Title:

Differential retention contributes to racial/ethnic disparity in U.S. academia

Authors:

Allison K. Shaw[1], Chiara Accolla[1], Jeremy M. Chacón[1], Taryn L. Mueller[1], Maxime Vaugeois[1], Ya Yang[2], Nitin Sekar[3], Daniel E. Stanton[1]

[1]Department of Ecology, Evolution and Behavior; University of Minnesota-Twin Cities, Saint Paul, MN 55108

[2]Department of Plant and Microbial Biology; University of Minnesota-Twin Cities, Saint Paul, MN 55108

[3]Wildlife and Habitats Division, WWF India, New Delhi, Delhi 110003, India

Abstract

16 Several racial and ethnic identities are widely understood to be under-represented within

academia, however, actual quantification of this under-representation is surprisingly limited.

18 Challenges include data availability, demographic inertia and identifying comparison points. We

use de-aggregated data from the U.S. National Science Foundation to construct a null model of

20 ethnic and racial representation in one of the world's largest academic communities. Making

comparisons between our model and actual representation in academia allows us to measure the

22 effects of retention (while controlling for recruitment) at different academic stages. We find that,

regardless of recruitment, failed retention contributes to mis-representation across academia and

24 that the stages responsible for the largest disparities differ by race and ethnicity: for Black and

Hispanic scholars this occurs at the transition from graduate student to postdoctoral researcher

26 whereas for Native American/Alaskan Native and Native Hawaiian/Pacific Islander scholars this

occurs at transitions to and within faculty stages. Even for Asian and Asian-Americans, often

28 perceived as well represented, circumstances are complex and depend on choice of baseline. Our

findings demonstrate that while recruitment continues to be important, retention is also a

30 pervasive barrier to proportional representation.

34 **Introduction**

36 Large segments of society are under-represented in academic Science and Engineering (S&E) [1,2]. For example, in 2017, 12% and 0.7% of the general U.S. population were Black and

38 American Indian/Alaskan Native respectively, compared to 10% and 0.5% of students graduating with a bachelor's degree, and 4% and 0.2% of tenured faculty [3]. Critically, the groups that are

40 most under-represented in S&E are the fastest growing in the U.S. population [2].

42 Understanding and addressing mis-representation (representation that differs from a baseline expectation of proportional representation) within academia is important for numerous reasons.

44 First, mis-representation of groups can indicate that access is not equitably distributed and that some groups have been excluded from academia [4,5]. Second, mis-representation can mean that

46 some of the best minds are excluded from academia [4]. Furthermore, because members of under-represented groups across various axes (gender, race, experience) can produce innovative

48 work at higher rates than those of well represented groups [6], current mis-representation may be lowering overall academic productivity. Third, although researchers individually have unique

50 perspectives (and thus biases), diversity across researchers can minimize collective bias and improve objectivity [4]. Finally, representation in academia can facilitate a virtuous cycle:

52 academics, as instructors and thought-leaders, are often role-models to those considering professional scholarship, so a diverse academic environment can help draw talent from all

54 segments of society/backgrounds [7].

56   A critical step in addressing mis-representation is determining where and when disparities occur.

Historically, U.S. academia has been primarily composed of White scholars with under-

58   represented minorities systematically excluded from the late 1800s through to the 1970s [2] .

Although U.S. academia (especially at the undergraduate stage) has become more diverse in the

60   past 40 years, most racial/ethnic groups are still under-represented compared to the general U.S.

population [2]. Mis-representation at any stage in academia can be driven by recruitment into --

62   as well as retention within -- that stage [8,9]. Past efforts to increase under-represented groups

have primarily focused on recruitment into the undergraduate stage, and have seen limited

64   success [10]. Increasingly there is a call for addressing factors that shape retention of under-

represented groups in academia post-undergrad [9–12].

66

Despite its widespread existence and importance, mis-representation across academia is

68   challenging to study for a number of reasons. First, defining an appropriate baseline for

racial/ethnic minorities can be challenging. Critically, U.S. demographics are continuously

70   changing [13], and yet academic training is a multi-decade process, which means comparisons of

current academia to current census data are ignoring a potential lag effect. This heterogeneity

72   obscures any clear targets for what diversity 'should' look like. Second, data are often lacking,

either on the number of individuals (e.g., low sample size of under-represented groups) or over

74   time (e.g., long enough data to look for temporal trends). Thus, many studies that aim to test for

race/ethnicity-based differences often lack the sample size or statistical power [14]. Finally,

76   analyses can only be as disaggregated as the categories underlying the data. Studies often lump

together several minority groups into a broad 'under-represented minority' (URM) category [15].

78

Here, we combine two approaches to overcome these hurdles and quantify mis-representation

80   across racial/ethnic groups and across academia. We leverage large national datasets collected by

the United States National Science Foundation (NSF) on the racial and ethnic composition of all

82   U.S. Science and Engineering academics from undergraduate students to tenured professors,

spanning 25 years for students and faculty (seven years for postdocs). We generate a baseline

84   expectation for the racial/ethnic composition of academia by developing a null model [16,17]

that dynamically accounts for historical changes in racial/ethnic compositions. Using these two

86   tools, first, we quantify what racial/ethnic composition we would expect to see in academia, in a

scenario where individuals of each race/ethnicity were equally likely to have an academic career

88   (the null model). Second, we determine to what degree the actual representation of each

racial/ethnic group in each stage of academia (e.g., doctoral student, professor) is higher, equal,

90   or lower than that predicted by the null model. This approach allows us to control for recruitment

and measure the effects of differential retention. Finally, we show that the deviance from the null

92   model differs by racial/ethnic group and by academic stage. Our results provide a novel

perspective on the status of diversity in academia, the critical role of retention, and the

94   challenges academics continue to face.

96   **Methods**

98   We constructed a model of academia (Fig. 1) in the United States as a series of stages with inputs

(from the previous stage) and outputs (to the next stage or move out of academia). We

100   parameterized our model structure with data collected by NSF for Science and Engineering fields

(Biological and agricultural sciences; Earth, atmospheric, and ocean sciences;

102   Mathematics/computer sciences; Physical sciences; Psychology; Social sciences; Engineering)

for the years 1991-2016. We used our model to generate simulated 'predictions' of the

104   representation we would expect of each federally categorized racial/ethnic group (Asian,

Black/African-American, Native Hawaiian/Pacific Islander, Hispanic/Latino, American

106   Indian/Alaskan Native, White, More Than One Race) in each stage of academia under the null

assumption of no race/ethnicity-based differences in retention. With our approach, we can

108   control for recruitment at one stage of academia and measure the effects of retention to future

stages. That is, what 'should' representation in academia look like if there were no race- or

110   ethnicity-based differences in tendency to move between stages or out of academia, and how

does actual representation differ?

112

Data

114   We used data compiled by the National Science Foundation (NSF) on the structure of academia

(number of scholars in each academic stage, time spent in each stage), the racial/ethnic

116   composition of scholars at each stage, and the approximate age distribution of scholars in each

stage (see Supplementary Material section 1, Figs. S1-S5, Tables S1-S2). Data on the number of

118   bachelors and PhD degrees came from the NSF reports on Science and Engineering Degrees [18]

and Women, Minorities, and Persons with Disabilities (WMPD) [3], data on the number of

120   graduate students and postdoctoral scholars came from the NSF Survey of Graduate Students and

Postdoctorates in Science and Engineering [19], and data on the number of assistant and tenured

122   professors came from the 2019 NSF report on Science and Engineering Indicators [20]. The

length of time in each stage came from the 2018 NSF report on Science and Engineering

124   Indicators [21] for graduate students, the NSF report on Postdoc Participation of Science,

Engineering, and Health Doctorate Recipients [22] for postdocs and the integrated data system

126 Scientists and Engineers Statistical Data System (SESTAT) for faculty.

128 Data on the racial/ethnic composition of undergraduate and PhD students as well as assistant and

tenured professors came from the WMPD reports [3]. Data on postdoctoral researchers (2010

130 onward) came from NSF Surveys of Graduate Students and Postdoctorates in Science and

Engineering [23], and data prior to 2010 was estimated as the average of representation in the

132 graduate student and assistant professor stages. The student data in the NSF WMPD reports only

includes racial/ethnicity data for U.S. citizens and permanent residents. To account for

134 international students, we used the NSF reports on Doctorate Recipients from U.S. Universities

[23] for data on the proportion of permanent vs temporary resident PhD recipients and the

136 racial/ethnic composition of temporary resident PhD recipients. Count data on the number of

scholars of each racial/ethnic group were converted to proportions and data were smoothed with

138 a 5-year window moving average.

140 Finally, we used NSF data on the approximate age range of scholars at each stage by pulling data

from the SESTAT database and determining the most representative ages of each stage. These

142 age ranges were: 15 to 24 years old (undergraduate students), 20 to 29 (graduate students), 25 to

39 (Ph.D. recipients), 25 to 44 (postdoctoral researchers), 30 to 49 (assistant professors) and 35

144 to 59 (tenured professors). We used this data to determine which subset of the general population

we should compare each academic stage to. We determined the racial composition of the age

146 class corresponding to each academic stage based on data from the National Center for Health

Statistics and the U. S. Census Bureau [13].

148

Model structure

150   We constructed a model of academia as a series of stages (Fig. 1), building on previously

developed methods [16]. We considered five academic stages: undergraduate students, graduate

152   students, postdoctoral researchers, assistant professors and tenured professors. We used the time

spent in each stage to estimate a turnover rate for that stage which, in combination with the

154   number of scholars in each stage, gave us an estimate of the number of scholars that would have

either transitioned from one stage to the next or transitioned outside of the system for each year

156   (see Supplementary Material section 2, Fig. S6, Table S3).

158   *Model simulation*

We simulated the flow of scholars through our null model of academia over time, assuming there

160   was no racial/ethnic bias in movement patterns of scholars (see Supplementary Material section

3, Fig. S7-S9, Tables S3-S4). We initialized model simulations in a given starting year t_0 with

162   NSF data on the racial/ethnic composition of each stage in that same year. For each year going

forward, we fed in NSF data on racial/ethnic composition at a particular stage (e.g.,

164   undergraduate students), and used our model to predict the racial/ethnic composition at the other

stages (e.g., graduate students). We simulated the model under four scenarios (based on turnover

166   rate and turnover type) to capture uncertainty in the details of transitions for faculty. For turnover

rate, we considered 'slow' (8 years spent as an assistant professor and 30 years as a tenured

168   professor) and 'fast' (5 years as assistant and 20 years as tenured) turnover rates. For turnover

types, we considered 'supply' (assistant professors achieved tenure at a specified rate, and excess

170   tenured professors were retired accordingly), and 'demand' (tenured professors retired at a

specified rate and excess assistant professors becoming tenured left academia) scenarios. To

172   consider the overall effects of retention, we initialized the model by setting the number of

scholars in each class to the data from 1991 ($t_0 = 1991$), and fed in racial/ethnic data at the

174   undergraduate stage (and at the PhD stage for international students, the earliest stage this data

was available; see above) each year until 2016, and measuring simulated output at all other

176   stages, for each of the four scenarios above. To consider the effects of retention within each stage

of academia, we again initialized the model with 1991 data, but fed in racial/ethnic data at each

178   stage and measuring the model output at the next stage (e.g, feed in graduate student data,

measure postdoc data), and then took the average output across each of the four scenarios.

180

Testing model predictions

182   To test our null hypothesis that there is no racial/ethnic bias in transitions within academia, we

compared the racial/ethnic composition predicted by our null model to the actual composition

184   from NSF data. To quantify relative representation, we used metric

$$\theta_i(t, k) = \frac{\hat{f}_i(t, k) - f_i(t, k)}{\hat{f}_i(t, k)}$$

186   [1]

188   where $\hat{f}_i(t, k)$ and $f_i(t,k)$ are the observed and simulated (respectively) fraction of individuals in

stage *i* at time *t* from racial/ethnic group *k* (see Supplementary Material, section 4). To measure

190   confidence in our results, we considered a 5% increase or decrease in each $\hat{f}_i(t, k)$ and $f_i(t,k)$

values, recalculated $\theta_i(t,k)$ for these, and mark this range of values with confidence intervals.

192

194 **Results**

196 First, we considered the effects of retention across a full academic career while controlling for recruitment at the undergraduate stage (Fig. 2). Our null model predicts that representation of

198 scholars in most groups is still changing over time, indicating that parity would not yet have been reached, even under a null model (Fig. 2 and Fig. S7, solid coloured lines). We also find that

200 increasing representation of non-White scholars (driven by changing undergraduate demographics) does not come with a decrease in the absolute number of White scholars; rather

202 this is driven by an overall increase in the absolute number of scholars in each stage (Fig. 3). Our null model predicts that representation of White scholars would be lower than levels actually

204 observed in academia while all other groups (including Asian scholars, who are not traditionally considered an under-represented minority [3]) would be higher than observed (Fig. 2, coloured

206 lines versus dots). These deviations indicate that race/ethnicity-based biases occur after graduating with a science or engineering undergraduate degree, suggesting differential retention

208 within academia.

210 Second, we can compare our model results to census data, allowing us to consider the effects of recruitment, although indirectly. Here, we compare model predictions (which assume recruitment

212 at the undergraduate stage and control for retention at other stages) with the U.S. general population census data (which included individuals who both were and were not 'recruited' into

214 academia). Our null model predicts that, even if retention were the same across racial groups, representation of White and Asian scholars in academia would still be higher than in the U.S.

216 general population while all other groups would still be lower (Fig. 2 and Fig. S7, black lines).

The differences between the racial composition of the null model and the general population
218 indicate differential recruitment into academia, showing that there are race/ethnicity-based biases
in entering academia. Intriguingly, taken together, our results indicate that Asian scholars can be
220 considered overrepresented in U.S. academia if the baseline for comparison is the U.S. general
population, but can be considered under-represented in U.S. academia if the baseline for
222 comparison is student degree recipients. This result is driven by the fact that many U.S. PhD
recipients are international students (temporary residents; Fig. S5a-b), and that 60+% of these
224 students are Asian scholars (Fig. S5d).

226 Third, we considered the effects of retention within each stage of academia (Fig. 4). Here, we
control for recruitment at each stage of academia and measure the effects of retention to each
228 subsequent stage. We quantify relative representation (driven by retention) as a metric θ, the
deviance from the null model in representation, where positive values ($\theta > 0$) indicate a group
230 has higher representation than the model predicts and negative values ($\theta < 0$) indicate a group
has lower representation than predicted. We find that θ varies by racial/ethnic group and by stage
232 transition within academia (Fig. 4). The transition from undergraduate degree to graduate student
is captured well by the null model ($\theta \approx 0$, i.e., little differential retention at this transition). The
234 biggest loss in representation (lowest retention) for Native American/Alaskan Native and
Hawaiian/Pacific Islander scholars occurs in the transition to being a faculty member and staying
236 within the faculty (Fig. 4). In contrast, the biggest loss in representation for Asian, Black and
Hispanic scholars occurs in the transition from graduate student to postdoctoral researcher, and is
238 the worst for Black representation (Fig. 4). Although temporal trends over 15 years show the
system is approaching parity ($\theta \rightarrow 0$) for some race/ethnicity and stage combinations, deviance

240   from parity is actually increasing for Black, Hawaiian and Native scholars in faculty positions

(Fig. S8).

242

Discussion

244

The novelty of our work is three-fold: we provide new findings on the patterns, causes, and

246   consequences of misrepresentation racial/ethnic groups within US Science and Engineering

academia. In terms of patterns, we present one of the most extensive assessments of

248   misrepresentation, by contrasting one of the world's largest public datasets on demographics of

scholars with a null model of representation. Past studies have demonstrated that some

250   racial/ethnic groups are misrepresented at some stages [15,24], or within some disciplines [17],

however it was not previously clear to what extent these observations scaled up to affect cross-

252   discipline patterns at the national level. Here, we quantitatively show that they do. The breadth

and resolution of our analysis allows us to separate effects by racial/ethnic group (rather than

254   lumping all non-White scholars together, as past studies have done), thus demonstrating that

retention at each academic stage differs by race/ethnicity. The representation patterns that we

256   uncover also highlight the importance of explicitly defining a baseline against which to measure

representation. For example, we find that although Asian representation in academia is higher

258   than in the general U.S. population, it is simultaneously lower than would be predicted based on

student demographics. Much of the racial/ethnic diversity in PhD recipients derives from

260   immigration rather than retention of minority scholars (Fig S5e-f). Most international PhD

students are from India and China [25] and 70% of foreign-born PhD doctorates stay in the U.S.

262   after receiving their degree [26], which fit with our finding that non-U.S. born Asian scholars are

a critical input into U.S. academia (Fig. S9).

264

In terms of causes, we demonstrate definitively that failed retention of Black, Indigenous, and

266   Hispanic scholars is a substantial contributor to misrepresentation in academia. We find that

although representation of non-White scholars in academia is increasing, it is doing so slower

268   than expected under our null model predictions. In other words, although recruitment into

academia at the undergraduate stage is numerically the largest driver of representation, it alone

270   does not explain the lack of parity. Our findings show that training diverse students is not

enough; there is a substantial drop in racial/ethnic representation between students (graduate and

272   undergraduate) and researchers (postdocs and faculty), and bias in retention appears to be

increasing in some cases (Fig. S8, transitions to faculty for Black and Native scholars and within

274   faculty for Native scholars). Overall, these results provide quantitative evidence to support calls

for increased focus on inclusion/retention along with recruitment [9–12]\cite{whittaker2014,

276   bach2006, callahan2017, puritty2017} and show that neither time, nor simple pushes to increase

recruitment are panaceas to this societal challenge.

278

The patterns and causes discussed above have a number of consequences. First, failed retention

280   in academia is most problematic for representation of Black and Indigenous scholars (Fig. 4);

thus, paths forward must draw on understanding the specific cultural context of these scholars as

282   well as the challenges and discrimination that they face within academia [27–29]. Second, our

finding that the most problematic transitions within academia vary by race and ethnicity

284   indicates that different racial/ethnic groups need support at different stages [8]. Thus, policy

change to address misrepresentation within academia must account for the interactive effects

286 between race/ethnicity and academic stage; a one-size-fits all solution is insufficient. Finally, it is

clear that faculty do not reflect the diversity of undergraduate students., limiting the number of

288 students who can 'see themselves' represented among their instructors [7].

290 There are three key future directions that could build on our study. First, future work could use

different null models to test factors acting prior to undergraduate degrees (K-12 education, or

292 within the undergraduate years), or to consider variants on the career trajectory we considered

(e.g., removing the postdoctoral stage, allowing for time spent in industry jobs between academic

294 positions, or more explicitly modelling variation in the time spent in different stages). Second,

our approach would be greatly complemented by the collection and analysis of longitudinal

296 datasets (tracking the same individuals over time)[1]. For example, as definitions of race/ethnicity

change over time, scholars may move between race/ethnicity categories [30]. Non-U.S. born

298 scholars similarly change categories: they are not counted by race/ethnicity while they are

temporary residents (e.g. as students; [31]), but 'become' minorities with permanent residency.

300 Longitudinal data would also help distinguish between the possible scenarios of high input and

low retention versus low input and high retention. Third, future work could explore our research

302 questions at different scales. One could ask whether representation of scholars by race/ethnicity

varies across fields within S&E as is true for gender [16]. For example, Asian scholars are under-

304 represented in Ecology even as they appear overrepresented in S&E [32]. The category 'Asian' is

incredibly broad, masking a huge amount of diversity itself [33]; different scholars having very

306 different experiences based on cultural background and history [34]. Adopting an intersectional

---

1 Personal correspondence with Karen Hamrick (NSF-National Center for Science and Engineering Statistics) on January 5, 2021, indicating that longitudinal versions of the NSCG and SDR data are in development and are planned for future release.

perspective will almost certainly change our understanding of representation [35], with many

308   axes of identity (e.g. economic background, religion, disability, sexual orientation, gender, etc)

also impacting recruitment and retention [36]. Women of colour are especially likely to face

310   distinct challenges that can be masked by considering gender and race/ethnicity separately [37].

Finally, future work could attempt to project how long it would take to reach equity in the future

312   under varying social and policy scenarios. While this may seem like a simple extension of our

model, a simplistic forecast would be misleading at best, as predicting future dynamics requires

314   assumptions about future changes in academic labour pools and US demographics. However, our

model can be adapted to provide a framework for evaluating different future scenarios and policy

316   outcomes.

318   How then do we solve current mis-representation in academia? To create solutions we must draw

on social, cognitive and psychological frameworks to understand the factors contributing to mis-

320   representation [9,38], to explicitly address the alignment of cultural identities with STEM

identities [28] , and to guide both intervention programs and their metrics of success [39]. It is

322   also critical to recognize that a low relative representation in a stage can be due to problems that

accumulated across earlier stages [40]; thus low representation at a particular stage may not be

324   best served by intervention at that or the previous stage alone. Measuring whether these

interventions are working will require that demographic data is collected consistently and

326   transparently [41]. Where possible, data should be disaggregated to fully understand patterns. For

example, motivational factors can vary by racial/ethnic group [27] and likely also differ with

328   time spent in the US, especially in formative years [42] and with socio-economic and cultural

background. Data collection that consistently accounts for both race/ethnicity and

330   nativity/residence time will result in clearer understanding than current methods based on

residency categories. Finally, recruitment and retention must both be addressed [43]. Recruitment

332   into academia is not the only problem and thus a focus on increasing numbers of minority

undergraduates is not enough [9]. Individuals in under-represented versus well-represented

334   groups can have different reasons for pursuing career avenues and thus potentially different

reasons for leaving academia [44].

336

Although many academics wish to think of academia as unbiased and point to biases in earlier

338   stages and recruitment into academia itself as driving disparities in academia [8,10], our findings

indicate this is not the case: retention within academia is critical too. Furthermore, recruiting

340   under-represented scholars into a system (academia) that is not equipped to retain them is likely a

set up for all-around failure. These findings show that neither time, nor simple pushes to increase

342   recruitment are panaceas to this societal challenge. Models identifying the impacts and extent of

these biases (such as we have presented here) are a necessary part of developing and evaluating

344   solutions. However, we should not lose sight of the fact that numbers in our model represent real

people. Much work remains to address these representation problems in order to build an

346   academia that truly reflects and realizes the potential of the society it aims to serve.

352

Funding

354　The paper was funded in part by startup funds from the University of Minnesota.

356　**Data, code and materials**

No new data was collected for this study, all data used is publicly available and linked to from

358　the supplementary information. The model code and results generated for this study are available

via GitHub (https://github.com/allisonkshaw/academiamodel) and will be deposited in Data

360　Dryad upon manuscript acceptance.

362　**Authors contributions**

Conceptualization (AKS, NS, DES); Methodology and Analysis (AKS, CA, JMC, MV, YY, NS,

364　DES); Software (AKS); Data collection, compilation and curation (AKS, CA, JMC, MV);

Writing – Leading (AKS); Writing – Supporting (CA, JMC, TLM, MV, YY, NS, DES);

366　Visualization (all authors); Project Leadership (AKS).

1. American Federation of Teachers. 2010 Promoting racial and ethnic diversity in the faculty: What higher education unions can do. , 42.

2. Institute of Medicine. 2011 Expanding Underrepresented Minority Participation: America's Science and Technology Talent at the Crossroads.

3. National Science Foundation, National Center for Science and Engineering Statistics. 2019 Women, Minorities, and Persons with Disabilities in Science and Engineering: 2019. Special Report NSF 19-304.

4. Intemann K. 2009 Why Diversity Matters: Understanding and Applying the Diversity Component of the National Science Foundation's Broader Impacts Criterion. *Soc. Epistemol.* **23**, 249–266.

5. Ong M, Wright C, Espinosa L, Orfield G. 2011 Inside the Double Bind: A Synthesis of Empirical Research on Undergraduate and Graduate Women of Color in Science, Technology, Engineering, and Mathematics. *Harv. Educ. Rev.* **81**, 172–209. (doi:10.17763/haer.81.2.t022245n7x4752v2)

6. Hofstra B, Kulkarni VV, Munoz-Najar Galvez S, He B, Jurafsky D, McFarland DA. 2020 The Diversity–Innovation Paradox in Science. *Proc. Natl. Acad. Sci.* **117**, 9284–9291. (doi:10.1073/pnas.1915378117)

7. Shin JEL, Levy SR, London B. 2016 Effects of role model exposure on STEM and non-STEM student engagement. *J. Appl. Soc. Psychol.* **46**, 410–427. (doi:10.1111/jasp.12371)

8. Sethna BN. 2011 Minorities in higher education: A pipeline problem? *Res. High. Educ.* **13**, 1–18.

9. Callahan CN, LaDue ND, Baber LD, Sexton J, van der Hoeven Kraft, KJ, Zamani-Gallaher EM. 2017 Theoretical Perspectives on Increasing Recruitment and Retention of Underrepresented Students in the Geosciences. *J. Geosci. Educ.* **65**, 563–576. (doi:10.5408/16-238.1)

10. Whittaker JA, Montgomery BL. 2014 Cultivating Institutional Transformation and Sustainable STEM Diversity in Higher Education through Integrative Faculty Development. *Innov. High. Educ.* **39**, 263–275. (doi:10.1007/s10755-013-9277-9)

11. Bach DJ, Barnett MA, Fuentes JD, Frey SC. 2006 Promoting Intellectual Community and Professional Growth for a Diverse Faculty. *Improve Acad.* **24**, 166–182. (doi:10.1002/j.2334-4822.2006.tb00457.x)

12. Puritty C *et al.* 2017 Without inclusion, diversity initiatives may not be enough. *Science* **357**, 1101–1102. (doi:10.1126/science.aai9054)

13.   United States Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS). 2020 Bridged-Race Population Estimates, United States July 1st resident population by state, county, age, sex, bridged-race, and Hispanic origin. Compiled from 1990-1999 bridged-race intercensal population estimates (released by NCHS on 7/26/2004); revised bridged-race 2000-2009 intercensal population estimates (released by NCHS on 10/26/2012); and bridged-race Vintage 2019 (2010-2019) postcensal population estimates (released by NCHS on 7/9/2020). Available on CDC WONDER Online Database. Accessed at http://wonder.cdc.gov/bridged-race-v2019.html on Oct 9, 2020.

14.   Grunspan DZ, Eddy SL, Brownell SE, Wiggins BL, Crowe AJ, Goodreau SM. 2016 Males Under-Estimate Academic Performance of Their Female Peers in Undergraduate Biology Classrooms. *PloS One* **11**, e0148405.

15.   Borrego M, Knight DB, Gibbs K, Crede E. 2018 Pursuing Graduate Study: Factors Underlying Undergraduate Engineering Students' Decisions. *J. Eng. Educ.* **107**, 140–163. (doi:10.1002/jee.20185)

16.   Shaw AK, Stanton DE. 2012 Leaks in the pipeline: separating demographic inertia from ongoing gender differences in academia. *Proc. R. Soc. Lond. B Biol. Sci.* **279**, 3736–3741. (doi:10.1098/rspb.2012.0822)

17.   Gibbs KD, Basson J, Xierali IM, Broniatowski DA. 2016 Decoupling of the minority PhD talent pool and assistant professor hiring in medical school basic science departments in the US. *eLife* **5**, e21393. (doi:10.7554/eLife.21393)

18.   National Science Foundation, National Center for Science and Engineering Statistics. 2015 Science and Engineering Degrees: 1966–2012. Detailed Statistical Tables NSF 15-326.

19.   National Center for Science and Engineering Statistics. 2018 Survey of Graduate Students and Postdoctorates in Science and Engineering.

20.   National Science Board, National Science Foundation. 2019 Higher Education in Science and Engineering. Science and Engineering Indicators 2020. NSB-2019-7.

21.   National Science Board. 2018 Science and Engineering Indicators 2018. NSB-2018-1.

22.   National Science Foundation, Division of Science Resources Statistics. 2008 Postdoc Participation of Science, Engineering, and Health doctorate Recipients.

23.   National Center for Science and Engineering Statistics, National Science Foundation. 2019 Doctorate Recipients from U.S. Universities: 2018. Special Report NSF 20-301.

24.   Durodoye R, Gumpertz M, Wilson A, Griffith E, Ahmad S. 2020 Tenure and Promotion Outcomes at Four Large Land Grant Universities: Examining the Role of Gender, Race, and Academic Discipline. *Res. High. Educ.* **61**, 628–651. (doi:10.1007/s11162-019-09573-9)

25. Kahn S, MacGarvie M. 2020 The impact of permanent residency delays for STEM PhDs: Who leaves and why. *Res. Policy* **49**, 103879. (doi:10.1016/j.respol.2019.103879)

26. Finn MG. 2010 Stay Rates of Foreign Doctorate Recipients from U.S. Universities, 2007.

27. McGee EO, White DT, Jenkins AT, Houston S, Bentley LC, Smith WJ, Robinson WH. 2016 Black engineering students' motivation for PhD attainment: passion plus purpose. *J. Multicult. Educ.* **10**, 167–193. (doi:10.1108/JME-01-2016-0007)

28. Collins KH. 2018 Confronting Color-Blind STEM Talent Development: Toward a Contextual Model for Black Student STEM Identity. *J. Adv. Acad.* **29**, 143–168. (doi:10.1177/1932202X18757958)

29. Allaire FS. 2019 Navigating Uncharted Waters: First-Generation Native Hawaiian College Students in STEM. *J. Coll. Stud. Retent. Res. Theory Pract.* **21**, 305–325. (doi:10.1177/1521025117707955)

30. Tienda M, Mitchell F. 2006 *Multiple Origins, Uncertain Destinies: Hispanics and the American Future*. National Academies Press.

31. Byrd WC, Dika SL, Ramlal LT. 2013 Who's in STEM? An Exploration of Race, Ethnicity, and Citizenship Reporting in a Federal Education Dataset. *Equity Excell. Educ.* **46**, 484–501. (doi:10.1080/10665684.2013.838485)

32. Kou-Giesbrecht S. 2020 Asian Americans: The Forgotten Minority in Ecology. *Bull. Ecol. Soc. Am.* **101**, e01696. (doi:10.1002/bes2.1696)

33. Ing M, Victorino C. 2016 Differences in Classroom Engagement of Asian American Engineering Students. *J. Eng. Educ.* **105**, 431–451. (doi:10.1002/jee.20126)

34. Kelly KR, Gunsalus A-JC, Gunsalus R. 2009 Social Cognitive Predictors of the Career Goals of Korean American Students. *Career Dev. Q.* **58**, 14–28. (doi:10.1002/j.2161-0045.2009.tb00170.x)

35. Gayles JG, Smith KN. 2018 Advancing Theoretical Frameworks for Intersectional Research on Women in STEM. *New Dir. Institutional Res.* **2018**, 27–43. (doi:10.1002/ir.20274)

36. Metcalf H, Russell D, Hill C. 2018 Broadening the Science of Broadening Participation in STEM Through Critical Mixed Methodologies and Intersectionality Frameworks. *Am. Behav. Sci.* **62**, 580–599. (doi:10.1177/0002764218768872)

37. Corneille M, Lee A, Allen S, Cannady J, Guess A. 2019 Barriers to the advancement of women of color faculty in STEM: The need for promoting equity using an intersectional framework. *Equal. Divers. Incl. Int. J.* **38**, 328–348. (doi:10.1108/EDI-09-2017-0199)

38.   Estrada M, Woodcock A, Hernandez PR, Schultz PW. 2011 Toward a model of social influence that explains minority student integration into the scientific community. *J. Educ. Psychol.* **103**, 206–222. (doi:10.1037/a0020743)

39.   Margherio C, Horner-Devine MC, Mizumori SJY, Yen JW. 2016 Learning to Thrive: Building Diverse Scientists' Access to Community and Resources through the BRAINS Program. *CBE—Life Sci. Educ.* **15**, ar49. (doi:10.1187/cbe.16-01-0058)

40.   Nelson RG, Rutherford JN, Hinde K, Clancy KBH. 2017 Signaling Safety: Characterizing Fieldwork Experiences and Their Implications for Career Trajectories. *Am. Anthropol.* **119**, 710–722. (doi:10.1111/aman.12929)

41.   Minefee I, Rabelo VC, Stewart IV OJC, Young NCJ. 2018 Repairing Leaks in the Pipeline: A Social Closure Perspective on Underrepresented Racial/Ethnic Minority Recruitment and Retention in Business Schools. *Acad. Manag. Learn. Educ.* **17**, 79–95. (doi:10.5465/amle.2015.0215)

42.   Roysircar G, Carey J, Koroma S. 2010 Asian Indian College Students' Science and Math Preferences: Influences of Cultural Contexts. *J. Career Dev.* **36**, 324–347. (doi:10.1177/0894845309345671)

43.   Bhalla N. 2019 Strategies to improve equity in faculty hiring. *Mol. Biol. Cell* **30**, 2744–2749. (doi:10.1091/mbc.E19-08-0476)

44.   Layton RL, Brandt PD, Freeman AM, Harrell JR, Hall JD, Sinche M. 2016 Diversity Exiting the Academy: Influential Factors for the Career Choice of Well-Represented and Underrepresented Minority Scientists. *CBE—Life Sci. Educ.* **15**, ar41. (doi:10.1187/cbe.16-01-0066)

370

372 **Figure 1.** Model schematic of academia as a series of stages, where individuals either move to

the next stage, or move outside of the system (academia) to other career paths. Black text

374 indicates NSF data, blue text indicates estimated data. The stages for graduate students (*G*) and

tenured professors (*T)* are split into sub-partitions (grey lines), representing pre- and post-exam

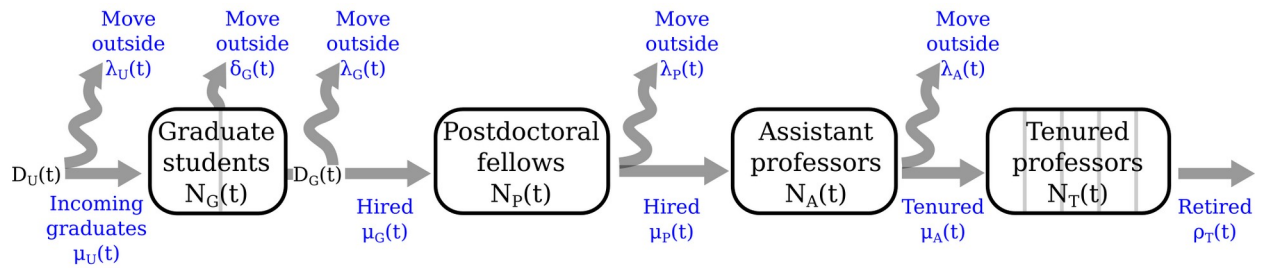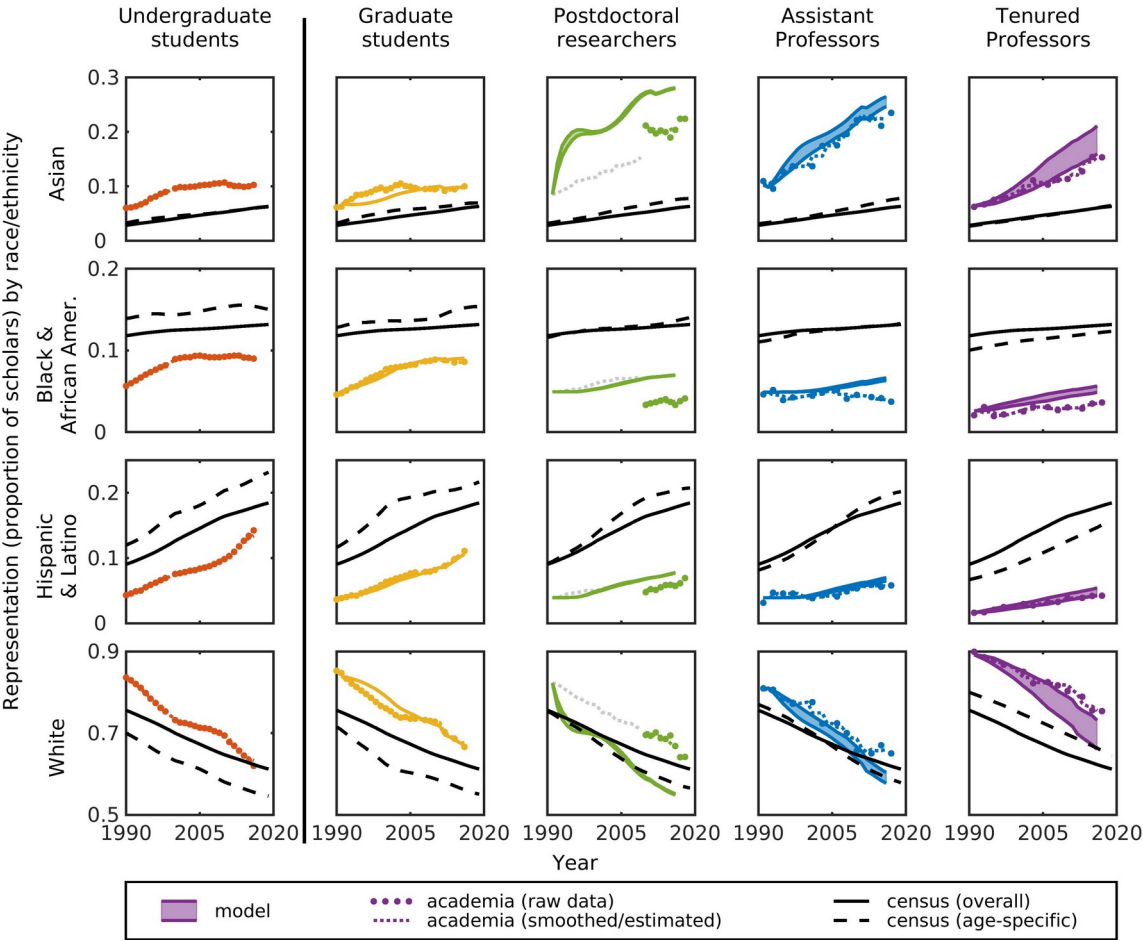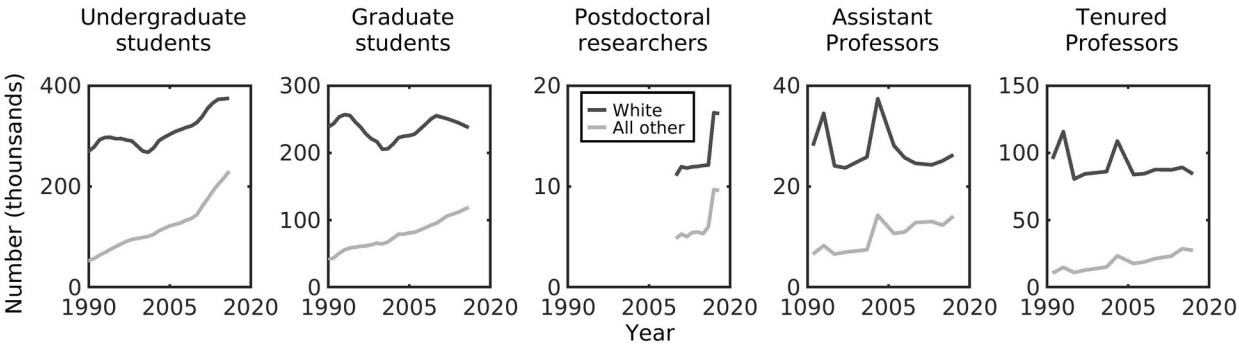376 stages for graduate students and equally spaced intervals for tenured professors.

Figure 2. The representation (i.e., the proportion of scholars in that stage that identify as that

380   race or ethnicity) of the four largest American race/ethnicity categories (rows) in each academic

stage (columns) over time comparing: null model predictions (coloured solid lines), academia

382   data (dots are raw data, dotted lines are smoothed data), and census data for the U.S. overall

population (black solid lines) and US age-specific population (black dashed line). Mismatch

384   between model and academia data indicate race/ethnicity-based biases of retention within

academia, mismatch between model and census indicates race/ethnicity-based biases in

386   recruitment into academia. Postdoc data before 2010 was unavailable, was estimated as the

average of the graduate student and assistant professor data, and is greyed out in the figure. (See

388   Figure S7 for additional race/ethnicity categories)

390

392   **Figure 3.** The absolute number (in thousands) of scholars that are White (solid black line) and all

other races/ethnicities (solid grey line) in each stage (panel) over time.

394

396   **Figure 4.** The relative representation (θ [eqn 1]; comparing data and the null model) over 15

years (1991-2016) of each race/ethnicity category through one of the transitions within

398   academia: undergraduate to graduate student (*U* to *G*), graduate student to postdoctoral

researcher (*G* to *P*) or to assistant professor (*G* to *A*), and assistant to tenured professor (*A* to *T*).

400   Positive or negative values indicate a race/ethnicity category faces correspondingly positive or

negative bias across that transition. Confidence intervals mark the range of θ values that result

402   from a 5% increase or decrease in representation in either the data or model.