

Edge interventions can mitigate demographic and prestige disparities in the Computer Science coauthorship network

Kate Barnes

kathryn.barnes@colorado.edu
University of Colorado Boulder
Boulder, Colorado, USA

Nayera Hasan

nhasan1@haverford.edu
Haverford College
Haverford, Pennsylvania, USA

Sorelle Friedler

sorelle@cs.haverford.edu
Haverford College
Haverford, Pennsylvania, USA

Mia Ellis-Einhorn

melliseinh@gmail.com
Haverford College
Haverford, Pennsylvania, USA

Mohammad Fanous

mfanous@haverford.edu
Haverford College
Haverford, Pennsylvania, USA

Carolina Chávez-Ruelas

carolina.chavezruelas@colorado.edu
University of Colorado Boulder
Boulder, Colorado, USA

Blair D. Sullivan

sullivan@cs.utah.edu
The University of Utah
Salt Lake City, Utah, USA

Abstract

Social factors such as demographic traits and institutional prestige structure the creation and dissemination of ideas in academic publishing. One place these effects can be observed is in how central or peripheral a researcher is in the coauthorship network. Here we investigate inequities in network centrality in a hand-collected data set of 5,670 U.S.-based faculty employed in Ph.D.-granting Computer Science departments and their DBLP coauthorship connections. We introduce algorithms for combining name- and perception-based demographic labels by maximizing alignment with self-reported demographics from a survey of faculty from our census. We find that women and individuals with minoritized race identities are less central in the computer science coauthorship network, implying worse access to and ability to spread information. Centrality is also highly correlated with prestige, such that faculty in top-ranked departments are at the core and those in low-ranked departments are in the peripheries of the computer science coauthorship network. We show that these disparities can be mitigated using simulated edge interventions, interpreted as facilitated collaborations. Our intervention increases the centrality of target individuals, chosen independently of the network structure, by linking them with researchers from highly ranked institutions. When applied to scholars during their Ph.D., the intervention also improves the predicted rank of their placement institution in the academic job market. This work was guided by an ameliorative approach: uncovering social inequities in order to address them. By targeting scholars for intervention based on institutional prestige, we are able to improve their centrality in the coauthorship network that plays a key role in job placement and longer-term academic success.

CCS Concepts

- Human-centered computing → Empirical studies in collaborative and social computing.

Keywords

network fairness, edge interventions, demographic inference, science of science, coauthorship networks

1 Introduction

Prestige plays a central role in structuring a wide variety of outcomes in academic systems. For instance, the most prestigious 20% of universities in the U.S. train roughly 80% of all tenured and tenure-track faculty in every field at U.S.-based Ph.D.-granting institutions [50]. Scientists at elite institutions publish more papers [30, 52], have larger research groups [59], and receive more grant funding and citations to their papers [30]. These prestige hierarchies are also remarkably stable over time [4, 31], and effectively amplify the spread of ideas originating from elite institutions [36]. To adapt to this context a famous quote by Theodosius Dobzhansky [14], it is reasonable to conclude that “*Nothing in academia makes sense except in the light of prestige hierarchies.*”

A strong prestige gradient is not by itself necessarily unfair. Existing analyses suggest that the kind of steep gradients observed in academia reflect a combination of differences in merit and various non-meritocratic factors or processes [13, 28, 37]. Hence, we can ask whether it may be possible to intervene in some natural way to mitigate non-meritocratic disparities, including systemic devaluation of scientists with gender or racial minority identities [12, 15, 43, 44]. Such an intervention could aim to improve the position of individuals outside of elite institutions or improve the job prospects of promising individuals from non-elite institutions. In highly collaborative fields like Computer Science, coauthorship on published papers is a key way that the community structures itself [39, 40]. This makes suggested collaborations a natural way to intervene in the academic system to improve fairness or reduce systemic disparities, e.g., by supporting fellowships for a collaborative project with a research group at another institution.

Contributions. Here, we study how to make edge interventions into a real-world faculty coauthorship network to improve an individual’s

position in the system or their predicted academic job placement. Specifically, we make the following contributions.

Computer science faculty coauthorship network. We conduct a census of 5,670 computer science (CS) faculty in 178 Ph.D.-granting departments at U.S. institutions, collecting for each faculty member, their name, current faculty rank, current institution, doctoral institution, and perceived gender and race based on public faculty websites and photographs (Section 3). We create a novel CS faculty coauthorship network from this census using DBLP data (Section 4), in which nodes are also labeled with inferred demographic values and several measures of institutional prestige. We make this high-quality, largely manually collected network and associated attribute data publicly available [2].

Demographics meta-labeling algorithms. We introduce a name-and-perception-based algorithm to augment this real-world coauthorship network with faculty gender and racial demographic information, to facilitate the study of demographic-based disparities in computer science. These algorithms are created and validated based on collected survey data on self-identification (Section 3.4), and then applied to the CS faculty coauthorship network.

Fairness findings. We analyze the CS faculty coauthorship network and its associated institutional prestige, gender, and racial labels, and we show that :

- (1) centrality within the coauthorship network correlates with institutional prestige, meaning that more central nodes are affiliated with more prestigious institutions (Section 5);
- (2) centrality also correlates with race, such that Ph.D. students and faculty from majority racial groups (White, East Asian, and South Asian) have significantly higher closeness centrality (Section 5.3) than those from minority racial groups; and,
- (3) doctoral pedigree and centrality within the coauthorship network are predictive of institutional prestige in faculty placement (Section 6.2).

Intervention algorithm. We introduce an intervention algorithm that selects—*without knowledge of the network itself*—a single edge to add to the network (i.e., creating a new coauthorship link). We show that when applied to Ph.D. students, this intervention increases both their closeness centrality and improves their predicted placement rank within the academic job market (Section 6.)

2 Related Work

While institutional prestige has enormous influence on many aspects of faculty life, faculty experience can vary dramatically with social identity. For instance, a broad literature indicates that academic life can be strongly gendered in some ways, with negative consequences [12, 43]. Recent work shows that women faculty leave tenured and tenure-track faculty positions at higher rates than men at all career stages, that gendered attrition is higher at less prestigious institutions [44], and that the shorter career lengths of women tend to reduce their overall number of scientific contributions [23]. A key factor in such attrition is gendered devaluation, both formally in summative assessments and informally in department life [45]. On the other hand, work on faculty hiring networks indicates that gender does not significantly correlate with differences in job placement prestige after the Ph.D. in most fields [50].

Racial minority faculty can also experience various forms of devaluation in the academy [15, 43, 45]. For instance, research shows they can be subjected to double standards in promotion evaluations [35] and to implicit biases during faculty hiring [54], and that their scholarship is funded at lower rates by grant-making agencies [22].

Race and gender also influence the structure of coauthorship networks [8, 55]. For instance, women researchers often have fewer distinct collaborators (smaller degree) than men [58], which can reduce their effective productivity and prominence over their career [32]. Additionally, researchers often exhibit gender or racial homophily, coauthoring with same-gender or same-race individuals at higher rates than expected by chance [19, 55]. We examine coauthorship homophily in Section 5.2. Recent work suggests that women in computer science working on artificial intelligence tend to occupy less central positions in collaboration networks [49], but no studies have considered the broader question of how race and/or gender relate to centrality in computer science at large, or whether centrality in coauthorship networks correlates with prestige.

Such network-driven demographic disparities have received increasing algorithmic attention, with recent work often focused on interventions that aim to achieve fairness in social networks (for a survey, see [41]). Definitions of fairness have largely focused on access to information propagated under some network flow model (often the independent cascade model), with the motivation that information and other resources (e.g., access to jobs) are often shared via social networks. Individual fairness definitions have aimed to maximize the minimum access of an individual to some information seeded in the network (e.g., [18]) or other individual notions of control, broadcast, or structural access in a network (e.g., [3]), while group fairness definitions have aimed to equalize such information access across demographic groups represented as node attributes [17, 46]. Intervention approaches have included changing who has access to information and direct intervention on the network structure via edge augmentation [3, 5, 7, 18, 56].

In this paper, we bring together these two lines of work; studying faculty with a focus on demographics and institutional prestige in the context of real-world networks, and algorithmic fairness interventions in a network setting.

3 Collection of Faculty Attributes and Demographic Inference

We introduce a novel, real-world data set describing current computer science faculty and their coauthorship relations. The data were hand-collected in 2023-24 directly from the websites of Ph.D.-granting Computer Science departments at U.S. institutions. These universities were identified based on a list from previous research [13]. The complete list of 178 institutions we use is given in Appendix C. Our coding and data procedures ensure the collection of high-quality attributes, including, in particular, gender and race demographic labels for the 5,670 faculty in our census. Race was categorized in alignment with present U.S. Census Bureau categories [9], with the exception of the U.S. Census Asian category which we expand to East Asian, South Asian (Indian / Indian subcontinent) and Southeast Asian. The terms “race” and “gender”, instead of “ethnicity” or “sex”, are used throughout to reflect the

socially constructed nature of the demographic categories we study. Institutional prestige, as studied here, is another dimension along which social hierarchies are perceived and enacted. We additionally augment the faculty data with their current and Ph.D. institutional ranks from U.S. News & World Report (USNWR)¹, CSRankings² and a measure derived from faculty hiring networks [50].

Information reflecting individuals' self-identified demographics and lived experience is hard to collect at scale. While survey responses are the gold standard [33], this information is not always broadly available, and can be subject to various sampling biases [16]. Researchers lacking self-reported data sometimes resort to perception or name-based inference for labeling demographics, which can diverge from self-identification, sometimes in systematic ways. However, the need for and utility of research on demographic disparities exists regardless of the availability of self-reported data. We introduce algorithmic tools which combine name- and perception-based labels, designed to maximize alignment with self-identification responses from a survey of faculty in our census. While the exact algorithms we apply are particular to the demographic distribution of our census of U.S.-based faculty, the methodology we introduce is itself general, and can be adapted to other studies with survey demographic information for a subset of their population.

This high-quality data set is a major contribution of the present work. Excluding protected survey information, our coauthorship network edge list and node metadata is available on GitHub for future studies of the social structures shaping collaboration in computer science [2].

3.1 Assembling a census of CS faculty

In the context of social analyses using potentially sensitive attributes, high-quality data is of great importance. To ensure the accuracy of our data, we hired undergraduate students during the 2023-24 academic year to conduct a census of tenured and tenure-track computer science faculty at Ph.D.-granting institutions based in the United States. Hand-collected information on each faculty member included their full name, email address, current and Ph.D. institution names, and current position titles—all collected from publicly available sources.

Data coders. Seven undergraduate coders were hired during the 2023-24 academic year to conduct the census. The hourly rate paid was determined by institutional policy at the Haverford College. Hired student coders were also given the opportunity to join the core research team, and one is an author on this paper.

Coding procedure. Coders were instructed to search for institutions by name and navigate to the Computer Science department faculty pages. Faculty members' information was collected only if they were listed as assistant professors, associate professors, professors, or distinguished professors. Individuals with other titles such as visiting professor or senior lecturer were not collected, as we assume these faculty are not members of the tenure track. To avoid spelling errors, coders copied and pasted faculty members'

names, emails, institution names, Ph.D. institution names, and titles directly from the webpage. If any of this information was not included on the department page, coders were instructed to check faculty members' personal websites. If a faculty member was listed at multiple institutions, coders rectified this ambiguity by referring to the scholar's curriculum vitae, keeping only information for the institution in which they were employed in the present academic year, 2023-24. This process refines a similar procedure developed by [13] for constructing a faculty census. The process was completed for all 178 institutions in the list of Ph.D. granting departments from [13], and resulted in a complete census of 5,670 U.S. computer science faculty. The specific coding procedure followed to collect this data was discussed in person with the hired coders as a group to help ensure consistency and clarify any confusions. The detailed procedure can be found in Appendix A.1.

An error analysis was conducted after the completion of the census of 5,670 faculty. We randomly sampled 400 individuals from the census to conduct this analysis. Information for these individuals was re-collected according to the same coding procedure and then compared with the previous collection, noting errors. This analysis revealed 11/400 faculty who should not have been collected because their professional titles (Teaching Professor, Adjunct Professor, Emeritus, etc.) were not in our list for inclusion, an error rate of 2.75%. Additionally, 4 individuals' current or Ph.D. institutions were incorrectly collected, an error rate of 1%.

3.2 Departmental prestige

In order to test the associations between prestige, demographics, and coauthorship in computer science (Section 6.2), we assigned each member of the CS faculty census an institutional prestige score based on their current institutional affiliation. Each of the 178 unique institutions in our data was annotated with departmental rankings specific to computer science from three sources: (i) the US News & World Report (USNWR) Computer Science graduate program ranking, (ii) CSRankings and (iii) a prestige measure derived from faculty hiring networks [50].

CSRankings ranks departments based on a weighted count of their publications in selected CS conferences. USNWR ranks graduate programs primarily based on reputation, as measured by survey. The [50] ranking measures a department's ability to place its graduates as faculty at other institutions, and is thus called a measure of "placement power". This network-based measure of prestige is highly correlated with both CSRankings (Pearson's $r = 0.77$, $p = 8.5 \times 10^{-53}$) and USNWR (Pearson's $r = 0.88$, $p = 1.5 \times 10^{-33}$) but has been shown to more accurately predict faculty placement [13]. Additionally, these correlations with the placement power measure are comparable to the correlation between CSRankings and USNWR (Pearson's $r = 0.88$, $p = 3.5 \times 10^{-53}$).

CSRankings and USNWR ranks were collected from their corresponding websites in November 2024, and are used to corroborate our findings according to the placement power measure of institutional prestige. Hand-collected university names from our data were linked with these three rankings using string matching techniques, resulting in complete coverage. Seven universities from our list were not listed in the placement power ranking; 11 were not found in CSRankings; 2 were not found in USNWR. No universities

¹<https://www.usnews.com/best-graduate-schools/top-science-schools/computer-science-rankings>

²<https://csrankings.org/#/fromyear/2014/toyear/2024/index?all&us>

were covered by none of these rankings. These specifics are given in Section A.7.

We define an *early career scholar* to be a faculty member in our data set whose first publication was 2010 or later. Our census contains 2041 such faculty who completed their Ph.D. degree in the U.S. Using this subsample of our faculty, we will test the importance of Ph.D. prestige in scholars' placement outcomes and the effect of our simulated edge intervention on targeted Ph.D. student scholars (Section 6.3). These individuals span 179 unique doctoral institutions, of which only one (Oregon Graduate Institute) could not be matched to the [50] placement power measure we use. The [50] ranking is calibrated using only data from U.S. Ph.D.-granting institutions, but an additional 223 early career scholars in our census received their Ph.D. degrees from international institutions. These individuals are included in an expanded early career cohort of 2264 individuals which is analyzed using CSRankings and USNWR rankings in supplemental results in Appendix B.1 and B.2.

3.3 Augmenting the census with demographic variables

Mislabeling gender or race attributes can have harmful impacts. In research contexts these include misrepresenting population sizes and underlying patterns of marginalization. To mitigate potentially harmful errors in our demographic labeling of faculty in our census, we consider three sources of demographic information: perception, name-based inference, and survey self-identification. While perception and name-based inference estimates may align with individuals' self-identified gender and race, these are not necessarily the same. Self-reports are the preferred method for obtaining demographic information [33], but often are not available, or are only available for a subset of a study's population. In Section 3.4, we introduce methods to combine perception and name-based inference methods to better match self-reported survey data. We first describe the collection process for race and gender data using perception, name-based inference, and self-reporting.

3.3.1 Perception of race and gender. Perception data about race and gender was collected based on photos as perceived by the hired undergraduate student coders (discussed in Section 3.1). The coding procedure (given in full in Appendix A.1) began by searching for faculty photos on department websites. If the faculty photo was not found there, coders were instructed to find another photo by searching, in order of preference, "firstname lastname CurrentInstitution", "firstname lastname linkedin" or "firstname lastname PhDInstitution". If the photo used was not from the department website, coders ensured it was the correct individual by verifying that their information aligned with what was previously collected (e.g., by looking at their institutions listed on LinkedIn to make sure they matched with the institutions recorded in our data).

Once an image had been found, coders used faculty photos to categorize scholars based on perceived race and gender. Coders recorded perceived gender as "Man", "Woman", "Non-Binary / Uncertain" or "No photo found". Perceived race was recorded in alignment with current U.S. Census Bureau [9] categories as "White",

"Black", "Latinx", "East Asian", "Southeast Asian", "South Asian (Indian / Indian subcontinent)", "Middle Eastern / North African", "Native American / Other Indigenous", "Native Hawaiian or Other Pacific Islander", "Multiracial or unsure" or "No photo found". Coders recorded their own name in a "perceived by" column. Based on self-identification of these seven coders, six are women, one is non-binary, one is White, two are Middle Eastern / North African, two are Hispanic, one is South Asian / Indian, and one is East Asian. While we did not have a large enough sample of student coders to analyze the differential impact of personal gender and racial identification on perception, we note these demographic characteristics since such an impact may exist.

To ensure inter-coder reliability, 400 faculty in our census were double-collected by different individuals. Of these 400, coders double-collected perceived race and gender labels for 283 individuals. Subsequent analysis showed 89% agreement (251/283) between coders on perceived race labels and 100% agreement (282/283) on perceived gender labels. The full details of the perception coding procedure are given in the Appendix in Section A.1.

3.3.2 Name-based inference of race and gender. In addition to perception labels, we employed four name-based gender (GenderGuesser [1], NamesOrFullNames [38], NonQuamGender [11] and WikiGendersort [6]) and five name-based race inference tools (EthnicolrWiki, EthnicolrCensus, EthnicolrFlorida [29], Ethnea [47], and EthnicSeer [48]). These inference tools assign demographic labels based on names alone, a method which is importantly fraught. Researchers using these methods run the risk of reifying the association between names and sociodemographic categories [20]. Names, simply strings of letters, do not have gender or race but rather are imbued with demographic signal through cultural consensus [11]. Thus, the association between a name and a particular race or gender is not fixed, but varies by culture and over time [21]. Some name-based algorithms seek to remedy this by including data from across time periods and allowing users to specify time as a variable. Importantly, however, name-based inference methods allow the automatic association of race and gender information to each person in the study.

The name-based inference tools we used report the uncertainty of the association between names and demographics via a probability value, indicating the strength of the gender or race signal of a particular name. The demographic labels and their associated probabilities are data-driven estimates. For example, NonQuam [11] labels the name "Sally" as gendered female with $p = 0.984$, indicating that on average 98.4% of occurrences of the name "Sally" in sources from NonQuam's database are gendered female. Despite the intrinsic limitations of name-based demographic labeling, high probability labels indicate strong alignment between the given name and an individual's gender [11] and race attributes [25], understood within the cultural consensus framework.

A major limitation of tools for name-based gender inference is that they exclusively return binary gender categories [42]. Therefore, name-based inference labels could not correctly label any non-binary faculty in our census. Additionally, racial categories returned by name-based inference methods did not always align with the U.S. census categories that we employ in our perception- and survey-collected data. Thus, we manually aligned the race labels

from name-based inference with our categories. EthnicolrFlorida was only used to label Latinx, Black, and White individuals due to reporting only one Asian category which did not match ours and returning no other categories. EthnicolrCensus was not used due to reporting only four, broad racial categories and performing poorly in alignment with survey self-reports. Ethnea, EthnicolrWiki and Ethnicseer returned many racial categories. The full details of how these were aligned with U.S. census categories are provided in Appendix Section A.2.

3.3.3 Survey self-identification of race and gender. Our IRB-approved³ survey was sent to all faculty in our census in the summer of 2024 using the emails collected as part of the data coding procedure. The survey requested self-reported demographic information. We received 820 responses, a response rate of 15%. Survey respondents provided their name, confirmed they were currently employed as tenured or tenure-track faculty, and selected a single gender attribute: male, female, non-binary or other. Respondents were also asked to identify their race/ethnicity by choosing one or multiple racial categories from the same list of options provided to coders. For both race and gender questions, respondents were provided an optional free-form text box to provide further information if desired. 55 individuals used this space to report other racial identities, predominantly to identify as Jewish, and 11 individuals self-described their gender. The complete text of the survey questionnaire is provided in Appendix A.3.

3.4 Meta-labeling algorithms for gender and race

Perception and name-based inference have their own limitations and raise particular ethical concerns. Socially-perceived race and gender do not always align with individual's lived experiences, and names are only associated with racial and gender categories through a process of cultural consensus. However, both of these methods capture some dimensions of race and gender as socially constructed categories, and it is these sociological formations we aim to study. In this section, we show significant alignment between demographic labels from perception and name-based inference and those that were self-reported in our survey data. Further, we introduce algorithmic tools which combine name- and perception-based labels, designed to maximize alignment with self-identification responses. The final labels applied by our gender and race meta-labeling algorithm align better with survey self-identification than either perception or any name-based inference tool alone. While the exact algorithms we apply are particular to the demographic distribution of our census, the general methodology can be used in future research.

With survey self-reports to validate against, we developed two meta-labeling algorithms to combine the perception-based demographic estimates made by coders with name-based estimates from automatic inference methods. Overall, perception labels aligned best with survey responses for both gender and race. For gender, 99% of perceived labels matched, compared to an average of 83%

³Approved by the Haverford College IRB dated June 17, 2024 under the title “Demographics and Faculty Co-Authorship Networks.”

for all name-based gender inference methods. For race, 91% of perceived labels matched survey self-identified race labels, compared to an average of 82% for all name-based race methods.

Both our race and gender meta-labelers proceed by first applying name-based inference methods to all names and accepting only high-confidence labels greater than some per-method threshold. Then, names with gender or race signal below our chosen thresholds are classified using the perception labels. Given the high alignment between perception and self-reported data, the idea behind first using name-based inference followed by perception is that future researchers could use this generalized methodology to start by using the cheaper and more scalable name-based inference methods and use the more expensive, hand-collected perception data only when necessary.

A general schematic for the race and gender meta-labelers described above is given in Algorithm 1. To determine the resulting gender and race meta-labelers, we search through a space of many algorithms determined by the specific choices of for loop orderings (Line 5) and thresholds (Line 7). We describe this process and results for race and gender meta-labelers next.

Algorithm 1 Demographic Meta-Labeler

Input: NameMap : name → (inferenceMethods, perceptionLabel),

inferenceMethods : demographicLabel → probability,
allDemographicLabels

Output: ResultMap: name → metaLabel

```

1: ResultMap ← ∅ // Initialize all MetaLabels as None.
2: for name ∈ NameMap do
3:   labeled ← False
4:   inferenceMethods, perceptionLabel ← NameMap[name]
5:   for L ∈ allDemographicLabels do
6:     for inference ∈ inferenceMethods do
7:       if inference[L] ≥ threshold and ¬labeled then
8:         ResultMap[Name] ← L
9:         labeled ← True
10:      end if
11:    end for
12:  end for
13:  if ¬labeled then
14:    ResultMap[Name] ← perceptionLabel
15:  end if
16: end for
```

3.4.1 Race Meta-Labeling. In determining the specific meta-labeling algorithm to use for racial inference, following the general pattern described in Algorithm 1, we must determine the order in which to apply racial labels (the order of iteration for Line 5). Using a brute force search, we examine all permutations of the racial labels used throughout our census, and choose the label ordering resulting in the meta-labeling algorithm with the best alignment with the survey data, measured as accuracy. The chosen order is critical to the algorithm since it can resolve false positives. A name misclassified as White in one ordering could be correctly classified in a different ordering which labels the name as belonging to the self-reported racial category before applying any White labels.

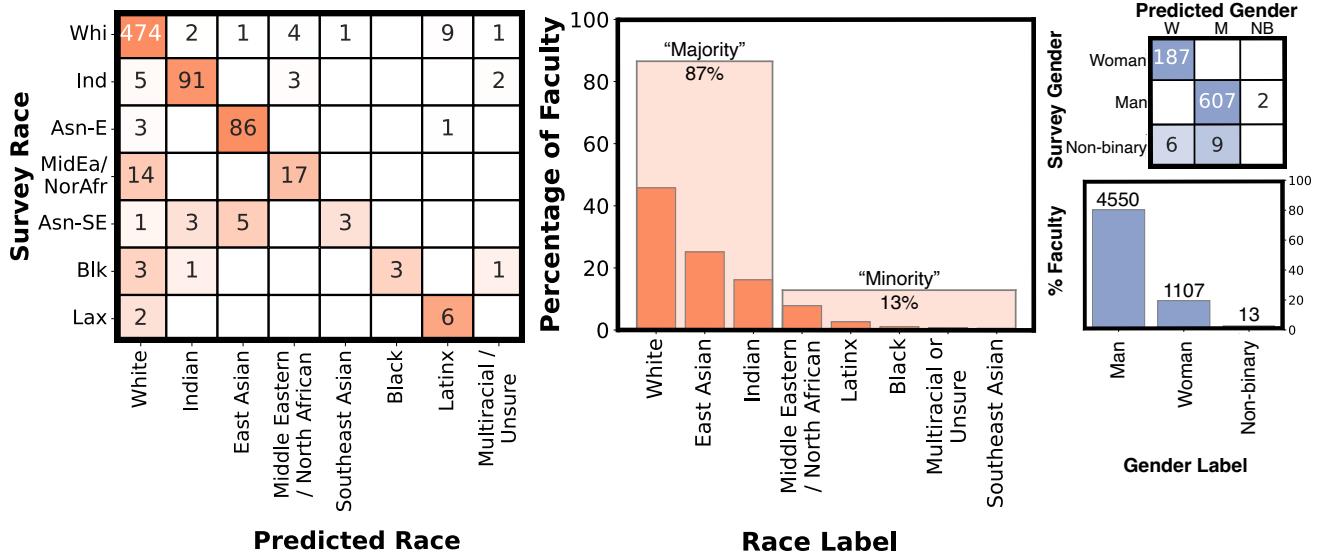


Figure 1: (Left) Confusion matrices showing agreement between our meta-labeling algorithms and survey self-reports. (Right) Distributions of demographic meta-labels for all 5,670 computer science faculty in our census.

The possible inference methods used for the race meta-labeler are EthniclWiki, EthniclFlorida [29], Ethnea [47] and perception-based labels. Preliminary comparisons with survey self-reports showed that these three automatic tools outperformed the other name-based methods for race inference. For each of these three tools, we selected separate confidence thresholds for every race (White, Indian, East Asian, Middle Eastern / North African, Black, Latinx, Southeast Asian, English and Jewish) by using the lowest threshold that still minimized false positives. For example, names of individuals who self-reported as Black were more commonly mislabeled English than White, and names labeled as Jewish by the inference tools were self-reported as Jewish, White or Middle Eastern. Thus we selected confidence thresholds such that no individuals who self-reported as Black were labeled English by Ethnea ($p \geq 0.99$). EthniclWiki mislabeled one individual who self-reported Black with $p = 0.99$, but a threshold of $p \geq 0.91$, captured 73 individuals with just this one false positive.

With these inference algorithms and associated thresholds chosen, the brute force search of permutations of eight racial categories (White, Indian, East Asian, Middle Eastern / North African, Black, Latinx, English and Jewish) resulted in 137 possible meta-labeling algorithms that tied with an accuracy of 92% (733 / 797) in comparison to the survey data. Most of these high-accuracy algorithms labeled White or English names first, due to strong alignment with survey self-reports for these categories. From among these top-performing permutations, we chose an algorithm which instead labeled non-White names first to reduce false positives in which an individual who does not identify as White is labeled as White. Pseudocode for the final race meta-labeler is given in Algorithm 2 in Appendix A.4, including the selected threshold values for name-based inference for each racial category. The final race meta-labeler

was validated using 5-fold cross-validation and through comparison with a different exhaustive methodology, also discussed in Appendix A.4.

Applying Algorithm 2 to 797 respondents who answered survey race questions resulted in 52 people classified with perception. Results for individuals who self-reported one race are shown in the confusion matrix in Figure 1. The race meta-labeler identifies White, Indian and East Asian names well, achieving group accuracies over 90% in all cases. It performs worse for minority groups, often misidentifying Middle Eastern / North African or Black names as White. Many of the people incorrectly predicted White who self-identify as Middle Eastern were due to names labeled as Jewish by Ethnea or EthniclWiki. These were individuals who wrote “Jewish” in the provided text box, but in some cases chose Middle Eastern / North African and in other cases White from our check list. We chose to label Ethnea and EthniclWiki Jewish names as White, but in many cases Middle Eastern / North African would also be appropriate. Out of the 55 respondents who reported multiple racial categories, our algorithm correctly classified them with one of their reported races in all but 2 cases.

3.4.2 Gender Meta-Labeling. For the gender meta-labeler, the key search space for Algorithm 1 is over possible thresholds per inference method (Line 7); we search for the thresholds which maximize accuracy of the resulting label with respect to survey self-reports. Our identified thresholds result in zero individuals who self-reported as men or women mislabeled by name-based inference.

We solely consider the NonQuamGender [11] inference method for the gender meta-labeler, since it out-performed the other name-based inference methods, achieving 87% accuracy overall and 98% accuracy on faculty names which were self-reported as men or women. East Asian names, known to have lower gender signal [11],

were disproportionately labeled incorrectly by NonQuam. Therefore, our gender meta-labeler requires higher confidence for labeling East Asian names.

A threshold of 0.75 applied to all faculty names in our survey data maximized alignment between NonQuam and self-identified labels, resulting in only six mislabeled self-identified men or women. All six of these individuals received East Asian as their top probability race label according to Ethnea. A NonQuam threshold of 0.85 for East Asian names resulted in 32 additional people classified with perception and captured 4 of these misclassified individuals. Two were still labeled incorrectly due to NonQuam gender signal below our threshold and a “no photo found” label from perception coders. The final gender meta-labeler first assigns NonQuam estimates gendered with $p > 0.85$ for those names identified as East Asian by Ethnea, and then assigns NonQuam estimates to all other names gendered with $p > 0.75$. Names which did not reach these confidence thresholds were labeled by perception. Complete pseudocode for the gender meta-labeler is provided in Algorithm 3 in the Appendix.

Applying the gender meta-labeler to 811 respondents who answered survey gender questions resulted in 117 people classified with perception. The algorithm did not misclassify any self-reported women and misclassified two self-reported men because no photo was found for perception labeling. The name-based inference methods we used for gender apply binary labels, therefore our algorithm can only label individuals as non-binary in the perception labeling step. Although 3 out of the 15 self-reported non-binary scholars were labeled as non-binary by coders, in our algorithm these names are first mislabeled as men or women by NonQuam, resulting in a misclassification of all 15 self-reported non-binary scholars. Comparisons between our gender meta-labeler and survey self-reports can be seen in Figure 1.

3.4.3 Results and Validation. We labeled all 5,670 individuals in our data using these race and gender meta-labeling algorithms (Algorithms 2 and 3). Our subsequent analyses use binary demographic variables classifying men vs. women or non-binary, and majority (White, Indian, East Asian) vs. minority (Middle Eastern / North African, Latinx, Black, Southeast Asian, Multiracial or unsure) race. For both gender and race, all minority labels combined made up less than 20% of individuals in the data. We use the term “minoritized” for both race and gender, to refer to these demographic groups who make up less than 20% of the population.

In agreement with distributions found in previous research on computer science [27], our faculty sample is primarily men with close to 20% women and 13 non-binary scholars. Racially, our computer science census is primarily White, East Asian and Indian scholars, with racial minorities making up 13% of the population combined. These distributions of our final demographic meta-labels are shown in Figure 1. Demographic distributions by title, for comparison with the Taulbee reports, are given in Appendix A.6. Our gender labels for the total population fall within 5% of Taulbee estimates for every category (full, associate, and assistant professors), and our race labels for the total population are within 10% of Taulbee estimates. Typically, Taulbee reported lower percentages of non-White faculty than we found, so we expect the difference is

in part due to Taulbee’s inclusion of “non-resident alien” and “residency unknown” categories which our label set does not include.

Our methods maximize alignment with self-identification in order to most appropriately label faculty in our census. However, the meta-labeling algorithms exclusively apply perception or name-based inference labels. These are social signals of race and gender, and therefore our final labels should be thought of as reflecting sociological formations over the population at large rather than the self-identification of particular individuals [20, 53].

4 Constructing CS Coauthorship Networks

Factors like institutional prestige and demographic traits structure social aspects of academic scholarship. In order to study these dynamics, we construct networks describing coauthorship relations among computer science faculty in our census. Nodes are the U.S. computer science faculty identified via our census (see Section 3). Two nodes are connected by an edge if they are listed as coauthors on a paper listed in the DBLP Computer Science Bibliography⁴, an online repository of publication data from major computer science venues. Each edge is associated with a publication year attribute and a weight representing the number of papers those coauthors published together in that year. The resulting network is made publicly available at [2].

4.1 Identifying faculty in DBLP

In order to match faculty from our census to the correct DBLP ID, we needed to disambiguate potentially matching names. Coders used both automated and manual techniques to collect DBLP IDs for every faculty member in our census. Where possible, DBLP IDs were automatically matched to faculty in the census by searching first initial, last name, and institution on DBLP and CSRankings; if all three matched, the DBLP ID was associated with that faculty member. If faculty members had multiple names listed in the DBLP-aliases file, publicly available through the CSRankings repository on GitHub⁵, all corresponding names were checked. All exact matches to faculty members’ names or CSRankings aliases were recorded, and close matches were checked manually before recording. For faculty with no matches according to this method, coders manually found publication lists on faculty members’ institutional websites. Then, coders chose a publication from this list and searched it by name in the publicly available “dblp.xml” file. This procedure resulted in complete coverage, including 2,124 faculty in our sample with multiple associated DBLP IDs. The full details of this procedure are in Appendix Section A.1.

4.2 Cumulative CS faculty coauthorship network

A cumulative coauthorship network, representing academic collaboration relationships between computer science faculty in our sample, was built using all faculty from our census and bibliometric data from DBLP. The network was constructed from all DBLP-indexed publications which list faculty from our census as authors. In total, the coauthorship network incorporates data from 3,652,370

⁴<https://dblp.org>

⁵<https://github.com/emeryberger/CSRankings/blob/gh-pages/dblp-aliases.csv>

journal articles and 3,563,465 publications in conference proceedings. Nodes in this network are faculty from our census, and each pair of nodes is connected by an undirected edge weighted with the number of their coauthored papers. The edges are annotated with a dictionary indicating how many papers faculty members coauthored together in each year.

The total network has $n = 5348$ nodes and $m = 35,551$ edges. Only 323 faculty in our census are not included in the network due to having no DBLP-indexed papers or no collaborations with other individuals in our census. Based on a manual examination of excluded faculty, we believe these are largely researchers in interdisciplinary computing who publish in venues that are not indexed by DBLP, e.g., *Science* or *PNAS*, or recent hires from international institutions who may not have yet collaborated with other U.S. computer science faculty.

4.3 CS faculty coauthorship networks by year

In addition to the cumulative collaboration network, built using all DBLP bibliometric data, we constructed collaboration networks for individuals' Ph.D. periods with the goal of understanding how a graduating students' position in the network shapes later career experiences. We only considered Ph.D. networks for *early career* faculty in our census, defined as those whose year of first publication was 2010 or later. There are a total of 2,041 such early career scholars in our census. The Ph.D. networks of current faculty who began publishing earlier than this would be poorly represented by our data which does not include retired faculty who may have been prominent members of the network during their Ph.D. periods.

For early career scholars, we can approximate the DBLP collaboration network at the time of their Ph.D. by identifying their year of first publication t_1 and building a network based on all publications in our DBLP bibliometric data with publication year $t \leq t_1 + 5$. Just as in the cumulative network, nodes in this network are faculty from our census, and each pair of nodes is connected by an undirected edge weighted with the number of their coauthored papers. We built one Ph.D. network for each of the years between 2015 and 2024. The Ph.D. periods of faculty whose year of first publication was $t_1 = 2010$ are represented by the 2015 Ph.D. network, and those with $t_1 = 2011$ are represented by the 2016 Ph.D. network, etc. These networks represent subgraphs of the cumulative DBLP publication data including only those papers which were published before or during the cutoff year. We will refer to the cumulative coauthorship network simply as the computer science network, and specify when results were calculated from the Ph.D. networks.

5 The Structure of Computer Science Collaboration Networks

Network centrality measures characterize differences in node position within a network, and are commonly interpreted as proxies for a node's structural importance or social influence [10, 51]. In a scientific coauthorship network, higher centrality individuals have greater or easier access to other parts of the network, and can act as bridges in receiving or distributing ideas. In contrast, lower centrality individuals have less or lower access to the same, which can be interpreted as a kind of epistemic dis-empowerment. Here, we

primarily focus on a node v_i 's closeness centrality $C(v_i)$, defined as

$$C(v_i) = \frac{n - 1}{\sum_{v_j \in V \setminus v_i} d(v_i, v_j)} , \quad (1)$$

where $d(v_i, v_j)$ is the shortest-path (geodesic) distance between nodes v_i and v_j in a graph with vertices V and where $|V| = n$. Thus, closeness centrality measures the inverse average path length to all other nodes in the network. We also report results for betweenness centrality, measuring how often a node falls on the shortest paths between other nodes, in Appendix B.1.

These centrality measures (closeness and betweenness) were calculated for all 5348 individuals in the cumulative coauthorship network. Additionally, we calculated the closeness and betweenness centrality of individuals during their Ph.D. periods for all 2,041 individuals whose year of first publication was $t_1 \geq 2010$. Ph.D. closeness and betweenness centrality were calculated with respect to the Ph.D. networks discussed above (Section 4.3).

Node centralities like closeness and betweenness tend to correlate with node degree, the number of connections in the network (number of collaborators). Hence, for comparison, we also calculated nodes' normalized degree,

$$D(v_i) = \frac{1}{n - 1} \sum_{v_j \in V \setminus v_i} A_{i,j} , \quad (2)$$

where $A_{i,j} = 1$ when nodes v_i and v_j have an edge between them, and is 0 otherwise. We use $D(v_i)$ to test the residual influence of faculty centrality in our census, accounting for their number of collaborators.

5.1 Faculty centrality and institutional prestige

We find a strong correlation between a faculty member's centrality in the CS coauthorship network and the prestige of their current institution—a finding not previously reported in studies of scientific coauthorship [49, 59].

Faculty in top-ranked CS departments are highly central in the network (higher closeness scores), whereas faculty at bottom ranked institutions are systematically more peripheral. This pattern appears as the strong linear relationship between closeness centrality and prestige rank ($R^2 = -0.81$) shown in Figure 2. The correlation is evidence that scientific coauthorships also align with the prestige hierarchy within computer science, in which coauthorships tend to be densely interconnected among computer scientists at the most prestigious institutions, and more diffuse among those at less prestigious institutions. Such a strong core-periphery pattern in computer science coauthorships may further reinforce the way prestige-structured faculty hiring networks can skew the spread of ideas in computer science, undervaluing good ideas that originate outside the core [36].

Moreover, this pattern is not sensitive to the particular measure of institutional prestige, and we find similar results when we use CSRankings or USNWR prestige scores. Figures showing the correlations between closeness centrality and prestige according to these university rankings are provided in Appendix B.1.

5.2 Homophily in coauthorship

Homophily is an empirical pattern in which members of the same social group are more likely to share social links with each other than

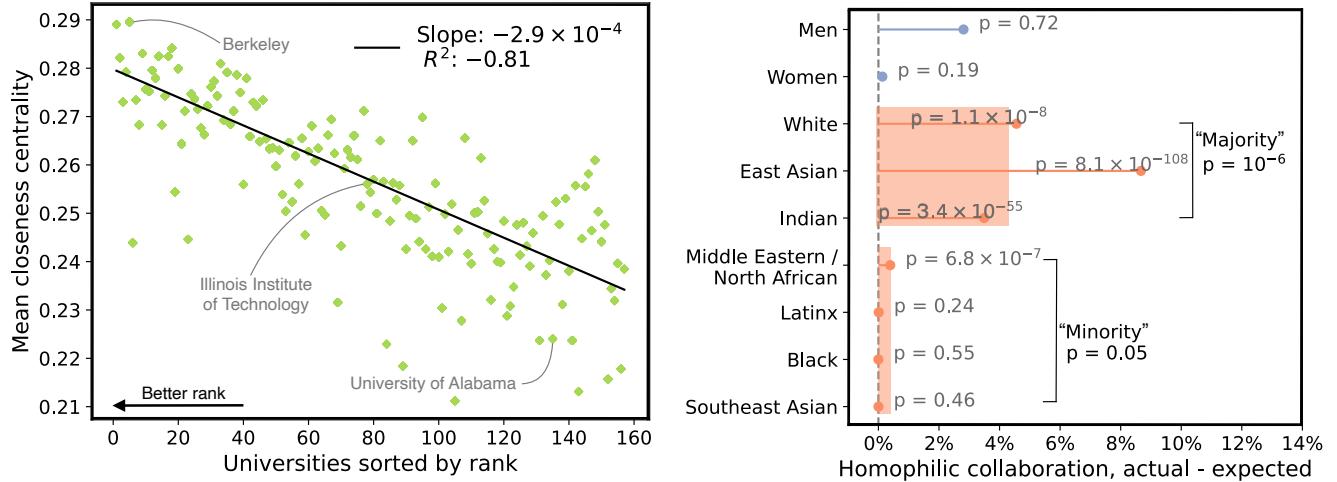


Figure 2: (Left) University prestige (rank, using the [50] measure of placement power; smaller score is more prestigious) versus the mean closeness centrality of faculty at that institution. (Right) Difference between observed frequency of homophilic collaborations in the CS coauthorship network, relative to its expectation under a permutation test that keeps the network fixed while shuffling node labels; p-values indicate the expected fraction of permutations with greater homophily under a Gaussian approximation of the sampled distribution.

with members of other groups. We measure the prevalence of race- and gender-based collaboration homophily in the CS coauthorship network by comparing the observed frequency of homophilic collaborations (same race, or same gender) to the expected frequency given random mixing with the observed demographic distributions in our network. We calculate the expected frequency under random mixing using a permutation model, in which the coauthorship network structure is fixed, and we measure the frequency of edges connecting nodes with the same label under random permutations of the demographic labels over the nodes. These comparisons between the observed and expected frequencies of homophilic collaborations are shown in Figure 2; p-values for these comparisons are calculated relative to a Gaussian approximation, parameterized by 1000 samples, of frequencies under the permutation model.

We find strong evidence of racial homophily among majority race groups (White, East Asian and Indian) in the CS coauthorship network (all $p < 0.05$). In contrast, the racial homophily rates among most minority race groups (Latinx, Black, Southeast Asian) are not significantly different from their expected values ($p > 0.05$), although we do observe a significant difference for the largest minoritized group (Middle Eastern / North African). The significant differences for majority racial groups and non-significant differences for most minority racial groups may reflect a lack of statistical power that hides homophily among individuals within these minority groups because of their overall low representation, or it could indicate genuine non-significance due to, e.g., limited practical options for homophilic coauthorships.

For both men and women, the observed rates of homophilic collaboration were not significantly different from what we expect under the null model. This negative result contrasts with previous results finding that men, but not women, exhibit homophily in their scientific collaborations [26, 34]. This may indicate that properly

measuring homophily in coauthorship networks requires controlling for the network structure itself, as under our permutation test, or it may reflect differences specific to computer science collaborations (i.e. gendered homophily may vary by discipline and/or region).

5.3 Disparities in network position

Women, scholars with minoritized race identities, and those from low-ranked institutions are less central in the CS coauthorship network than are men, scholars in majority race groups, and those from high-ranked institutions. Comparisons for each of these pairs of groups, along with p-values from pooled t-tests, can be seen in Figure 3. We found similar patterns for betweenness centrality and degree, which are shown in Appendix B.1. The same patterns of disparity shown in Figure 3 were also significant for these other network metrics for all demographics except gender, for which the betweenness centrality of women and men did not significantly differ (pooled t-test, $p = 0.31$).

In addition to disparities in the network centrality of demographic groups shown in Figure 3, we found that individuals with minoritized race identities had on average fewer collaborators ($p = 1.2 \times 10^{-4}$) and published less ($p = 4.6 \times 10^{-5}$) than individuals from the majority. In pairwise comparisons, these disparities also held for gender, replicating the results of [58]. However, when we controlled for researcher's academic age, measured as the number of years since first publication, this result was no longer significant. This indicates that the apparent gender difference in degree and productivity is driven by a proportionally greater number of early-career women scholars. Previous research has shown that the predominance of women scholars in early-career stages is largely due to increased hiring of women in recent years [27]. To a lesser

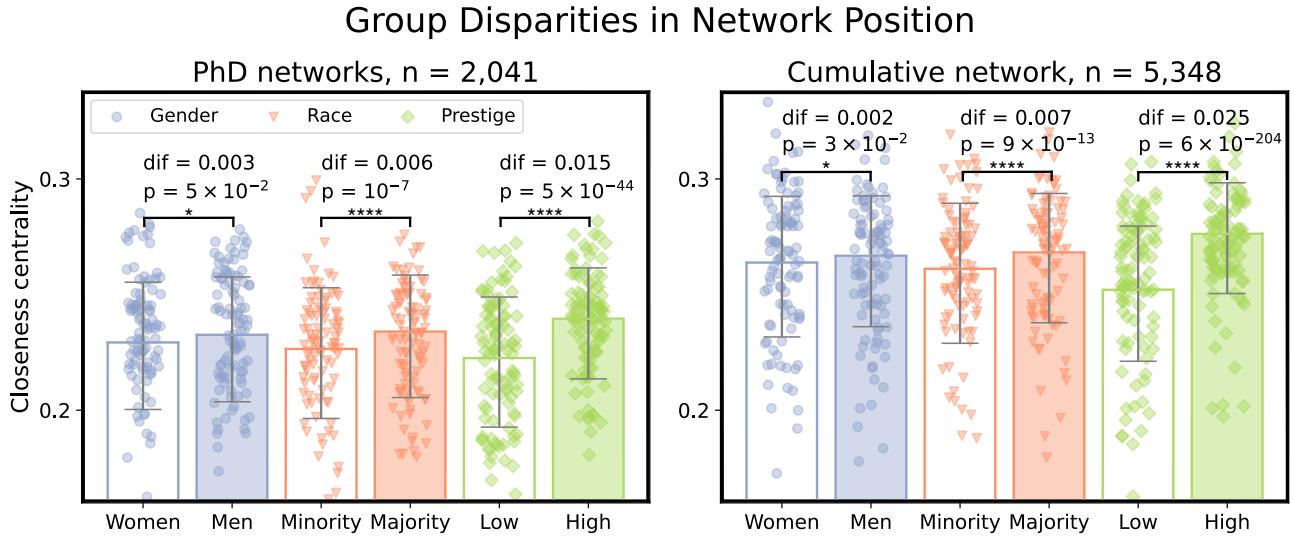


Figure 3: Comparison of closeness centralities by demographic label and by institutional prestige, showing significant differences (pooled t-tests) between women and men, between racial minority and majority groups, and between low and high prestige institutions (split at the median rank), for (Left) Ph.D. networks and (Right) the cumulative CS coauthorship network.

degree, this is explained by higher rates of attrition among women than men in U.S. faculty [27, 44].

We found no correlation between our demographic variables and institution rank for either the Ph.D. or current institutions of faculty in our census. This negative result implies that the disparities in closeness centrality presented in Figure 3 cannot be explained by the institutional career trajectories of minoritized scholars. Why we see these demographic disparities remains an open question.

Making sense of gender and race as causal variables is highly debated [57]. Thus, the question “does one’s race/gender cause them to be peripherally located in the collaboration network” is perhaps better phrased in terms of social beliefs about these demographic traits or institutions of sexism/racism. An alternative explanation for the lower centrality of minoritized groups comes from the network modeling literature. [24] shows that under certain conditions of network growth, homophily in link formation can lead members of minority groups to have lower degrees, on average, than members of majority groups. We find evidence of race- (majority $p = 1.3 \times 10^{-6}$, minority $p = 0.05$) but not gender-based (men $p = 0.72$, women $p = 0.19$) homophily in our CS coauthorship network; however, it is not clear whether the simulation results from [24] hold for real-world networks, or whether our CS network exhibits enough homophily (i.e., the observed difference between the actual and expected frequency of homophilic collaborations is great enough) to explain the size of the differences in centrality we see between minority and majority groups.

6 Intervening to Address Demographic and Prestige Disparities

We now introduce a simple network intervention that can help mitigate disparities in the centrality of individuals from minority racial groups at low prestige institutions. The simulated intervention adds

one single edge to the coauthorship network by pairing a target individual with certain characteristics with a sponsor individual, e.g., by providing a research fellowship to an individual to spend a semester in the research group of a sponsor. In practice, a new collaboration may result in multiple new coauthorship edges; we focus on the single-edge case as a minimal change to the network. We consider this intervention in two contexts. First, we show that such an intervention can increase a target’s centrality in the coauthorship network, which we interpret as improving their relative status in the research community. Second, we show that when applied to scholars in their Ph.D. periods, the resulting improvements in centrality can be expected to improve their placement outcomes on the faculty job market.

6.1 Suggested collaborations improve target centrality

In general, a peripheral node’s centrality can be improved by linking it with a highly central node in the network. We propose a simple intervention along these lines to improve the position of peripheral nodes in the CS coauthorship network. However, an intervention on individuals with low centrality is impractical because of the analysis required to identify people based on their network position. Instead, our intervention targets the population of individuals with minoritized race identities who are currently employed at lower-ranked institutions, whom our results (Fig. 3) suggest are statistically more likely to be less central than average in the coauthorship network compared to faculty with other social identities at the same institutions.

To begin, we bin institutions into five groups: institutions ranked from 21–37, 38–60, 61–83, 84–114 or 115–157, and we define target groups for each bin as the corresponding institutions’ set of minority race faculty in our census. We define the sponsor group

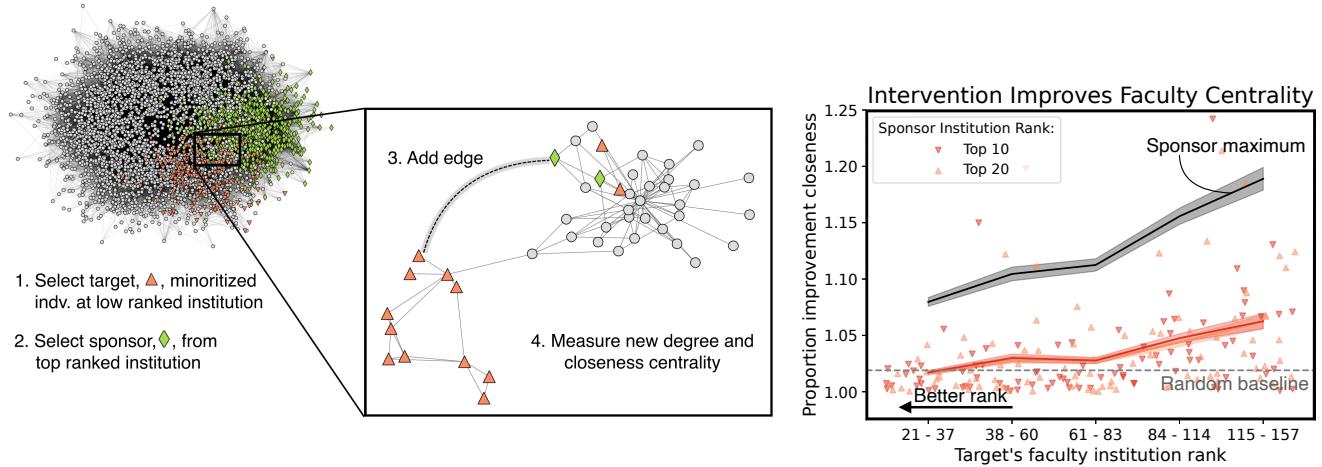


Figure 4: (Left) Schematic of the proposed intervention to improve the centrality of target individuals. **(Right)** This simulated intervention results in increases in the closeness centrality of target individuals with minoritized race identities at low prestige current institutions (rank using [50] placement power). Our interventions (red lines) are compared to greedy selection of sponsor who would maximally improve target's closeness (black line) and to a random baseline.

of faculty as either all faculty at top-10 ranked institutions or all faculty at top-20 ranked institutions. For every individual in a target group, we apply the simulated intervention shown in Fig. 4. First, we randomly select a sponsor from the sponsor group and add an edge connecting the current target to the selected sponsor. Finally, we measure the post-intervention closeness centrality of both target and sponsor. Notably, both targets and sponsors are selected with no knowledge of the pre-intervention coauthorship network. Instead, by selecting sponsors from top-ranked institutions, it becomes statistically more likely that they are also highly central in the network.

The simulated intervention results in a positive improvement in the centrality of every individual, as shown in Fig. 4. Both target and sponsor groups experience an increase in their closeness centrality as a result of the intervention, although this change is far larger for the target group. Moreover, the positive slope in Fig. 4, shows that the network intervention has a greater effect for scholars at lower ranked institutions. We find little difference in post-intervention centralities for whether the sponsor group is faculty at top-10 or top-20 ranked institutions.

By constraining the intervention to one edge, we can further determine the maximum improvement possible and a random expected baseline change in centrality. The random baseline was calculated as the average improvement in centrality resulting from adding one edge between a randomly selected pair of individuals from the entire network. We also searched over all possible edge additions in the cumulative CS coauthorship network for the edge that resulted in the maximum improvement in the centrality of a node. This edge connected a scholar from a rank 130 institution to a scholar from a rank 5 institution and resulted in a proportional improvement of 1.94 in the former scholar's closeness centrality, highlighting the importance of institutional prestige to network centrality. In addition to this absolute maximum, Fig. 4 compares

to a greedy maximum defined within the scope of our intervention. Instead of connecting each individual in the target group to a sponsor chosen randomly from the sponsor group, we select the sponsor from that group that results in the largest possible increase in the given target individual's closeness centrality. Importantly, such an intervention would not be possible without knowledge of the coauthorship network structure, but it provides an upper bound for our intervention defined without this knowledge.

6.2 Predicting job placement based on Ph.D. prestige, centrality and degree

Beyond the inherent benefits of increasing network centrality, we conjecture that centrality plays a practical role in improving individuals' career prospects. We hypothesize that when an early career scholar enters the job market, it matters how many professional connections they have and to whom they connect. We test this hypothesis by using linear regression models to predict the rank of a hiring institution for each faculty in the census, based on the rank of their Ph.D. institution, network degree (number of coauthors), and closeness centrality. The models are trained on the data from 2,041 early career scholars in our census, using the degree and closeness centrality calculated from their Ph.D. networks, described in Section 4. Strictly interpreted, this analysis assumes that the current institutions in our census are the same institution that first hired that particular faculty, and that the Ph.D. centrality we calculated accurately represents their network position at that time. Given that most faculty stay at their first university job [50], we find this to be a reasonable assumption.

The results show that while Ph.D. prestige is the most important factor in predicting placement, network degree and network centrality both also have a significant effect (Fig. 5). The first model (M1) shows how current institution rank depends on Ph.D. institution rank and Ph.D. closeness centrality. The positive coefficient

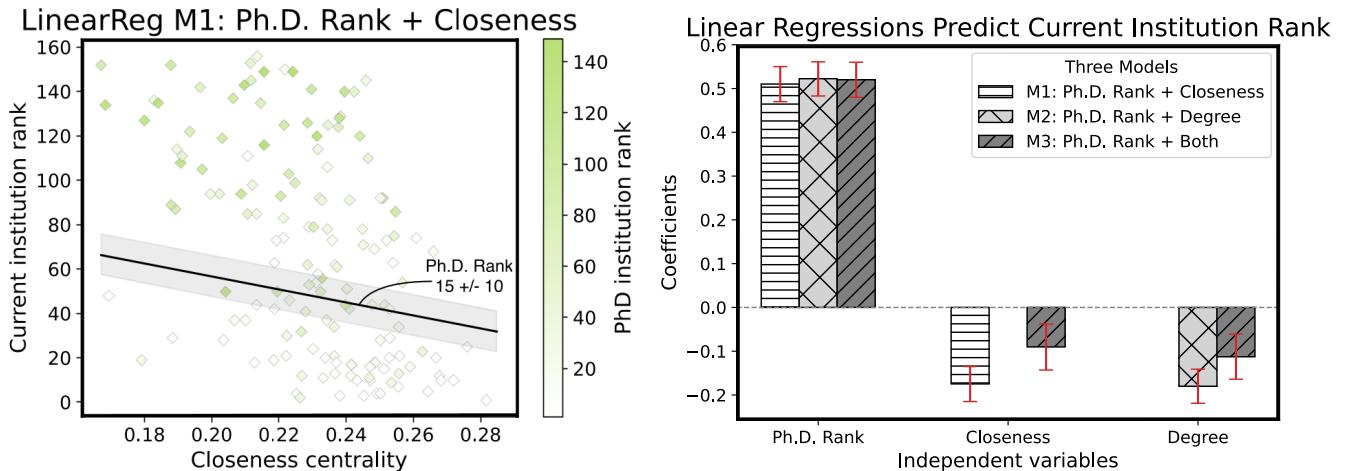


Figure 5: Model M1 line of best fit through a random subset of 150 points from fit data (Left) and standardized coefficients (Right) of three linear regression models predicting placement rank based on Ph.D. rank and network characteristics.

for Ph.D. rank shows that individuals from low-ranked Ph.D. institutions are likely to be hired at low-ranked institutions, and vice versa for high rank. The negative coefficient for closeness centrality shows that more central nodes tend to be placed at better institutions. Although Ph.D. rank is more than twice as important to the model, closeness is still a significant predictor of current institution rank. The second model (M2), which predicts placement prestige using Ph.D. prestige and degree, replicates these results.

Taken together, the M1 and M2 models show that individuals' position and number of connections within the coauthorship network are individually predictive of placement outcomes, above and beyond the predictive value of Ph.D. prestige. Centrality measures like closeness are also known to correlate with network degree. This co-linearity is evident in the decreased importance of both factors in the third model (M3), which predicts current institution rank based on Ph.D. rank, network degree and closeness centrality. However, while degree has a mediating effect on closeness in M3, both factors remain significant predictors of current institution rank, of comparable importance. These results indicate that in addition to how many coauthors (degree), for predicting placement in the faculty hiring market, it matters which specific coauthors (closeness) a researcher has.

The idea that it matters who you collaborate with is corroborated by the intervention results in Fig. 4, where pairing targets from minoritized racial groups at more peripheral institutions with sponsor individuals at central institutions disproportionately improves the centrality of the target group. Further, selecting targets and sponsors using minoritized race and institutional prestige is substantially more effective than pairing randomly chosen individuals.

6.3 Improving Ph.D. centrality and predicted rank of placement institution

The correlation between centrality and placement institution rank (Fig. 5) suggests that intervening to improve the centrality of early

career scholars may have material benefits, improving the rank of institution at which they are placed in the academic job market.

We test this possibility with our second intervention experiment applied to the Ph.D. coauthorship networks (Section 4). This intervention again uses minoritized race and lower Ph.D. institution rank to define the target populations, and we again bin institutions into 5 bins by their Ph.D. rank, ensuring an equal number of Ph.D. scholars in each bin. For each target individual in a given bin, a random sponsor is selected from the sponsor group, again defined as faculty either at top-10 or at top-20 institutions by rank. The random sponsor is accepted if their year of first publication is at least 5 years prior to the target individual's year of first publication, which ensures that they are more senior than the target individual; if not, we draw a new random sponsor. We then add the new coauthorship edge and calculate the post-intervention closeness centrality of both target and sponsor in the the Ph.D. coauthorship network.

The improvement in centrality resulting from this intervention is compared to a random baseline and to a greedy maximum, using the same strategy as our first intervention experiment (Fig. 4). Notably, the proportional increase in target Ph.D. centrality (Fig. 6) is greater than the improvement found in the first experiment. Ph.D. centrality is calculated using fewer coauthorships—only the first five years of the target individual's publication record—meaning that one added coauthorship has a bigger impact on their network centrality at this time than it does in their later career. As before, we find that the intervention has the largest effect for scholars at the lowest ranked Ph.D. institutions, and the choice of sponsor group (top-10 or top-20) makes little difference to targets' improvements in centrality.

Improving one's centrality in the academic network has inherent benefits, such as increased ability to distribute one's scientific ideas, and implies an improvement in the rank of the institution a scholar is hired at. We estimated the improved ranks of target individual's placement institutions using M1 with the updated closeness centrality of target individuals post-intervention. These estimates

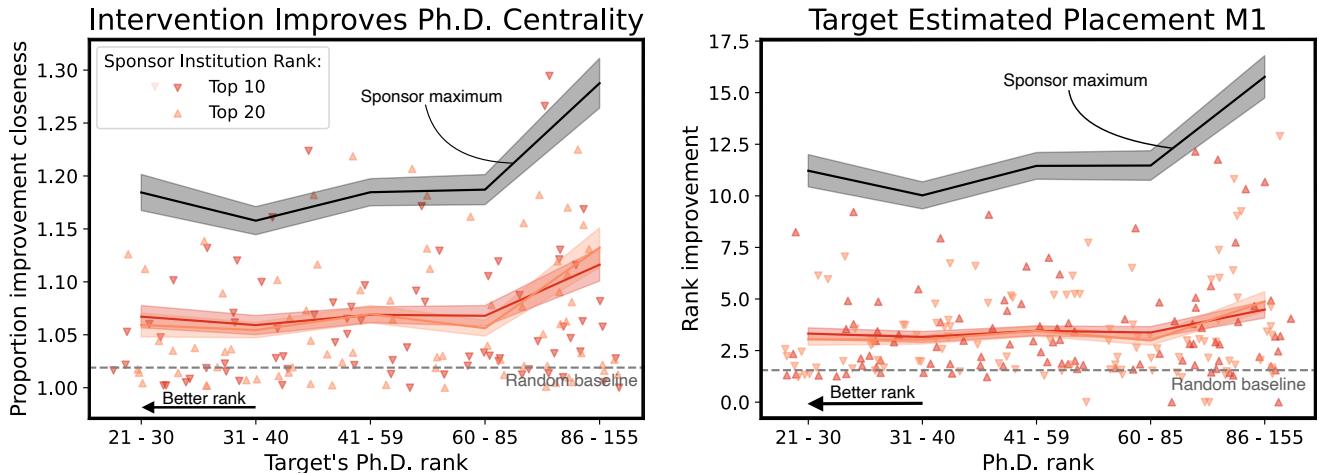


Figure 6: Simulated collaboration adds one edge between sponsor from top-ranked institution and racially minoritized target from low-ranked Ph.D. institution (rank using [50] placement power). The intervention increases the (Left) centrality of target individuals and improves the (Right) estimated rank of their placement institutions on the academic job market. Comparisons to greedy selection of sponsor who would maximally improve target’s closeness (black line) and to random baseline.

place target scholars at a higher ranked institution than their current institution. This is shown in Fig. 6, along with the estimated improvement in placement rank from the greedy maximum improvement in centrality that could result from our intervention.

Although the simulated intervention suggests a causal relationship between targets’ Ph.D. centrality and placement institution ranks, our results are not causal. In order to make that claim, the intervention would need to be carried out in practice, e.g., implemented via a fellowship program. We observe that since computer science papers typically have more than two coauthors, any new collaborations generated by such an intervention program would likely improve the target Ph.D. student’s closeness centrality more than our simulated interventions. Additional benefits would include the establishment during the fellowship period of new professional relationships with faculty beyond new coauthors, which could also improve their academic job market placement. That is, we expect that our experiments underestimate the real-world improvement in both the centrality and placement from intervention.

7 Discussion and Conclusion

The analyses presented were carried out using a novel data set, hand-collected from Computer Science department and faculty websites in the academic year 2023-24. We make this data set available for reuse by the community [2]. These data address the dearth of demographic information available to previous research, by combining name-based inference and perception to generate demographic meta-labels that align with survey self-reports.

We found that centrality in the computer science coauthorship network increases as institutional prestige increases, and identified demographic disparities in the closeness centrality of individuals in the CS coauthorship network. By suggesting a single collaboration (edge intervention in the network), we were able to improve the centrality of target individuals with minoritized race identities at

lower-ranked institutions. This intervention is lightweight and realistic since it did not require access to coauthorship network information, using only institutional prestige and identity data. Furthermore, when applied to Ph.D. student scholars, our simulated intervention improves the predicted rank of targets’ placement institutions on the academic job market.

These results expand on evidence from previous literature showing that differences in institutional prestige drive many factors related to academic success [28, 50, 52, 59]. In parallel with the findings of [36], which showed that prestige amplifies the spread of scientific ideas originating at top universities, the pattern we find suggests that peripheral individuals are epistemically dis-empowered in receiving and distributing scientific knowledge. Our results (Section 6.2) indicate that Ph.D. scholars’ network attributes have practical consequences in determining the rank of their placement institutions on the academic job market, and that while node degree explains a portion of this, network position has residual predictive power even when degree is accounted for.

In intervening to mitigate these disparities in network centrality, and to potentially improve job placement in the faculty hiring market, we designed a network intervention that could be realistically implemented, e.g. through a targeted fellowship program. We note that our intervention experiment on Ph.D. placement was based on a causal assumption between centrality and job placement. Assessing this causal assumption would be a valuable direction for future work. Furthermore, of course, the success of a scientific collaboration depends on the specific people involved, and is more complicated than we assume in our network intervention. At the same time, our intervention was intentionally minimal—adding only a single new coauthorship edge—while many collaborations today involve many more authors, which would add a corresponding number of new edges. Hence, our estimates of the improvement of such interventions may be conservative.

All of our results are limited to tenured and tenure-track faculty in Ph.D. granting computer science departments at U.S. institutions, which excludes the effects or importance of international coauthorships and coauthorships with researchers at non-academic institutions, e.g. industry researchers. We also note that the demographics of CS faculty differ from other disciplines. Hence, the specific meta-labeling algorithms we develop and apply to our data (Section 3.4) should not be used for other populations. However, the methodological framework is general, and can be adapted to other settings for future work. The correlation between prestige and researcher centrality in the coauthorship is a novel observation, and future work should investigate its prevalence in other fields.

Guided by an ameliorative approach, we uncovered demographic and prestige inequities in network centrality for the purpose of addressing these disparities. We do not know why gender and race disparities in network centrality causally exist. An ethical risk of publishing results reporting correlations between demographics and other variables is to reify these associations. Gender and race are sociological formations that change over time. Future work may aim to provide the causal explanations our study lacks by appealing to social beliefs about such demographic traits or to institutions of sexism/racism. Another direction for future work is to investigate why the gender disparity we found is smaller than the corresponding racial disparity in network centrality (Section 5.3), and whether this is explained by changes over time.

Future work could test whether our proposed intervention works in practice, either using longitudinal data or a real-world program implementation. Such a program may help to reduce observed inequities in the creation and dissemination of scientific knowledge, and experimentally discover the causal factors involved in differential network position and academic career outcomes.

Acknowledgments

This work was supported in part by NSF Award IIS-1955321 (S.F., M.E.E., N.H. and M.F.), 2219609 (A.C., C.C.R. and K.B.) and IIS-1956286 (B.D.S.).

References

- [1] Shlomo Argamon, Moshe Koppel, Jonathan Fine, and Anat Rachel Shimoni. 2003. Gender, genre, and writing style in formal written texts. *Text & talk* 23, 3 (2003), 321–346.
- [2] Kate Barnes, Mia Ellis-Einhorn, Carolina Chavez-Ruelas, Nayera Hasan, Mohammad Fanous, Blair D. Sullivan, Sorelle Friedler, and Aaron Clauset. 2025. Edge interventions can mitigate demographic and prestige disparities in the Computer Science coauthorship network. <https://github.com/k-barnes/cs-us-faculty-collaboration-network>.
- [3] Ashkan Bashardoust, Sorelle Friedler, Carlos Scheidegger, Blair D Sullivan, and Suresh Venkatasubramanian. 2023. Reducing Access Disparities in Networks using Edge Augmentation. In *Proceedings of the 2023 ACM conference on fairness, accountability, and transparency*. 1635–1651.
- [4] Michael N. Bastedo and Nicholas A. Bowman. 2010. U.S. News & World Report College Rankings: Modeling Institutional Effects on Organizational Reputation. *American Journal of Education* 116 (2010), 163–183.
- [5] Ruben Becker, Gianlorenzo D’Angelo, and Sajjad Ghobadi. 2023. Improving fairness in information exposure by adding links. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 14119–14126.
- [6] Nicolas Bérubé, Gita Ghiasi, Maxime Sainte-Marie, and Vincent Larivière. 2020. Wiki-Gendersort: Automatic gender detection using first names in Wikipedia. (2020).
- [7] Aditya Bhaskara, Alex Crane, Shweta Jain, Md Mumtahin Habib Ullah Mazumder, Blair D Sullivan, and Prasanth Yalamanchili. 2024. Optimizing Information Access in Networks via Edge Augmentation. *arXiv preprint arXiv:2407.02624* (2024).
- [8] Gecia Bravo-Hermsdorff, Valkyrie Fels, Emily Ray, Lee M. Gunderson, Mary E. Helander, Joana Maria, and Yael Niv. 2019. Gender and collaboration patterns in a temporal scientific authorship network. *Applied Network Science* 4 (2019), 112.
- [9] U.S. Census Bureau. 2024. Updates to Race/Ethnicity Standards for Our Nation. <https://www.census.gov/about/our-research/race-ethnicity/standards-updates.html>
- [10] Ronald S Burt. 2004. Structural holes and good ideas. *American journal of sociology* 110, 2 (2004), 349–399.
- [11] Ian Van Buskirk, Aaron Clauset, and Daniel B. Larremore. 2023. *An Open-Source Cultural Consensus Approach to Name-Based Gender Classification*. Proceedings of the International AAAI Conference on Web and Social Media. doi:10.1609/icwsm.v17i1.22195
- [12] B.J. Casad, J.E. Franks, C.E. Garasky, M.M. Kittleman, Roesler A.C., Hall D.Y., and Z. W. Petzel. 2021. Gender inequality in academia: Problems and solutions for women faculty in STEM. *J Neurosci Res.* 99 (2021), 13–23.
- [13] Aaron Clauset, Samuel Arbesman, and Daniel B. Larremore. 2015. Systematic inequality and hierarchy in faculty hiring networks. *Science advances* 1, 1 (2015).
- [14] Theodosius Dobzhansky. 1973. Nothing in Biology Makes Sense Except in the Light of Evolution. *The American Biology Teacher* (1973).
- [15] C. H. Dupree and C. M. Boykin. 2021. Racial Inequality in Academia: Systemic Origins, Modern Challenges, and Policy Recommendations. *Policy Insights from the Behavioral and Brain Sciences* 8 (2021), 11–18.
- [16] Dirk M Elston. 2021. Participation bias, self-selection bias, and response bias. *Journal of the American Academy of Dermatology* (2021).
- [17] Golnoosh Farnad, Behrouz Babaki, and Michel Gendreau. 2020. A unifying framework for fairness-aware influence maximization. In *Companion Proceedings of the Web Conference 2020*. 714–722.
- [18] Benjamin Fish, Ashkan Bashardoust, Danah Boyd, Sorelle Friedler, Carlos Scheidegger, and Suresh Venkatasubramanian. 2019. Gaps in information access in social networks. In *The World Wide Web Conference*. 480–490.
- [19] Richard B. Freeman and Wei Huang. 2015. Collaborating with People Like Me: Ethnic Coauthorship within the United States. *Journal of Labor Economics* 33 (2015), S289–S318.
- [20] Vagrant Gautam, Arjun Subramonian, Anne Lauscher, and Os Keyes. 2024. Stop! in the name of flaws: Disentangling personal names and sociodemographic attributes in NLP. *arXiv preprint arXiv:2405.17159* (2024).
- [21] Sally Haslanger. 2000. Gender and Race:(What) Are They?(What) Do We Want Them to Be? 1. *Nous* 34, 1 (2000), 31–55.
- [22] Travis A. Hoppe, Aviva Litovitz, Kristine A. Willis, Rebecca A. Meseroll, Matthew J. Perkins, Alison F. Davis B. Ian Hutchins and, Michael S. Lauer, Hannah A. Valentine, James M. Anderson, and George M. Santangelo. 2019. Topic choice contributes to the lower rate of NIH awards to African-American/black scientists. *Science Advances* 5 (2019).
- [23] Junming Huang, 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 Science USA* 117, 9 (2020), 4609–4616.
- [24] Fariba Karimi, Mathieu Génois, Claudia Wagner, Philipp Singer, and Markus Strohmaier. 2018. Homophily influences ranking of minorities in social networks. *Scientific reports* 8, 11077 (2018), 1–12.
- [25] Diego Kozlowski, Dakota S Murray, Alexis Bell, Will Hulsey, Vincent Larivière, Thema Monroe-White, and Cassidy R Sugimoto. 2022. Avoiding bias when inferring race using name-based approaches. *Plos one* 17, 3 (2022).
- [26] Marek Kwiek and Wojciech Roszka. 2021. Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics* 15, 3 (2021).
- [27] Nicholas LaBerge, Kenneth Hunter Wapman, Aaron Clauset, and Daniel B Larremore. 2024. Gendered hiring and attrition on the path to parity for academic faculty. *eLife* 13 (2024).
- [28] Nicholas LaBerge, Kenneth Hunter Wapman, Allison C. Morgan, Sam Zhang, Daniel B. Larremore, and Aaron Clauset. 2022. Subfield Prestige and Gender Inequality among U.S. Computing Faculty. *Commun. ACM* 65, 12 (2022), 46–55.
- [29] Suriyan Laohaprapanon, Gaurav Sood, and Bashar Naji. 2022. *ethnicolr: Predict Race and Ethnicity From Name*. <https://github.com/appeler/ethnicolr>.
- [30] Vincent Larivière, Benoit Macaluso, Éric Archambault, and Yves Gingras. 2010. Which scientific elites? On the concentration of research funds, publications and citations. *Research Evaluation* 19, 1 (2010), 45–53.
- [31] Eun Lee, Aaron Clauset, and Daniel B. Larremore. 2012. The dynamics of faculty hiring networks. *EPJ Data Science* 10, 48 (2012).
- [32] Weihua Li, Sam Zhang, Zhiming Zheng, Skyler J. Cranmer, and Aaron Clauset. 2022. Untangling the network effects of productivity and prominence among scientists. *Nature Communications* 13 (2022), 4907.
- [33] Jeffrey W. Lockhart, Molly M. King, and Christin Munsch. 2023. Name-based demographic inference and the unequal distribution of misrecognition. *Nature Human Behaviour* 7, 7 (2023), 1084–1095.
- [34] Andrea Löther and Frederike Freund. 2024. Gender-Based Homophily in International Research Collaborations. *Social Sciences* 13, 549 (2024).

- [35] Theodore Masters-Waage, Christiane Spitzmueller, Ebenezer Edema-Sillo, Ally St. Aubin, Michelle Penn-Marshall, Erika Henderson, Peggy Lindner, Cynthia Werner, Tracey Rizzuto, and Juan Madera. 2024. Underrepresented minority faculty in the USA face a double standard in promotion and tenure decisions. *Nature Human Behaviour* 8, 11 (2024), 2107–2118.
- [36] Allison C. Morgan, Dimitrios J. Economou, Samuel F. Way, and Aaron Clauset. 2018. Prestige drives epistemic inequality in the diffusion of scientific ideas. *EPJ Data Science* 7, 40 (2018), 40.
- [37] Allison C. Morgan, Nicholas LaBerge, Daniel B. Larremore, Mirta Galesic, Jennie E. Brand, and Aaron Clauset. 2022. Socioeconomic roots of academic faculty. *Nature Human Behaviour* 6 (2022), 1625–1633.
- [38] Namesor. [n. d.]. Namstor, name checker for gender, origin and ethnicity determination. <https://namstor.app>. year accessed: 2024.
- [39] M. E. J. Newman. 2004. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the National Academy of Sciences USA* 101 (2004), 5200–5205.
- [40] Emre Sarigöl, René Pfitzner, Ingo Scholtes, Antonios Garas, and Frank Schweitzer. 2014. Predicting scientific success based on coauthorship networks. *EPJ Data Science* 9 (2014).
- [41] Akrati Saxena, George Fletcher, and Mykola Pechenizkiy. 2024. Fairsna: Algorithmic fairness in social network analysis. *Comput. Surveys* 56, 8 (2024), 1–45.
- [42] Morgan Klaus Scheuerman and Jed R Brubaker. 2018. Gender is not a Boolean: towards designing algorithms to understand complex human identities. In *In Participation+ Algorithms Workshop at CSCW 2018*.
- [43] I. H. Settles, M. K. Jones, N. T. Buchanan, and K. Dotson. 2021. Epistemic exclusion: Scholar(ly) devaluation that marginalizes faculty of color. *Journal of Diversity in Higher Education* 14, 4 (2021), 493–507.
- [44] Katie Spoon, Nicholas LaBerge, Kenneth Hunter Wapman, Sam Zhang, Allison C Morgan, Mirta Galesic, Bailey K Fosdick, Daniel B Larremore, and Aaron Clauset. 2023. Gender and retention patterns among US faculty. *Science advances* 9, 42 (2023).
- [45] Katie Spoon, Joanna Mendy, Maria Martinez, Mirta Galesic, Daniel B Larremore, Aaron Clauset, and Lauren A Rivera. 2024. Gendered devaluation underlies faculty retention. (2024). Preprint, socarxiv/g6xwk_v1.
- [46] Ana-Andreea Stoica and Augustin Chaintreau. 2019. Fairness in social influence maximization. In *Companion Proceedings of The 2019 World Wide Web Conference*, 569–574.
- [47] Vette L. Torvik and Sneha Agarwal. 2016. *Ethnea: an instance-based ethnicity classifier based on geo-coded author names in a large-scale bibliographic database*. Proceedings of the International AAAI Conference on Web and Social Media.
- [48] Pucktada Treeratpituk and C. Lee Giles. 2012. Name-ethnicity classification and ethnicity-sensitive name matching. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 26. 1141–1147.
- [49] Madalina Vlasceanu, Miroslav Dudík, and Ida Momennejad. 2022. Interdisciplinarity, Gender Diversity, and Network Structure Predict the Centrality of AI Organizations. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) (FAccT '22). 1–10.
- [50] Kenneth Hunter Wapman, Sam Zhang, Aaron Clauset, and Daniel B. Larremore. 2022. Quantifying hierarchy and dynamics in US faculty hiring and retention. *Nature* 610, 7930 (2022), 120–127.
- [51] Stanley Wasserman and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge University Press.
- [52] Samuel F Way, 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 (2019), 10729–10733.
- [53] Sasha Weitman. 1981. Some methodological issues in quantitative onomastics. 181–196 pages.
- [54] Damani K White-Lewis. 2020. The facade of fit in faculty search processes. *The Journal of Higher Education* 91, 6 (2020), 833–857.
- [55] Kjersten Bunker Whittington, Molly M. King, and Isabella Cingolani. 2024. Structure, status, and span: gender differences in co-authorship networks across 16 region-subject pairs (2009–2013). *Scientometrics* 129 (2024), 147–179.
- [56] Dennis Robert Windham, Caroline J Wendt, Alex Crane, Madelyn J Warr, Freda Shi, Sorelle A Friedler, Blair D Sullivan, and Aaron Clauset. 2024. Fast algorithms to improve fair information access in networks. *arXiv preprint arXiv:2409.03127* (2024).
- [57] James Woodward. 2003. *Making things happen: A theory of causal explanation*. Oxford university press.
- [58] Xiao Han T Zeng, Jordi Duch, Marta Sales-Pardo, João AG Moreira, Filippo Radicchi, Haroldo V Ribeiro, Teresa K Woodruff, and Luis A Nunes Amaral. 2016. Differences in collaboration patterns across discipline, career stage, and gender. *PLoS biology* 14, 11 (2016).
- [59] Sam Zhang, Kenneth Hunter Wapman, Daniel B Larremore, and Aaron Clauset. 2022. Labor advantages drive the greater productivity of faculty at elite universities. *Science advances* 8, 46 (2022).

A Research Methods

A.1 Coding Procedure

The following coding procedure was used to collect faculty demographic information.

- (1) Find the institution's CS faculty page. Google the institution and navigate to the CS department website listing all faculty.
- (2) For each faculty member:
 - (a) Check the faculty member's title.
 - (i) Continue to the next steps for the following titles:
 - Assistant Professor
 - Associate Professor
 - Professor
 - Distinguished Professor
 - (ii) Do NOT collect for the following titles:
 - Visiting Assistant Professor
 - Lecturer
 - Senior Lecturer
 - Teaching Assistant Professor
 - Teaching Associate Professor
 - Teaching Professor
 - Research Assistant Professor
 - Research Associate Professor
 - Research Full Professor
 - Emeritus Professor
 - Affiliated Faculty
 - Professor of the Practice
 - Assistant Professor, Lecturer
 - Adjunct Faculty
 - (b) Check the faculty is a professor in Computer Science.
 - (i) Do not include affiliated faculty that are not directly Computer Science Professors.
 - (c) Copy the faculty member's name.
 - (d) Check if the person is already included in the spreadsheet
 - (i) Command + F the name with different variations:
 - (A) If middle initial:
 - With the middle initial and without.
 - (ii) If they are included but listed at a different institution:
 - (A) Google search [Firstname lastname DifferentInstitution]. Look for a personal homepage or department page
 - (B) In the following order, check if:
 - The person indicates they moved to the new institution
 - CV includes previous appointments, including both their appointment at new institution and different institution.
 - Photo included matches the photo on the new institution faculty page.
 - If any of the above checks passes: Add the current institution in the “UpdateInstitution” column.

- (C) Include their title in the “title” column in the spreadsheet.
- (D) If the title is XXX name in front of any of the titles listed in step ai, mark “Yes” in the “Named Position” column.
- (iii) If they are not included:
- (A) Add a row to the spreadsheet (Double click, add row below, to group together with colleagues)
 - (B) Do NOT add anything in the “CensusPersonID” column. This is left intentionally blank.
 - (C) Paste their name into the “name” column. Note: it’s important to use copy/paste here and below to avoid spelling errors.
 - (D) Copy and paste the person’s email address into the email column
 - If this is not easily accessible / not listed on the site, leave the column blank.
 - (E) Copy and paste their current Institution name in the corresponding column.
 - (F) Copy and paste their PhD institution in the PhD Institution column
 - (G) Include their title in the “title” column in the spreadsheet.
 - (H) If the title is XXX name in front of any of the titles listed in step ai, mark “Yes” in the “Name Positiiton” column.
 - (I) If any of the above are not included on the main department page, check the faculty member’s homepage / personal website.
- (e) Perceived race and gender:
- (i) Find a photo of the person. Check if a photo is included on the website found in step 1.
 - (A) If the department page does not include a photo, try other sites (but make sure you’re getting the right person). Add the website where the photo is found to the spreadsheet in the “notes” column. In order of preference the websites and queries (indicated by square brackets, where university, first-name, etc are substituted by the relevant information for the person) to try are:
 - Firstname lastname CurrentUniversity
 - Firstname lastname linkedin
 - Double check it’s the right person by checking the university affiliation
 - Firstname lastname PhDInstitution
 - (ii) Perceived gender. Use your impression of the person based on their photo to fill in the perceived gender column. Options are:
 - Man
 - Woman
 - Non-binary/Uncertain
 - No photo found
- (iii) Perceived race. Use your impression of the person based on their photo to fill in the perceived race column. Options are:
- White
 - Black
 - Latinx
 - East Asian
 - Southeast Asian
 - Indian / Indian subcontinent
 - Middle Eastern / North African
 - Native American / other Indigenous
 - Native Hawaiian or Other Pacific Islander
 - Multiracial or unsure
 - No photo found
- (iv) Code pronouns. In the website you found in step 2 or c1, see if the individual refers to themselves in the third person or otherwise indicates their own pronouns.
- (A) If a bio is not included in the website you found in step b or c1, but a personal website is linked, see if the individual refers to themselves in third person in their personal website.
 - (B) Do NOT google search for additional/external articles not written by the faculty member.
 - (C) Code pronouns those in the pronouns column. Options are:
 - he/him
 - she/her
 - they/them
 - he/them or they/he
 - she/them or they/she
 - Neo pronouns or other
 - No pronouns found
 - (v) Include your name in the “perceived by” column.
 - (vi) (OPTIONAL) Add any notes about the data as you enter it.
 - (vii) If the faculty member is noted as deceased/retired (or Professor Emeritus) on the website you found in step 2 or c1, mark “yes” in the “in the original dataset but not to collect” column.
 - (viii) If faculty member is noted as left academia / no longer a Professor in CS, mark “yes” in the “in the original dataset but not to collect” column.
- In addition to the above procedure, coders were given the following procedure to match faculty to DBLP IDs.
- Before the next step:
 - (1) First Name initial , Last Name, and institution match with CSrankings
 - (2) First Name initial , Last Name match with CSrankings
 - (3) First Name initial , Last Name match with DBLP
 - Only run for the professors whose matches could not be found automatically using csrankings or dblp.xml.

- (1) (only do once) Create a google spreadsheet with the following fields:
 - name
 - institution
 - website
 - dblp_id
- (2) Copy and paste the professor name and the institution name into Google
 - (a) Go to their institution website
 - (b) Add the institution website to your spreadsheet
 - (c) Find their publications list
 - (d) Copy and paste a publication title into a search query for dblp.xml
 - (i) (only do once) download dblp.xml from <https://dblp.org/>
 - (ii) in terminal, open the dblp.xml using: less dblp.xml
 - (iii) within less, to open a search query type: /
 - (iv) paste the publication name after you type / and press enter
 - (v) wait a long time
 - (e) Copy and paste the author name that matches the person you're looking for from that publication list of authors - this is the professor's dblp id - into the google spreadsheet
 - (f) If multiple publication names are not found in the dblp.xml for a professor, alternatively, go to dblp.org, check for the professor's name (verify it is the right professor using one of their publications on their website) or for one of their publications, and copy the dblp id into the google spreadsheet.
 - (g) If publication are not linked in their faculty homepage, check their google scholars webpage (the one verified with the university we have)

A.2 Amending name-based inference categories

The below shows how race categories from the name-based inference methods we employed were matched to our expanded U.S. census race categories.

EthnicolrWiki:

- White: GreaterEuropean (British), GreaterEuropean (Eastern European), GreaterEuropean (Jewish), GreaterEuropean (WestEuropean, French), GreaterEuropean (WestEuropean, Germanic), GreaterEuropean (WestEuropean, Italian), GreaterEuropean (WestEuropean, Nordic)
- East Asian: Asian (GreaterEastAsian, EastAsian), Asian (GreaterEastAsian, Japanese)
- South Asian (Indian / Indian subcontinent): Asian (Indian SubContinent)
- Middle Eastern / North African: GreaterAfrican (Muslim)
- Latinx: GreaterEuropean (WestEuropean, Hispanic)
- Black: GreaterAfrican (Africans)
- Native Hawaiian or Pacific Islander: None
- Southeast Asian: None

Ethnicseer:

- White: eng, ger, frn, ita, rus
- East Asian: chi, jap, kor
- South Asian (Indian / Indian subcontinent): ind
- Middle Eastern / North African: mea
- Latinx: spa
- Black: None
- Native Hawaiian or Pacific Islander: None
- Southeast Asian: vie

Ethnea:

- White: ENGLISH, GERMAN, FRENCH, ITALIAN, SLAV, NORDIC, DUTCH, HUNGARIAN, ROMANIAN, ISRAELI
- East Asian: CHINESE, JAPANESE, KOREAN, MONGOLIAN
- South Asian (Indian / Indian subcontinent): INDIAN
- Middle Eastern / North African: ARAB, TURKISH, GREEK
- Latinx: HISPANIC, CARIBBEAN
- Black: AFRICAN
- Native Hawaiian or Pacific Islander: POLYNESIAN
- Southeast Asian: VIETNAMESE, THAI, INDONESIAN

A.3 Survey text

This survey asks you about your demographics and parental educational history and takes about 1 minute to complete.

The data will be used to inform an analysis of the influence of faculty demographics and socioeconomic status on advancement, productivity, and coauthorship patterns. The goal of this research is to support equity by better understanding how demographics relate to and shape faculty experiences and connections such as coauthorship relationships in faculty networks, with the broader aim of promoting equal access to information for all groups, especially those currently underrepresented within computer science.

Your responses will be used for the purposes of this study and the data may also be used for future research also focusing on demographics and socioeconomic status and faculty coauthorship networks. Such research could occur indefinitely in the future, and individuals will not be alerted to each new publication.

The study "Demographics and Faculty Co-Authorship Networks" (IRB approved) is carried out by researchers from Haverford College. This study is supported by the National Science Foundation (NSF IIS-1955321).

Your participation is voluntary and you may stop at any time. Your information will be handled confidentially, but can not be truly anonymous given its association with a coauthorship network, and there is a risk of data leakage. The specific data stored about you indefinitely will include: name, title, email, institution, PhD institution, gender, race, parental degree attainment, parental faculty status, and publications with coauthorship information. The self-reported information from this survey will not be released publicly; any published information will be based on aggregated data.

If you have further questions about the research or your rights as a research participant, please contact Prof. Sorelle Friedler via sorelle@cs.haverford.edu. You may also address any concerns to the chairperson of Haverford College's IRB (a committee with oversight over human subject research) via hc-irb@haverford.edu.

To proceed, please check the box next to the following statement:

- I have read and understood the consent form and give my permission to participate in this study.
- (1) Please provide the full name under which you publish academic work.
 - (2) Are you currently employed as a tenured or tenure-track faculty member in a department that grants PhDs in computer science?
 - Yes
 - No
 - (3) How do you identify your gender?
 - Male
 - Female
 - Non-Binary
 - Other
 - Free-form text box
 - (4) How do you identify your race / ethnicity? (Select one or more)
 - White
 - Black
 - Latinx
 - East Asian
 - Southeast Asian
 - South Asian (Indian / Indian subcontinent)
 - Middle Eastern / North African
 - Native American or other Indigenous
 - Native Hawaiian or Pacific Islander
 - Other
 - Free-form text box
 - (5) What is the highest educational attainment of any of your parents?
 - Some high school
 - High school
 - Some college
 - College degree
 - Masters degree
 - Doctorate (professional, e.g., JD, MD)
 - Doctorate (research, e.g., Ph.D.)
 - (6) Have any of your parents ever been a tenured or tenure-track professor at a PhD-granting institution?
 - Yes
 - No

A.4 Race Meta-Labeling

In the brute force method discussed in the main text, 137 of the race orderings tied with an accuracy of 733/797 (92%) survey respondents labeled correctly. Of these winning algorithms, we selected Algorithm 2 to apply to the entire data set of 5,670 computer science faculty members. Every name which was not classified automatically with name-based inference was classified using perception. The ordering of race categories, thresholds for each race and name-based inference method applied in Algorithm 2 are given in Table 1.

The brute force method described in the text was corroborated by comparison to an exhaustive search for the best algorithm. This methodology at every step checks the accuracy of every name-based inference tool (checking all ethnicities and all thresholds) and chooses the classification which results in the least false positives.

Race	Order	Ethnea	EthnicolrWiki	EthnicolrFlorida
Southeast Asian	1st	0.90	-	-
South Asian	2nd	0.67	0.57	-
Jewish	3rd	0.45	0.61	-
English	4th	0.99	0.91	-
East Asian	5th	0.69	0.66	-
White	6th	0.74	0.74	0.97
MidEast/NorAfr	7th	0.99	0.77	-
Latinx	8th	0.95	0.72	0.90
Black	9th	-	-	0.77

Table 1: Thresholds and ordering of race categories in meta-labeling algorithm.

Using this method, 156 people can be classified with 0 false positives. 646 more people are classified automatically by allowing for either 0 or 1 false positives. Any further automatic classifications required accepting 2 or more false positives and performed worse than applying perception labels to the same names. Therefore, the remaining 113 people were classified with perception. This exhaustive search resulted in an overall accuracy of 734/797 (92%). Thus, the results of the algorithm found using the exhaustive search approach are comparable to Algorithm 2 found by brute force, indicating that we are achieving the maximum accuracy possible given our survey data set.

Algorithm 2 found by the brute force method described in the text is preferable to the algorithm found by exhaustive search because it requires classifying less people using perception and it is less likely to overfit the survey data. This is evident by the results of a 5-fold cross-validation. The mean training accuracy for 5 folds was 91.97% for the exhaustive search method and 92.28% for the brute force method, whereas their mean test accuracies were 89.59% and 92.22% respectively.

A.5 Gender Meta-Labeling

Algorithm 3 was used to label the gender of all 5,670 faculty in our census. The NonQuamGender thresholds used are given in Table 2.

Gender	East Asian	Order	NonQuamGender Threshold
Woman	Yes	1st	0.85
Man	Yes	2nd	0.85
Woman	No	3rd	0.75
Man	No	4th	0.75

Table 2: Thresholds for gender meta-labeling algorithm.

A.6 Comparisons with the Taulbee Reports

In Tables 3–6 below, the Taulbee report’s “other” column includes teaching professors, other instructors, researchers and postdocs.

A.7 University Prestige Rankings

Not every institution used in this study (see Appendix C) is ranked by every prestige measure:

Algorithm 2 Race Meta-Labeler

Input: NameMap: → (Ethnea → (raceLabel, probability), EthnicolrWiki → (raceLabel, probability), EthnicolrFlorida → (raceLabel, probability), perceptionLabel)
Output: ResultMap: name → metaLabel

```

1: for name ∈ NameMap do
2:   labeled ← False
3:   Ethnea, EthnicolrWiki, EthnicolrFlorida, perceptionLabel ← NameMap[name]
4:   if Ethnea[“RaceLabel”] == “Southeast Asian” and Ethnea[“p”] ≥ 0.90 then
5:     ResultMap[Name] ← “Southeast Asian”
6:     labeled ← True
7:   else if (Ethnea[“RaceLabel”] == “Indian / Indian subcontinent” and Ethnea[“p”] ≥ 0.67) then
8:     ResultMap[Name] ← “Indian / Indian subcontinent”
9:     labeled ← True
10:  else if (EthnicolrWiki[“RaceLabel”] == “Indian / Indian subcontinent” and EthnicolrWiki[“p”] ≥ 0.57) then
11:    ResultMap[Name] ← “Indian / Indian subcontinent”
12:    labeled ← True
13:  else if (Ethnea[“RaceLabel”] == “Jewish” and Ethnea[“p”] ≥ 0.45) then
14:    ResultMap[Name] ← “White”
15:    labeled ← True
16:  else if (EthnicolrWiki[“RaceLabel”] == “Jewish” and EthnicolrWiki[“p”] ≥ 0.61) then
17:    ResultMap[Name] ← “White”
18:    labeled ← True
19:  else if (Ethnea[“RaceLabel”] == “English” and Ethnea[“p”] ≥ 0.99) then
20:    ResultMap[Name] ← “White”
21:    labeled ← True
22:  else if (EthnicolrWiki[“RaceLabel”] == “English” and EthnicolrWiki[“p”] ≥ 0.91) then
23:    ResultMap[Name] ← “White”
24:    labeled ← True
25:  else if (Ethnea[“RaceLabel”] == “East Asian” and Ethnea[“p”] ≥ 0.69) then
26:    ResultMap[Name] ← “East Asian”
27:    labeled ← True
28:  else if (EthnicolrWiki[“RaceLabel”] == “East Asian” and EthnicolrWiki[“p”] ≥ 0.66) then
29:    ResultMap[Name] ← “East Asian”
30:    labeled ← True
31:  else if (Ethnea[“RaceLabel”] == “White” and Ethnea[“p”] ≥ 0.74) then
32:    ResultMap[Name] ← “White”
33:    labeled ← True
34:  else if (EthnicolrWiki[“RaceLabel”] == “English” and EthnicolrWiki[“p”] ≥ 0.74) then
35:    ResultMap[Name] ← “White”
36:    labeled ← True
37:  else if (EthnicolrFlorida[“RaceLabel”] == “White” and EthnicolrWiki[“p”] ≥ 0.97) then
38:    ResultMap[Name] ← “White”
39:    labeled ← True
40:  else if (Ethnea[“RaceLabel”] == “Middle Eastern / North African” and Ethnea[“p”] ≥ 0.99) then
41:    ResultMap[Name] ← “Middle Eastern / North African”
42:    labeled ← True
43:  else if (EthnicolrWiki[“RaceLabel”] == “Middle Eastern / North African” and EthnicolrWiki[“p”] ≥ 0.77) then
44:    ResultMap[Name] ← “Middle Eastern / North African”
45:    labeled ← True

```

Algorithm 2 Race Meta-Labeler

```

46: else if (Ethnea[“RaceLabel”] == “Latinx” and Ethnea[“p”]  $\geq 0.95$ ) then
47:     ResultMap[Name]  $\leftarrow$  “Latinx”
48:     labeled  $\leftarrow$  True
49: else if (EthniclrWiki[“RaceLabel”] == “Latinx” and EthniclrWiki[“p”]  $\geq 0.72$ ) then
50:     ResultMap[Name]  $\leftarrow$  “Latinx”
51:     labeled  $\leftarrow$  True
52: else if (EthniclrFlorida[“RaceLabel”] == “Latinx” and EthniclrFlorida[“p”]  $\geq 0.90$ ) then
53:     ResultMap[Name]  $\leftarrow$  “Latinx”
54:     labeled  $\leftarrow$  True
55: else if (EthniclrFlorida[“RaceLabel”] == “Black” and EthniclrFlorida[“p”]  $\geq 0.77$ ) then
56:     ResultMap[Name]  $\leftarrow$  “Black”
57:     labeled  $\leftarrow$  True
58: end if
59: if  $\neg$ labeled then
60:     ResultMap[Name]  $\leftarrow$  perceptionLabel
61: end if
62: end for

```

Algorithm 3 Gender Meta-Labeler

Input: NameMap: \rightarrow (NonQuam \rightarrow (genderLabel, probability), Ethnea \rightarrow (raceLabel, probability), perceptionLabel)
Output: ResultMap: name \rightarrow metaLabel

```

1: // Initialize all MetaLabels as None.
2: for name  $\in$  NameMap do
3:     labeled  $\leftarrow$  False
4:     NonQuam, Ethnea, perceptionLabel  $\leftarrow$  NameMap[name]
5:     if Ethnea[“RaceLabel”] == “East Asian” then
6:         if (NonQuam[“GenderLabel”] == “Female” and NonQuam[“p”]  $\geq 0.85$ ) then
7:             ResultMap[Name]  $\leftarrow$  “Woman”
8:             labeled  $\leftarrow$  True
9:         else if (NonQuam[“GenderLabel”] == “Male” and NonQuam[“p”]  $\geq 0.85$ ) then
10:            ResultMap[Name]  $\leftarrow$  “Man”
11:            labeled  $\leftarrow$  True
12:        end if
13:    else
14:        if (NonQuam[“GenderLabel”] == “Female” and NonQuam[“p”]  $\geq 0.75$ ) then
15:            ResultMap[Name]  $\leftarrow$  “Woman”
16:            labeled  $\leftarrow$  True
17:        else if (NonQuam[“GenderLabel”] == “Male” and NonQuam[“p”]  $\geq 0.75$ ) then
18:            ResultMap[Name]  $\leftarrow$  “Man”
19:            labeled  $\leftarrow$  True
20:        end if
21:    end if
22:    if  $\neg$ labeled then
23:        ResultMap[Name]  $\leftarrow$  perceptionLabel
24:    end if
25: end for

```

	Full	Associate	Assistant	Other	Total
Male	2049 (81%)	1108 (78%)	1257 (72%)	2207 (72%)	6621 (76%)
Female	455 (18%)	315 (22%)	501 (28%)	865 (28%)	2136 (24%)
Non-binary	9 (1%)	0 (0%)	3 (0%)	6 (0%)	18 (0%)
Unknown	215	84	119	185	603
Total	2728	1507	1880	3263	9378

Table 3: Gender of current US faculty as reported in the 2024 Taulbee report (Table F6).

	Full	Associate	Assistant	Distinguished	Total (n=5670)
Male	2043 (84%)	1112 (80%)	1217 (75%)	178 (82%)	4550 (80%)
Female	388 (16%)	273 (20%)	406 (25%)	40 (18%)	1107 (20%)
Non-binary	1	3	4	0	8
No photo	1	2	2	0	5
Total	2433	1390	1629	218	5670

Table 4: Gender of current US faculty in our data, replicating the Taulbee calculations.

	Full	Associate	Assistant	Other	Total
Nonresident Alien	56 (2%)	39 (3%)	269 (17%)	290 (10%)	654 (8%)
American Native	34 (2%)	5 (1%)	31 (2%)	12 (0%)	82 (1%)
Asian	735 (32%)	437 (34%)	616 (38%)	565 (20%)	2353 (32%)
Black	28 (1%)	29 (2%)	32 (2%)	79 (0%)	168 (2%)
Pacific Islander	0 (0%)	1 (0%)	2 (0%)	2 (0%)	5 (0%)
White	1335 (58%)	689 (53%)	538 (33%)	1582 (57%)	4144 (52%)
Multiracial	15 (2%)	11 (1%)	18 (1%)	12 (0%)	56 (0%)
Hispanic	52 (2%)	37 (3%)	50 (3%)	92 (3%)	231 (2%)
Unknown	57 (3%)	53 (4%)	64 (4%)	138 (5%)	312 (3%)
Total	2728	1507	1880	2772	9378

Table 5: Race of current US faculty as reported in the 2024 Taulbee report (Table F7).

	Full	Associate	Assistant	Distinguished	Total (n=5670)
South Asian	409 (17%)	177 (13%)	287 (18%)	44 (20%)	917 (16%)
East Asian	460 (19%)	349 (25%)	585 (36%)	32 (15%)	1426 (25%)
Southeast Asian	14 (1%)	5 (0%)	13 (1%)	1 (1%)	33 (1%)
Asian (all)	883 (36%)	531 (38%)	885 (54%)	77 (35%)	2376 (42%)
Black	11 (1%)	18 (2%)	30 (2%)	1 (1%)	60 (1%)
MidEast/NorAfr	143 (6%)	130 (10%)	162 (10%)	9 (4%)	444 (8%)
White	1320 (55%)	658 (47%)	491 (31%)	126 (57%)	2595 (46%)
Latinx	58 (2%)	38 (3%)	51 (3%)	4 (2%)	151 (3%)
Multiracial	13 (0%)	7 (0%)	8 (0%)	1 (1%)	29 (0%)
No photo	4 (0%)	6 (0%)	1 (0%)	0 (0%)	11 (0%)
Total	2433	1390	1629	218	5670

Table 6: Race of current US faculty in our data, replicating the Taulbee calculations.

- 11 universities were not found in CSrankings: University of Tennessee Chattanooga, Catholic University of America, Claremont Graduate University, Santa Clara University, Southern Illinois University-Carbondale, University of Toledo, University of Colorado Denver, University of Maine, University of Mississippi, University of Nevada, Wright State University.
- 2 universities were not found in USNWR: University of Tennessee Chattanooga, University of Wisconsin Milwaukee.

B Results

B.1 Network description

Figure 7 shows correlation between closeness centrality and university rank according to CSRankings and USNWR.

Figure 8 shows demographic and prestige disparities in degree and closeness centrality.

B.2 Interventions

Figure 9 shows the results of our basic intervention (Section 6.1) on the closeness centrality and predicted placement of faculty in our census when target and sponsor individuals are defined by their USNWR and CSRankings institutional rank.

Figure 10 shows the results of our Ph.D. intervention (Section 6.3) on the closeness centrality and predicted placement of Ph.D. scholars when target and sponsor individuals are defined by their USNWR institutional rank.

Figure 11 shows the results of our Ph.D. intervention (Section 6.3) on the closeness centrality and predicted placement of Ph.D. scholars when target and sponsor individuals are defined by their CSRankings institutional rank.

C Universities

The list of institutions used in this study:

University of Utah, Auburn University, Univ. of Illinois at Urbana-Champaign, University of South Carolina, University of South Florida, California Institute of Technology, Carnegie Mellon University, Indiana University, University of Toledo, New York University, Oakland University, Princeton University, Northwestern University, University of Massachusetts Amherst, University of New Mexico, Arizona State University, University of Texas at Austin, Virginia Commonwealth University, Iowa State University, New Jersey Institute of Technology, Florida State University, Texas A&M University, Boston University, Brandeis University, Brigham Young University, Brown University, Georgia Institute of Technology, University of Michigan, Univ. of California-Irvine, Univ. of California-Berkeley, Columbia University, Case Western Reserve University, George Mason University, Catholic University of America, Claremont Graduate University, Clarkson University, Clemson University, University of Florida, Florida Atlantic University, College of William and Mary, North Carolina State University, Colorado School of Mines, Colorado State University, Yale University, Cornell University, University of Texas at Dallas, Dartmouth College, DePaul University, Drexel University, University of Maryland-College Park, Duke University, Florida Institute of Technology, Florida International University, University of North Carolina-Charlotte, University of Missouri-Kansas City, Virginia Tech, State University of

- 7 universities in our data were not found in the hiring based prestige ranking: Oakland University, DePaul University, Michigan Technological University, New Mexico Tech, University of Colorado at Colorado Springs, University of Massachusetts Boston, University of Nebraska-Omaha.

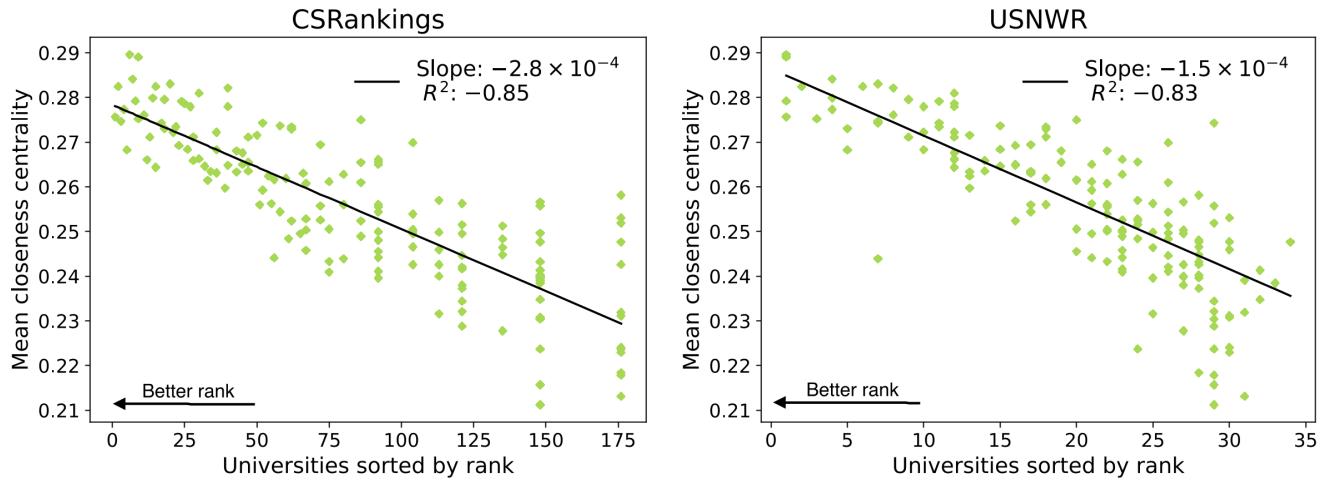


Figure 7: University prestige (rank, using the [50] measure of placement power; smaller score is more prestigious) versus the mean closeness centrality of faculty at that institution.

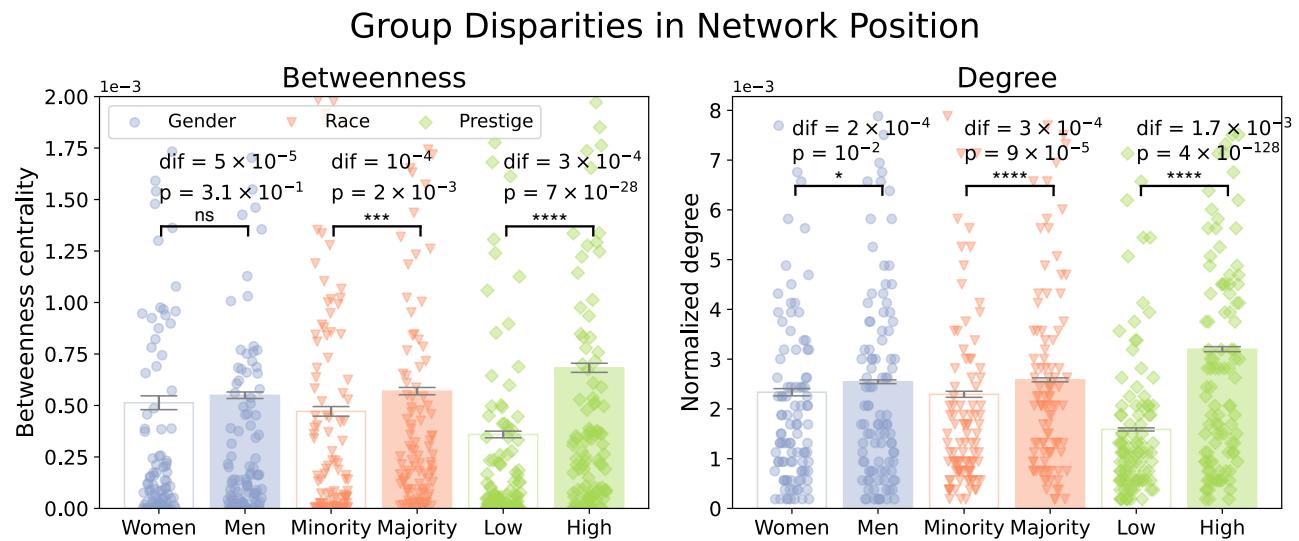


Figure 8: The p-values reported are from pooled t-tests comparing the betweenness centrality (left) and degree (right) of women and men; majority and minority race groups; and individuals currently employed at universities ranked above or below the median in our data. Standard error of the mean (SEM) error bars.

New York-Albany, George Washington University, University of Colorado Boulder, Georgia State University, University of Texas at San Antonio, Northeastern University, Univ. of California-Santa Barbara, Harvard University, Illinois Institute of Technology, University of Illinois at Chicago, Pennsylvania State University, University of Virginia, Binghamton University, Johns Hopkins University, Rice University, Kansas State University, Kent State University, Lehigh University, Louisiana State University, Missouri University of Technology, Michigan State University, Michigan Technological University, Western Michigan University, University of Cincinnati, Mississippi State University, Stonybrook University, Vanderbilt

University, Massachusetts Institute of Technology, Montana State University, University of North Texas, Naval Postgraduate School, University of Southern Mississippi, New Mexico Tech, New Mexico State University, Stanford University, North Dakota State University, Nova Southeastern University, Purdue University, Ohio State University, University of Georgia, Ohio University, Oklahoma State University, Old Dominion University, Oregon Health & Science University, Oregon State University, University of Wisconsin-Madison, Univ. of California-Santa Cruz, Portland State University, Rensselaer Polytechnic Institute, Rochester Institute of Technology, Rutgers

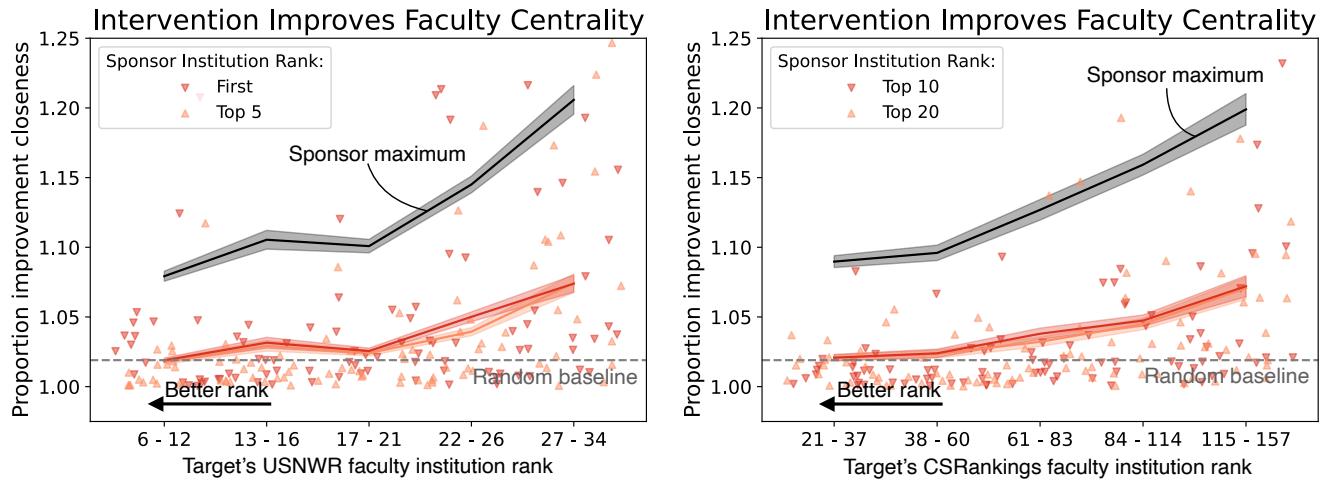


Figure 9: This simulated intervention results in increases in the closeness centrality of target individuals with minoritized race identities at low prestige current institutions (rank using Wapman et al. [2022] placement power). Our interventions (red lines) are compared to greedy selection of sponsor who would maximally improve target's closeness (black line) and to a random baseline.

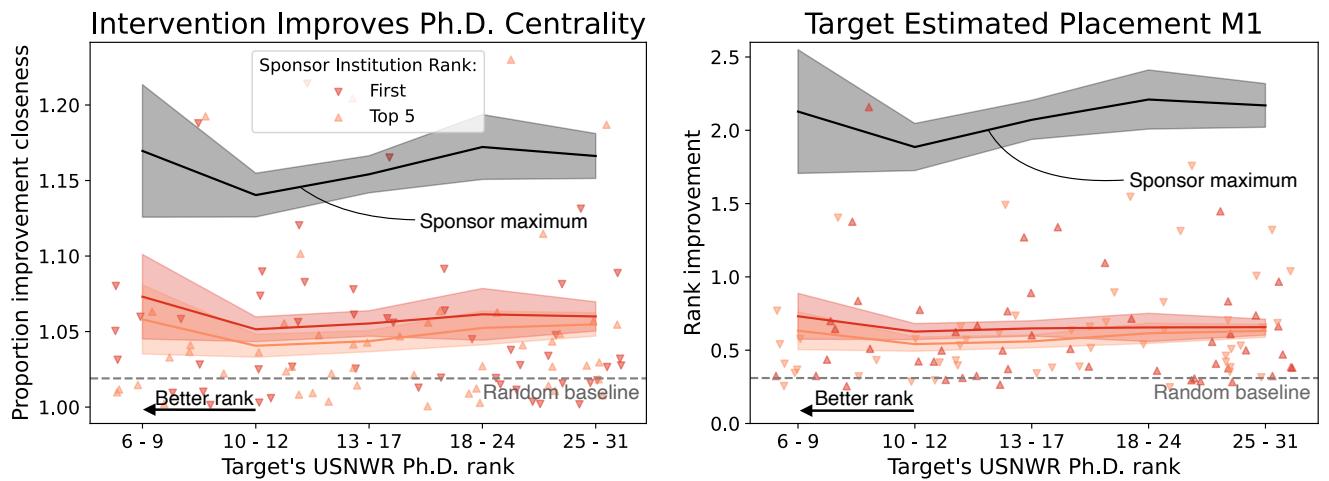


Figure 10: USNWR: Simulated collaboration increases the centrality of target individuals at low ranked institutions and improves the estimated rank of their placement institutions on the academic job market.

University, Santa Clara University, Southern Illinois University-Carbondale, Southern Methodist University, Univ. of California-San Diego, University of Washington, University at Buffalo, Stevens Institute of Technology, Syracuse University, Univ. of California-Riverside, Temple University, Texas Tech University, Tufts University, Univ. of California-Davis, University of Southern California, University of Chicago, Univ. of California-Los Angeles, University of Alabama-Birmingham, University of Tennessee-Chattanooga, University of Alabama, University of Alabama-Huntsville, University of Arizona, University of Arkansas, University of Arkansas-Little Rock, University of Central Florida, University of Colorado-Denver, University of Colorado at Colorado Springs, University of

Connecticut, University of Delaware, University of Denver, University of Hawaii at Manoa, University of Houston, University of Idaho-Moscow, University of Pennsylvania, University of Iowa, University of Kansas, University of Massachusetts Lowell, University of Kentucky, Washington University in St. Louis, University of Louisiana-Lafayette, University of Louisville, University of Massachusetts Boston, University of Maine, Univ. of Maryland-Baltimore County, University of Miami, University of Memphis, University of Minnesota, University of Mississippi, University of Missouri, University of Nebraska, University of Texas at Arlington, University of Nebraska-Omaha, University of Nevada Las Vegas, University of

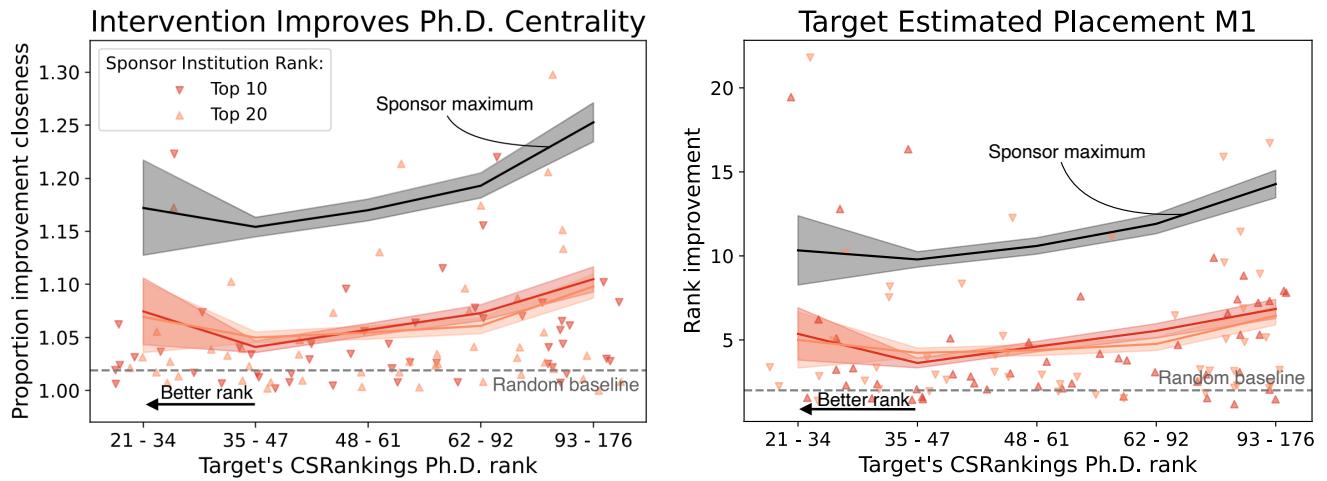


Figure 11: CSRankings: Simulated collaboration increases the centrality of target individuals at low ranked institutions and improves the estimated rank of their placement institutions on the academic job market.

Nevada, University of New Hampshire, University of North Carolina, Worcester Polytechnic Institute, University of Notre Dame, University of Oklahoma, University of Oregon, University of Pittsburgh, University of Rhode Island, University of Rochester, Wright

State University, University of Tennessee, University of Texas-El Paso, University of Tulsa, University of Wisconsin-Milwaukee, University of Wyoming, Utah State University, Washington State University, Wayne State University