Labour Economics 94 (2025) 102725

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Unveiling citation bias in economics: Taste-based discrimination against Chinese-authored papers☆

Xiaoliang Yang [a], Peng Zhou [b,*]

[a] Zhongnan University of Economics and Law, China
[b] Cardiff University, D47, Aberconway Building, Colum Drive, Cardiff CF10 3EU, United Kingdom

ARTICLE INFO

JEL classification:
J15
J16
J44
J71

Keywords:
Ethnic Discrimination
Citation
Publication
Knowledge Diffusion

ABSTRACT

We present evidence for taste-based discrimination against Chinese first authors in economic citations. We utilize a gravity model of citations and interpret the bias as a negative effect of "cultural distance". After controlling for quality as well as author-, paper-, and journal-specific attributes, publications with a Chinese first author receive 14 % less citations. Coauthoring with non-Chinese does not mitigate the discrimination at all. While being affiliated with a US-based institute slightly reduces the bias by dampening the perceived "Chineseness", it is not big enough to offset the discriminatory effect. Moreover, the COVID pandemic exacerbated the discriminatory effect. The forensic analysis narrowed down the source of discrimination to non-Chinese top economists from non-US affiliations.

1. Introduction

Ethnic discrimination has been extensively investigated in labor markets (Charles & Guryan, 2008) and consumption markets (Christensen & Timmins, 2022). However, little has been done in ideas markets. On the supply side, researchers publish articles to generate academic impacts, typically measured by citations (Rubin & Rubin, 2021). Ethnic discrimination by citers on the demand side can lead to disparities in impact and hinder knowledge diffusion.

Using descriptive statistics from the Scopus database, economic articles with Chinese first authors receive significantly fewer citations (Fig. 1). The proportion of Chinese first-authored papers remains stable throughout the sample years (ranging from 14.5 % to 16.6 %), and the age of the papers across ethnicities is similar (Chinese = 1438 days V.S. non-Chinese = 1420 days), ruling out the timing of publication as a contributing factor. Since only articles published in top journals[1] are kept in the sample, it is highly improbable that all leading economics journals unanimously and persistently publish low-quality articles written by Chinese authors. Given the rigorous quality control standards of these journals, we can reasonably dismiss systematic quality differences as the primary cause of the observed citation disparity. After controlling for a comprehensive set of journal-, paper-, and author-related attributes, the most plausible explanation is ethnic discrimination.

Examples of ethnic animus are not hard to find within civil society, such as the rise of Islamophobia after the 911 attacks and the escalation of anti-Asian hate incidents during the COVID pandemic. Nevertheless, it is anticipated that the academic community would exhibit lower levels of ethnic discrimination, considering that researchers are expected to possess the ability to think independently, free from media influence, political propaganda, and stereotypical prejudice. So, we tried all the econometric techniques an applied economist can think of—controlling for possible factors, correcting for various biases, trying different estimation methods, and limiting to different subsamples—to reject the hypothesis of ethnic discrimination. Despite all these efforts, the coefficient of the Chinese dummy still stands firmly negative. Eventually, we

☆ We would like to dedicate this paper to Xiaoliang Yang's wife, Yingying Shi, in gratitude for her unwavering support, and to the loving memory of their late son, Youzhi Yang. We are also grateful for the constructive comments from Prof. James Foreman-Peck (Cardiff University) and Prof. Stefano DellaVigna (UC Berkeley). We acknowledge the financial support from the Hubei Provincial Department of Education Philosophy and Social Science Research Project (24G094) and the Major Program of National Fund of Philosophy and Social Science of China (23ZDA097).Data will be made available on request.

* Corresponding author.
E-mail addresses: xiaoliang@zuel.edu.cn (X. Yang), zhoup1@cardiff.ac.uk (P. Zhou).

[1] We only keep articles published in Grade-4* and Grade-4 journals in the Academic Journal Guide published by the Association of Business Schools (ABS).

https://doi.org/10.1016/j.labeco.2025.102725
Received 3 April 2024; Received in revised form 11 April 2025; Accepted 20 April 2025
Available online 21 April 2025
0927-5371/

gave up trying to falsify the hypothesis and turned to understanding the neglected ethnic discrimination in economic research.

The first contribution of this paper is empirical. We fill a critical gap in the literature of ethnic discrimination in academic research. The ideal identification strategy is to compare the citation gap between two papers that are identical in all aspects except for the authors' ethnicities. However, conducting experiments like this is evidently impractical—it violates the principle against plagiarism. While identifying discriminatory effects using observed data may face challenges related to unobservable information, this concern has been mitigated to the minimum. Following the well-established convention in citation research (e.g., Medoff, 2003; McCabe & Snyder, 2015), we control all possible author-, paper-, journal-specific factors. Specifically, we restrict our sample to leading economic journals to ensure the quality of articles and to prevent preferential treatment towards Chinese authors due to unobservable academic networks. Even if unobservable networks do exist, Chinese authors would be disadvantaged as a minority group and must overcome a higher quality barrier to publish in leading journals (mainly in the US and Europe), making our estimate a lower bound. Surviving various robust checks, ethnic discrimination against Chinese first authors is found in economic research. Our results indicate that little can be done by Chinese economists to mitigate the taste-based ethnic discrimination unless they affiliate themselves with a US institution. Additionally, by employing a difference-in-differences method (event study), we demonstrate that the discrimination against Chinese economists has intensified amidst the COVID pandemic.

The second contribution is theoretical. We provide a novel gravity model framework to explain citations. Ethnicity is interpreted as part of the "cultural distance" among researchers, so the distance is negatively related to the likelihood to be cited by non-Chinese authors and positively related to that by Chinese authors. We find evidence for both, so it implies that the Chinese authors have also engaged in some sort of reverse discrimination.

The next section comprehensively reviews the literature based on which we develop four testable hypotheses. Section 3 discusses the empirical strategy of the paper including the data, the measure, and the model. Section 4 presents the estimated discriminatory effect with robustness tests and heterogeneity checks. Section 5 investigates the dampening effect of "Chineseness dilutors" and the exacerbating effect of the pandemic. Building on these findings, we further narrow down the discriminators by a forensic analysis in Section 6. Section 7 concludes.

2. Literature review

Economic literature on discrimination has proliferated across two dimensions: various types of discrimination (e.g., taste-based, statistical) and diverse "domains" of discrimination (e.g., labor markets, consumption markets, justice systems, academia). This section reviews relevant papers in the two dimensions, develops some testable hypotheses, and discusses alternative data strategies.

2.1. Types of discrimination

Economic research on discrimination can be traced back to *The Economics of Discrimination* by Becker (1957), who discussed racial discrimination based on taste, or taste-based discrimination (TD). TD occurs when decision-makers hold prejudice or animus against certain groups and treat them differently despite incurring higher costs. For instance, an employer with such bias may require extra compensation to hire a black employee who is otherwise identical to a white counterpart (Becker, 1971). Another primary type is statistical discrimination (SD), which occurs when decision-makers utilize observable characteristics of groups like ethnicity, race, gender, and appearance as a proxy to infer unobserved attributes like ability and productivity (Arrow, 1972; Phelps, 1972). In contrast to TD, SD does not necessarily involve personal prejudice against specific groups. It arises due to asymmetric information rather than subjective preferences (Lang & Lehmann, 2012).

Distinction between the two primary types of discrimination is crucial, as they imply entirely opposite policies. To address TD, it is advisable to withhold as much personal information (e.g., gender, ethnicity) as possible from the decision-makers to avoid prejudice. On the contrary, to deal with SD, it is suggested to disclose as much personal information (e.g., education, criminal record) as possible to the decision-makers to mitigate asymmetric information (Bohren et al., 2019). Thus, empirical literature has devoted much effort to distinguish between the two types and documented pervasive evidence for both TD and SD.

For example, TD is found to contribute to gender homophily in research (Gallen & Wasserman, 2023), racial wage differentials (Charles & Guryan, 2008), ethnic preferences in coworker selection (Hedegaard & Tyran, 2018), and abuses of force by police officers (Fryer, 2019). Evidence for SD is discovered in police vehicle searches (Knowles et al., 2001), pupil grade improvements (Botelho et al., 2015), sentencing in courts (Kleinberg et al., 2018; Arnold et al., 2018), delivery service (Castillo et al., 2013) and employers' recruitment screening (Agan &

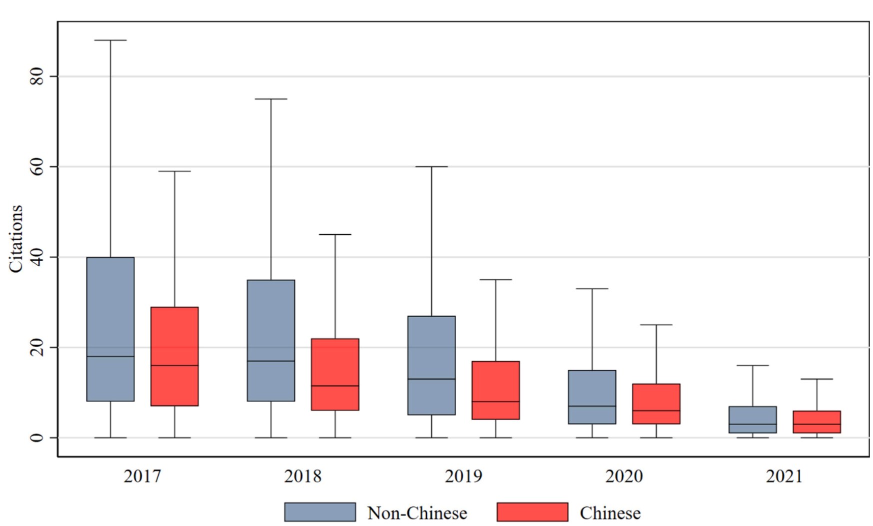

Fig. 1. Boxplot of Citations by Author Ethnicity and Publication Year.
Notes: Statistics in the boxplot include maximum, minimum, median, and quartiles. Database: Scopus. Fields: Economics, Econometrics, and Finance. Journals: AJG Grade-4 and 4*.

Starr, 2018; Doleac & Hansen, 2020). SD is typically tested by a connection between ethnic/racial/gender disparities and asymmetric information, or a reduction in these disparities after asymmetric information is reduced.

Nevertheless, it is not always straightforward to separate TD from SD because asymmetric information not only exists between discriminators and discriminated-against groups, but also lies between discriminators and researchers (Brock et al., 2012). The two asymmetries spark an interesting debate on racial discrimination in police vehicle searches. Grogger & Ridgeway (2006) put forward the hypothesis of "veil of darkness" that police officers have difficulty in distinguishing races at night, and they find no evidence for TD in police vehicle search by comparing the search rates between days and nights. Conversely, Horrace & Rohlin (2016) point out the neglect of nighttime lighting and conclude contrasting evidence. The debate arises because the researchers do not know what the discriminators know, which makes the distinction between TD and SD difficult.

Fortunately, in our case it is straightforward to determine the type of discrimination against Chinese authors. On the one hand, citers have full information on the published papers, so there is no asymmetric information between the discriminators (citers) and the discriminated-against group (Chinese authors). On the other hand, the published papers are also accessible by anyone, so there is no asymmetric information between the researchers and the discriminators either. The symmetric information set refutes the condition for SD. Citations are supposed to be guided by substantial relevance and research quality rather than author ethnicity, so any ethnic gap in citations ceteris paribus suggests TD. Given the argument above and the observation in Fig. 1, we propose the hypothesis of ethnic discrimination [H1].

[H1] A paper with Chinese first author receives fewer citations due to TD.

2.2. Domains of discrimination

Empirical research on discrimination has expanded to various domains such as labor markets, consumption markets, justice systems, and academia. As each field evolves, empirical studies go beyond whether discrimination exists and pay attention to how discrimination is mitigated and exacerbated in different contexts (moderation effects).

Discrimination in labor markets received the earliest research attention. It can occur during job seeking (Bertrand & Mullainathan, 2004), while at work (Glover et al., 2017), in career promotion (Stewart, 1983), and across generations (Lang & Spitzer, 2020). In addition, moderation effects in discrimination are also investigated. For example, ethnic discrimination can be exacerbated by the disparities in the costs of acquiring skills during childhood (Neal & Johnson, 1996). Fictious résumés are usually utilized to study the discriminatory effect and moderation effects like neighborhood quality (Giuliano et al., 2009) and anti-discriminatory policy (Agan & Starr, 2018). In general, discrimination in labor markets is mainly TD, so promoting competition in labor markets can mitigate ethnic and racial biases (Becker, 1957).

Consumption markets represent another major domain. Ethnic biases have been found in house rentals (Yinger, 1986), house sales (Christensen & Timmins, 2022), and mortgage applications (Tootell, 1996). Discrimination in housing markets is mainly SD, so a mitigating factor is to disclose more signals such as nonsmoking behavior and respectable occupations (Ewens et al., 2014). Similar discriminatory effects are found in B&B bookings (Edelman et al., 2017), used car markets (Zussman, 2013), taxi services (Castillo et al., 2013) and retail stores (Leonard et al., 2010). External events such as outbreaks of violence between Israelis and Palestinians can exacerbate discriminatory effects (Bar & Zussman, 2017), while market competition can mitigate discrimination.

Justice systems, in contrast to markets, cannot introduce competition to reduce discrimination. Econometricians identify various moderation effects such as the race of the police on duty (Antonovics & Knight, 2009), reassignment of officers (Goncalves & Mello, 2021), nighttime lighting (Horrace & Rohlin, 2016), and forms of force (Fryer, 2019). In the downstream of the justice system, the fairness of court trials has also been challenged. Identified moderation effects include the racial composition of the jury pool (Anwar et al., 2012), experience of judges (Arnold et al., 2018), and media coverage (Rehavi & Starr, 2014). Due to the information asymmetry between researchers and decision-makers, it is difficult to distinguish between SD and TD.

Academia is a domain with less information asymmetries. For example, female students in course grading tend to be discriminated against by male graders (Carrell et al., 2010; Jansson & Tyrefors, 2022), but this can be mitigated by gender-balanced academic councils (De Paola & Scoppa, 2015). Gender discrimination is also extended to evaluation of research outputs (Krawczyk & Smyk, 2016), application of research grants (Bukstein & Gandelman, 2019), teaching evaluation (Keng, 2020), and nomination of committee members (Card et al., 2022), but the discrimination can again be alleviated by a more gender-balanced decision panel. Nevertheless, these studies mainly concentrate on gender discrimination. Literature on ethnic discrimination in academia is scanty (Blackaby & Frank, 2000; Dee, 2004). The most recent attempt to study ethnic discrimination in academic publications is Liu et al. (2023). Their study, which examines total citations per author rather than per article as our paper presents, concludes that white authors receive more citations. However, their comparative analysis based on sample average lacks controls for journal-, paper-, and author-related attributes that influence citations, as well as the quality of the articles themselves. To the best of our knowledge, our paper presents one of the first attempts to examine ethnic discrimination in economic publications on a per-article basis.[2]

As summarized in [H1], the ethnic bias in citations is TD against Chinese authors, so diluting the "Chineseness" should dampen the discriminatory effect. An author from ethnic minority (Chinese) faces two mitigating strategies: to dilute the Chineseness of the *paper* (e.g., working with non-Chinese coauthors) and to dilute the Chineseness of the *author* (e.g., working in non-Chinese affiliations). The two Chineseness dilutors result in two dampening hypotheses:

[H2A] Coauthoring with non-Chinese dampens the discriminatory effect.
[H2B] Affiliation outside China dampens the discriminatory effect.

Moreover, the level of discrimination is not constant; rather, it can shift after certain events. Ethnic discrimination against Muslim immigrants increased significantly after the 911 attacks (Gould & Klor, 2016), and such discrimination even spread to other minorities such as Hispanics (McConnell & Rasul, 2021). The Brexit has elevated the perception of discrimination against non-EU immigrants in the UK (Rienzo, 2024). The most serious global emergency in the recent years was undoubtedly the COVID pandemic and it was closely tied to Chinese. Research has documented a sudden antipathy towards Chinese and other Asians following the COVID outbreak (Lu & Sheng, 2022). The Academic community is part of society, so the pandemic may exacerbate ethnic discrimination in research. On this account, we propose the exacerbating hypothesis.

[H3] The COVID pandemic exacerbates the discriminatory effect.

2.3. Data strategy

These hypotheses above can be formally tested by data. In the empirical literature on discrimination, three types of data are commonly

[2] Two recent working papers (Qiu et al., 2023; 2024) also find "home bias" in citations, but they focus on natural sciences (e.g., chemistry).

used: experimental data, survey data, and observational data.

As a popular method of primary data research, behavioral experiments have the advantage of controlling personal attributes in identifying the discriminatory effect. One widely used approach is "audit experiments", where "auditors" of different ethnicities are trained to perform specific tasks, and researchers detect discrimination based on reactions of unwitting respondents when interacting with auditors. Examples of this approach include Yinger (1986), Fershtman & Gneezy (2001), Hedegaard & Tyran (2018) and Martinez de Lafuente (2021). A criticism of this approach is that auditors are aware of the experiment, so the observed behavior may differ from the actual behavior in real circumstances. Besides, auditors can be different in various aspects apart from ethnicity (Neumark, 2012). To address these discrepancies, "correspondence experiments" create fictious profiles with the same personal attributes except for ethnicity (disclosed by minority-sounding names), which are sent to unwitting respondents to elicit their reactions (e.g., callback for job interview). Examples of this approach include Bertrand & Mullainathan (2004), Zussman (2013), Ewens et al. (2014), Edelman et al. (2017), Agan & Starr (2018) and Berson et al. (2020). The limitation of this approach is that callbacks are influenced by other real profiles sent to the employers at the same time, which cannot be controlled in the experiments. To summarize, behavioral experiments are not suitable for studying ethnic bias in citations because it is difficult to verify the expressed willingness of citing Chinese-authored papers (as in audit experiments) and infeasible to create fake papers for economists to cite (as in correspondence experiments).

Another prevalent method of primary data research is surveys. Some papers use surveys as the only data source (Stewart, 1983; Neal & Johnson, 1996; Tootell, 1996; Blackaby & Frank, 2000; Charles & Guryan, 2008; Bar & Zussman, 2017). Others use surveys as a complement to other secondary data sources (Glover et al., 2017). A common issue for all primary data methods is that respondents may conceal their true discriminatory propensity. Even if respondents are honest, the reliability of the responses is also questionable because the environmental factors surrounding actual decisions cannot be perfectly mapped into the questionnaire (Heckman, 1998).

In contrast, observational data from records of actual activities are widely available in various domains such as labor markets (Kahn & Sherer, 1988; Giuliano et al., 2009; Pierre-Philippe et al., 2016), consumption markets (Leonard et al., 2010; Rubinstein & Brenner, 2014), justice systems (Knowles et al., 2001; Grogger & Ridgeway, 2006; Anwar et al., 2012; Horrace & Rohlin, 2016; Arnold et al., 2018; Fryer, 2019; Goncalves & Mello, 2021), and academia (Dee et al., 2004; Botelho et al., 2015). These secondary data reflect the actual rather than claimed decisions, but identification of the discriminatory effect is always a challenge.

3. Empirical methodology

To combine the advantages of primary and secondary data, we adopt a mixed-method approach in data collection. It utilizes both observational data and survey data to measure actual and perceived Chinese ethnicity. Following the discussion of the two datasets, we also propose an empirical gravity model of citations where the discriminatory effect is interpreted as one of the "distance" effects.

3.1. Data on chinese surnames

Chinese surnames offer an ideal case for ethnicity identification due to the unique linguistic features of the Chinese language. The Romanization spelling system of Chinese ("Pinyin") was developed in 1950s to bridge the gap between Chinese scripts and other languages. It is an artificially designed spelling system, so unlike natural languages, there is a strict set of mapping rules between pronunciation and Pinyin spelling. For example, the pronunciation /i:/[3] is always spelled as "i" in Pinyin and any "i" in Pinyin spelling is always pronounced as /i:/ (e.g., surnames Bi, Di, Ji). By contrast, in English, the pronunciation of /i:/ can be spelled as "ee" (e.g., tree) or "ea" (e.g., tea) among many others, while "ea" can also be pronounced as /ɪə/ (e.g., idea) or /ɛ/ (e.g., bread). Featured spelling combinations in Pinyin such as "zh" (e.g., surnames Zhang, Zhou, Zhu) are unusual in other languages, making them easy to spot.

Moreover, each Chinese character has only one syllable composed of one consonant and one vowel (sometimes there is no consonant). The single-syllable structure of characters makes Chinese surnames stand out. For example, in the famous Durbin-Wu-Hausman test, "Wu" is a Chinese surname as it contains only one consonant "W" followed by one vowel "u". The single-syllable structure results in a small number of distinct sounds or "limited phonemic inventory", so homophones are very common.[4] In fact, four homophonic surnames (巫吴武戊) share the same Romanized spelling of "Wu". Homophones substantially reduce the number of distinctive Pinyin spellings of Chinese surnames, which further facilitates the identification.

Additionally, Chinese surnames maintain a stable set thanks to the cultural continuity and ethnic homogeneity. Most Chinese surnames nowadays can be traced back to 4000 years ago (Bai & Kung, 2022). Surnames carry clan lineage which holds significant value in the Chinese culture, so people rarely change their surnames. By the Song dynasty (960–1279 CE), Chinese surnames had evolved into a relatively fixed set (Liu et al., 2012). According to the official statistics, there are about 6000 Chinese surnames nowadays, and the top 100 surnames account for 86 % of the population (China's Ministry of Public Security, 2021). The limited set of popular Chinese surnames provides another condition for distinguishing Chinese from others.

The first author of an article, considered to benefit the most (Einav & Yariv, 2006; Ray & Robson, 2018), is the subject of this study. Following Kerr (2008), Kerr & Lincoln (2010), and Bai & Kung (2022), we utilize a comprehensive list of Chinese surnames to determine the ethnicity of authors (Zhu, 2009). However, a typical non-Chinese may not know the comprehensive list of Chinese surnames. A more realistic assumption is that only some very popular Chinese surnames such as "Chen" and "Wang" are known to a typical non-Chinese. To reflect this reality, we develop two measures of "Chineseness" alternative to the actual Chinese surname dummy (C). One is based on the *census* distribution of Chinese surnames ($\overline{C}$), and the other is based on the *perceived* identification of Chinese surnames ($\widehat{C}$).

For $\overline{C}$, we quantify the Chineseness of a surname in terms of its relative popularity in the Chinese census (China's Ministry of Public Security, 2021). It results in a continuous measure between 0 (non--Chinese surnames) and 1 (most populous Chinese surnames). Following Spenkuch et al. (2018), we define $\overline{C}_k \equiv (K-k)/(K-1)$ to measure the Chineseness of the k-th popular surname in the census where $k \in [1, K]$. All non-Chinese surnames are ranked in the last place (K), so $\overline{C}_K = 0$.

For $\widehat{C}$, we sent a survey to 43 non-Chinese economists (our colleagues and academic friends in the US, the UK, and the EU) and asked them to judge if the Chinese surnames seem Chinese to them. It results in another continuous measure between 0 (no one thinks it is Chinese) and 1 (all respondents think it is Chinese). $\widehat{C}$ is a subjective cognition of Chinese surnames, which can be different from the census-based measure $\overline{C}$. For

[3] The notations like /i:/ and /ɛ/ in this paper are based on the International Phonetic Alphabet (IPA).

[4] To compensate for the limited phonemic inventory, Chinese evolved to a tonal language, in which the pitch or tone of a character can change its meaning. In contrast, stress-timed languages like English are characterized by irregular intervals between stressed syllables of words.

example, popular Chinese surnames “Cui” and “Ma” were not picked up by most respondents (Fig. 2).

A scatter plot demonstrates the relationship between $\overline{C}$ (census-based measure) and $\widehat{C}$ (survey-based measure) in Fig. 2. For the top ten surnames in the two measures, five of them are in common. We run a Spearman test under the null hypothesis “$\overline{C}$ and $\widehat{C}$ are independent”, which is rejected at the 0.001 significance level. The positive regression line suggests that, despite some discrepancies, the actual and perceived Chinese surnames are generally consistent.

3.2. Data on published articles

The primary dependent variable, citations (y), is commonly used to measure the research impact of a paper (Furman & Stern, 2011). It is defined as the cumulative citations up to the data collection day (February 1st, 2023), so we have a cross-sectional dataset. The choice of cut-off date has no effect on the regression coefficients since the intercept term can absorb the difference (Davis, 2009). An alternative approach is to compare citations of articles with the same age (e.g., Gaulé & Maystre, 2011), but it must forgo substantial observations of articles published on different dates. Data on another dependent variable (to be examined in Section 5), publication time (T), are extracted from the paper texts or the corresponding journal website. Collected information includes dates of receipt, acceptance, online publication, and offline publication. Publication time T can be defined as pre-acceptance time (between submission date and acceptance date) or post-acceptance time (between acceptance date and publication date, online or offline whichever is earlier) or the sum of the two.

Control variables are based on information retrieved from paper texts, authors, and journals. Most variables are directly observable, but some need to be constructed. For example, we match each author to the Web of Science list of top 50,000 economists to capture author ranking. We then generate a dummy for being ranked in the top 250 (top 0.5 %) to control for the “star effect” (Gaulé & Maystre, 2011). Another constructed control variable is paper type based on textual analysis of the title, keywords, and abstract (Hamermesh, 2018). After removing editorials, errata, and comments, we have seven paper types: review, theoretical, econometric theory, numerical simulation, empirical, experimental, and mixed. The former four belong to theory-oriented types while the latter three are empirics-oriented.

For quality control, we restrict our sample to reputable economics journals rated as Grade-4* and Grade-4 in the Academic Journal Guide (AJG, also known as the ABS).[5] The list includes *American Economic Review (AER), Econometrica (ECM), Quarterly Journal of Economics (QJE), Journal of Political Economy (JPE), Review of Economic Studies (RES), Biometrika, Econometric Theory, Economic Journal, International Economic Review, Journal of Econometrics, Journal of Economic Theory, Journal of International Economics, Journal of Labor Economics, Journal of Monetary Economics, Journal of the American Statistical Association, Journal of the European Economic Association, Journal of the Royal Statistical Society Series B: Statistical Methodology, RAND Journal of Economics,* and *Review of Economics and Statistics*. Scopus is used as the database to retrieve publication information, including the title, author (s), abstract, keywords, citations, open access status, financial support information, and page number. The collected data are cross-checked with Web of Science to ensure validity. Missing information is manually filled out using data on journal websites. Table 1 summarizes descriptive statistics of the variables in the empirical model. Specifically, out of the final sample (7559 articles), 15.8 % have Chinese first authors, but only about 10 % would have been perceived as Chinese ethnicity according to perceived measures.

3.3. Empirical model

A linear regression model of citations can be written as Eq. (1), where C is equal to 1 if the first author has a Chinese surname and $\mathbf{x}$ is a vector of control variables.

$$y_i = \delta C_i + \mathbf{x}_i'\boldsymbol{\beta} + \epsilon_i \tag{1}$$

In practice, the $\ln(1+y)$ form is often used to interpret coefficients as elasticities and to avoid observation loss when $y = 0$ (Card & DellaVigna, 2013; Galasso & Schankerman, 2015; Rubin & Rubin, 2021). Despite its popularity, the use of $\ln(1+y)$ receives extensive criticism for its inconsistent estimates especially when y is small (Cohn, et al., 2022). Noting that citations are non-negative, integer-valued, and skewed towards zero, the negative binomial (NB) regression (2) is a more appropriate specification than (1). It includes Poisson distribution as a special case where the mean is equal to the variance, i.e., the overdispersion parameter $\alpha = 0$ in the Gamma distribution Γ.

$$\Pr(y = y_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}, \text{ where } \lambda_i = e^{\delta C_i + \mathbf{x}_i'\boldsymbol{\beta} + \epsilon_i}, \ e^{\epsilon_i} \sim \Gamma\left(\frac{1}{\alpha}, \alpha\right) \tag{2}$$

Vector $\mathbf{x}$ contains control variables following the literature on citations:

- author-related attributes: number of authors (Medoff, 2003; Rubin & Rubin, 2021), ranking of authors (Gaulé & Maystre, 2011), affiliation of authors (Li & Yi, 2020).
- paper-related attributes: paper age (McCabe & Snyder, 2015), number of pages (Card & DellaVigna, 2014; Rubin & Rubin, 2021), open access (Davis, 2009), paper type (Angrist et al., 2017), topic, special issue (Conlon et al., 2006).
- journal-related attributes: journal fixed effects (Drivas & Kremmydas, 2020).

An important control variable is paper age (A) which is defined as days between citation retrieval date and publication date. It can be used to normalize the “exposure” of a paper to its citing pool in the NB regression. To do so, rewrite the right-hand side term $\beta \ln A = \ln A + \widetilde{\beta}\ln A$ where $\widetilde{\beta} \equiv \beta - 1$. The first term ($\ln A$) effectively transforms the dependent variable on the left-hand side from gross citations $\ln y$ to average citations: $\ln y - \ln A = \ln(y/A) = \ln\overline{y}$. The second term ($\widetilde{\beta}\ln A$) allows for a flexible effect $\widetilde{\beta}$ of paper age on average citations.

Another perspective is to treat citations as “trade” of knowledge, so a paper citing another paper is a unilateral “flow” of knowledge from paper i to papers in its citing pool. The original gravity model sheds light on how trade flow is influenced by “distances” like cultural distance (Guisoet al., 2009), language distance (Foreman-Peck & Zhou, 2015), and political relations (Zhou et al., 2021). Citations can be viewed as a “trade” of knowledge, so the gravity model also provides a useful theoretical lens through which we can better understand the empirical results in the next three sections. The general form of the gravity model of citations can be written as:

$$y_i = f(M_i, \overline{M}_i, \Delta_i, G), \text{ where :} \tag{3}$$

- the “mass” of paper i (M_i): number of authors, ranking of authors, affiliation of authors, number of pages.
- the “mass” of the citing pool of paper i ($\overline{M}_i$): open access, paper age.
- the “distances” between paper i and its citing pool (Δ_i): **Chinese first author**, paper type, topic, special issue.
- the “gravity constant” (G): journal fixed effects.

[5] AJG is an academic journal guide published by the Association of Business Schools (ABS), in which journals are rated Grade-4*, 4, 3, 2 and 1. It is widely accepted by research institutes (explicitly or implicitly) as the criteria to evaluate research outputs.

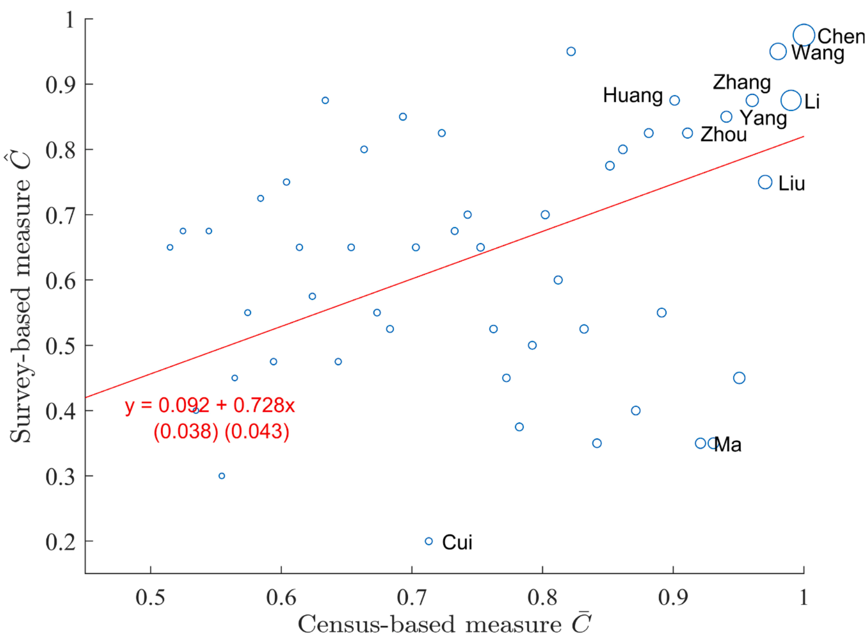

Fig. 2. Relationship between census-based and survey-based measures.
Notes: The size of circles indicates the frequency of surnames in our sample. "Chen" is the most frequent surname which appears 109 times as first authors in our sample (7559 papers in total and 1194 first-authored by Chinese).

Specifically, the gravity model suggests that authors with shorter cultural distance (e.g., between the US and the UK) tend to cite each other more than those with greater cultural distance (e.g., between the US and China). Similarly, it is also expected that Chinese authors tend to cite Chinese authors more than non-Chinese authors for the same reason, which implies a form of reverse discrimination. It is referred to as "home bias in citations" by Qiu et al. (2024). Furthermore, we argue that the effect of cultural distance reflects taste-based discrimination as stated in [H1]. This is because, in principle, citations should be based on the content of the paper, rather than the ethnicity of the authors. Note that there is no asymmetric information regarding the content of a published paper, so it cannot be statistical discrimination. If an author still prefers to cite papers written by a particular ethnic group with shorter cultural distance, then this is clearly taste-based.

4. The discriminatory effect

In this section, we begin with the baseline results, which are then challenged by various robustness tests. After testing the hypothesis of ethnic discrimination [H1], we then try to understand how Chinese authors are discriminated against by heterogeneity regressions. A very important point is re-stated here. In this study, we are *not* comparing a published paper with a rejected paper. We are *not* either comparing a paper published by *American Economic Review* with one published by *China Economic Review*. Citations are compared ceteris paribus among published papers after controlling a comprehensive set of factors.

4.1. Baseline

In the baseline results, the "treatment" (C: first author having a Chinese surname) is included as an intercept effect, so slopes/coefficients of other regressors are assumed to be homogenous for both Chinese and other ethnic groups. The estimated coefficients from Eqs. (1) and (2) under this assumption are reported in Table 2. As a special case of NB regression, the Poisson regression is also presented in column (2)'. To facilitate interpretation and comparison, we calculate the marginal effects beneath the coefficients. The three sets of estimation results are qualitatively consistent, so the NB regression is used as the main baseline result in the following analysis (this choice is justified in Subsection 4.2).

To identify the discriminatory effect, we consider a comprehensive set of control variables. The findings on these control variables in Table 2 are consistent with the literature on citations. Start with the *author-related* factors. One additional coauthor can enhance the citations by 1.35 times, which may reflect the division-of-labor effect in knowledge production. We find that coauthoring has become a secular trend in economic research—on average there are 2.47 authors in a paper (Table 1) and only 17.6 % papers are sole-authored in our sample. Nonetheless, the benefit derived from the *quantity* of authors is much smaller than that derived from the *quality* of authors. The top 250 economists (the ISI Web of Science list) can attract 6.33 more citations, which are known as the "star effect" (Gaulé & Maystre, 2011). Affiliations in the US significantly promote citations thanks to the high reputation and wide networks.

Turning to the *paper-related* factors, articles with an older age (A), a longer length (pages), and open access are more exposed to the citing pool, so they have higher citations. These findings are consistent with the general literature on citations (Davis, 2009; Rubin & Rubin, 2021). Moreover, the topic of the article can also significantly affect citability. We undertook textual analysis of titles, abstracts, and keywords. Papers on "China" and "United States" respectively receive 10.5 and 3.9 more citations than those on neither. Thus, China-related topics should attract more citations, *not* less. Given that Chinese authors are more likely to study China-related topics (6.37 %) compared to non-Chinese authors (1.27 %), the lower citations of articles written by Chinese authors become even more puzzling. Another control variable related to papers' topics is the special issue dummy. It is found that papers published in special issues receive more citations, because these papers usually belong to extensively discussed topics. In addition, the models also control for article types (e.g., theoretical, empirical, review, etc.) and journal-specific fixed effects.

After accounting for all relevant factors, the coefficient (δ) of the Chinese dummy (C) remains significantly negative. The implied marginal effect of C represents the discriminatory effect defined as $E[y|C = 1] - E[y|C = 0]$. All three models yield similar results. For example, in the NB regression, articles with a Chinese first author receive, on average,

Table 1
Descriptive statistics.

	Mean			S.D.	Range
	Chinese (N=1194)	Non-Chinese (N=6365)	All (N=7559)	All	All
Dependent variables					
Citations (y)	14.068	21.153	20.034	38.061	[0,1416]
Publication time in days (T)	658.90	730.39	716.88	363.98	[2,5687]
Waiting days before acceptance	564.92	612.39	603.44	345.58	[1,5632]
Waiting days after acceptance	93.284	113.28	109.61	138.31	[1926]
Independent variables					
Chinese first author (C)			0.1580	0.3647	{0,1}
Census-based Chineseness ($\widetilde{C}$)			0.1035	0.2775	{0,1}
Survey-based Chineseness ($\widehat{C}$)			0.1074	0.2633	{0,1}
No. of authors	2.6834	2.4317	2.4715	1.0813	[1,18]
Top 250 economist	0.0503	0.0624	0.0605	0.2383	{0,1}
Affiliations in China	0.2345	0.0022	0.0389	0.1934	{0,1}
Affiliations in the US (D_{US})	0.5226	0.4885	0.4938	0.5000	{0,1}
Affiliations in others (D_{NN})	0.2429	0.5093	0.4673	0.4990	{0,1}
Non-Chinese coauthor (D_{CO})	0.4506		0.9132	0.2816	{0,1}
Paper age in days (A)	1438.0	1420.1	1422.9	520.74	[412,2759]
Pages	21.372	28.503	27.377	13.582	[1130]
Open access (O)	0.0687	0.1321	0.1221	0.3274	{0,1}
Special issue	0.0879	0.0987	0.0970	0.2959	{0,1}
Topics related to China	0.0637	0.0127	0.0208	0.1426	{0,1}
Topics related to the US	0.2839	0.1703	0.1883	0.3909	{0,1}
Funding	0.6658	0.5219	0.5446	0.4980	{0,1}
Theory-oriented paper types	0.2010	0.2894	0.2754	0.4468	{0,1}
COVID (E)	0.2764	0.3032	0.2990	0.4578	{0,1}

Notes: The date of publication ranges from 2016–2021. In the range column, square brackets indicate ranges of continuous variables, and curly brackets indicate ranges of discrete/integer variables. Due to incomplete information on publication timelines for some journals, the sample size for "Publication time in days" is 5554 (with 1049 papers authored by Chinese first authors), for "Waiting days before acceptance" it is 5564 (with 1049 Chinese first authors), and for "Waiting days after acceptance" it is 5778 (with 1059 Chinese first authors).

Table 2
Baseline results.

	(1) Linear	(2) NB	(2)' Poisson
Chinese first author (C)			
coefficient (δ)		**-0.1558*****	-0.1711***
		(0.0338)	(0.0090)
marginal effect	-2.5146**	**-3.1103*****	-3.4280***
	(1.1729)	**(0.6785)**	(0.1811)
No. of authors			
coefficient		0.0675***	0.0729***
		(0.0109)	(0.0024)
marginal effect	1.1759***	1.3477***	1.4603***
	(0.3823)	(0.2200)	(0.0485)
Top 250 economists			
coefficient		0.3171***	0.4009***
		(0.0464)	(0.0084)
marginal effect	11.9934***	6.3311***	8.0318***
	(1.6746)	(0.9365)	(0.1693)
Affiliation in US			
coefficient		0.1854***	0.1877***
		(0.0235)	(0.0056)
marginal effect	3.2562***	3.7023***	3.7601***
	(0.8263)	(0.4746)	(0.1119)
Paper age (ln$A$)			
coefficient ($\widetilde{\beta}$)		0.5927***	0.6053***
		(0.0285)	(0.0083)
marginal effect	10.7965***	11.8342***	12.1266***
	(0.9827)	(0.6511)	(0.1701)
Paper length (ln($pages$))			
coefficient		0.6762***	0.7103***
		(0.0312)	(0.0074)
marginal effect	12.5485***	13.5004***	14.2306***
	(1.0788)	(0.6764)	(0.1536)
Open access			
coefficient		0.2646***	0.2383***
		(0.0348)	(0.0073)
marginal effect	6.7189***	5.2839***	4.7734***
	(1.2352)	(0.7048)	(0.1467)
China-related topics			
coefficient		0.4409***	0.3982***
		(0.0800)	(0.0152)
marginal effect	9.8411***	10.4810***	9.3904***
	(2.8433)	(2.3273)	(0.4284)
US-related topics			
coefficient		0.1879***	0.1379***
		(0.0431)	(0.0081)
marginal effect	3.7479***	3.9111***	2.8392***
	(1.5069)	(0.9494)	(0.1730)
Special issue			
coefficient		0.3928***	0.2371***
		(0.0443)	(0.0105)
marginal effect	4.2307***	7.8434***	4.7499***
	(1.5281)	(0.9020)	(0.2102)
Paper type dummies	Yes	Yes	Yes
Journal dummies	Yes	Yes	Yes
Constant	Yes	Yes	Yes
Overdispersion (α)		0.8191***	
		(0.0141)	
Observations	7559	7559	7559

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %. To be comparable with NB/Poisson regressions, the dependent variable of column (1) is standardized as (citations/exposure) × (average citations). Marginal effects are average marginal effects at original values of independent variables.

3.11 fewer citations than those with a non-Chinese first author. It corresponds to a $1-\exp(-0.1558)=14.4\ \%$ reduction in citations. This ethnic bias is unjustifiable because citations should reflect the relevance and quality of the paper, not the ethnicity of the author(s). SD does not apply in this context since there is no ambiguity regarding the content of a paper once it has been published. Therefore, the disparity in citations is attributed to TD as stated in [H1].

We can also interpret the baseline results using the gravity model of citations in Eq. (3). As a paper's "mass" M_i rises (e.g., number of authors, ranking of authors, affiliated in the US, number of pages), the citations increase. Similarly, as the "mass" of its citing pool $\overline{M}_i$ expands cross-sectionally (open access) and temporally (paper age), the citations also grow. "Distances" among papers Δ_i can be affected by topic relevance, issue popularity, and type similarity. Specifically, author ethnicity (C in our case) creates a further "cultural distance" due to lack of trust (Guisoet al., 2009), so the "exchange" (citations in our case) is reduced.

It is noted that the base group of the regression ($C = 0$) also contains other ethnicities potentially subject to discrimination. For example, some Indian surnames like Gupta, Khan, and Banerjee are also easily identifiable. If Indian authors are also discriminated against, then the base group would have lower citation counts than what would be expected without discrimination. Therefore, our estimate of δ represents a lower bound of the discrimination against Chinese authors.

4.2. Robustness

To test the robustness of [H1], the baseline results are challenged in various dimensions, but the hypothesis of ethnic discrimination passes all the tests.

First, we check the robustness of [H1] to **specification bias.** Table 2 indicates that the conclusion holds true across different specifications (linear, NB, and Poisson). Arguably, NB regression is a more appropriate specification than linear regression in modeling count variables like citations. The NB regression is also empirically superior to the restricted Poisson regression in terms of predictive accuracy (Fig. 3). The restriction hypothesis $\alpha = 0$ can be rejected at a high significance level, so the NB regression is justified to be the baseline. Besides, zero-inflated models (ZIP and ZINB) are not appropriate in this context because zero citations are not excessive (7.28 %) and there is no threshold effect of being cited for the first time.

Second, we evaluate **bias from unobservables** in the spirit of Altonji et al. (2005). They provide a measure (selection ratio) to assess the extent of potential bias stemming from unobservable factors—how much stronger selection on unobservable factors, relative to selection on observable factors, must be to explain away the full estimated effect. Following Nunn & Wantchenkon (2011), the key coefficient $\hat{\delta}$ in the baseline result is re-estimated with restricted controls ($\hat{\delta}^R$), and the selection ratio can be calculated as $|\hat{\delta}/(\hat{\delta}^R - \hat{\delta})|$. In Table 3, we start with a simple regression with the key regressor C only. The small selection bias (0.5948) in column (1) suggests that the estimate (-0.4177) is very likely to be subject to bias from unobservables. Nevertheless, inclusion of journal dummies in column (2) removes the bias. Further inclusion of other observables, such as paper type in column (3), author-related factors in column (4), and paper-related factors in column (5), do not change the conclusion. The selection ratios range from 10 to 66, suggesting that, to attribute the entire baseline estimate to selection effects, selection on unobservables would have to be at least 10 times greater than selection on observables. Therefore, we believe that the bias from unobservables is unlikely.

Third, we check the robustness of [H1] to **endogeneity bias**. Most regressors are exogenous by definition. For example, the key variables like the first author's surname (C) and paper age (A) are both exogenously determined. However, the open-access status of a paper (O) is an endogenous decision made by authors subject to availability of research fund (the excluded instrument). To correct for possible endogeneity bias, we apply Lee (1978)'s two-part method and endogenize the decision of open access by a probit selection Eq. (4).

$$\Pr(O_i = 1) = \Phi(\mathbf{z}_i'\gamma), \text{ where } \Phi(\cdot) \text{ is the normal cumulative distribution function} \tag{4}$$

The vector $\mathbf{z}$ includes research fund and the control variables $\mathbf{x}$ in Eq. (2). The Inverse Mill's Ratio (IMR) is computed based on the selection Eq. (4) and added in the outcome Eq. (2) as an additional regressor. The estimation results are listed in column 2 of Table 4. The coefficient of IMR is not significant, suggesting that the estimated discriminatory effect is not liable to endogeneity bias.

Fourth, we check the robustness of [H1] to **selection bias**. Most top journals in our sample have a tradition of ordering authors alphabetically, and some even strictly impose this policy (e.g., *Review of Economic Studies*). Thanks to the tradition/policy, only 14.78 % papers in our sample do not follow the alphabetic order (which may cause selection bias), close to the ratio provided by Ray & Robson (2018). Therefore, the author order in our sample is predominantly random. If we discard 14.78 % non-alphabetically ordered papers, the discriminatory effect becomes even stronger (column 3 of Table 4).

Fifth, we check the robustness of [H1] to **measurement bias**. We have shown in Subsection 3.1 that the two alternative measures ($\overline{C}$ and

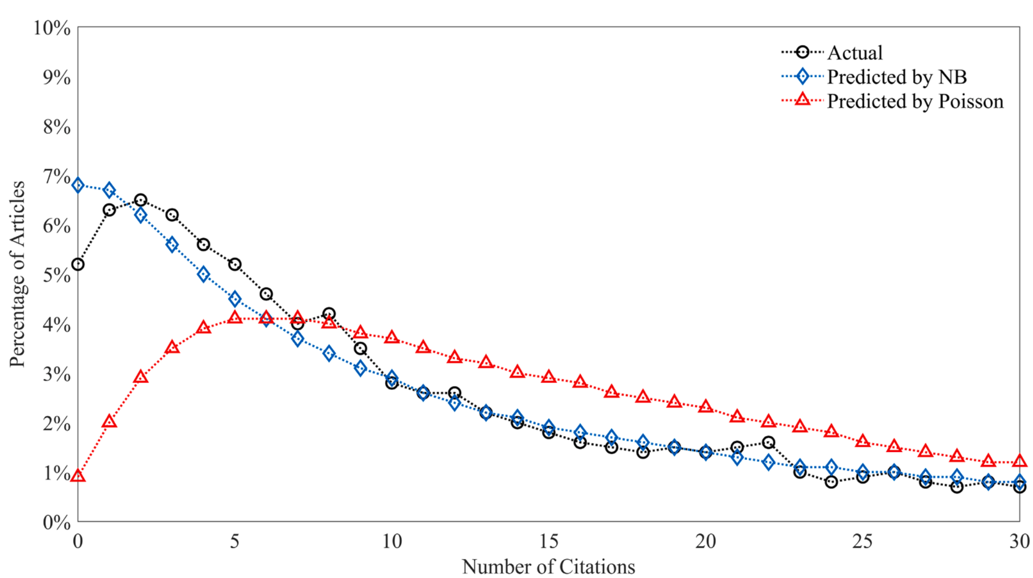

Fig. 3. Predicted citations based on the NB and Poisson regressions.

Table 3
Using selection on observables to assess bias from unobservables.

	(1)	(2)	(3)	(4)	(5)
Chinese first author (C)	-0.4177*** (0.0358)	-0.1581*** (0.0350)	-0.1479*** (0.0347)	-0.1586*** (0.0342)	-0.1410*** (0.0342)
No. of authors				0.0810*** (0.0115)	
Top 250 economists				0.3258*** (0.0485)	
Affiliation in US				0.2286*** (0.0241)	
Paper age ($\ln A$)					0.6118*** (0.0287)
Paper length ($\ln(pages)$)					0.6950*** (0.0314)
Open access					0.2608*** (0.0350)
China-related topics					0.3982 (0.0152)
US-related topics					0.1379 (0.0081)
Special issue					0.4507*** (0.0444)
Paper type dummies	No	No	Yes	Yes	Yes
Journal dummies	No	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes
Selection Ratio	0.5948	65.961	19.867	55.795	10.514
Observations	7559	7559	7559	7559	7559

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %. All selection ratios are calculated based on the baseline estimate $\hat{\delta} = -0.1558$ (Table 2).

Table 4
Robustness test against endogeneity and measurement biases.

	(1) Baseline	(2) Endogeneity	(3) Selection	(4) Measure 1	(5) Measure 2
Chinese (C)	-0.1558*** (0.0338)	-0.1540*** (0.0339)	-0.1884*** (0.0406)		
Census-based ($\overline{C}$)				-0.1424*** (0.0434)	
Survey-based ($\widehat{C}$)					-0.1961*** (0.0463)
IMR		0.1486 (0.1256)			
Control variables	Yes	Yes		Yes	Yes
Observations	7559	7559	6442	7559	7559
Model	NB	Two-Part	NB	NB	NB

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %.

$\widehat{C}$) are generally consistent but they also contain different information. The census-based measure $\overline{C}$ reflects the statistical popularity of Chinese surnames, while the survey-based measure $\widehat{C}$ reflects the subjective perception. Both are continuous measures, so they contain richer information on the Chineseness compared to the dummy variable (C). As shown in the final two columns of Table 4, the estimated discriminatory effect is stronger if $\widehat{C}$ is used instead of C or $\overline{C}$. This is because the measure $\widehat{C}$ can more accurately capture the perceived Chineseness of surnames. For example, well-known Chinese surnames like "Chen", "Wang", and "Zhang" are more likely to be targeted by discriminators. Surnames like "Ma" are easy to be confused with non-Chinese surnames (e.g., "Mar" in Arabic and Jewish, "Maa" in Finnish, and "May" in English), so the discriminatory effect is weaker. This finding entrenches the hypothesis of ethnic discrimination. It is also in line with the gravity model of citations (3) as the continuous measures ($\overline{C}$ and $\widehat{C}$) can better capture the cultural distance than the discrete dummy (C).

Sixth, a possible concern is that the estimated discrimination effect could reflect the fact that Chinese first-authored papers are skewed to recent years, giving them less time to accumulate citations.[6] To address this concern, we performed three additional tests.[7] (i) We added an interaction term between the Chinese dummy (C) and paper age ($\ln A$) to assess any moderating effect. The interaction term, 0.0729, is insignificant, suggesting that paper age does not confound the discrimination effect. (ii) We updated the citation data from February 2023 to January 2025 to extend the exposure period for all papers, and the estimated coefficient of C remains negative and significant at -0.1539***. (iii) We created a subsample that excludes Chinese-authored papers published after 2020 and non-Chinese-authored papers published before 2018. This selection artificially prolongs the exposure time of Chinese-authored papers, while the coefficient is still significantly negative at -0.1649***. Overall, these findings confirm that the discrimination

[6] We thank the anonymous reviewer for raising this concern. In fact, as shown in Table 1, the average age of Chinese first-authored papers (1,438 days) is even slightly longer than that of non-Chinese first-authored papers (1,420 days). So, moderation effect of paper age, if any, will strengthen rather than weaken our conclusion.

[7] The related empirical results can be found in Table A2 in the appendix.

Table 5
Robustness test against symmetric treatment effects.

	Baseline	RA	RA	NNM
ATE	-3.1103*** (0.6785)	-3.8598*** (0.9749)	-4.5027*** (0.8720)	-4.4279*** (1.0924)
ATET	-	-2.4166** (1.0360)	-2.3369** (1.0397)	-3.1496** (1.4947)
ATEU	-	-4.1306*** (1.0283)	-4.9090*** (0.9171)	-4.6677*** (1.1635)
Model	NB	Linear	Poisson	Mahalanobis

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %. In the NNM approach, all control variables in the models (1) and (2) are used to compute the distance metrics.

effect is not an artefact of differences in paper age.

Apart from the above-mentioned tests, the baseline results are also robust to other common empirical biases. For example, one concern is **survivorship bias**—the papers published by Chinese authors are already handpicked by editors under a "Chinese quota" as in international trade and school admission (Utar, 2018). If this is true, then our conclusions are just strengthened rather than weakened.[8]

In sum, there is strong evidence for ethnic discrimination against Chinese authors in citations. The next question is how they are discriminated against.

4.3. *Heterogeneities*

The baseline results assume that marginal effects are homogeneous for Chinese and non-Chinese authored papers. To relax this assumption, two methods are often used. Regression Adjustment (RA) allows for parametric heterogeneities between the two groups, while Nearest-Neighbor Matching (NNM) is a more flexible, nonparametric method to compare closest counterparts in the two groups. Either way, the Average Treatment Effect (ATE) can be distinguished for the treated (ATET) and for the untreated (ATEU) using the differences between the actual and counterfactual outcomes ("treatment" = Chinese first author).

Specifically, RA utilizes an estimated regression model to simulate counterfactual outcomes. This means that for Chinese first-authored papers, RA predicts what their citations would have been had the first author been non-Chinese, and vice versa for non-Chinese first-authored papers. ATET ($\equiv E[Y_{C=1} - Y_{C=0}|C = 1]$) in the treated group is similar to ATE ($\equiv E[Y_{C=1} - Y_{C=0}]$), but it uses only the papers observed in the treated group ($C = 1$). While ATEU ($\equiv E[Y_{C=1} - Y_{C=0}|C = 0]$) in the control/untreated group uses only the subjects observed in the untreated group ($C = 0$).

In contrast, the NNM method works by identifying the "nearest" match for each paper based on the distance between attributes of papers. For each Chinese first-authored paper, we find the non-Chinese first-authored paper that has the most similar covariate profile based on the Mahalanobis distance. The citation difference between the two matched papers is then calculated, and the average of these differences provides an estimate of the ATET. Conversely, to compute the ATEU, we reverse the process by taking each non-Chinese first-authored paper as the "treated" and searching its nearest Chinese counterpart.

Table 5 compares the baseline results with the RA and NNM results. It is shown that the overall ATEs are stronger when asymmetric treatment effects are allowed. Papers first-authored by Chinese could have received 2–3 more citations (equivalent to 17 %-18 %) if they were first-authored by non-Chinese (ATET), and papers first-authored by non-Chinese could have received 4–5 less citations (equivalent to 21 %-24 %) if they were first-authored by Chinese (ATEU). In other words, Chinese authors are "unlucky" to have Chinese surnames, but non-Chinese authors could have been even "unluckier" if they have Chinese surnames.

Table 6
Heterogeneous treatment effects.

	Sample	Coefficient (δ)	Marginal Effect	Observations
	Baseline (all sample)	-0.1558*** (0.0338)	-3.1103*** (0.6785)	7559
(1)	Journal = not grade-4* (i.e., grade-4)	-0.1426*** (0.0374)	-2.1307*** (0.5608)	5561
(2a)	Affiliation = not in China	-0.1494*** (0.0112)	-3.0375*** (0.7608)	7265
(2b)	Affiliation = in the US	-0.1805*** (0.0481)	-4.519*** (1.1934)	3733
(3)	Topic = not related to China	-0.1757*** (0.0346)	-3.4634*** (0.6857)	7402
(4)	Type = not theoretical (i.e., empirical)	-0.1893*** (0.0380)	-4.0942*** (0.8281)	5477

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %. Grade-4* journals are AER, Annals of Statistics, ECM, JPE, QJE, and RES. Specification: Baseline NB regression.

The discriminatory effect can be heterogenous not only between Chinese and non-Chinese but also within Chinese. Essentially, there are two dimensions along which the effect can differ, i.e., (i) *who* is discriminated against, and (ii) *what* is discriminated against. These questions can be answered by exploring heterogeneities over different subsamples. For (i), we separate the data by journal grades and author affiliations to see if discrimination distinguishes where papers are published and where authors are located. For (ii), we use subsamples of topics and types to see if discrimination is selective on specific contents written by Chinese.

According to the row (1) in Table 6, papers published in AJG grade-4 journals (i.e., excluding the grade-4* journals) have a similar discriminatory effect to the baseline results in Table 2. The homogeneous discriminatory effects indicate that the ethnic bias is purely taste-based rather than quality-based.

A similar discriminatory effect is found for authors in non-Chinese affiliations in row (2a) suggesting that discrimination targets *people*, not *affiliations*. If Chinese affiliations are also discriminated against, then the discriminatory effect in row (2a) should be significantly lower than the baseline. However, it is not the case, not even for those in US affiliations (2b). This is because Chinese authors have already been identified by their surnames, so affiliations do not provide additional information for ethnic discriminators. This finding implies that, if a non-Chinese author works at a Chinese institution, then s/he is not discriminated against (e.g., Philip Dybvig worked in China before he won the Nobel Prize). On the contrary, if a Chinese author works at a non-Chinese institution, even at a US institution, s/he is still subject to discrimination (e.g., Yiran Fan worked at the University of Chicago, but his JPE paper published in July/2021 received zero citation up to May/2023).

According to row (3), the ethnic bias is weaker in topics related to China. The gravity model of citations provides a theoretical lens to understand this result. As a form of knowledge diffusion, citations are negatively associated with the cultural distance (Δ_i). The between-groups distance is bigger than within-group distances (Liu et al., 2021), so non-Chinese authors tend to under-cite Chinese-authored papers. In contrast, Chinese-authored papers on China-related topics are less affected by this bias, because it is more like "intranational trade" than "international trade"—Chinese authors cite Chinese-authored

[8] Let alone that the quota per se might be a type of "institutional discrimination" (Small & Pager, 2020). A consequence of such quota is that these published papers by Chinese authors must have relatively higher quality to compete against other Chinese peers for a slimer chance of publication. Even so, we observed that the citations of these higher-quality papers by Chinese authors are still lower. In this case, the baseline results provide a lower bound estimate of discriminatory effect.

papers on China. Even for non-Chinese authors who study China-related topics, it is costly to discriminate against Chinese authors because 48.4 % of papers on China are written by Chinese in our sample. The cost of not citing those papers can partially offset the taste-based discriminatory effect. However, these positive and negative feedback mechanisms seem to trap Chinese authors in China-related topics (2.08 % in our sample), which result in segmented, immobile “ideas markets”.

Finally, we distinguish the sample by research type in row (4) of Table 6, which shows that the discriminatory effect in empirical papers is relatively bigger. Sample statistics indicate that empirical papers receive an average of 21.8 citations, significantly more than theoretical papers (13.6 citations), possibly due to the greater readability of empirical research. Despite the fact that empirical papers make up a larger proportion of articles with Chinese first authors (79.90 %) compared to those with non-Chinese first authors (71.06 %) (Table 1), the higher citation potential of empirical papers does not help narrow down the citation gap. This further supports the evidence of citation bias against Chinese first authors.

5. The moderating effects

After testing the discriminatory effect, this section explores how the effect can be mitigated [H2] and exacerbated [H3]. Findings on these moderating effects are informative for developing strategies to counter discrimination.

5.1. Dampening effect

As discussed in Section 2, there are two ways of diluting the “Chineseness” of their papers to mitigate the ethnic discrimination. One is to collaborate with non-Chinese co-authors, and the other is to have non-Chinese affiliations. We use one of the following Chineseness dilutors (D) to interact with the key dependent variable (C) in the baseline regression:

- $D_{CO} = 1$ if the first author has non-Chinese coauthors, and 0 otherwise.
- $D_{AF}^{(NN)} = 1$ if the first author works in a non-Chinese, non-US affiliation, and 0 otherwise.
- $D_{AF}^{(US)} = 1$ if the first author works in a US affiliation, and 0 otherwise.

Table 7 reports the estimated dampening effects of the Chineseness dilutors. On the one hand, coauthoring with non-Chinese researchers (D_{CO}) does not help much as long as the first author is still Chinese (column 1), in line with Qiu et al. (2023)’s finding. In academic referencing, if a paper has more than three authors (accounting for 53 % cases in our subsample), then it simply becomes “xxxx et al.” where xxxx is a Chinese surname. The role of non-Chinese coauthors in mitigating discrimination is therefore small. On the other hand, the dampening effect of non-Chinese affiliations (D_{AF}) is selective. If the Chinese author has a US affiliation, then discrimination against him/her can be significantly weaker, but the dampening effect is not significant if he/she works in countries other than the US (columns 2 and 3). Given that the US has the most top universities in the world, the dampening effect can be derived from both reputation and network. However, this dampening effect of US affiliations is not big enough to offset the discriminatory effect. To show this, we run a regression based on the subsample of US affiliation observations. This approach allows for a direct comparison between Chinese authors affiliated with US institutions and non-Chinese authors affiliated with US institutions, without any confounding effects from non-US affiliation samples (Qiu et al., 2023). The results (Table 6 Row (2b)) show that the coefficient of the Chinese dummy remains significantly negative even among US affiliated economists (-0.1805***). Therefore, despite the dampening effect of US affiliations, the discriminatory effect is still significant.

Table 7
Dampening effects.

	(1)	(2)	(3)
Chinese first author (C)			
coefficient (δ)	-0.1754*** 0.0426)	-0.1250*** (0.0379)	-0.2367*** (0.0417)
marginal effect	-3.5027*** (0.8535)	-2.4965*** (0.7580)	-4.7248*** (0.9257)
$C \times D_{CO}$			
coefficient	0.0429 (0.0569)		
marginal effect	0.8560 (1.1370)		
$C \times D_{AF}^{(NN)}$			
coefficient		-0.1016 (0.0669)	
marginal effect		-2.0293 (1.3371)	
$C \times D_{AF}^{(US)}$			
coefficient			0.1581*** (0.0584)
marginal effect			3.1562*** (1.1660)
Observations	7559	7559	7559

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %. Specification: Baseline NB regression plus interactive terms of $C \times D$.

5.2. Exacerbating effect

The previous subsection discusses strategies to dampen discrimination, while this subsection investigates shocks that exacerbate it. The COVID pandemic provides an ideal event study to apply Difference-in-Differences (DID) as it is strongly exogenous to citations. The media have reported that Asians, especially Chinese, have experienced more ethnic discrimination in the US and Europe since the pandemic. The literature has investigated how the pandemic exacerbates ethnic discrimination in social life (Lu & Sheng, 2022). Whether the pandemic exacerbates discrimination in academia [H3], however, is novel.

To intuitively demonstrate the difference, we first estimate the baseline NB specification using the pre- and post-COVID subsamples. The first two columns of Table 8 show that the discriminatory effect is indeed larger in the post-COVID subsample. To formally test whether the discriminatory effect is significantly exacerbated, we define an event dummy E, which is equal to 1 if the paper is published after March 11th, 2020 (the official date of the pandemic declared by WHO) and

Table 8
Exacerbating effect of COVID on the discriminatory effect.

	Subsample		Full sample	
	Pre-COVID	Post-COVID	Slope Dummy	Slope Dummies
Chinese (C): δ	-0.1257*** (0.0393)	-0.2947*** (0.0789)	-0.1302*** (0.0400)	-0.1183*** (0.0405)
COVID (E)				2.6267*** (0.9273)
Interactive ($C \times E$)			-0.1598** (0.0726)	-0.1684** (0.0836)
Control Variables	Yes	Yes	Yes	Yes
Control Variables $\times$ E	No	No	No	Yes
Observations	4714	1922	6636	6636

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %. Specification: Baseline NB regression. Despite not changing the results, open access articles (12.3 % of the full sample) are dropped because the pandemic dummy E is highly collinear with the collective open access agreements between JISC and major publishers (e.g., Elsevier, Wiley, Springer) in 2020–21.

0 otherwise. This date has been widely used by the literature to identify the COVID impact (e.g., Haan et al., 2022). In the “slope dummy” column of Table 8, we add an interactive term $C \times E$ to test the moderating effect of E on δ. In the “slope dummies” column, we allow for interactive terms for all slopes. In both cases, citations of papers with Chinese first authors are substantially lower after the COVID pandemic ($E = 1$), so evidence supports [H3] that ethnic discrimination has expanded from social life to academia.

6. Who are the discriminators?

The previous two sections have found supportive evidence for the hypothesis that Chinese authors are discriminated against. This section attempts to reveal who are the discriminators. For this purpose, we adopt the four-step workflow in forensic analysis: case briefing, evidence investigation, lead follow-up, and case resolution.

6.1. *Case briefing*

First, we overview the “case” of ethnic discrimination and identify all “suspects” in different phases of academic publication. As shown in Fig. 4, after an author submits the paper to a journal, there are four “suspects” for potential discriminators, i.e., editors, reviewers, publishers, and readers. In what follows, we will investigate the evidence for the four “leads” respectively.

6.2. *Evidence investigation*

Editors and *Reviewers.* If editors or reviewers discriminate against Chinese, then two consequences should follow. On the one hand, Chinese first authors tend to face a higher rejection rate. On the other hand, they tend to face longer publication time. The two consequences are correlated outcomes of ethnic discrimination, so we investigate both as “corroborative evidence” or “cross-check” (a common investigative strategy).

[Higher Rejection Rate] A higher quality threshold can be imposed either in the initial or the final decisions (E1 and E3 in Fig. 4). Editors can also choose the “right” reviewers (E2 and R in Fig. 4) to manipulate the decision. Unfortunately, we cannot directly test (E1)-(E3) and (R) data on rejected papers are not publicly available. Nevertheless, we can investigate the indirect evidence by comparing the ratio of Chinese-authored paper (C) and the ratio of Chinese-surnamed economists (N). We use the Web of Science top 50,000 economists to calculate $N = 8.53$ %. The average ratio of C in our sample is 15.80 %, which turns out to be even higher than N (Fig. 5). The right-tailed T test of the hypothesis of $C > N$ is significantly rejected (t-statistic = 8.3739). Therefore, there is no evidence that editors or reviewers discriminate against Chinese authors by setting a higher rejection rate or a higher quality threshold.

[Longer Publication Time] We do have data on pre-acceptance waiting time (in days) for published papers, which can be used to test if editors deprioritize Chinese-authored papers. The result of the Gamma regression of pre-acceptance review time is shown in the first column of Table 9. There is no significant difference in review time between papers authored by Chinese and non-Chinese. Thus, there is “corroborative evidence” that editors and reviewers are not the discriminators. Compared to ethnicity, more important factors to reduce review time include number of authors (division-of-labor effect), affiliation in the US, and special issue dummy.

Publishers. Similarly, publishers could prolong the production time by deprioritizing Chinese-authored papers (P in Fig. 4). The Gamma regression in the middle column of Table 9 suggests no significant difference in post-acceptance production time. Academic publishers like Elsevier, Wiley, and Springer Nature centrally manage the production process of a large number of articles from different journals. The production queue is formed automatically, so fairness is assured by the streamlined process. Therefore, publishers are innocent too.

Readers. After excluding the above “suspects”, readers are the only possible discriminators. Paper quality has been strictly assured by the editorial and reviewing processes of the top journals in our sample, so there should not be systematic quality difference in papers between Chinese-authored and non-Chinese-authored papers. Relevance-wise, the evidence shows that a higher proportion of Chinese-authored papers embark on China-related topics, while these topics attract greater citations in general (Table 2). Therefore, lower citations of Chinese-authored papers reinforce the hypothesis of ethnic discrimination [H1] and that it is pure taste-based discrimination.

6.3. *Lead follow-up*

Readers are discriminators, but this answer is crude. With the information at hand, we can follow the lead further to narrow down the discriminators in three directions, i.e., Chinese versus non-Chinese, top economists versus non-top economists, and US affiliations versus non-US

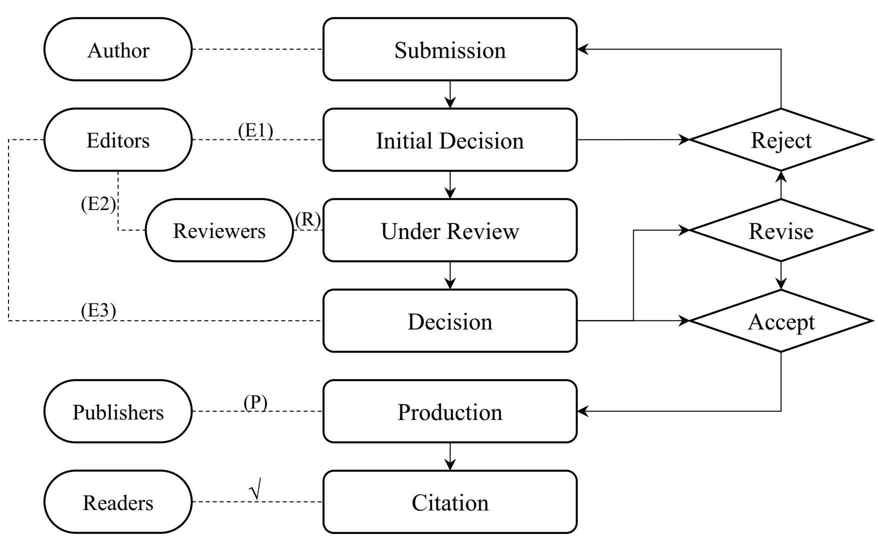

Fig. 4. Stakeholders and phases of academic publication.
Notes: Round boxes = stakeholders; square boxes = stages; diamond boxes = conditions; solid lines with arrows = directions of publication; dash lines = influences of stakeholders.

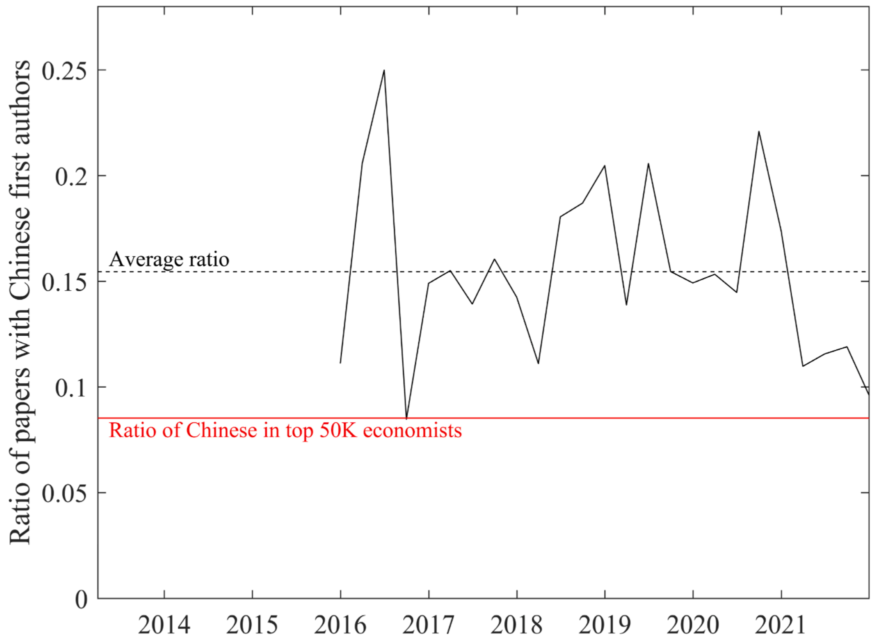

Fig. 5. Ratio of Chinese-authored papers vs. ratio of Chinese economists.

Table 9
Publication time.

	Pre-acceptance review time+	Post-acceptance production time=	Total publication time
Chinese first author (C)			
coefficient	-0.0267 (0.0190)	-0.0222 (0.0357)	-0.0195 (0.0173)
marginal effect	-16.134 (11.501)	-2.4341 (3.9189)	-13.985(12.427)
No. of authors	-0.0168*** (0.0064)	-0.0052 (0.0125)	-0.0182*** (0.0059)
Top 250 economists	-0.0256 (0.0325)	-0.0458 (0.0580)	-0.0273 (0.0295)
Affiliation in US	-0.0398*** (0.0143)	0.0366 (0.0261)	-0.0297** (0.1299)
Paper length (ln($pages$))	0.0371* (0.0213)	0.0355 (0.0385)	0.0349* (0.0194)
Open access (O)	-0.0413* (0.0235)	-0.0031 (0.0420)	-0.0429** (0.0213)
China-related topics	0.0861* (0.0486)	0.0451 (0.0890)	0.0806* (0.0441)
US-related topics	0.0614** (0.0290)	0.0081 (0.0511)	0.0517** (0.0263)
Special issue	-0.3192*** (0.0271)	0.1001** (0.0484)	-0.3194*** (0.0247)
Paper type dummies	Yes	Yes	Yes
Journal dummies	Yes	Yes	Yes
Constant	Yes	Yes	Yes
Observations	5564	5778	5554

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %. Only coefficients are reported for control variables. Standard errors in parentheses. Gamma regression is used because time (in days) is always positive.

affiliations.

[Chinese versus non-Chinese] In the literature on discrimination, it has been found that ethnic discrimination may well occur within the same ethnic group. For example, Zussman (2013) shows that both Jews and Arab second-hand car buyers discriminate against Arab sellers because the asymmetric information between buyers and sellers tends to cause statistical discrimination. In contrast, there is no asymmetric information between readers and published papers, so the discrimination is purely taste-based. According to the gravity model of citations, we can

Table 10
References of Chinese and non-Chinese authored papers.

Sample	No. of refs			No. of Chinese refs			Ratio
	Mean	S.D.	Range	Mean	S.D.	Range	
Chinese author	40.92	17.17	{0121}	7.99	6.87	{0,47}	19.52 %
Non-Chinese author	46.45	21.83	{0228}	1.87	2.90	{0,27}	4.02 %
All	45.58	21.26	{0228}	2.83	4.42	{0,47}	6.22 %

Table 11
Ratio of Chinese-authored references.

	(1)	(2)	(3)
Chinese (C)	**0.5349*** (0.0178)**		
Census-based ($\overline{C}$)		**0.5590*** (0.0229)**	
Survey-based ($\widehat{C}$)			**0.6434*** (0.0245)**
No. of authors	0.0311*** (0.0065)	0.0368*** (0.0066)	0.0325*** (0.0067)
Top 250 economists	-0.1081*** (0.0298)	-0.1196*** (0.0300)	-0.1115*** (0.0299)
Affiliation in US	**0.0647*** (0.0149)**	**0.0716*** (0.0152)**	**0.0678*** (0.0152)**
Paper age (ln$A$)	0.0557*** (0.0180)	0.0470*** (0.0183)	0.0502*** (0.0183)
Paper length (ln($pages$))	-0.0042 (0.0230)	-0.0060 (0.0236)	0.0041 (0.0232)
Open access	-0.0523** (0.0236)	-0.0754*** (0.0243)	-0.0716*** (0.0239)
China-related topics	0.5894*** (0.0567)	0.6537*** (0.0573)	0.6216*** (0.0558)
US-related topics	0.0979*** (0.0284)	0.0980*** (0.0294)	0.0990*** (0.0299)
Special issue	-0.0363 (0.0292)	-0.0418 (0.0302)	-0.0380 (0.0303)
Type dummies	Yes	Yes	Yes
Journal dummies	Yes	Yes	Yes
Constant	Yes	Yes	Yes
Observations	7467	7467	7467

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %.

predict that non-Chinese authors are more likely to discriminate against Chinese, while Chinese authors are less likely to discriminate against Chinese.

To intuitively test this prediction, we use the references of papers in our sample to extract information on who cites whom. As shown in Table 10, Chinese-authored papers account for 6.22 % in the references of all papers. This ratio is 19.52 % for Chinese-authored papers, but only 4.02 % for non-Chinese-authored papers.

To formally explain the differences in the ratio, we use the fractional regression to estimate the citation effects of Chinese authorship measures ($C, \overline{C}, \widehat{C}$) and other factors in Table 11. It is shown that the ratio of Chinese-authored references is significantly higher in Chinese-authored papers, while the ratio is significantly lower in non-Chinese-authored papers. Again, the difference is neither because of quality (assured by the top journals) nor because of relevance (controlled by topics and types), but purely taste-based. However, the differences can be argued in both ways—it is of course possible that non-Chinese authors under-cite Chinese-authored papers due to discrimination, while it is also possible that Chinese authors over-cite Chinese-authored papers due to favoritism or reverse discrimination (Kim, 2007). According to the gravity model of citations, both can happen because "cultural distance" can go either way.

[Top Economists versus non-top Economists] Another finding from Table 11 is that the top 250 economists are less likely to cite Chinese-authored papers. As a result, the voice of Chinese economists is overlooked by the most influential economists. In fact, the representation of Chinese economists in the Top 250 list is only 0.8 %, much lower than the overall ratio (8.53 %). There seems to be a glass ceiling for Chinese economists to go beyond the minority community to the mainstream stage.

[US Affiliations versus Other Affiliations] It is consistent with the gravity model of citations that Chinese authors tend to cite more Chinese-authored papers due to the shorter cultural distance (coefficients of $C, \overline{C}, \widehat{C}$ in Table 11). Nevertheless, it is interesting that authors from US affiliations also cite more Chinese-authored papers ("affiliation in US" in Table 11). One explanation is that increasingly intense interactions between the two countries since 2016 (e.g., the trade war, COVID) boost citations of Chinese-authored, China-related research. To some extent, the rising influence of China helps increase academic impacts of Chinese economists, but it also creates tensions and exacerbates discrimination against Chinese as we found in Subsection 5.2. If both Chinese authors and authors from US affiliations cite more Chinese-authored papers, then authors from non-US affiliations must be the main source of discrimination.

Admittedly, special events such as the US-China trade war (2018–2019) are likely to heighten academic interest of US researchers in China-related research. This topical relevance can contribute to the observed increase in citations of Chinese-authored papers by US-affiliated authors (Table 11). However, this possibility does not necessarily disprove the discriminatory practices by non-US-affiliated authors. To show robustness of our conclusion, we performed two further tests. First, we ran our regressions for the subsample of articles not mentioning China in titles, abstracts, or keywords. The discriminatory effect attributed to non-US affiliations remained statistically significant, suggesting that the bias is not solely driven by topic-specific citations. Second, we re-estimated the model for the subsample of articles before 2018 (before the trade war), we still observed lower citation rates for Chinese authored papers by non-US-affiliated authors.[9] This indicates that the discrimination pattern predates the special events and is not an artifact of the crisis-driven citation surge.

6.4. Case resolution

Based on the baseline results and new evidence, we conclude that non-Chinese top economists from non-US affiliations are the main source of discrimination. Nevertheless, it is also likely that Chinese authors from Chinese affiliations have been engaged in reverse discrimination. Overall, the discriminatory effect dominates the reverse discriminatory effect, and the rise of China in international politics and the world economy has influenced academic impacts of Chinese authors.

This finding can be understood through the gravity model of citations. Eq. (3) suggests that citations are more likely to occur between authors with shorter cultural distance. Arguably, non-Chinese top economists from non-US affiliations are primarily based in Europe. They have a shorter cultural distance from the West than from China, so they tend to cite papers by European or American economists, resulting in Chinese-authored papers being under-cited.

7. Conclusion

Using articles published in top economic journals, we find evidence for the hypothesis [H1] that taste-based ethnic discrimination exists in citing papers written by Chinese first authors. The hypothesis is robust under various estimation methods, specifications, causality assumptions, and measures. We note that the "control group" may also contain other discriminated-against minorities and discriminators may not spot all Chinese authors in the "treated group". Both oversampling of the control group and undersampling of the treated group make the estimated discriminatory effect a **lower bound** of the true level, which further reinforces the hypothesis. Collaborations with non-Chinese coauthors do not mitigate discrimination as long as the first author has a Chinese surname [H2A]. The dampening effect of non-Chinese affiliations is not significant [H2B]. We apply a DID method (event study) to show that the COVID pandemic further reduced the citations of Chinese authors [H3], so the academic community is not free from a stronger ethnic discrimination during the pandemic. In addition, we follow the four-step workflow of forensic analysis to narrow down the discriminators. It is found that non-Chinese top economists from non-Chinese, non-US affiliations are the main source of discrimination, but evidence also shows that Chinese authors have been engaged in reverse discrimination against non-Chinese authors. We theoretically interpret the discrimination and reverse discrimination as "distance" effects based on the gravity model of citations.

The findings fill an important gap (ethnic discrimination in academic research) in the discrimination literature. Ethnic discrimination in citations can impede knowledge diffusion and undermine research equality. The findings provide useful insights and potential impacts for researchers, editors, and publishers.

For researchers, evidence suggests that taste-based discrimination is difficult to counter solely through the efforts of the discriminated-against group. Coauthoring with non-Chinese researchers, while retaining first authorship, has been found to be ineffective [H2A]. Similarly, affiliating with institutions outside China does not appear to resolve the issue [H2B]. Pandemic-induced discrimination also remains largely beyond the control of individuals [H3]. However, it has been shown that simply revealing instances of discrimination can help reduce its occurrence in academia (Boring & Philippe, 2021). We hope that, once published, our paper will raise awareness among non-Chinese researchers of this phenomenon and contribute to the reduction of ethnic discrimination within the academic community.

For editors and publishers, we propose a recommendation regarding citation practices aimed at addressing ethnic bias in citations. Given that ethnic bias in citations is taste-based rather than statistical discrimination, it is suggested that information disclosure on ethnicity should be minimized (Laouénan & Rathelot, 2022). One potential way of doing it

[9] The subsample includes 2,753 papers published before July 6, 2018, the date when the US-China trade war started. The NB regression results of the baseline model show that the estimated coefficient, δ, is -0.1321*** (0.0520), with a p-value of 0.011.

is to use initials of authors rather than their surnames when cited. For example, instead of citing "Becker (1957)", one can use the initials of surname ("Becker") and forename ("Gary") or "BG (1957)". For coauthored papers, we can simply use the initials of surnames. This style of citation has precedent in academic literature, but it is usually reserved for well-known papers (e.g., the "HO" model in trade, the "BGG" model in macroeconomics). Another potential solution is to adopt the use of numerical citation indices (1, 2, 3, …) in place of author names, a practice that is already widespread in the natural sciences. Leading journals such as *Nature, Science*, and *Cell* have long employed this referencing style. There is no reason why similar conventions cannot be adopted in social sciences. By removing the visibility of author names from citations, this method could help reduce the potential for ethnic bias and promote a more objective evaluation of research contributions based solely on content. A more radical approach is to use codified author IDs (such as ORCID) instead of author names on published papers. The increased level of author anonymization raises the costs of author identification and taste-based discrimination, potentially shifting the focus from "who writes the paper" to "what is written". However, discrimination is not the only issue that needs to be addressed in scientific research. Such an extreme change may introduce more problems than it resolves.

We acknowledge that implementing such a change may be challenging in the short term, given the entrenched nature of citation norms and practices. Systemic changes take time and must be approached with care. By encouraging discussion around this issue, we hope to inspire more inclusive practices in scientific research and knowledge diffusion.

CRediT authorship contribution statement

Xiaoliang Yang: Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization, Project administration, Resources, Software, Validation, Visualization. **Peng Zhou:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Conceptualization, Data curation, Resources, Software, Supervision, Validation, Visualization.

Declaration of competing interest

None.

Appendix A

Table A1 presents the results with additional adjustments. Column (2) incorporates controls for research field categories based on JEL classifications. Column (3) clusters standard errors by the first author's identity.

Table A2.

Table A1
Results with research field controls and clustered standard errors.

	(1) Baseline	(2) with JEL	(3) Clustered SE
Chinese first author (C)			
coefficient (δ)	**-0.1558*** (0.0338)**	**-0.1565*** (0.0337)**	**-0.1565*** (0.0550)**
marginal effect	**-3.1103*** (0.6785)**	**-3.1292*** (0.6758)**	**-3.1292*** (0.1121)**
No. of authors			
coefficient	0.0675*** (0.0109)	0.0663*** (0.0109)	0.0663*** (0.0197)
marginal effect	1.3477*** (0.2200)	1.3247*** (0.2202)	1.3247*** (0.3964)
Top 250 economists			
coefficient	0.3171*** (0.0464)	0.3146*** (0.0462)	0.3146*** (0.0604)
marginal effect	6.3311*** (0.9365)	6.2887*** (0.9344)	6.2887*** (1.2156)
Affiliation in US			
coefficient	0.1854*** (0.0235)	0.1867*** (0.0234)	0.1867*** (0.0369)
marginal effect	3.7023*** (0.4746)	3.7315*** (0.4736)	3.7315*** (0.7495)
Paper age ($\ln A$)			
coefficient ($\tilde{\beta}$)	0.5927*** (0.0285)	0.6002*** (0.0284)	0.6002*** (0.0684)
marginal effect	11.8342*** (0.6511)	11.9986*** (0.6511)	11.9986*** (1.3990)
Paper length ($\ln(pages)$)			
coefficient	0.6762*** (0.0312)	0.6741*** (0.0312)	0.6741*** (0.0582)
marginal effect	13.5004*** (0.6764)	13.4764*** (0.6776)	13.4764*** (1.2540)
Open access			
coefficient	0.2646*** (0.0348)	0.2615*** (0.0346)	0.2615*** (0.0391)
marginal effect	5.2839*** (0.7048)	5.2280*** (0.7019)	5.2280*** (0.7802)
China-related topics			
coefficient	0.4409*** (0.0800)	0.4511*** (0.0800)	0.4511*** (0.0829)
marginal effect	10.4810*** (2.3273)	10.7951*** (2.3544)	10.7951*** (2.4415)
US-related topics			
coefficient	0.1879*** (0.0431)	0.1869*** (0.0431)	0.1869*** (0.0501)
marginal effect	3.9111*** (0.9494)	3.8911*** (0.9480)	3.8911*** (1.1082)
Special issue			
coefficient	0.3928*** (0.0443)	0.3867*** (0.0441)	0.3867*** (0.1028)
marginal effect	7.8434*** (0.9020)	7.7301*** (0.8982)	7.7301*** (2.0876)
Paper type dummies	Yes	Yes	Yes
Journal dummies	Yes	Yes	Yes
Research Field (JEL)	No	Yes	Yes
Constant	Yes	Yes	Yes
Overdispersion (α)	0.8191*** (0.0141)	0.8071*** (0.0139)	0.8071*** (0.0305)
Observations	7559	7559	7559

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %.

Table A2
Robustness tests for the effect of paper age.

	(1) Moderating Test	(2) Updated Citations	(3) Selected Subsample
Chinese first author (C)			
coefficient (δ)	-0.1592*** (0.0341)	-0.1539*** (0.0344)	-0.1649*** (0.0391)
marginal effect	-3.1770*** (0.6823)	-4.8956*** (1.0996)	-2.9061*** (0.6912)
Paper age ($\ln A$)			
coefficient ($\widetilde{\beta}$)	0.5833*** (0.0304)	0.4343*** (0.0455)	0.6315*** (0.0357)
marginal effect	11.6418*** (0.6831)	13.8150*** (1.5091)	11.1261*** (0.7127)
$C \times \ln A$			
coefficient	0.0729 (0.0807)		
marginal effect	1.4558 (1.6102)		
Control variables	Yes	Yes	Yes
Paper type dummies	Yes	Yes	Yes
Journal dummies	Yes	Yes	Yes
Constant	Yes	Yes	Yes
Overdispersion (α)	0.8190*** (0.0141)	0.8905*** (0.0142)	0.8128*** (0.0159)
Observations	7559	7559	5988

Notes: Standard errors in parentheses. Statistical significance: *** 1 %, ** 5 %, * 10 %. Column (2) reports the results based on updated citation data up to January 2025. Column (3) reports the results based on a subsample that excludes Chinese-authored papers published after 2020 and non-Chinese-authored papers published before 2018.

Data availability

Data will be made available on request.

References

Agan, A., Starr, S., 2018. Ban the box, criminal records, and racial discrimination: a field experiment. Q. J. Econ. 133 (1), 191–235.
Altonji, J.G., Elder, T.E., Taber, C.R., 2005. Selection on observed and unobserved variables: assessing the effectiveness of Catholic schools. J. Polit. Econ. 113 (1), 151–184.
Angrist, J., Azoulay, P., Ellison, G., Hill, R., Lu, S.F., 2017. Economic research evolves: fields and styles. Am. Econ. Rev. 107 (5), 293–297.
Antonovics, K., Knight, B.G., 2009. A new look at racial profiling: evidence from The Boston Police Department. Rev. Econ. Stat. 91 (1), 163–177.
Anwar, S., Bayer, P., Hjalmarsson, R., 2012. The impact of jury race in criminal trials. Q. J. Econ. 127 (2), 1017–1055.
Arnold, D., Dobbie, W., Yang, C.S., 2018. Racial bias in bail decisions. Q. J. Econ. 133 (4), 1885–1932.
Arrow, K.J. 1972. Models of job discrimination. In Racial Discrimination in Economic Life, edited by Pascal, A.H., 83–102. Lexington, MA: D.C. Heath.
Bai, Y., Kung, J.K., 2022. Surname distance and technology diffusion: the case of the adoption of maize in late imperial China. J. Econ. Growth 27 (4), 569–607.
Bar, R., Zussman, A., 2017. Customer discrimination: evidence from Israel. J. Labor. Econ. 35 (4), 1031–1059.
Becker, G.S., 1957. The Economics of Discrimination. Chicago University Press, Chicago.
Becker, G.S., 1971. The Economics of Discrimination, 2nd ed. University of Chicago Press, Chicago.
Berson, C., Laouénan, M., Valat, E., 2020. Outsourcing recruitment as a solution to prevent discrimination: a correspondence study. Labour. Econ. 64, 101838.
Bertrand, M., Mullainathan, S., 2004. Are Emily and Greg more employable than Lakisha and Jamal: a field experiment on labor market discrimination. Am. Econ. Rev. 94 (4), 991–1013.
Blackaby, D., Frank, J., 2000. Ethnic and other minority representation in UK academic economics. Econ. J. 110 (464), 293–311.
Bohren, J., Imas, A., Rosenberg, M., 2019. The dynamics of discrimination: theory and evidence. Am. Econ. Rev. 109 (10), 3395–3436.
Boring, A., Philippe, A., 2021. Reducing discrimination in the field: evidence from an awareness raising intervention targeting gender biases in student evaluations of teaching. J. Public Econ. 193, 104323.
Botelho, F., Madeira, R.A., Rangel, M.A., 2015. Racial discrimination in grading: evidence from Brazil. Am. Econ. J.: Appl. Econ. 7 (4), 37–52.
Brock, W.A., Cooley, J., Durlauf, S.N., Navarro, S., 2012. On the observational implications of taste-based discrimination in racial profiling. J. Econ. 166 (1), 66–78.
Bukstein, D., Gandelman, N., 2019. Glass ceilings in research: evidence from a national program in Uruguay. Res. Policy. 48 (6), 1550–1563.
Card, D., DellaVigna, S., 2013. Nine facts about top journals in economics. J. Econ. Lit. 51 (1), 144–161.
Card, D., DellaVigna, S., Funk, P., Iriberri, N., 2022. Gender differences in peer recognition by economists. Econometrica 90 (5), 1937–1971.
Card, D., 2014. Page limits on economics articles: evidence from two journals. J. Econ. Perspect. 28 (3), 149–167.
Carrell, S.E., Page, M.E., West, J.E., 2010. Sex and science: how professor gender perpetuates the gender gap. Q. J. Econ. 125 (3), 1101–1144.
Castillo, M., Petrie, R., Torero, M., Vesterlund, L., 2013. Gender differences in bargaining outcomes: a field experiment on discrimination. J. Public Econ. 99, 35–48.
Charles, K.K., Guryan, J., 2008. Prejudice and wages: an empirical assessment of Becker's the economics of discrimination. J. Polit. Econ. 116 (5), 773–809.
China's Ministry of Public Security, 2021. National name report, 2000. China's Ministry of Public Security. https://www.gov.cn/xinwen/2021-02/08/content_5585906.htm. accessed March 5, 2023.
Christensen, P., Timmins, C., 2022. Sorting or steering: the effects of housing discrimination on neighborhood choice. J. Polit. Econ. 130 (8), 2110–2163.
Cohn, J.B., Liu, Z., Wardlaw, M.I., 2022. Count (and count-like) data in finance. J. financ. econ. 146 (2), 529–551.
Conlon, D.E., Morgeson, F.P., McNamara, G., Wiseman, R.M., Skilton, P.F., 2006. From the editors: examining the impact and role of special issue and regular journal articles in the field of management. Acad. Manag. J. 49 (5), 857–872.
Davis, P.M., 2009. Author-choice open-access publishing in the biological and medical literature: a citation analysis. J. Am. Soc. Inf. Sci. Technol. 60 (1), 3–8.
De Paola, M., Scoppa, V., 2015. Gender discrimination and evaluators' Gender: evidence from Italian Academia. Economica 82 (325), 162–188.
Dee, T.S., 2004. Teachers, race, and student achievement in a randomized experiment. Rev. Econ. Stat. 86 (1), 195–210.
Doleac, J.L., Hansen, B., 2020. The unintended consequences of "ban the box": statistical discrimination and employment outcomes when criminal histories are hidden. J. Labor. Econ. 38 (2), 321–374.
Drivas, K., Kremmydas, D., 2020. The Matthew effect of a journal's ranking. Res. Policy. 49 (4), 103951.
Edelman, B., Luca, M., Svirsky, D., 2017. Racial discrimination in the sharing economy: evidence from a field experiment. Am. Econ. J.: Appl. Econ. 9 (2), 1–22.
Einav, L., Yariv, L., 2006. What's in a surname? The effects of surname initials on academic success. J. Econ. Perspect. 20 (1), 175–187.
Ewens, M., Tomlin, B., Wang, L.C., 2014. Statistical discrimination or prejudice? A large sample field experiment. Rev. Econ. Stat. 96 (1), 119–134.
Fershtman, C., Gneezy, U., 2001. Discrimination in a segmented society: an experimental approach. Q. J. Econ. 116 (1), 351–377.
Foreman-Peck, J., Zhou, P., 2015. Firm-level evidence for the language investment effect on SME exporters. Scott. J. Polit. Econ. 62 (4), 351–377.
Fryer, R.G., 2019. An empirical analysis of racial differences in police use of force. J. Polit. Econ. 127 (3), 1210–1261.
Furman, J.L., Stern, S., 2011. Climbing atop the shoulders of giants: the impact of institutions on cumulative research. Am. Econ. Rev. 101 (5), 1933–1963.
Galasso, A., Schankerman, M., 2015. Patents and cumulative innovation: causal evidence from the courts. Q. J. Econ. 130 (1), 317–370.
Gallen, Y., Wasserman, M., 2023. Does information affect homophily? J. Public Econ. 222, 104876.
Gaulé, P., Maystre, N., 2011. Getting cited: does open access help? Res. Policy. 40 (10), 1332–1338.
Giuliano, L., Levine, D.I., Leonard, J., 2009. Manager race and the race of new hires. J. Labor. Econ. 27 (4), 589–631.
Glover, D., Pallais, A., Pariente, W., 2017. Discrimination as A self-fulfilling prophecy: evidence from French grocery stores. Q. J. Econ. 132 (3), 1219–1260.
Goncalves, F., Mello, S., 2021. A few bad apples? Racial bias in policing. Am. Econ. Rev. 111 (5), 1406–1441.
Gould, E.D., Klor, E.F., 2016. The long-run effect of 9/11: terrorism, backlash, and the assimilation of Muslim immigrants in the West. Econ. J. 126 (597), 2064–2114.
Grogger, J., Ridgeway, G., 2006. Testing for racial profiling in traffic stops from behind a veil of darkness. J. Am. Stat. Assoc. 101 (475), 878–887.
Guiso, L., Sapienza, P., Zingales, L., 2009. Cultural biases in economic exchange? Q. J. Econ. 124 (3), 1095–1131.
Haan, P., Peichl, A., Schrenker, A., Weizsäcker, G., Winter, J., 2022. Expectation management of policy leaders: evidence from COVID-19. J. Public Econ. 209, 104659.

Hamermesh, D.S., 2018. Citations in economics: measurement, uses, and impacts. J. Econ. Lit. 56 (1), 115–156.

Heckman, J.J., 1998. Detecting discrimination. J. Econ. Perspect. 12 (2), 101–116.

Hedegaard, M.S., Tyran, J., 2018. The price of prejudice. Am. Econ. J.: Appl. Econ. 10 (1), 40–63.

Horrace, W.C., Rohlin, S.M., 2016. How dark is dark? Bright lights, big City, racial profiling. Rev. Econ. Stat. 98 (2), 226–232.

Jansson, J., Tyrefors, B., 2022. Grading bias and the leaky pipeline in economics: evidence from Stockholm University. Labour. Econ. 78, 102212.

Kahn, L.M., Sherer, P.D., 1988. Racial differences in professional basketball players' Compensation. J. Labor. Econ. 6 (1), 40–61.

Keng, S.H., 2020. Gender bias and statistical discrimination against female instructors in student evaluations of teaching. Labour. Econ. 66, 101889.

Kerr, W.R., 2008. Ethnic scientific communities and international technology diffusion. Rev. Econ. Stat. 90 (3), 518–537.

Kerr, W.R., Lincoln, W.F., 2010. The supply side of innovation: H-1B visa reforms and U. S. Ethnic invention. J. Labor. Econ. 28 (3), 473–508.

Kim, K., 2007. Favoritism and reverse discrimination. Eur. Econ. Rev. 51 (1), 101–123.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., Mullainathan, S., 2018. Human decisions and machine predictions. Q. J. Econ. 133 (1), 237–293.

Knowles, J., Persico, N., Todd, P., 2001. Racial bias in motor vehicle searches: theory and evidence. J. Polit. Econ. 109 (1), 203–229.

Krawczyk, M., Smyk, M., 2016. Author's gender affects rating of academic articles: evidence from an incentivized, deception-free laboratory experiment. Eur. Econ. Rev. 90, 326.

Lang, K., Spitzer, A.K., 2020. Race discrimination: an economic perspective. J. Econ. Perspect. 34 (2), 68–89.

Lang, K., Lehmann, J.K., 2012. Racial discrimination in the labor market: theory and empirics. J. Econ. Lit. 50 (4), 959–1006.

Laouénan, M., Rathelot, R., 2022. Can information reduce ethnic discrimination? Evidence from Airbnb. Am. Econ. J.: Appl. Econ. 14 (1), 107–132.

Lee, L., 1978. Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables. Int. Econ. Rev. (Philadelphia) 19 (2), 415–433.

Leonard, J.S., Levine, D.I., Giuliano, L., 2010. Customer discrimination. Rev. Econ. Stat. 92 (3), 670–678.

Li, W., Yi, J., 2020. Alphabetical author order, intellectual collaboration and high-skilled migration. Econ. J. 131 (635), 1250–1268.

Liu, A., Lu, C., Wang, Z., 2021. Does cultural distance hinder exports? A comparative study of China and the United States. Econ. Model. 105, 105668.

Liu, F., Rahwan, T., AlShebli, B., 2023. Non-White scientists appear on fewer editorial boards, spend more time under review, and receive fewer citations. Proc. Natl. Acad. Sci. 120 (13), e2215324120.

Liu, Y., Chen, L., Yuan, Y., Chen, J., 2012. A study of surnames in China through isonymy. Am. J. Phys. Anthropol. 148 (3), 341–350.

Lu, R., Sheng, S.Y., 2022. How racial animus forms and spreads: evidence from the Coronavirus Pandemic. J. Econ. Behav. Organ. 200, 82–98.

Martinez de Lafuente, D., 2021. Cultural assimilation and ethnic discrimination: an audit study with schools. Labour. Econ. 72, 102058.

McCabe, M.J., Snyder, C.M., 2015. Does online availability increase citations? Theory and evidence from A panel of economics and Business journals. Rev. Econ. Stat. 97 (1), 144–165.

McConnell, B., Rasul, I., 2021. Contagious animosity in the field: evidence from the Federal Criminal Justice System. J. Labor. Econ. 39 (3), 739–785.

Medoff, M.H., 2003. Collaboration and the quality of economics research. Labour. Econ. 10 (5), 597–608.

Neal, D.A., Johnson, W.R., 1996. The role of premarket factors in black-white wage differences. J. Polit. Econ. 104 (5), 869–895.

Neumark, D., 2012. Detecting discrimination in audit and correspondence studies. J. Hum. Resour. 47 (4), 1128–1157.

Nunn, N., Wantchekon, Leonard., 2011. The slave trade and the origins of mistrust in Africa. Am. Econ. Rev. 101 (7), 3221–3252.

Phelps, E.S., 1972. The statistical theory of racism and sexism. Am. Econ. Rev. 62 (4), 659–661.

Pierre-Philippe, C., Bruno, D., Morgane, L., Alain, T., 2016. Customer discrimination and employment outcomes: theory and evidence from the French labor market. J. Labor. Econ. 34 (1), 107–160.

Qiu, S., Steinwender, C., Azoulay, P., 2023. Who stands on the shoulders of Chinese (Scientific) giants? Evidence from chemistry. NBER Working Paper Series WP30772.

Qiu, S., Steinwender, C., Azoulay, P., 2024. Paper Tiger? Chinese science and home bias in citations. NBER Working Paper Series WP32468.

Ray, D., Robson, A., 2018. Certified random: A new order for coauthorship. Am. Econ. Rev. 108 (2), 489–520.

Rehavi, M.M., Starr, S.B., 2014. Racial disparity in federal criminal sentences. J. Polit. Econ. 122 (6), 1320–1354.

Rienzo, C., 2024. Trick or treat? The Brexit effect on immigrants' Mental health in the United Kingdom. Eur. Econ. Rev. 162, 104660.

Rubin, A., Rubin, E., 2021. Systematic bias in the progress of research. J. Polit. Econ. 129 (9), 2666–2719.

Rubinstein, Y., Brenner, D., 2014. Pride and prejudice: using ethnic-sounding names and inter-ethnic marriages to identify labour market discrimination. Rev. Econ. Stud. 81, 389–425, 1(286).

Small, M.L., Pager, D., 2020. Sociological perspectives on racial discrimination. J. Econ. Perspect. 34 (2), 49–67.

Spenkuch, J.L., Montagnes, B.P., Magleby, D.B., 2018. Backward induction in the wild? Evidence from sequential voting in the US Senate. Am. Econ. Rev. 108 (7), 1971–2013.

Stewart, M.B., 1983. Racial discrimination and occupational attainment in Britain. Econ. J. 93 (371), 521–541.

Tootell, G.M.B., 1996. Redlining in Boston: do mortgage lenders discriminate against neighborhoods? Q. J. Econ. 111 (4), 1049–1079.

Utar, H., 2018. Workers beneath the floodgates: low-wage import competition and workers' adjustment. Rev. Econ. Stat. 100 (4), 631–647.

Yinger, J., 1986. Measuring racial discrimination with fair housing audits: caught in the act. Am. Econ. Rev. 76 (5), 881–893.

Zhou, B., Zhang, Y., Zhou, P., 2021. Multilateral political effects on outbound tourism. Ann. Tour. Res. 88, 103184.

Zhu, Y., 2009. Hundred Family Surnames, Chinese Edition. China translation and publishing Corporation, Beijing.

Zussman, A., 2013. Ethnic discrimination: lessons from the Israeli online market for used cars. Econ. J. 123 (572), F433–F468.