

“From Global Indicators to Local Applications”

7-9 September 2022 | Granada, Spain

#STI22GRX

Applying an Intersectional Lens to Author Composition at Women’s Colleges, Historically Black Colleges and Universities, and Hispanic Serving Institutions in the United States

Diego Kozlowski*, Sonia Doshi**, Amara Rangwala**, Cassidy R. Sugimoto**, Vincent Larivière***, and Thema Monroe-White****

*diego.kozlowski@uni.lu

Faculty of Science, Technology and Medicine, University of Luxembourg, 6 avenue de la Fonte, Esch-Sur-Alzette, L-4364 (Luxembourg)

** sdoshi62@gatech.edu; arangwala8@gatech.edu; sugimoto@gatech.edu

School of Public Policy, Georgia Institute of Technology (USA)

*** vincent.lariviere@umontreal.ca

École de bibliothéconomie et des sciences de l’information, Université de Montréal, Montréal, QC H3T 1N8 (Canada)

**** tmonroewhite@berry.edu

Department of Technology, Entrepreneurship & Data Analytics, Berry College, Mount Berry, GA 30149 (USA)

Introduction

Institutions of higher education (HEIs) are not created equal. In the United States (U.S.) context, HEIs can vary dramatically with respect to their mission and focus, with implications on the composition of the faculty, staff, and student body. The institutional context has implications on the research funding received, collaboration opportunities, teaching expectations, the availability and access to departmental, college, and institutional supports, and professional service obligations placed on pre-tenure and tenured faculty (Dundar & Lewis, 1998; Bland et al., 2005; NRC, 2010). Institutional context may also influence the productivity level of faculty in terms of publications and grants, as evaluation criteria by institution can vary widely, based on different weights assigned to publications, teaching, and service, respectively. Furthermore, inclusive cultures vary between and within HEIs. Faculty, students, and staff of varying socio-demographic identities (i.e., first generation, low-income, race/ethnicity, gender, disability status, sexual orientation etc.) can experience varied levels of inclusion by college, school, department, program, and classroom environment. Previous work has demonstrated strong alignment between diversity in the scientific workforce and an expansion of the knowledge base, with higher prevalence of topics that align with and reflect the lived experiences of minoritized populations (Kozlowski et al., 2022; Sugimoto et al., 2019). It is unknown, however, whether and how knowledge production varies as a function of the composition of the scientific workforce within the institution with which faculty are affiliated.

This article adopts an intersectional lens to describe the distribution of authors by race and gender in three types of institutions: Historically Black Colleges and Universities (HBCU), Hispanic Serving Institutions (HSI), and Women's Colleges. These institutions are uniquely focused on demographics of students; however, little is known about the authors that emerge

from these different types of institutions. To begin to understand the relationship between institution and knowledge production, we examine composition using four measurements: 1) with respect to the proportion of U.S. authors by race, gender and year on each type of institution; 2) relative over/under representation relative to the population of authors; 3) the ratio with respect to the U.S. Census; and, 4) the ratio with respect to the number of PhD graduates. Understanding the distribution of authors by race and gender sheds new light on the role these specific institutions (i.e., HBCUs, HSIs and women's colleges) play in diversifying the scientific landscape of the U.S.

Background

Women's colleges in the U.S. were founded in 1836 to increase access to higher education (Harwarth et al. 1997). The founders of these institutions were advocates of women's intellectual abilities and promoted their participation in society and in practical trades (ibid.). The development of women's colleges was also fueled by difficulties faced by women for entering higher education. Such barriers included legal restrictions, quotas, administrative regulations, and anti-nepotism rules (Parker, 2016). One such example is the quota system enacted by Stanford university between 1933 and 1973, under which one woman was accepted for every three men accepted (ibid.). Women's colleges are particularly notable for the equity provided for women when compared to co-educational institutions during the 19th century and continue to model this behavior of equity for women's education in modern society where access is not considered a significant issue (Langdon, 2001). In 1960, there were 230 women's colleges which has now been reduced to 47 women's colleges within the U.S. and Canada during 2014 (Women's College Coalition, n.d.). Women's colleges are the only type of HEI where women account for the majority of the faculty (61.4%) (Sax, 2014). By comparison, women represent 39.1% of faculty at American public institutions and 37.7% of private institutions in the U.S. (ibid.). Despite women being proportionally more present in women's colleges, there is a sizable difference at the full professor level: 39.8% of men reached that position compared with 28.5% of women (ibid.). The racial composition of faculty at women's colleges is 86% is White, with only 3.2% identifying as Black/African Americans (ibid.). The mission to focus on women's education also shapes the curriculum: the Higher Education Research Institute's triennial Faculty Survey shows that women's colleges have the highest assignment of readings on issues related to women, gender, race, and ethnicity and that the faculty research interests similarly focus on these issues (ibid.). These data align, in part, with our study demonstrating alignment with characteristics of the authors and the topics they study (Kozlowski et al., 2022).

Ten years after *Brown vs. Board of Education*, in 1965, the Higher Education Act defined Historically Black Colleges and Universities (HBCUs) as "institutions of higher learning" dedicated to the "education of Black Americans" and the preparation of Black men and women to enter the workforce. The history of these institutions is rooted in the systematic legal and social exclusion of Black people from formal U.S. educational institutions. As of Fall 2017, there are currently 101 HBCU institutions, representing 2.3% of degree-granting post-secondary institutions within the U.S. More than half (59%) of HBCUs only offer undergraduate degrees, while 41% award graduate degrees, and 28% grant doctoral degrees. Despite their small size HBCUs graduate a disproportionate number of Black students: 20-30 percent of all bachelor's degrees earned by Black students are from HBCUs (Sauders & Nagle, 2018; Lovett, 2015). These graduates are majority women, though the faculty have slightly more men than women (Lovett, 2015). As of 2001, Black faculty constituted, on average, less than 60% of the faculty at HBCUs (Lovett, 2015). Salaries are also depressed, compared to other HEIs, which leads to issues in faculty recruitment and retention (Lovett, 2015). Most

HBCUs (89%) are concentrated within the Southern region of the U.S which aligns with the relative overrepresentation of Black people in the Southern U.S States (Tamir, 2021). HBCUs fostered a sense of racial pride and uplift which cultivated an affirmational environment where Black students felt safe to use their education for upward social mobility. HBCUs enforce the notion of using education as a tool for liberation from the persistent traumas invoked by structural racism and racial oppression and count among their graduates many leaders of the modern U.S. Civil Rights Movement--such as Thurgood Marshall, Rosa Parks, Martin Luther King Jr, and John Lewis.

Minority serving institutions (MSIs) are defined by the Office of Civil Rights as institutions that have at least 50% enrollment of minoritized populations—including undergraduate and graduate students, as well as part-time and full time students. Therefore, while women's colleges and HBCUs take a mission-based approach, MSIs are categorized using a threshold approach (and therefore maintain their status based on student enrollment). MSIs include HBCUs, Predominantly Black Institutions (PBIs), Tribal Colleges and Universities (TCUs), and Hispanic Serving Institutions (HSIs) (U.S. Department of Interior, n.d.; Legal Information Institute, n.d.). MSIs, by virtue of the composition of their student body, serve low-income, underrepresented students of color (U.S Department of Interior, n.d.). In 2012, MSIs enrolled approximately 3.6 million undergraduates (Center for Minority Serving Institutions, 2014). Of these students, over half of the enrolled students qualified for Pell Grants, which is an American federal subsidy for students who display exceptional financial need (Center for Minority Serving Institutions, 2014; U.S Department of Education, n.d.). In comparison, only 31% of all students within the U.S. displayed enough financial need to qualify for a pell grant (Center for Minority Serving Institutions, 2014). There are currently about 700 MSIs in the U.S. but this is a number with regular fluctuation based on marginalized students' enrollment (National Academies of Science, Engineering, and Medicine, 2019). The faculty at MSIs are known to have greater diversity which by some scholars has been attributed to the institutional culture (Blake, 2018). Professors who were interviewed regarding their motivations for teaching at MSIs stated that it was predominantly to make a difference in their community, as opposed to the traditional motivations of academic freedom or autonomy (Blake, 2018).

Data & Methods

Our dataset consists of 4,865,896 articles with U.S. affiliations published between 2008 and 2020 and indexed in the Web of Science (WOS), and the 4,285,954 corresponding U.S. authors affiliated with these articles. These articles were associated with 261,336 distinct institution name strings, which were cleaned to assign papers to specific universities. The cleaning process of institutions consisted of two tasks: first, normalizing the multiple names by which the name of the same university appears in WOS; and, second, building a crosswalk between institutions' names as they appear in WOS and in the Carnegie list of institutions. Both tasks were first conducted algorithmically, and then checked manually. After cleaning, the final dataset consists of 4,076,167 articles, 3,130,039 U.S. authors, and 560 universities, which covers 84% of articles and 73% of authors contained in the full U.S. dataset. Out of the 560 universities, 23 are HBCUs, 154 are HSIs, and 7 are women's colleges. The lack of coverage of all institutions may in part be due to a low signal in WoS for many HBCUs, WCs, and MSIs. For example, we took a manual approach to retrieving all articles with Tribal Colleges, another important mission-driven institution in the U.S. categorized by the Carnegie classification. However, the low volume of retrieved articles reduces the validity of the application of the racial disambiguation algorithm for these institutions. Therefore, this classification type was removed from the analysis. This is a limitation of the present work.

Following Kozlowski et. al. (2022a), authors of those papers were assigned a race based on the association between their family names and race found in the US census data (USCB, 2016). Gender was inferred using authors' given names, based on the method presented in Larivière et al. (2013). Gender is considered in a binary way, as other genders can only be assigned through self-identification. This is a limitation of algorithmic approaches. As shown in Kozlowski et. al. (2022b), the demographics of US authors have a different distribution than those of the 2010 US census, with an under-representation of Black and Latinx authors. Therefore, using the 2010 US as a source for the name-based racial inference can potentially overestimate these groups, as has been recently shown by LaBerge et. al. (2022). The reason for this is that Black and White identities tend to share family names, given the historical legacy of assigning names to human property during the hundreds of years of institutionalized chattel slavery in the U.S. As our results show below, there are reasons to believe these methods hold face validity for a number of analyses, but can nevertheless overestimate the proportion of Black and Latinx authors and, therefore, can be considered as an upper bound. Furthermore, the automatic inference of racial groups impede us to acknowledge Native Americans and individuals with more than one race.

According to the latest U.S. Census (2020), 61.6% of the population identified as White, 12.4% as Black or African American, and 17.1% as Hispanic or Latino. Given the unequal size of these populations, we examine not only the absolute proportions, but also three other metrics. First, we compute the relative over/under representation of a given race and gender identity on each university type, as the average proportion by race and gender on that institution type and year, with respect to their mean proportion on that year, across all institutions. This metric allows us to understand in which types of institutions specific groups are relatively more present within the population of authors. Secondly, we calculate the ratio between the proportion of authors by race and gender each year with respect to the proportion of that race and gender in the 2010 U.S. Census. This comparison allows one to better understand the alignment between the population size of an identity, and their participation in scientific publishing. Nevertheless, it is also known that the migration patterns in academia are different from those in the general population, which could generate a misleading impression, especially regarding Asian authors. Therefore, in our third approach, we compare the number of authors by race and gender each year and institution type with the number of PhD graduates by race and gender between 2010 and 2019, using the Survey of Earned Doctorates (SED)¹ data. This will produce an expected value from those who are trained in scientific publishing within the U.S.

Results

White authors represent the dominant author group across all institution types and within HBCUs, HSIs, and women's colleges (Figure 1). White men are dominant in all cases except women's colleges, where White women are in the majority. There is, however, a trend towards diversification: the proportion of White men across all institutions has dropped from 40% in 2008 to 35% in 2018. (As a comparison, the proportion of the U.S. that identified as White dropped from 72.4% in 2010 to 61.6% in 2020, according to the U.S. Census Bureau.) The decreasing trend is not observed across all types of institutions: the proportion of White men in HBCUs and women's colleges remains stable over the decade.

¹ <https://nces.nsf.gov/pubs/nsf21308/data-tables>

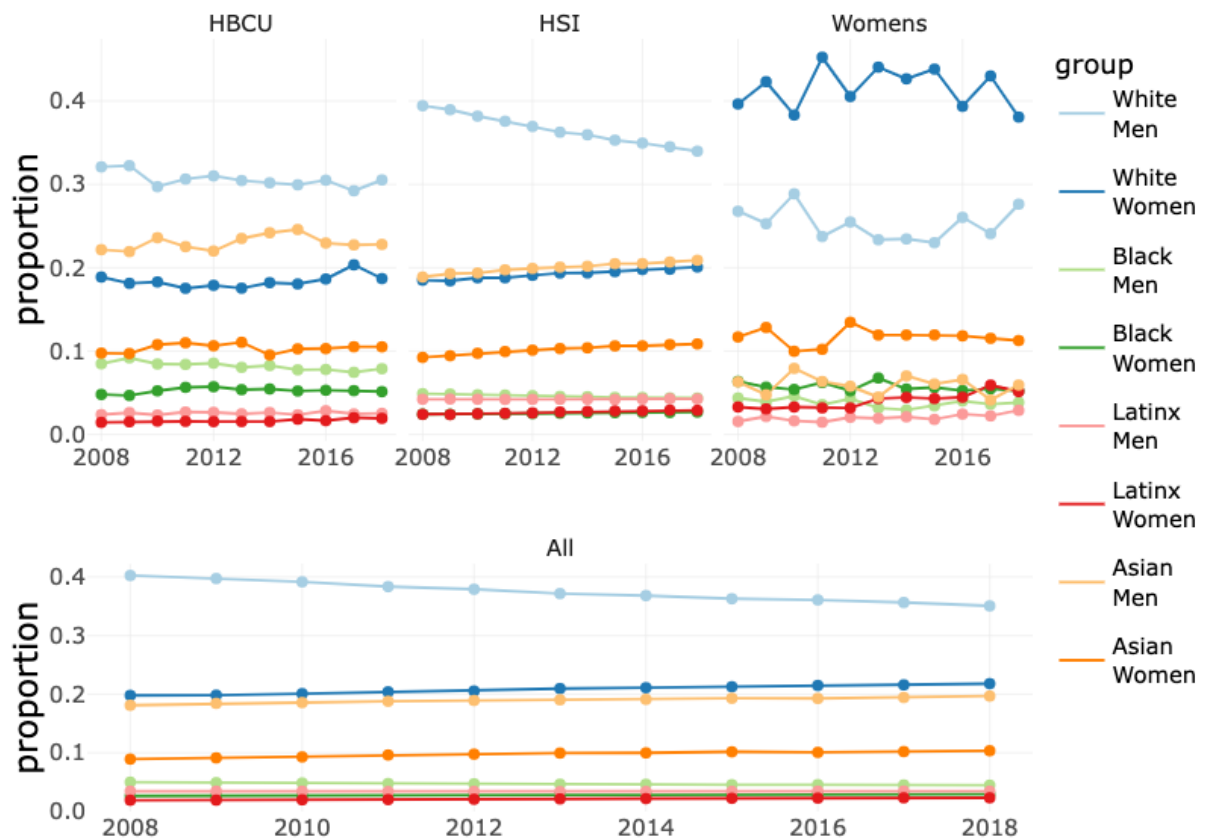


Figure 1. Proportion of authors by race and gender over time. HBCU: historically Black Colleges and Universities, HSI: Hispanic Serving Institutions, Womens: Womens' colleges, All: all institutions on our dataset.

When taking into account the relative observance of these groups across the population of U.S. authors, we observe that minority groups are overrepresented in HBCUs, HSIs, and women's colleges. Black men and women are overrepresented at HBCUs, Latinx men and women at HSIs, and women at women's colleges (Figure 2). These results show that the demographic characteristics of authors at these institutions are consistent with mission-oriented and threshold-based alignment with particular social groups. The results also provide a secondary methodological validation, suggesting that the automatic inference of race and gender of

authors manage to capture aggregate trends. It is worth noting that Asian authors also have a high participation at both HBCU and HSIs.

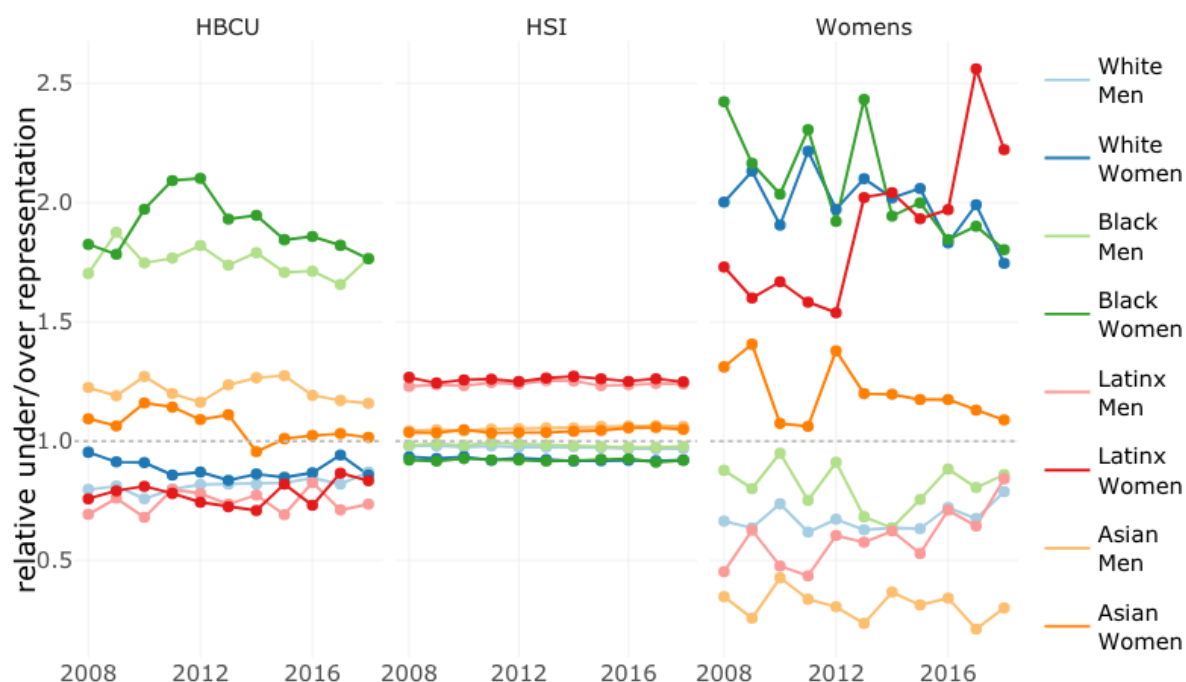


Figure 2. Relative over/under representation of groups by race and gender, relative to their participation in the overall dataset. HBCU: Historically Black Colleges and Universities, HSI: Hispanic Serving Institutions, Womens: Womens' colleges.

In order to assess the over/under representation of demographic groups relative to their proportion of the population, we compared the proportion of authors by race and gender in each year to their proportion in the 2010 U.S. Census (Figure 3). These results show that the proportion of Asian men among authors is eightfold their proportion in the U.S. population, and fourfold in the case of Asian women. Nevertheless, the interpretation of these results should be done carefully. Migration patterns in academia are not equal to those in the general population, and they also differ by race and gender. SED data shows that between 2010 and 2019, 80% of Asian men and 69% of Asian women graduated with a temporary visa, suggesting that this population is largely foreign-born. In all other groups, temporary visa holders are a minority². Therefore, it is not correct to infer that the overrepresentation of Asian authors is due to structural inequalities that privilege Asian-Americans in their access to academia, but due to specific global migration patterns. When considering the proportion of permanent residents that graduate in each race and gender group, the overrepresentation of Asian men and women decrease to 1.4 and 1.2 respectively. Of course, this is an imperfect proxy, given that the

² <https://sciencebias.uni.lu/app>

proportion of temporary visa holders among highly-skilled workers varies dramatically from the proportion in the general population.

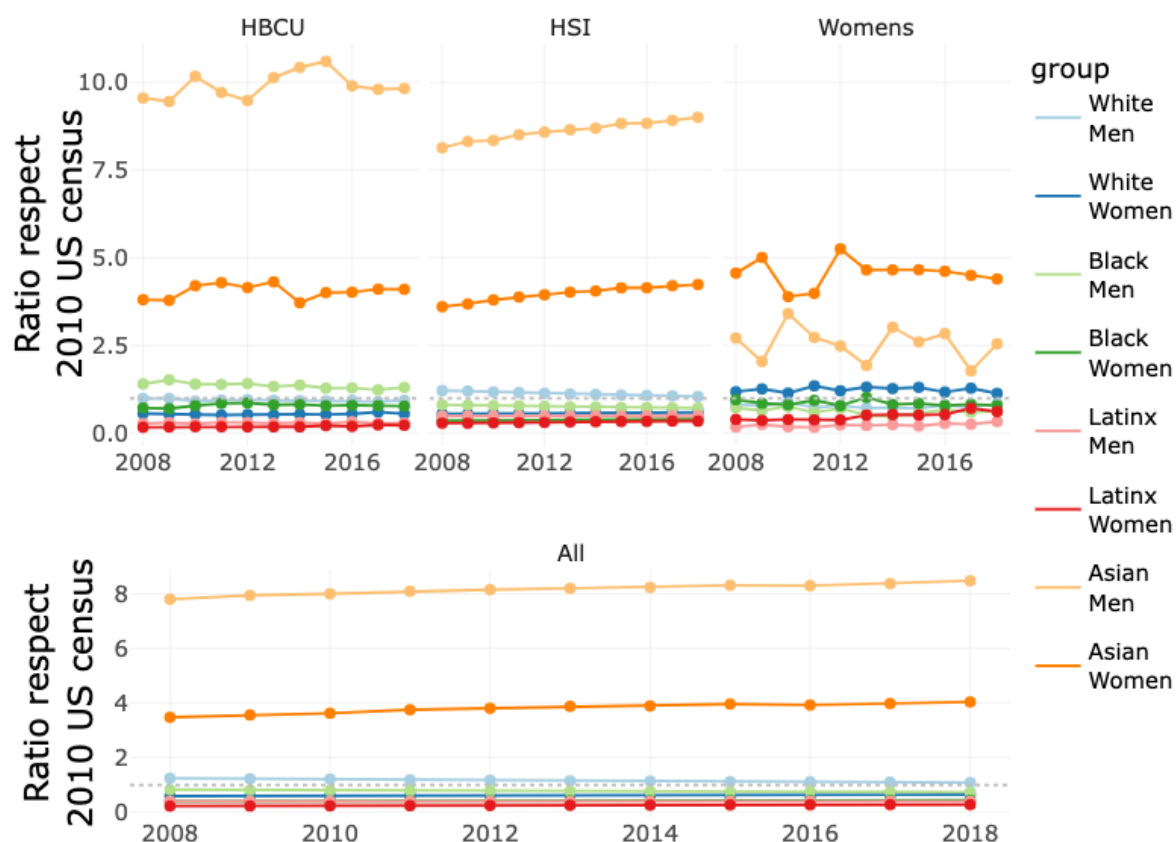


Figure 3. Ratio of authors by race and gender, relative to the U.S. census. HBCU: historically Black Colleges and Universities, HSI: Hispanic Serving Institutions, Womens: Womens' colleges, All: all institutions on our dataset.

The complexities of comparing with the population--with and without respect to immigration status--may be solved by setting expected values from the population of doctoral graduates. Compared with PhD graduates, Black men are overrepresented in all types of institutions (Figure 4), which is likely due to the overestimation of Black authors in name-based racial inference (LaBerge et. al, 2022). This overestimation should be considered an upper bound of the representation of Black authors in science—if they appear to be underrepresented, then the reality is likely lower. The overestimation, however, generates problems when comparing to other statistics, like doctoral graduates: given that the number of Black men PhD graduates is small, the overrepresentation analyses are highly volatile. These results, therefore, should be interpreted with caution. That said, the degree of overrepresentation in HBCUs is higher than in other institutions, which suggests that the method (despite limitations), remains an informative signal of representation. Other nuances can also be observed. For example, White men have a larger relative presence in HBCU than Black women, and Latinx authors are the

most underrepresented group with respect to their proportion among the U.S. population, even in HSI.

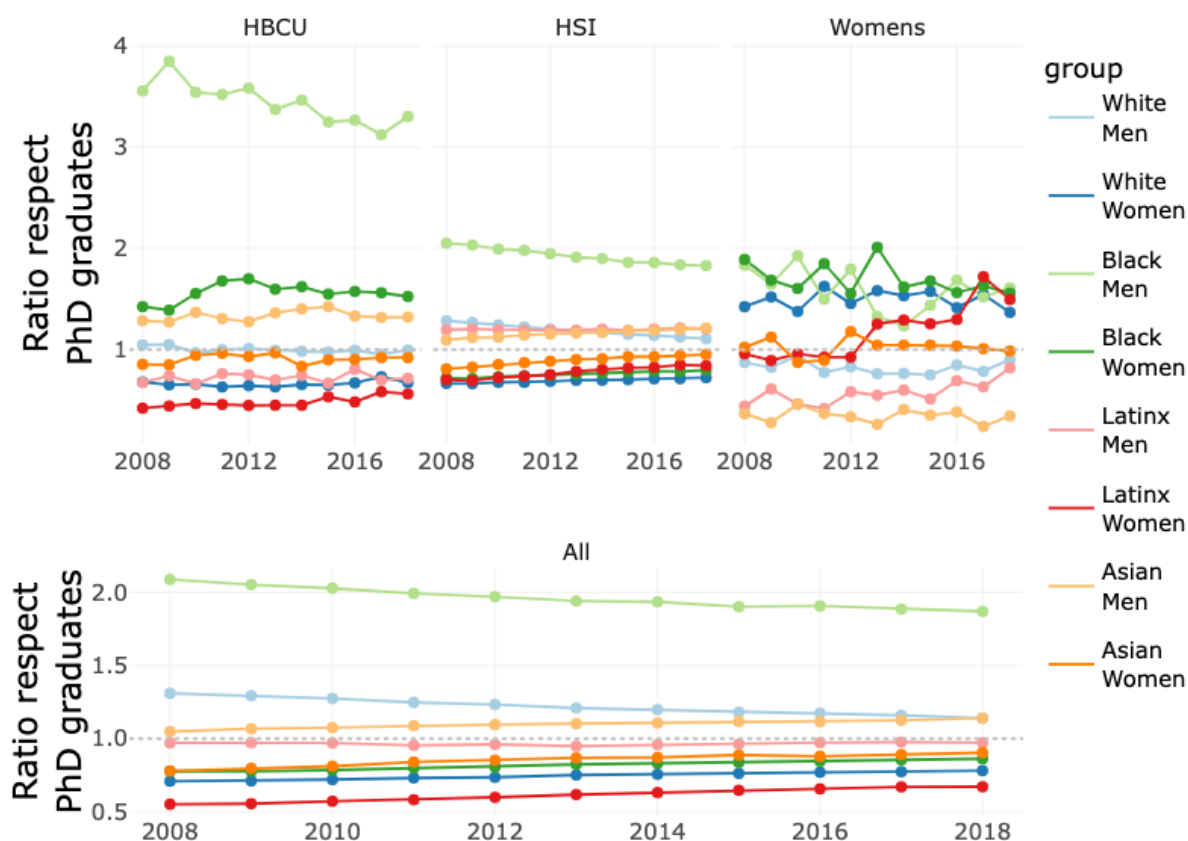


Figure 4. Ratio of authors by race and gender, relative to the number of PhD graduates. HBCU: historically Black Colleges and Universities, HSI: Hispanic Serving Institutions, Womens: Womens' colleges, All: all institutions on our dataset.

Discussion

Our analysis reinforces the dominance of White authors across institution types, even those with missions or thresholds that emphasize other racial groups. The most striking observation from our institutional analysis was the proportional overrepresentation of White and Asian men and women as published authors at HBCUs (see Figure 2). This statistic raises questions regarding support needed to facilitate the work of Black academics, and the persistent structural barriers within the higher education industry that lead to the continued underrepresentation of Black women and men. Future research should explore these findings with survey and/or interview data to unpack these patterns; and identify the racialized and gendered, motivational and institutional factors that drive publication rates by race/ethnicity and gender of authors at these institutions over time.

Despite the overwhelming majority of White authors, our analysis does suggest that HBCUs, HSIs, and women's colleges overrepresent authors from their target populations to a degree that is higher than expected among all authors. However, normalization by U.S. permanent residents shows an overrepresentation of Asian authors and the analysis by doctoral graduates suggests an overrepresentation of Black men. These analyses suggest that the data are highly sensitive to normalization, with each approach reflecting a different element of inequities in the U.S.

scientific workforce. This work sets a platform for future work on the effect of institutional composition on knowledge production.

References

Agency for Healthcare Research and Quality (n.d.) “Racial / Ethnic Minorities”

<https://www.ahrq.gov/topics/racial-ethnic-minorities.html>.

Albritton, Travis J. “Educating Our Own: The Historical Legacy of Hbcus and Their Relevance for Educating a New Generation of Leaders.” *The Urban Review* 44, no. 3 (2012): 311–31. <https://doi.org/10.1007/s11256-012-0202-9>.

Blake, D. (2018, April 16) “Want a Diverse Faculty? Learn from Minority Serving Institutions,” *Diverse: Issues In Higher Education*.
<https://www.diverseeducation.com/institutions/msis/article/15102346/want-a-diverse-faculty-learn-from-minority-serving-institutions>.

Bracey, Earnest N. “The Significance of Historically Black Colleges and Universities (HBCUs) in the 21st Century: Will Such Institutions of Higher Learning Survive?” *American Center for Minority Serving Institutions* (2014) “Rutgers GSE CMSI - A Brief History of MSIs.”. <https://cmsi.gse.rutgers.edu/content/brief-history-msis>.

Harwarth, Irene, Elizabeth DeBra, and Mindi Maline (1997) *Women’s Colleges in the United States: History, Issues, & Challenges*. DIANE Publishing.

“History of HBCUs.” Thurgood Marshall College Fund. Accessed April 30, 2022.
<https://www.tmcf.org/history-of-hbcus/>.

Kozlowski D, Murray DS, Bell A, Hulsey W, Larivière V, et al. (2022) Avoiding bias when inferring race using name-based approaches. *PLOS ONE* 17(3): e0264270.
<https://doi.org/10.1371/journal.pone.0264270>

Kozlowski, D., Larivière, V., Sugimoto, C. R., & Monroe-White, T. (2022b). Intersectional inequalities in science. *Proceedings of the National Academy of Sciences*, 119(2).

LaBerge, N., Wapman, K. H., Morgan, A. C., Zhang, S., Larremore, D. B., & Clauset, A. (2022). Subfield prestige and gender inequality in computing. *arXiv preprint arXiv:2201.00254*.

Langdon, E. (January 2001) “Women’s Colleges Then and Now: Access Then, Equity Now.” *Peabody Journal of Education* 76, no. 1: 5–30.
https://doi.org/10.1207/S15327930PJE7601_02.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature*, 504(7479), 211–213.

Legal Information Institute (n.d.) “Definition: Minority-Serving Institution from 7 CFR § 3430.302 | LII / Legal Information Institute.”
<https://www.law.cornell.edu/definitions/index.php?width=840&height=800&iframe=true&de>

f_id=b7a95e98388e0f4b7387b9dbc92e16b5&term_occur=999&term_src=Title:7:Subtitle:B:Chapter:XXXIV:Part:3430:Subpart:G:3430.303\.

National Academies of Science, Engineering, and Medicine (2019) “Minority Serving Institutions: America’s Underutilized Resource for Strengthening the STEM Workforce” at *NAP.Edu*. <https://doi.org/10.17226/25257>.

Parker, P. (Spring 2015) “The Historical Role of Women in Higher Education.” *Administrative Issues Journal* 5, no. 1: 3–14. <https://doi.org/10.5929/2015.5.1.1>.

Sax, L. (2014) “Who Teaches at Women’s Colleges .” University of California, Los Angeles. https://www.womenscolleges.org/sites/default/files/report/files/main/wcc_faculty_report_final.pdf

Sugimoto, C.R. Ahn, Y.-., Smith, E., Macaluso, B., & Lariviere, V. (2019). Factors affecting sex-related reporting in medical research: a cross-disciplinary bibliometric analysis. *Lancet*, 393: 550-559.

US Census Bureau, Frequently Occurring Surnames from the 2010 Census. (USCB, 2016). https://www.census.gov/topics/population/genealogy/data/2010_surnames.html. Accessed 21 December 2021.

U.S Department of Education (n.d.) “Federal Student Aid.” <https://studentaid.gov/understand-aid/types/grants/pell>.

U.S. Department of Interior (n.d.) “Minority Serving Institutions Program,” <https://www.doi.gov/pmb/eeo/doi-minority-serving-institutions-program>.

Women’s College Coalition (n.d.) “History | Women’s College Coalition.” <https://www.womenscolleges.org/history>.

Dundar, H., & Lewis, D. R. (1998). Determinants of research productivity in higher education. *Research in higher education*, 39(6), 607-631.

Bland, C. J., Center, B. A., Finstad, D. A., Risbey, K. R., & Staples, J. G. (2005). A theoretical, practical, predictive model of faculty and department research productivity. *Academic Medicine*, 80(3), 225-237.

National Research Council. (2010). Gender differences at critical transitions in the careers of science, engineering, and mathematics faculty. National Academies Press.

Saunders, K. M., & Nagle, B. T. (2018). HBCUs Punching above Their Weight: A State-Level Analysis of Historically Black College and University Enrollment and Graduation. Frederick D. Patterson Research Institute, UNCF.

Tamir, C. (2021). The growing diversity of Black America. Pew Research Center's Social & Demographic Trends Project. Retrieved from <https://policycommons.net/artifacts/1434774/the-growing-diversity-of-black-america/2055963/> on 01 May 2022. CID: 20.500.12592/80xzf5.