Diversity in Academia *

Anjali Adukia
Elisa Xi Chen
Anna Gifty Opoku-Agyeman
Marieke Kleemans
Christian Valencia

December 11, 2025
Very preliminary. Please do not cite or distribute.

Abstract

Higher education institutions are centers of influence and knowledge production. Faculty produce this knowledge and determine the education imparted on students. But who are these individuals and do their backgrounds matter? In this study, we seek to understand both (1) the identities of these arbiters of higher education in terms of visible phenotypic features and (2) the relationship between these identities and faculty outcomes, including tenure and promotion, and the research they pursue. Specifically, we use computational tools from computer vision and natural language processing to extract information about the race, skin color, and gender of over 337,000 faculty members from 354 U.S. institutions. We document several stylized facts about representation and career progression in U.S. academia. Faculty are disproportionately male, with men accounting for roughly 64% of faculty, and predominantly White, with 65% identified as White. We document pronounced differences in skin tone by gender: female faculty are significantly lighter-skinned than their male counterparts, a pattern that holds within racial groups and persists across faculty rank and academic discipline. Examining raw differences in career progression, we find that non-White faculty have substantially lower tenure rates, while gender differences are more pronounced in promotion to full professorship.

*We thank Shivangi Ambardar, Daniel Barros De Aquino Fonseca, Matt Bonci, Paula Dastres Gallardo, Mish Mercer, and Aneesha Mishra for helping prepare the data. We thank Trishala Thakur for her help with the models. We thank Carlos Cavalcante and Zhou Chen for research assistance. This project greatly benefited from conversations with seminar participants at the University of Illinois at Urbana-Champaign. All errors are our own.

I Introduction

Higher education institutions are centers of influence and knowledge production. Faculty members are key agents of this knowledge production and diffusion (Blackwell, 1988; Kezar and Eckel, 2008; Morgan et al., 2022; Williams, 2023; Baker and Koedel, 2024). Specifically, faculty meaningfully influence instruction, curriculum development, mentorship, governance, public scholarship, and the intellectual climate across academic institutions and beyond (Antonio, 2002; Smith, 2024; Antman et al., 2025). But who are these individuals and do their backgrounds matter?

In this study, we seek to understand two primary questions: (1) What are the visible phenotypic identities of faculty members in research institutions of higher education, in terms of race, skin color, and gender? (2) What is the relationship between these identities and faculty outcomes, including tenure and promotion, and the research they ultimately pursue?

To conduct this analysis, we use individual faculty level data from the Academic Analytics Research Center (AARC) from 2011 to 2023. The data include 337,000 academics covering the universe of tenured and tenure-track faculty at 354 academic institutions across the United States. For each faculty member, employed between 2011 and 2023, we observe university and department of employment, academic rank (assistant, associate, full), discipline, the year and institution where they earned their doctorate, annual publication count, and total citation count. The dataset also contains information for all academic disciplines in the United States.

While we had personnel data on these faculty members, we needed to systematically collect information on demographics; however, individual-level data on demographics was not available in existing data sources. To systematically collect these data, we use computational tools from computer vision and natural language processing to extract information about the race, skin color, and gender using image feature classification models built using convolutional neural networks and surname-matching predictions.

Our analysis highlights persistent disparities in representation, appearance, and career progression among faculty in U.S. research institutions. Faculty remain predominantly White, and men constitute a majority across all racial groups. Beyond race and gender representation, we document robust differences in skin tone by gender: women faculty are consistently lighter-skinned than men, both on average and within the same racial groups. This suggest that skin tone captures a distinct dimension of identity beyond race alone. These patterns persist across academic rank and discipline, and across a range of image-based measures. We further document sizable disparities in career progression, with non-White faculty exhibiting substantially lower tenure rates than White faculty, and gender gaps emerging more strongly in promotion to full professorship. Together, these findings underscore the

importance of considering skin tone, alongside race and gender, in understanding inequality within U.S. academia.

Beyond documenting these trends, we examine whether scholars' demographic identities, including race, gender, and skin tone, systematically shape the topics and methodological approaches they pursue in their academic work. To do so, we implement a hybrid strategy. We use a multi-label XGBoost framework to explore nonlinear relationships and intersectional patterns in how identity relates to topic choice. By documenting how representation within academia structures the intellectual labor of scholarship, this study sheds light on the mechanisms through which epistemic inequality persists in academic knowledge production. The survival analysis we conduct provides insight into how promotion and retention differ across different demographic groups.

Faculty play a critical role in the production of higher education. Our findings highlight disproportionalities in patterns in the representation of race, skin color, and gender among faculty in the top research institutions in the U.S., which can have implications for institutional effectiveness and student outcomes (Blackwell, 1988; Morgan et al., 2022; Carrell and Kurlaender, 2023; Stansbury and Schultz, 2023; Kleemans and Thornton, 2023; Stansbury and Rodriguez, 2024; Abramitzky et al., 2024; Koffi et al., 2024; Ginther et al., 2025).

This study makes several contributions. First, we document previously unknown representation in the academy by showing the state and evolution of race, skin color, and gender across faculty members across 347 research institutions in the United States from 2011 to 2023.[1] After measuring the demographic trends within academia, we then explore the relationship between demographic identity and faculty tenure, promotion, and research scholarship. This study is among the first to provide demographic trends for faculty that include skin color as a measure of identity.[2] We also expand the literature on measurement of identities by drawing in image-based feature classification models and using different name-based algorithms to create scaleable and systematic measures for race and skin color, as applied in the academic context. Previous literature typically relies only on one specific name-based predictive algorithm that assigns either race or gender to an individual. Furthermore, we extend evidence on how faculty identity intersects with career progression within the academy and what implications faculty identity has for scholarship.

The rest of this preliminary working paper proceeds as follows. In Section II, we discuss related background. In Section III, we discuss the data we draw from. In Section IV, we discuss the methods we use to extract data from images and text. In Section V, we discuss our findings.

[1] We use the following racial categories: Asian, Black, Latine, and White.
[2] Exceptions are: Wapman et al. (2022) and Spoon et al. (2023).

II Background

Faculty serve as the central agents of knowledge production and diffusion within universities and colleges. Their responsibilities extend beyond instruction to include research, curriculum development, mentorship, institutional governance, and shaping the intellectual climate of their departments and the broader academic community (Antonio, 2002; Smith, 2009; Antman et al., 2025). Faculty influence the scope and content of educational programs, guide research priorities, contribute to the recruitment and retention of students, and collectively uphold academic standards. Through advising and mentorship, they assist students in their academic, professional, and personal development (Gurin et al., 2002).

Accordingly, the composition and activities of faculty play a decisive role in shaping institutional effectiveness and student outcomes (Blackwell, 1988; Morgan et al., 2022; Baker and Koedel, 2024). Faculty and administrative leadership in higher education encompasses roles such as department chairs, deans, provosts, and university presidents. The demographic composition and diversity of individuals in these positions have been shown to impact institutional priorities, governance, and responsiveness to societal needs (Kezar and Eckel, 2008; Williams, 2023).

Different policies have been adopted in US college campuses to improve diversity. For undergraduate student admissions, affirmative action policies have been placed for several years and evoked divided opinions (Arcidiacono and Lovenheim, 2016; Bleemer, 2022). Another policy implemented by US universities is the appointment of chief diversity officers, and while diversity increased over recent years, it cannot be attributed to the appointment of these authorities (Bradley et al., 2022). Ultimately, efficiency, performance and productivity are key to the production of research and teaching in academia. Hiring biases in either direction could prevent the efficient allocation of professors to faculty positions and dampen scientific progress.

Research indicates that leadership teams with varied backgrounds and experiences can improve decision-making due to a broader range of perspectives, enhance adaptability in changing environments, and promote policies that may be more inclusive of diverse stakeholders (Kim and Starks, 2016; Dugan, 2024; Bernile et al., 2018). Leadership diversity has also been associated with greater institutional legitimacy and the ability to mentor a wider set of emerging leaders, reflecting the constituencies served by the institution (Smith, 2009).

Diversity among faculty across racial, ethnic, gender, and disciplinary lines has been examined for its effects on teaching, research, and student experiences. Studies have found that faculty from diverse backgrounds tend to introduce varied perspectives into classroom discussions, enhance critical thinking, and broaden research agendas (Antonio, 2002; Umbach, 2006). Moreover, past evidence suggests that an educator's identity can impact student

performance and classroom management (Harlow, 2003; Dee, 2004; Edmonds, 2022). For students, exposure to a diverse faculty has been linked to increased academic engagement, more inclusive curricula, and better preparation for participation in increasingly multifaceted workplaces and societies (Gurin et al., 2002; Egalite and Kisida, 2015). Research also suggests that retaining a diverse faculty enables institutions to address a wider array of societal and scholarly topics, broadening institutional impact (Jayakumar, 2008).

There is also substantial research suggesting that diversity can enhance performance. One hypothesis is that the presence of diverse individuals in groups can change group-dynamics and behavior of individual members. Bayer and Rouse (2016) document that opinions about policy in economics matters differentially by gender. Hoogendoorn et al. (2013) find that mixed gender groups display more mutual monitoring and produce better outcomes. A similar case has been made for ethnic diversity in academia. Freeman and Huang (2015) find that papers written by authors belonging to a different ethnicity have more citations and impact than papers where authors belong to the same ethnicity. On the other hand, Calder-Wang et al. (2021) studies performance of entrepreneurial teams in universities, and find that randomly assigned teams that balance on ethnicity/race have significantly worse performance than self-selected teams. They point that the process of how teams are formed is relevant towards improvement of performance. Documenting and understanding the state of diversity in academia is relevant to potential improvements in performance and productivity in the profession.

In universities, faculty composition is key for educational and professional outcomes. Carrell and Kurlaender (2023) find that targeted feedback from professors to underrepresented students can lead to gains in grades and spillovers to other classes. The interaction of students and faculty of underrepresented groups can lead to better educational outcomes. Fairlie et al. (2014) find that underrepresented minority students' dropout and grade gaps compared to majority students fall when taking classes with underrepresented minority instructors. For faculty, researchers continue to find disparities in representation across race, gender, and class have implications for career progression and recognition in research or otherwise (Bradley et al., 2022; Stansbury and Schultz, 2023; Abramitzky et al., 2024; Koffi et al., 2024; Stansbury and Rodriguez, 2024; Ginther et al., 2025).

The concept of role modeling is also central to social learning theory and is applicable in a wide variety of contexts, including educational settings. Role models can influence individuals' aspirations, self-concept, and persistence towards goals (Ladson-Billings, 1995; Dee, 2004; Edmonds, 2022; Riley, 2024). Evidence shows that when students encounter mentors or instructors whom they perceive as relatable or aspirational – due to shared backgrounds, interests, or experiences – they may experience increased motivation, engagement, and re-

silience (Ladson-Billings, 1995; Gurin et al., 2002; Gershenson et al., 2022). Role models can also provide critical guidance in navigating institutional structures and professional pathways , which can have meaningful impacts on student outcomes (Dynan and Rouse, 1997; Porter and Serra, 2020; Martini and Urueña, 2025).While professional qualifications and expertise form the principal basis of role models in academia, visible physical attributes such as skin color may also influence students' perceptions of belonging, the signaling of inclusivity, and the salience of diversity within academic environments (Porter and Serra, 2020; Haynes, 2021).

In the context of the United States, the societal impact of colorism or skin color bias cannot be overstated. Historically, differences in skin tone have informed treatment and perceptions of marginalized groups, especially darker skinned Black individuals (Keith and Herring, 1991; Bodenhorn, 2015; Reece, 2018). The ability to characterize different skin tones can begin at a young age and can impact how children view themselves within social hierarchies informed by racial identity Spencer (2010); Pauker et al. (2016). Present day colorism still persists evident in popular media and the labor market Harrison et al. (2008); Leath and Mims (2023). Today, lighter-skinned individuals earn higher incomes, work at more prestigious jobs, and face less labor market discrimination than their darker-skinned counterparts (Hersch, 2006; Hochschild and Weaver, 2007; Glenn, 2009; Marira and Mitra, 2013). Furthermore, sociologists have shown that lighter skinned people are perceived as more physical attractive and intelligent as compared to dark skinned individuals (Goldsmith et al., 2007; Glenn, 2009; Hannon, 2015; Monk Jr et al., 2021; Monk Jr, 2021). Globally, darker skinned individuals face more discrimination as compared to lighter-skinned people (Glenn, 2009).

These broader trends in colorism are evident in educational environments. For example, Hunter (2016) finds that colorism can influence teacher's expectations, peer status, familial involvement, and disciplinary practices in K-12 schools. Similar work in higher education suggests that for some students the presence of faculty who share visible characteristics such as skin color can enhance psychological safety, diminish feelings of marginalization, and affirm diverse identities within academic spaces (Smith, 2009; Stanley, 2006). Moreover, visible diversity can serve as a tangible indication of institutional commitment to openness and equity, which may affect both current and prospective students' perspectives on an institution (Umbach, 2006). There is also substantial research on the varieties of implicit and explicit factors that influence academic recruitment and promotion processes. Among these, social and professional networks, perceptions of competence and fit, and subjective evaluations can play important roles. Studies suggest that, in some settings, differences in skin color may intersect with these processes, influencing who is recruited or selected for

leadership and faculty positions (Turner, 2002; Joseph and Hirshfield, 2011). Such effects may arise through network homophily (the tendency for individuals to associate with similar others), implicit biases in the evaluation of merit and leadership potential, and cultural taxation, where faculty of color take on disproportionate mentoring or service responsibilities (Stanley, 2006; Joseph and Hirshfield, 2011). These dynamics can affect representation at both junior and senior levels, and have implications for career advancement and retention within academia.

The persistence of colorism in who is deemed attractive also shapes beauty standards, representation in popular media, and the marriage market, especially for marginalized women (Sahay and Piran, 1997; Egbeyemi, 2019). In the Black community, Black women across skin tone see clear social benefits to having lighter skin in terms of their own desirability (Harper and Choma, 2019). When surveyed, Southeast Asian dark-skinned women in Canada reported that they desired lighter skin and light-skinned women among this group (Sahay and Piran, 1997). These observations are consistent with Hamilton et al. (2009) showing that in marriage markets where there are less "marriageable" men, lighter-skinned women are preferred over darker-skinned women. This finding is consistent when you consider the intersection of beauty, desirability, and colorism. Multiple studies also discuss how colorism can lead individuals to report an overall dissatisfaction with their skin color and a desire to change their skin tone through skin bleaching and lightening practices (Keith et al., 2010; Bond and Cash, 1992).

Current evidence suggests that beauty and attractiveness are associated with better job outcomes, higher earnings, and favorable performance evaluations in a variety of professional settings, including academia (Hale et al., 2023; Alkusari et al., 2024; Gründler et al., 2024). In the academy, there is evidence that a one standard deviation increase faculty-rated attractiveness is associated with an increased chance of women holding an academic position (Alkusari et al., 2024). Pairing these findings with how colorism also shapes beauty standards, especially for women, may explain how skintone trends evolve within academic disciplines and across the academy. Thus, our paper's contribution to the literature is novel in that it is the first to document how colorism and race shape faculty representation within higher education in the United States, measuring how skin tone varies across time along the dimensions of academic rank, prestige, and research productivity.

III Data

Faculty Information. Our faculty level data come from the Academic Analytics Research Center (AARC). AARC provides information of all tenured and tenure-track faculty members at 354 higher education institutions in the United States from 2011 to 2023 covering 111

different fields. The data includes information of 337,395 professors with annual data on their university and department of employment, position rank (assistant, associate, or full professor), publications, citations, grants and, institution and year they received their Ph.D. degree. AARC also provides a classification of academic fields.

Documenting race and ethnicity poses a challenge, as such information is typically not directly available. To overcome this challenge, we collected digital images for faculty members using a targeted Google Image search. The methodology section provides detailed information on how we collected the images.

To supplement information of academic institutions, we collect online data of university rankings using three different sources: QS rankings, US News, and Times. In addition, we use the Carnegie Classification of Institutions of Higher Education that classifies universities as high research institutions or less intensive research institutions. We also collect information of university location from online sources. We collect city, state (if applicable), and country of each institution that granted a PhD degree to the individuals in our sample.

Naturally, our data are limited to including individuals once they become faculty. Our findings do not account for the selection into becoming a faculty member in the first place.

IV Methodology

IV.A Race and Gender Predictions

Race and gender are not directly observed in faculty directories, so we constructed these measures from publicly available digital images. First, we collected faculty headshots using a targeted Google Image search. Specifically, we queried "faculty name + institution name + discipline" to ensure that the top search results corresponded to the intended individual. We manually verified the results.

We then implemented a two-step classification pipeline combining (i) FairFace, a state-of-the-art computer vision model for race and gender prediction, with (ii) Ethnicolr, which infers race/ethnicity probabilistically using name-based models. These outputs were combined to generate consistent race and gender predictions.

FairFace's race and gender classification model is trained on 108,501 facial images from the YFCC-100M Flickr dataset and is currently the state-of-the-art off-the-shelf model with the most balanced training distribution across racial groups (Karkkainen and Joo, 2021). It achieves an in sample overall accuracy of 95.7% for gender prediction and 80.0% for race prediction, using seven race categories: White, Black, Indian, East Asian, Southeast Asian, Middle Eastern, and Latino (Karkkainen and Joo, 2021). To assess performance in our context, we validated a subsample of predicted race assignments using self-reported race data from Texas College Salaries. The TX self-reported race file includes the categories "White,"

"Hispanic or Latino," "Black or African American," "Asian," "American Indian or Alaska Native," "Native Hawaiian or Other Pacific Islander," "Two or More Races," "Not Provided," and missing values. Our image-based predictions aligned extremely well for Black and White faculty, but performed less accurately for Asian and Hispanic individuals. To improve classification for these groups, we supplemented image-based predictions with Ethnicolr, a name-based algorithm trained on U.S. Census surnames, Florida voter registration records, and Wikipedia data. Incorporating Ethnicolr substantially increased accuracy for Asian and Hispanic faculty when benchmarked against Texas self-reported race. Because the Texas dataset does not include a Middle Eastern category, we cannot validate predictions for this group; therefore, we code individuals predicted as Middle Eastern as "Unknown" in our final dataset.

IV.B Skin Color Predictions

There are multiple approaches to predicting skin color from images. Pixel-based methods use basic color features, like average RGB values, to classify skin tone. While easy to compute, these methods struggle with differences in lighting, shadows, and facial structure. In contrast, Convolutional Neural Networks (CNNs) learn useful features directly from raw images. By applying filters across small regions of the image, CNNs can pick up on patterns like texture, shading, and the spatial layout of facial features. This makes them more effective at handling variation in skin tone appearance caused by differences in lighting and background environment.

IV.B.1 Pixel-Based Skin Color Estimation

To construct a pixel-based measure of skin color, we follow procedures similar to those in Adukia et al. (2023). We use the Google MediaPipe Face Mesh model, which detects 478 three-dimensional facial landmarks. Using these landmarks, we segment the skin regions by isolating areas of the face containing skin and removing non-skin features such as the eyes, eyebrows, and mouth. We then construct convex hulls around the eyes, eyebrows, and mouth, render the pixels within these regions transparent, and compute the average RGB value of the remaining skin pixels. This average color is subsequently converted from RGB space into the perceptually uniform CIE Lab* color space, and we use the L^* channel as a continuous measure of skin lightness ranging from 0 (darkest) to 100 (lightest).

IV.B.2 CNN-Based Skin Color Estimation

We also trained a CNN model to predict skin color. The first step in this process is to assemble a labeled training dataset with skin tone annotations that resemble the characteristics of our target population: faculty at U.S. universities. To do so, we randomly sampled

2,000 images from the 86,744-image FairFace training dataset (Karkkainen and Joo, 2021). This dataset is well suited to our purposes because it is drawn from the Yahoo YFCC100M Flickr dataset and was explicitly curated to achieve balanced representation across major racial and ethnic groups in the United States (Karkkainen and Joo, 2021).

To label skin tones in the sampled images, we recruited six graduate research assistants from the University of Chicago and the University of Illinois at Urbana-Champaign. Each image was evaluated using three scales: (1) the New Immigrant Survey (NIS) Skin Color Scale, a 10-point scale widely used in the social sciences ranging from 1 (darkest) to 10 (light); see Figure 1); (2) the Fenty Beauty foundation scale, which provides 57 distinct skin tone shades from darkest to lightest (see Figure 2); and (3) a categorical "words" scale classifying individuals as "light," "medium," or "dark." Three raters labeled each image using the NIS and words scales, and the remaining three used the Fenty and words scales.

For the categorical words scale, we used the majority rule to determine the ground-truth label: if at least three raters agreed on a classification, we adopted that label as the definitive annotation. These labeled images formed the training set for our CNN model, which was designed to predict all three skin tone scales from raw image inputs using a shared multi-task architecture.

We implemented a multi-output convolutional neural network (CNN) in PyTorch to perform joint regression on the NIS, Fenty, and Words scales. The model consists of three convolutional layers with ReLU activations and 2×2 max-pooling. The number of feature maps increases across layers (32, 48, and 64) to allow the network to learn progressively more complex visual representations. The extracted feature maps are flattened and passed through a fully connected layer with 64 hidden units, followed by a dropout layer ($p = 0.3$), and a final output layer of three units corresponding to the three target scales. A sigmoid activation function is applied to constrain outputs to the $[0, 1]$ interval. Target values are normalized prior to training to ensure comparability across scales.

To improve generalization, we applied standard data augmentation techniques, including random horizontal flips, slight rotations ($\pm 5^\circ$), and mild brightness and contrast jitter. The data were randomly split into training (80%) and validation (20%) subsets. We trained the model using the Adam optimizer (learning rate = 0.001), mean squared error loss, and a batch size of 8. Training continued for up to 20 epochs, with early stopping triggered if validation loss failed to improve for five consecutive epochs. The best-performing model, based on validation loss, was saved for final evaluation.

V Patterns of Faculty Diversity

V.A Summary Statistics

Table 1 reports the summary statistics of the panel data from 2011 to 2023. The panel covers 354 universities with individual information on 337,395 tenured and tenure-track faculty members. The table presents the share of faculty by different demographic characteristics and its standard deviations. In terms of race, 65% of faculty are White, 16% Asian, 4% Latine/Hispanic, 3% Black, and 14% are unclassified in our data. Non white faculty add up to 35%, and underrepresented minority faculty (Black and Latino Hispanic) represent only 7% of total faculty share. In terms of gender, 36% of faculty are female.

Our data contains information on the rank of faculty for every year. In the panel data, 24% of the faculty are Assistant Professors, 30% Associate Professors, and 45% Full Professors. Importantly, we observe when faculty are promoted from one rank to the other, or whenever they are not in the panel dataset anymore. The share of faculty in STEM fields, defined by the discipline, is 44%. In our data we also observe publications for faculty members and citations. On average, a faculty in our sample has 37.3 publications, and the average number of citations per published article is 40.78.

V.B Empirical Results

We begin by documenting basic patterns of representation by race and gender. Figures 3 and 4 present the distribution of faculty across racial groups separately for men and women, in levels and percentages. Across all racial categories, men constitute a majority of faculty, while women remain underrepresented. At the same time, the racial composition of faculty differs by gender, with women representing a larger share of Black and Hispanic faculty than of White faculty. These patterns serve as a descriptive benchmark for understanding subsequent differences in skin tone and career progression.

We next turn to differences in skin tone by gender. Figure 5 shows the unconditional distribution of skin tint for male and female faculty, revealing a clear rightward shift for women, indicating lighter skin on average. Importantly, this pattern persists when conditioning on race. Figures 6 through 9 show skin-tone distributions by race and gender using multiple measures, including pixel-based tint, the New Immigrant Survey (NIS) scale, the Fenty Foundation Scale, and a categorical words-based measure. Across all racial groups and measurement approaches, female faculty are consistently lighter-skinned than their male counterparts. These differences are therefore not driven by a higher share of White women in the faculty population, but instead reflect within-race gender differences in skin tone.

The gender gradient in skin tone is remarkably stable across academic contexts. Figure 10 shows that women are lighter-skinned than men at every faculty rank while Figure 11

demonstrates that the same pattern holds in both STEM and non-STEM fields. Figure 12 further illustrates that these differences persist across academic disciplines. Taken together, these results point to skin tone as a distinct dimension along which gender differences manifest in academia, separate from race composition alone. The observed patterns resonate with a broader literature on colorism and gendered beauty standards, which documents that lighter skin is socially valued and particularly consequential for women across labor market and social settings.

Finally, we examine raw differences in career progression by race and gender. Figure 13 shows that non-White faculty have substantially lower tenure rates than White faculty across the career cycle, with gaps reaching up to seven percentage points at peak tenure ages. Gender differences in tenure rates are smaller on average, but become more pronounced at later career stages. As shown in Figure 14, promotion to full professorship exhibits sizable gender gaps, with men promoted at higher rates than women. Together, these results document persistent disparities in academic career outcomes by race and gender, alongside pronounced and robust gender differences in skin tone across the faculty population.

References

Abramitzky, R., Greska, L., Pérez, S., Price, J., Schwarz, C., and Waldinger, F. (2024). Climbing the ivory tower: How socio-economic background shapes academia. Technical report, National Bureau of Economic Research.

Adukia, A., Eble, A., Harrison, E., Runesha, H. B., and Szasz, T. (2023). What we teach about race and gender: Representation in images and text of children's books. *The Quarterly Journal of Economics*, 138(4):2225–2285.

Alkusari, H., Gupta, N. D., and Etcoff, N. (2024). In the eye of the promoter? How faculty ratings of attractiveness matter for junior academic careers. *Journal of Economic Behavior & Organization*, 220:645–659.

Antman, F. M., Qu, S., Weinberg, B. A., and Logan, T. D. (2025). The long-run impacts of mentoring underrepresented minority groups in economics. In *AEA Papers and Proceedings*, volume 115, pages 522–528. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Antonio, A. L. (2002). Faculty of color reconsidered: Reassessing contributions to scholarship. *The Journal of Higher Education*, 73(5):582–602.

Arcidiacono, P. and Lovenheim, M. (2016). Affirmative action and the quality–fit trade-off. *Journal of Economic Literature*, 54(1):3–51.

Baker, S. P. and Koedel, C. (2024). Diversity trends among faculty in STEM and non-STEM fields at selective public universities in the US from 2016 to 2023. *Humanities and Social Sciences Communications*, 11(1):1–9.

Bayer, A. and Rouse, C. E. (2016). Diversity in the economics profession: A new attack on an old problem. *Journal of Economic Perspectives*, 30(4):221–242.

Bernile, G., Bhagwat, V., and Yonker, S. (2018). Board diversity, firm risk, and corporate policies. *Journal of Financial Economics*, 127(3):588–612.

Blackwell, J. E. (1988). Faculty issues: The impact on minorities. *The Review of Higher Education*, 11(4):417–434.

Bleemer, Z. (2022). Affirmative action, mismatch, and economic mobility after California's Proposition 209. *The Quarterly Journal of Economics*, 137(1):115–160.

Bodenhorn, H. (2015). *The color factor: The economics of African-American well-being in the nineteenth-century South*. Oxford University Press.

Bond, S. and Cash, T. F. (1992). Black beauty: Skin color and body images among african-american college women. *Journal of Applied Social Psychology*, 22(11):874–888.

Bradley, S. W., Garven, J. R., Law, W. W., and West, J. E. (2022). The impact of chief diversity officers on diverse faculty hiring. *Southern Economic Journal*, 89(1):3–36.

Calder-Wang, S., Gompers, P. A., and Huang, K. (2021). Diversity and performance in entrepreneurial teams. Technical report, National Bureau of Economic Research.

Carrell, S. E. and Kurlaender, M. (2023). My professor cares: Experimental evidence on the role of faculty engagement. *American Economic Journal: Economic Policy*, 15(4):113–141.

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, 86(1):195–210.

Dugan, J. P. (2024). *Leadership Theory: Cultivating Critical Perspectives*. Jossey-Bass.

Dynan, K. E. and Rouse, C. E. (1997). The underrepresentation of women in economics: A study of undergraduate economics students. *The Journal of Economic Education*, 28(4):350–368.

Edmonds, L. (2022). Role models revisited: HBCUs, same-race teacher effects, and Black student achievement.

Egalite, A. J. and Kisida, B. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 62:59–69.

Egbeyemi, A. (2019). Shedding light on colorism: How the colonial fabrication of colorism impacts the lives of African American women. *Journal of Integrative Research & Reflection*, 2(2):14–25.

Fairlie, R. W., Hoffmann, F., and Oreopoulos, P. (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review*, 104(8):2567–2591.

Freeman, R. B. and Huang, W. (2015). Collaborating with people like me: Ethnic coauthorship within the united states. *Journal of Labor Economics*, 33(S1):S289–S318.

Gershenson, S., Hart, C. M., Hyman, J., Lindsay, C. A., and Papageorge, N. W. (2022). The long-run impacts of same-race teachers. *American Economic Journal: Economic Policy*, 14(4):300–342.

Ginther, D. K., Kahn, S., and Milakhina, D. (2025). Same as it ever was: Gender, race, and ethnicity differences in promotion for academic economists. In *AEA Papers and Proceedings*, volume 115, pages 195–201. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Glenn, E. N. (2009). *Shades of difference: Why skin color matters*. Stanford University Press.

Goldsmith, A. H., Hamilton, D., and Darity, W. (2007). From dark to light: Skin color and wages among African-Americans. *Journal of Human Resources*, 42(4):701–738.

Gründler, K., Potrafke, N., and Wochner, T. (2024). The beauty premium of politicians in office. *Journal of Economic Behavior & Organization*, 217:298–311.

Gurin, P., Dey, E. L., Hurtado, S., and Gurin, G. (2002). Diversity and higher education: Theory and impact on educational outcomes. *Harvard Educational Review*, 72(3):330–366.

Hale, G., Regev, T., and Rubinstein, Y. (2023). Do looks matter for an academic career in economics? *Journal of Economic Behavior & Organization*, 215:406–420.

Hamilton, D., Goldsmith, A. H., and Darity Jr, W. (2009). Shedding "light" on marriage: The influence of skin shade on marriage for black females. *Journal of Economic Behavior & Organization*, 72(1):30–50.

Hannon, L. (2015). White colorism. *Social Currents*, 2(1):13–21.

Harlow, R. (2003). " race doesn't matter, but...": The effect of race on professors' experiences and emotion management in the undergraduate college classroom. *Social psychology quarterly*, pages 348–363.

Harper, K. and Choma, B. L. (2019). Internalised white ideal, skin tone surveillance, and hair surveillance predict skin and hair dissatisfaction and skin bleaching among African American and Indian women. *Sex Roles*, 80(11):735–744.

Harrison, M. S., Reynolds-Dobbs, W., and Thomas, K. M. (2008). Skin color bias in the workplace: The media's role and implications toward preference. *Racism in the 21st century: An empirical analysis of skin color*, pages 47–62.

Haynes, A. E. (2021). *Colorism Experiences of Non-White Women Leaders in Higher Education*. Nova Southeastern University.

Hersch, J. (2006). Skin-tone effects among african americans: Perceptions and reality. *American Economic Review*, 96(2):251–255.

Hochschild, J. L. and Weaver, V. (2007). The skin color paradox and the American racial order. *Social forces*, 86(2):643–670.

Hoogendoorn, S., Landman, R., van Kooten, J., and Schreuder, M. (2013). Integrated network management amsterdam: control approach and test results. In *16th international ieee conference on intelligent transportation systems (itsc 2013)*, pages 474–479. IEEE.

Hunter, M. (2016). Colorism in the classroom: How skin tone stratifies African American and Latina/o students. *Theory Into Practice*, 55(1):54–61.

Jayakumar, U. M. (2008). Can higher education meet the needs of an increasingly diverse society? *Harvard Educational Review*, 78(4):615–651.

Joseph, T. D. and Hirshfield, L. E. (2011). 'why don't you get somebody new to do it?' race and cultural taxation in the academy. *Ethnic and Racial Studies*, 34(1):121–141.

Karkkainen, K. and Joo, J. (2021). Fairface: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pages 1548–1558.

Keith, V. M. and Herring, C. (1991). Skin tone and stratification in the Black community. *American journal of sociology*, 97(3):760–778.

Keith, V. M., Lincoln, K. D., Taylor, R. J., and Jackson, J. S. (2010). Discriminatory experiences and depressive symptoms among African American women: Do skin tone and mastery matter? *Sex Roles*, 62(1):48–59.

Kezar, A. and Eckel, P. (2008). Advancing diversity agendas on campus: Examining transactional and transformational presidential leadership styles. *International Journal of Leadership in Education*, 11(4):379–405.

Kim, D. and Starks, L. T. (2016). Gender diversity on corporate boards: Do women contribute unique skills? *American Economic Review*, 106(5):267–271.

Kleemans, M. and Thornton, R. (2023). Fully promoted: The distribution and determinants of full professorship in the economics profession. In *AEA Papers and Proceedings*, volume 113, pages 467–472. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Koffi, M., Pongou, R., and Wantchekon, L. (2024). The color of ideas: Racial dynamics and citations in economics. Technical report, National Bureau of Economic Research.

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 32(3):465–491.

Leath, S. and Mims, L. (2023). A qualitative exploration of Black women's familial socialization on controlling images of Black womanhood and the internalization of respectability politics. *Journal of Family Studies*, 29(2):774–791.

Marira, T. D. and Mitra, P. (2013). Colorism: Ubiquitous yet understudied. *Industrial and Organizational Psychology*, 6(1):103–107.

Martini, C. A. and Urueña, V. (2025). Can role models influence female students' decision to participate in the labor market? evidence from a field experiment. *Economic Development and Cultural Change*, 73(2):875–907.

Monk Jr, E. P. (2021). The unceasing significance of colorism: Skin tone stratification in the United States. *Daedalus*, 150(2):76–90.

Monk Jr, E. P., Esposito, M. H., and Lee, H. (2021). Beholding inequality: Race, gender, and returns to physical attractiveness in the united states. *American Journal of Sociology*, 127(1):194–241.

Morgan, A. C., LaBerge, N., Larremore, D. B., Galesic, M., Brand, J. E., and Clauset, A. (2022). Socioeconomic roots of academic faculty. *Nature human behaviour*, 6(12):1625–1633.

Pauker, K., Williams, A., and Steele, J. R. (2016). Children's racial categorization in context. *Child development perspectives*, 10(1):33–38.

Porter, C. and Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3):226–254.

Reece, R. L. (2018). Genesis of us colorism and skin tone stratification: Slavery, freedom, and mulatto-black occupational inequality in the late 19th century. *The Review of Black Political Economy*, 45(1):3–21.

Riley, E. (2024). Role models in movies: The impact of queen of katwe on students' educational attainment. *Review of Economics and Statistics*, 106(2):334–351.

Sahay, S. and Piran, N. (1997). Skin-color preferences and body satisfaction among south asian-canadian and european-canadian female university students. *The Journal of Social Psychology*, 137(2):161–171.

Smith, D. G. (2009). *Diversity's Promise for Higher Education: Making it Work*. Johns Hopkins University Press.

Smith, D. G. (2024). *Diversity's promise for higher education: Making it work*. Jhu Press.

Spencer, M. B. (2010). White and black children biased toward lighter skin. *CNN. Retrieved March*, 8:2020.

Spoon, K., LaBerge, N., Wapman, K. H., Zhang, S., Morgan, A. C., Galesic, M., Fosdick, B. K., Larremore, D. B., and Clauset, A. (2023). Gender and retention patterns among us faculty. *Science Advances*, 9(42):eadi2205.

Stanley, C. A. (2006). *Faculty of Color: Teaching in Predominantly White Colleges and Universities*, pages 3–24. Jossey-Bass.

Stansbury, A. and Rodriguez, K. (2024). The class gap in career progression: Evidence from us academia. Technical report, Mimeo.

Stansbury, A. and Schultz, R. (2023). The economics profession's socioeconomic diversity problem. *Journal of Economic Perspectives*, 37(4):207–230.

Turner, C. S. (2002). Barriers to the success of faculty of color. *The Journal of Higher Education*, 73(1):43–68.

Umbach, P. D. (2006). The role of faculty in college students' sociopolitical awareness. *Review of Higher Education*, 30(1):37–61.

Wapman, K. H., Zhang, S., Clauset, A., and Larremore, D. B. (2022). Quantifying hierarchy and dynamics in us faculty hiring and retention. *Nature*, 610(7930):120–127.

Williams, D. A. (2023). *Strategic diversity leadership: Activating change and transformation in higher education*. Routledge.

Tables and Figures

Figure 1: the New Immigrant Survey (NIS) Skin Color Scale

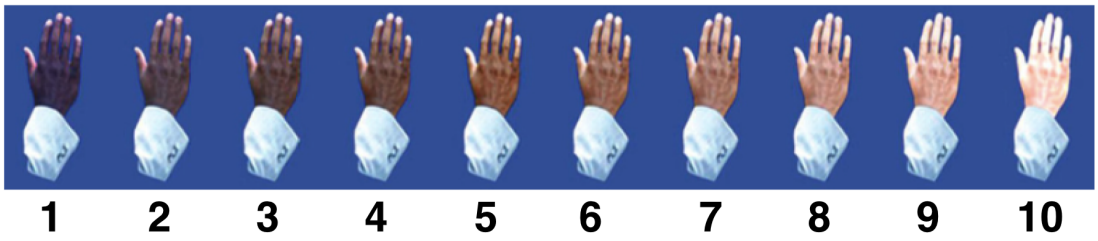


Figure 2: the Fenty Beauty Foundation Skin Color Scale

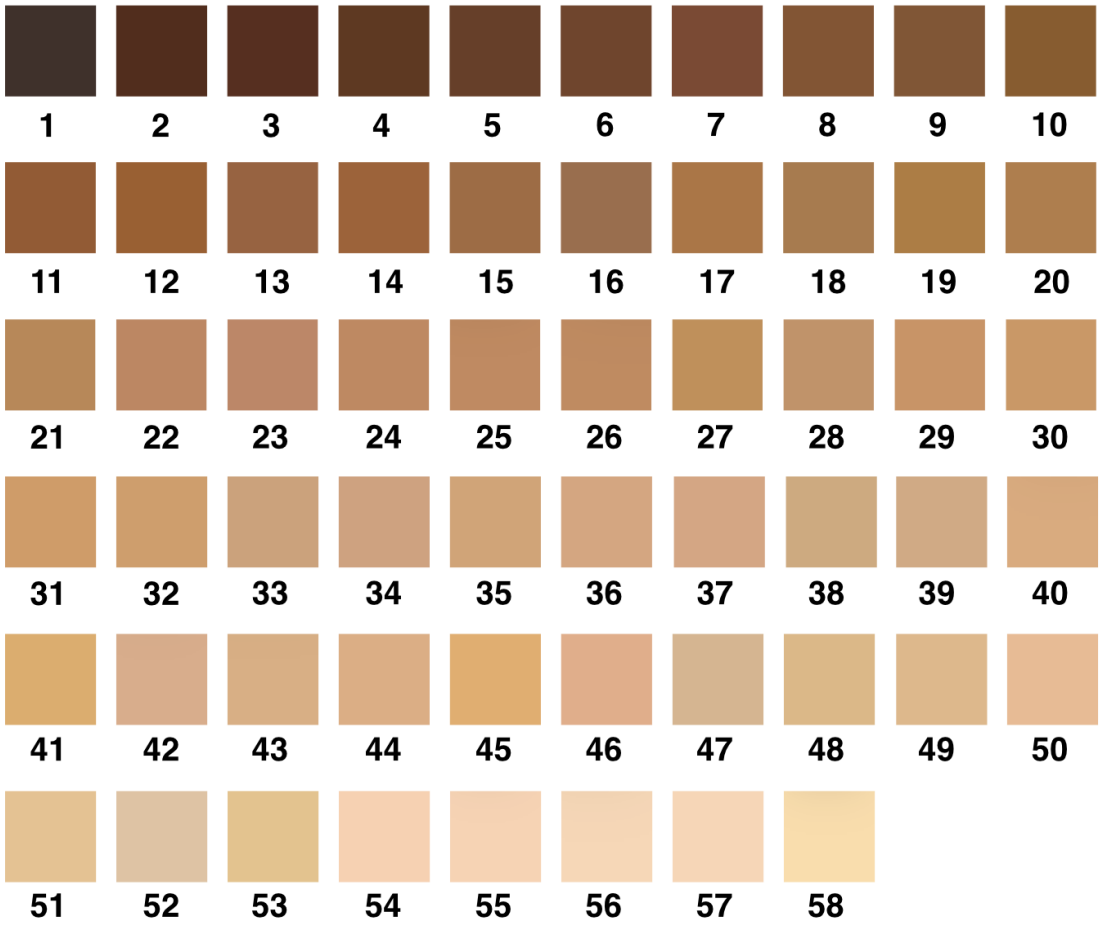


Table 1: Summary Statistics. Panel Data 2011-2023

Number of Universities	354	
Number of Faculty	337,395	
	Mean	Std. Dev
Share of non White Faculty	0.35	0.48
Share of Underrepresented Minority	0.07	0.26
Share of Asian Faculty	0.16	0.37
Share of Black Faculty	0.03	0.18
Share of Latino Hispanic Faculty	0.04	0.20
Share of White Faculty	0.65	0.48
Share of Unknown Race Faculty	0.14	0.35
Share of Female Faculty	0.36	0.48
Share of Assistant Professors	0.24	0.43
Share of Associate Professors	0.30	0.46
Share of Full Professors	0.45	0.50
Share of Faculty in STEM	0.44	0.50
Publications per Faculty Member	37.28	74.67
Citations per Published Article	40.78	184.66

Figure 3: Summary Statistics for Race and Gender

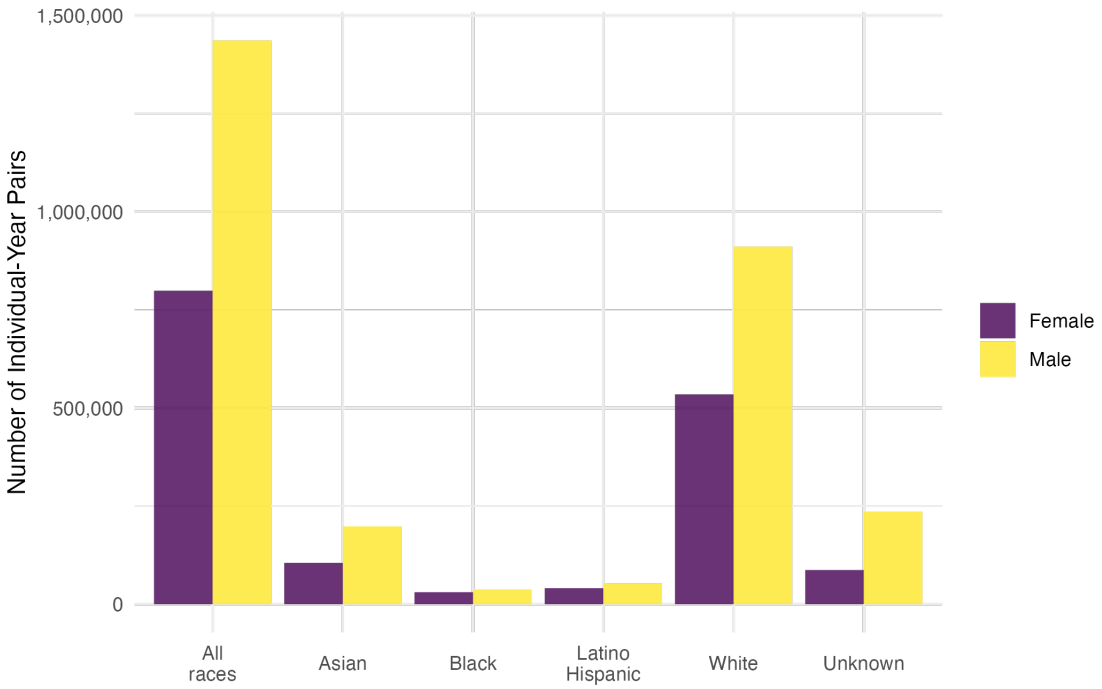


Figure 4: Summary Statistics for Race and Gender as a Percentage

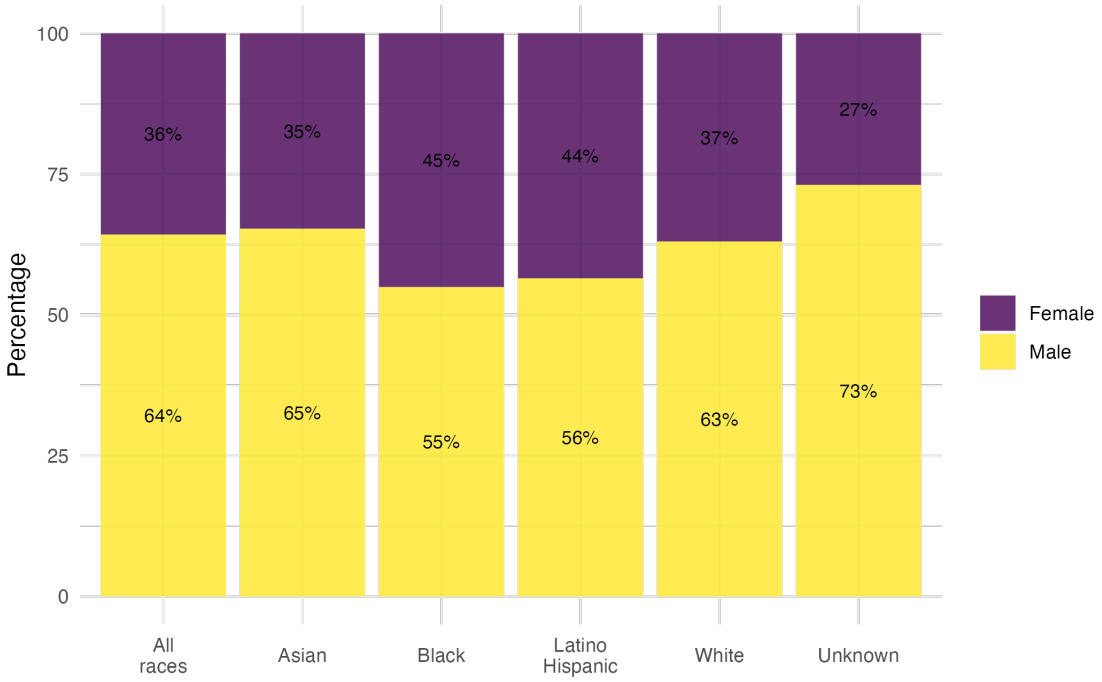


Figure 5: Density of Skin Tint for all Faculty by Gender

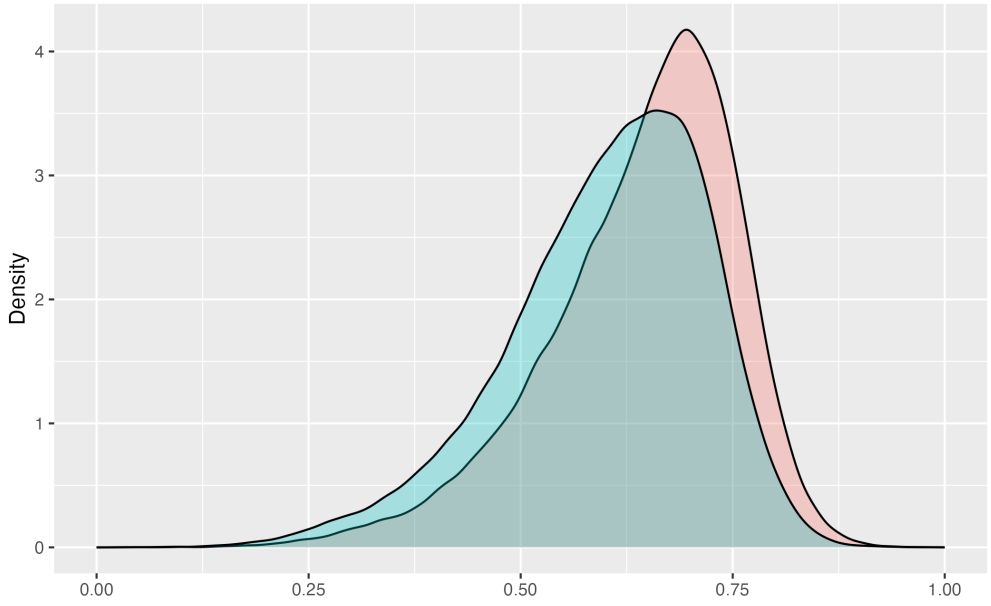


Figure 6: Density of Skin Tint: Race by Gender

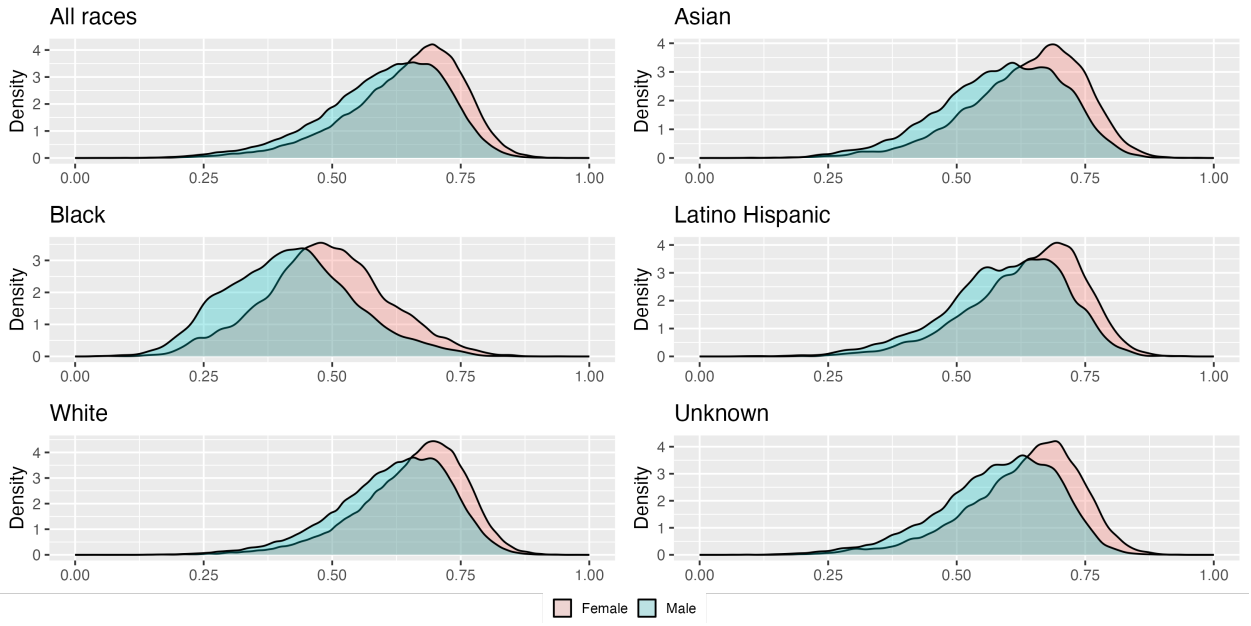


Figure 7: New Immigrant Survey (NIS) Scale: Race by Gender

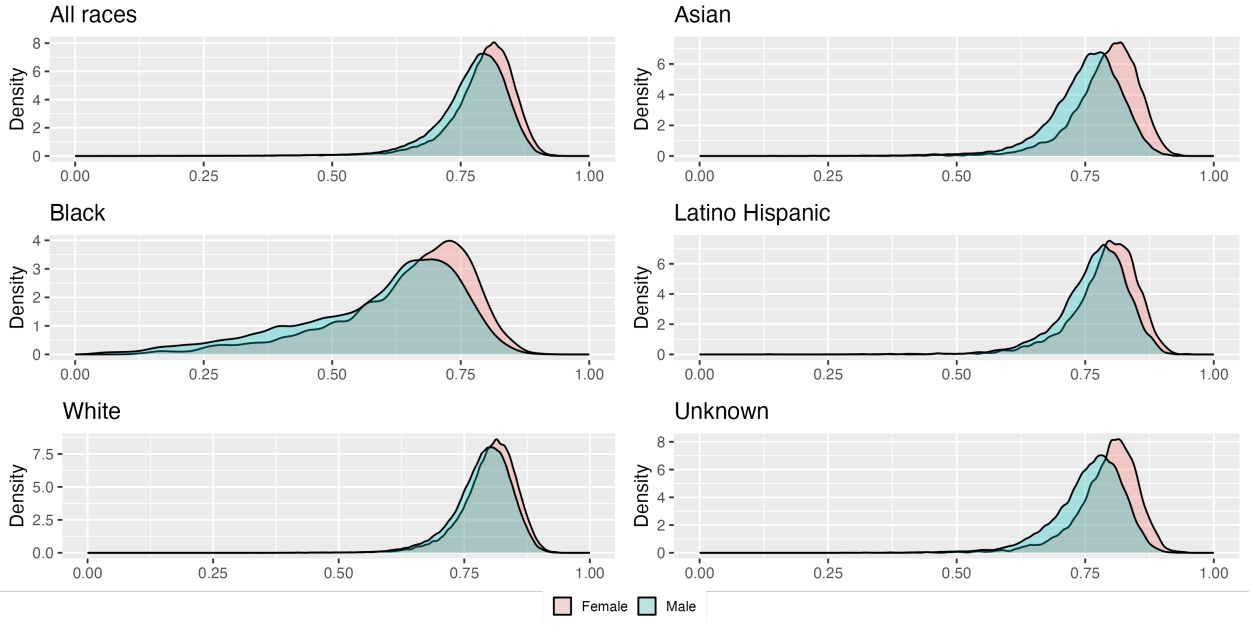


Figure 8: Fenty Scale: Race by Gender

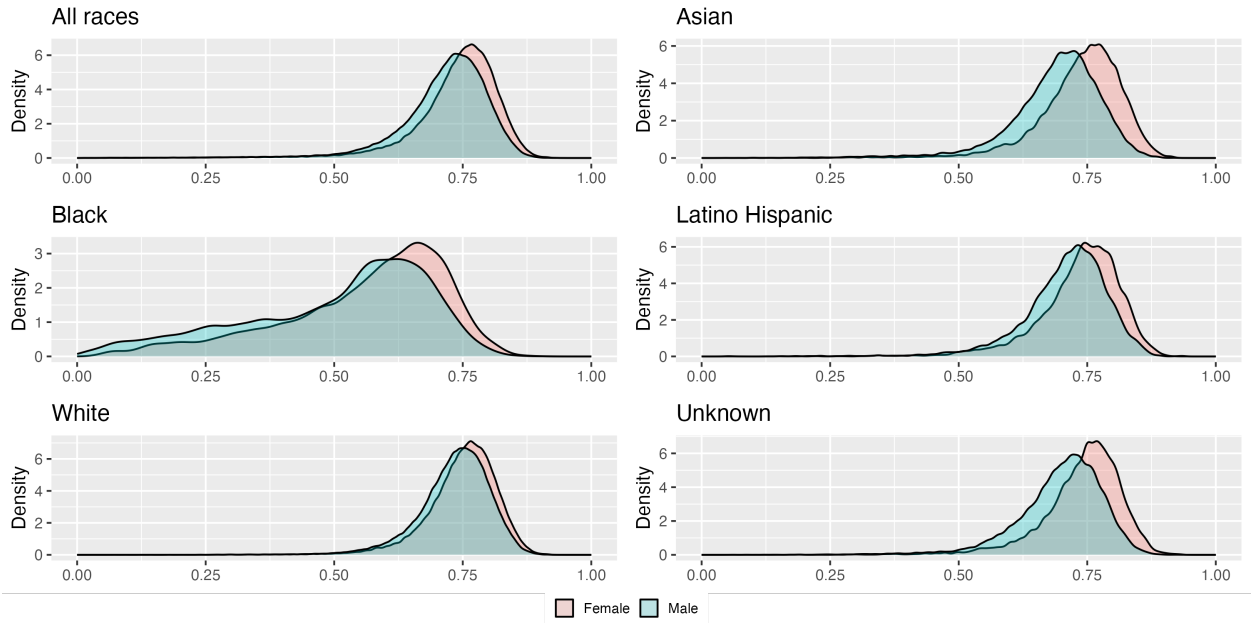


Figure 9: Word Scale: Race by Gender

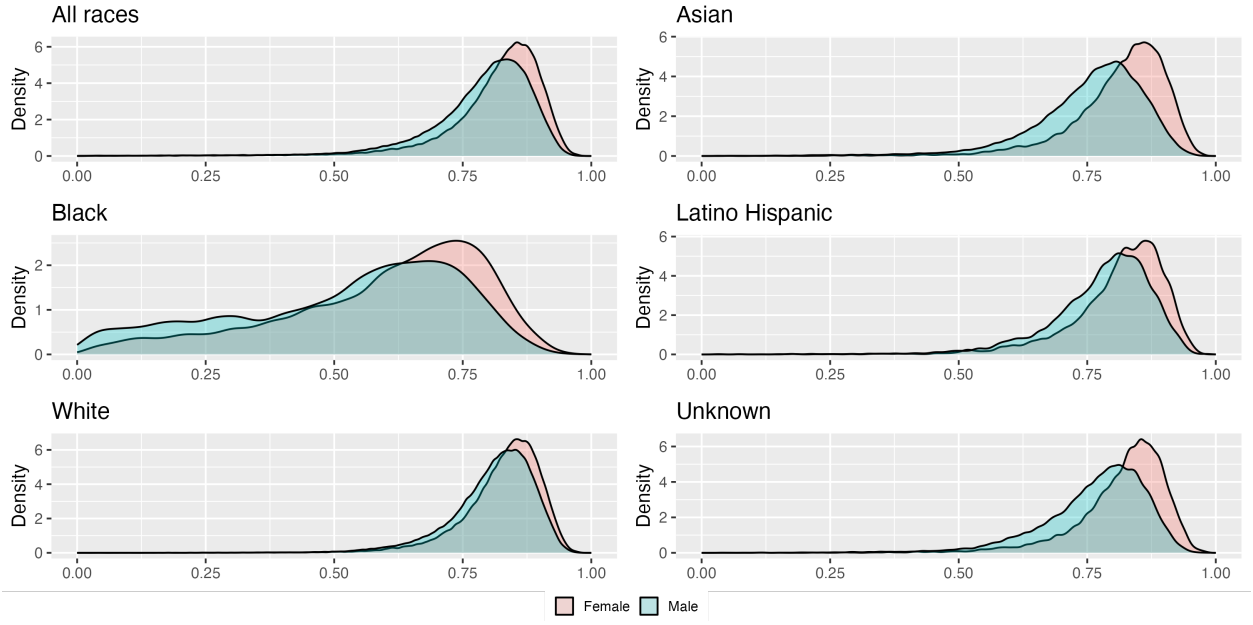


Figure 10: Skin Tint: Rank by Gender

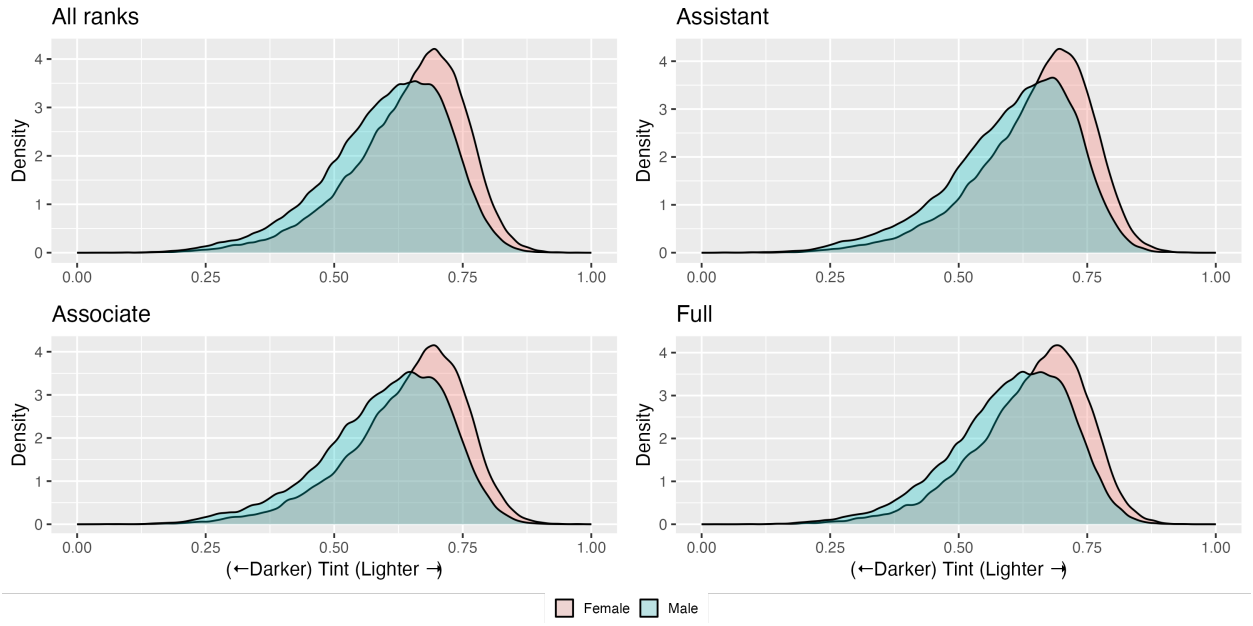


Figure 11: Skin Tint: STEM by Gender

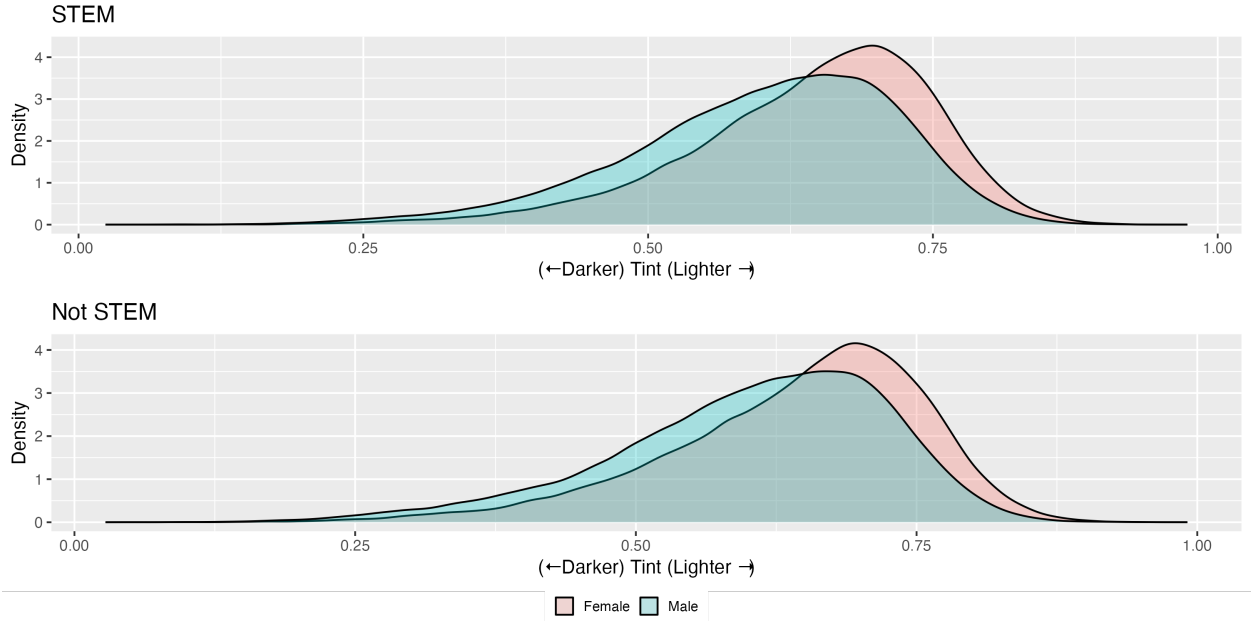


Figure 12: Skin Tint: Discipline by Gender

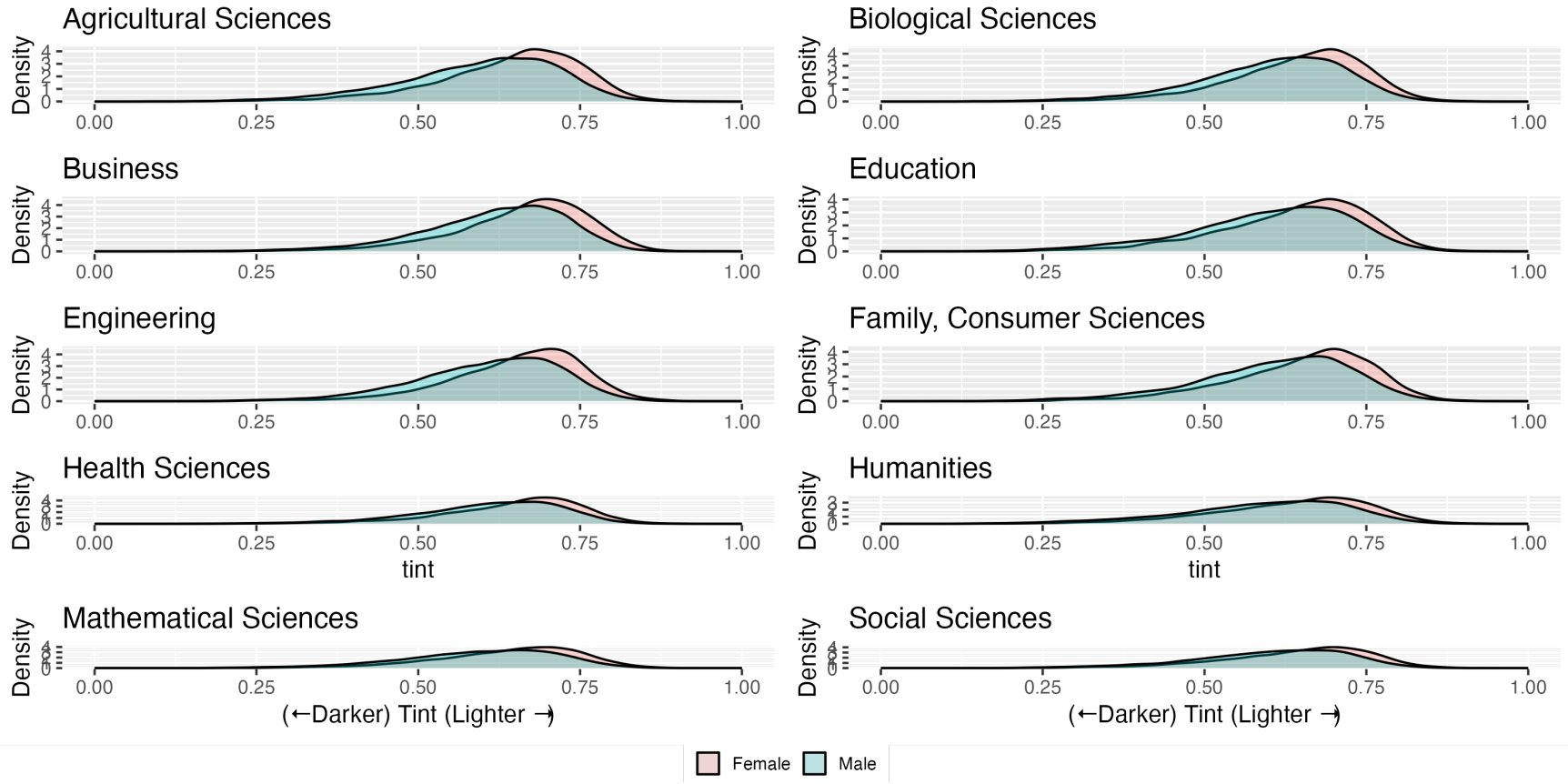
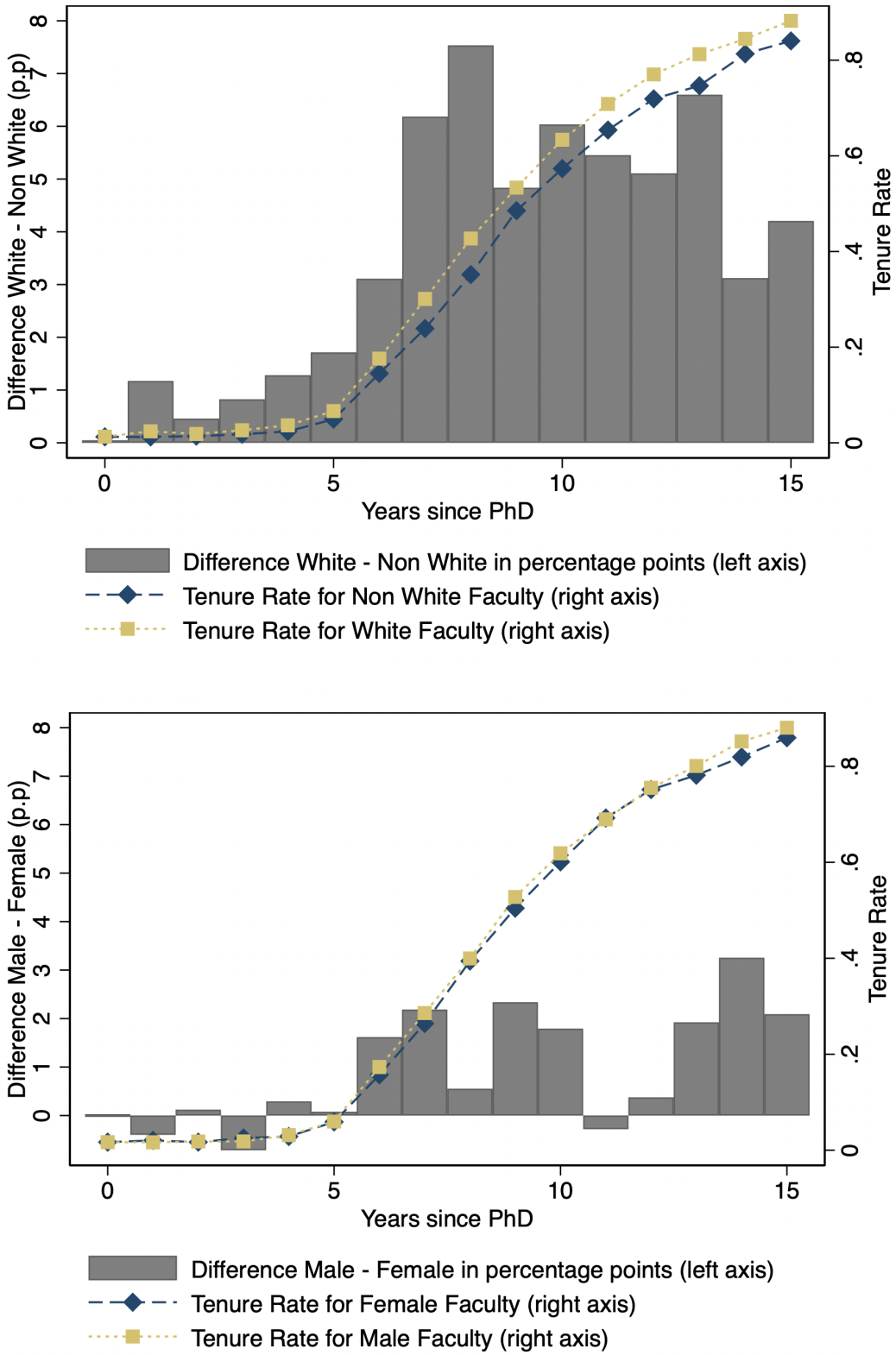
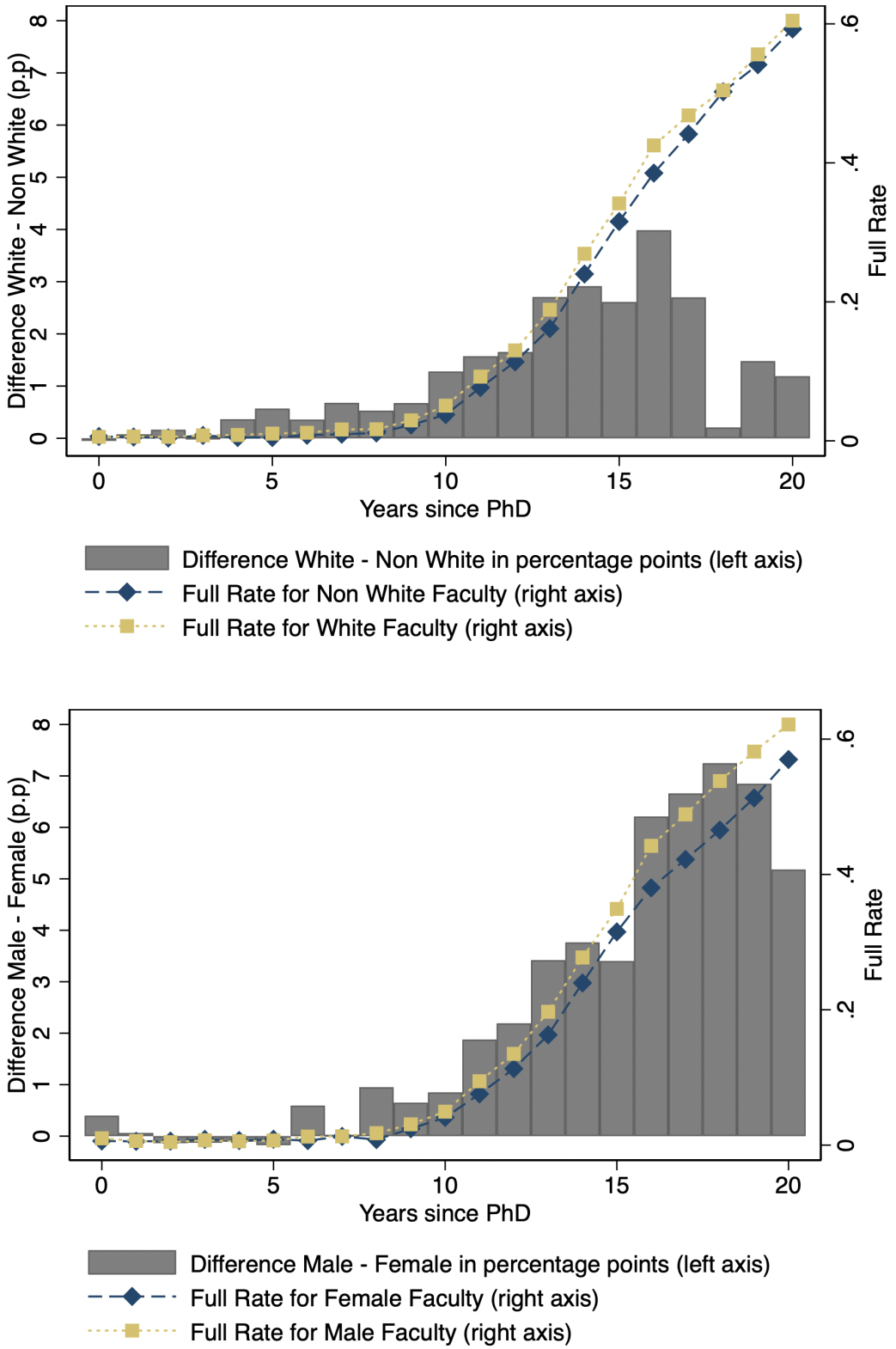


Figure 13: Differences in Tenure Rates



* *Note:* Author's calculation based on AARC data.

Figure 14: Differences in Rates of Promotion to Full Professorship



* *Note:* Author's calculation based on AARC data.

A Labeling exercise

We randomly selected 2,000 images from the 86,744 images available in the FairFace training data set (Karkkainen and Joo, 2021) to conduct a labeling exercise. Table 2 presents a comparison of the attributes between the full sample and our randomly selected sub-sample. The sub-sample means are statistically identical to the ones in the full sample.

We decided to replace 68 images with cropped versions of them when the images had two or more faces and it was not clear which face needed labels.

Table 2: Attributes of Fairface dataset and subsample

	Full sample n = 86,744		Sub-sample n = 2,000	
	Mean	SD	Mean	SD
Gender				
Female	0.470	0.499	0.452	0.498
Male	0.530	0.499	0.548	0.498
Age				
0-2	0.021	0.142	0.023	0.148
3-9	0.120	0.325	0.114	0.318
10-19	0.105	0.306	0.101	0.301
20-29	0.295	0.456	0.300	0.459
30-39	0.222	0.416	0.227	0.419
40-49	0.124	0.329	0.120	0.326
50-59	0.072	0.258	0.070	0.255
60-69	0.032	0.176	0.036	0.185
70 or more	0.010	0.098	0.010	0.099
Race				
Black	0.141	0.348	0.155	0.362
East Asian	0.142	0.349	0.132	0.338
Indian	0.142	0.349	0.140	0.347
Latino-Hispanic	0.154	0.361	0.162	0.368
Middle-Eastern	0.106	0.308	0.103	0.305
Southeast Asian	0.124	0.330	0.113	0.317
White	0.191	0.393	0.195	0.396

The labeling exercise was carried by 6 enumerators, 3 were presented images to label the skin color using a NIS scale and 3 using a Fenty scale. Labelers completed the exercise of labeling skin tone, gender, race and skin darkness for 2,000 images.[3]

[3]Out of the 2,000 images, a few were omitted likely due error. We excluded the skipped images to trained the model.