the PhD connection coefficient for the second case, and the sum of all three coefficients corresponds to the third case.

This is useful to separate out to what extent the estimate on the advisor connection may reflect the graduate's own connections due to overlap. The advisor connection estimates slightly increase compared to the main text (from 0.68 to 0.73). When both the advisor and PhD graduate are connected, the estimated effect is somewhat smaller when considering the interaction ($0.73 + 1.8 - 1.03 = 1.5$ instead of $0.684 + 1.07 = 1.754$ in the main text). And lastly, the estimated coefficient for the case when only the PhD is connected is substantially larger. However, this is quantitatively irrelevant, given it applies to a very small number of cases. Thus, adding the interaction term does not substantially alter the main results regarding advisor connections. Still, we consider the interaction when calculating the overall predicted role of connections.

Table 10: Matching results with interaction between PhD and advisor connection

Dependent Variable:	Match formed				
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Advisor connection	0.772	0.794	0.790	0.791	0.730
	(0.025)	(0.026)	(0.026)	(0.027)	(0.024)
PhD connection	1.68	1.86	1.87	1.86	1.80
	(0.055)	(0.068)	(0.069)	(0.069)	(0.069)
Advisor connection × PhD connection	-0.899	-1.04	-1.05	-1.04	-1.03
	(0.066)	(0.079)	(0.080)	(0.080)	(0.082)
Fixed-effects					
PhD Class×University ID		Yes	Yes	Yes	Yes
Advisor Citation Decile×University Citation Decile×Field			Yes	Yes	Yes
Pre Graduation Productivity×University Citation Decile×Field				Yes	Yes
Connections Decile×University Citation Decile×Field				Yes	Yes
Varying Slopes					
Max similarity to faculty members ×Field					Yes
Avg. similarity to faculty members ×Field					Yes
Fit statistics					
Observations	650,527	650,527	650,527	650,527	650,527

B.2.6 Similarity descriptives and coefficients

In Table 11, we report the coefficients on the similarity controls that we omitted in Table 2. The coefficients for the topic similarity controls are shown in the first two columns. In columns 3 and 4 we show the standard deviation of the similarity measures. We scale the coefficients by the sample standard deviation of the similar measures in columns 5 and 6 to compare them to the estimated coefficients on network connections. The one-standard-deviation scaled coefficients are almost all positive and range from approximately zero to just above one. Thus, they are of similar magnitude as the estimated coefficients on advisor connections. Importantly, the estimates show that the topic similarity measures capture information relevant to matching.

Table 11: Matching: Coefficients on similarity measures and their standard deviation

Field	(1) Max similarity	(2) Avg similarity	(3) SD Max	(4) SD Avg	(5) Coef × SD Max	(6) Coef × SD Avg
Art	-0.140	13.861	0.113	0.076	-0.016	1.060
	(2.841)	(9.263)				
Biology	2.081	9.474	0.170	0.077	0.354	0.729
	(0.164)	(0.630)				
Business	1.383	3.108	0.115	0.102	0.159	0.318
	(0.317)	(0.750)				
Chemistry	1.371	12.277	0.125	0.061	0.172	0.748
	(0.438)	(1.538)				
Computer Science	2.156	7.022	0.130	0.119	0.280	0.832
	(0.470)	(1.706)				
Economics	0.920	4.040	0.102	0.086	0.094	0.346
	(0.518)	(1.478)				
Engineering	0.990	6.963	0.111	0.048	0.110	0.337
	(0.313)	(1.072)				
Environmental Science	2.542	4.742	0.147	0.108	0.375	0.511
	(0.883)	(2.042)				
Geography	1.557	3.599	0.113	0.087	0.176	0.315
	(1.051)	(3.236)				
Geology	2.331	5.101	0.151	0.086	0.351	0.440
	(0.741)	(2.074)				
History	0.397	4.226	0.122	0.106	0.048	0.447
	(1.077)	(2.158)				
Materials Science	1.941	17.823	0.115	0.062	0.223	1.107
	(0.594)	(3.128)				
Mathematics	2.567	10.469	0.123	0.106	0.316	1.107
	(0.397)	(1.063)				
Medicine	2.555	6.314	0.129	0.079	0.329	0.502
	(0.272)	(1.114)				
Philosophy	-1.023	21.934	0.090	0.059	-0.092	1.296
	(3.847)	(11.192)				
Physics	5.125	5.769	0.174	0.102	0.891	0.591
	(0.839)	(1.261)				
Political Science	1.296	1.510	0.131	0.119	0.170	0.179
	(0.735)	(1.118)				
Psychology	2.603	5.584	0.142	0.174	0.369	0.974
	(0.247)	(0.534)				
Sociology	-0.714	4.670	0.114	0.088	-0.081	0.412
	(0.781)	(1.755)				

Notes: The coefficients for the topic similarity controls are shown in the first two columns. They are estimated in the regression presented in the last column in Table 2 Panel A. The standard errors of coefficients are in parentheses. In columns 3 and 4 we show the standard deviations of the similarity measures. The "Coef × SD" columns multiply the coefficient by one standard deviation of the similarity measure.

B.2.7 Matching regression without PhD-class × Hiring University fixed effect

Our main analysis of matching between PhD graduates and universities includes a fixed effect at the level of PhD class×Hiring University pair. This implies that only pairs of PhD class×Hiring University with variation in connection status and outcome status contribute to the estimation of the connection coefficient γ. For the remaining cases, the fixed effect for PhD class×Hiring University perfectly explains the outcome, and thus, they do not contribute to the estimation of γ. That is why the main analysis of matching in the paper uses 650,000 dyads out of approximately 30,000,000 possible dyads. A dyad is a pair of PhD graduate and hiring university.

Here, we repeat the analysis with separate fixed effects for PhD class and hiring university. This allows the inclusion of almost all dyads in the regression, as now it is only required that there is variation in connection status and outcomes within the hiring university and separately within PhD class. This is true for most possible dyads.

$$P(d_{i,j} = 1) = \frac{e^{(\phi_j + \beta X_{i,j} + \gamma A_{i,j})}}{\sum_{j'} e^{\phi_{j'} + \beta X_{i,j'} + \gamma A_{i,j'}}}. \tag{6}$$

The results from the regression with just the fixed effect for the hiring university × Field × 5-year window are shown in Table 12 in column 2. We also add a fixed effect for the interaction between the rank of the PhD granting university and the rank of the hiring university in column 3. Similar to the main results in the paper, in Table 2, the estimated coefficient is stable across specifications (starting with column 2). However, the point estimate is larger compared to the main results. This can be due to the different fixed effects or the larger sample. To determine which explanation applies, we show the results with similarity controls for the extended sample in column 5 and the smaller sample used in the main analysis in column 6. The estimate in column 6 is very close to the main estimate. This suggests that the fixed effect for the interaction between PhD class and hiring university does not alter the estimate much but that the estimate depends on the sample. In particular, the estimate is larger in the extended sample. Thus, our main estimates are conservative regarding the role of co-author connections for

placement.

Table 12: Multinomial Logit without interacted PhD class×Hiring University fixed effect

Dependent Variable:	Match formed					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Advisor connection	2.09	1.21	1.20	1.16	1.11	0.664
	(0.055)	(0.031)	(0.031)	(0.031)	(0.029)	(0.021)
PhD connection	1.45	1.06	1.05	1.00	0.939	1.01
	(0.046)	(0.054)	(0.054)	(0.053)	(0.052)	(0.040)
Fixed-effects						
University ID×Field×5-Year Window		Yes	Yes	Yes	Yes	Yes
PhD University Citation Decile×University Citation Decile×Field			Yes	Yes	Yes	Yes
Advisor Citation Decile×University Citation Decile×Field				Yes	Yes	Yes
Pre Graduation Productivity×University Citation Decile×Field				Yes	Yes	Yes
Connections Decile×University Citation Decile×Field				Yes	Yes	Yes
Varying Slopes						
Max similarity to faculty members×Field					Yes	Yes
Avg similarity to faculty members×Field					Yes	Yes
Fit statistics						
Observations	28,564,555	28,564,555	28,564,555	28,564,555	28,564,555	650,519

Notes: Unit of observation is a pair of PhD graduate and a hiring university. See equation (2) for the multinomial logit specification. Clustered (PhD university×Field×5 year window + hiring university) standard-errors in parentheses

B.3 Post-PhD outcomes

B.3.1 Post-PhD outcomes without controls

In the main text, we presented results for most post-PhD outcomes only with controls. Here, we present the results without the controls for pre-graduation productivity of the PhD student and their advisor. These additional results show the controls' role in explaining the differences between connected and non-connected hires. Pre-graduation productivity predictors are informative about citations and the number of papers post-PhD, but less so for the remaining outcomes. This is also true for the coefficients on connections, i.e., they hardly change for the remaining outcomes.

Table 13: Post-PhD outcomes of connected vs. not-connected hires without controls

Panel A: Comparison of post-PhD outcomes with Class Fixed Effect

Dependent Variables:	N Cites PhD graduate	N papers	Co-authors First Affil	Same Affil PhD+6yrs	Any output PhD+6yrs	N Cites of First Affil
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Advisor connection	0.238	0.109	0.204	-0.143	-0.002	0.523
	(0.021)	(0.011)	(0.019)	(0.013)	(0.004)	(0.080)
PhD connection	0.232	0.195	-0.014	-0.514	-0.008	0.019
	(0.030)	(0.016)	(0.032)	(0.032)	(0.006)	(0.041)
Fixed-effects						
PhD Class	Yes	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Pseudo R^2	0.43	0.24	0.34	0.04	0.01	0.74
Observations	82,141	82,279	77,484	79,895	81,771	82,092

Panel B: Comparison of post-PhD outcomes with Destination Fixed Effect

Dependent Variables:	N Cites PhD graduate	N papers	Co-authors First Affil	Same Affil PhD+6yrs	Any output PhD+6yrs
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Advisor connection	0.117	0.058	0.036	-0.059	-0.003
	(0.022)	(0.009)	(0.013)	(0.012)	(0.004)
PhD connection	0.208	0.177	-0.008	-0.469	-0.011
	(0.026)	(0.014)	(0.028)	(0.037)	(0.007)
Fixed-effects					
Field×5 Year Window	Yes	Yes	Yes	Yes	Yes
Hiring University Id×Field × 5 year window	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Pseudo R^2	0.51	0.32	0.36	0.06	0.01
Observations	81,660	82,279	67,755	75,833	80,463

Panel C: Comparison of post-PhD outcomes with Class and Destination Fixed Effect

Dependent Variables:	N Cites PhD graduate	N papers	Co-authors First Affil	Same Affil PhD+6yrs	Any output PhD+6yrs
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Advisor connection	0.080	0.049	0.029	-0.067	-0.006
	(0.023)	(0.010)	(0.017)	(0.013)	(0.005)
PhD connection	0.218	0.179	-0.006	-0.485	-0.010
	(0.029)	(0.018)	(0.036)	(0.039)	(0.007)
Fixed-effects					
PhD Class	Yes	Yes	Yes	Yes	Yes
Hiring University Id×Field × 5 year window	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Pseudo R^2	0.63	0.40	0.44	0.08	0.02
Observations	81,614	82,279	66,551	74,276	80,151

Notes: We repeat the analysis presented in Table 3 without the pre-graduation productivity controls.

B.3.2 Time trends in post-PhD outcomes of connected vs. non-connected graduates

In table 14, we report the results from estimating equation (5). We test whether the differences between connected and non-connected graduates has changed over time.

We report the results that include fixed effects for both the destination and the class. For all the outcomes, we find no significant changes in the gap between graduates placed through the network and graduates not placed through the network.

Table 14: Time trends in post-PhD outcomes of connected and non-connected graduates

Dependent Variables:	N Cites PhD graduate	N papers	Co-authors First Affil	Same Affil PhD+6yrs	Any output PhD+6yrs
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Advisor connection $\times(t-1990)$	0.003	0.002	0	0.0008	0.0003
	(0.003)	(0.002)	(0.003)	(0.002)	(0.0008)
PhD connection $\times(t-1990)$	0.005	-0.002	0.008	0.009	0.004
	(0.005)	(0.002)	(0.005)	(0.005)	(0.001)
Fixed-effects					
PhD Class	Yes	Yes	Yes	Yes	Yes
Hiring University Id×Field×5 year window	Yes	Yes	Yes	Yes	Yes
Pre Graduation Productivity×Field	Yes	Yes	Yes	Yes	Yes
Advisor Citation Decile×Field	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes
$(t-1990)$ ×Field	Yes	Yes	Yes	Yes	Yes
Advisor connection×Field	Yes	Yes	Yes	Yes	Yes
PhD connection×Field	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Pseudo R^2	0.68	0.43	0.46	0.09	0.02
Observations	81,614	82,279	66,551	74,276	80,151

Notes: Results from estimating equation (5). Values smaller than $1e-5$ reported as 0.

B.3.3 Topical Fit and post-PhD productivity

We repeat the analysis of the results in the main text regarding post-PhD productivity, but we now also control for two measures of topical fit between the PhD graduate and the first affiliation.

The results shown in Table 15 are very similar to the results in the main text. We again show the results from the main text in columns 1-3 and repeat the same analysis controlling for topical fit in columns 4-6. The results show that while topical fit predicts output, it only explains a small part of the differences in output between connected and not-connected hires. The point estimates for the gap in output between connected and not-connected hires are slightly smaller. Still, the difference in coefficients is small economically, approximately 1-2 log points, compared to baseline differences of 9-26 log points. Further, standard errors are of similar magnitude as the differences across specifications, about 2 log points; thus, the coefficients are not significantly different. Nevertheless, topical fit predicts post-PhD output; the reported pseudo R^2 increases by about 0.03-0.05 depending on the specification.

Table 15: Post-PhD outcomes of connected vs. not-connected hires and topic similarity

Panel A: Comparison of post-PhD outcomes with Class Fixed Effect						
Dependent Variable:	N Cites PhD graduate					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Advisor connection	0.238	0.172	0.121	0.217	0.157	0.106
	(0.021)	(0.019)	(0.019)	(0.021)	(0.019)	(0.019)
PhD connection	0.232	0.027	0.045	0.212	0.010	0.029
	(0.030)	(0.030)	(0.029)	(0.031)	(0.030)	(0.029)
Fixed-effects						
PhD Class	Yes	Yes	Yes	Yes	Yes	Yes
Subfield	Yes	Yes	Yes	Yes	Yes	Yes
Pre Graduation Productivity×Field		Yes	Yes		Yes	Yes
Advisor Citation Decile×Field			Yes			Yes
Varying Slopes						
Max similarity to faculty members×Field				Yes	Yes	Yes
Avg similarity to faculty members×Field				Yes	Yes	Yes
Fit statistics						
Pseudo R^2	0.43	0.51	0.52	0.44	0.51	0.52
Observations	82,141	82,141	82,141	82,141	82,141	82,141

Panel B: Comparison of post-PhD outcomes with Destination Fixed Effect

Dependent Variable:	N Cites PhD graduate					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Advisor connection	0.117	0.059	-0.013	0.118	0.061	-0.011
	(0.022)	(0.020)	(0.021)	(0.021)	(0.020)	(0.020)
PhD connection	0.208	0.014	0.039	0.203	0.009	0.035
	(0.026)	(0.025)	(0.025)	(0.026)	(0.025)	(0.025)
Fixed-effects						
Hiring University Id×Field × 5 year window	Yes	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes	Yes
Pre Graduation Productivity×Field		Yes	Yes		Yes	Yes
Advisor Citation Decile×Field			Yes			Yes
Varying Slopes						
Max similarity to faculty members×Field				Yes	Yes	Yes
Avg similarity to faculty members×Field				Yes	Yes	Yes
Fit statistics						
Pseudo R^2	0.51	0.57	0.58	0.51	0.57	0.58
Observations	81,660	81,660	81,660	81,660	81,660	81,660

Notes: We show again columns (1) to (3) of Table 3 and in columns (4) to (6) we add controls for topic similarity. We omit Panel C for brevity, which shows very similar results.

B.3.4 Heterogeneity by Field

The main results reported in Table 3 reflect the average gap between connected and not-connected hires in outcomes across all fields. Here we present the main results, columns (1) and (3) of Panel B in Table 3 across fields. Figure 9 shows that the main results are roughly similar across fields, that is network hires tend to be positively selected within hiring universities. And observable information eliminates most of the productivity premium. Social sciences tend to have a positive productivity premium of network hires even after controls, while for physics, engineering and math the point estimates turn negative.

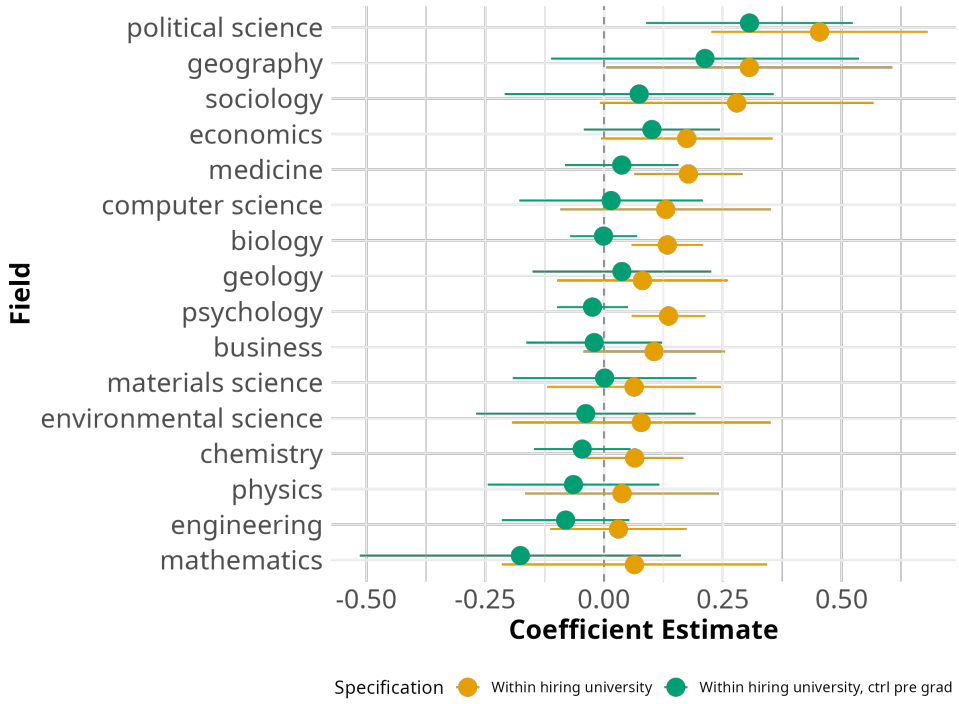


Figure 9: Post-PhD Citations: Advisor Connection coefficient by Field

B.3.5 Robustness Paper Counts

The main analysis of post-PhD outcomes of PhD graduates uses documents independent of (i) their language and (ii) whether there are multiple versions of the same document recorded in MAG.

First, we use the `fasttext` package in Python (Joulin et al., 2016) to predict the language of a document based on its title and repeat the analysis only for English-language documents.

Second, we restrict the documents to those that are the main documents of a Paper Family. A Family is defined by MAG, and FamilyIds are assigned to documents of which multiple versions exist, e.g., a pre-print and a published version (Sinha et al., 2015). We keep only documents that are the sole or the main document of their Family.

We report the results in Table 16. The results are almost unchanged when keeping only English language papers (columns 2 and 5) and keeping only the main document of a Paper Family (columns 3 and 6). Thus, our results are robust against concerns (i) regarding translations of documents and (ii) double-counting pre-prints and published

papers.

Table 16: Post-PhD outcomes robustness Paper Types

Panel A: Comparison of post-PhD outcomes with Class Fixed Effect

Dependent Variables:	N Cites	N Cites only English	N Cites only Main Paper	N Papers	N Papers only English	N Papers only Main Paper
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Advisor connection	0.121	0.121	0.118	0.061	0.060	0.060
	(0.019)	(0.019)	(0.019)	(0.010)	(0.010)	(0.010)
PhD connection	0.045	0.045	0.044	0.053	0.053	0.053
	(0.029)	(0.029)	(0.030)	(0.015)	(0.015)	(0.015)
Fixed-effects						
PhD Class	Yes	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Graduation Productivity×Field	Yes	Yes	Yes	Yes	Yes	Yes
Advisor Citation Decile×Field	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Pseudo R^2	0.52	0.51	0.51	0.30	0.30	0.30
Observations	82,141	82,131	82,141	82,279	82,274	82,279

Panel B: Comparison of post-PhD outcomes with Destination Fixed Effect

Dependent Variables:	N Cites	N Cites only English	N Cites only Main Paper	N Papers	N Papers only English	N Papers only Main Paper
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Advisor connection	-0.013	-0.014	-0.016	0.018	0.017	0.018
	(0.021)	(0.021)	(0.021)	(0.009)	(0.009)	(0.009)
PhD connection	0.039	0.040	0.038	0.048	0.048	0.047
	(0.025)	(0.025)	(0.026)	(0.013)	(0.013)	(0.013)
Fixed-effects						
Hiring University Id×Field × 5 year window	Yes	Yes	Yes	Yes	Yes	Yes
Subfield (MAG lvl 1)	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Graduation Productivity×Field	Yes	Yes	Yes	Yes	Yes	Yes
Advisor Citation Decile×Field	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Pseudo R^2	0.58	0.58	0.58	0.37	0.37	0.37
Observations	81,660	81,608	81,657	82,279	82,250	82,279

Notes: Robustness check of results in Table 3, Panel A and B, restricting the analysis to only English documents and the main/sole document in a MAG Family of papers.

B.3.6 Robustness network controls

In our main analysis, we consider co-author connections of collaborations up to 20 years before the graduation year of the PhD student. This implies that more senior advisors have had more time to form co-author connections. Thus, it would be possible that more senior advisors are better at mentoring PhD students, implying their students perform better. However, they also have larger networks, thus implying that their students are more likely to end up at a connected affiliation. Hence, network hires outperform because they overrepresent graduates with more senior, better mentors. This performance premium would be there even if they were placed outside their advisor's network. To account for this, we repeated the analysis of the post-PhD outcomes and controlled for the network size. The results are unchanged; thus, attributes that co-vary with the size of the network should not be a threat to our main analysis. The results are shown in Table 17.

Table 17: Post-PhD outcomes robustness network controls

Dependent Variables:	N Cites PhD graduate			N papers	Co-authors First Affil	Same Affil PhD+6yrs	Any output PhD+6yrs
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
Advisor connection	0.117	0.055	-0.015	0.010	-0.0001	-0.069	-0.009
	(0.022)	(0.020)	(0.019)	(0.009)	(0.014)	(0.012)	(0.005)
PhD connection	0.208	-0.010	0.022	0.018	-0.093	-0.475	-0.033
	(0.026)	(0.025)	(0.025)	(0.013)	(0.027)	(0.036)	(0.007)
Fixed-effects							
Hiring University Id×Field × 5 year window	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Subfield	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Graduation Productivity×Field		Yes	Yes	Yes	Yes	Yes	Yes
PhD Connections Decile×Field		Yes	Yes	Yes	Yes	Yes	Yes
Advisor Citation Decile×Field			Yes	Yes	Yes	Yes	Yes
Advisor Connections Decile×Field			Yes	Yes	Yes	Yes	Yes
Fit statistics							
Observations	81,660	81,660	81,660	82,279	67,755	75,833	80,463
Pseudo R^2	0.51	0.57	0.59	0.36	0.38	0.06	0.01

Notes: Robustness check of results in Table 3. We add a control for the size of the network to account for the possibility that PhD graduates with more connected advisors outperform independently of whether they end up being a connected hire.

C Regression Framework - Matching

C.1 Multinomial Logit and Poisson

We show that the multinomial logit model in equation (2) is equivalent to a Poisson regression in the sense that it delivers the same parameter estimates. To do that, we show that the likelihood of the quasi-maximum likelihood Poisson regression is the same as that of the multinomial logit if one includes a fixed effect for each graduate in the Poisson regression. Notation and derivation of the likelihood of the multinomial logit and Poisson models follow Wooldridge (2010). The well-known equivalence of the two likelihoods, once one includes a fixed effect for each choice set, follows Guimaraes, Figueirdo and Woodward (2003).

Consider an additive random utility model, that gives rise to the multinomial logit setup, where a (potential) match between graduate i and university j has value

$$v_{ij} + \varepsilon_{ij} = \alpha_{c(i),j} + \beta X_{ij} + \gamma A_{ij} + \frac{1}{\kappa}\varepsilon_{ij}, \quad \varepsilon_{ij} \sim \text{Gumbel}.$$

Then the choice probability of graduate i choosing j out of potential universities $j' = 1, \ldots, J$ follows

$$P(\text{i match j}) = \frac{e^{v_{ij}}}{\sum_{j'=1}^{J} e^{v_{ij'}}}.$$

The multinomial logit likelihood for a set of graduates $i = 1, \ldots, N$ and universities $j = 1, \ldots, J$ is given by:

$$\mathcal{L}_{ML} = \prod_{i=1}^{N} \prod_{j=1}^{J} P(\text{i match j})^{d_{ij}},$$

where d_{ij} is a binary indicator that takes the value 1 if graduate i matches with university j and 0 otherwise. Taking the log of this likelihood function, we obtain the log-likelihood:

$$\log \mathcal{L}_{ML} = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} \log P(\text{i match j}),$$

Substituting the expression for P(i match j), the log-likelihood becomes:

$$\log \mathcal{L}_{ML} = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} \left(v_{ij} - \log \sum_{j'=1}^{J} e^{v_{ij'}} \right).$$

Now, consider the Poisson likelihood with a fixed effect at the individual level i. The expected number of matches for a pair i, j follows

$$\mu_{ij} = e^{\phi_i + v_{ij}},$$

and thus the log likelihood for observation i,j follows

$$l_{ij} = d_{ij} \log(\mu_{ij}) - \mu_{ij},$$

the log likelihood for individual i then follows

$$l_i = \sum_{j=1}^{J} d_{ij} \log(\mu_{ij}) - \mu_{ij},$$

and the overall log likelihood for the sample

$$\log \mathcal{L}_P = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} \log(\mu_{ij}) - \mu_{ij}$$

Let's plug in the expected number of matches

$$\log \mathcal{L}_P = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} \left(\phi_i + v_{ij} \right) - e^{\phi_i + v_{ij}}.$$

Regroup terms to get

$$\log \mathcal{L}_P = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{ij} v_{ij} + d_{ij} \phi_i - e^{\phi_i + v_{ij}}.$$

Now concentrate out the individual fixed effect as Guimaraes, Figueirdo and Woodward (2003). The first order condition with respect to ϕ_i is

$$\partial_{\phi_i} \log \mathcal{L}_P = \sum_{j=1}^{J} d_{ij} - \sum_j e^{\phi_i + v_{ij}} = 0,$$

so ϕ_i solves

$$e^{\phi_i} = \frac{1}{\sum_{j=1}^{J} e^{v_{ij}}},$$

which we plug back into the Poisson likelihood:

$$\log \mathcal{L}_P = \sum_{i=1}^{N} \sum_{j=1}^{J} \left[d_{ij} v_{ij} + d_{ij} \log \left(\frac{1}{\sum_{j'=1}^{J} e^{v_{ij}}} \right) - \frac{e^{v_{ij}}}{\sum_{j'=1}^{J} e^{v_{ij}}} \right].$$

Using that $\sum_{j=1}^{J} \frac{e^{v_{ij}}}{\sum_{j'=1}^{J} e^{v_{ij}}} = 1$, that is each graduate matches with one university, we obtain

$$\log \mathcal{L}_P = \sum_{i=1}^{N} \sum_{j=1}^{J} \left[d_{ij} \left(v_{ij} - \log \sum_{j'=1}^{J} e^{v_{ij}} \right) \right] - NJ.$$

Thus, $\log \mathcal{L}_P + const = \log \mathcal{L}_{ML}$, i.e. the likelihoods are equal up to a constant. It follows directly that the estimates from either approach are identical. □

C.2 Matching models and Poisson Regression

The connection between matching models based on Choo and Siow (2006) and Poisson regression is nicely summarized by Theorem 4 in Galichon and Salanié (2024). They show that a Poisson regression can be used to estimate matching models based on Choo and Siow (2006), as the Poisson regression is equivalent to popular moment matching estimators.

In the following, we illustrate that the equilibrium allocation of Dupuy and Galichon (2014), a version of Choo and Siow (2006) with continuous attributes, is consistent with a Poisson regression. Note that we do not aim to estimate a matching model's full set of parameters but to document differences in matching patterns between connected and unconnected pairs while employing an approach consistent with systematic matching.

Consider the problem of matching PhD graduates i, whose characteristics can be summarized by a vector x_i, and universities j, whose characteristics can be summarized

by a vector y_j. Define v_{ij} as the systematic surplus of a match between graduate i and university j. Further, assume that surplus does not directly depend on the identities, that is $v_{i,j} = v(x_i, y_j)$.

In combination with standard assumptions on idiosyncratic tastes, following a Gumbel distribution, one obtains logit demands for either side of the market and an equilibrium density of matches that is log-linear in systematic surplus. Theorem 1 in Dupuy and Galichon (2014) summarizes the equilibrium allocation, and therein, equation 9 describes the equilibrium density $\pi(x, y)$:

$$\pi(x, y) = \exp\left(\frac{v(x, y) - a(x) - b(y)}{\sigma}\right). \tag{7}$$

This result is important, as it implies that i) a Poisson regression/multinomial logit estimation is consistent with the functional form of equilibrium matching and ii) including a graduate's fixed effect accounts for $a(x)$ and including a fixed effect for each choice option accounts for $b(y)$.

To clarify those points, consider (7) and parametrize $v(x_i, y_j) = \tilde{\beta}_{c(i),j} + \delta X_{ij} + \gamma A_{ij}$. Then the probability of matching follows

$$P(d_{ij} = 1) = \pi(x_i, y_j) = \exp\left(\frac{\tilde{\beta}c(i), j + \beta X_{ij} + \gamma A_{ij} - a(x_i) - b(y_j)}{\sigma}\right) \tag{8}$$

$$\alpha_i = -a(x_i) \tag{9}$$

$$\beta_{c(i),j} = \tilde{\beta}c(i), j - b(y_j) \tag{10}$$

$$P(d_{ij} = 1) = \exp\left(\frac{\alpha_i + \beta_{c(i),j} + \beta X_{ij} + \gamma A_{ij}}{\sigma}\right), \tag{11}$$

which implies our empirical strategy estimates $\frac{\gamma}{\sigma}$. This is the correct estimand, as $\frac{\gamma}{\sigma}$ tells us how much the probability to match changes with a connection $A_{i,j}$ conditional on attributes and equilibrium values.

C.3 Counterfactual Matching probability - Multinomial logit

Based on the estimated multinomial logit model, we quantify how much the predicted probability of matching any of the connected universities would change if the connections were absent.

The predicted probability of graduate i and university j matching is

$$\hat{P}(d_{i,j}=1) = \frac{e^{\hat{\beta}_{c(i),j}+\hat{\gamma}A_{i,j}+\hat{\delta}X_{i,j}}}{\sum_{j'} e^{\hat{\beta}_{c(i),j'}+\hat{\gamma}A_{i,j'}+\hat{\delta}X_{i,j'}}}, \tag{12}$$

where $A_{i,j}$ is an indicator variable that takes value one if the graduate i is connected to university j and zero otherwise.

The connected set of universities is $J_{i,A=1} = \{j : A_{i,j} = 1\}$ and the unconnected set $J_{i,A=0}$ is defined analogously. We rewrite the denominator by splitting the sum into the connected and unconnected set

$$\hat{P}(d_{i,j}=1) = \frac{e^{\hat{\beta}_{c(i),j}+\hat{\gamma}A_{i,j}+\hat{\delta}X_{i,j}}}{\sum_{j'\in J_{i,A=1}} e^{\hat{\gamma}} e^{\hat{\beta}_{c(i),j'}+\hat{\delta}X_{i,j'}} + \sum_{j'\in J_{i,A=0}} e^{\hat{\beta}_{c(i),j'}+\hat{\delta}X_{i,j'}}}. \tag{13}$$

We consider a connected university and calculate how much the predicted probability would change if one removes all connections.

$$\hat{P}_{i,j}^{A=0} = \frac{e^{\hat{\beta}_{c(i),j}+\hat{\delta}X_{i,j}}}{\sum_{j'\in J_{i,A=1}} e^{\hat{\beta}_{c(i),j'}+\hat{\delta}X_{i,j'}} + \sum_{j'\in J_{i,A=0}} e^{\hat{\beta}_{c(i),j'}+\hat{\delta}X_{i,j'}}} \tag{14}$$

The predicted probability to match with any university in the connected set $J_{A=1}$ is

$$\hat{P}_{i,j\in J_{A=1}} = \sum_{j\in J_{A=1}} \frac{e^{\hat{\beta}_{c(i),j}+\hat{\gamma}A_{i,j}+\hat{\delta}X_{i,j}}}{\sum_{j'} e^{\hat{\beta}_{c(i),j'}+\hat{\gamma}A_{i,j'}+\hat{\delta}X_{i,j'}}}. \tag{15}$$

Next we split the sum in the denominator into the connected and unconnected set as well and take $\hat{\gamma}A_{i,j}$ outside the sum, as A_i, j is constant within the connected, unconnected set, respectively. Further, the denominator is constant across j. Thus, we can take the

sum of the numerator and divide by the denominator afterwards

$$\hat{P}_{i,j\in J_{A=1}} = e^{\hat{\gamma}} \frac{\sum_{j\in J_{A=1}} e^{\hat{\beta}_{c(i),j}+\hat{\delta}X_{i,j}}}{e^{\hat{\gamma}} \sum_{j'\in J_{i,A=1}} e^{\hat{\beta}_{c(i),j'}+\hat{\delta}X_{i,j'}} + \sum_{j'\in J_{i,A=0}} e^{\hat{\beta}_{c(i),j'}+\hat{\delta}X_{i,j'}}} \tag{16}$$

$$\hat{p}_{1,i} = \sum_{j\in J_{A=1}} e^{\hat{\beta}_{c(i),j}+\hat{\delta}X_{i,j}} \tag{17}$$

$$\hat{p}_{0,i} = \sum_{j\in J_{i,A=0}} e^{\hat{\beta}_{c(i),j}+\hat{\delta}X_{i,j}} \tag{18}$$

$$\hat{P}_{i,j\in J_{A=1}} = \frac{e^{\hat{\gamma}}\hat{p}_{1,i}}{e^{\hat{\gamma}}\hat{p}_{1,i} + \hat{p}_{0,i}}. \tag{19}$$

Now we calculate the predicted probability to match any affiliation in the connected set when all connections would be absent:

$$P^{A=0}_{i,j\in J_{A=1}} = \sum_{j\in J_{A=1}} \frac{e^{\hat{\beta}_{c(i),j}+\hat{\delta}X_{i,j}}}{\sum_{j'\in J_{i,A=1}} e^{\hat{\beta}_{c(i),j'}+\hat{\delta}X_{i,j'}} + \sum_{j'\in J_{i,A=0}} e^{\hat{\beta}_{c(i),j'}+\hat{\delta}X_{i,j'}}} \tag{20}$$

$$\hat{P}^{A=0}_{i,j\in J_{A=1}} = \frac{\hat{p}_{1,i}}{\hat{p}_{1,i} + \hat{p}_{0,i}} \tag{21}$$

Now we express $\hat{P}^{A=0}_{i,j\in J_{A=1}}$ as a function of predicted probabilities

$$e^{-\hat{\gamma}}\hat{P}_{i,j\in J_{A=1}} = \frac{\hat{p}_{1,i}}{e^{\hat{\gamma}}\hat{p}_{1,i} + \hat{p}_{0,i}} \tag{22}$$

$$\hat{P}_{i,j\in J_{A=0}} = \frac{\hat{p}_{0,i}}{e^{\hat{\gamma}}\hat{p}_{1,i} + \hat{p}_{0,i}} \tag{23}$$

$$\frac{e^{-\hat{\gamma}}\hat{P}_{i,j\in J_{A=1}}}{e^{-\hat{\gamma}}\hat{P}_{i,j\in J_{A=1}} + \hat{P}_{i,j\in J_{A=0}}} = \frac{\hat{p}_{1,i}}{\hat{p}_{1,i} + \hat{p}_{0,i}} \tag{24}$$

$$\hat{P}^{A=0}_{i,j\in J_{A=1}} = \frac{e^{-\hat{\gamma}}\hat{P}_{i,j\in J_{A=1}}}{e^{-\hat{\gamma}}\hat{P}_{i,j\in J_{A=1}} + \hat{P}_{i,j\in J_{A=0}}} \tag{25}$$

To approximate the role of co-author connections for placement in the connected set we use the last line and evaluate it at the sample shares of hires in the connected set combined with the estimates for γ.

References

American Council on Education. 2021. "Carnegie Classification of Institutions of Higher Education. `https://carnegieclassifications.acenet.edu/wp-content/uploads/2023/02/CCIHE2021-PublicData.xlsx` (accessed June 2022)."

Azoulay, Pierre, Christopher C Liu, and Toby E Stuart. 2017. "Social influence given (partially) deliberate matching: Career imprints in the creation of academic entrepreneurs." *American Journal of Sociology*, 122(4): 1223–1271.

Bailey, Martha J, Connor Cole, Morgan Henderson, and Catherine Massey. 2020. "How well do automated linking methods perform? Lessons from US historical data." *Journal of Economic Literature*, 58(4): 997–1044.

Baruffaldi, Stefano, Fabiana Visentin, and Annamaria Conti. 2016. "The productivity of science & engineering PhD students hired from supervisors' networks." *Research Policy*, 45(4): 785–796.

Beaman, Lori, and Jeremy Magruder. 2012. "Who gets the job referral? Evidence from a social networks experiment." *American Economic Review*, 102(7): 3574–93.

Bergé, Laurent. 2018. "Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm." *CREA Discussion Papers*, , (13).

Brown, David G. 1965. *Academic labor markets.*

Brown, Meta, Elizabeth Setren, and Giorgio Topa. 2016. "Do informal referrals lead to better matches? Evidence from a firm's employee referral system." *Journal of Labor Economics*, 34(1): 161–209.

Burks, Stephen V, Bo Cowgill, Mitchell Hoffman, and Michael Housman. 2015. "The value of hiring through employee referrals." *The Quarterly Journal of Economics*, 130(2): 805–839.

Choo, Eugene, and Aloysius Siow. 2006. "Who marries whom and why." *Journal of political Economy*, 114(1): 175–201.

Clauset, Aaron, Samuel Arbesman, and Daniel B Larremore. 2015. "Systematic inequality and hierarchy in faculty hiring networks." *Science advances*, 1(1): e1400005.

Cornell, Bradford, and Ivo Welch. 1996. "Culture, information, and screening discrimination." *Journal of political Economy*, 104(3): 542–571.

Dupuy, Arnaud, and Alfred Galichon. 2014. "Personality traits and the marriage market." *Journal of Political Economy*, 122(6): 1271–1319.

Dustmann, Christian, Albrecht Glitz, Uta Schönberg, and Herbert Brücker. 2016. "Referral-based job search networks." *The Review of Economic Studies*, 83(2): 514–546.

Eliason, Marcus, Lena Hensvik, Francis Kramarz, and Oskar Nordström Skans. 2022. "Social connections and the sorting of workers to firms." *Journal of Econometrics*.

Fiegener, Mark K. 2010. "Numbers of Doctorates Awarded Continue to Grow in 2009; Indicators of Employment Outcomes Mixed." National Science Foundation InfoBrief NSF 11-305. Accessed: 2024-11-13.

Fortunato, Santo, Carl T Bergstrom, Katy Börner, James A Evans, Dirk Helbing, Staša Milojević, Alexander M Petersen, Filippo Radicchi, Roberta Sinatra, Brian Uzzi, et al. 2018. "Science of science." *Science*, 359(6379): eaao0185.

Freeman, Richard B, Ina Ganguli, and Raviv Murciano-Goroff. 2014. "Why and wherefore of increased scientific collaboration." In *The changing frontier: Rethinking science and innovation policy*. 17–48. University of Chicago Press.

Galenianos, Manolis. 2013. "Learning about match quality and the use of referrals." *Review of Economic Dynamics*, 16(4): 668–690.

Galichon, Alfred, and Bernard Salanié. 2024. "Estimating separable matching models." *Journal of Applied Econometrics*, 39(6): 1021–1044.

Gaulé, Patrick, and Mario Piacentini. 2013. "Chinese graduate students and US scientific productivity." *Review of Economics and Statistics*, 95(2): 698–701.

Gregg, Forest, and Derek Eder. 2022. "Dedupe, version 2.0.11. URL: https://github.com/dedupeio/dedupe."

Guimaraes, Paulo, Octávio Figueirdo, and Douglas Woodward. 2003. "A tractable approach to the firm location decision problem." *Review of Economics and Statistics*, 85(1): 201–204.

Hadlock, Charles J, and Joshua R Pierce. 2021. "Hiring your friends: Evidence from the market for financial economists." *ILR Review*, 74(4): 977–1007.

Harzing, Anne-Wil, and Satu Alakangas. 2017. "Microsoft Academic: Is the phoenix getting wings?" *Scientometrics*, 110(1): 371–383.

Heath, Rachel. 2018. "Why do firms hire using referrals? Evidence from Bangladeshi garment factories." *Journal of Political Economy*, 126(4): 1691–1746.

Hensvik, Lena, and Oskar Nordström Skans. 2016. "Social networks, employee selection, and labor market outcomes." *Journal of Labor Economics*, 34(4): 825–867.

Huang, Junming, Alexander J Gates, Roberta Sinatra, and Albert-László Barabási. 2020. "Historical comparison of gender inequality in scientific careers across countries and disciplines." *Proceedings of the National Academy of Sciences*, 117(9): 4609–4616.

Hug, Sven E., and Martin P. Brändle. 2017. "The coverage of Microsoft Academic: Analyzing the publication output of a university." *Scientometrics*, 113(3): 1551–1571.

Hug, Sven E., Michael Ochsner, and Martin P. Brändle. 2017. "Citation analysis with Microsoft Academic." *Scientometrics*, 111(1): 371–378.

Jones, Benjamin F. 2009. "The burden of knowledge and the "death of the renaissance man": Is innovation getting harder?" *The Review of Economic Studies*, 76(1): 283–317.

Joulin, Armand, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2016. "Bag of Tricks for Efficient Text Classification." *arXiv preprint arXiv:1607.01759*.

Kramarz, Francis, and David Thesmar. 2013. "Social networks in the boardroom." *Journal of the European Economic Association*, 11(4): 780–807.

Malmgren, R Dean, Julio M Ottino, and Luís A Nunes Amaral. 2010. "The role of mentorship in protégé performance." *Nature*, 465(7298): 622–626.

Martín-Martín, Alberto, Mike Thelwall, Enrique Orduna-Malea, and Emilio Delgado López-Cózar. 2020. "Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations." *arXiv preprint arXiv:2004.14329*.

Ma, Yifang, Satyam Mukherjee, and Brian Uzzi. 2020. "Mentorship and protégé success in STEM fields." *Proceedings of the National Academy of Sciences*, 117(25): 14077–14083.

Microsoft. 2021. "Microsoft Academic Graph [Data set]. Version from 2021-05-24." Microsoft Academic [Distributor].

Missouri Census Data Center. 2020. "ZIP Code Lookup, complete list in excel file. `https://mcdc.missouri.edu/applications/zipcodes/ZIP_codes_2020.xls` (accessed June 2022)."

Montgomery, James D. 1991. "Social networks and labor-market outcomes: Toward an economic analysis." *The American economic review*, 81(5): 1408–1418.

Organisation for Economic Co-operation and Development. 2007. "Revised Field of Science and Technology (FOS) Classification in the Frascati Manual. `https://www.oecd.org/science/inno/38235147.pdf` (accessed June 2022)."

Pallais, Amanda, and Emily Glassberg Sands. 2016. "Why the referential treatment? Evidence from field experiments on referrals." *Journal of Political Economy*, 124(6): 1793–1828.

Proquest. 2023. "Proquest Dissertation & Theses Global. https://about.proquest.com/en/products-services/pqdtglobal/ (accessed March 2023)."

Rajkumar, Karthik, Guillaume Saint-Jacques, Iavor Bojinov, Erik Brynjolfsson, and Sinan Aral. 2022. "A causal test of the strength of weak ties." *Science*, 377(6612): 1304–1310.

Rees, Albert. 1966. "Information networks in labor markets." *The American Economic Review*, 56(1/2): 559–566.

Rose, Michael E, and Suraj Shekhar. 2023. "Adviser Connectedness and Placement Outcomes in the Economics Job Market." *Labour Economics*, 102397.

San, Shmuel. 2022. "Who Works Where and Why: The Role of Social Connections in the Labor Market."

Simon, Curtis J, and John T Warner. 1992. "Matchmaker, matchmaker: The effect of old boy networks on job match quality, earnings, and tenure." *Journal of labor economics*, 10(3): 306–330.

Sinha, Arnab, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, Bo-June Hsu, and Kuansan Wang. 2015. "An overview of microsoft academic service (mas) and applications." 243–246.

Smeets, Valérie, Frédéric Warzynski, and Tom Coupé. 2006. "Does the academic labor market initially allocate new graduates efficiently?" *Journal of Economic Perspectives*, 20(3): 161–172.

Terviö, Marko. 2011. "Divisions within academia: Evidence from faculty hiring and placement." *Review of Economics and Statistics*, 93(3): 1053–1062.

Topa, Giorgio. 2011. "Chapter 22 - Labor Markets and Referrals." In . Vol. 1 of *Handbook of Social Economics*, , ed. Jess Benhabib, Alberto Bisin and Matthew O. Jackson, 1193–1221. North-Holland.

Visser, Martijn, Nees Jan van Eck, and Ludo Waltman. 2021. "Large-scale comparison of bibliographic data sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic." *Quantitative Science Studies*, 2(1): 20–41.

Wang, Kuansan, Zhihong Shen, Chiyuan Huang, Chieh-Han Wu, Darrin Eide, Yuxiao Dong, Junjie Qian, Anshul Kanakia, Alvin Chen, and Richard Rogahn. 2019. "A review of microsoft academic services for science of science studies." *Frontiers in Big Data*, 2: 45.

Wang, Kuansan, Zhihong Shen, Chiyuan Huang, Chieh-Han Wu, Yuxiao Dong, and Anshul Kanakia. 2020. "Microsoft academic graph: When experts are not enough." *Quantitative Science Studies*, 1(1): 396–413.

Way, Samuel F, Allison C Morgan, Daniel B Larremore, and Aaron Clauset. 2019. "Productivity, prominence, and the effects of academic environment." *Proceedings of the National Academy of Sciences*, 116(22): 10729–10733.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data.* MIT press.

Wuchty, Stefan, Benjamin F Jones, and Brian Uzzi. 2007. "The increasing dominance of teams in production of knowledge." *Science*, 316(5827): 1036–1039.