NBER WORKING PAPER SERIES

THE COLOR OF IDEAS:
RACIAL DYNAMICS AND CITATIONS IN ECONOMICS

Marlène Koffi
Roland Pongou
Leonard Wantchekon

Working Paper 33150
http://www.nber.org/papers/w33150

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2024

We thank Joshua Angrist, Cynthia Bansak, Patrick Bayer, Matthew Notowidigdo, Zach Ward, participants at the ASSA 2024, the 2024 Africa Meeting of the Econometric Society, the Conference on Race, Racism, and Structural Inequality, the 2024 annual meeting of the Canadian Economics Association, the 2024 annual meeting of Société canadienne de science économique, and the 2024 AERIP/ASHE/NEA Summer Conference, and seminar participants at York University and Duke University, for insightful feedback. We thank Christopher Snyder and Mark McCabe for sharing their data on JSTOR online availability with us. We also thank Landon Liu and Ashna Zaman for their excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Color of Ideas: Racial Dynamics and Citations in Economics
Marlène Koffi, Roland Pongou, and Leonard Wantchekon
NBER Working Paper No. 33150
November 2024
JEL No. A14, I23, J15

ABSTRACT

This paper investigates the existence of racial disparities in the dissemination of ideas using the paper citation network in economics. Exploiting a comprehensive dataset of over 330,000 publications from 1950 to 2021, combined with manually collected data from the CVs of thousands of economists, we document that papers authored by non-White scholars (Black, Hispanic, or Asian) receive 5.1% to 9.6% fewer citations than those authored by White scholars. The citation gap remains or even amplifies with increasing author seniority and conventional quality indicators and is especially pronounced for Black authors. Moreover, papers authored by non-White scholars are less likely to serve as citation bridges and are less often cited by highly cited papers as measured by the centrality indexes, limiting both their direct and indirect influence. Our analysis indicates that this disparity is not attributable to differences in research quality, author ability, or visibility. Rather, it is largely driven by homophily in citation patterns and racial clusters in networks, where scholars tend to cite authors from their racial group. These findings can be rationalized by a simple theoretical model where citation costs and peer-review preferences influence citation behavior. Then, we provide suggestive evidence that reducing information friction—thereby lowering the cost of citing—could reduce the racial citation gap by up to 50%. Finally, using natural language processing, we highlight the complementarity across racial groups in research and discuss potential losses from racial barriers to idea diffusion.

Marlène Koffi
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 3G7
and NBER
marlene.koffi@utoronto.ca

Roland Pongou
University of Ottawa
120 University
Ottawa, ON K1N 6N5
Canada
rpongou@uottawa.ca

Leonard Wantchekon
Department of Politics
Princeton University
321 Bendheim Hall
Princeton, NJ 08544
and NBER
lwantche@princeton.edu

1 Introduction

Racial inequality in education and labor markets remains a significant issue around the world. In the United States, for instance, as of 2022, only 25% of Hispanics in the labor force held a bachelor's degree or higher, compared to 34% of Blacks, 44% of Whites, and 68% of Asians. Additionally, African Americans face an unemployment rate nearly twice that of Whites (6.1% compared to 3.2%), while both Hispanics and Blacks continue to earn significantly less than their White and Asian counterparts.[1] These broader social disparities are also evident in academia, despite its status as a field requiring high levels of skill and selective entry. In 2022, Black and Latino scholars comprised only 13% of the academic workforce, compared to 72% of White scholars, with even greater underrepresentation in STEM fields (Kozlowski et al. [2022]). This raises the question of whether racial disparities in academia persist beyond the entry stage.

This paper addresses this question by examining racial disparities in the diffusion and dissemination of ideas within economics. Specifically, we document patterns and dynamics of citations as they relate to racial categories. Analyzing citation flows is important for several reasons: citations are used to quantify research quality and influence, as well as the knowledge transmission (Angrist et al. [2020]). In addition, citations predict important career outcomes, including salary (Hamermesh et al. [1982], Hilmer et al. [2015]), employment at prestigious institutions (Ellison [2013]), grants, and prestigious awards (Fiala and Tutoky [2017], Card et al. [2020, 2022], Hager et al. [2023]). Moreover, focusing on a single discipline, such as economics, allows for a more precise and nuanced understanding of the dynamics and mechanisms behind the propagation of scientific ideas. But also, the field of economics is particularly illustrative regarding diversity issues. For example, PhDs in economics are among the least diverse in STEM fields in terms of ethnic and racial representation (Bayer and Rouse [2016]). This has raised growing concerns about the persistent underrepresentation of minority groups in the economics profession, as highlighted in panel discussions and articles presented at meetings of the American Economic Association (AEA) and the Econometric Society.[2] Yet, very little large-scale and rigorous evidence of racial disparities and inequitable practices exist in the profession.[3]

We present a series of analyses aimed at addressing this gap in the literature. First, we undertake an important data collection effort, combining existing structured data and hand-collected data. Specifically, we use publication records and bibliometric information from over 200 economics journals, encompassing approximately 330,000 publications from 1950 to 2021. This dataset represents 221,249 economists. We also manually collect and extract information from the curricula vitae of nearly 4,500 scholars. Using a combination of algorithmic procedures and manual verification, we identify the race of each author.[4] Based on this, we infer the racial composition of each team of

[1] https://www.bls.gov/opub/reports/race-and-ethnicity/2022/.

[2] Allied Social Science Association (ASSA) annual meetings, January 3, 2020, in San Diego "How Can Economics Solve Its Race Problem?"; Econometric Society 2020 world congress.

[3] Koffi and Wantchekon [2024] present evidence of the under-representation of Africans in the NBER.

[4] We acknowledge that this method does not account for self-identification, which may differ from external perceptions of race based on visual cues and information about origins.

authors. We then construct a detailed set of control variables by leveraging article characteristics and creating an annual publication profile for each author in the database. Exploiting the full bibliographic references of articles, we refine the analysis and categorize papers by journals of publications, affiliations, fields, and methodological approaches. We also build citation flows over time and cumulative citations received up to a given year for each publication and author, providing a robust measure of the impact and perceived quality of a researcher's work. Furthermore, by analyzing the acknowledgments sections of articles and conducting an extensive manual data collection of the CVs of authors publishing in the top five journals over time, resulting in nearly 100,000 observations, we extract information on attendance at seminars and conferences to construct a measure of authors' active outreach efforts. This comprehensive data collection strategy helps mitigate concerns about omitted variable bias, particularly in the context of racial analysis, where not all inequalities necessarily reflect bias.

We find evidence of a racial citation gap in economics, despite significant progress in the racial diversity of authorship over time (in 2021, just over half of papers were authored exclusively by White authors, compared to nearly 95% in the 1950s). The estimated average unconditional citation gap is 13.9 log points. After controlling for observable paper characteristics—such as the number of authors, field, journal, year of publication, and author prominence—papers with at least one non-White author (Black, Hispanic, or Asian) receive 5.1 log points fewer citations than papers authored solely by White individuals. We replicate our analysis by focusing only on papers published in top journals to address the concern that the racial citation gap may be driven by lower-quality publications. Restricting the sample to articles published in a set of 35 high-impact journals (as identified by Card et al. [2020]), 16 high-impact journals (as identified by Koffi [2021b]), or the traditional top five journals, we continue to observe a persistent and, in some cases, larger, citation gap. For instance, the gap for the top five journals is 9.6 log points, compared to 5.1 log points in the original full sample. We further disaggregate the analysis by non-White racial categories, revealing that Asian researchers experience a substantial citation penalty, estimated at 7.1 to 8.5 log points in the top five and 16 high-impact journals, respectively while the penalty is 7 and 14.7 log points for Hispanic authors. Yet, the largest citation penalty is experienced by Black researchers, where the gap is approximately double that of Asian researchers in both the top five and 16 high-impact journals.

Our analysis also indicates a significant heterogeneity in the racial citation gap along several dimensions: it varies widely across affiliations, methodological approaches, and subfields. The gap is nearly twice as large for papers authored by individuals affiliated with top-tier institutions compared to those from mid- or low-tier institutions (12 log points versus 6 log points). It is almost negligible in empirical fields but more pronounced in areas such as microeconomics and international economics. Encouragingly, the racial citation gap has narrowed over time, standing at just over 3% in the most recent decade.

We conduct several robustness checks, beginning with an examination of the role of unobservables. While we cannot entirely dismiss the potential influence of unobservables in lower-ranked journals, this hypothesis is strongly challenged when considering the top five journals.

Next, the racial citation gap may, in part, be explained by the concentration of non-White scholars in fields that tend to receive fewer citations or by their focus on research topics that attract less attention from other researchers. To address the argument that the racial citation gap may result from differences in research areas and topics across racial groups, we introduce a large set of fixed effects using two-digit and three-digit JEL codes. To further refine our analysis, we apply natural language processing techniques to the abstracts of the papers and extract a set of 768 continuous value vectors. This method effectively captures topic-level differences, as shown in Koffi [2021b]. If anything, our results indicate that incorporating granular field fixed effects into the model increases the estimated baseline coefficient by up to 40%. Moreover, the citation penalty persists when using alternative definitions of the team's racial composition or when controlling for text readability, which also accounts for whether the author is a native English speaker or not, which might impact the text's complexity. The penalty also remains robust when using alternative definitions of the citation variable, such as top percentile, citation counts, or excluding zero citations.

While analyzing citation counts is important, it fails to capture other dimensions of a paper's scholarly relevance. As an alternative approach, we assess research dissemination by incorporating centrality measures. Centrality highlights a paper's significance within a citation network, with one class of measures evaluating a node's direct and indirect links (such as eigenvector, Katz, PageRank, and degree centrality), and another focusing on its pivotal position within the network (betweenness centrality). Across all five measures, we find that papers authored by non-White scholars score lower, suggesting they are not only less cited overall but also less likely to serve as bridges between other papers and less frequently cited by highly influential works. Additionally, we examine another centrality measure, the clustering coefficient, which captures how likely a paper's neighbors are to cite each other, forming a tightly knit group. Our analysis shows that papers authored by non-White scholars exhibit higher clustering coefficients, indicating that these papers are more often located within cohesive subgroups of the citation network. This suggests a more limited dissemination of their ideas.

Overall, our findings could be unexpected, particularly given the common assumption that scholars affiliated with prestigious institutions and publishing in top journals typically produce higher-quality research or enjoy greater visibility. To further explore the sources of the citation penalty faced by non-White-authored papers, we investigate several potential explanations, including racial differences in paper quality and authors' ability, homophily, bias, and information cost.

We examine the role of differential author ability through several approaches. First, we discuss the possibility that affirmative action policies, or the perceived application of such policies, could account for the citation gap. However, given the strong citation gap observed in the 1950s and the fact that Asian authors also experience this penalty (despite affirmative policies being perceived by detractors as favoring Black and Latinx scholars), we find this explanation unlikely to account for the gap. Second, we examine the effect of publication history. In traditional belief-based models, accumulating more signals of ability should reduce the original gap as agents update their priors. We find that previous publications only reduce the citation gap for authors at the extreme end of the publication record distribution. Specifically, only non-White authors with at least 20 overall publications, or at least 11 publications in the top 35 or top 16 high-impact journals and 6 publications in the top 5 journals, experience a significant reduction in the citation gap compared to their White counterparts. Third, we use authors' citation history at the time of publication as a proxy for author quality. While the total number of citations accumulated by the authors prior to a paper's publication positively influences citations, controlling for this variable does not eliminate the citation gap. Similarly, an analysis controlling for the cumulative citations of both the paper and the authors before a given year still reveals a persistent gap, despite the endogeneity of those controls and their likelihood to induce a downward bias estimate. Therefore, our analysis does not support the hypothesis that the citation penalty faced by papers with at least one Black, Hispanic, or Asian author is driven by differential author ability.

Next, we find that the racial citation gap is entirely driven by homophily in citation patterns. Our results reveal a pronounced homophilic trend: papers authored by at least one Black, Hispanic, or Asian scholar receive a 23% citation premium from other non-White scholars but face a 30% citation penalty from White scholars. This homophily effect fully explains the previously observed variations across fields and over time. Importantly, this effect is not attributable to geographical factors, as the patterns hold for both within-continent and across-continent citations.

To explore the sources of this persistent behavior, we support our findings with a simple theoretical model that examines the citation behavior of an author navigating a multiethnic scholarly environment while facing a random referee. In this setting, scholars tend to be more familiar with works from other scholars from their own racial group, making it less costly to cite co-ethnic authors and easier to recognize the value of their work. The model predicts a homophilic citation pattern of similar magnitude across ethnic groups if citation costs (such as the cost of acquiring information or bias) are the only influencing factor. However, homophilic citation becomes less pronounced for minority groups when the possibility of selecting referees and referee preferences for diversity are considered. Nonetheless, in both scenarios, papers authored by minority scholars receive fewer citations than papers of comparable quality authored by scholars of comparable ability from the majority group. Further analysis of the model reveals interesting heterogeneities in the racial citation gap. Specifically, a comparative static analysis shows that the gap is more pronounced for higher-quality publications and among higher-ability authors. Our empirical findings align with all of these theoretical predictions.

We investigate the role of differential author visibility in contributing to the racial citation gap. Increased visibility can enhance citation rates by reducing information costs. To assess this, we use our extensive hand-collected CV database to control for a wide range of variables that likely impact an author's visibility and network reach. These variables include total seminar and conference counts, attendance at prestigious conferences, seminar invitations from top-ranked institutions, visiting positions at top-tier institutions, editorial roles, and membership in prestigious organizations (such as NBER, JPAL, CEPR, BREAD, IZA, and CIFAR), as well as holding a PhD from a top-ranked institution. We find that seminars and conferences play a limited role in mitigating the citation gap, primarily because they are highly correlated with author prominence, a factor already accounted for in our empirical framework. Together, those factors explain only 17% of the variation in the racial citation gap, with the majority of the variation stemming from racial disparities in membership within prestigious associations.

To differentiate between genuine information costs and bias as potential sources of the citation gap as highlighted by the model, we examine the role of collaboration with non-coethnic authors as well as the impact of online platforms like JSTOR, in improving awareness and access to research papers. We find that collaborating with a non-coethnic author tends to shape the perceptions and increases citation to non-coethnic groups. Similarly, we also uncover suggestive evidence that the introduction of online platforms such as JSTOR has significantly reduced the homophily effect, particularly decreasing the citation penalty imposed by White scholars on papers authored by non-White scholars by as much as 50%.

Overall, our paper demonstrates that reducing information friction between groups can connect the different racial networks and expand the dissemination of work authored by scholars across racial categories.

Finally, we discuss the potential losses to scientific progress arising from racial barriers in the diffusion of ideas. By applying natural language processing techniques, we show that the research produced by scholars from different racial groups exhibits some degree of complementarity. Such complementarity—where ideas from diverse backgrounds enrich and build upon each other—can drive innovation and produce better science. Therefore even aside from ethnic collaborations as in Freeman and Huang [2015], reducing racial barriers to the diffusion of ideas could amplify scientific progress and its applications in policy and economics.

2 Related literature

Our paper contributes to the literature on identity-based inequalities in academic research. It is among a handful of studies that systematically examine how the dissemination of academic ideas and contributions varies by the racial characteristics of authors and document possible explanations both from the citing and the cited authors. Within the field of economics, while the literature on gender inequalities in academic research has expanded significantly in recent years, there is much less information on racial inequalities.

Two papers have examined the racial citation gap, but they mainly focus on the economics of race and crime. Mason et al. [2022] find that articles published in top economics journals on topics related to race and crime (e.g., identity, police use of force, and mass incarceration) by Black scholars receive fewer citations than similar articles authored by non-Black scholars. They also find that articles are cited less if published in *The Review of Black Political Economy* (RBPE)—the flagship journal of the National Economic Association (the organization of Black economists in the United States)—than if published in other journals. These findings are confirmed in a companion paper by the same authors (Mason et al. [2023]). In the latter, they also find that articles authored by Black scholars or published in RBPE are more likely to document the presence of racial discrimination in various outcomes. Interestingly, research in the crime literature authored by Black scholars or published in RBPE is undervalued by mainstream economics, and despite evidence that Black economists have persistently engaged on economic policy issues related to race and other relevant topics such as poverty, education, and public finance (Francis et al. [2022]). Our scope, analysis, and findings significantly differ from those of these studies. We include articles on all the topics covered in economics and focus on all non-White groups.

Other works on racial and ethnic inequalities within the economics profession examine the underrepresentation of minority scholars and scholars living in developing countries. While bias exists due to institutional barriers, Bayer and Rouse [2016] argue that women and minority groups are underrepresented in the field of economics as a whole. Analyzing surveys conducted among Black, Latinx, and Native American people at various stages along the economics career trajectory, Bayer et al. [2020] find that bias, hostile climate, and lack of information and good mentoring are among factors that explain the underrepresentation of members of minority groups in economics. Koffi et al. [2024] document racial inequalities in economic publishing and find notable differences in authorship between White and non-White authors, with Black, Hispanic, and Asian economists being significantly underrepresented relative to their proportion of the overall population. They find the authorship underrepresentation of non-White individuals to be the highest for Black economists. Amarante et al. [2022] document evidence showing that researchers from developing countries—a main supplier of minority scholars within the economics profession—are significantly underrepresented in research on development economics in terms of conference presentation, publication in top development journals, and citation of published papers. Importantly, the underrepresentation of scholars from developing countries in research is not unique to economics; Thalmayer et al. [2021] find that less than 1% of both lead researchers and research samples in top psychology journals are from Africa. Alarmed by these observations, Mughogho et al. [2023] argue that the underrepresentation of scholars from developing countries in behavioral research limits the validity of scientific evidence and undermines the ability of policymakers to design sound policies for these countries. Greenspon and Rodrik [2021] reach similar conclusions regarding the global distribution of authorship in economics. They find that the representation of developing countries authors in top economics journals is extremely low and has not significantly improved over time. In this paper, we document the increase in authorship from racially diverse profiles over time. However, despite this upward trend, we show that the recognition of their work through citations remains insufficient.

Our paper is also linked to research documenting citation patterns in economics. Hamermesh [2018] finds that the fraction of citations to empirical work in economics from top journals has significantly increased since 1990. Lutmar and Reingewertz [2021] provide insight into general citation counts for articles written by in-group members versus out-group members, where in-group papers share the journal's institutional affiliation. They find a level of academic in-group bias in economics in some top 5 journals. Other works have focused on the insularity of economics and its influence on other disciplines. Using bibliometric data from 1995-1997, Pieters and Baumgartner [2002] argue that economics is an insular discipline in that economics journals cite very few papers from other social sciences. Based on bibliometric data from 2000-2009, Fourcade et al. [2015] reach a similar conclusion; they analyze citation flows in top journals in economics, political science, and sociology and find that economics is more insular than political science and sociology. Angrist et al. [2020] look at a broader set of economics journals, and find that while economics is more insular than political science and sociology, it is less insular than psychology and is comparable to anthropology. Our research adds to this observation by showing that the racial gap in citations is independent of the journal quality. In general, we find a significant racial gap in scholarly influence and document the characteristics of papers that experience high racial citation penalty.

Our research complements the growing literature on gender inequalities in the economics profession. Using bibliometric data from papers published in top economics journals, Koffi [2021a] finds that female-authored papers receive fewer citations from top-tier journals and from men than similarly situated papers authored by men. Koffi [2021b] uses machine learning to establish similarities between papers and construct relevant counterfactual citations, and finds that papers receive 10% fewer citations if authored solely by women than if authored by men; although most of the variation in citation was due to the author publication history. She also finds a strong in-group bias where female authors cite more male authors and vice-versa. Card et al. [2020] find that papers written by women are more likely to be rejected by editors and referees than if they were seeking to maximize citations. Using data on interactions between presenters and their audience in seminars in leading economics departments and conferences, Dupas et al. [2021] find that women presenters are treated differently compared to similarly situated men; not only do women receive more questions than men during a seminar, but questions asked of women tend to be more hostile. Other studies showing unfavorable outcomes for women economists focus on turn-around times for papers submitted to economics journals (Hengel [2022]), prestige rankings (Zacchia [2021]), career trajectory, tenure and promotion (Ginther and Kahn [2004], Sarsons [2017], Chen et al. [2022]), and treatments on online forums (Wu [2020]). Comparable studies on racial inequalities in the economics profession have yet to be conducted.

Racial inequalities have also been documented outside of economics (Kozlowski et al. [2022]). Much of the focus has been on racial bias within other STEM fields where two key points regarding this topic in existing literature are observed. Bertolero et al. [2020] document racial and ethnic imbalances in the reference lists of papers published in the top five neuroscience journals over the last 25 years. Reference lists tended more often to include papers with a White person as the first and last author. The authors suggest that the amount by which this is observed is far too large for there to be no link between race and ethnicity when referencing. Other works document the existence of implicit biases in the hiring process and funding. Eaton et al. [2020] highlight the racial biases involved in post-doctoral candidate evaluation in the fields of physics; White and Asian candidates were regarded as more competent and hireable than Black and Latinx candidates. Also, researchers of color may get lower funding than White researchers. Taffe and Gilpin [2021] and Ginther et al. [2011] show that grant applications submitted to the National Institute of Health (NIH) in the U.S. were less likely to fund Black Principal Investigators (PIs) than White PIs. Black applicants reported fewer papers, had fewer citations, and appeared in journals with lower impact factors, factors that influence their ability to obtain funding, implying there may be an inherent racial barrier. Qiu et al. [2022] find that Chinese authors receive fewer citations from US researchers. Our paper significantly contributes to this literature by advancing the understanding of the sources of this gap and proposing alternative solutions to address it.

3 Data description

3.1 Data collection

We collected publication records from more than 200 journals using the Web of Science (WoS) database, Econlit, and the Microsoft Academic Graph (MAG) database—sources widely recognized as some of the largest repositories of academic research in economics. WoS serves as the core database, Econlit provides information on Journal of Economic Literature (JEL) codes, and MAG assists with name disambiguation.

The collected information was subsequently organized into a new database. The sample includes the top five general-interest journals widely recognized in the field, namely, *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economics Studies*, along with other leading field and high-impact journals. The complete list of journals is provided in Appendix Table F.

The resulting database comprises full-length articles and excludes various types of content, including proceedings papers, comments, articles shorter than three pages, book reviews, bibliographical items, editorial material, letters, and corrections. The databases were then merged using the Digital Object Identifier (DOI) or other attributes, such as title, journal name, publication year, first-page number, or abstract. To facilitate this process, the variables related to the journal name and article title were standardized to lowercase, and all punctuation and white spaces were removed. Duplicate articles were eliminated through manual review.

In total, the database contains 329,338 papers. Each paper is accompanied by a range of associated information, including the title, journal of publication, publication date, authors' names, authors' institutions, number of citations, DOI, abstract, keywords, JEL codes, and references. The citation data was downloaded between March 2021 and May 2021.

3.2 Race coding

3.2.1 Manual and algorithmic predictions

The determination of the author's race in this study involves a combination of automated procedures, particularly using U.S. Census data and the *ethnicolr* package, as well as manual data collection efforts. The *ethnicolr* package is based on the work of Sood and Laohaprapanon [2018], who used U.S. Census data and Florida Voter Registration data to construct a race classifier based on last names and/or first names.

The race determination protocol is outlined as follows. First, the authors' names were retrieved from the bibliographic data. To minimize the risk of errors, we initially use the prediction based on the last name alone, followed by a prediction using both the last name and first name. The following conditions were imposed: (1) agreement between the two predictions, and (2) the probability of assigning a given racial label to an individual must be higher than 0.7. We also present alternative specifications based on other cutoffs and conditions.

As shown in Hofstra et al. [2020], the algorithm is quite accurate for identifying Whites and Asians (and, to some extent, Hispanics) but less so for Blacks. In fact, the Black category has the highest error rate due to name assimilation resulting from colonial and slavery heritage. To address this limitation, we supplemented the data with a list of Black economists from the National Economic Association. We manually verified and traced the networks (e.g., co-authors, Ph.D. universities) of these Black economists to identify additional Black economists.

Overall, the database includes 221,249 authors. Race assignments were achieved for approximately 80.5% of the authors in the dataset, among which 70.2% were identified as Whites, 23.8% as Asians, 5.4% as Hispanics, and the remaining approximately 0.6% as Blacks.

Furthermore, due to the skewed racial composition of teams, with more White authors and fewer non-White authors, we adopted a conservative measure of team racial diversity. Specifically, the racial composition of each co-author team was categorized into three distinct groups: "All-White," indicating that all co-authors are White; "At least one Black/Hispanic/Asian," indicating that the co-authorship includes at least one non-White author; and "Other races," for cases where at least one author's race could not be classified into the conventional categories. This latter category typically includes individuals of Arabic or Persian descent or those with mixed perceived ethnicities whose names do not clearly align with a single category. We also present results using alternative definitions of team racial composition.

3.2.2 Measurement errors

Race coding based on last names is likely to capture racial likelihood or perceived race. While imposing cutoff constraints may increase the alignment between likelihood and actual race, it does not eliminate uncertainty or the potential for measurement errors.

To address this, we manually collected data from the CV of a subsample of the authors to verify whether the findings based on algorithmic predictions are, on average, consistent with those based on hand-checked data (see section 3.3 below for details). These results are presented in the robustness section of the paper.[5] Therefore, throughout the paper, we employ a simplified terminology and refer to an author's predicted race as their race.

3.3 Hand-collected CV data

Acknowledging that authors' backgrounds and visibility through visiting positions, seminars, and conferences may have an impact on citations, we conducted a thorough manual collection of the curriculum vitae of authors publishing in the top journals. We focused on authors with at least one top-five publication since 1990. There were 5,150 authors in this category. From the hand-collected CVs, we extracted information on demographic background and characteristics including date of birth, marital status, and number of children at the time of data collection, whenever available. We also gathered data on PhD, MA, and bachelor's degree-granting institutions and the year each degree was granted. Whenever possible, information about the names of PhD advisors was also gathered. Additionally, we collected annual data on the different institutional affiliations of each author since the year of PhD completion. More precisely, data was gathered on main affiliations, visiting appointments, and membership in esteemed research organizations such as the NBER, JPAL, CEPR, BREAD, IZA, and CIFAR. Moreover, we collected annual information on each author's academic rank, editorial activities, and rank at every journal with which they are associated. We further compiled details regarding attendance at seminars and workshops, and annual data about fellowships and awards received since obtaining the PhD. Finally, we manually verified each author's gender and race. Following every author every year since their PhD was completed resulted in almost 100,000 observations.

3.4 Other variables

Aside from race, several key constructed variables in this study merit discussion.

To obtain granular information on citation patterns, we collected the bibliographical references included in each published paper. We can then track the citations over time, and match the characteristics of the citing paper to those of the cited paper. JEL code information is available in Econlit, although not for all articles. For some articles, text information matched to JEL codes was available. We used the consortium mapping of JEL codes to JEL descriptions from the AEA website to match JEL descriptions, as used in the EconLit database, to the JEL codes. However, the field classification was still missing for about a third of the data. To address this, we used Koffi [2021b]'s field classification algorithm to predict the primary field of the remaining articles.[6]

[5] Even when the predicted probability matches the perceived race, there remains a distinction between perceived race and self-identification, which is not captured in our analysis.

[6] The algorithm is based on text classifications using information from the title, abstracts, and keywords. We also employed a similar classifier to identify the methodological style of the paper. Readers can refer to Koffi [2021b] for

The article database was converted to track each researcher's publications. To match the three databases at the author level, conditional on an article-level match, we used authors' last names and positions. A disambiguation procedure, similar to that in Card et al. [2022], was employed. A yearly publication history for each author was then compiled, enabling the creation of variables representing their publication trajectories, including lifetime and date-specific publications. Article-level metrics such as the most prolific author, average publications per author, and total publications across journal tiers were calculated. The first publication date for each author was also tracked to identify the most senior author on a paper.

Author institutions were identified following a procedure similar to Hengel [2022] and Koffi [2021b]. Indeed, given that each paper can potentially have multiple authors from different affiliations, we ranked the affiliations based on their total number of publications in high-impact journals.

3.5 Summary statistics

Table I presents the summary statistics for the dataset, which includes a total of 329,231 papers. Of the total, 178,411 papers (54.2%) were authored exclusively by White teams. In contrast, 83,242 papers (25.3%) included at least one Black, Hispanic, or Asian author, with 2,754 papers involving at least one Black scholar, 20,983 papers including at least one Hispanic scholar, and 59,505 papers featuring at least one Asian scholar. Additionally, 67,578 papers (20.5%) were authored by individuals from other racial groups. In particular, in Figure I, there has been a steady growth in the proportion of published papers including non-White authors over time. The growth becomes particularly noticeable from the 1990s onwards. In the set of 35 high-impact journals, their share even reached 50% as of 2021.

Papers authored by all-White teams tend to have fewer authors, averaging 1.7 authors per paper. In contrast, papers with at least one Black, Hispanic, or Asian author average 2.43 authors per paper, while papers from other racial groups average 2.19 authors per paper.

Regarding publication records, all-White author teams had the highest average number of publications across various journal tiers as of 2021. For example, all-White teams averaged 19.97 publications in the top 200 journals, compared to 19.33 for papers with at least one non-White author and 15.62 for papers from other racial groups. A similar trend is observed in top-tier journals (Top 35, Top 16, and Top 5), where all-White teams consistently had higher averages compared to teams with non-White authors. The differences are statistically significant.

The dataset also provides insight into the distribution of papers across different academic fields. Papers authored by all-White teams were most concentrated in microeconomics (19.14%) and other empirical fields (26.21%). Teams with at least one non-White author showed a higher concentration in finance (19.27%) and econometrics (11.84%). Non-White authors tend to produce a higher proportion of empirical papers and a lower proportion of theoretical papers compared to their White counterparts.

further details on the classification algorithms.

Finally, the dataset tracks the institutional affiliations of the authors, ranking them into low-, mid-, and top-tier institutions. All-White teams had the highest proportion of papers affiliated with top-tier institutions (14.3%), compared to 13.37% for teams with at least one non-White author and 11.01% for papers from other racial groups.

4 Empirical strategy and results

4.1 Empirical framework

We test for how citation patterns differ depending on the racial composition of a paper's authors. Assume that paper p was published in year t and cited by articles published in a given subset of journals. Then, our outcome variable, the total number of citations (using the inverse hyperbolic sine function), denoted $C_{p,t}$, received by paper p is given by the following linear model:

$$C_{p,t} = \beta_0 + \beta_R R_p + \beta_3 X_p + \gamma_{jt} + \lambda_f + \epsilon_{p,t} \tag{1}$$

In this study, the primary variable of interest, R_p, defines the racial composition of the authors of article p. In the baseline specification, this indicator is defined as follows: it is equal to 0 if the article has only White co-authors (all-White); 1 if the article has at least one co-author who is Black, Hispanic, or Asian; and 2 if all the authors in the article cannot be considered White and the article does not have any Black, Hispanic, or Asian authors (see Section 3.2.1). In Equation 1, the coefficient of interest, β_R captures the difference in the number of citations received by article p when it includes at least one non-White co-author, compared to when it is authored exclusively by an all-White team.

To account for potential heterogeneity between articles, we include a range of control variables represented by the vector X_p. These controls encompass characteristics such as the number of authors, the affiliation of the article, and the academic achievements of the authors. The number of authors can influence both the racial diversity of the article and the likelihood of citation, as a larger network may increase citation probability. Including authors' affiliations helps control for differences in visibility or perceived quality of the authors. Similarly, considering the academic reputation of the authors allows us to condition the prestige and prominence of the authors. We also include journal times year fixed effects, denoted by γ_{jt}, to control for factors related to the journals and their publication years. Additionally, we incorporate dummy subdomains using the primary field of the paper, represented by λ_f, to account for variations across economic subfields.

4.2 Main findings

The estimation results of Equation 1 above are presented in Table II. We use various specifications and controls across the columns. Column (1) includes no controls. Column (2) incorporates controls for the number of authors. Column (3) additionally controls for authors' academic performance.

Column (4) introduces year-fixed effects, while Column (5) includes affiliation rank fixed effects. Column (6) adds field-fixed effects, and Column (7) controls for all the aforementioned variables, including journal-year fixed effects.

The table shows a significant citation gap between papers exclusively authored by White scholars and those with non-White authors, even after controlling for team size, prior publications (as a productivity signal), affiliations, primary field, and journal-year of publications. Articles with at least one Black, Hispanic, or Asian author consistently received 5.1 to 25.6 log points fewer citations than those authored solely by White scholars, with the unconditional gap being around 13.9 log points.

Three variables significantly influence the estimates. First, increasing team size tends to widen the citation gap, increasing the difference by 12 log points. Non-White scholars typically have larger team sizes, which should attract more citations. However, these diverse teams do not receive as many citations as similarly sized all-White teams. In contrast, publication history reduces the gap, explaining 10 log points, while publication years account for at least 9 log points.[7] In the sample, non-White scholars tend to have fewer publications as of 2021, and their publication years are, on average, 10 years more recent. Citations typically accumulate over time, contributing to the observed citation disparity.

Overall, in our preferred specification in Column (7), the estimated citation penalty for articles with at least one Black, Hispanic, or Asian author remains, on average, 5.1 log points or approximately 5%. Similarly, the citation penalty for the other races category averages 4.4 log points.

Furthermore, due to the substantial heterogeneity in the journals included in our baseline estimation, it is unclear if the citation gap persists in a subset of high-quality journals. For instance, the citation gap might be driven by lower-quality journals, while higher-quality journals could be immune to this gap. However, Table III shows that the citation penalty for non-White scholars is not confined to lower-ranked journals. We estimate a gap of around 9% (9.6 log points) for the top five journals in economics, 8% (8.5 log points) for a set of top 16 journals used in Koffi [2021b], and 8% (8.6 log points) for a set of 35 high-impact journals as per Card et al. [2020].[8]

Lastly, to gain deeper insight into the underlying dynamics, we separately categorized the papers based on whether they included at least one Black, Hispanic, or Asian scholar, with these categories being mutually exclusive. Figure II shows that the citation penalty persists across all racial categories and journal ranks. However, the penalty is more pronounced for Black scholars in the top 35, top 16, and top 5 high-impact journals, with respective coefficients for Black-authored papers of 10%, 17%, and 18%.

[7]We accounted for the number of publications by the most prolific author at the time of the paper's publication and, as of 2021, the final year of the database's data collection. This dual approach acknowledges that an author's current prominence and their prominence at the time of publication can impact citation frequency. The results show that prominence up to 2021 positively affects citation rates, while prominence at the time of publication appears to have a negative impact when controlled for prominence up to 2021.

[8]These results contrast with findings in gender analyses within economics, where publication history tends to vanish or even reverse the gap (Card et al. [2020]; Koffi [2021a]).

4.3 Heterogenous effects

First, we examine how the racial citation gap varies according to the authors' affiliation. We categorize papers based on whether their most prolific authors belong to a top 15 institution ("Top Tier"), a top 15-100 institution ("Mid Tier"), or an institution ranked below the top 100 ("Low Tier"). Figure III shows that the citation penalty for non-White scholars is twice as large in top-tier affiliations compared to mid and lower-tier affiliations. In other words, non-White scholars in top-ranked institutions face a stronger citation penalty compared to their White counterparts, and this penalty is more pronounced than that experienced by non-White scholars in mid- or lower-ranked universities compared to other White scholars in similarly ranked institutions.

We then estimate the model across different subfields and methodology styles within economics, as shown in Figure IV. The citation gap for papers with at least one non-White author is evident in almost all fields. The most severe penalties occur in microeconomics and international economics (slightly exceeding -10 log points), followed by econometrics, labor and education, macroeconomics, and finance. Moreover, the presence of at least one minority author consistently results in a citation gap impact in mostly theoretical (-0.079; std. err. = 0.007) and mostly structural (-0.061; std. err. = 0.024) fields. The citation gap for papers with non-White authors is almost five times higher in theoretical fields than the gap in empirical fields (-0.016; std. err. = 0.009).

In Table IV, we estimate a citation penalty for non-White authors of 5 log points with a single author and 7.2 log points with two authors. For papers with more than three authors, we estimate a 3.2 log-point penalty.

Finally, an important aspect of analyzing heterogeneity is determining whether the observed citation penalty persists in contemporary settings. Specifically, it is important to examine whether the citation gap has narrowed over time with the increasing presence of teams that include at least one non-White author. Figure V shows that while the negative impact of having non-White authors on article citations continues, the penalty has decreased since the 1990s. For articles published before 1989, the citation penalty was approximately between 8 log points and 13 log points. Post-1989, the penalty is closer to -5 log points.

4.4 Robustness

In this section, we show that our results are robust to a wide range of alternative specifications.

4.4.1 Selection on unobservables

One key possible concern related to the estimated citation penalty for the non-White author team in economics is the possible effect of unobservable variables that could be correlated with race. As shown in Table II, the fluctuation in the coefficient after including the full set of control variables suggests that the gap may reflect the influence of some unobservable variables correlated with race. To address this, we estimate the Oster [2019] bound for our main specifications in Table II. The identified set uses parameters $\delta = 1$ and $R^2_{\max} = \min(1, 2R^2_c - R^2_u)$, where R^2_c represents the R^2 resulting from regressing the dependent variable on observable factors, and R^2_u signifies the R^2 from regressing the dependent variable exclusively on the race variable. The adjusted coefficient is given by $\beta = \beta_{\text{controlled}} - \delta \cdot [\beta_{\text{uncontrolled}} - \beta_{\text{controlled}}] \cdot \frac{R^2_{\max} - R^2_c}{R^2_c - R^2_u}$.

Using Table V—Columns (1) and (2), we find that the identified set for the coefficient of non-White authored papers is $[-0.051, 0.026]$ with $R^2_{\max} = 1$. Noticing that $R^2_{\max} = 1$ could be an extreme assumption, for an $R^2_{\max}$ lower than 0.78, the upper bound is lower than -0.01 log points. Similarly, with these extreme parameter values ($R^2_{\max} = 1$), the identified set for the coefficient of the variable "at least one Black/Hispanic/Asian" likely includes zero for the subsample of the top 35 and top 16 journals.

A particularly intriguing finding is that for all racial category definitions, Oster's bounds are excluding zeros for the publications in the top five journals. The identified sets are $[-0.319, -0.200]$ for teams with at least one Black scholar, $[-0.231, -0.147]$ for teams with at least one Hispanic scholar, $[-0.071, -0.063]$ for teams with at least one Asian scholar, $[-0.12, -0.096]$ for teams with at least one Black/Hispanic/Asian member, and $[-0.141, -0.073]$ for teams with neither identified Black/Hispanic/Asian/White member.

So, while we could not exclude the potential effect of unobservable variables, the results show that at least at the very top, they are unlikely to be the driven mechanisms.

4.4.2 Finer field controls

We conduct robustness checks by adding more granular field controls, as presented in Table VI. Column (1) shows the baseline specification with all controls, restricting the sample to papers with available JEL codes. Similar to the baseline, we find a racial citation gap of approximately -5 log points for papers with at least one Black, Hispanic, or Asian author compared to papers with all-White authors. Additionally, there is a penalty of 3.5 log points for papers whose authors are neither classified as Black, Hispanic, Asian, nor White. Columns (2) to (4) alternatively control for the 1-digit, 2-digit, and 3-digit JEL codes, yielding citation gaps of 7.0, 7.0, and 6.5 log points, respectively.

In Column (5), we use a more granular field representation by employing advanced natural language processing algorithms, such as BERT, to generate an embedding representation (a 768-dimensional vector) of the papers.[9] This allows for finer topic determination, resulting in a citation gap of 6.9 log points. Similarly, in Column (8), which controls for the 3-digit JEL code, methodology style, and embedding representations, the gap is still estimated at 6.4 log points.

We conclude that the negative citation premium for papers with a non-White author persists, even with more granular field controls.[10]

[9] Bidirectional Encoder Representations from Transformers (BERT) is a large language model introduced by Devlin [2018]. BERT revolutionized natural language processing by using a bidirectional transformer model, allowing for a deeper contextual understanding of language. See Koffi [2021b] for more discussion on the use of those algorithms.

[10] Our analysis also reveals that papers with more empirical content are more likely to be cited than those with more

4.4.3 Measurement errors on race

Algorithmic methods for determining an individual's race rely on name-based classifiers, which may not accurately capture racial identity due to the diversity of names within racial groups, and the effects of cultural mixing (and also the complexities of self-identification). They are, therefore, inherently prone to measurement error. Consequently, such methods may introduce bias and noise into the analysis. This measurement error is especially problematic in studies of racial disparities, as misclassification could obscure the true extent of racial gaps or, conversely, exaggerate differences.

To validate the results from the algorithmic classification, we conducted a manual verification of the authors in our dataset. Given the large database size, which includes over 220,000 unique authors, we limited the manual verification to authors who published after 1990 and have at least one publication in a top 5 journal. For the 5,149 authors meeting this criterion, we manually verified their race by reviewing their personal and institutional websites and searching for photos or information about their countries of origin.[11]

Our findings from the manual verification in Appendix Table C.1 confirm the presence of racial citation bias, with non-White authors being cited less frequently than White authors. However, algorithmic predictions estimated a bias approximately 2 log points higher than manual verification results, though this difference is not statistically significant. In our most conservative specification, the bias is estimated at 7.5 log points, with a standard deviation of 2.9 log points.

4.4.4 Beyond direct connections: the use of centrality measures

In the previous section, we measured a paper's ability to disseminate a research idea using the number of citations it has received. However, scholars view some citations as more important than others, and the influence of a paper is not only measured on the basis of the number of citations it has received but also by the number of "influential" papers it has inspired. We analyze the role of authors' races in determining the influence of their scholarly works.

To measure paper importance, we assume that published papers form a directed network where a paper is linked to another if it cites the latter. In this context, papers are "nodes," and citation links are "edges." In graph theory, node importance is measured using various centrality indices, categorized into two broad classes. The first class assesses importance based on direct and indirect links to other nodes (e.g., degree, eigenvector, Katz, and Pagerank centrality, and clustering coefficient). The second class focuses on positions pivotal to the network's connectivity (e.g., betweenness centrality). We consider all six measures to assess a paper's importance.

Degree centrality counts the number of direct citations a paper receives. Betweenness centrality measures a paper's importance based on how often it acts as a bridge along the shortest path between two other papers. A paper with high betweenness centrality often mediates the influence of an earlier

theoretical content, consistent with the findings of Angrist et al. [2020].

[11] Although this method helps in reducing the potential effect of measurement error, it is still different from self-identification.

article on a later one, providing explanations, extensions, or applications of previous ideas or serving as a literature review summarizing original findings. Eigenvector centrality, on the other hand, assigns scores to papers based on the idea that citations from highly cited papers are more valuable than those from less cited papers. Katz centrality and PageRank centrality are variations of eigenvector centrality. Katz centrality measures the number of citations a paper receives, and its indirect links through citations, with closer neighbors weighted more heavily. PageRank centrality is similar but differs in how it weights indirect connections. The clustering coefficient quantifies how close the neighbors of a paper are to forming a complete graph, indicating how tightly knit they are. The respective formulas of these indices are presented in Appendix B.

We estimate the baseline equation using degree centrality, betweenness centrality, eigenvector centrality, Katz centrality, and PageRank centrality as outcome variables. The results, presented in Table VII (Columns 1 to 5), show that papers with at least one non-White author score significantly lower on all measures compared to papers by all-White teams.

Therefore, the citation penalty for papers authored by non-White scholars extends beyond direct citations to encompass various measures of a paper's importance within a citation network. This includes indirect citations and the paper's position within a network of interconnected papers. These findings indicate that the citation disparity is not merely a matter of fewer direct citations, but also reflects a broader trend in how these papers are integrated and disseminated within the scholarly community.

Column (6) of Table VII presents the clustering coefficient. We observe that papers by non-White authors tend to have higher clustering coefficients, meaning that papers citing non-White authors are also likely to cite each other, forming a tightly interconnected group. This suggests that these papers are more likely to be situated in a cohesive subgroup of the citation network, indicating a limited dissemination of their ideas.

4.4.5 Other robustness checks

In additional robustness checks, we use alternative definitions of race, as presented in Table VIII. Columns (1) and (2) define race using predictions from the US census and Florida voter registration, respectively. Column (3) uses Florida voter registration data for the author's first name, if available, and the US census database if the first name is not available. Columns (4), (6), and (7) use a combined approach with varying probability thresholds and manual checks for Black economists, with Column (4) being similar to the baseline estimation in Table II. Columns (5), (8), and (9) vary in probability thresholds and exclude manual checks for Black economists. The racial citation gap estimate shows little variation across these definitions, with the penalty for papers with a non-White author ranging from around 3.8 log points (Column 1) to 5.1 log points (Column 4). These findings again show that the definition of race does not drive our results. In Table IX, we also present results using the share of authors by racial categories and different categorical variables, which exhibit a similar pattern.

In Table X, we analyze how the effect of race on citation is impacted by the readability index, which measures how difficult a text is to read. This analysis addresses the concern that papers authored by non-White scholars might be harder to read, particularly if English is not their native language, potentially leading to fewer citations. We do not find that controlling for the readability score of the paper substantially changes the results.

Given the recent debate on the possible incorrectness of the use or interpretation of the inverse hyperbolic sine function (Bellemare and Wichman [2020],Chen and Roth [2024]), we show that the racial citation gap estimation is robust to varying the citation variable's functional form, as presented in Table XI. We consider alternative definitions: the simple citation count (Column 1), the logarithm of citations plus one (Column 2), the Poisson transformation (Column 3), binary indicators for the top 2%, 5%, and 10% of citations, adjusted for the time variation (Columns 4-6), and excluding zero-citation papers (Column 7). Our findings consistently show that papers with at least one non-White author face a citation penalty. For example, they receive 4.5 fewer citations (45% of the median citation of 10 or 12% of the mean citation of 37) and are 20% less likely to be in the top 2% of the citation distribution.

It is possible that our controls for publication are not sufficiently granular to capture variations due to the authors' prominence. In Table XII, we add detailed controls based on the most prolific authors by separately counting their number of publications in various top journals at the time of publication and up to the final year of our sample. Similarly, we also consider the average prominence of the authors on the paper and the total prominence (sum of all publications) of the authors on the paper. The results are not significantly affected by these changes.

Finally, we exclude self-citations from the total citation count, as shown in Appendix Table C.2, and the qualitative results remain consistent.

5 Possible sources of the racial citation gap

The racial citation gap exhibits the following characteristics: it persists in top-tier journals and institutions, it cannot be attributed to differences in the prominence or recognition of authors across racial groups, it is almost negligible in empirical fields, and it has significantly diminished over time. This reduction may be attributed to improved dissemination and increased awareness of research by non-White scholars, as well as the growing prominence of empirical fields, where the citation gap tends to be smaller (Angrist et al. [2020]). Nevertheless, the continued citation gap in recent years might suggest that non-White researchers may still encounter challenges in gaining visibility for their work, as proposed by Doleac et al. [2021], or that there may be differences in (perceived) quality. In this section, we explore potential sources of the citation gap.

5.1 Sorting based on research areas and topics

Recent evidence in patenting, medical sciences and some fields of social sciences and humanities indicates that scientists tend to conduct their research based on their socio-demographic backgrounds (Einiö et al. [2019]; Koning et al. [2021]; Kozlowski et al. [2022]). If this trend extends to the social sciences, we would expect to see a concentration of individuals in certain research topics based on characteristics such as ethnicity or race. If some research fields receive fewer citations than others, as shown in the empirical/theoretical comparison in Section 4.3, then non-White researchers might be concentrated in fields with lower citation rates. This could explain the citation gap observed even for articles published in top economics journals.

However, despite the possible existence of these differences, as illustrated in Table VI, accounting for an extremely detailed research and topic categorization tends, if anything, to increase the average gap by 20%-40%.

5.2 Differential quality and impact

5.2.1 Editorial-based quality differential

If citations serve as a proxy for quality, albeit an imperfect one, then a lower citation count might indicate a lower-quality paper. Tables III and V reveal that the differences are more pronounced for publications in top economics journals (or for publications by authors affiliated with top-ranked institutions, and the differences are persistent even conditioning on the authors' publication records). This contradicts the argument that papers from non-White scholars are of lower quality (at least using those additional proxies of quality). However, consider a publisher that accepts a paper if the expected quality of the paper meets a certain threshold $\bar{q}$. If a policy similar to affirmative action, which has been criticized for potentially lowering the threshold for certain racial groups (Arcidiacono et al. [2023]), is implicitly applied in academic publication, or if it is perceived that such a policy exists, then the threshold (or the perceived threshold) $\bar{q}_{nonWhite}$ would be lower than $\bar{q}_{White}$, suggesting that papers authored exclusively by White scholars would be of higher quality, or perceived as such.

While there is no direct evidence of such phenomena in academic publishing, the emphasis on diversity and inclusion initiatives may influence perceptions and beliefs about the quality of papers authored by non-White scholars. This could inadvertently have the same effect as if such a policy were in place, resulting in a lower perceived quality threshold for non-White authors' papers ($\bar{q}_{nonWhite} < \bar{q}_{White}$).

However, two arguments suggest that this scenario is unlikely. First, our data extends back to the 1950s, a period when positive action policies were less probable. During this time, non-White scholars who successfully entered the publication world would likely have been at least as competent as their White counterparts ($\bar{q}_{nonWhite} \geq \bar{q}_{White}$). Second, if anything, criticisms of differential quality-based admission policies have typically focused on Blacks and Latinx, with some evidence, as previously mentioned, indicating that Asians may face discrimination under such policies. Therefore, we would not expect papers with Asian authors to be perceived as of lower quality.

Overall, it seems unlikely that the editorial-based quality differential favoring non-White authors drives the observed results.

5.2.2 Noisy ability signal and information updating

If agents have a lower prior on non-White scholars' ability and anticipate that papers by non-White scholars are, on average, of lower quality, then traditional belief-based bias models, based on the mean quality, will imply that they will update as they receive clear signals about a non-White scholar's ability.

We investigate this by analyzing how additional publications affect the citation penalty experienced by non-White researchers. Figure VI illustrates the estimation results. Overall, the effect of additional publications on mitigating the citation penalty is non-monotonic. Strikingly, among authors with only one publication in a top-five journal, the citation penalty for non-White scholars is generally higher than that for authors with no prior publications. This suggests that an additional publication, even in top journals, tends to benefit all-White scholar teams more than teams that include at least one Black, Hispanic, or Asian member.

Moreover, within a dataset encompassing over 200 journals, non-White researchers see a reduction in the citation penalty when they have at least 20 prior publications in total, 11 publications across all 35 high-impact journals, or at least 6 publications in all 16 high-impact journals or the top 5 journals. Authors with publications in all 35 high-impact journals or all 16 high-impact journals need at least 11 prior publications in the set of top 35 journals or 6 publications in the top 16 or top 5 journals to observe a noticeable decrease in the citation gap. In the case of the top 5 journals, at least 2-5 additional prior publications are required to reduce the citation penalty, with a significant reduction occurring at a threshold of 6 prior publications.

Therefore, while additional publications appear to have an impact, their significance in reducing the gap clearly depends on the type of journal, even when controlling for the journal of the paper under consideration. More importantly, this effect does not become evident until a critical mass of publication records is reached.

Typically, factors such as affiliation ranking, journal of publication, and an author's established prominence are considered indicators of quality. However, our study finds that even when these factors are taken into account, they do not reduce the citation gap between White and non-White scholars, except at the extreme end of the distribution of publication records.

5.2.3 Past citations as a proxy for quality and potential impact

To substantiate the argument that the quality differential or impact is likely to play a minor role in explaining the citation gap, we use the authors' citation history at the time of publication as a proxy for author quality and, consequently, the potential quality of the article. Specifically, for each article, we calculate the average number of citations accumulated by the authors prior to the publication of the article.

Table XIII presents the results of the regression analysis, incorporating a non-linear form of the team's past citations into the baseline equation from section 4.1. The variable "At least one Black/Hispanic/Asian" consistently shows a significant negative coefficient across all journal groups, indicating a citation penalty for papers authored by these groups. However, this penalty is reduced by 1 to 3 percentage points compared to the baseline. Specifically, in the more than 200 journals category, the penalty is 4.3 log points, while it is 6.8 log points in the top 35 journals, 6 log points in the top 16 journals, and 8.3 log points in the top 5 journals. Similarly, the "Other races" category also displays negative and significant coefficients across all journal groups. We also observe that past citations are strong and positive predictors of the citations for current articles.

In Appendix Table C.3, we further refine our analysis by disaggregating the outcome variable to consider the annual citation count of each article. Specifically, for each year, we calculate the number of citations an article receives by extracting references from our database. In addition to the control variables previously considered, we introduce several new controls: cumulative previous citations (citations accumulated by the article up to, but not including, time t); the average number of publications in the top 5, top 16, and top 35 journals; as well as total publications by the authors at time t. We also control for the average cumulative previous citations of the authors up to time t, excluding time t and the current article citations. This specification aims to account for both the quality of the authors and the perceived quality of the paper prior to receiving citations at time t. The trend remains consistent: teams with at least one Black, Hispanic, or Asian member consistently show a citation penalty compared to teams composed entirely of White authors. The estimates are relatively lower because it is an annual reduction.

Overall, even after controlling for all these factors, the tables illustrate a persistent citation disadvantage for articles authored by non-White researchers. This disadvantage remains significant even when accounting for citation history, which is endogenous and likely to introduce downward bias.

5.3 Citation patterns and racial similarity

We investigate the extent to which racial similarities between authors and citing scholars influence citation patterns. The results, detailed in Table XIV, distinguish between citations from non-White and White scholars, revealing a pronounced homophilic pattern. Specifically, the analysis shows that papers authored by at least one Black, Hispanic, or Asian individual receive a citation premium of approximately 23% from non-White scholars compared to papers authored exclusively by White scholars. Conversely, this same group (papers authored by at least one Black, Hispanic, or Asian individual) experiences a citation penalty of around 30% from White scholars. These findings indicate a significant homophily effect, where citation practices are more likely to be influenced by the racial composition of the author team.

Further investigation into the homophily effect is presented in Figure VII, which depicts citation patterns over time and across different fields based on the race of the authors of the citing paper. The figure shows a consistent citation premium for papers authored by non-White scholars when cited by non-White scholars, evident across all time periods, with a notable increase around the 1990-99 period. In contrast, citation patterns from White scholars consistently demonstrate a penalty for papers authored by non-White scholars. These trends indicate that the reduction in the citation penalty over time is largely attributable to the increasing presence of non-White scholars in various fields and the corresponding rise in citations from non-White authors to their peers. This suggests that the observed changes are not necessarily indicative of a broad shift in citation practices across races, but rather reflect growing intra-group support.

This pattern remains consistent when analyzing citation patterns by field and methodology style. Even in empirical fields, which exhibit the lowest citation penalty for teams with at least one Black, Hispanic, or Asian member compared to all-White teams, the trend persists: non-White scholars are more likely to cite papers authored by non-White researchers than White scholars are. The magnitudes are also quite significant, with penalties consistently exceeding 25 log points and premiums above 15 log points. We also examine subfields with the highest proportion of non-White scholars, as shown in Appendix Figure IX, and find that the homophilic pattern remains evident even in these areas.

Race may also be associated with spatial and geographic differences. In other words, what might appear as a citation preference for individuals of the same race could instead be due to geographic factors that facilitate access to articles by individuals of the same racial group or to topic similarities arising from shared locations. To test this hypothesis, we analyze the institutions of the authors, focusing on a sample where at least one author in the paper is based in the United States (US) who are also cited by other US-based authors. The choice of the United States is significant, given the country's central role in the profession. The results, presented in Appendix Tables C.4 and C.5, show that the findings remain virtually unchanged, suggesting that homophily is unlikely to be driven by geographic differences.

5.4 Preferences/bias, cost of information and distributional effect

5.4.1 A simple conceptual framework

The origins of this homophily pattern are ambiguous and may be attributed to preference or bias, such as a preference for intraracial ties or a differential assessment of the quality of work from other racial groups. Homophily could also result from information asymmetry between groups, manifested in the relative ease of knowing the work of similar individuals versus the higher cost of being aware of the work of those from different racial backgrounds (different social networks). For instance, in Section 4.4.4, we show that the citation network of non-White researchers is narrower than that of all-White research teams, and non-White researchers are more likely to be cited by individuals who also cite each other. Ultimately, apparent homophily could also result from distributional effects, where both the cost of missing a strong referee and the likelihood of obtaining a referee from the majority group are higher.

Consider the following setup: There are two ethnic groups of authors, G_1, the minority group, and G_2, the majority group, where the size of G_1 is lower than the size of G_2. Let author i cite s_1 papers authored by members of the minority group and s_2 papers authored by members of the majority group.

Missing an important referee can be costly, either because referees are typically key figures in the relevant literature, and overlooking a significant paper could diminish the perceived contribution, or due to potential editorial bias (Wilhite and Fong [2012], Fong and Wilhite [2017]). It follows that the perceived quality of a paper partially depends on the combination of cited papers (s_1, s_2) and the average quality of these papers. This view relies on the assumption that ideas originating from authors with different identities and lived experiences are generally complementary (Page [2008]). However, scholars are assumed to be more familiar with papers written by others within their own group, motivated by ties such as shared nationality, regional conferences, or other connections. For this reason, it may be easier to evaluate and recognize the value of papers authored by scholars who belong to one's racial networks. Therefore, the perceived value of a paper can be defined as $\alpha[aln(s_1) + bln(s_2)]$ for a referee from the minority group and $\alpha[bln(s_1) + aln(s_2)]$ for a referee from the majority group, where a and b are positive, $a + b = 1$, $a \geq b$, and $\alpha \in (0, 1)$ is the average quality of the cited papers. The fact that $a \geq b$ means that it could be easier for a referee from the minority (resp. majority) group to recognize the value of citations to papers authored by minority (resp. majority) scholars; however, a minority referee and a majority referee are identical with respect to the easiness with which each recognizes the value of citations to papers authored by co-ethnics. Consequently, the average perceived quality of a paper is:

$$v(s_1, s_2) = \frac{n_1}{n_1 + n_2}\alpha[aln(s_1) + bln(s_2)] + \frac{n_2}{n_1 + n_2}\alpha[bln(s_1) + aln(s_2)]$$

where n_1 and n_2 are the number of referees from groups G_1 and G_2, respectively.

Writing a paper is costly. For example, citing papers accurately incurs costs, including those associated with building a literature review, staying informed about research developments, and participating in professional networks. Let $c(s_1, s_2)$ represent the total cost incurred by the author i of writing a paper. Again, owing to the fact that authors are assumed to be more familiar with papers written by others within their own group, the marginal cost of citing a paper written by a co-ethnic author is lower than that of citing a paper by a non-co-ethnic author. In addition, we assume that authors invest more effort in a paper they would like to publish in a top journal, meaning that the cost of writing a paper increases in the quality of the journal targeted by the author. However, higher-ability authors incur a lower cost, *ceteris paribus*. The cost of writing a paper can be defined as:

$$c(s_1, s_2) = h(q, r)g(s_1, s_2)$$

. where $q > 0$ is the quality of the journal targeted by the author, $r > 0$ is the author's ability, $h(q, r) > 0$ is a function that is increasing in q and decreasing in r, and $g(s_1, s_2)$ is an increasing function of the number of citations to papers authored by scholars from the minority group (s_1) and those from the majority group (s_2). Without loss of generality, we assume that $g(s_1, s_2)$ is linear. The function $h(q, r)$ measures how hard it is to publish in a journal of quality q, given an author's ability r. Then, the multiplicative expression in the cost function is a reasonable assumption, as an author aiming for a top-tier journal will typically invest more time in carefully reading and citing a paper to ensure accuracy, compared to when targeting a lower-ranked journal. This is because a more precise citation enhances the quality of the paper. Additionally, a higher-ability author will exert less effort in utilizing a paper for citation than a lower-ability author. Therefore, the marginal cost of writing or utilizing a paper increases when the target journal is of higher quality, and it decreases for authors with greater ability.[12]

The author i's maximization problem therefore consists of the following:

$$\max_{(s_1, s_2)} \quad u(s_1, s_2) = v(s_1, s_2) - c(s_1, s_2)$$

The resolution and details of this problem are provided in Appendix E considering different functional forms. The solution typically suggests a tradeoff between the cost of citing and the likelihood of selecting a referee from the minority group. If costs primarily drive the results, we observe a homophilic pattern of similar magnitude across racial groups. However, if the likelihood of selecting a referee is a factor, the homophilic pattern is weaker for the minority group.

Specifically, the homophily observed in this study indicates that the minority group cites members of their own group less frequently than the majority group cites its own members. The ratio of estimates characterizing this homophily ($\left|\frac{\beta_{\text{nonWhite by nonWhite}}}{\beta_{\text{nonWhite by White}}}\right|$) ranges from as low as 0.6 to as high as 0.95. This suggests that factors from the refereeing process might play a significant role in citation decisions. However, the cost effect is also notably strong, especially in broad empirical fields, where each group tends to cite within its own cluster. While data limitations prevent us from directly testing factors related to the refereeing process, we provide evidence suggesting that information costs and perception account for a substantial portion of the citation gap.

5.4.2 On the effects of seminars and conferences attendances, visiting positions, various affiliations, and editorial roles

Seminar and conference attendance, visiting appointments, and affiliation to prestigious research institutions can significantly affect citation rates by increasing research visibility and dissemination, ultimately reducing the information cost. These events provide a platform for scholars to present their work to a broader audience, fostering direct engagement with peers who may later cite their research.

[12] A natural candidate for the function $h(q, r)$ is $\frac{q}{r}$, though other functions could also be considered.

To test this hypothesis, we first extracted data on seminar and conference invitations from the database. We then created variables relating to the total count of seminars and conferences, attendance at prestigious conferences (e.g., NBER), and invitations to top-ranked institutions. Since this information is only available for papers referencing potential funding, the results may be subject to selection bias. Therefore, we essentially compare the coefficients' magnitude by analyzing the base coefficient without controls for seminars and conferences in this subset of the database.

The results, presented in Table XV, show that variables related to seminars and conferences account for, at most, 0.006 log points. The coefficients are insignificant, likely due to the small sample size.

Moreover, recognizing the incomplete listing of seminars at the paper level and the fact that citations may also be influenced by the visibility of the authors at various seminars and conferences (providing insight into the author's broader academic pipeline), we undertook an extensive manual collection and gathered detailed CV information from the authors wherever available, on their participation in such events.[13] The authors here have at least one top 5 after 1990, as in our manually checked sample for race. Based on this data, we created variables capturing the dissemination opportunities of the authors' work. The results show that these variables explain only a marginal portion of the coefficient, approximately 0.002 log points (Table C.6).

Additionally, we leverage the full scope of our hand-collected database to enhance the model by incorporating additional variables likely to influence author visibility (therefore reducing the cost of information): editorial roles, visiting positions, membership in prestigious organizations (NBER, JPAL, CEPR, BREAD, IZA, and CIFAR), and holding a PhD from a top-tier institution. Combined with seminar variables, this comprehensive model accounts for 17% of the variation in the racial citation gap, with most of the variation attributable to disparities in membership within prestigious associations.[14]

5.4.3 On the effects of cross-groups collaboration: a potential shift in perceptions

In trying to understand the cost function in our conceptual framework, we rely on the literature on contact theory. This literature explores how interactions between members of different groups can influence the biases and perceptions the majority group holds toward a minority. The theory suggests that intergroup contact can help diminish stereotypes and promote greater understanding, and has been tested in diverse fields experimental (Boisjoly et al. [2006], Carrell et al. [2019], Dahl et al. [2021] among others).

[13]In many papers, we encounter the phrase, "We thank numerous seminar and conference participants…" without explicitly listing the names of the seminars and conferences.

[14]Overall, seminars and conferences marginally contribute to explaining the citation gap, though this does not suggest that seminars and conference attendance are not key determinants for citation patterns. In fact, the positive signs of the relevant variables indicate that participation in these events does, in fact, increase citations. Furthermore, univariate regressions (Appendix Table C.7) reveal that seminar and conference variables explain nearly 8.3 log points of the racial citation gap. However, participation in these events is also indicative of article quality or author prominence, factors that the other control variables already account for, leaving a marginal role for this additional variable despite its inherent relevance.

In this study, we test contact theory within the context of academic publications. To mitigate potential selection effects, we use an event study framework, where the "event" is an author's first collaboration with an out-group (non-coethnic) coauthor. We then analyze all subsequent publications with only coethnic authors, truncating the sample at the next non-coethnic collaboration. The outcome variable is the number of listed bibliographical references (asinh) where the authors are from the out-group. We include the following control variables: authors' fixed effects, year of publication fixed effects, team size, fixed effects for publications count at date t, and the proportion of non-White scholars in the field of research over time.

Figure VIII presents the estimation results. Prior to collaborating with a non-coethnic author, the reference listing patterns of authors who eventually collaborate with a non-coethnic author are similar to those of authors who do not, except among White authors in panel (a). However, following the initial collaboration, we observe an increase in both the share and number of references by White authors to non-White authors, as well as an increase in the number (but not the share) of references by non-White authors to White authors. This effect persists, remaining above the pre-trend, although it diminishes relative to the initial impact. On average, the number of listed bibliographical references to non-coethnic authors increases by 8%, with the share of listed bibliographical references by White scholars to papers authored by non-White scholars rising by 5-10%, while the share of listed bibliographical references from non-White scholars to White scholars may decrease by up to 3%.

One might argue that this shift simply reflects an expansion in relevant literature from the collaboration. We counter this by showing that the effect endures even when subsequent papers do not fall within the same sub-field (measured by three-digit JEL codes) as the initial collaborative work (X).

Overall, this suggests that collaboration may influence beliefs and perceptions; however, as prior research indicates, a lasting effect may require repeated collaborations (Dahl et al. [2021]).

5.4.4 On the effects of online availability

A substantial body of works has shown that open access enhances the use of scientific knowledge (Furman and Stern [2011], Williams [2013], Murray et al. [2016], Staudt [2020], Bryan and Ozcan [2021], Biasi and Moser [2021]). This supports the hypothesis that open access reduces the cost of accessing scientific knowledge and promotes its dissemination. In economics, for instance, McCabe and Snyder [2015] found that online availability can increase citations by up to 10%, particularly through platforms like JSTOR.

Based on our conceptual framework, if the marginal cost of citing a non-co-ethnic author is higher than the one of a co-ethnic author, then a shock that increases the availability of papers, as explored by McCabe and Snyder [2015], could reduce access costs and help narrow the citation gap. However, there is also a risk that such a shock might disproportionately benefit the co-ethnic ties, potentially reinforcing existing disparities. In the conceptual framework, for example, the cost of information could not be disentangled from preferences or biases in citation. Therefore, while increasing research availability might reduce citation inequalities, the distribution of these benefits across racial groups requires careful consideration. We explore this by leveraging the data from McCabe and Snyder [2015].

We use an event-study framework to explore the dynamic of the citation gap between papers with at least one Black, Hispanic, or Asian author and papers with only White authors before and after the introduction of JSTOR. We regress the number of citations to a paper on a binary indicator for whether the paper has at least one non-White author, years dummies (starting with the year of the paper's publication) relative to one year prior to the introduction of JSTOR, an interaction term between the race dummy and each year dummy, and all the baseline controls. Specifically, we estimate the following model:

$$C_{p,j,t} = \lambda_0 + \lambda_1 R_p + \sum_{t=-m}^{m} \lambda_{2,t} JSTOR_{j,t} + \sum_{t=-m}^{m} \lambda_{3,t} R_p * JSTOR_{j,t} + X'_{p,t}\pi + \alpha_p + \delta_t + \sigma_j + \varepsilon_{p,c,t} \quad (2)$$

All variables are defined as in equation (2), with the exception that the new variables $JSTOR_{j,t}$ are dummy variables indicating the number of years relative to the introduction of JSTOR for journal j. Our parameters of interest are $\lambda_{3,t}$, where we allow the effect of JSTOR on the citation gap to vary by year. In equation (2), m defines the window of years under consideration. The model includes a comprehensive set of fixed effects, including paper-fixed effects, publication year-by-citation-year fixed effects, and journal-fixed effects. The reference period is $t = -1$, representing the year before JSTOR's introduction. Thus, the parameter of interest reflects the change in the citation gap for each year before and after the introduction of JSTOR.

The results are presented in a series of graphs in Appendix Figures XI and XII, which depict event-study analyses of the impact of JSTOR's introduction on the citation gap between non-White and White scholars. To estimate this effect, we examine citations for each article by year, starting from the year of its publication. The key identifying assumption is that there are no shocks related to the introduction of JSTOR that affect differently citations of White and non-White scholars.

The graphs indicate that the introduction of JSTOR contributed to a reduction in the citation gap, especially when examining citations originating from White scholars. The results strongly suggest that JSTOR is reducing the citation gap between White and non-White scholars, particularly when citations originate from White scholars. In the more saturated models, which include article fixed effects, publication-year-by-citation year fixed effects, and journal fixed effects, the introduction of JSTOR increases the relative citations of non-White scholars compared to White scholars by up to 2 log points annually in the four years following JSTOR's introduction. For the pull of citations and the citations originating from non-White authors, an effect observed in period -2 complicates a definitive interpretation.

When analyzing the aggregate effects (Appendix Table C.8), the citation gap between White and non-White scholars decreases by 2.2 to 3 log points when citations come from all-White teams after the introduction of JSTOR. This reduction corresponds to a narrowing of the initial citation gap (5.7

log points in column 7) by 39% to 53%.[15]

Based on these estimates, it appears that up to 60% of the citation gap could be attributed to information friction, specifically the lack of awareness of non-coethnic articles.

5.4.5 Discussion: On the potential losses of racial diffusion of ideas

In this section, we discuss potential losses to idea diffusion and innovation that could be caused by racial barriers. The idea that life experiences shape research perspectives has long been discussed in sociology and psychology. To examine this, we analyze research topics using three-digit JEL codes. Within each JEL code, we cluster papers into groups of relatively small average size (200 observations in the baseline) based on their embedding representations. For each cluster, we treat racial categories as subgroups and analyze the semantic differences across these groups. This aims to analyze more than research topics but the specific angle they pursue within a homogenous cluster of research questions.

First, we analyze the linguistic proximity of each racial category to others. To do this, we employ an inverse term frequency-inverse document frequency (TF-IDF) metric, grouping articles by racial category. We refrain from using advanced similarity techniques relying on embeddings because all representations address the same research questions; our goal is to capture variation in terminology within these questions. Figure XIII in the Appendix shows the distribution of similarity, using the all-White author category as the reference group. We observe substantial variability in similarity across groups. Specifically, the distribution for teams including Black scholars is left-skewed, with similarity values often dropping below 0.2. This suggests notable differences in terminology—and potentially in the approach to research questions—across racial groups. Similar patterns emerge when comparing teams with Hispanic or Asian researchers, where similarity is not concentrated near 1.

Panel (d) of Appendix Figure XIII presents an additional finding: the degree of variability in approaches to research questions with respect to the share of non-White researchers follows an inverted U-shape. In other words, as the pool of researchers working on a given question becomes more heterogeneous, so does the variety in topical approaches to a given research question.

To better illustrate these findings, we remove the most commonly used words (those defining the core research questions) within each cluster under a three-digit JEL code and analyze the remaining terms across racial categories to capture the specific angles addressed by each racial group. We provide examples from the database in Appendix Figures XIV and XV. Panel A, based on JEL code J15 (Economics of Minorities, Race, Indigenous Peoples, and Immigrants; Nonworkplace Discrimination), focuses on research related to child education. Here, all-White teams emphasize school environments and the broader racial gap, while teams with non-White scholars concentrate on class size and peer

[15] The introduction of JSTOR resulted in a relative increase in citations of non-White scholars compared to all-White citations by 0.003 to 1.5 log points (0 to 62% of the initial -2.4 log points- column 1). The aggregate effects suggest a relative increase in citations of White scholars compared to non-White scholars when citations originate from non-White scholars, though this may be influenced by pre-trend effects observed in the event study analysis.

effects.

In Panel B, teams with non-White members discuss teenage pregnancy and examine the impacts of various programs on issues related to women's and girls' education. For women's health, non-White scholars focus on preventive measures like screening and abortion, while White scholars emphasize the role of insurance, birth weight, and variation across states and income levels. In Panel D, while White scholars focus on unemployment, non-White scholars discuss disability and income volatility.

In Panel E, agricultural policy and farmers' markets are analyzed from distinct perspectives: White scholars focus on crop prices and public sector roles, whereas non-White scholars emphasize environmental impacts, with attention to cocoa farming, certification, and food reserves. Similarly, when studying Chinese exports, White scholars often examine competition effects on firms, while non-White scholars take a broader approach, addressing technological processes, productivity, and innovation.

Overall, our analysis reveals that scholars from different racial backgrounds bring distinct perspectives to research questions, resulting in varied angles of investigation and emphasis. In line with Page [2008], our findings indicate that by failing to build on the unique ideas contributed by scholars of various racial backgrounds, we risk overlooking valuable insights that could enhance economic theory, public policy, and social progress.

6 Conclusion

Our analysis reveals significant differences in the way research produced by White and non-White scholars is perceived and cited in economics. Papers with at least one non-White author receive 5.1% to 9.6% fewer citations. While the citation penalty is relatively consistent across racial minority groups in the full sample of papers, it tends to be larger for Black authors in papers published in top journals. These findings are robust even after controlling for a wide range of paper characteristics, such as the number of authors, field, journal, and year of publication, as well as prestige and quality-related factors like author affiliation, academic prominence, disaggregated fields, and prior citations. The findings are also robust to controlling for additional variables such as seminar and conference attendance, attendance at prestigious conferences, seminar invitations from top-ranked institutions, visiting appointments in top-tier institutions, editorial roles, and membership in prestigious organizations. Over time, the racial citation gap has narrowed, reflecting an increase in non-White co-authorship and a shift towards empirical economic research over the past three decades (Angrist and Pischke [2010]).

The analysis of centrality to assess paper influence in citation networks reveals that papers authored by non-White scholars not only receive fewer direct citations, but are also less likely to act as a bridge along the shortest citation path connecting two other papers and are less likely to be cited by other highly cited papers. The analysis also implies that papers that cite the works of non-White economists are generally not highly cited, which denies them any indirect influence.

To explain the racial citation gap in economics, we explore various factors. Notably, we find insufficient evidence to suggest that differences in article or author quality account for the gap. However, our analysis reveals significant group bias, with non-White authors more likely to cite other non-White authors, and White authors more likely to cite other White authors. In a theoretical framework, we demonstrate that this homophily can arise from factors such as the cost of citation—whether due to a genuine lack of awareness or bias—as well as potential preferences among referees for diversity. Finally, we show that reducing information costs could decrease the citation gap by up to 50%. Overall, we may be missing valuable opportunities for the advancement of economic research by not building upon the distinct ideas contributed by each racial group. This raises the importance of connecting those distinct networks.

References

Amarante, V., Burger, R., Chelwa, G., Cockburn, J., Kassouf, A., McKay, A., and Zurbrigg, J. (2022). Underrepresentation of developing country researchers in development research. *Applied economics letters*, 29(17):1659–1664.

Angrist, J., Azoulay, P., Ellison, G., Hill, R., and Lu, S. F. (2020). Inside job or deep impact? extramural citations and the influence of economic scholarship. *Journal of Economic Literature*, 58(1):3–52.

Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives*, 24(2):3–30.

Arcidiacono, P., Kinsler, J., and Ransom, T. (2023). What the students for fair admissions cases reveal about racial preferences. *Journal of Political Economy Microeconomics*, 1(4):615–668.

Bayer, A., Hoover, G. A., and Washington, E. (2020). How you can work to increase the presence and improve the experience of black, latinx, and native american people in the economics profession. *Journal of Economic Perspectives*, 34(3):193–219.

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.

Bellemare, M. F. and Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, 82(1):50–61.

Bertolero, M. A., Dworkin, J. D., David, S. U., Lloreda, C. L., Srivastava, P., Stiso, J., Zhou, D., Dzirasa, K., Fair, D. A., Kaczkurkin, A. N., et al. (2020). Racial and ethnic imbalance in neuroscience reference lists and intersections with gender. *BioRxiv*, pages 2020–10.

Biasi, B. and Moser, P. (2021). Effects of copyrights on science: Evidence from the wwii book republication program. *American Economic Journal: Microeconomics*, 13(4):218–260.

Boisjoly, J., Duncan, G. J., Kremer, M., Levy, D. M., and Eccles, J. (2006). Empathy or antipathy? the impact of diversity. *American Economic Review*, 96(5):1890–1905.

Bryan, K. A. and Ozcan, Y. (2021). The impact of open access mandates on invention. *Review of Economics and Statistics*, 103(5):954–967.

Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2020). Are referees and editors in economics gender neutral? *The Quarterly Journal of Economics*, 135(1):269–327.

Card, D., DellaVigna, S., Funk, P., and Iriberri, N. (2022). Gender differences in peer recognition by economists. *Econometrica*, 90(5):1937–1971.

Carrell, S. E., Hoekstra, M., and West, J. E. (2019). The impact of college diversity on behavior toward minorities. *American Economic Journal: Economic Policy*, 11(4):159–182.

Chen, J., Liu, Q., and Kim, M. (2022). Gender gap in tenure and promotion: Evidence from the economics ph. d. class of 2008. *Southern Economic Journal*, 88(4):1277–1312.

Chen, J. and Roth, J. (2024). Logs with zeros? some problems and solutions. *The Quarterly Journal of Economics*, 139(2):891–936.

Dahl, G. B., Kotsadam, A., and Rooth, D.-O. (2021). Does integration change gender attitudes? the effect of randomly assigning women to traditionally male teams. *The Quarterly journal of economics*, 136(2):987–1030.

Devlin, J. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Doleac, J. L., Hengel, E., and Pancotti, E. (2021). Diversity in economics seminars: who gives invited talks? In *AEA Papers and Proceedings*, volume 111, pages 55–59. American Economic Association

2014 Broadway, Suite 305, Nashville, TN 37203.

Dupas, P., Modestino, A. S., Niederle, M., Wolfers, J., et al. (2021). Gender and the dynamics of economics seminars. Technical report, National Bureau of Economic Research.

Eaton, A. A., Saunders, J. F., Jacobson, R. K., and West, K. (2020). How gender and race stereotypes impact the advancement of scholars in stem: Professors' biased evaluations of physics and biology post-doctoral candidates. *Sex Roles*, 82:127–141.

Einiö, E., Feng, J., and Jaravel, X. (2019). Social push and the direction of innovation. *Available at SSRN 3383703*.

Ellison, G. (2013). How does the market use citation data? the hirsch index in economics. *American Economic Journal: Applied Economics*, 5(3):63–90.

Fiala, D. and Tutoky, G. (2017). Pagerank-based prediction of award-winning researchers and the impact of citations. *Journal of Informetrics*, 11(4):1044–1068.

Fong, E. A. and Wilhite, A. W. (2017). Authorship and citation manipulation in academic research. *PloS one*, 12(12):e0187394.

Fourcade, M., Ollion, E., and Algan, Y. (2015). The superiority of economists. *Journal of Economic Perspectives*, 29(1):89–114.

Francis, D. V., Hardy, B. L., and Jones, D. (2022). Black economists on race and policy: Contributions to education, poverty and mobility, and public finance. *Journal of Economic Literature*, 60(2):454–493.

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.

Furman, J. L. and Stern, S. (2011). Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *American Economic Review*, 101(5):1933–1963.

Ginther, D. K. and Kahn, S. (2004). Women in economics: moving up or falling off the academic career ladder? *Journal of Economic perspectives*, 18(3):193–214.

Ginther, D. K., Schaffer, W. T., Schnell, J., Masimore, B., Liu, F., Haak, L. L., and Kington, R. (2011). Race, ethnicity, and nih research awards. *Science*, 333(6045):1015–1019.

Greenspon, J. and Rodrik, D. (2021). A note on the global distribution of authorship in economics journals. Technical report, National Bureau of Economic Research.

Hager, S., Schwarz, C., and Waldinger, F. (2023). Measuring science: Performance metrics and the allocation of talent. *Available at SSRN 4472441*.

Hamermesh, D. S. (2018). Citations in economics: Measurement, uses, and impacts. *Journal of Economic Literature*, 56(1):115–156.

Hamermesh, D. S., Johnson, G. E., and Weisbrod, B. A. (1982). Scholarship, citations and salaries: Economic rewards in economics. *Southern economic journal*, pages 472–481.

Hengel, E. (2022). Publishing while female: Are women held to higher standards? evidence from peer review. *The Economic Journal*, 132(648):2951–2991.

Hilmer, M. J., Ransom, M. R., and Hilmer, C. E. (2015). Fame and the fortune of academic economists: How the market rewards influential research in economics. *Southern Economic Journal*, 82(2):430–452.

Hofstra, B., Kulkarni, V. V., Munoz-Najar Galvez, S., He, B., Jurafsky, D., and McFarland, D. A. (2020). The diversity–innovation paradox in science. *Proceedings of the National Academy of Sciences*, 117(17):9284–9291.

Koffi, M. (2021a). Gendered citations at top economic journals. In *AEA Papers and Proceedings*, volume 111, pages 60–64. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Koffi, M. (2021b). Innovative ideas and gender inequality. Technical report, Working Paper Series.

Koffi, M., Pongou, R., and Wantchekon, L. (2024). Racial inequality and publication in economics. *AEA Papers and Proceedings*, 114:300–304.

Koffi, M. and Wantchekon, L. (2024). Racial justice from within? diversity and inclusion in economics. *Monograph of Econometric society*.

Koning, R., Samila, S., and Ferguson, J.-P. (2021). Who do we invent for? patents by women focus more on women's health, but few women get to invent. *Science*, 372(6548):1345–1348.

Kozlowski, D., Larivière, V., Sugimoto, C. R., and Monroe-White, T. (2022). Intersectional inequalities in science. *Proceedings of the National Academy of Sciences*, 119(2):e2113067119.

Lutmar, C. and Reingewertz, Y. (2021). Academic in-group bias in the top five economics journals. *Scientometrics*, 126(12):9543–9556.

Mason, P., Myers, S. L., Simms, M., Lai, Y., and Liu, X. (2023). Research note: Citation bias and the economics of race and crime literature. *The Review of Black Political Economy*, page 00346446231212717.

Mason, P. L., Myers Jr, S. L., and Simms, M. (2022). Racial isolation and marginalization of economic research on race and crime. *Journal of Economic Literature*, 60(2):494–526.

McCabe, M. J. and Snyder, C. M. (2015). Does online availability increase citations? theory and evidence from a panel of economics and business journals. *Review of Economics and Statistics*, 97(1):144–165.

Mughogho, W., Adhiambo, J., and Forscher, P. S. (2023). African researchers must be full participants in behavioural science research. *Nature Human Behaviour*, 7(3):297–299.

Murray, F., Aghion, P., Dewatripont, M., Kolev, J., and Stern, S. (2016). Of mice and academics: Examining the effect of openness on innovation. *American Economic Journal: Economic Policy*, 8(1):212–252.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Page, S. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies-new edition*. Princeton University Press.

Pieters, R. and Baumgartner, H. (2002). Who talks to whom? intra-and interdisciplinary communication of economics journals. *Journal of Economic Literature*, 40(2):483–509.

Qiu, S., Steinwender, C., and Azoulay, P. (2022). Who stands on the shoulders of chinese (scientific) giants? evidence from chemistry. Technical report, National Bureau of Economic Research.

Sarsons, H. (2017). Recognition for group work: Gender differences in academia. *American Economic Review*, 107(5):141–145.

Sood, G. and Laohaprapanon, S. (2018). Predicting race and ethnicity from the sequence of characters in a name. *arXiv preprint arXiv:1805.02109*.

Staudt, J. (2020). Mandating access: assessing the nih's public access policy. *Economic policy*, 35(102):269–304.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of econometrics*, 225(2):175–199.

Taffe, M. A. and Gilpin, N. W. (2021). Racial inequity in grant funding from the us national institutes of health. *Elife*, 10:e65697.

Thalmayer, A. G., Toscanelli, C., and Arnett, J. J. (2021). The neglected 95% revisited: Is american psychology becoming less american? *American Psychologist*, 76(1):116.

Wilhite, A. W. and Fong, E. A. (2012). Coercive citation in academic publishing. *Science*, 335(6068):542–543.

Williams, H. L. (2013). Intellectual property rights and innovation: Evidence from the human genome. *Journal of Political Economy*, 121(1):1–27.

Wu, A. H. (2020). Gender bias among professionals: an identity-based interpretation. *Review of Economics and Statistics*, 102(5):867–880.

Zacchia, G. (2021). What does it take to be top women economists? an analysis using rankings in repec. *Review of Political Economy*, 33(2):170–193.

Tables

Table I: Summary table

	All White	At least one black/hispanic/asian	Other races	Total
N	178,411	83,242	67,578	329,231
Number of authors	1.700	2.433	2.192	1.986
Publications as of 2021 (top 200)	19.968	19.333	15.617	18.914
Publications as of 2021 (top 35)	7.368	6.442	4.549	6.555
Publications as of 2021 (top 16)	3.670	3.346	2.173	3.281
Publications as of 2021 (top 5)	1.624	0.995	0.815	1.299
Journal rank (%)				
200 journals	70.041	74.352	77.684	72.700
Top 35	16.030	14.661	12.955	15.053
Top 16	7.310	7.881	5.713	7.126
Top 5	6.619	3.107	3.648	5.121
Fields (%)				
Econometrics	10.685	11.835	9.478	10.734
Microeconomics	19.135	15.025	17.527	17.742
Macroeconomics	8.139	6.688	7.648	7.663
International Economics	6.355	7.205	6.832	6.673
Finance	13.250	19.267	15.389	15.245
Labour & Education	13.191	10.095	11.986	12.143
Other (empirical) fields	26.213	27.043	28.629	26.924
Miscellaneous	3.031	2.841	2.511	2.875
Style (%)				
Mostly theoretical	65.291	55.934	60.986	61.987
Mostly empirical	29.286	35.188	33.495	31.677
Mostly structural	5.423	8.877	5.519	6.336
Affliation categories (%)				
Low Tier	63.595	63.482	68.441	64.561
Mid Tier	22.107	23.147	20.550	22.050
Top Tier	14.298	13.371	11.010	13.388

This table shows the summary statistics for the set of papers in the database.

Table II: Citations and race: baseline estimation

	Outcome variable: Citation (asinh)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
At least one Black/Hispanic/Asian	-0.139***	-0.256***	-0.150***	-0.059***	-0.058***	-0.071***	-0.051***
	(0.007)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.005)
Other races	-0.166***	-0.244***	-0.132***	-0.076***	-0.059***	-0.062***	-0.044***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.006)
Number of authors		0.159***	0.146***	0.228***	0.222***	0.191***	0.096***
		(0.005)	(0.005)	(0.008)	(0.008)	(0.007)	(0.004)
Prior publications at the time of publication			-0.378***	-0.141***	-0.165***	-0.150***	-0.120***
			(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Publications as of 2021			0.551***	0.266***	0.249***	0.250***	0.234***
			(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
Observations	329231	329231	329231	329231	329231	329231	329231
Adj. R-sqr	0.002	0.011	0.091	0.262	0.282	0.303	0.526
Year fixed effects				Y	Y	Y	
Affiliation fixed effects					Y	Y	Y
Field fixed effects						Y	Y
Journal-Year fixed effects							Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table III: Citation and race: effect across journal types

	Outcome variable: Citation (asinh)			
	More than 200 journals	Top 35 journals (Card et al. [2020])	Top 16 journals (Koffi [2021b])	Top 5 journals
	(1)	(2)	(3)	(4)
At least one Black/Hispanic/Asian	-0.051***	-0.086***	-0.085***	-0.096***
	(0.005)	(0.010)	(0.015)	(0.028)
Other races	-0.044***	-0.050***	-0.055***	-0.072**
	(0.006)	(0.012)	(0.018)	(0.031)
Number of authors	0.096***	0.093***	0.124***	0.108***
	(0.004)	(0.006)	(0.009)	(0.015)
Prior publications at the time of publication	-0.120***	-0.123***	-0.167***	-0.211***
	(0.003)	(0.006)	(0.009)	(0.015)
Publications as of 2021	0.234***	0.291***	0.348***	0.415***
	(0.003)	(0.006)	(0.009)	(0.015)
Observations	329231	89881	40322	16860
Adj. R-sqr	0.526	0.455	0.485	0.413
Affiliation fixed effects	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Column (1) reports the results of Table II- Column (7). Column (2) restricts the sample to 35 high-impact journals identified by Card et al. [2020]. Column (3) limits the sample to 16 high-impact journals according to Koffi [2021b]. Column (4) restricts the sample to the top five economics journals: *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table IV: Heterogeneous effect: Number of authors

	Outcome variable: Citation (asinh)		
	Solo Author	Two Authors	More than Two authors
	(1)	(2)	(3)
At least one Black/Hispanic/Asian	-0.050***	-0.072***	-0.032***
	(0.011)	(0.008)	(0.009)
Other races	-0.053***	-0.057***	-0.033***
	(0.010)	(0.009)	(0.010)
Observations	122779	122380	84072
R-sqr	0.458	0.535	0.613
Publications and team size controls	Y	Y	Y
Affiliation fixed effects	Y	Y	Y
Field fixed effects	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, focusing on the heterogeneous results depending on the number of authors. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table V: Selection on unobservables- Estimation with and without the full set of controls

	Outcome variable: Citation (asinh)							
	> 200 journals		Top 35 journals		Top 16 journals		Top 5 journals	
	Uncontrolled	Controlled	Uncontrolled	Controlled	Uncontrolled	Controlled	Uncontrolled	Controlled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At least one Black/Hispanic/Asian	-0.139***	-0.051***	-0.211***	-0.086***	-0.231***	-0.085***	-0.068*	-0.096***
	(0.007)	(0.005)	(0.014)	(0.010)	(0.021)	(0.015)	(0.037)	(0.028)
	Disaggregation							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At least one Black	-0.147***	-0.067***	-0.107	-0.110**	-0.131	-0.186***	0.053	-0.200*
	(0.031)	(0.021)	(0.068)	(0.046)	(0.100)	(0.069)	(0.166)	(0.120)
At least one Hispanic	-0.218***	-0.062***	-0.351***	-0.073***	-0.323***	-0.070***	-0.055	-0.147***
	(0.012)	(0.008)	(0.026)	(0.018)	(0.041)	(0.027)	(0.066)	(0.044)
At least one Asian	-0.111***	-0.046***	-0.176***	-0.089***	-0.209***	-0.085***	-0.079*	-0.071**
	(0.008)	(0.006)	(0.015)	(0.012)	(0.023)	(0.017)	(0.043)	(0.032)
Other races	-0.166***	-0.044***	-0.259***	-0.050***	-0.304***	-0.055***	-0.287***	-0.073**
	(0.008)	(0.006)	(0.015)	(0.012)	(0.024)	(0.018)	(0.039)	(0.031)
Observations	329231	329231	89881	89881	40322	40322	16860	16860
R-sqr	0.002	0.526	0.005	0.455	0.006	0.485	0.003	0.413
Publications and team size controls		Y		Y		Y		Y
Affiliation fixed effects		Y		Y		Y		Y
Field fixed effects		Y		Y		Y		Y
Journal-Year fixed effects		Y		Y		Y		Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Uneven columns represent the regression model without control variables. Even columns represent the regression model with the full set of control variables. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table VI: Robustness: Adding granular field controls

	Outcome variable: Citation (asinh)							
	Baseline Papers with JEL codes	1-digit JEL code	2-digit JEL code	3-digit JEL code	768 Embeddings	Style	3-digit JEL code + Style	3-digit JEL code + Style+ Embeddings
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At least one Black/Hispanic/Asian	-0.049***	-0.070***	-0.070***	-0.065***	-0.069***	-0.059***	-0.063***	-0.064***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)
Other races	-0.035***	-0.041***	-0.040***	-0.038***	-0.048***	-0.043***	-0.036***	-0.037***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.005)	(0.006)	(0.007)	(0.006)
Probability of theoretical content						-0.414***	-0.261***	0.028
						(0.019)	(0.026)	(0.051)
Probability of empirical content						0.046**	0.201***	0.054
						(0.019)	(0.026)	(0.059)
Observations	208045	208045	208045	208045	317649	317649	207118	207118
R-sqr	0.538	0.540	0.545	0.553	0.566	0.536	0.559	0.591
Publications and team size controls	Y	Y	Y	Y	Y	Y	Y	Y
Affiliation fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, adding granular controls for the fields and style of methodology. The reference category for the race variable is *All-White*, which represents papers written by only White authors. The sample size in this table differs from that in Table II due to variations in data availability: not all papers have their JEL codes available (Columns 1 to 4), not all papers have their abstracts available (Columns 5 and 6), and not all papers have both JEL codes and abstracts (Columns 7 and 8). Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table VII: Robustness: Centrality measures

	Outcome variable: Standardized centrality measures					
	In-degree	Betweeness	Eigenvector	Katz	PageRank	Clustering
	(1)	(2)	(3)	(4)	(5)	(6)
At least one Black/Hispanic/Asian	-0.027***	-0.020***	-0.006***	-0.016***	-0.024***	0.041***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.003)	(0.004)
Other races	-0.025***	-0.013***	-0.005*	-0.013***	-0.020***	0.014***
	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)
Observations	329231	329231	329231	329231	329231	314706
Adj. R-sqr	0.174	0.111	0.034	0.084	0.128	0.105
Publications and team size controls	Y	Y	Y	Y	Y	Y
Affiliation fixed effects	Y	Y	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y	Y	Y

This table shows the relationship between the centrality measures and the racial composition of the paper's authors, adding granular controls for the fields and style of methodology. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table VIII: Robustness: Alternative race definitions

	Outcome variable: Citation (asinh)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
At least one Black/Hispanic/Asian	-0.038***								
	(0.005)								
At least one Black/Hispanic/Asian		-0.038***							
		(0.005)							
At least one Black/Hispanic/Asian			-0.043***						
			(0.005)						
Other races			-0.040***						
			(0.006)						
At least one Black/Hispanic/Asian				-0.051***					
				(0.005)					
Other races				-0.044***					
				(0.006)					
At least one Black/Hispanic/Asian					-0.044***				
					(0.005)				
Other races					-0.052***				
					(0.006)				
At least one Black/Hispanic/Asian						-0.048***			
						(0.005)			
Other races						-0.022***			
						(0.008)			
At least one Black/Hispanic/Asian							-0.044***		
							(0.006)		
Other races							-0.024***		
							(0.005)		
At least one Black/Hispanic/Asian								-0.042***	
								(0.006)	
Other races								-0.051***	
								(0.005)	
At least one Black/Hispanic/Asian									-0.050***
									(0.005)
Other races									-0.024***
									(0.008)
Observations	329231	329231	329231	329231	329231	329231	329231	329231	329231
R-sqr	0.526	0.526	0.526	0.526	0.526	0.526	0.526	0.526	0.526
Publications and team size controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Affiliation fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, using various race definitions. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Columns (1) and (2) define race using prediction from the US census and Florida voter registration, respectively. There is no threshold restriction, so the category "Other races" is not included. Column (3) employs Florida voter registration if the author's first name is available and the US census database for last names only. Column (4) is the baseline estimation in Table II. Black economists are additionally identified through manual checks and a list from the National Economic Association. Column (5) is similar to Column (4) but without the probability threshold. Column (6) also uses a combination of US census and Florida voter registration data with only a 0.7 probability threshold, including manual checks for Black economists. Column (7) is like Column (6) with a 0.9 probability threshold. Column (8) is akin to Column (4) but excludes manual checks for Black economists. Column (9) resembles Column (6) but also omits manual checks for Black economists. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table IX: Robustness: Alternative race definitions (2)

	Outcome variable: Citation (asinh)		
	(1)	(2)	(3)
Share of Black/Hispanic/Asian	-0.049***		
	(0.007)		
Share of other races	-0.068***		
	(0.007)		
All Black/Hispanic/Asian		-0.042***	
		(0.008)	
At least one Black/Hispanic/Asian (not all)		-0.057***	
		(0.007)	
Other races		-0.044***	
		(0.006)	
All Black/Hispanic/Asian			-0.043***
			(0.008)
One Black/Hispanic/Asian and one White			-0.044***
			(0.009)
50% White (not teams of 2)			-0.021***
			(0.008)
At least one Black/Hispanic/Asian (not all)			-0.063***
			(0.008)
Other races			-0.060***
			(0.007)
Observations	329231	329231	329231
R-sqr	0.526	0.526	0.526
Publications and team size controls	Y	Y	Y
Affiliation fixed effects	Y	Y	Y
Field fixed effects	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, using various race definitions. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Column (1) uses the share of authors identified as Black, Hispanic, and Asian and the share of authors that are neither Black, Hispanic, Asian, nor White. Column (2) is a categorical variable equal to 0 for all-White authors, 1 for all Black, Hispanic, or Asian authors, 2 for teams with at least one Black, Hispanic, or Asian author (excluding teams where all authors are Black, Hispanic, or Asian), and 3 for other races. Column (3) is a categorical variable equal to 0 for all-White authors, 1 for all Black, Hispanic, or Asian authors, 2 for teams with one Black, Hispanic, or Asian author and one White author, 3 for teams where 50% of the authors are White and 50% are non-White, 4 for teams with at least one Black, Hispanic, or Asian author (excluding teams in the previous categories), and 5 for other races. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table X: Robustness: Readability index

	Outcome variable: Citation (asinh)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At least one Black/Hispanic/Asian	-0.051***	-0.054***	-0.055***	-0.055***	-0.054***	-0.054***	-0.055***	-0.058***
	(0.005)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Other races	-0.044***	-0.044***	-0.044***	-0.044***	-0.044***	-0.044***	-0.044***	-0.044***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Flesh reading index			-0.003***					-0.002***
			(0.000)					(0.000)
Flesh Kincaid index				0.011***				0.033***
				(0.001)				(0.003)
Smog index					0.016***			0.012***
					(0.001)			(0.001)
Gunning Fog index						0.005***		-0.035***
						(0.001)		(0.002)
Dale-Shall index							-0.018***	0.002
							(0.002)	(0.002)
Observations	329231	317649	317649	317649	317649	317649	317649	317649
R-sqr	0.526	0.530	0.531	0.531	0.531	0.531	0.531	0.532
Publications and team size controls	Y	Y	Y	Y	Y	Y	Y	Y
Affiliation fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, adding controls for the readability of the abstract of the paper. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Columns (1) and (2) are related to the baseline estimation without restriction and with restriction on papers for which we could compute the readability scores. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table XI: Robustness: Functional form of the citation variable

	Citation count	log(1+citation)	Poisson model	Top 2 percentile of citation distribution	Top 5 percentile of citation distribution	Top 10 percentile of citation distribution	Remove papers with 0 citation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
At least one Black/Hispanic/Asian	-4.484***	-0.048***	-0.126***	-0.004***	-0.006***	-0.010***	-0.058***
	(0.720)	(0.005)	(0.020)	(0.001)	(0.001)	(0.001)	(0.005)
Other races	-4.402***	-0.041***	-0.118***	-0.002***	-0.005***	-0.008***	-0.048***
	(0.759)	(0.005)	(0.021)	(0.001)	(0.001)	(0.001)	(0.005)
Observations	329231	329231	328852	329231	329231	329231	293435
R-sqr	0.086	0.520		0.101	0.160	0.201	0.440
Publications and team size controls	Y	Y	Y	Y	Y	Y	Y
Affiliation fixed effects	Y	Y	Y	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y	Y	Y	Y

This table shows the relationship between the citation and race, using alternative functional forms of the dependent variable. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Columns (4) to (6), in particular, estimate a linear probability model where the outcome variable is 1 if the paper's citations are in the top 2% (Column 4), top 5% (Column 5), or top 10% (Column 6) of the citation distribution and 0 otherwise. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table XII: Robustness: Alternative measures of authors' prominence

	Outcome variable: Citation (asinh)			
	>200	Top 35	Top 16	Top 5
	Measures based on the more prolific author			
	(1)	(2)	(3)	(4)
At least one Black/Hispanic/Asian	-0.053***	-0.078***	-0.067***	-0.091***
	(0.005)	(0.010)	(0.015)	(0.028)
Other races	-0.036***	-0.042***	-0.040**	-0.059*
	(0.006)	(0.012)	(0.018)	(0.031)
Observations	329231	89881	40322	16860
Adj. R-sqr	0.529	0.459	0.487	0.417
	Measures based on the average number of publications			
	(1)	(2)	(3)	(4)
At least one Black/Hispanic/Asian	-0.063***	-0.088***	-0.082***	-0.102***
	(0.006)	(0.011)	(0.016)	(0.028)
Other races	-0.038***	-0.055***	-0.060***	-0.085***
	(0.006)	(0.012)	(0.018)	(0.031)
Observations	329231	89881	40322	16860
Adj. R-sqr	0.524	0.451	0.480	0.405
	Measures based on the total number of publications			
	(1)	(2)	(3)	(4)
At least one Black/Hispanic/Asian	-0.054***	-0.079***	-0.077***	-0.103***
	(0.005)	(0.011)	(0.016)	(0.029)
Other races	-0.037***	-0.035***	-0.030	-0.062*
	(0.005)	(0.012)	(0.019)	(0.033)
Observations	329231	89881	40322	16860
Adj. R-sqr	0.533	0.463	0.492	0.419
Affiliation fixed effects	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, emphasizing the effect of seniority. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Seniority is measured by the authors' publication records and based on the maximum number of publications, the average number of publications, and the total number of publications by the team. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table XIII: Citation and race: control for past citations

	Outcome variable: Citation (asinh)			
	More than 200 journals	Top 35 journals (Card et al. [2020])	Top 16 journals (Koffi [2021b])	Top 5 journals
	(1)	(2)	(3)	(4)
At least one Black/Hispanic/Asian	-0.043***	-0.068***	-0.060***	-0.083***
	(0.006)	(0.010)	(0.015)	(0.028)
Other races	-0.037***	-0.039***	-0.039**	-0.058*
	(0.006)	(0.012)	(0.018)	(0.031)
Past citations (1)	0.125***	0.098***	0.104***	0.067
	(0.010)	(0.021)	(0.034)	(0.052)
Past citations (1-3.6)	0.197***	0.170***	0.151***	0.124***
	(0.010)	(0.020)	(0.031)	(0.048)
Past citations (3.6-9)	0.265***	0.240***	0.228***	0.166***
	(0.010)	(0.020)	(0.029)	(0.046)
Past citations (9-20)	0.348***	0.348***	0.341***	0.291***
	(0.012)	(0.021)	(0.030)	(0.047)
Past citations (20-47)	0.412***	0.427***	0.435***	0.313***
	(0.013)	(0.023)	(0.032)	(0.050)
Past citations (47-128)	0.508***	0.521***	0.566***	0.449***
	(0.015)	(0.025)	(0.034)	(0.054)
Past citations (>128)	0.655***	0.659***	0.705***	0.503***
	(0.017)	(0.030)	(0.041)	(0.063)
Constant	2.606***	2.797***	2.897***	3.513***
	(0.042)	(0.084)	(0.101)	(0.159)
Observations	329231	89881	40322	16860
Adj. R-sqr	0.529	0.458	0.489	0.416
Affiliation fixed effects	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors while adding controls for authors' past citations. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Column (1) reports the results of Table II- Column (7). Column (2) restricts the sample to 35 high-impact journals identified by Card et al. [2020]. Column (3) limits the sample to 16 high-impact journals according to Koffi [2021b]. Column (4) restricts the sample to the top five economics journals: *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*. Standard errors are reported in parentheses. The "past citations" variable is a categorical variable based on the decile of the average team citations distribution. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table XIV: Citations and the race of the authors of the citing paper

	Outcome variable: Citation (asinh)							
	> 200 journals		Top 35 journals		Top 16 journals		Top 5 journals	
	From non-White	From White	From non-White	From White	From non-White	From White	From non-White	From White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At least one Black/Hispanic/Asian	0.232***	-0.303***	0.211***	-0.338***	0.212***	-0.323***	0.215***	-0.279***
	(0.004)	(0.004)	(0.010)	(0.010)	(0.015)	(0.015)	(0.027)	(0.028)
Other races	0.012***	-0.180***	0.008	-0.185***	0.014	-0.165***	0.029	-0.144***
	(0.004)	(0.005)	(0.010)	(0.011)	(0.017)	(0.017)	(0.028)	(0.029)
Observations	329231	329231	89881	89881	40322	40322	16860	16860
R-sqr	0.419	0.460	0.420	0.440	0.427	0.413	0.428	0.367

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, emphasizing the race of the authors of the citing paper. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Columns (1), (3), (5), and (7) restrict the citation count to citations from only teams including non-White scholars. Columns (2), (4), (6), and (8) restrict the citation count to citations from only teams, including all-White scholars. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table XV: Effect of seminars and conference attendance

	Outcome variable: Citation (asinh)							
	> 200 journals		Top 35 journals		Top 16 journals		Top 5 journals	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At least one Black/Hispanic/Asian	-0.022	-0.020	-0.087***	-0.083***	-0.063*	-0.057	-0.100	-0.094
	(0.015)	(0.015)	(0.025)	(0.025)	(0.038)	(0.037)	(0.067)	(0.067)
Other races	-0.037**	-0.037**	-0.026	-0.027	0.017	0.013	0.009	0.008
	(0.017)	(0.017)	(0.028)	(0.027)	(0.044)	(0.043)	(0.072)	(0.071)
Seminars/conferences count		0.010***		0.011***		0.009***		-0.002
		(0.002)		(0.002)		(0.003)		(0.005)
Prominent conferences		0.015		0.006		0.046		-0.010
		(0.034)		(0.039)		(0.052)		(0.094)
Seminars in top-tier affiliations		0.024***		0.015		0.003		0.034*
		(0.009)		(0.010)		(0.013)		(0.019)
Observations	22352	22352	8170	8170	3538	3538	1192	1192
Adj. R-sqr	0.648	0.649	0.635	0.637	0.629	0.631	0.656	0.656
Publications and team size controls	Y	Y	Y	Y	Y	Y	Y	Y
Affiliation fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, emphasizing the effect of seminars and conference attendance. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Uneven columns represent the regression model without control variables. Even columns represent the regression model with the full set of control variables. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figures

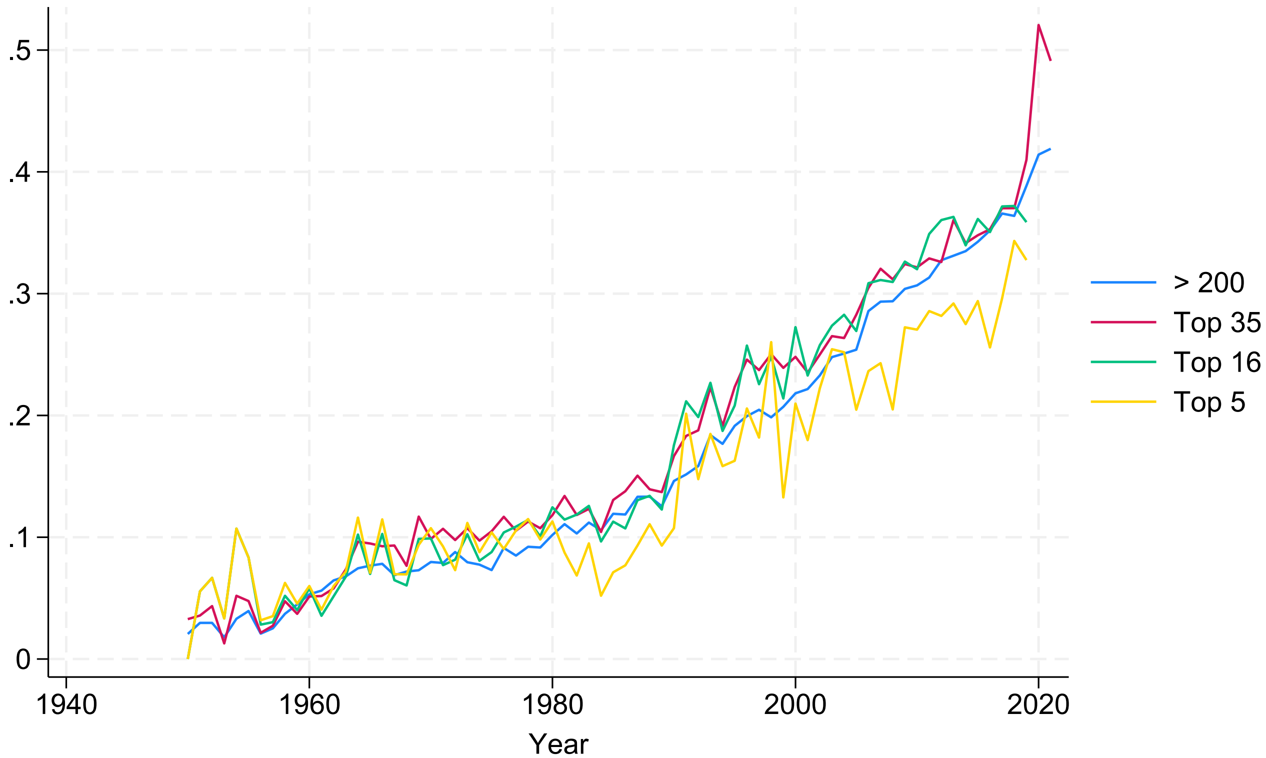


Figure I:
Share of publications with at least Black, Hispanic, or Asian over time

The figure presents evolution over time of the share of papers with at least one Black, Hispanic, or Asian for different sets of journals (whole sample with more than 200 journals, a set of 35 high-impact journals as per Card et al. [2020], for a set of top 16 journals used in Koffi [2021b], and the traditional top 5 high impact journals.

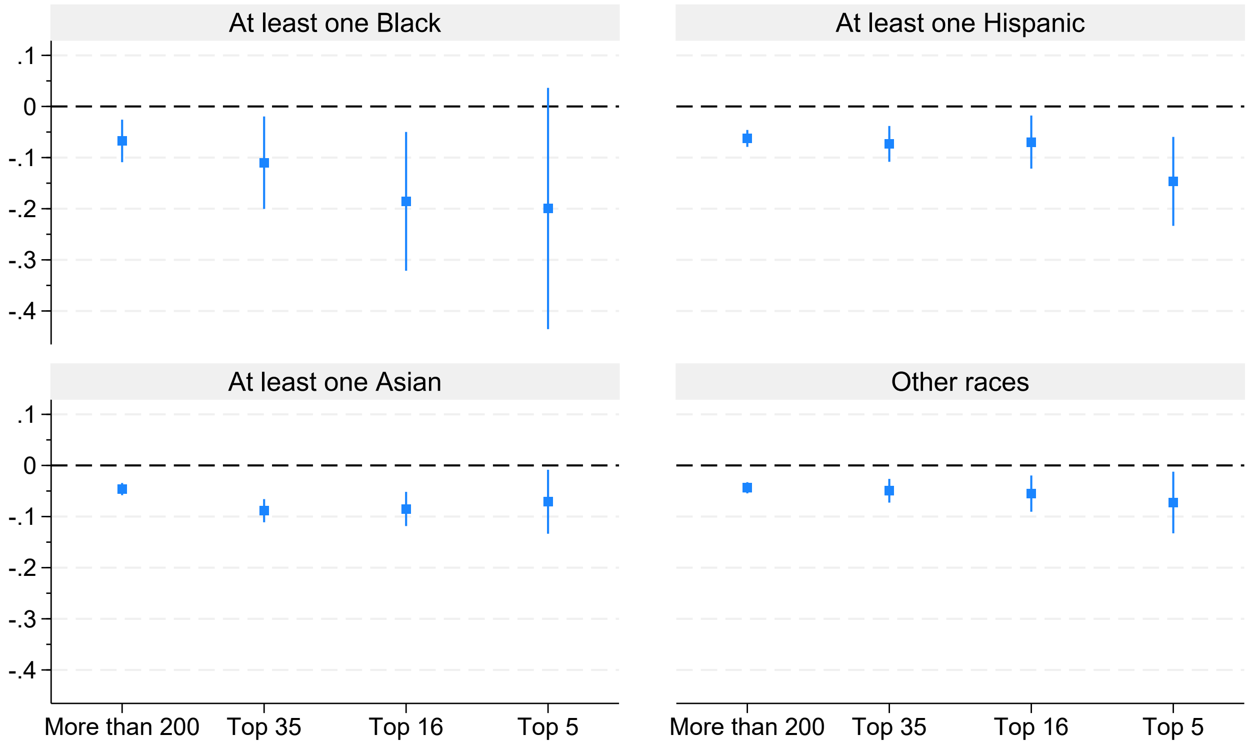


Figure II:
Citation and race

The figure presents the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors. The reference category for the race variable is *All-White*, which represents papers written by only White authors. The figure shows the estimation results by specific racial groups and journal rank.

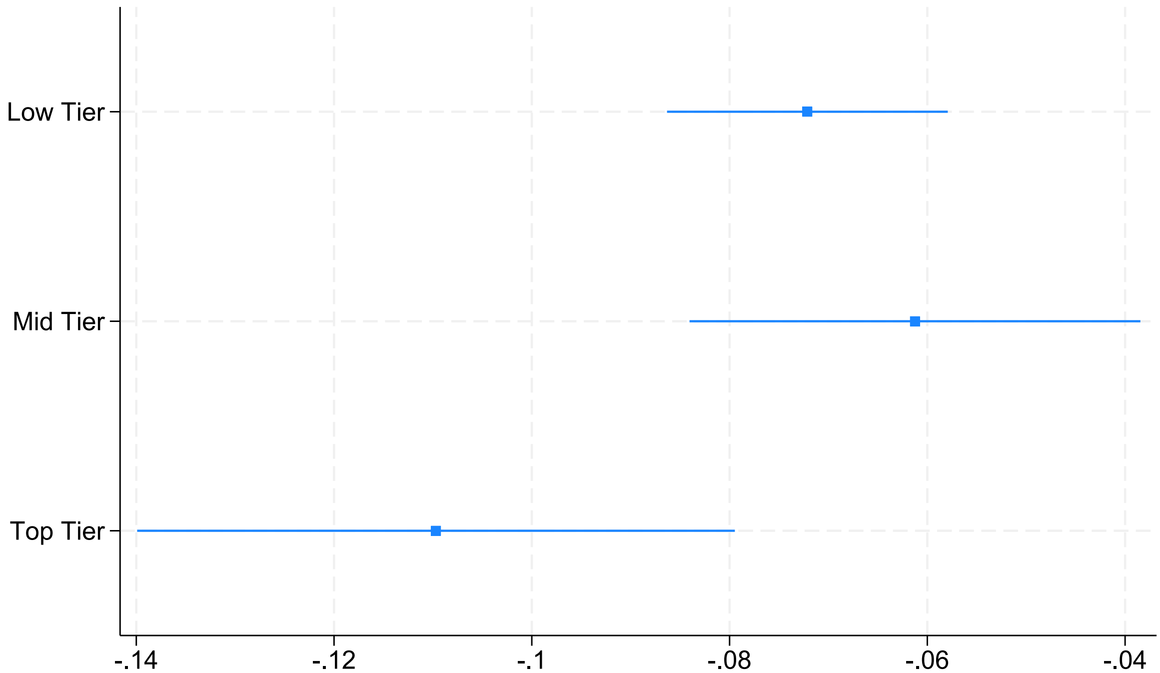


Figure III:
Citation and race: heterogeneity by affiliations

The figure presents the relationship between citation and race, focusing on the heterogeneous results depending on the authors' affiliations rankings.

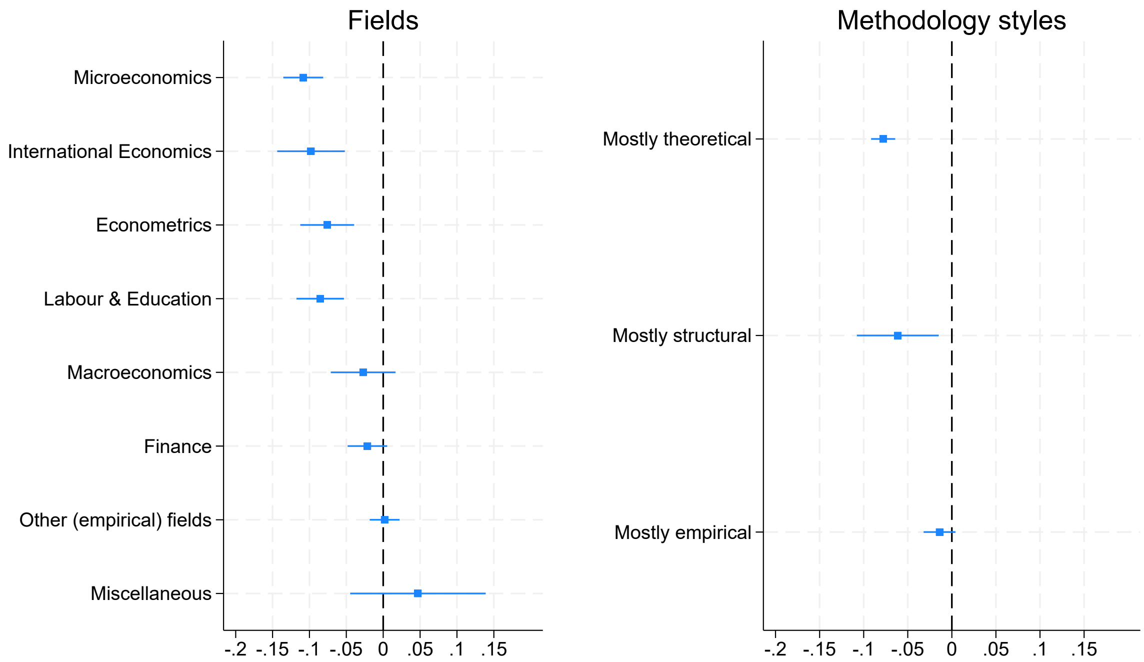


Figure IV:
Citation and race: heterogeneity by fields and methodology styles

The figure presents the relationship between citation and race, focusing on the heterogeneous results depending on the primary field and the methodology style of the paper.

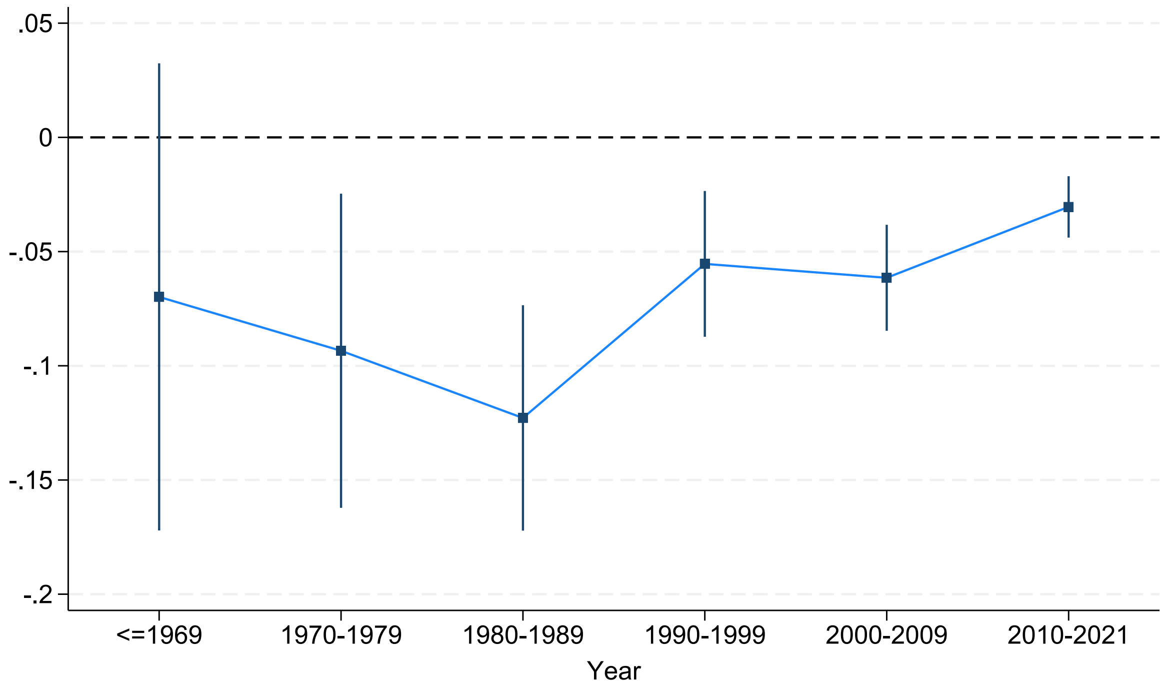
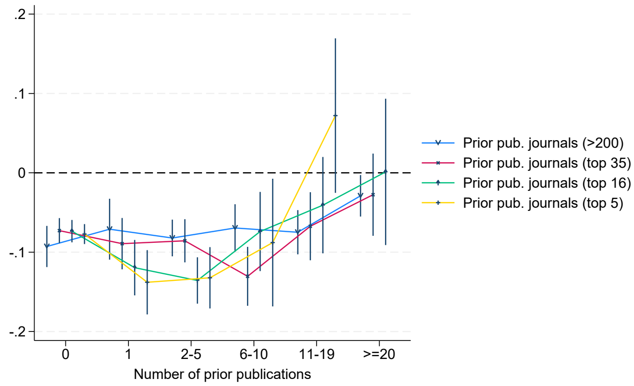
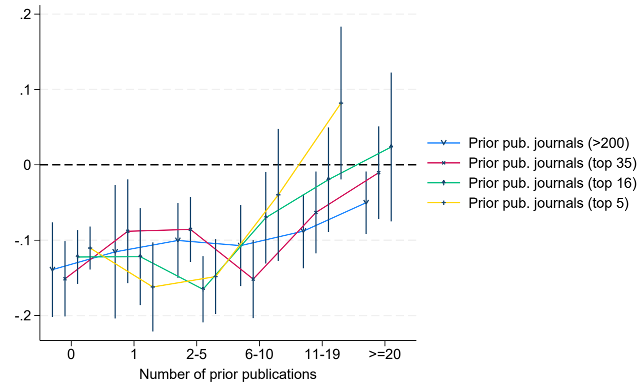


Figure V:
Citation and race: Evolution over time

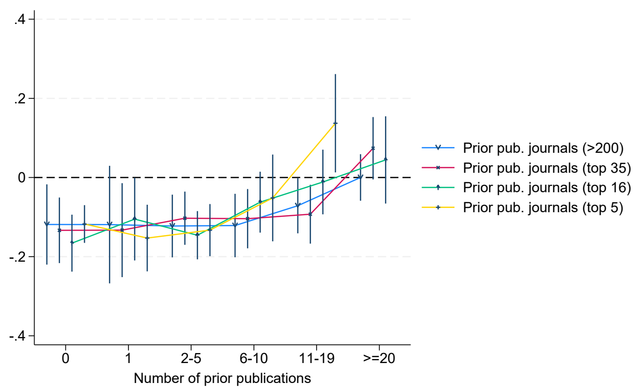
The figure presents the relationship between citation and race, focusing on the evolution over time.



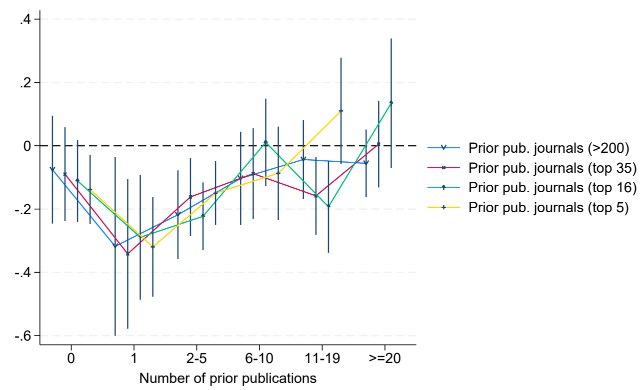
(a) Gap in more than 200 journals



(b) Gap in top 35 journals



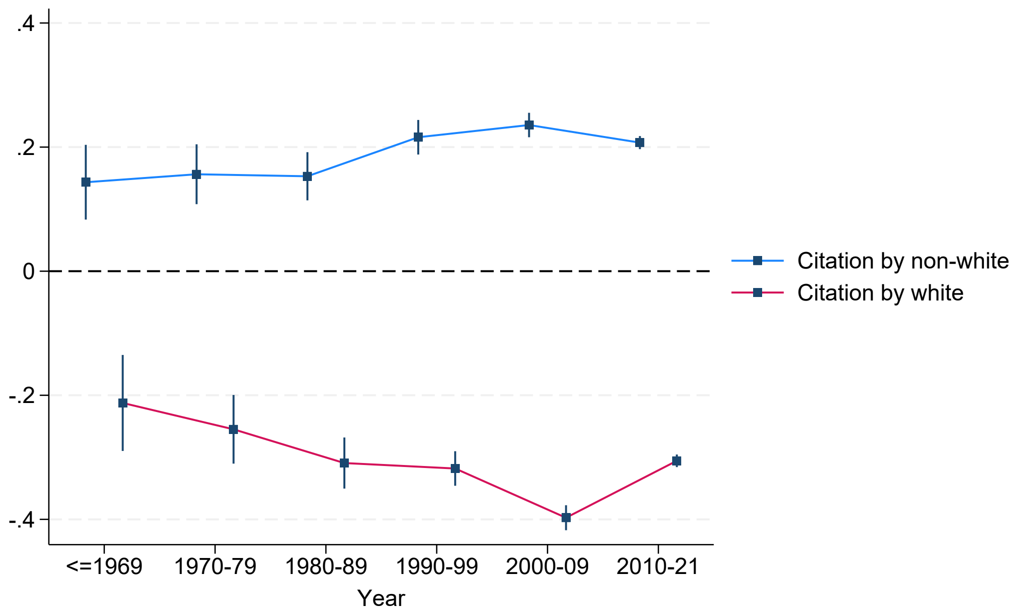
(c) Gap in top 16 journals



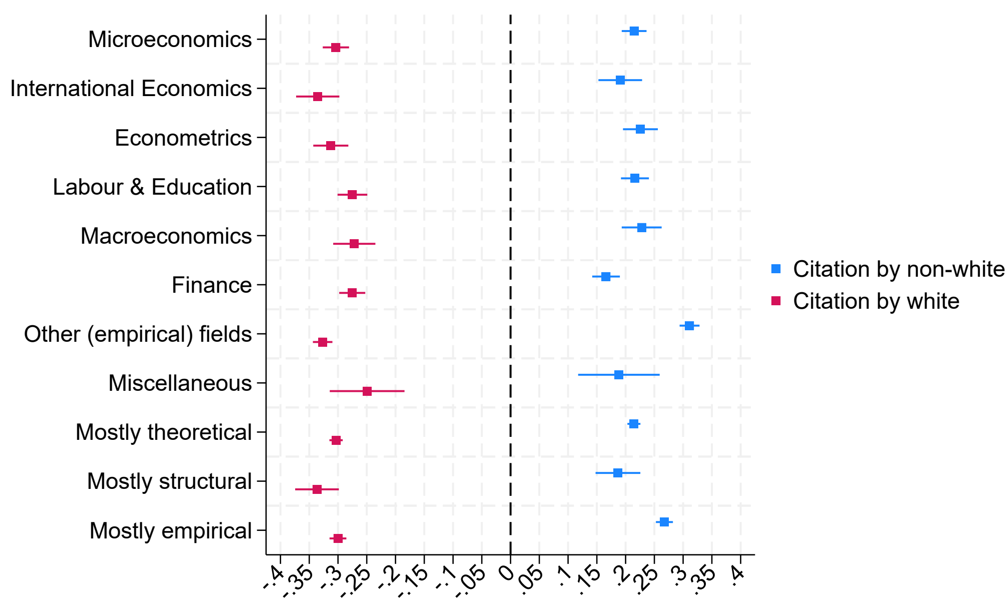
(d) Gap in top 5 journals

Figure VI:
Citation, race and prior publications

The figure presents the relationship between the citation (inverse hyperbolic sine—asinh), the outcome variable, and the racial composition of the paper's authors, the independent variable of interest, focusing on the effect of an additional publication. The reference category for the race variable is *All-White*, which represents papers written by only White authors. The plots display the estimate for teams with at least one Black/Hispanic/Asian author by comparison to a team with all-White authors.



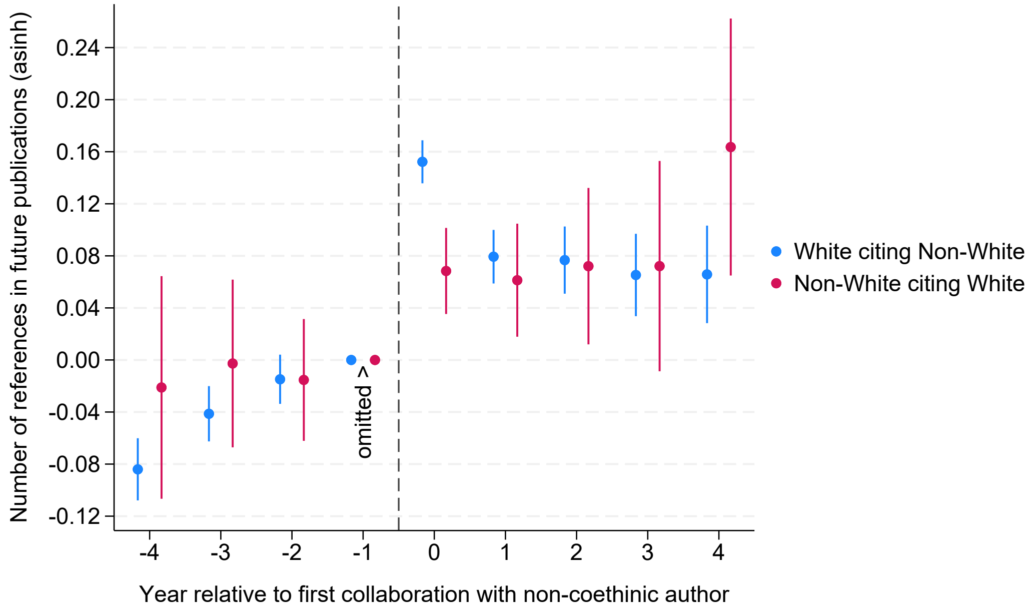
(a) Non-White/All-White citation penalty over premium over time



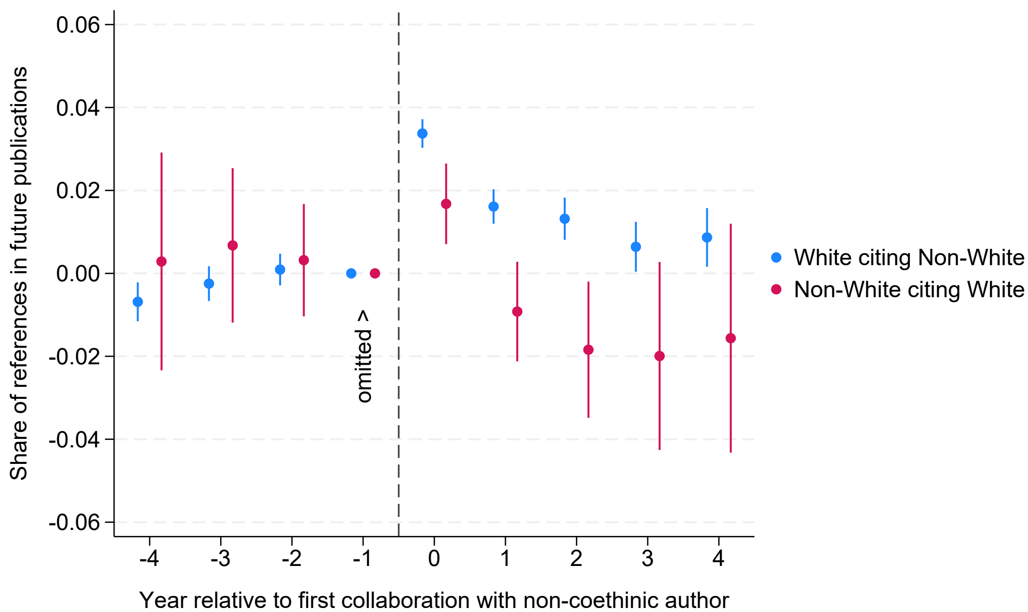
(b) Non-White/All-White citation penalty across fields and methodology

Figure VII:
Citation and race of the authors in the citing paper

The figure presents the relationship between citation and race, focusing on the racial composition of the authors of the citing paper. The reference category for the race variable is *All-White*, which represents papers written by only White authors. The plots display the estimate for teams with at least one Black/Hispanic/Asian author by comparison to a team with all-White authors.



(a) Number of listed bibliographical references



(b) Share of listed bibliographical references

Figure VIII:
Effect of collaboration with a non-coethnic author

The figure presents the event study showing the impact of the collaboration with a non-coethnic author on the number and the share of listed bibliographical references to papers with authors of the out-group. To estimate this effect, we consider an author-level estimation and analyze all subsequent publications with only coethnic authors following the first collaboration with a non-coethnic author. We truncate the sample to the next non-coethnic collaboration.

Appendix For Online Publication
The Color of Ideas: Racial Dynamics and Citations in Economics

Marlène Koffi & Roland Pongou & Leonard Wantchekon

A Name disambiguation

A sequential process then parsed each name into first, middle, and last names, with thorough checks at each step to minimize errors.

When complete first, middle, and last names are available, they are considered to belong to the same author. Subsequently, if both first and last names are present, along with an initial middle name, and only one candidate is identified, these names are merged. Following this, cases in which names comprise only the initials of the first and middle names are examined. Finally, cases in which middle names are absent are addressed. In such instances, the frequency of the first and last name in the database is analyzed, considering the relative proportion of the first name in the database and utilizing surname classifications from the US census data. For instance, a name like "Smith," which is very popular, is excluded from the process, and disambiguation was manually checked whenever possible.

Chinese names were systematically excluded during this process, and the disambiguation of those names occurred primarily through MAG disambiguation. While disambiguation procedures are inherently imperfect, this multi-staged process enables the creation of a comprehensive and meaningful publication history for the authors.

B Centrality measures

We compute measures of centrality to assess the importance of papers. Centrality measures account for different notions of importance in a (citation) network; for example, we might only be interested in counting the number of citations received by a paper, or we might also be interested in how citing papers are cited by other papers. We consider six well-known notions of centrality measures: degree centrality, betweenness centrality, eigenvector centrality, Katz centrality, Pagerank centrality, and clustering coefficient.

To compute the values of these centrality measures, we assume that all papers form a network of citations, where a paper i is linked to another paper j if i has received a citation from j. Clearly, this is a directed network as a paper j may cite i without the converse being true. Assume there are n published papers. The citation network among these papers can be represented by an adjacent matrix $A=(a_{ij})$, where for any nodes i and j, the matrix entry or cell $a_{ij}=1$ if i has received a citation from j and 0 if not. Below, we formally define the six centrality measures we use.

- The degree centrality of a paper i, denoted by χ_i, is the total number of citations that i has received from other papers: $\chi_i = \sum\limits_{j=1}^{n} a_{ij}$.

- The betweenness centrality of a paper i, denoted by b_i, is the fraction of the shortest paths between two other papers passing through i: $b_i = \sum_{j,k} \frac{\sigma^i_{jk}}{\sigma_{jk}}$, where σ_{jk} is the total number of shortest paths from agents j to k (with i different from both j and k), and σ^i_{jk} is the total number of shortest paths that pass through paper i.

- The eigenvector centrality is a measure of prestige that reflects the notion that being cited by other highly cited papers is important. The eigenvector centrality of a paper i, denoted by ν_i, measures the extent to which i is cited by other highly cited papers. It is given by the formula $\nu_i = \frac{1}{e} \sum_{j=1}^{n} a_{ij}\nu_j$, where e is the largest eigenvalue and ν (a column vector with n entries) the corresponding eigenvector of the adjacency matrix A: $A\nu = e\nu$.

- The Katz centrality, similar to the eigenvector centrality, values citations from papers that are also cited; however, having distant neighbors penalize a node's score by some factor α. The Katz centrality of a paper i, denoted by K_i, is defined as $K_i = \sum_{k=1}^{\infty} \sum_{j=1}^{n} \alpha^k (A^k)_{ji}$, where $(A^k)_{ji}$ is the number of k degree connections between papers i and j and α is an attenuation factor (strictly between 0 and 1) chosen to be smaller than the inverse of the largest eigenvalue of the adjacency matrix A. The fact α^k is a decreasing function of k reflects the fact that the Katz centrality weights a link to a more distant paper less.

- The Pagerank centrality is also similar to the eigenvector centrality. The PageRank centrality of a paper i, denoted by PR_i, is recursively defined as: $PR_i = \alpha \sum_j a_{ji} \frac{PR_j}{\chi_j} + \frac{1-\alpha}{n}$, where χ_j is the number of citations received by j, and α is a dampening factor that is positive and strictly lower than 1. item The clustering coefficient of a paper i, denoted by C_i, is a measure of the extent to which papers that cite paper i tend to cite each other, leading to a tightly knit group of papers. Let $N_i = \{j : a_{ij} = 1\}$ be the set of papers that cite paper i, and let $|N_i|$ be the cardinality of this set. The clustering coefficient of paper i is given by: $C_i = \frac{|\{a_{jk}: a_{jk}=1, j,k \in N_i\}|}{|N_i||N_i-1|}$.

To process our data from a panel data into a network structure, we first create a table of directed links so that every row observation is some article j citing some article i, and then convert this to a *directed* network object using the python library NetworkX; this is then converted to a Networkit object. Because our data is in a node-to-node panel format, a first transformation to a NetworkX graph object is necessary prior to the conversion to a Networkit directed graph. Then, using the generated directed network object, we calculate the above centrality measures to be accurate at the 6 significant figure level using built-in functionality with the NetworkX library. In particular, we were concerned with processing times for loading the dataset, shortest path calculations, PageRank calculations as they most relate to our work, and built-in capabilities to multi-thread the process.

C Additional Tables

C.1 Citations and race: baseline estimation on manually checked sample

	Outcome variable: Citation (asinh)						
	Checked race on checked sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
At least one Black/Hispanic/Asian	-0.335***	-0.303***	-0.235***	-0.084***	-0.100***	-0.094***	-0.075**
	(0.050)	(0.051)	(0.046)	(0.031)	(0.030)	(0.030)	(0.029)
Mixed races	-0.440	-0.392	-0.330	-0.093	-0.247	-0.214	-0.255
	(0.557)	(0.547)	(0.498)	(0.247)	(0.220)	(0.211)	(0.202)
Observations	5544	5544	5544	5544	5544	5544	5544
Adj. R-sqr	0.008	0.010	0.182	0.629	0.643	0.660	0.684
	Algorithmic-based race on checked sample						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
At least one Black/Hispanic/Asian	-0.357***	-0.318***	-0.252***	-0.132***	-0.144***	-0.129***	-0.099***
	(0.054)	(0.055)	(0.050)	(0.034)	(0.033)	(0.033)	(0.033)
Other races	-0.386***	-0.356***	-0.225***	-0.013	-0.023	-0.016	0.023
	(0.061)	(0.062)	(0.057)	(0.038)	(0.037)	(0.036)	(0.036)
Observations	5544	5544	5544	5544	5544	5544	5544
Adj. R-sqr	0.011	0.013	0.182	0.629	0.644	0.660	0.684
Team size controls		Y	Y	Y	Y	Y	Y
Publications controls			Y	Y	Y	Y	Y
Year fixed effects				Y	Y	Y	
Affiliation fixed effects					Y	Y	Y
Field fixed effects						Y	Y
Journal-Year fixed effects							Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors. The estimation compared the sample where the race of the authors was checked when we consider the manually annotated race to the case of algorithmic-based determination of the race. The reference category for the race variable is *All-White*, which represents papers written by only White authors. "Mixed races" refers specifically to individuals identified as belonging to more than one racial category. In contrast, "Other races" is used to capture uncertainty in classification as described in Section 3.2.1. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.2 Regression excluding self-citations

	Outcome variable: Citation (asinh)			
	> 200 journals	Top 35 journals	Top 16 journals	Top 5 journals
	(1)	(2)	(3)	(4)
At least one Black/Hispanic/Asian	-0.045***	-0.077***	-0.082***	-0.099***
	(0.006)	(0.012)	(0.018)	(0.030)
Other races	-0.042***	-0.044***	-0.052***	-0.071**
	(0.006)	(0.013)	(0.020)	(0.033)
Observations	329231	89881	40322	16860
Adj. R-sqr	0.518	0.480	0.513	0.444
Affiliation fixed effects	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, excluding self-citations from the total citation count. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.3 Citation and race: paper-time level regression

	Outcome variable: Citation (asinh)							
	> 200 journals		Top 35 journals		Top 16 journals		Top 5 journals	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At least one Black/Hispanic/Asian	-0.012***	-0.012***	-0.018***	-0.019***	-0.023***	-0.028***	-0.032***	-0.037***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Other races	-0.005***	-0.004***	-0.004***	-0.005***	0.002	-0.001	0.007***	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
Observations	6606106	6606106	2251757	2251757	1070057	1070057	570892	570892
Adj. R-sqr	0.406	0.406	0.471	0.471	0.508	0.509	0.525	0.527
Publications and team size controls	Y	Y	Y	Y	Y	Y	Y	Y
Affiliation fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y
Paper cumulative citation up to time t	Y	Y	Y	Y	Y	Y	Y	Y
Publications in top 5, top 16, top 35, overall at time t	Y	Y	Y	Y	Y	Y	Y	Y
Average cumulative citations of the authors up to time t		Y		Y		Y		Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Columns (1) and (2) use the full baseline sample. Columns (3) and (4) restrict the sample to 35 high-impact journals identified by Card et al. [2020]. Columns (5) and (6) limit the sample to 16 high-impact journals according to Koffi [2021b]. Column (7) and (8) restrict the sample to the top five economics journals: *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics*, and *Review of Economic Studies*. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4 US-based publications and citations

	Outcome variable: Citation (asinh)			
	> 200 journals	Top 35 journals	Top 16 journals	Top 5 journals
	(1)	(2)	(3)	(4)
At least one Black/Hispanic/Asian	-0.056***	-0.083***	-0.061***	-0.098***
	(0.009)	(0.013)	(0.019)	(0.033)
Other races	-0.043***	-0.046***	-0.021	-0.002
	(0.010)	(0.016)	(0.022)	(0.037)
Observations	118111	45690	21886	7258
Adj. R-sqr	0.539	0.530	0.572	0.559
Affiliation fixed effects	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Here, we analyze the institutions of the authors, focusing on a sample where at least one author in the paper is based in the United States (US), and those papers are also cited by other US-based authors. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.5 US-based publications and citations: Homophily

	Outcome variable: Citation (asinh)							
	> 200 journals		Top 35 journals		Top 16 journals		Top 5 journals	
	From non-White	From White	From non-White	From White	From non-White	From White	From non-White	From White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
At least one Black/Hispanic/Asian	0.206***	-0.306***	0.179***	-0.338***	0.199***	-0.296***	0.182***	-0.287***
	(0.008)	(0.008)	(0.013)	(0.013)	(0.019)	(0.018)	(0.034)	(0.033)
Other races	0.006	-0.184***	0.019	-0.178***	0.057**	-0.119***	0.095**	-0.072**
	(0.008)	(0.008)	(0.015)	(0.015)	(0.023)	(0.022)	(0.038)	(0.036)
Observations	118111	118111	45690	45690	21886	21886	7258	7258
R-sqr	0.450	0.483	0.452	0.489	0.447	0.475	0.465	0.463

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, emphasizing the race of the authors of the citing paper. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Columns (1), (3), (5), and (7) restrict the citation count to citations from only teams, including non-White scholars. Columns (2), (4), (6), and (8) restrict the citation count to citations from only teams including all-White scholars. Here, we analyze the institutions of the authors, focusing on a sample where at least one author in the paper is based in the United States (US), and those papers are also cited by other US-based authors. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.6 Effect of seminars and conference attendance: Author level seminars and conferences

	Outcome variable: Citation (asinh)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
At least one Black/Hispanic/Asian	-0.087***	-0.086***	-0.086***	-0.085***	-0.091***	-0.075**	-0.072**
	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
Other races	-0.064	-0.050	-0.034	-0.038	-0.029	0.057	0.053
	(0.242)	(0.239)	(0.238)	(0.238)	(0.234)	(0.234)	(0.233)
Seminars/conferences count at time of publication		0.007***		0.003			0.002
		(0.002)		(0.003)			(0.003)
Prominent conferences at time of publication		0.045**		0.021			0.020
		(0.018)		(0.021)			(0.021)
Seminars in top-tier affiliations at time of publication		-0.012		-0.004			-0.008
		(0.010)		(0.011)			(0.011)
Seminars/conferences count 1 year before paper publication			0.010***	0.008*			0.008*
			(0.004)	(0.004)			(0.004)
Seminars in top-tier affiliations 1 year before paper publication			-0.002	-0.002			-0.003
			(0.012)	(0.012)			(0.012)
Prominent conferences 1 year before paper publication			0.010	0.005			0.006
			(0.021)	(0.022)			(0.022)
Seminars/conferences count 2 years before paper publication			0.001	0.001			0.002
			(0.003)	(0.003)			(0.003)
Seminars in top-tier affiliations 2 years before paper publication			-0.016	-0.016			-0.022*
			(0.012)	(0.012)			(0.012)
Prominent conferences 2 years before paper publication			0.017	0.011			0.003
			(0.022)	(0.023)			(0.023)
Seminars/conferences count 3 years before paper publication			-0.004	-0.004			-0.003
			(0.003)	(0.003)			(0.003)
Seminars in top-tier affiliations 3 years before paper publication			0.002	0.002			-0.004
			(0.011)	(0.011)			(0.011)
Prominent conferences 3 years before paper publication			0.060***	0.059**			0.051**
			(0.023)	(0.023)			(0.023)
Seminars/conferences count 4 years before paper publication			0.001	0.001			0.001
			(0.003)	(0.003)			(0.003)
Seminars in top-tier affiliations 4 years before paper publication			-0.025**	-0.024*			-0.022*
			(0.012)	(0.013)			(0.012)
Prominent conferences 4 years before paper publication			-0.008	-0.009			-0.005
			(0.023)	(0.023)			(0.023)
Number of editorial positions before paper publication					0.001**	0.001**	0.001**
					(0.001)	(0.001)	(0.001)
Any top 5 journals editorial positions before paper publication					0.170***	0.140***	0.137***
					(0.045)	(0.045)	(0.045)
Any top 35 journals editorial positions (not top 5) before paper publication					0.120***	0.092**	0.084*
					(0.043)	(0.043)	(0.043)
Total number of other visiting positions before paper publication					-0.010	-0.015	-0.023*
					(0.013)	(0.013)	(0.013)
Total number of visiting positions at top-tier instititions before paper pub.					0.065***	0.060***	0.058***
					(0.017)	(0.016)	(0.017)
Any authors from a top-tier PhD instititon					0.055	0.037	0.041
					(0.038)	(0.038)	(0.038)
Any authors member of NBER, JPAL, CEPR, BREAD, IZA, CIFAR						0.138***	0.132***
						(0.020)	(0.020)
Observations	4427	4427	4427	4427	4427	4427	4427
Adj. R-sqr	0.705	0.706	0.707	0.707	0.709	0.712	0.713
Publications and team size controls	Y	Y	Y	Y	Y	Y	Y
Affiliation fixed effects	Y	Y	Y	Y	Y	Y	Y
Field fixed effects	Y	Y	Y	Y	Y	Y	Y
Journal-Year fixed effects	Y	Y	Y	Y	Y	Y	Y

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, emphasizing the effect of seminars and conference attendance. The reference category for the race variable is *All-White*, which represents papers written by only White authors. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.7 Effect of seminars and conference attendance: Univariate

	Outcome variable: Citation (asinh)			
	(1)	(2)	(3)	(4)
At least one Black/Hispanic/Asian	-0.305***	-0.253***	-0.225***	-0.222***
	(0.056)	(0.054)	(0.053)	(0.053)
Other races	0.196	0.119	0.069	0.069
	(0.621)	(0.561)	(0.555)	(0.555)
Seminars/conferences count before paper publication		-0.059***		-0.019***
		(0.005)		(0.006)
Prominent conferences before paper publication		0.161***		0.096**
		(0.037)		(0.041)
Seminars in top-tier affiliations before paper publication		0.057***		0.029
		(0.022)		(0.022)
Seminars/conferences count 1 before paper publication			-0.042***	-0.033***
			(0.009)	(0.009)
Seminars in top-tier affiliations 1 before paper publication			0.027	0.017
			(0.024)	(0.024)
Prominent conferences 1 before paper publication			0.000	0.000
			(.)	(.)
Seminars/conferences count 2 before paper publication			-0.027***	-0.023**
			(0.009)	(0.009)
Seminars in top-tier affiliations 2 before paper publication			0.008	0.001
			(0.024)	(0.024)
Prominent conferences 2 before paper publication			0.000	0.000
			(.)	(.)
Seminars/conferences count 3 before paper publication			-0.010	-0.008
			(0.007)	(0.007)
Seminars in top-tier affiliations 3 before paper publication			-0.006	-0.009
			(0.024)	(0.024)
Prominent conferences 3 before paper publication			0.000	0.000
			(.)	(.)
Seminars/conferences count 4 before paper publication			-0.033***	-0.030***
			(0.012)	(0.011)
Seminars in top-tier affiliations 4 before paper publication			-0.002	-0.006
			(0.026)	(0.026)
Prominent conferences 4 before paper publication			0.000	0.000
			(.)	(.)
Observations	4427	4427	4427	4427
Adj. R-sqr	0.006	0.063	0.091	0.092

This table shows the relationship between the citation (inverse hyperbolic sine—asinh) and the racial composition of the paper's authors, emphasizing the effect of seminars and conference attendance.

C.8 Effect of JSTOR

	Outcome variable: Citation (asinh)								
	All citations			Citations from non-White			Citations from White		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
At least one Black/Hispanic/Asian	-0.024***	0.000	0.000	0.034***	0.000	0.000	-0.057***	0.000	0.000
	(0.001)	(.)	(.)	(0.001)	(.)	(.)	(0.001)	(.)	(.)
Other races	-0.010***	0.000	0.000	0.004***	0.000	0.000	-0.023***	0.000	0.000
	(0.001)	(.)	(.)	(0.001)	(.)	(.)	(0.001)	(.)	(.)
After JSTOR	-0.029***	0.005***	0.012***	0.011***	0.020***	0.022***	-0.036***	-0.012***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
At least one Black/Hispanic/Asian × After JSTOR	0.015***	0.007***	0.003	-0.010***	-0.029***	-0.031***	0.022***	0.030***	0.027***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Other races × After JSTOR	0.005***	0.003	-0.001	-0.005***	-0.008***	-0.009***	0.014***	0.020***	0.017***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Observations	3172893	3172893	3172893	3172893	3172893	3172893	3172893	3172893	3172893
Adj. R-sqr	0.417	0.621	0.633	0.275	0.513	0.517	0.305	0.514	0.526
Journal fixed effects	Y	Y	Y				Y	Y	Y
Year of publication by citation year fixed effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
Article fixed effects				Y	Y	Y	Y	Y	Y

This table shows the impact of JSTOR's introduction on the citation gap between non-White and White scholars. To estimate this effect, we examine citations for each article by year, starting from the year of its publication. The reference category for the race variable is *All-White*, which represents papers written by only White authors. The regressions without articles fixed effects include controls for the number of authors, the fields, and the affiliations. All the regressions include controls for the authors' prominence as measured by the number of publications at a given time in a given type of journals (top 5, top 16, top 35, all), the previous paper cumulative citations up to time t. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Additional Figures

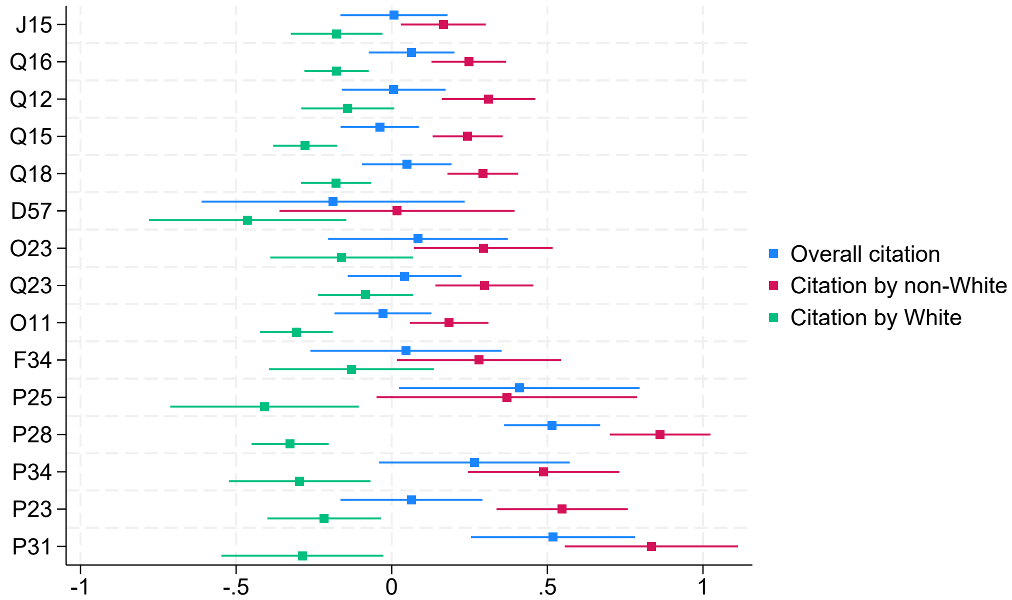
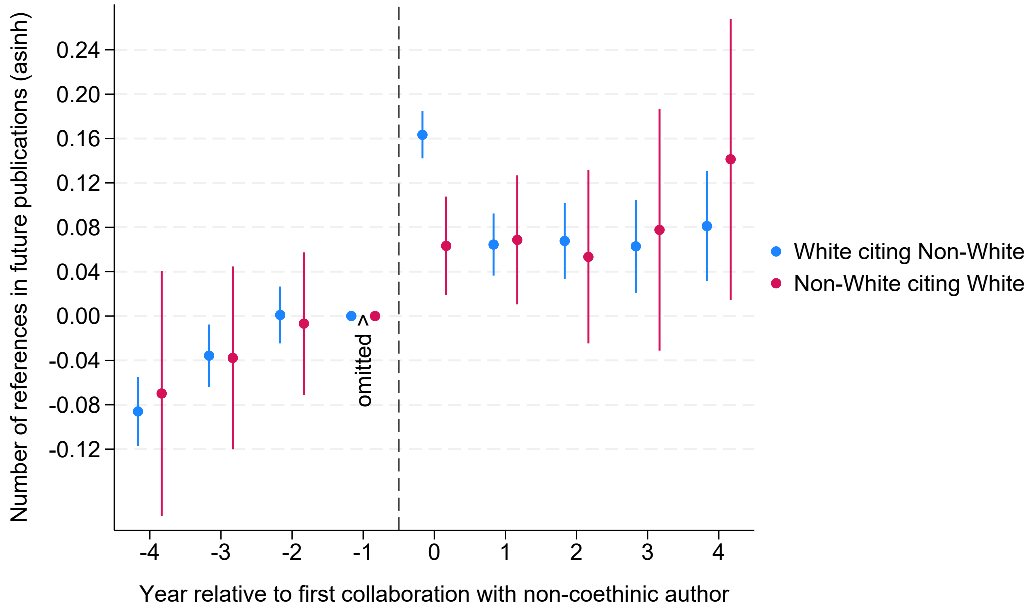


Figure IX:
Citation and race of the authors in the citing paper

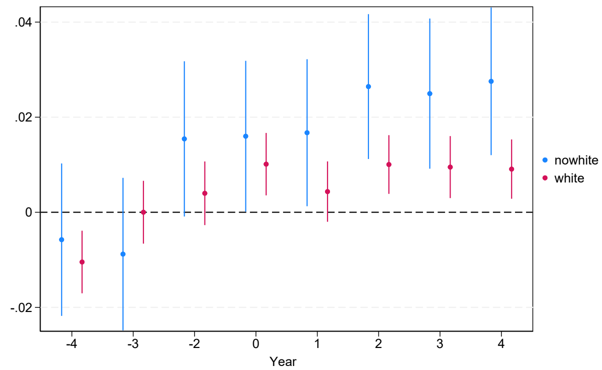
The figure presents the relationship between citation and race for the three-digit JEL code fields with higher fractions of non-White scholars, focusing on the racial composition of the authors of the citing paper. The reference category for the race variable is *All-White*, which represents papers written by only White authors. The plots display the estimate for teams with at least one Black/Hispanic/Asian author by comparison to a team with all-White authors.



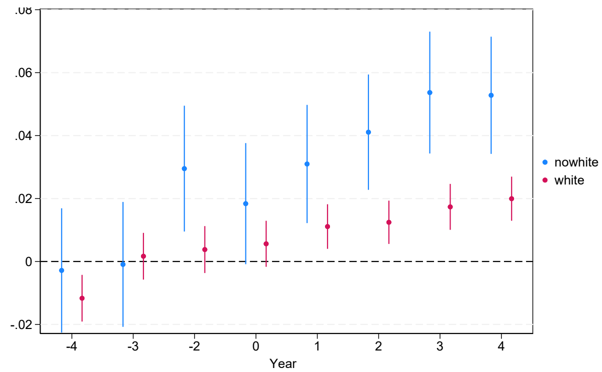
(a) Number of listed bibliographical references (Papers not in the same field)

Figure X:
Effect of collaboration with a non-coethnic author

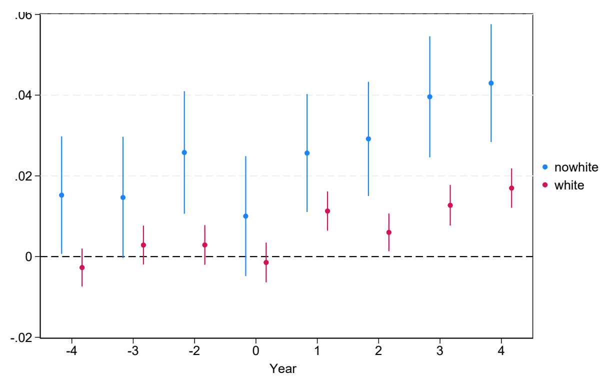
The figure presents the event study showing the impact of the collaboration with a non-coethnic author on the number and the share of listed bibliographical references to papers with authors of the out-group. To estimate this effect, we consider an author-level estimation and analyze all subsequent publications with only coethnic authors following the first collaboration with a non-coethnic author. We truncate the sample to the next non-coethnic collaboration.



(a) Citations by teams with all-White scholars



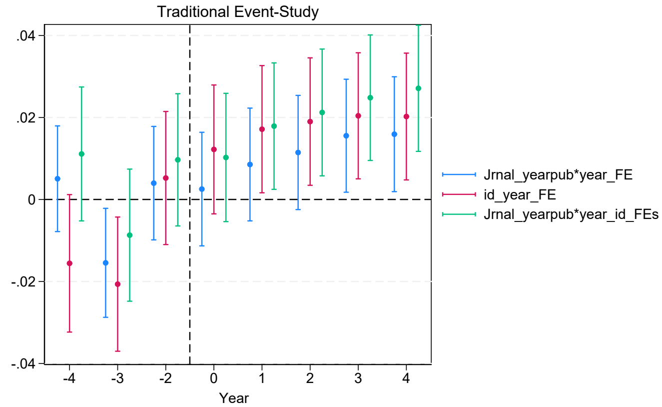
(b) All citations



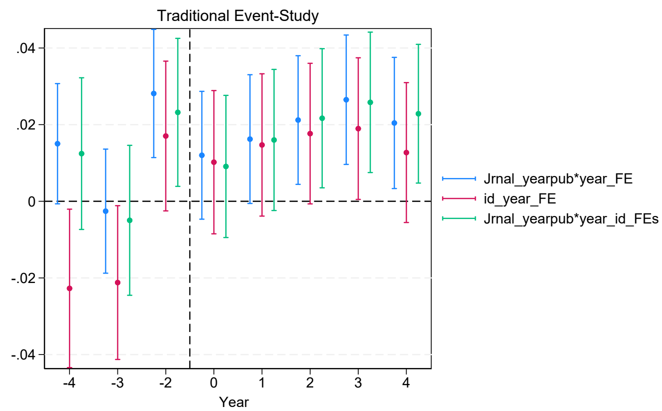
(c) Citations by teams with non-White scholars

Figure XI:
Effect of JSTOR's introduction adjusting for heterogenous treatment effects (Sun and Abraham [2021])

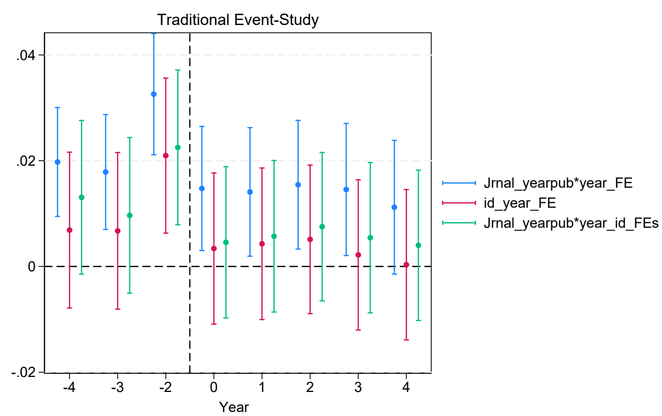
The figure presents the event study showing the impact of JSTOR's introduction on the citation gap between non-White and White scholars. To estimate this effect, we examine citations for each article by year, starting from the year of its publication. The plots display the estimate for teams with at least one Black/Hispanic/Asian author before the introduction of JSTOR and after the introduction of JSTOR (blue bars) and the estimate for teams with all-White scholars before the introduction of JSTOR and after the introduction of JSTOR, using one period before the introduction as a reference for time comparison.



(a) Citations by teams with all-White scholars



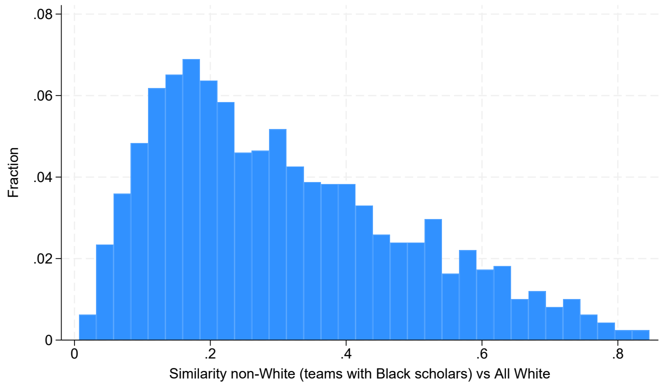
(b) All citations



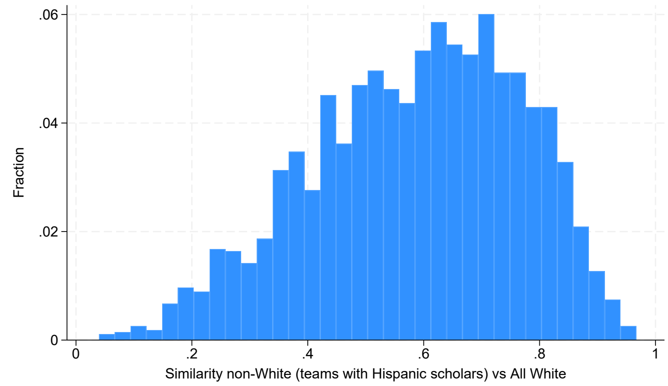
(c) Citations by teams including non-White scholars

Figure XII:
Effect of JSTOR's introduction (not adjusting for heterogenous treatment effects)

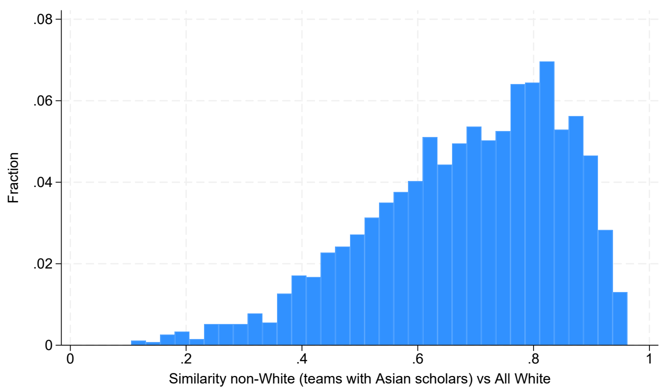
The figure presents the event study showing the impact of JSTOR's introduction on the citation gap between non-White and White scholars. To estimate this effect, we examine citations for each article by year, starting from the year of its publication. The reference category for the race variable is *All-White*, which represents papers written by only White authors. The plots display the estimate for teams with at least one Black/Hispanic/Asian author before the introduction of JSTOR and after the introduction of JSTOR compared to a team with all-White authors, using one period before the introduction as a reference for time comparison. The blue bars represent estimates that include journal fixed effects and publication year-by-citation year fixed effects. The red bars show estimates using article-fixed effects and citation-year-fixed effects. The green bars combine article-fixed effects, journal-fixed effects, and publication year-by-citation-year fixed effects.



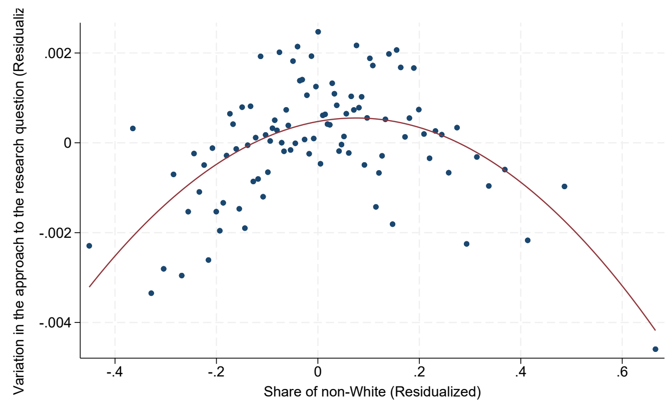
(a) Similarity distribution: Black versus White



(b) Similarity distribution: Hispanic versus White



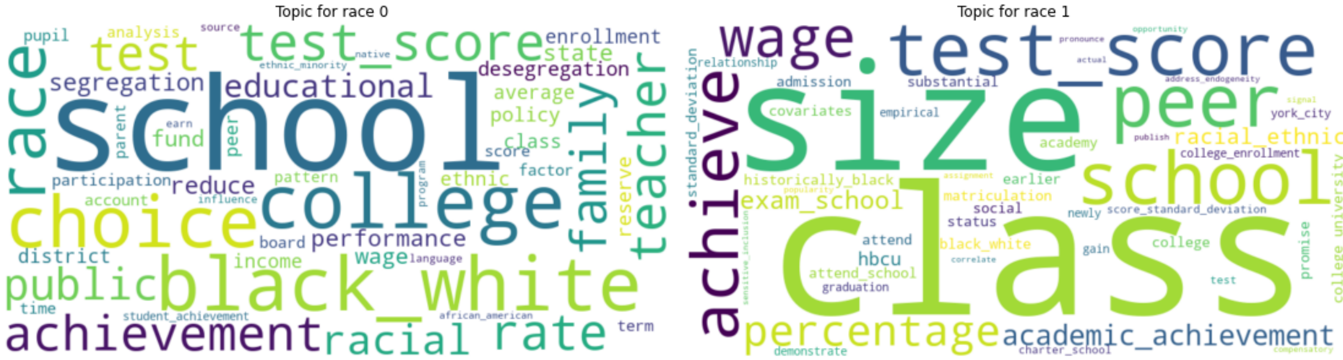
(c) Similarity distribution: Asian versus White



(d) Residualized plot

Figure XIII:
Racial categories and variation in the approach to the research question

The figure illustrates variations in research approaches across racial categories. Within each JEL code, papers are clustered into groups averaging 200 observations based on embedding representations. For each cluster, racial categories are treated as subgroups, and semantic differences are analyzed using the term frequency-inverse document frequency (TF-IDF) metric. All-White teams is used as the reference category. Figures (a), (b), and (c) show the distribution of this similarity metric between teams with Black, Hispanic, and Asian scholars, respectively, and all-White teams. Figure (d) examines the similarity distance between a paper and the cluster centroid within each cluster. It then computes the standard deviation of these similarity scores within a cluster and correlates this with the proportion of non-White scholars in the cluster. The similarity metric and the proportion of non-White scholars are calculated over time and then residualized using cluster and time-fixed effects to generate the plot.



White

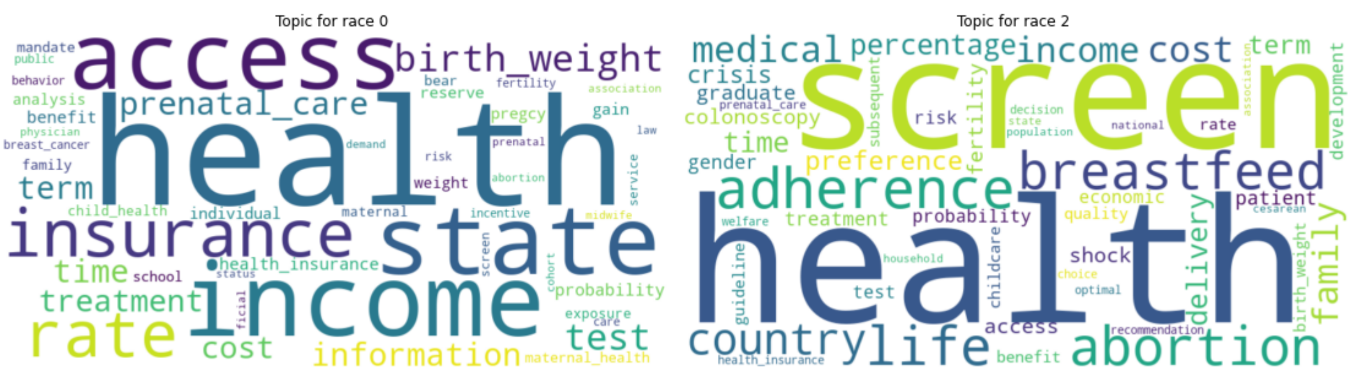
Non-White

Panel A: Topic: Economics of Minorities; Cluster: Child, student, education

White

Non-White

Panel B: Topic: Economics of Gender; Cluster: Women education



White

Non-White

Panel C: Topic: Economics of Gender; Cluster: Women health

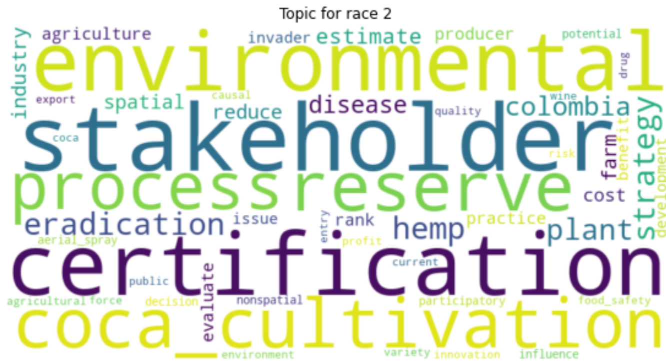
Figure XIV:
Examples: Racial categories and variation in the topic angle

The figure presents examples of the word clouds of each cluster by different racial categories.

White

Non-White

Panel D: Topic: Wage level and structure; Cluster: Worker earning



White

Non-White

Panel E: Topic: Agricultural Policy; Cluster: Country farmer market

White

Non-White

Panel F: Topic: International Trade; Cluster: China's export

Figure XV:
Examples: Racial categories and variation in the topic angle

The figure presents examples of the word clouds of each cluster by different racial categories.

E A simple model

E.1 Author identity, homophily, and the racial citation gap

In this section, we solve the author's maximization problem stated in section 5.4.1. The model has clear implications for the racial citation gap and shows how this outcome is driven by homophily in citation patterns. The model also sheds light on some of the heterogeneous effects of race on citation found in the empirical analysis.

As with producer theory's input cost, we assume that the function $g(s_1, s_2)$ is a linear function of the input factors s_1 and s_2: $g(s_1, s_2) = c_1 s_1 + c_2 s_2$. This implies that the total cost function $c(s_1, s_2) = h(q, r)g(s_1, s_2)$ is linear as well. Fixing $h(q, r)$, the parameter c_1 (resp. c_2) is therefore the marginal cost of exploiting and citing a paper authored by a scholar from the minority group (resp. majority group). While we solve the maximization problem without accounting for author identity, when analyzing how citation behavior differs by identity, we follow the notation introduced in section 5.4.1 and assume that $c_1 = c_1^1$ and $c_2 = c_2^1$ for an author from the **minority group**, and $c_1 = c_1^2$ and $c_2 = c_2^2$ for an author from the **majority group**, where $c_i^{i'}$ represents the marginal cost for authors in group i' to cite authors from group i, with $i, i' = 1, 2$.

The author's maximization problem stated in section 5.4.1 is equivalent to:

$$\max_{(s_1, s_2)} \quad \alpha \frac{n_1 a + n_2 b}{n} ln(s_1) + \alpha \frac{n_1 b + n_2 a}{n} ln(s_2) - h(q, r)(c_1 s_1 + c_2 s_2)$$

For simplicity, we assume that the numbers of cited papers from minority and majority authors, s_1 and s_2, are positive reals. Let $k_1 = n_1 a + n_2 b$ and $k_2 = n_1 b + n_2 a$. The first-order conditions (F.O.Cs) for interior solutions are:

$$\partial s_1 : \frac{\alpha k_1}{n s_1} = h(q, r)c_1 \tag{2}$$

$$\partial s_2 : \frac{\alpha k_2}{n s_2} = h(q, r)c_2 \tag{3}$$

From Equation 2 and Equation 3, we deduce that:

$$s_2^* = \frac{k_2 c_1}{k_1 c_2} s_1^* \tag{4}$$

It follows that

$$s_1^* = \frac{\alpha k_1}{h(q, r) n c_1} \tag{5}$$

$$s_2^* = \frac{\alpha k_2}{h(q, r) n c_2} \tag{6}$$

$$s_1^* - s_2^* = \left[\frac{k_1 c_2 - k_2 c_1}{k_1 c_2} \right] s_1^* = \frac{\alpha (k_1 c_2 - k_2 c_1)}{h(q, r) n c_1 c_2} \tag{7}$$

Recall that $c_1 = c_1^1$ and $c_2 = c_2^1$ for an author from the minority group; similarly, $c_1 = c_1^2$ and $c_2 = c_2^2$ for an author from the majority group. Then, for a **minority author**, Equation 7, which gives the difference in the number of citations to papers authored by minority and majority scholars, becomes:

$$s_1^{1*} - s_2^{1*} = \frac{\alpha(k_1 c_2^1 - k_2 c_1^1)}{h(q,r) n c_1^1 c_2^1} \tag{8}$$

For a **majority author**, Equation 7 becomes:

$$s_1^{2*} - s_2^{2*} = \frac{\alpha(k_1 c_2^2 - k_2 c_1^2)}{h(q,r) n c_1^2 c_2^2} \tag{9}$$

We maintain the assumption that the marginal cost of citing a paper written by a co-ethnic author is lower than that of citing a paper by a non-co-ethnic author: $c_i^i < c_i^{i'}$ for $i \neq i'$. At the same time, we assume that a minority author and a majority author are *symmetric* in terms of their costs, so that the predictions of the model are not driven by differential citation costs between minority and majority authors. Therefore, let $c_1^1 = c_2^2 = \kappa$ and $c_2^1 = c_1^2 = \gamma$. We know by assumption that κ is smaller than γ ($\kappa < \gamma$) because of the aforementioned assumption that the cost of citing a paper written by a co-ethnic author is smaller than the cost of citing a paper written by an author from a different ethnic group. We next show that homophilic behavior in citation patterns differs according to the ethnic identity of an author.

The racial citation gap in a paper written by a **minority author** is:

$$s_1^{1*} - s_2^{1*} = \frac{\alpha(k_1 \gamma - k_2 \kappa)}{h(q,r) n \kappa \gamma} \tag{10}$$

This quantity could be positive (**homophily**), meaning that the number of citations to papers authored by minority scholars exceeds the number of citations to papers authored by scholars from the majority group, or negative (**heterophily**), depending on the relative size of the minority group.

For a **majority author**, the racial citation gap is:

$$s_1^{2*} - s_2^{2*} = \frac{\alpha(k_1 \kappa - k_2 \gamma)}{h(q,r) n \kappa \gamma} \tag{11}$$

This quantity is clearly negative (**homophily**) because it can be shown that the numerator is negative.

We next show that, all else equal, authors from the majority group display a greater level of homophily in citation than authors from the minority group. This means that the minority group cites members of their own group ($s_1^{1*} - s_2^{1*}$) less frequently than the majority group cites its own members ($s_2^{2*} - s_1^{2*}$). We do this by comparing the expressions in Equations 10 and 11. If we assume that ($s_1^{1*} - s_2^{1*}$) is negative, which would mean that minority authors are heterophilous, then the conclusion follows immediately. Assume instead that $s_1^{1*} - s_2^{1*}$ is positive. Then, it suffices to show that $(s_2^{2*} - s_1^{2*}) - (s_1^{1*} - s_2^{1*}) \geq 0$. One can easily show that:

$$(s_2^{2*} - s_1^{2*}) - (s_1^{1*} - s_2^{1*}) = \frac{\alpha(k_2 - k_1)(\kappa + \gamma)}{h(q,r) n \kappa \gamma} \tag{12}$$

It is obvious that the right-hand side of Equation 18 is positive given that $k_2 - k_1 = (a - b)(n_2 - n_1) \geq 0$ as $a \geq b$ and $n_2 \geq n_1$. The prediction that homophily is larger among majority authors compared to minority authors is consistent with our empirical findings.

Are papers authored by minority authors less frequently cited than papers of similar quality authored by majority authors? If one assumes that the pool of referees is identical to the pool of authors, which is a reasonable assumption in a peer-review system, then the **average** difference in the number of citations to papers authored by minority scholars and papers authored by majority scholars is:

$$E(s_1^* - s_2^*) = \frac{n_1}{n}[\frac{\alpha(k_1\gamma - k_2\kappa)}{h(q,r)n\kappa\gamma}] + \frac{n_2}{n}[\frac{\alpha(k_1\kappa - k_2\gamma)}{h(q,r)n\kappa\gamma}] = \frac{\alpha(n_1^2 - n_2^2)(a\gamma - b\kappa)}{h(q,r)n^2\kappa\gamma} \tag{13}$$

This quantity is clearly negative since $n_1^2 - n_2^2 \leq 0$ (because $n_2 \geq n_1$) and $a\gamma - b\kappa > 0$ (because $a \geq b$ and $\gamma > \kappa$). This means that works authored by minority scholars receive fewer citations, on average, than works authored by majority scholars. This prediction is consistent with our empirical findings.

E.1.1 Bounding the effects of referee preference/bias

In this section, we bound the effect of referee preference for diversity, analyze how referee bias affects author homophily in citation behavior, and determine whether strategic citation (Angrist et al. [2020]) could arise in response to such bias. In the absence of referee bias, citations to works authored by co-ethnic scholars are given the same weight as citations to works authored by non-co-ethnic scholars; that is, $a = b$ in the objective function. However, in the presence of extreme bias, a referee only finds works authored by co-ethnic scholars to be informative; in this case, $a = 1$ and $b = 0$.

In the **absence of referee bias** ($a = b$), we have $k_1 = k_2$, and it follows from Equations 10 and 11 that:

$$s_1^{1*} - s_2^{1*} = -(s_1^{2*} - s_2^{2*}) = \frac{\alpha(\gamma - \kappa)}{2h(q,r)\kappa\gamma} \tag{14}$$

This quantity implies that when referee bias is minimal, homophilic behavior in citation does not vary by author identity; minority and majority authors display the same level of homophily.

The **average** racial citation gap in this case is:

$$E(s_1^* - s_2^*) = \frac{n_1}{n}[\frac{\alpha(\gamma - \kappa)}{2h(q,r)\kappa\gamma}] + \frac{n_2}{n}[\frac{\alpha(\kappa - \gamma)}{2h(q,r)\kappa\gamma}] = \frac{n_1 - n_2}{n}[\frac{\alpha(\gamma - \kappa)}{2h(q,r)\kappa\gamma}] \tag{15}$$

This is negative, suggesting that works authored by minority scholars receive fewer citations than works authored by majority scholars. We can see from Equation 15 that this is driven by the differential size of the minority and majority groups ($n_1 - n_2$) and the differential cost of citing co-ethnics relative to non-co-ethnics ($\gamma - \kappa$). In fact, as the difference in the size of the two groups decreases (i.e., $n_1 - n_2$ tends to zero), the racial citation gap tends to zero as well. Similarly as the differential citation cost decreases, the citation gap decreases as well.

In the **presence of extreme referee bias** ($a = 1$ and $b = 0$), we can show from Equation 10 that the racial citation gap in a paper authored by a **minority author** is:

$$s_1^{1*} - s_2^{1*} = \frac{\alpha(n_1\gamma - n_2\kappa)}{h(q,r)n\kappa\gamma} \tag{16}$$

As for the general case, this quantity can be positive (**homophily**) or negative (**heterophily**) depending on the relative size of the minority group. In particular, if the size of the minority group is sufficiently small relative to the majority group, minority authors will tend to cite works authored by scholars from the majority group more than the works of co-ethnic authors.

From Equation 11, the racial citation gap in a paper of a **majority author** is:

$$s_1^{2*} - s_2^{2*} = \frac{\alpha(n_1\kappa - n_2\gamma)}{h(q,r)n\kappa\gamma} \tag{17}$$

As for the general case, this is always negative, meaning that scholars from the majority group tend to cite works authored by co-ethnic scholars more (**homophily**). Comparing Equations 16 and 17 clearly shows that the intensity of homophily is larger for majority authors than for minority authors. In fact, the difference in homophily between a majority author and a minority author is:

$$(s_2^{2*} - s_1^{2*}) - (s_1^{1*} - s_2^{1*}) = \frac{\alpha(n_2 - n_1)(\kappa + \gamma)}{h(q, r)n\kappa\gamma} \tag{18}$$

This quantity is clearly positive, which is consistent with our empirical results.

The **average** racial citation gap is:

$$E(s_1^* - s_2^*) = \frac{n_1}{n}[\frac{\alpha(n_1\gamma - n_2\kappa)}{h(q, r)n\kappa\gamma}] + \frac{n_2}{n}[\frac{\alpha(n_1\kappa - n_2\gamma)}{h(q, r)n\kappa\gamma}] = \frac{\alpha[(n_1^2 - n_2^2)]}{h(q, r)n^2\kappa\gamma} \tag{19}$$

This is negative because the numerator is negative, again showing that papers written by minority scholars receive fewer citations on average. The racial citation gap is driven both the referee bias (or preference for diversity) and the differential size of the minority and the majority groups. Comparing Equations 15 and 19 therefore yields the relative size of the racial citation bias that is due to extreme referee bias:

$$\frac{E(s_1^* - s_2^*)_{a=1,b=0}}{E(s_1^* - s_2^*)_{a=b}} = \frac{2\gamma}{\gamma - \kappa} \tag{20}$$

Given that this quantity is larger than 1, it shows that the combination of differential group size and referee bias leads to a larger racial citation gap than just the differential group size.

We have seen that in the absence of referee bias, both minority and majority authors are homophilic to the same extent. However, when referee bias is present, majority authors tend to be more homophilic than minority authors. This implies that part of the racial citation gap in papers authored by minority scholars can be attributable to **strategic citation** (Angrist et al. [2020]). How big is the size of the racial citation gap due to strategic citation among minority authors? One can calculate it by taking the difference of the absolute value of the racial citation gap between majority and minority authors.

In the presence of **homophily** among minority authors, the racial citation gap attributable to strategic citation is equal to:

$$(s_1^{1*} - s_2^{1*})_{strategic} = |s_1^{2*} - s_2^{2*}| - |s_1^{1*} - s_2^{1*}| = \frac{\alpha[(n_2 - n_1)(\kappa + \gamma)]}{h(q, r)n\kappa\gamma} \tag{21}$$

In the presence of **heterophily** among minority authors, the racial citation gap attributable to strategic citation is equal to:

$$(s_1^{1*} - s_2^{1*})_{strategic} = |s_1^{2*} - s_2^{2*}| - |s_1^{1*} - s_2^{1*}| = \frac{\alpha[n(\kappa + \gamma)]}{h(q, r)n\kappa\gamma} \tag{22}$$

Remark that this quantity also measures the extent to which majority authors are more homophilic than minority authors. It is positive regardless of whether minority authors are homophilic or heterophilic, with the size of the citation gap due to strategic citation being clearly larger when heterophily prevails. Our empirical findings corroborate these predictions, as we have found that homophilic citation is higher for papers written by an all-White team than papers with at least one minority author.

E.1.2 The effects of citation costs, group size, and referee preferences

Does the differential cost of citing co-ethnics relative to non-co-ethnics alone drive the citation gap? The model suggests that the answer is no. In fact, from Equation 13, one can see that, as γ tends to κ (that is, the difference between the cost of citing co-ethnic authors versus non-co-ethnic authors is minimal), the racial citation gap $E(s_1^* - s_2^*)$ tends to $\frac{\alpha(n_1^2 - n_2^2)(a-b)}{h(q,r)n^2\kappa}$, which is negative. Similarly, referee preference for diversity alone does not drive the citation gap. In fact, when the referee is neutral ($a = b$), we have seen that the racial citation gap is negative (Equation 15). The differential group size ($n_1 - n_2$) seems to play a significant role in explaining the gap. However, its effect only appears when referee bias is present. It follows from the analysis that relative citation costs, differential group size, and referee preferences for diversity play complementary roles in determining the racial citation gap.

While the differential cost of citing co-ethnics relative to non-co-ethnics solely does not explain the racial citation gap, our empirical findings suggest that it partly explains homophily in citation behavior. In fact, if the differential citation cost is sufficiently small (γ tends to κ), then Equation 10 implies that minority authors cite works authored by majority authors more frequently than works from co-ethnics. This is not consistent with our empirical analysis. Similarly, our empirical findings suggest that both referee preference and the differential size of racial groups are important at explaining homophily in citation patterns. In fact, Equation 18, which shows the difference in homophily between authors from the majority and minority groups, implies that if referees are neutral ($a = b$) and/or the size difference between the groups is very small, then minority and majority authors will display the same level of homophily, which is not what we find empirically.

E.1.3 Other predictions of the model

In this section, we state other predictions of the model that are supported by our empirical analysis.

- From Equations 13, 15, and 19, the racial citation gap ($E(s_1^* - s_2^*)$) becomes more pronounced as paper quality (α) increases; this implies that the citation penalty experienced by minority scholars is larger for their high quality papers than for their low quality papers.[16]

- From Equations 13, 15, and 19, the racial citation gap ($E(s_1^* - s_2^*)$), becomes more pronounced for papers written by higher ability authors (r); this implies that the citation penalty is experienced more strongly by minority scholars who are more prominent.

- From Equations 13, 15, and 19, citing papers published in higher quality journals (q) discriminate less against the works of minority scholars; in other words, more prominent journals discriminate less against minority scholars.[17]

- From Equations 13, 15, and 19, the racial citation gap ($E(s_1^* - s_2^*)$) becomes more pronounced as the relative size of the minority group decreases. This prediction has an implication for how the racial citation gap varies across different fields (e.g., **theory vs. empirics**); it implies that in fields where minority scholars are significantly underrepresented, the racial citation gap is larger. The reduction of the gap over the past three decades that we find in the data

[16] If one assumes that higher quality papers are published in higher ranked journals, this prediction implies that the works of minority scholars published in more prominent journals suffer a higher citation penalty. Our empirical findings corroborate this prediction.

[17] This is different from saying that the racial gap in citations to papers published in higher quality journals is smaller.

is attributable to a greater representation of minority scholars in the empirical field, which corroborates the model's prediction.

E.2 A modification of the model

In this section, we show a modification of the model where author ability (r) instead enhances the quality of the paper, similar to the well-known total factor productivity in the Cobb-Douglas production function. Assume that a paper cites (s_1, s_2) papers of average quality α. Assume that the perceived value of this paper is $\alpha(aln(s_1) + bln(s_2)) + ln(r)$ to a scholar from the minority group and $\alpha(bln(s_1) + aln(s_2) + ln(r))$ to a scholar from the majority group, where a and b are positive, $a + b = 1$, and $a \geq b$. The average perceived value of the paper to a random referee is therefore:

$$\frac{n_1}{n_1 + n_2}\alpha[aln(s_1) + bln(s_2)] + \frac{n_2}{n_1 + n_2}\alpha[bln(s_1) + aln(s_2)] + ln(r)$$

Taking the exponential transform of this expression and letting $n = n_1 + n_2$, the overall average perceived quality of the paper is a function of the author's ability (r) and a combination of the paper's references, based on the notion that scholarly work is often influenced by existing literature:

$$v(s_1, s_2) = r \cdot s_1^{\alpha\frac{n_1 a + n_2 b}{n}} \cdot s_2^{\alpha\frac{n_1 b + n_2 a}{n}}$$

This is the traditional Cobb-Douglas production function. Assume that the total cost of producing the paper is defined as in the previous model, with the exception of the omission of the parameter capturing the author's quality:

$$c(s_1, s_2) = l(q)(c_1 s_1 + c_2 s_2)$$

where $l(q)$ is positive and increasing in q. The function $l(q)$ measures how hard it is to publish in a journal of quality q.

The author therefore solves the following maximization problem:

$$\max_{(s_1, s_2)} \quad u(s_1, s_2) = r s_1^{\frac{\alpha(n_1 a + n_2 b)}{n}} s_2^{\frac{\alpha(n_1 b + n_2 a)}{n}} - l(q)(c_1 s_1 + c_2 s_2)$$

As previously, let $k_1 = n_1 a + n_2 b$ and $k_2 = n_1 b + n_2 a$. It is easy to remark that this maximization problem is equivalent to:

$$\max_{(s_1, s_2)} \quad u(s_1, s_2) = r s_1^{\frac{\alpha k_1}{n}} s_2^{\frac{\alpha k_2}{n}} - l(q)(c_1 s_1 + c_2 s_2)$$

The F.O.Cs imply the following solution:

$$s_1^* = \left(\frac{\alpha k_1 r}{nl(q)c_1}\right)^{\frac{1}{1-\alpha}} \left(\frac{k_2 c_1}{k_1 c_2}\right)^{\frac{\alpha k_2}{(1-\alpha)k_1}} \tag{23}$$

$$s_2^* = \left(\frac{\alpha k_2 r}{nl(q)c_2}\right)^{\frac{1}{1-\alpha}} \left(\frac{k_1 c_2}{k_2 c_1}\right)^{\frac{\alpha k_1}{(1-\alpha)k_2}} \tag{24}$$

$$s_1^* - s_2^* = \left[\frac{(k_1 c_2 - k_2 c_1)}{k_1 c_2}\right] s_1^* \tag{25}$$

Assuming, as previously that a minority author and a majority author are symmetric in terms of their costs, we have $c_1^1 = c_2^2 = \kappa$ and $c_2^1 = c_1^2 = \gamma$. We note that κ is smaller than γ.

We find that author identity determines homophilic behavior, and that the findings are similar to those of the previous model. For a **minority author**, the racial citation gap is:

$$s_1^{1*} - s_2^{1*} = \frac{(k_1\gamma - k_2\kappa)}{k_1\gamma} s_1^* \tag{26}$$

This quantity could be positive (**homophily**) or negative (**heterophily**) depending on the relative size of the minority group.

For a **majority author**, the racial citation gap is:

$$s_1^* - s_2^* = \frac{(k_1\kappa - k_2\gamma)}{k_1\kappa} s_1^* \tag{27}$$

This is always negative (**homophily**).

The **average** racial citation gap is therefore:

$$E(s_1^* - s_2^*) = \frac{n_1}{n}[\frac{(k_1\gamma - k_2\kappa)}{k_1\gamma} s_1^*] + \frac{n_2}{n}[\frac{(k_1\kappa - k_2\gamma)}{k_1\kappa} s_1^*] \tag{28}$$

This quantity is negative, which means that works authored by minority scholars receive fewer citations that similar quality works authored by scholars from the majority group. This prediction is corroborated by our empirical findings.

Equations (21) and (22) also imply that the extent of homophily is larger for authors that belong to the majority group than for authors that belong to the minority group. This prediction is also consistent with our empirical findings. Overall, the analysis shows that the main insights of the theoretical analysis are not too sensitive to functional forms.

F List of Journals

ACCOUNTING ORGANIZATIONS AND SOCIETY	1,194
ACTA STRUCTILIA	1
AGRICULTURAL ECONOMICS	1,581
AMERICAN ECONOMIC JOURNAL-APPLIED ECONOMICS	421
AMERICAN ECONOMIC JOURNAL-ECONOMIC POLICY	407
AMERICAN ECONOMIC JOURNAL-MACROECONOMICS	336
AMERICAN ECONOMIC JOURNAL-MICROECONOMICS	393
AMERICAN ECONOMIC REVIEW	5,273
AMERICAN JOURNAL OF AGRICULTURAL ECONOMICS	4,920
AMERICAN LAW AND ECONOMICS REVIEW	194
ANNALS OF ECONOMICS AND FINANCE	303
APPLIED ECONOMIC PERSPECTIVES AND POLICY	299
APPLIED ECONOMICS	8,354
APPLIED ECONOMICS LETTERS	5,857
ASIAN JOURNAL OF LAW AND ECONOMICS	123
American Economic Journal-Economic Policy	2
B E JOURNAL OF ECONOMIC ANALYSIS & POLICY	915
B E JOURNAL OF MACROECONOMICS	469
BALTIC JOURNAL OF ECONOMIC STUDIES	850
BELL JOURNAL OF ECONOMICS AND MANAGEMENT SCIENCE	46
BROOKINGS PAPERS ON ECONOMIC ACTIVITY	438
CAMBRIDGE JOURNAL OF ECONOMICS	1,729
CANADIAN JOURNAL OF ECONOMICS-REVUE CANADIENNE D ECONOMIQUE	1,777
CESIFO ECONOMIC STUDIES	355
CHINA ECONOMIC REVIEW	1,280
CRITICAL FINANCE REVIEW	32
ECOLOGICAL ECONOMICS	5,046
ECONOMETRIC REVIEWS	581
ECONOMETRIC THEORY	1,355
ECONOMETRICA	3,605
ECONOMETRICS JOURNAL	339
ECONOMIC AND POLITICAL STUDIES-EPS	162
ECONOMIC DEVELOPMENT AND CULTURAL CHANGE	1,474
ECONOMIC INQUIRY	2,539
ECONOMIC JOURNAL	3,450
ECONOMIC POLICY	322
ECONOMIC RECORD	1,508
ECONOMIC SYSTEMS	422
ECONOMIC SYSTEMS RESEARCH	326
ECONOMIC THEORY	2,255
ECONOMICA	1,973
ECONOMICS & POLITICS	188
ECONOMICS LETTERS	11,175
ECONOMICS OF EDUCATION REVIEW	1,641
ECONOMICS OF GOVERNANCE	171
ECONOMICS OF INNOVATION AND NEW TECHNOLOGY	652
ECONOMICS OF TRANSITION	504
ECONOMIST	637
EMERGING MARKETS REVIEW	477
EMPIRICAL ECONOMICS	1,820
ENERGY ECONOMICS	4,329
ENERGY JOURNAL	1,089
ENERGY POLICY	10,968
ENVIRONMENT AND DEVELOPMENT ECONOMICS	659
ENVIRONMENTAL & RESOURCE ECONOMICS	1,933

EUROPEAN ECONOMIC REVIEW	3,411
EUROPEAN FINANCIAL MANAGEMENT	531
EUROPEAN JOURNAL OF POLITICAL ECONOMY	833
EXPERIMENTAL ECONOMICS	583
EXPLORATIONS IN ECONOMIC HISTORY	916
FEDERAL RESERVE BANK OF ST LOUIS REVIEW	381
FINANCE A UVER	166
FINANCE A UVER-CZECH JOURNAL OF ECONOMICS AND FINANCE	482
FINANCE AND STOCHASTICS	542
FINANCIAL MANAGEMENT	1,476
FISCAL STUDIES	416
FOUNDATIONS AND TRENDS IN ENTREPRENEURSHIP	78
GAMES AND ECONOMIC BEHAVIOR	2,566
HEALTH ECONOMICS	2,274
IMF ECONOMIC REVIEW	225
IMF STAFF PAPERS	298
INDUSTRIAL & LABOR RELATIONS REVIEW	1,445
INDUSTRIAL AND CORPORATE CHANGE	716
INTERNATIONAL BUSINESS REVIEW	1,105
INTERNATIONAL ECONOMIC REVIEW	2,304
INTERNATIONAL FINANCE	157
INTERNATIONAL JOURNAL OF CENTRAL BANKING	431
INTERNATIONAL JOURNAL OF FINANCE & ECONOMICS	821
INTERNATIONAL JOURNAL OF FORECASTING	1,807
INTERNATIONAL JOURNAL OF GAME THEORY	1,144
INTERNATIONAL JOURNAL OF INDUSTRIAL ORGANIZATION	1,717
INTERNATIONAL LABOUR REVIEW	1,779
INTERNATIONAL REVIEW OF ENVIRONMENTAL AND RESOURCE ECONOMICS	111
INTERNATIONAL TAX AND PUBLIC FINANCE	753
IZA JOURNAL OF LABOR POLICY	157
JOURNAL OF ACCOUNTING & ECONOMICS	973
JOURNAL OF ACCOUNTING AND PUBLIC POLICY	565
JOURNAL OF ACCOUNTING RESEARCH	1,593
JOURNAL OF AFRICAN ECONOMIES	623
JOURNAL OF AGRICULTURAL ECONOMICS	1,359
JOURNAL OF APPLIED ECONOMETRICS	1,330
JOURNAL OF APPLIED ECONOMICS	303
JOURNAL OF BANKING & FINANCE	5,246
JOURNAL OF BUSINESS	1,465
JOURNAL OF BUSINESS & ECONOMIC STATISTICS	1,597
JOURNAL OF BUSINESS VENTURING	1,111
JOURNAL OF COMPARATIVE ECONOMICS	1,205
JOURNAL OF CONSUMER RESEARCH	2,199
JOURNAL OF CORPORATE FINANCE	1,566
JOURNAL OF DEVELOPMENT ECONOMICS	2,555
JOURNAL OF DEVELOPMENT STUDIES	2,614
JOURNAL OF ECONOMETRICS	3,623
JOURNAL OF ECONOMIC BEHAVIOR & ORGANIZATION	4,644
JOURNAL OF ECONOMIC DYNAMICS & CONTROL	3,474
JOURNAL OF ECONOMIC GEOGRAPHY	648
JOURNAL OF ECONOMIC GROWTH	260
JOURNAL OF ECONOMIC HISTORY	1,853
JOURNAL OF ECONOMIC INEQUALITY	312
JOURNAL OF ECONOMIC ISSUES	1,910
JOURNAL OF ECONOMIC LITERATURE	743
JOURNAL OF ECONOMIC PERSPECTIVES	1,262
JOURNAL OF ECONOMIC PSYCHOLOGY	1,783
JOURNAL OF ECONOMIC SURVEYS	563

JOURNAL OF ECONOMIC THEORY	3,882
JOURNAL OF ECONOMICS & MANAGEMENT STRATEGY	928
JOURNAL OF EMPIRICAL FINANCE	827
JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT	1,976
JOURNAL OF EVOLUTIONARY ECONOMICS	713
JOURNAL OF FINANCE	4,001
JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS	2,436
JOURNAL OF FINANCIAL ECONOMETRICS	265
JOURNAL OF FINANCIAL ECONOMICS	2,481
JOURNAL OF FINANCIAL INTERMEDIATION	506
JOURNAL OF FINANCIAL MARKETS	471
JOURNAL OF FINANCIAL SERVICES RESEARCH	474
JOURNAL OF FINANCIAL STABILITY	610
JOURNAL OF GLOBAL ECONOMIC ANALYSIS	45
JOURNAL OF HEALTH ECONOMICS	1,795
JOURNAL OF HOUSING ECONOMICS	560
JOURNAL OF HUMAN CAPITAL	173
JOURNAL OF HUMAN RESOURCES	1,307
JOURNAL OF INDUSTRIAL ECONOMICS	1,170
JOURNAL OF INTERNATIONAL BUSINESS STUDIES	1,699
JOURNAL OF INTERNATIONAL ECONOMICS	2,402
JOURNAL OF INTERNATIONAL FINANCIAL MARKETS INSTITUTIONS & MONEY	807
JOURNAL OF INTERNATIONAL MANAGEMENT	395
JOURNAL OF INTERNATIONAL MONEY AND FINANCE	2,332
JOURNAL OF LABOR ECONOMICS	1,028
JOURNAL OF LAW & ECONOMICS	1,138
JOURNAL OF LAW ECONOMICS & ORGANIZATION	751
JOURNAL OF LEGAL STUDIES	857
JOURNAL OF MACROECONOMICS	2,032
JOURNAL OF MANAGEMENT STUDIES	1,664
JOURNAL OF MATHEMATICAL ECONOMICS	1,828
JOURNAL OF MONETARY ECONOMICS	1,955
JOURNAL OF MONEY CREDIT AND BANKING	2,143
JOURNAL OF POLITICAL ECONOMY	3,044
JOURNAL OF POPULATION ECONOMICS	1,051
JOURNAL OF PRODUCTIVITY ANALYSIS	741
JOURNAL OF PUBLIC ECONOMICS	3,549
JOURNAL OF REAL ESTATE FINANCE AND ECONOMICS	1,017
JOURNAL OF REGIONAL SCIENCE	1,444
JOURNAL OF REGULATORY ECONOMICS	754
JOURNAL OF RETAILING	1,670
JOURNAL OF RISK AND INSURANCE	1,810
JOURNAL OF RISK AND UNCERTAINTY	740
JOURNAL OF TAXATION	2,367
JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION	7,941
JOURNAL OF THE ASSOCIATION OF ENVIRONMENTAL AND RESOURCE ECONOMISTS	232
JOURNAL OF THE EUROPEAN ECONOMIC ASSOCIATION	887
JOURNAL OF THE JAPANESE AND INTERNATIONAL ECONOMIES	619
JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES B-STATISTICAL METHODOLOGY	2,030
JOURNAL OF TRANSPORT ECONOMICS AND POLICY	1,056
JOURNAL OF URBAN ECONOMICS	1,557
JOURNAL OF WORLD BUSINESS	965
KYKLOS	1,336
LABOUR ECONOMICS	1,240
M&SOM-MANUFACTURING & SERVICE OPERATIONS MANAGEMENT	662

MACROECONOMIC DYNAMICS	1,161
MACROECONOMICS ANNUAL 2002	5
MANAGEMENT SCIENCE	6,254
MATHEMATICAL FINANCE	681
NATIONAL TAX JOURNAL	2,035
NBER MACROECONOMICS ANNUAL	86
NBER MACROECONOMICS ANNUAL 1997	10
NBER MACROECONOMICS ANNUAL 1998	5
NBER MACROECONOMICS ANNUAL 1999	6
NBER MACROECONOMICS ANNUAL 2000	5
NBER MACROECONOMICS ANNUAL 2001	6
NBER MACROECONOMICS ANNUAL 2003	8
NEW ENGLAND REVIEW-MIDDLEBURY SERIES	548
ORGANIZATION SCIENCE	1,597
OXFORD BULLETIN OF ECONOMICS AND STATISTICS	1,543
OXFORD ECONOMIC PAPERS-NEW SERIES	2,167
OXFORD REVIEW OF ECONOMIC POLICY	981
POLITICAL ANALYSIS	469
PORTUGUESE ECONOMIC JOURNAL	158
PUBLIC CHOICE	3,528
QME-QUANTITATIVE MARKETING AND ECONOMICS	245
QUANTITATIVE ECONOMICS	284
QUARTERLY JOURNAL OF ECONOMICS	2,391
RAND JOURNAL OF ECONOMICS	1,339
REAL ESTATE ECONOMICS	690
REGIONAL SCIENCE AND URBAN ECONOMICS	1,859
REGIONAL STUDIES	3,166
RESEARCH POLICY	3,082
RESOURCE AND ENERGY ECONOMICS	793
REVIEW OF ECONOMIC DYNAMICS	883
REVIEW OF ECONOMIC STUDIES	2,547
REVIEW OF ECONOMICS AND STATISTICS	4,499
REVIEW OF ENVIRONMENTAL ECONOMICS AND POLICY	183
REVIEW OF FINANCE	538
REVIEW OF FINANCIAL ECONOMICS	369
REVIEW OF FINANCIAL STUDIES	1,825
REVIEW OF INCOME AND WEALTH	964
REVIEW OF INTERNATIONAL ECONOMICS	829
REVIEW OF NETWORK ECONOMICS	146
REVIEW OF WORLD ECONOMICS	563
SCANDINAVIAN JOURNAL OF ECONOMICS	1,150
SMALL BUSINESS ECONOMICS	1,809
SOCIAL CHOICE AND WELFARE	1,674
SOUTHERN ECONOMIC JOURNAL	2,888
STATA JOURNAL	809
STRATEGIC MANAGEMENT JOURNAL	2,549
THEORETICAL ECONOMICS	266
THEORY AND DECISION	1,639
TRANSPORTATION	1,522
TRANSPORTATION RESEARCH PART A-POLICY AND PRACTICE	1
WATER RESOURCES RESEARCH	1
WELTWIRTSCHAFTLICHES ARCHIV-REVIEW OF WORLD ECONOMICS	1,478
WORLD BANK ECONOMIC REVIEW	760
WORLD BANK RESEARCH OBSERVER	262
WORLD DEVELOPMENT	5,566