Lack of peer reviewer diversity advantages scientists from wealthier countries

James M. Zumel Dumlao and Misha Teplitskiy*
University of Michigan School of Information

February 26, 2025

Authors from low- and middle-income countries often have lower acceptance rates in academic journals, leading to their ideas receiving less attention. One hypothesized contributor to this disparity is the limited geographical diversity of peer reviewers. If reviewers favor submissions from their own country (*i.e.* homophily), and authors from certain countries are disproportionately reviewed by same-country reviewers, this creates a "geographical representation bias" favoring those authors. Using administrative data from the Institute of Physics Publishing (IOP), encompassing metadata on 204,718 submissions to 60 STEM journals between 2018-2022, we find strong evidence of this bias. Comparing different reviewers of the same manuscript, we find that those from the same country as the corresponding author are 4.78 percentage points more likely to review positively, and authors from countries well-represented in the reviewer pool (e.g., USA, China, India) are 8-9 times more likely to be evaluated by same-country reviewers. An instrumental variables analysis using an exogenous policy change that created variation in which submissions were anonymized shows that anonymizing submissions does not significantly reduce homophily. These findings provide support for efforts to diversify reviewer pools, and suggest that there are no "quick fixes" for geographical representation bias.

* To whom correspondence should be addressed: Misha Teplitskiy, tepl@umich.edu

Introduction

Authors from low- and middle-income countries often have lower acceptance rates in academic journals across various fields (Peng, Lakhani, and Teplitskiy 2021; Smith et al. 2023), leading to their ideas receiving less attention from the scientific community. (Gomez, Herman, and Parigi 2022; Qiu, Steinwender, and Azoulay 2023). The epistemic consequences of these publishing disparities may be substantial. Scientists' choices of which problems to research are frequently correlated with their identities (Kozlowski et al. 2022) and, in particular, their geographic location (Fry 2023). Consequently, geographic barriers to publishing may perpetuate blindspots in the visible corpus of knowledge, shape what problems are viewed as pursuit-worthy, and make it harder for peripheral scientists to achieve priority.

Some causes of global disparities in publishing are well understood. First, authors from lower-income countries may have less access to the equipment, networks, and other resources needed for cutting-edge work in resource-intensive fields (Burns and Fox 2017; Collyer 2018) and face language barriers (Amano et al. 2023; NHB Editorial Board 2023). Consequently, their papers may have lower quality or fit for English-language outlets. Second, authors from underrepresented demographic groups may study topics that are viewed less favorably by evaluators in overrepresented groups (Chen et al. 2022; Hoppe et al. 2019) and such topic mismatch may affect geographically underrepresented scientists as well. Third, some research finds that peer reviewers and editors show bias toward high-status authors and work from wealthier countries (Fox, Meyer, and Aimé 2023; Fox and Paine 2019; Harris et al. 2017; Kowal et al. 2022; Nielsen and Andersen 2021; Sun, Barry Danfa, and Teplitskiy 2022).

Whereas these causes are related to individual-level author characteristics like resources, status, and country, here we focus on a different and fundamentally relational cause related to peer reviewer diversity. Specifically, we investigate geographical concordance between authors and reviewers. The concordance or mismatch may result in what we call "geographical representation bias" if two conditions are met. The first is *homophily*—reviewers favoring work from authors that are similar or close to them in some dimension. We focus on homophily by country, sometimes referred to as "home bias" (Qiu, Steinwender, and Azoulay 2024; Wright, Koning, and Khanna 2023). Reviewers from the same country as authors, *i.e*. same-country reviewers (SCRs), may exhibit favoritism through mechanisms including having more information about the submission or its authors, similar tastes to the authors, and identity-based biases like nationalism (see Mechanisms in Background Section). For instance, reviewers may be more likely to know same-country authors better than other authors, and give them the benefit of the doubt when the claims in a manuscript are incompletely described (Heesen and Romeijn 2019). The second condition is *differential access to homophily*—systematic differences in how likely authors are to be evaluated by homophilous reviewers. Both conditions are necessary for a lack of evaluator diversity to result in the geographical representation bias. If reviewers exhibit homophily (condition 1 is met) but all authors have equal access to SCRs (condition 2 not met),

then no group is disadvantaged relative to another. And a lack of demographic diversity among reviewers (condition 2 is met) does not by itself create a representation bias if they review similarly (condition 1 not met).

Geographical representation bias represents a Matthew Effect (Merton 1968) at the structural, rather than individual, level. Peer reviewers for particular outlets are likely selected disproportionately from those publishing in these outlets. Consequently, authors from these wealthier countries who have historically produced more research will have greater representation in the reviewer pool and benefit from homophily more often. In this way, past success improves future success.

Despite widespread belief that representation bias affects science, existing evidence for it is mixed and inconclusive. The key challenge has been that peer review data are typically sensitive and proprietary. In cases where data were available, some analyses found suggestive evidence but could not control for relevant confounding factors like submission quality (Gaston and Smart 2018; Link 1998; Murray et al. 2018). Others found null evidence but lacked sufficient statistical power (Tomkins, Zhang, and Heavlin 2017). Without strong evidence that reviewers review differently depending on their concordance with authors, it is possible that diversifying reviewers, as many have called for, would do little for the publication prospects of underrepresented authors. And as difficult as it has been to establish geographical representation bias in the first place, to our knowledge no studies have examined ways to reduce this bias.

Here, we overcome the confounding and statistical power challenges of prior work using large-scale administrative peer review data on 204,718 submissions to 60 English-language journals from January 2018 to December 2022 from the Institute of Physics Publishing (IOP). Submissions were matched to countries using the self-reported country of the corresponding author. Most submissions were reviewed single-anonymously, where authors' identities were visible to reviewers. The granularity of the data enabled us to fully control for manuscript quality and other attributes using manuscript fixed effects, *i.e.*, effectively comparing different reviewers of the same manuscript. We found strong, and arguably conclusive, evidence of country homophily: reviewers from the same country as authors (same-country reviewers or SCRs) gave positive recommendations on manuscripts about 4.78 p.p. more often than non-SCRs on the same manuscript. Authors from low- and lower-middle-income countries were less than half as likely as those from high-income countries to be assigned to SCRs, proportionate to their representation in the reviewer pool. Taken together, we conclude that representation bias is present in our data, even when fully controlling for differences in manuscript quality and other attributes.

How can institutions remedy geographical representation bias? Since there are two necessary conditions for the bias, a remedy could work by equalizing access to homophily (Garisto 2022; Mohanti et al. 2023; Murray et al. 2018; Smith et al. 2023), which would in effect spread homophily out more evenly. An alternative method is to reduce country homophily in the first

place. Numerous studies have shown that double-anonymization (*i.e.,* hiding authors' identities from reviewers) is effective at reducing a variety of reviewer biases (Blank 1991; Fox et al. 2023; Fox and Paine 2019; Harris et al. 2017; Kowal et al. 2022; Nielsen and Andersen 2021). However, to our knowledge no study has conclusively shown whether such policies reduce country homophily or considered them as a remedy for low reviewer diversity. The effectiveness of anonymization is important not only practically but conceptually. Double-anonymization should affect evaluations only through identity-related mechanisms, such as a nationalism bias. To test for the effect of anonymization on homophily we leverage a policy change among IOP journals that enabled authors to select into double-anonymous peer review (Banks 2020). The rollout of the anonymization policy was staggered across journals in an arbitrary order, making the policy an attractive instrumental variable for anonymization. In the instrumental variables analysis, the effect of anonymization on homophily was substantively small and statistically non-significant. In other words, hiding authors' identities (including geographic information) did little to reduce country homophily.

Existing evidence and mechanisms

Country homophily

Existing evidence of country homophily

Country homophily – evaluators favoring products, people, or ideas from their own country – is well-studied outside of science (McPherson, Smith-Lovin, and Cook 2001; Rivera and Owens 2021; Snellman and Solal 2023). For example, Wright et al. (Wright et al. 2023) find that judges of global startups are systematically more likely to recommend startups from their home region and provide evidence that this preference is due to information differences rather than differences in taste or evaluator biases. Verlegh (Verlegh 2007) discusses several studies finding that consumers prefer domestic over foreign products, arguing that the homophily stems from concerns for the domestic economy and an in-group identity bias.

In science, Rubin et al. (Rubin, Rubin, and Segal 2023) provide some evidence that editors from civil-law European countries are more favorable toward work from their own country, while those from the common-law countries are not. They attribute this pattern to a difference in individuals' orientation to institutions in these regions. It is also suggested that the European editors' same-country preference is not merit-based, since those same-country manuscripts they publish receive fewer citations than their common-law counterparts.

Our closest intellectual antecedents are (Link 1998), (Murray et al. 2018), (Gaston and Smart 2018), and (Tomkins et al. 2017). Link explored whether peer reviewers favor US vs. non-US submissions to a medical journal. She found that while both USA and non-USA reviewers favored submissions from the USA over those from non-USA, USA reviewers favored them more strongly. Murray *et al*. studied reviewer-author gender and country homophily in a setting

where reviewers evaluate manuscripts as a team. They found higher acceptance rates for manuscripts where at least one member of the reviewer team was from the same country as the last or corresponding author. Gaston and Smart analyzed journals in two disciplines and found that reviewers from the same region as the corresponding author are more likely to accept invitations to review. They also show that manuscripts with reviewers and authors from the same region have higher acceptance rates. Higher acceptance may be related to positive reviews specifically from same-country reviewers (SCRs), but neither Murray *et al*. nor Gaston and Smart confirm whether this is the case. While the evidence is suggestive, it is inconclusive because of two plausible confounders—differences in submission quality and differences in reviewers' standards. Manuscripts of a higher quality may come from countries with more reviewers, creating an association between SCRs and positivity. Furthermore, reviewers' standards of evaluation may differ systematically by country.

Tomkins *et al*. conduct an experiment on anonymizing submissions to a computer science conference. They find that, controlling for manuscript quality using scores given by double-anonymous reviewers, there is no statistically significant association of SCRs with positivity. However, only 146 submissions in their data had SCRs, so the null effect could have been due to low statistical power.

In sum, non-science contexts have shown country homophily, while in science the evidence is mixed and inconclusive. Nevertheless, several mechanisms may drive homophily, which we consider next.

Mechanisms

Country homophily might operate via three mechanisms: information, tastes, and biases. We conceptualize information and tastes as mechanisms related to the content of what is being evaluated (manuscripts), whereas biases are related to the identities of evaluators and creators of the evaluated objects (reviewers and authors). Therefore, hiding authors' identities should only affect reviewer homophily through the biases mechanism.

By "information" we refer to evaluators' private knowledge of the work and authors they review. Reviewers may have additional information about research from their own country for several reasons. They may share professional connections with the authors, such as professional societies and co-authors. The lower travel costs associated with geographic proximity may also increase exposure to authors' work at seminars and conferences. They may also work on problems or use methods that are geographically concentrated. The additional information reviewers may have about same-country research can translate into more favorable reviews, theoretically, due to Bayesian updating—reviewers make assessments of ideas in a manuscript by updating their prior assessments of those ideas, based on information outside the manuscript. If additional information is associated with holding an even slightly positive prior, the updated assessment will tend to be more positive (Heesen and Romeijn 2019). Empirically, several studies have

found expertise to be associated with more positive assessments (Li 2017; Zinovyeva and Bagues 2015), although in settings without self-selection into reviewing, the relationship may be reversed (Boudreau et al. 2016).

By "tastes" we refer to evaluators' partiality toward the substance and style of a work. A reviewer may have particular tastes regarding research topics or methods, or even writing style. The existence of cultural zones and local tastes is often discussed outside of science (Chon, Barnett, and Choi 2003; Inglehart and Carballo 1997). Location determines the resources available in the production of knowledge (Carlino and Kerr 2015) and lay community perceptions of innovation (Ozgun and Broekel 2021), which gives reason to suspect that geography is a factor in judging science. However, the universal and disinterested character of science makes the notion of geographically clustered tastes controversial in the evaluation of supposedly objective research (Merton 1973). While evidence on regional tastes among academics is sparse, at least one account provides testimony that research tastes cluster by location (Collyer 2018). SCRs may have research tastes similar to those of the submitting authors, resulting in systematically positive reviews from the former.

By "biases," we refer to evaluators' explicit or implicit affinity toward particular researchers based on their identity. These biases may be cognitive and unconscious, or they may be strategic and conscious. For example, a reviewer with a conflict of interest may be strategically biased in their evaluation if they are aware of an author's identity. Randomly hiding authors' identities is a common strategy for uncovering these biases. Several studies have detected biases among editors and reviewers along social dimensions like prestige (Blank 1991; Nielsen et al. 2021; Tomkins et al. 2017) and geography (Fox et al. 2023; Harris et al. 2017; Kowal et al. 2022; Link 1998). These biases are often conceptualized as a hierarchy with dominant and marginalized groups where the dominant group is reliably favored by all types of reviewers (*e.g.,* reviewers all prefer work from wealthy countries). Instead, we focus on how reviewers might be systematically more positive based on their relationship to authors. As noted above, authors of one country may have more professional ties with SCRs than non-SCRs, inducing conflicts of interest. Or perhaps seeing an author's country triggers nationalistic biases in SCRs and not in non-SCRs. In either case, biases could explain SCR positivity.

Differential access to same-country reviewers

Authors from different countries are likely to be assigned to SCRs at different rates. First, much discussion surrounding the diversification of editors and reviewers centers around homophily in the formation of professional networks. Editors, who are typically from a small set of wealthier countries (Altman and Cohen 2021; Liu, Rahwan, and AlShebli 2023; Smith et al. 2023), may rely on their professional networks to fill editorial boards and find reviewers (Brogaard, Engelberg, and Parsons 2014; Murray et al. 2018). Consequently, the pool of evaluators may be disproportionately composed of researchers based in those countries. This result may occur entirely unintentionally. If a journal's pool of reviewers is composed primarily of individuals

who have submitted and published in the journal in the past, that pool will overrepresent countries submitting the most work, such as China, the United States of America, and India. Consequently, if editors assign reviewers by choosing them from the existing pool and based only on expertise – ostensibly the most fair way – authors from poorly represented countries will have lower chances of being assigned to SCRs. To summarize, there is increasing evidence that the geographical distributions of those who submit manuscripts to academic journals and those who evaluate them are diverging. Attention has focused primarily on editors rather than reviewers, although it is implied that reviewer demographics will correlate with editor demographics. The present study focuses on reviewers directly.

Empirical setting

Partnering with the Institute of Physics Publishing (IOP), which is based in the United Kingdom, we use metadata on all submissions and reviews from January 2018 to December 2022. The metadata comprise 63 English-language journals in the physical sciences (3 journals were ultimately excluded from the analytic sample due to insufficient complete observations). The data did not include the content of the submissions or the peer reviews. The analysis was undertaken with a data use agreement and approved by the University of Michigan Institutional Review Board, Protocol #HUM00194927.

Authors and reviewers from roughly 150 countries are represented in the analytic sample. Thus, the setting for this study is truly global. Tables S1 and S2 display summary statistics for variables used in the model specifications for the data on realized reviews and invited reviewers, respectively. 30.25% of the 787,055 invitations to review resulted in a realized review. 113,232 of 204,718 (55.31%) submissions received at least one review (many were desk-rejected). 70.40% of 238,097 first-round reviews were positive recommendations, *i.e.,* accept or revise. The analytical sample is restricted to submissions that underwent external peer review, *i.e.,* not desk rejected. Of the invited and realized reviewers in the first review round, 11.94% and 14.67% were SCRs, respectively. 24.49% of submitted manuscripts involved at least one SCR.

Figure S1 shows the number of manuscripts and journals in the analytic sample by year. The number of journals increased each year until peaking in 2021 with 60 journals. Country-level distributions for number of submissions, number of unique realized reviewers, acceptance rates, and reviewer positivity and response rates are available in Figure S2. The top three countries by both number of submissions and unique reviewers were China, India, and the United States of America. Acceptance rates were highest for countries in Europe, Northern America, and Oceania. Of the 204,718 submissions in the raw data, 68,905 (33.66%) were from high-income countries (HIC), 84,222 (41.14%) were from upper-middle-income countries (UMIC), and 51,591 (25.20%) were from low- and lower-middle income countries (LLMIC) (see *Income Categories* in Measures and Methods). Figure S3 aggregates raw acceptance rates by the country

income groups. Submissions from HIC authors are accepted more than twice as often as those from LLMIC authors, corroborating work finding large geographic disparities in acceptances.

Measures and Methods

Measures

Review Positivity

Different journals gave reviewers slightly different options for recommendations, resulting in 36 distinct options. Following previous literature (Teplitskiy et al. 2022), we recoded these options into a binary indicator *Review Positivity* where 0 indicates Reject and 1 indicates Accept or Revise-and-resubmit. See "Positive and Not-Postive Codes for Reviewer Response Categories" in SI Appendix for the exact mapping. Following previous literature, our analyses of reviewer Positivity focus on only the first round of review. This is because in second and subsequent rounds reviewers may influence each other, which would violate assumptions of independence in our statistical analyses. Additionally, the objective of the reviews may shift, for example from filtering to improving a submission likely to be accepted.

Income Categories

Author and reviewer self-reported countries were aggregated into three income categories based on GDP per capita according to the 2023 fiscal year World Bank Country and Lending Groups. The groups are *High-income countries* (HIC, *e.g.,* USA), *Upper-middle-income countries* (UMIC, *e.g.,* China), and *Low- and Lower-middle-income countries* (LLMIC, *e.g.,* India). The full list of countries is displayed in "Countries by Income Group" in the SI Appendix.

Country of submissions

To assign submissions, which are mostly team-authored, to a single country we use the self-reported country of the corresponding author. In "Submission Country Robustness Checks" in the SI Appendix, we consider the sensitivity of our results to two other plausible operationalizations. First, we rerun the analysis after defining SCR as a match between reviewer and lead author (as opposed to corresponding author). Second, we restrict our original analysis to submissions with authors all from the same country or region. Regions are defined according to the UN geoscheme as shown in "Countries by Region" in the SI Appendix. We find that estimates of SCR positivity (Tables S7 and S8) and access to SCRs (Table S10) in these robustness checks are highly similar to those in our main analysis.

Methods

For simplicity and to enable estimation with the rich set of fixed effects, we use linear probability models with cluster-robust standard errors. In the specifications included below, manuscripts are indexed by *i,* reviewers by *j*, and time by *t*. Fixed effects are denoted as journal γ, manuscript α, reviewer invite or submission year-month τ, reviewer ρ, and ε is the error term. For regressions without manuscript fixed effects we include a control for team size, as prior work has shown it to be highly predictive of positive reviews (Teplitskiy et al. 2022; Thelwall et al. 2023). Standard errors are clustered at the level of the fixed effects in each case.

Establishing homophily using linear models with manuscript and reviewer fixed effects

A key challenge in prior work was potential confounding of the association between reviewers' geography and positivity by submission quality or other characteristics. To address this challenge, we rely on *manuscript* fixed effects. This research design effectively compares reviews from different reviewers of the same manuscript, accounting for any differences in manuscript quality. Additionally, we use *reviewer* fixed effects to account for potentially idiosyncratic reviewing standards.

Equations 1 and 2 below are used to estimate SCR Positivity (Table S3). Equation 1 does not include *manuscript* fixed effects, and therefore does not control fully for manuscript quality and other unobserved characteristics. It does include journal and reviewer fixed effects, and controls for logged team size, anonymization status, and corresponding author country. This specification is used as a baseline for comparison with prior work and our preferred specification – Equation 2. Equation 2 includes manuscript and reviewer fixed effects. Manuscript-level fixed effects absorb the effects of characteristics that do not vary across reviewers of a manuscript, such as team size, anonymization status, corresponding author country. Estimates from Equation 2 are reported as our preferred estimate of reviewers' country homophily. Additionally, we use similar specifications to measure the association between SCR status and whether reviewers accept the invitation to review (Table S3).

$$Positivity_{ij} = \gamma_i + \rho_j + \beta_1 SCR_{ij} + \beta_2 ln(TeamSize)_i + \beta_3 Anon_i + \beta_4 AuthCountry_i + \varepsilon_{ij} \quad (1)$$

$$Positivity_{ij} = \alpha_i + \rho_j + \beta_1 SCR_{ij} + \varepsilon_{ij} \quad (2)$$

Furthermore, we explore whether country homophily is exhibited by all or just some groups of reviewers with a secondary analysis that interacts SCR status with corresponding author country income group. Equation 2A below adds an interaction term to Equation 2 to estimate the heterogeneity of SCR positivity across country income groups. Country income categories were coded via dummy coding with high-income countries (HIC) as the reference category, and all

categories were included in a single regression. The results of this subanalysis are in Table S4, and the total effects (main plus interaction) are presented in Figure 1.

$$Positivity_{ij} = \alpha_i + \rho_j + \beta_1 SCR_{ij} + \beta_2 SCR_{ij} * IncomeCat_i + \varepsilon_{ij} \quad (2A)$$

Establishing differential access to homophily using linear models

We operationalize an author's "access to country homophily" or "SCR access" as the probability that a reviewer of their paper is an SCR. Equations 3A and 3B below are specifications that estimate geographic differences in invitation rates and SCR access. We estimate the likelihood of this outcome by corresponding author country (Fig. 2A) or country income group (Figs. 2B and S4, Table S9), expressed as categorical variables, with journal fixed effects and controls for demeaned log team size and anonymization status. This specification was used for the set of invited reviewers in the case of invitation rates (Fig. S4) and of realized reviewers in the case of access to SCRs (Fig. 2). Statistically significant differences among the geographic categories would indicate an uneven distribution of access to country homophily.

$$SCR_{ij} = \gamma_i + \beta_1 AuthCountry_i + \beta_2 ln(TeamSize)_i + \beta_3 Anon_i + \varepsilon_{ij} \quad (3A)$$

$$SCR_{ij} = \gamma_i + \beta_1 IncomeCat_i + \beta_2 ln(TeamSize)_i + \beta_3 Anon_i + \varepsilon_{ij} \quad (3B)$$

Establishing the causal effect of double-anonymization using instrumental variables

Physical sciences journals have historically been single-anonymous. In 2020 IOP announced a shift to offer authors double-anonymized peer review at journals in their portfolio (Banks 2020). A handful of IOP journals were already offering double-anonymization at the time of the announcement. The policy encouraged but did not force authors to make their submissions anonymous. In our dataset, 8,742 (7.72%) submitters selected the double-anonymous option. In consultation with the last author (MT), IOP staff made the anonymization policy available across its journals in an arbitrary order (*i.e.,* unrelated to author and reviewer characteristics) on different days in almost all cases. In a few cases, journals were rolled out at similar times if submitted manuscripts were commonly transferred across them. Transferring journals were grouped so that, if a manuscript was submitted for double-anonymous review and was not accepted at one journal, authors could be recommended to submit at a similar journal that also offered double-anonymous review.

To identify the effect of anonymization, it is tempting to compare evaluations across anonymized and non-anonymized submissions. Equation 4 below makes this comparison between subsets of SCRs, while still controlling for the rich set of fixed effects used earlier, particularly manuscript fixed effects.

$$Positivity_{ij} = \alpha_i + \rho_j + \beta_1 SCR_{ij} + \beta_2 SCR_{ij} * Anon_i + \varepsilon_{ij} \quad (4)$$

However, this naive comparison, presented in the second column of Table S6 ("OLS"), may yield a biased estimate of the effect because certain types of authors may have systematically chosen or avoided anonymizing.

To overcome the endogeneity challenge, we leverage the staggered and quasi-random anonymization option rollout and instrumental variables (IV). IV accounts for the self-selection problem in *actual* anonymization (the endogenous treatment variable) by using the quasi-random variation generated in it by an auxiliary, non-self-selected variable (the instrument), which in our case is the anonymization option. Intuitively, IV estimates the effect of being offered the anonymization option (intention-to-treat effect, or ITT) and scales it by the proportion of authors who choose to anonymize when given the option, thereby isolating the causal effect of anonymization among those who comply.

We restricted our analysis to papers that received reviews (55.72% of all submissions), *i.e.,* editors did not desk-reject. Although this restriction means that our analysis is conditioning on a post-treatment outcome, it is unlikely to bias our results because neither the instrument nor the treatment affected the information *editors* saw, and therefore should not have affected their desk-rejection decisions. Consistent with this argument, we found that the instrument did not affect desk-rejection probability (see Table S5).

There are several other assumptions required for IV estimates to be valid in this setting (see "Instrumental Variable Approach" in SI Appendix). The most critical is the exclusion restriction, which states that the instrument (Anonymization option) should influence the outcome (review Positivity) only through the treatment (Anonymization), and not through other channels, such as influencing authors' decisions to submit or how editors desk reject or assign reviewers. This assumption cannot be validated directly. Equation 5 below displays the specification used to test for some threats to the exclusion restriction as discussed in the SI Appendix. This specification examines the effect of anonymization option availability on four outcomes separately and includes fixed effects for journal γ and submission year-month τ. The first three outcomes pertain to the submissions that went to review: logged team size, probability of a submission being reviewed by at least one SCR, and the total number of realized reviewers. The fourth outcome is the probability of not being desk rejected (*i.e.,* going to review), which includes submissions not included in the rest of our analyses. The results from these validity tests are presented in Table S5.

$$Outcome_{ijt} = \gamma_i + \tau_t + \beta_1 PolicyAvail_{it} + \varepsilon_{ij} \quad (5)$$

If policy availability is a valid instrument for anonymization status, then policy availability interacted with SCR status is a valid instrument for anonymization interacted with SCR status (Balli and Sørensen 2013), which is our effect of interest. Thus, our instrumental variables

approach addresses two endogenous variables (i.e., anonymization and anonymization interacted with SCR status) with two instruments (i.e., policy availability and policy availability interacted with SCR status). This is implemented in a two-stage least squares (2SLS) procedure as described in Equations 6A, 6B and 7 below. Since anonymization status does not vary within-manuscript, comparisons must be made across manuscripts, thus precluding the use of manuscript fixed effects in this analysis. Equations 6A and 6B are the first stages, where each of the two endogenous variables—anonymization status and anonymization interacted with SCR status—are separately regressed on two instruments (*i.e.,* policy availability and the interaction between policy availability and same-country status), journal and reviewer fixed effects, a reviewer invite year-month fixed effect τ, and controls for logged team size and author country. Equation 7 is the second stage, where review Positivity is regressed on the predicted values of the endogenous variables and the same covariates as in the first-stage regressions. The results of this secondary analysis are in the third column of Table S6 ("IV").

First Stages:

$$Anon_{ijt} = \gamma_i + \tau_t + \rho_j + \beta_1 PolicyAvail_i + \beta_2 SCR_{ij} + \beta_3 PolicyAvail_i * SCR_{ij} + \beta_4 ln(TeamSize)_i + \beta_5 AuthCountry_i + \varepsilon_{ijt} \quad (6A)$$

$$Anon * SCR_{ijt} = \gamma_i + \tau_t + \rho_j + \beta_1 PolicyAvail_i + \beta_2 SCR_{ij} + \beta_3 PolicyAvail_i * SCR_{ij} + \beta_4 ln(TeamSize)_i + \beta_5 AuthCountry_i + \varepsilon_{ijt} \quad (6B)$$

Second Stage:

$$Positivity_{ijt} = \gamma_i + \tau_i + \rho_j + \beta_1 SCR_{ij} + \beta_2 \widehat{Anon}_{ijt} + \beta_3 \widehat{Anon * SCR}_{ijt} + \beta_4 ln(TeamSize)_i + \beta_5 AuthCountry_i + \varepsilon_{ijt} \quad (7)$$

From a policy perspective, the effect of the anonymization option on evaluations is interesting in itself. Double-anonymization can be costly to enforce, so a voluntary policy represents a potential approach for balancing goals of bias reduction and enforcement effort. The policy's effect can be interpreted as the intent-to-treat (ITT) effect, which only requires that the anonymization option is randomly assigned. Equation 8 below is the specification for the Intent-to-Treat (ITT) analysis, which takes a similar form to Equation 7. Here, we still have journal, reviewer invite year-month, and reviewer fixed effects, as well as controls for team size and author country, but we replace anonymization with the availability of the voluntary anonymization policy to estimate its effect on SCR Positivity (β3). Results for this subanalysis are in the first column of Table S6 ("ITT").

$$Positivity_{ijt} = \gamma_i + \tau_t + \rho_j + \beta_1 PolicyAvail_i + \beta_2 SCR_{ij} + \beta_3 PolicyAvail_i * SCR_{ij}$$

$$+ \beta_4 ln(TeamSize)_i + \beta_5 AuthCountry_i + \varepsilon_{ij} \quad (8)$$

Substantive significance of reducing geographical representation bias: back-of-the-envelope analysis

We estimate how diversifying the review pool would affect inequality in acceptance rates using a simple back-of-the-envelope calculation. Equation 9 is the specification used to estimate the association between the number of SCRs on a review panel and likelihood of acceptance for a back-of-the-envelope calculation. Note that the outcome and SCR are indexed at the manuscript-level *i*. SCR is not a binary variable in this specification, but rather the number of SCRs associated with a submitted manuscript. This specification includes journal fixed effects and controls for logged team size and corresponding author country.

$$Accepted_i = \gamma_i + \beta_1 SCR_i + \beta_2 ln(TeamSize)_i + \beta_3 AuthCountry_i + \varepsilon_i \quad (9)$$

Taking the estimates from this specification, we present a thought experiment proxying reviewer diversification policies where all submissions have an average level of SCR access. We do not assume that this policy is feasible to implement, but using such an extreme example can give a sense of the bounds of diversification's effectiveness.

Results

Is there country homophily in peer review?

We find that SCRs had a 4.78 p.p. ($p < 0.0001$) higher Positivity compared to non-SCRs evaluating the same submission. Figure 1 shows the "Overall" estimate and each of the three country income groups separately. Table S3 shows the full regression table.

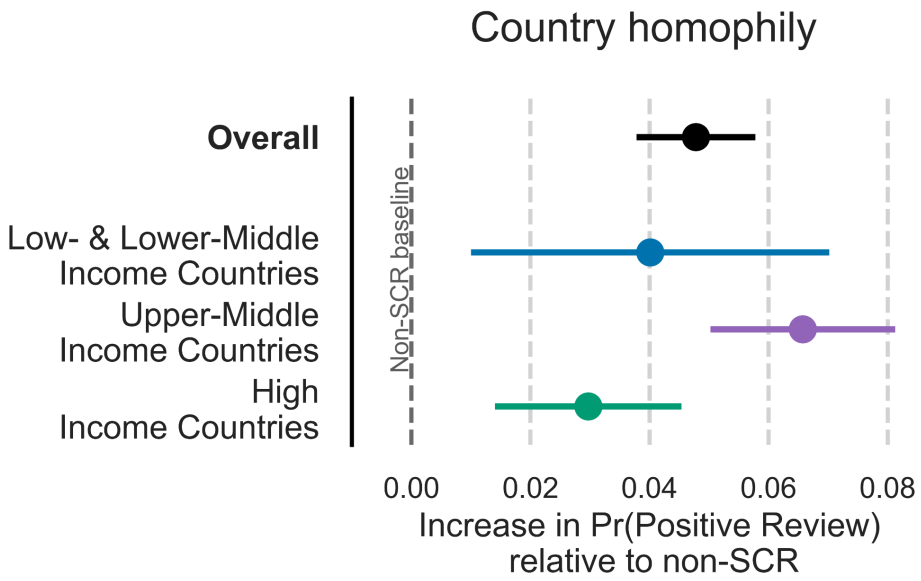


Fig. 1 SCRs are more likely to give positive reviews compared to non-SCRs on the same manuscript. *Marginal probability that a review is positive in overall data and for each country income group. The line at zero represents the positivity of non-SCRs. See Equations 2 and 2A for the model specifications. Standard errors were clustered at the level of the fixed effects and error bars are 95% confidence intervals.*

The intensity of SCR homophily varied by country income group, with the upper-middle income country (UMIC) group showing the largest estimates, but SCR homophily estimates were positive and significant for all (Table S4). We emphasize that these relationships between SCR status and positivity are not confounded by manuscript quality, because including manuscript fixed effects in the models ensures that *comparisons are between reviewers on the same manuscript.* Nor are they confounded by reviewers' standards, since reviewer fixed effects absorb individual reviewers' baseline positivity rates. The results establish the first necessary condition for geographical representation bias – homophily.

Are there geographical differences in access to homophily?

First, we investigate how SCR access is associated with the geographical composition of the reviewer pool. In principle, editors could assign reviewers to avoid or seek out SCRs, leading to a low or high SCR frequency for all countries, and therefore a low correlation between reviewer pool representation and SCR frequency. In contrast, assigning reviewers without regard to their country would lead to a high correlation, which is what our data reflect. Fig. 2A shows that SCR access for a country's authors is almost entirely predictable from that country's representation in the overall reviewer pool (Pearson corr. = 0.9761). The correlation between a country's number of *invited* reviewers and SCR access was similarly high (Pearson corr. = 0.9592). Three countries

– United States, China, and India – have outsized presence in the reviewer pool (and high SCR probabilities), so we consider them separately in subsequent analyses.

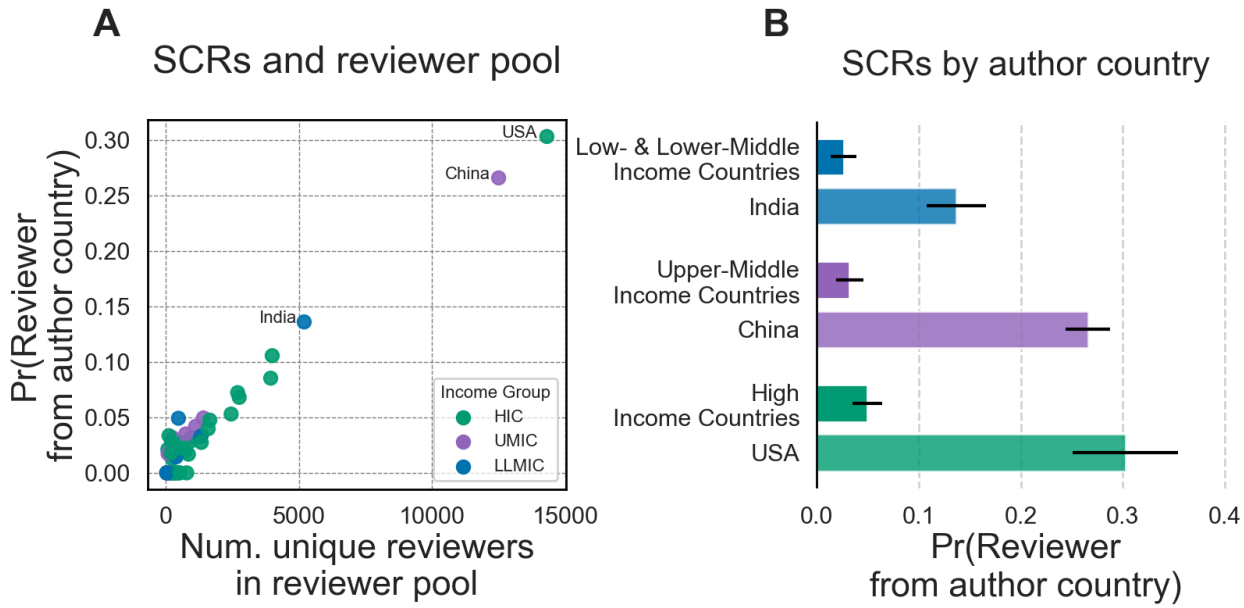


Fig. 2. Authors from countries well-represented in the reviewer pool are most likely to be reviewed by same-country reviewers. ***A****: Points are countries of corresponding authors. Countries with less than 100 submissions were excluded. SCR access (y-axis) estimated as described in Equation 3A. Estimates with p > 0.05 (some negative) were set to zero.* ***B****: Wealthier country groups have higher SCR access, excluding the largest submitter from each group (India, China, USA). Among those largest submitters (lighter bars), the wealthier ones have higher SCR access. See Equation 3B for the model specification, and Table S9 in SI Appendix for regression tables.*

The high correlation is consistent with editors inviting reviewers on criteria unrelated to their country, *e.g.,* expertise, resulting in as-if random sampling from the pool of previous reviewers. This suggests that differential access to country homophily is a structural outcome of the collective behavior of editors, reviewers, authors submitting works, as well as organizational policies and procedures like reviewer selection and recruitment practices.

Figure 2B shows that access to homophily varies substantially across and within country income groups. HIC authors receive SCRs over twice as often as LLMIC authors. USA-, China-, and India-based authors have the highest chance that a reviewer is SCR—30.27%, 26.59%, and 13.69%, respectively. The SCR rates of these well-represented countries are roughly eight to nine times higher relative to other countries in their respective income groups. For instance, LLMIC authors not from India have a 2.66% likelihood a review is from their country, the lowest rate among the country income categories (Table S9). Similar disparities can be seen in reviewer invitations (Fig. S4). These results establish the second necessary condition for geographical representation bias – differential access to homophily.

Does anonymization reduce geographical representation bias?

Using instrumental variables, we find that hiding author identities did not cause a significant reduction in homophily (among those submissions complying with the anonymization option). The estimated effect of anonymization on SCR Positivity, relative to non-SCRs, was to decrease it by 0.67% (p=0.0742) (see β_3 in Equation 7 and IV in the regression Table S6). For robustness, we also estimated SCR Positivity separately for voluntarily anonymized and non-anonymized submissions, using manuscript and reviewer fixed effects, and found that the two estimates were not statistically distinguishable (β_2=-0.0263 in Equation 4, p=0.1381, OLS in Table S6). The introduction of the anonymization option itself (*i.e.,* the ITT effect) did not change SCR Positivity significantly either (β_3=-0.0100 in Equation 8, p=0.3213, ITT in Table S6). The unresponsiveness of country homophily to the visibility of author information suggests that it is not driven by identity-based biases, pointing toward information or tastes mechanisms.

Substantive significance

Having established that geographic representation bias exists, we now perform a back-of-the-envelope analysis of how mitigating this bias would affect publication prospects for different authors. SCRs are associated with 6.78 p.p. ($p < 0.0001$) higher acceptance rates per SCR on the reviewer panel. If USA authors' SCR access was moved down from the observed 30.57% (Fig. 2A) to the dataset average of 14.68% (Table S1), they can be expected to experience a drop in acceptance rates of roughly 1.09 percentage points, *i.e.,* $(0.3071 - 0.1468) \times 0.0678$, from 70.71% to 69.63%. Conversely, consider Vietnam, an LLMIC with little representation in the reviewer pool. Raising Vietnamese authors' SCR access from the empirical 0.4% to the hypothetical 14.68% is expected to increase its acceptance rate from 60.92% to 61.88%, a 0.96 p.p. increase.

This exercise suggests that substantial reviewer diversification would have fairly modest effects, increasing acceptance by roughly 1 percentage point for the most disadvantaged groups. Importantly, diversification does not make everyone better off, as it would reduce acceptance rates of advantaged groups by the same amount.

Discussion

This study investigated geographical representation bias in scientific peer review—a bias hypothesized to result from the lack of geographical diversity among reviewers. There are two necessary conditions for geographical representation bias: reviewers need to systematically favor research from their own countries, *i.e.*, "country homophily", and authors from different countries need to have different chances of being assigned to such favorable reviewers, *i.e.* "access to homophily." We found strong evidence of homophily. Reviewers from the same country as authors (same-country reviewers or SCRs) were 4.78 p.p. more positive compared to

non-SCRs on the same manuscript. These reviewers were also 4.74 p.p. more likely to respond to invitations to review (Table S3), possibly making them more attractive to editors. To our knowledge, this study is the first to demonstrate author-reviewer country homophily conclusively by ruling out key confounders like submission topic or quality. This was made possible by comparing different reviewers of the same manuscript using manuscript fixed effects.

We also found strong evidence that authors from different countries had dramatically different access to homophily. Authors in USA, China, and India had the highest chance that a reviewer is an SCR—30.27%, 26.59%, and 13.69%, respectively. Low- and lower-middle-income countries (LLMIC) authors not from India had only a 2.66% chance of SCRs, the lowest rate among the country income categories (Table S9). Although editors appear to invite reviewers based on criteria unrelated to geography, which is arguably a fair, non-discriminatory practice, the practice advantages authors from countries with many reviewers in the reviewer pool, as it results in them being reviewed by SCRs more often. By analogy with the racial discrimination literature (Small and Pager 2020), the representation bias is structural and not directly linkable to any individual editor or other actor.

While the country homophily we observed in peer review was substantial, it is important to note that mitigating representation bias is not a "great equalizer" of geographical disparities in publishing. Our back-of-the-envelope calculations suggest that representation bias accounts for roughly 2 percentage points of the up to 30-percentage point disparity in acceptance rates between country income groups (Fig. S3).

We also considered double-anonymization as a strategy to reduce representation bias. Double-anonymization has proven effective in reducing prestige bias (Blank 1991; Fox et al. 2023; Huber et al. 2022; Okike et al. 2016; Sun et al. 2022; Tomkins et al. 2017), but evidence on whether it reduces country homophily was lacking until now. A policy enabling authors to choose double-anonymous review (instead of the usual single-anonymous) was rolled out in a staggered and arbitrary manner, allowing us to conduct an instrumental variables analysis, and thus controlling for self-selection into double-anonymization. We used policy availability and availability interacted with SCR status as instruments for anonymization and anonymization interacted with SCR status. We found that anonymization had no significant effect on SCRs' Positivity (Table S6).

Several limitations and open questions remain. First, an overwhelming majority of submissions are from teams of researchers, but we focus on the country of the corresponding author's institution, assuming that the corresponding author is most relevant to how the submission is evaluated. We find our results are robust when focusing on first authors, single-country teams, and regions instead of countries (Tables S7, S8, and S10, see "Submission Country Robustness Checks" in SI Appendix). Nevertheless, the measure does not account for researchers who have moved countries. If authors' and reviewers' self-reported countries are not reflective of their information and tastes, our estimate of country homophily might be conservative (*i.e.,*

downwardly biased). Future analyses with more detailed data on researchers' geographic trajectories may help elucidate the sub-mechanisms of homophily.

Second, it is important to study how editors interpret reviews from SCR and non-SCR reviewers. If country homophily is publicized, editors may discount SCR recommendations when making final decisions, and thereby reducing the effect of representation bias on final decisions. Additionally, as large as our data are, they comprise only one area of science and one language. For example, if non-English-language journals attract authors and reviewers who are more geographically concentrated, geographic homophily among reviewers may not get activated and this type of representation bias may not occur. Intensity of representation bias may also vary across fields based on their level of consensus. The physical sciences, the setting of this study, have relatively high consensus on establishing scientific validity, so the "room" for representation bias is arguably smallest there. In fields with less consensus, representation bias may be larger.

Last, those who choose to review could differ systematically from those who do not. For example, reviewers might be more biased due to expectations of reciprocity (*e.g.,* cartel behavior) or because senior researchers have different time constraints compared to junior researchers. Such self-selection occurs in our data and likely in many journal peer review settings. However, settings with less self-selection, such as grant review panels where panelists do not choose which proposals to review, might show less homophily, and therefore less representation bias.

In sum, our results establish geographical representation bias. To put it simply, the success of scientific papers is determined not only by *what* is in them but by *who* evaluates them. Peer reviewers who are geographically close to a paper's authors are systematically more favorable, and authors from wealthier countries that produce more science are more likely to be reviewed by such favorable reviewers. Anonymizing submissions, which appears to reduce other types of biases, was ineffective at reducing country homophily, and therefore representation bias, in our setting. This suggests that scientific institutions are unlikely to reduce geographical representation bias with "quick fixes," and instead need to invest in long-term reviewer diversification efforts. However, our back-of-the-envelope calculation shows that even such efforts would do little to equalize the sizable disparities in acceptance rates by geography, as even giving all authors similar access to favorable reviewers would change the rates by about 1%.

Data Availability

Deidentified data and replication code can be retrieved from Deep Blue Data (Dumlao and Teplitskiy 2024).

Acknowledgments

We are indebted to the Institute of Physics Publishing, and particularly Kim Eggleton and Mikka Pers, for supporting this project. We also thank Jason Owen-Smith, Scott Page, and Daniel Romero for extensive feedback on this project. We also thank the audiences of the following seminars and meetings for many useful suggestions and insights: Institute of Physics Publishing, Annual Meeting of the Society for the Social Studies of Science, New York University Abu Dhabi Social Research and Public Policy, International Conference on the Science of Science and Innovation, University of Michigan Computational Social Science Working Group, and University of Michigan Social Behavioral and Experimental Economics seminar.

We thank *Schmidt Futures* and the *Science for Progress Initiative* for financial support.

References

Altman, Micah, and Philip N. Cohen. 2021. *Openness and Diversity in Journal Editorial Boards*. *preprint*. SocArXiv. doi: 10.31235/osf.io/4nq97.

Amano, Tatsuya, Valeria Ramírez-Castañeda, Violeta Berdejo-Espinola, Israel Borokini, Shawan Chowdhury, Marina Golivets, Juan David González-Trujillo, Flavia Montaño-Centellas, Kumar Paudel, Rachel Louise White, and Diogo Veríssimo. 2023. "The Manifold Costs of Being a Non-Native English Speaker in Science" edited by U. Dirnagl. *PLOS Biology* 21(7):e3002184. doi: 10.1371/journal.pbio.3002184.

Balli, Hatice Ozer, and Bent E. Sørensen. 2013. "Interaction Effects in Econometrics." *Empirical Economics* 45(1):583–603. doi: 10.1007/s00181-012-0604-2.

Banks, Michael. 2020. "IOP Publishing Pledges to Implement Double-Blind Peer Review." *Physics World*.

Blank, Rebecca M. 1991. "The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from The American Economic Review." *The American Economic Review* 81(5):1041–67.

Boudreau, Kevin J., Eva C. Guinan, Karim R. Lakhani, and Christoph Riedl. 2016. "Looking Across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science." *Management Science* 62(10):2765–83. doi: 10.1287/mnsc.2015.2285.

Brogaard, Jonathan, Joseph Engelberg, and Christopher A. Parsons. 2014. "Networks and Productivity: Causal Evidence from Editor Rotations." *Journal of Financial Economics* 111(1):251–70. doi: 10.1016/j.jfineco.2013.10.006.

Burns, C. Sean, and Charles W. Fox. 2017. "Language and Socioeconomics Predict Geographic Variation in Peer Review Outcomes at an Ecology Journal." *Scientometrics* 113(2):1113–27. doi: 10.1007/s11192-017-2517-5.

Carlino, Gerald, and William R. Kerr. 2015. "Agglomeration and Innovation." Pp. 349–404 in *Handbook of Regional and Urban Economics*. Vol. 5. Elsevier.

Chen, Christine Yifeng, Sara S. Kahanamoku, Aradhna Tripati, Rosanna A. Alegado, Vernon R. Morris, Karen Andrade, and Justin Hosbey. 2022. "Systemic Racial Disparities in Funding Rates at the National Science Foundation." *eLife* 11:e83071. doi: 10.7554/eLife.83071.

Chon, Bum Soo, George Barnett, and Young Choi. 2003. "Clustering Local Tastes in Global Culture: The Reception Structure of Hollywood Films." *Sociological Research Online* 8(1):1–12. doi: 10.5153/sro.781.

Collyer, Fran M. 2018. "Global Patterns in the Publishing of Academic Knowledge: Global North, Global South." *Current Sociology* 66(1):56–73. doi: 10.1177/0011392116680020.

Dumlao, James M. Zumel, and Misha Teplitskiy. 2024. "Double-Anonymous Peer Review Mitigates Country Homophily and the Harms of Low Reviewer Diversity: Deidentified Data and Replication Code."

Fox, Charles W., Jennifer Meyer, and Emilie Aimé. 2023. "Double‐blind Peer Review Affects Reviewer Ratings and Editor Decisions at an Ecology Journal." *Functional Ecology* 37(5):1–14. doi: 10.1111/1365-2435.14259.

Fox, Charles W., and C. E. Timothy Paine. 2019. "Gender Differences in Peer Review Outcomes and Manuscript Impact at Six Journals of Ecology and Evolution." *Ecology and Evolution* 9(6):3599–3619. doi: 10.1002/ece3.4993.

Fry, Caroline V. 2023. “Local Researchers, Local Problems.”
Garisto, Daniel. 2022. “ArXiv.Org Reaches a Milestone and a Reckoning.” *Scientific American*, January 10.
Gaston, Thomas, and Pippa Smart. 2018. “What Influences the Regional Diversity of Reviewers: A Study of Medical and Agricultural/Biological Sciences Journals: Regional Diversity of Reviewers.” *Learned Publishing* 31(3):189–97. doi: 10.1002/leap.1155.
Gomez, Charles J., Andrew C. Herman, and Paolo Parigi. 2022. “Leading Countries in Global Science Increasingly Receive More Citations than Other Countries Doing Similar Research.” *Nature Human Behaviour* 6(7):919–29. doi: 10.1038/s41562-022-01351-5.
Harris, Matthew, Joachim Marti, Hillary Watt, Yasser Bhatti, James Macinko, and Ara W. Darzi. 2017. “Explicit Bias Toward High-Income-Country Research: A Randomized, Blinded, Crossover Experiment Of English Clinicians.” *Health Affairs* 36(11):1997–2004. doi: 10.1377/hlthaff.2017.0773.
Heesen, Remco, and Jan-Willem Romeijn. 2019. “Epistemic Diversity and Editor Decisions: A Statistical Matthew Effect.” *Philosophers’ Imprint* 19(39):1–20.
Hoppe, Travis A., Aviva Litovitz, Kristine A. Willis, Rebecca A. Meseroll, Matthew J. Perkins, B. Ian Hutchins, Alison F. Davis, Michael S. Lauer, Hannah A. Valantine, James M. Anderson, and George M. Santangelo. 2019. “Topic Choice Contributes to the Lower Rate of NIH Awards to African-American/Black Scientists.” *Science Advances* 5(10):eaaw7238. doi: 10.1126/sciadv.aaw7238.
Huber, Jürgen, Sabiou Inoua, Rudolf Kerschbamer, Christian König-Kersting, Stefan Palan, and Vernon L. Smith. 2022. “Nobel and Novice: Author Prominence Affects Peer Review.” *Proceedings of the National Academy of Sciences* 119(41):e2205779119. doi: 10.1073/pnas.2205779119.
Inglehart, Ronald, and Marita Carballo. 1997. “Does Latin America Exist? (And Is There a Confucian Culture?): A Global Analysis of Cross-Cultural Differences.” *PS: Political Science & Politics* 30(1):34–47.
Kowal, Marta, Piotr Sorokowski, Emanuel Kulczycki, and Agnieszka Żelaźniewicz. 2022. “The Impact of Geographical Bias When Judging Scientific Studies.” *Scientometrics* 127(1):265–73. doi: 10.1007/s11192-021-04176-7.
Kozlowski, Diego, Vincent Larivière, Cassidy R. Sugimoto, and Thema Monroe-White. 2022. “Intersectional Inequalities in Science.” *Proceedings of the National Academy of Sciences* 119(2):e2113067119. doi: 10.1073/pnas.2113067119.
Li, Danielle. 2017. “Expertise versus Bias in Evaluation: Evidence from the NIH.” *American Economic Journal: Applied Economics* 9(2):60–92. doi: 10.1257/app.20150421.
Link, Ann M. 1998. “US and Non-US Submissions: An Analysis of Reviewer Bias.” *JAMA* 280(3):246–47. doi: 10.1001/jama.280.3.246.
Liu, Fengyuan, Talal Rahwan, and Bedoor AlShebli. 2023. “Non-White Scientists Appear on Fewer Editorial Boards, Spend More Time under Review, and Receive Fewer Citations.” *Proceedings of the National Academy of Sciences* 120(13):e2215324120. doi: 10.1073/pnas.2215324120.
McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. “Birds of a Feather: Homophily in Social Networks.” *Annual Review of Sociology* 27(1):415–44. doi: 10.1146/annurev.soc.27.1.415.
Merton, Robert K. 1968. “The Matthew Effect in Science: The Reward and Communication Systems of Science Are Considered.” *Science* 159(3810):56–63. doi:

10.1126/science.159.3810.56.
Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. The University of Chicago Press.
Mohanti, Bidhu Kalyan, Anusheel Munshi, Biplab Sarkar, Atul Sharma, and Surya Vs Deo. 2023. "Peer Reviewers from Low- and Middle-Income Countries(LMIC) for Open Access Journals in Oncology Can Improve the Equity in Cancer Research and Clinical Trials." *Journal of Cancer Policy* 36:100419. doi: 10.1016/j.jcpo.2023.100419.
Murray, Dakota, Kyle Siler, Vincent Larivière, Wei Mun Chan, Andrew M. Collings, Jennifer Raymond, and Cassidy R. Sugimoto. 2018. *Author-Reviewer Homophily in Peer Review*. *preprint*. Scientific Communication and Education. doi: 10.1101/400515.
NHB Editorial Board. 2023. "Scientific Publishing Has a Language Problem." *Nature Human Behaviour* 7(7):1019–20. doi: 10.1038/s41562-023-01679-6.
Nielsen, Mathias Wullum, and Jens Peter Andersen. 2021. "Global Citation Inequality Is on the Rise." *Proceedings of the National Academy of Sciences* 118(7):e2012208118. doi: 10.1073/pnas.2012208118.
Nielsen, Mathias Wullum, Christine Friis Baker, Emer Brady, Michael Bang Petersen, and Jens Peter Andersen. 2021. "Weak Evidence of Country- and Institution-Related Status Bias in the Peer Review of Abstracts." *eLife* 10:e64561. doi: 10.7554/eLife.64561.
Okike, Kanu, Kevin T. Hug, Mininder S. Kocher, and Seth S. Leopold. 2016. "Single-Blind vs Double-Blind Peer Review in the Setting of Author Prestige." *JAMA* 316(12):1315. doi: 10.1001/jama.2016.11014.
Ozgun, Burcu, and Tom Broekel. 2021. "The Geography of Innovation and Technology News - An Empirical Study of the German News Media." *Technological Forecasting and Social Change* 167:1–16. doi: 10.1016/j.techfore.2021.120692.
Peng, Hao, Karim Lakhani, and Misha Teplitskiy. 2021. "Acceptance in Top Journals Shows Large Disparities across Name-Inferred Ethnicities."
Qiu, Shumin, Claudia Steinwender, and Pierre Azoulay. 2023. "NBER WORKING PAPER SERIES WHO STANDS ON THE SHOULDERS OF CHINESE (SCIENTIFIC) GIANTS? EVIDENCE FROM CHEMISTRY." *NBER* Working Paper 30772.
Qiu, Shumin, Claudia Steinwender, and Pierre Azoulay. 2024. *Paper Tiger? Chinese Science and Home Bias in Citations*. w32468. Cambridge, MA: National Bureau of Economic Research. doi: 10.3386/w32468.
Rivera, Lauren A., and Jayanti Owens. 2021. "Glass Floors and Glass Ceilings: Sex Homophily and Heterophily in Job Interviews." *Social Forces* 99(4):1363–93. doi: 10.1093/sf/soaa072.
Rubin, Amir, Eran Rubin, and Dan Segal. 2023. "Editor Home Bias?" *Research Policy* 52(6):1–15. doi: 10.1016/j.respol.2023.104766.
Small, Mario L., and Devah Pager. 2020. "Sociological Perspectives on Racial Discrimination." *Journal of Economic Perspectives* 34(2):49–67. doi: 10.1257/jep.34.2.49.
Smith, Olivia M., Kayla L. Davis, Riley B. Pizza, Robin Waterman, Kara C. Dobson, Brianna Foster, Julie C. Jarvey, Leonard N. Jones, Wendy Leuenberger, Nan Nourn, Emily E. Conway, Cynthia M. Fiser, Zoe A. Hansen, Ani Hristova, Caitlin Mack, Alyssa N. Saunders, Olivia J. Utley, Moriah L. Young, and Courtney L. Davis. 2023. "Peer Review Perpetuates Barriers for Historically Excluded Groups." *Nature Ecology & Evolution* 7(4):512–23. doi: 10.1038/s41559-023-01999-w.
Snellman, Kaisa, and Isabelle Solal. 2023. "Does Investor Gender Matter for the Success of

Female Entrepreneurs? Gender Homophily and the Stigma of Incompetence in Entrepreneurial Finance." *Organization Science* 34(2). doi: 10.4337/9781849807630.00024.

Sun, Mengyi, Jainabou Barry Danfa, and Misha Teplitskiy. 2022. "Does Double‐blind Peer Review Reduce Bias? Evidence from a Top Computer Science Conference." *Journal of the Association for Information Science and Technology* 73(6):811–19. doi: 10.1002/asi.24582.

Teplitskiy, Misha, Hao Peng, Andrea Blasco, and Karim R. Lakhani. 2022. "Is Novel Research Worth Doing? Evidence from Peer Review at 49 Journals." *Proceedings of the National Academy of Sciences* 119(47):e2118046119. doi: 10.1073/pnas.2118046119.

Thelwall, Mike, Kayvan Kousha, Mahshid Abdoli, Emma Stuart, Meiko Makita, Paul Wilson, and Jonathan Levitt. 2023. "Why Are Coauthored Academic Articles More Cited: Higher Quality or Larger Audience?" *Journal of the Association for Information Science and Technology* 74(7):791–810. doi: 10.1002/asi.24755.

Tomkins, Andrew, Min Zhang, and William D. Heavlin. 2017. "Reviewer Bias in Single- versus Double-Blind Peer Review." *Proceedings of the National Academy of Sciences* 114(48):12708–13. doi: 10.1073/pnas.1707323114.

Verlegh, Peeter W. J. 2007. "Home Country Bias in Product Evaluation: The Complementary Roles of Economic and Socio-Psychological Motives." *Journal of International Business Studies* 38(3):361–73. doi: 10.1057/palgrave.jibs.8400269.

Wright, Nataliya Langburd, Rembrand Koning, and Tarun Khanna. 2023. "Judging Foreign Startups." *Strategic Management Journal* smj.3490. doi: 10.1002/smj.3490.

Zinovyeva, Natalia, and Manuel Bagues. 2015. "The Role of Connections in Academic Promotions." *American Economic Journal: Applied Economics* 7(2):264–92. doi: 10.1257/app.20120337.

SI Appendix

This PDF file includes:
Supporting Text
Figures S1 to S4
Tables S1 to S10
SI References

Supporting Text

Instrumental Variable Approach: We conduct a secondary, exploratory analysis of whether double-anonymization would be an effective policy to mitigate geographical representation bias. Unlike the homophily analysis, anonymization status does not vary within-manuscript, so comparisons must be made across manuscripts. However, naively comparing evaluations across anonymized and non-anonymized submissions may give a biased estimate of the anonymization effect, because certain types of authors may have systematically chosen or avoided anonymization. To address this concern, we use the instrumental variables (IV) approach. IV accounts for the self-selection problem by leveraging an auxiliary variable (the instrument), which for this study is the availability of the anonymization option.

As mentioned in Materials and Methods, this IV analysis is enabled by the introduction of a voluntary anonymization policy across IOP's journals. Ideally, this analysis would simply require a t-test for the significance of the interaction effect of anonymization and SCR status, β3 in the simplified structural Equation 10 below.

$$Positivity_{ij} = \beta_1 Anon_i + \beta_2 SCR_{ij} + \beta_3 Anon_i * SCR_{ij} + \epsilon_{ij} \quad (10)$$

However, since anonymization status is likely endogenous due to self-selection or other omitted variables, the interaction between anonymization and SCR status would also be endogenous. We use the availability of the option to anonymize as an instrumental variable to estimate the causal effect of anonymization on reviewer Positivity (IV in Table S6). For anonymization policy availability to be a valid instrument, three assumptions must be met: exogeneity, relevance, and the exclusion restriction.

The exogeneity criterion implies that the timing of policy availability across journals should be orthogonal to the error term of the structural model (i.e., regressing the endogenous, independent variable and instrument on the dependent variable together). In other words, policy availability should be independent of any unobserved variables correlated with review Positivity. This assumption is supported by the data-generating process, since IOP's staggered rollout was done in an arbitrary order in consultation with the last author [MT]. This quasi-random rollout means the availability of the anonymization option assigned unrelatedly or exogenously to reviewers' behavior. Self-selection into treatment (anonymization) is one form of endogeneity between the treatment and the outcome (review Positivity). Using randomized assignment (anonymization option) as an instrument sidesteps the issue of selection into treatment because now we are estimating the effect of assignment, which we assume is only correlated with the treatment and not with any other determinants (see discussion on exclusion restriction below), on the outcome. Randomized assignment also enables a valid ITT analysis.

The relevance assumption implies that the ability to anonymize should have a strong effect on authors' choice to anonymize. Conventionally, strength is measured using the F-statistic of

the instruments in the first-stage regressions. With two endogenous variables, we conduct separate F-tests for the joint significance of our two instruments on the two first-stage regressions. In the first-stage regression predicting anonymization status (Equation 6A), the inclusion of the two instruments was statistically significant with an F-stat of 11.17. More important to estimating the interaction effect of interest, the F-stat of the instruments in the first-stage regression predicting the interaction between anonymization and SCR status is 33.25. Thus, the instruments meet the conventional standard for instrument strength of F>10 in both first stages.

The exclusion restriction in this context implies that the availability of the anonymization policy should influence review Positivity only through anonymization. One way the exclusion restriction might be violated is if the availability of the policy also affects the characteristics of submitting authors or reviewers, for example if some authors choose not to submit because the policy was introduced. This is a separate selection concern from self-selection into treatment, in which the groups *assigned* to treatment (option available) and control (not available) are statistically different on observables and unobservables. A previous study tested for whether the policy was associated with changes in two productivity characteristics of submitting authors or invited reviewers—their publications and accumulated citations—and found no statistically significant changes (Smirnova, Romero, and Teplitskiy 2022), consistent with the restriction. Additionally, we examine how the instrument affects four additional variables at the manuscript-level—logged team size (number of authors on submission), probability of a submission being reviewed by at least one SCR, number of realized reviewers, and probability of a submission being desk rejected—using the model specification in Equation 5. Table S5 shows that the anonymization option introduction did not have a statistically significant effect on these manuscript characteristics, further supporting this assumption. Still, the exclusion restriction is an assumption, and these tests do not rule out all possible threats to validity. Regarding author and reviewer characteristics, we are unable to test for differences between those who can and cannot anonymized within productivity strata. For example, if reviewers exhibit a systematic gender bias, it is possible that the option to anonymize may attract authors of one gender with similar productivity to those in the existing author pool. However, if the gender bias is not fully eliminated through anonymization, then policy availability may shift reviewer Positivity by changing the author pool (i.e., not through anonymization alone). Regarding manuscript-level differences, the anonymization option may have attracted lower-quality submissions since authors would be less likely to see negative reputational effects from reviewers, even though editors can still see authors' identities. In this case, reviewers may be more negative for the group of manuscripts with the anonymization policy available through a channel other than anonymization.

By making the three assumptions, the estimates from our analysis are interpreted as the Local Average Treatment Effect (LATE), also known as the Average Treatment Effect (ATE) among compliers. Observations can be conceptually categorized into four groups: Always-takers, never-takers, defiers, and compliers.

Always-takers receive the treatment whether they are assigned to it or not. In this study, we should have no always-takers (i.e., one-sided non-compliance) because submitting authors could not choose to anonymize unless the policy was available to them. There were 18 observations of submissions that were anonymized before the policy was available, but these are possibly administrative errors and were removed for analysis. Never-takers do not receive the treatment regardless of their assignment. When anonymization was an option, those who did not

anonymize—most submitters in our sample—are considered never-takers. However, we cannot observe the proportion of those who submitted when anonymization was not an option who were never-takers. We assume the proportion is similar to that on the anonymized (treated) side. Defiers are those who would receive the opposite treatment from what they are assigned (i.e., anonymize if option is not available or don't anonymize if option is available). For the instrumental variables approach, it must be assumed that defiers do not exist in our sample. Compliers are those who receive the treatment that they are assigned.

Conceptually, the LATE is the effect of assignment (anonymization policy availability) on the outcome (Positivity) scaled by the proportion of observations in the complier group. Among the submissions that received reviews, 28,666 had the anonymization option (assigned to treatment) and 84,566 did not (assigned to control). 8,742 submissions were anonymized so the proportion of compliers on the assigned to treatment side is 30.5%. A potential concern is the low take-up rate (~8% overall, ~30% when option was available), since our causal estimates apply only to authors who are induced to anonymize. This set of authors may vary with different policy implementations.

Additionally, we restricted our analysis to papers that received reviews (55.72% of all submissions), i.e. editors did not desk-reject. This restriction is unlikely to bias our results because neither the instrument nor the treatment affected the information editors see, and therefore should not have affected their desk-rejection decisions. Consistent with this argument, we found that the instrument did not affect desk-rejection probability (Table S5).

Positive and Not-Positive Codes for Reviewer Response Categories: The categorical responses coded as 0 (i.e., recommendation to reject) include, "Reject (average work that adds little new knowledge to the field)", "Reject (contains basic errors and faulty judgements)", "Unsatisfactory revision", "Reject (adds little new knowledge to the field; may be publishable in an archival journal)", "Reject (contains basic errors and faulty judgements; not publishable in any journal)", "Reject but may resubmit (may be publishable in this journal following these substantial revisions)", "Reject", "Reject as a Letter but may be resubmitted for consideration as a regular paper after expansion and/or revision", "Reject ", "Reject as a Letter but can be considered as a regular paper in current form", "Reject as a letter but may be resubmitted for consideration as a regular paper after expansion and/or revision", "Reject (not publishable even with revisions)", "Reject as a fast track communication but may be resubmitted for consideration as a regular paper after expansion and/or revision", "Reject (adds little knowledge to the field; may be publishable in an archival journal)", "Article is fundamentally flawed", "Reject as a frontline article but may be resubmitted for consideration as a regular paper after expansion and/or revision", "Reject as a fast track communication but can be considered as a regular paper in current form", "Reject as a letter but can be considered as a regular paper in current form", "Manuscript is fundamentally flawed", and "Reject as a frontline article but can be considered as a regular paper in current form." While some of these categories hint at possible acceptance in the future, they require the manuscript to be resubmitted which would appear in the data as a separate manuscript submission.

The categorical responses coded as 1 (i.e., recommendation to accept) include, "Accept", "Amendments required before acceptance", "Major revision and re-submission recommended (with these substantial revisions I expect that the article would make a significant contribution to the literature)", "Major revision required before acceptance", "Revisions required before acceptance (only changes to the text are required)", "Revisions are required", "Major revisions

required before acceptance", "Accept after optional revisions", "Accept in current form", "Revisions required before acceptance", "Article requires revisions to meet accepted scientific standards", "Article displays scientific rigour - further revisions are optional", "Manuscript displays scientific rigour - further revisions are optional", "Manuscript requires revisions to meet accepted scientific standards", "Article displays significant scientific rigour", and "Manuscript displays significant scientific rigour."

Countries by Income Group: High-income countries (HIC): Australia, Austria, Bahrain, Belgium, Brunei Darussalam, Canada, Chile, Cocos (Keeling) Islands, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Gibraltar, Greece, Hong Kong, Hungary, Iceland, Ireland, Isle of Man, Israel, Italy, Japan, Korea (Republic of), Kuwait, Latvia, Lithuania, Luxembourg, Macao, Malta, Martinique, Monaco, Netherlands, New Caledonia, New Zealand, Norway, Oman, Panama, Poland, Portugal, Puerto Rico, Qatar, Romania, Saudi Arabia, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, Taiwan, Trinidad and Tobago, United Arab Emirates, United Kingdom, United States Minor Outlying Islands, United States of America, Uruguay

Upper-middle-income countries (UMIC): Albania, Argentina, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, China, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, Fiji, French Guiana, Gabon, Georgia, Guatemala, Iraq, Jamaica, Jordan, Kazakhstan, Libya, Macedonia (the former Yugoslav Republic of), Malaysia, Mauritius, Mexico, Moldova (Republic of), Montenegro, Namibia, Paraguay, Peru, Russian Federation, Serbia, South Africa, Thailand, Turkey

Low- and Lower-middle-income countries (LLMIC): Afghanistan, Algeria, Bangladesh, Benin, Bolivia (Plurinational State of), Burkina Faso, Burundi, Cameroon, Congo, Congo (Democratic Republic of the), Côte d'Ivoire, Egypt, Ethiopia, Ghana, India, Indonesia, Iran (Islamic Republic of), Kenya, Korea (Democratic People's Republic of), Kyrgyzstan, Lebanon, Lesotho, Madagascar, Malawi, Mali, Mauritania, Mongolia, Morocco, Mozambique, Myanmar, Nepal, Niger, Nigeria, Pakistan, Palestine, State of, Papua New Guinea, Philippines, Rwanda, Senegal, Somalia, Sri Lanka, Sudan, Swaziland, Syrian Arab Republic, Tajikistan, Tanzania, Tanzania, United Republic of, Togo, Tunisia, Uganda, Ukraine, Uzbekistan, Venezuela (Bolivarian Republic of), Viet Nam, Yemen, Zambia, Zimbabwe

Submission Country Robustness Checks: First, we investigate whether using the lead author's country as the country of a submission changes our findings. In our data, lead and corresponding authors are most often from the same country, and even more often from the same region. Thus, we do not anticipate different findings using the lead author instead of the corresponding author. Table S7 shows that estimates of SCR homophily are similar whether corresponding or lead author country is used as reference. The number of observations differ between these models because corresponding authors were required to self-report their country upon submission whereas lead authors were not. Comparing the first two models in Table S10, differences in estimates of access to SCRs appear trivial between using corresponding or lead author. Therefore, evidence provided of geographical representation bias is robust when lead author country is substituted for corresponding author country in determining where a submission is from.

Second, we use two subsets to validate the country of the corresponding author as a proxy for the country of a submitted manuscript. The regional team subset contains submissions with

authors all from the same region (see Countries by Region), and the national team subset consists of submissions with authors all from the same country. A minority of submissions in our data are international, and even fewer span more than one region, so we do not expect findings to change using these subsets. Table S8 confirms that estimates of SCR homophily are similar between the national, regional, and full data. Again, Table S10 confirms that estimates of access to SCRs are similar in these subsets, where the relevant comparison is between the first model to the third and fourth models. We conclude that our main findings are robust when looking only at national or regional teams, who might be expected to have the stronger SCR effects.

Countries by Region: Africa: Algeria, Burkina Faso, Burundi, Cameroon, Congo, Congo (Democratic Republic of the), Côte d'Ivoire, Egypt, Ethiopia, Gabon, Ghana, Kenya, Lesotho, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Somalia, South Africa, Sudan, Swaziland, Tanzania, United Republic of, Togo, Tunisia, Uganda, Zambia, Zimbabwe

Asia: Afghanistan, Armenia, Azerbaijan, Bahrain, Bangladesh, Benin, Brunei Darussalam, China, Cyprus, Georgia, Hong Kong, India, Indonesia, Iran (Islamic Republic of), Iraq, Israel, Japan, Jordan, Kazakhstan, Korea (Democratic People's Republic of), Korea (Republic of), Kuwait, Kyrgyzstan, Lebanon, Macao, Malaysia, Mongolia, Myanmar, Nepal, Oman, Pakistan, Palestine, State of, Philippines, Qatar, Saudi Arabia, Singapore, Sri Lanka, Syrian Arab Republic, Taiwan, Tajikistan, Thailand, Turkey, United Arab Emirates, Uzbekistan, Viet Nam, Yemen

Europe: Albania, Austria, Belarus, Belgium, Bosnia and Herzegovina, Botswana, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Gibraltar, Greece, Hungary, Iceland, Ireland, Isle of Man, Italy, Latvia, Lithuania, Luxembourg, Macedonia (the former Yugoslav Republic of), Malta, Moldova (Republic of), Monaco, Montenegro, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Ukraine, United Kingdom

Latin America and the Caribbean: Argentina, Bolivia (Plurinational State of), Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, French Guiana, Guatemala, Jamaica, Martinique, Mexico, Panama, Paraguay, Peru, Puerto Rico, Trinidad and Tobago, Uruguay, Venezuela (Bolivarian Republic of)

Northern America: Canada, United States of America

Oceania: Australia, Cocos (Keeling) Islands, Fiji, New Caledonia, New Zealand, Papua New Guinea, United States Minor Outlying Islands

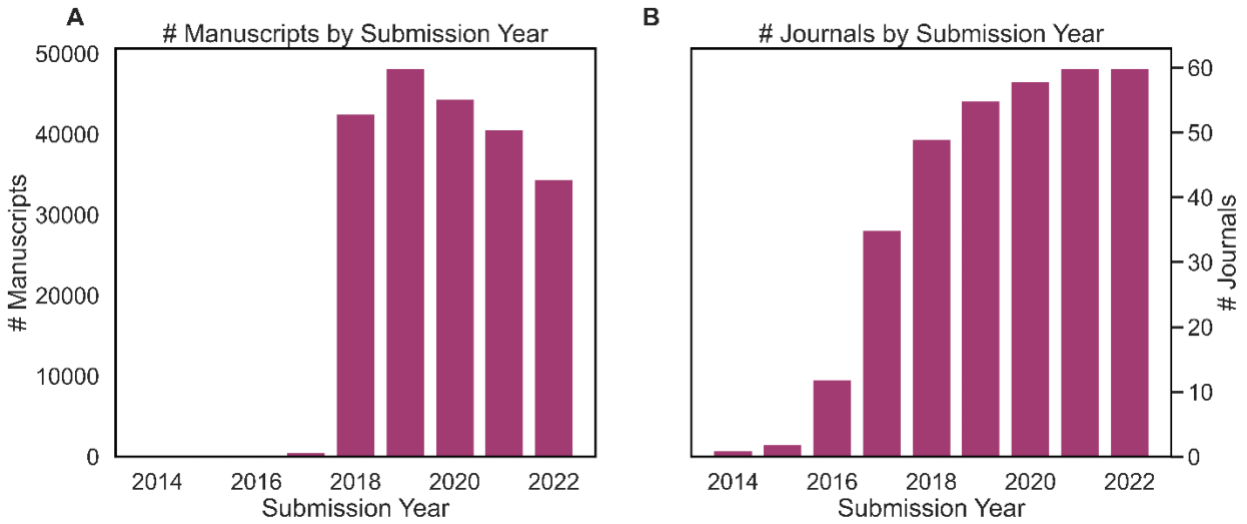


Fig. S1.
Number of manuscripts and journals by year. A: The modal manuscript received a final decision in 2019. B: Journals increased every year until reaching 60 in 2021.

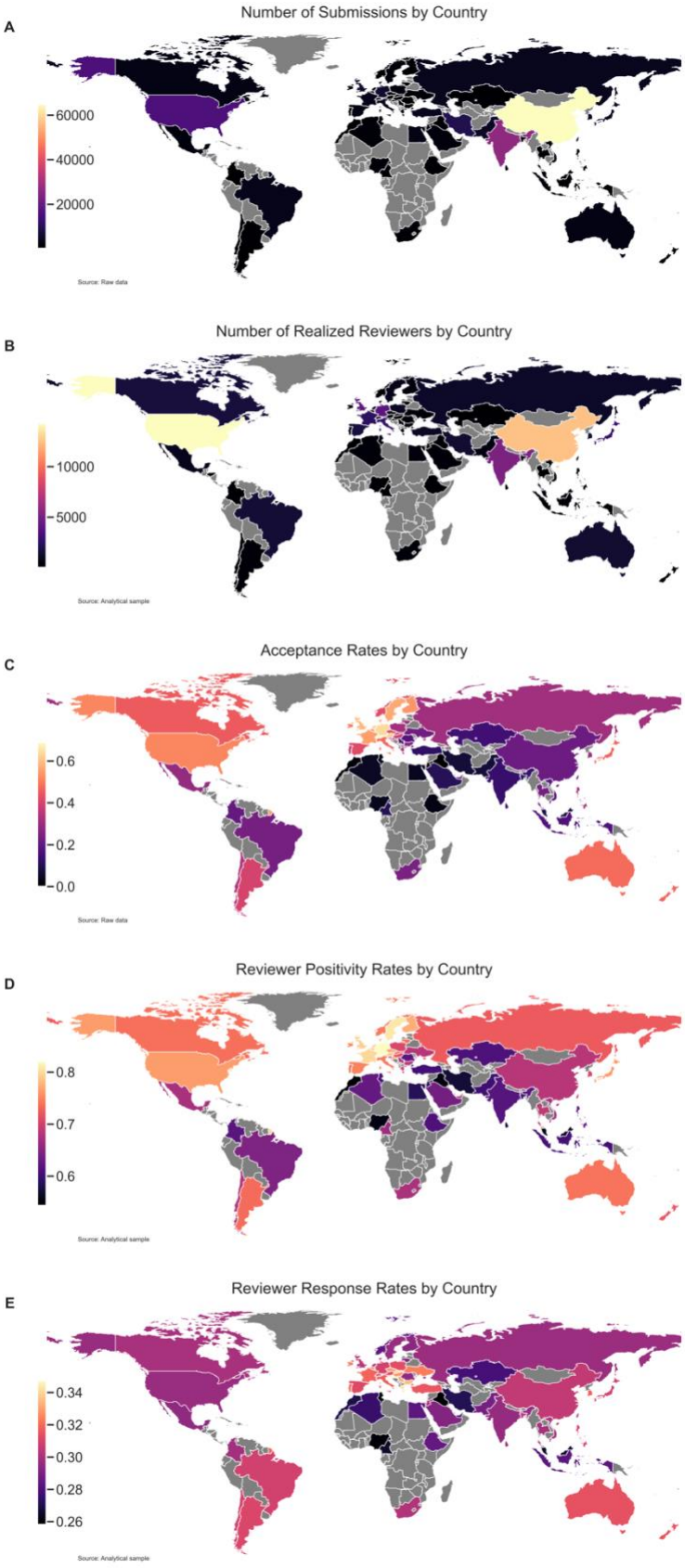

A
Number of Submissions by Country
60000
40000
20000
Source: Raw data
B
Number of Realized Reviewers by Country
10000
5000
Source: Analytical sample
C
Acceptance Rates by Country
0.6
0.4
0.2
0.0
Source: Raw data
D
Reviewer Positivity Rates by Country
0.8
0.7
0.6
Source: Analytical sample
E
Reviewer Response Rates by Country
0.34
0.32
0.30
0.28
0.26
Source: Analytical sample

Fig. S2.
Key variables by country. Countries with no data or with less than 100 submissions are in gray. **A:** Number of submissions. **B:** Number of unique realized reviewers. **C:** Acceptance rates. **D:** Reviewer positivity (recommendation to accept) rates. **E:** Reviewer response rates.

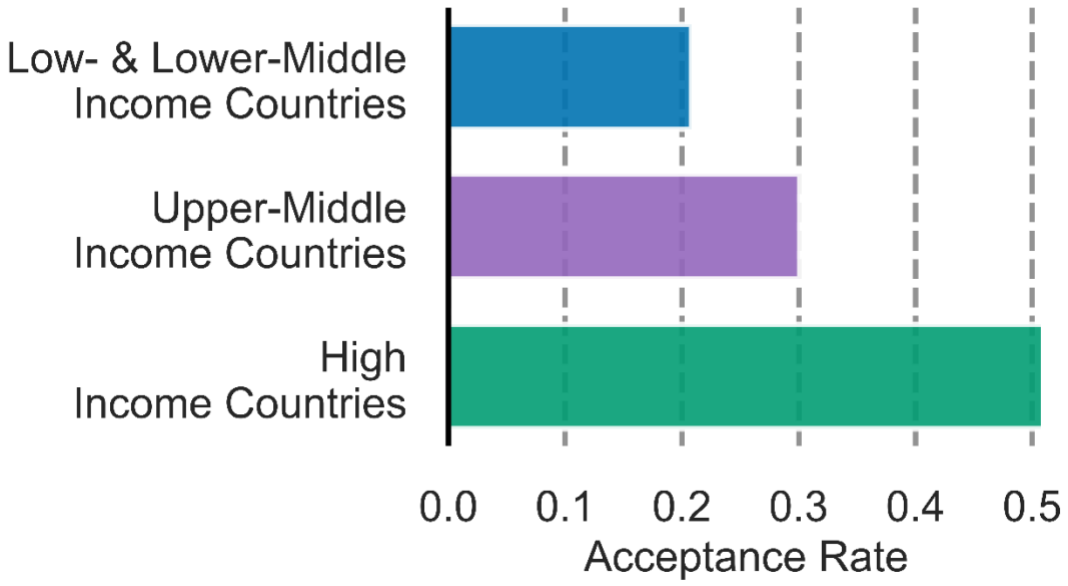


Fig. S3.
Acceptance rates by country income group from the raw data. Acceptance rate is directly related to GDP per capita.

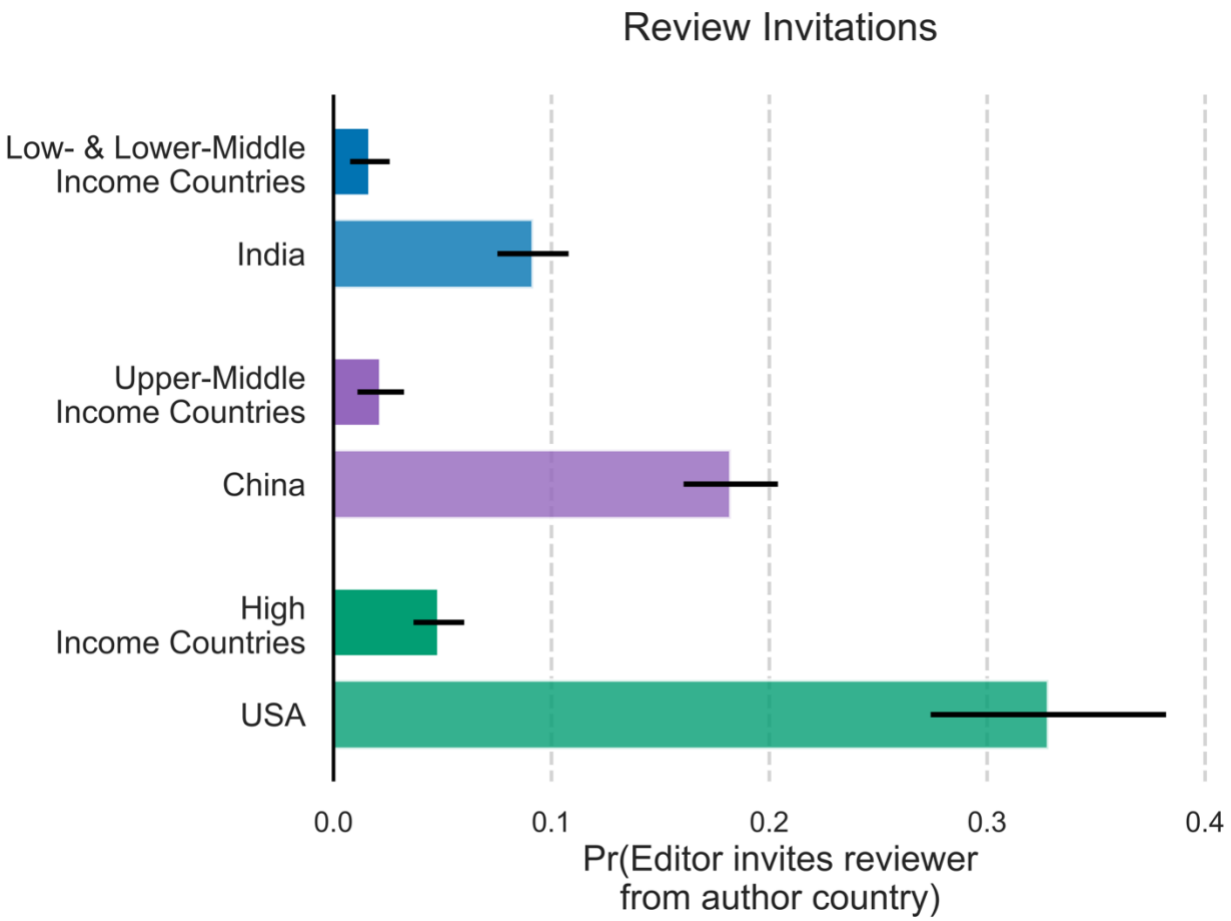


Fig. S4.
Countries well-represented in the reviewer pool are more likely to get SCRs. Wealthier countries of similar representation have more SCR access. See Equation 3B in Model Specifications for formal model. Regression estimates for SCR invitation and access by country income groups both with and without the three most represented countries is in Table S9. Editor invitation rates of SCRs by country income group, separating out the country with the most reviewers in each group. Realized reviewer counterpart is in Fig. 2B.

Table S1.
Summary statistics for data on realized reviews.

Variable	obs	unique	mean	stdev	median	min	max
Journal	238,097	60					
Manuscript	238,097	113,232					
Author Country	238,028	146					
Reviewer Country	237,562	142					
Submission Year	238,097	5	2019.86	1.35	2020	2018	2022
Decision Year	237,890	6	2020.05	1.34	2020	2018	2023
accepted	238,097	2	0.6234	0.4845		0.0	1.0
rev.positive	238,097	2	0.7040	0.4565		0.0	1.0
SCR	238,097	2	0.1467	0.3539		0.0	1.0
anonymized	238,097	2	0.0574	0.2326		0.0	1.0
national	238,036	2	0.7429	0.4370		0.0	1.0
regional	238,092	2	0.8307	0.3750		0.0	1.0
team.size	238,092		4.7942	3.3715	4.0	1.0	99.0

Table S2.
Summary statistics for data on invited reviews.

Variable	obs	unique	mean	stdev	median	min	max
Journal	787,055	60					
Manuscript	787,055	114,069					
Author Country	786,812	146					
Reviewer Country	781,508	160					
Submission Year	787,055	5	2019.78	1.34	2020	2018	2022
Decision Year	786,464	6	2019.99	1.32	2020	2018	2023
accepted	787,055	2	0.6348	0.4815		0.0	1.0
rev.responded	787,055	2	0.3025	0.4593		0.0	1.0
SCR	787,055	2	0.1194	0.3243		0.0	1.0
anonymized	787,055	2	0.0561	0.2300		0.0	1.0
national	786,855	2	0.7415	0.4378		0.0	1.0
regional	787,023	2	0.8285	0.3770		0.0	1.0
team.size	787,023		4.7911	3.3416	4.0	1.0	99.0

Table S3.
Linear probability models measuring response and Positivity bonuses of SCRs.

	Dependent Variable			
	Submitted Review		Review Positive	
SCR	0.0476*** (0.0037)	0.0474*** (0.0023)	0.0506*** (0.0061)	0.0478*** (0.0051)
Team Size Control	0.0089*** (0.0024)		0.0495*** (0.0031)	
Anonymity Control	-0.0596*** (0.0171)		-0.0348*** (0.0053)	
Auth. Country Control	Y	N	Y	N
Journal FE	Y	N	Y	N
Manuscript FE	N	Y	N	Y
Reviewer FE	Y	Y	Y	Y
Observations	787,023	787,023	238,009	238,014

Table S4.
Heterogenous estimates for benefits of SCRs across country income groups.

	Dependent Var: Review is Positive
SCR	0.0297***
	(0.0080)
SCR*UMIC	0.0361**
	(0.0116)
SCR*LLMIC	0.0104
	(0.0174)
Manuscript FE	Y
Reviewer FE	Y
Observations	237,805

Table S5.
The effect of anonymization option rollout on characteristics of submissions

Dependent Variable	ln(Team Size)	SCR on Panel	Number of Reviewers	Not Desk Rejected
PolicyAvail	-0.1697 (0.0993)	0.0022 (0.0083)	0.0072 (0.0225)	-0.0197 (0.02457)
Submission Year-Month FE	Y	Y	Y	Y
Journal FE	Y	Y	Y	Y
Observations	113,194	113,194	113,194	204,723

Table S6.
Effect of Anonymization on Benefits of SCRs.

	Dependent Var: Review is Positive		
	ITT	**OLS**	**IV**
SCR	0.0539*** (0.0081)	0.0495*** (0.0048)	0.0516*** (0.0078)
PolicyAvail	-0.0170 (0.0092)		
SCR*PolicyAvail	-0.0100 (0.0100)		
Anon			-0.0370 (0.0186)
SCR*Anon		-0.0274 (0.0180)	-0.0067 (0.0204)
Team Size Control	0.0490*** (0.0046)		0.0486*** (0.0044)
Auth. Country Control	Y	N	Y
Journal FE	Y	N	Y
Manuscript FE	N	Y	N
Reviewer FE	Y	Y	Y
Invite Year-Month FE	Y	N	Y
Observations	238,009	238,009	238,009

Table S7.
Comparing Estimates using Corresponding or Lead Author to identify country of submission.

Review is Positive	Corresponding	Lead
SCR	0.0478*** (0.0051)	0.0475*** (0.0051)
Manuscript FE	Y	Y
Reviewer FE	Y	Y
Observations	238,014	237,794

Table S8.
Comparing Estimates among National, Regional, and All Teams.

Review is Positive	All Teams	National	Regional
SCR	0.0478*** (0.0051)	0.0410*** (0.0063)	0.0414*** (0.0058)
Manuscript FE	Y	Y	Y
Reviewer FE	Y	Y	Y
Observations	238,014	176,806	197,711

Table S9.
Linear probability models measuring invitation rates and access to country homophily by corresponding author geography.

	Dependent Variable			
	Invited Reviewer is SCR		Realized Reviewer is SCR	
HIC	0.1163***	0.0484***	0.1101***	0.0497***
	(0.0097)	(0.0059)	(0.0093)	(0.0075)
UMIC	0.1458***	0.0217***	0.2111***	0.0323***
	(0.0100)	(0.0054)	(0.0107)	(0.0069)
LLMIC	0.0666***	0.0167***	0.0987***	0.0266***
	(0.0061)	(0.0046)	(0.0089)	(0.0064)
USA		0.3281***		0.3027***
		(0.0276)		(0.0263)
China		0.1823***		0.2659***
		(0.0110)		(0.0111)
India		0.0916***		0.1369***
		(0.0083)		(0.0149)
Team Size Control	0.0047	0.0013	0.0096**	0.0020
	(0.0026)	(0.0019)	(0.0032)	(0.0024)
Anonymity Control	0.0118	0.0174**	0.0102	0.0154
	(0.0078)	(0.0080)	(0.0097)	(0.0099)
Journal FE	Y	Y	Y	Y
Observations	786,232	786,232	237,823	237,823

Table S10.
Comparing Access between All Teams using Corresponding Author and National and Regional Teams and All Team using Lead Author.

	Dependent Variable			
	Corresponding Author/ All Teams	Lead Author	National Teams	Regional Teams
HIC	0.1101***	0.0982***	0.1275***	0.1124***
	(0.0093)	(0.0117)	(0.0112)	(0.0101)
UMIC	0.2111***	0.1969***	0.2224***	0.2186***
	(0.0107)	(0.0105)	(0.0106)	(0.0107)
LLMIC	0.0987***	0.0844***	0.1100***	0.1056***
	(0.0089)	(0.0090)	(0.0090)	(0.0090)
Team Size Control	0.0096**	0.0094**	0.0244***	0.0148***
	(0.0032)	(0.0033)	(0.0037)	(0.0034)
Anonymity Control	0.0102	0.0101	0.0106	0.0106
	(0.0097)	(0.0113)	(0.0115)	(0.0105)
Reviewer FE	Y	Y	Y	Y
Observations	237,823	237,812	176,667	197,552

SI References

Balli, Hatice Ozer, and Bent E. Sørensen. 2013. “Interaction Effects in Econometrics.” *Empirical Economics* 45(1):583–603. doi: 10.1007/s00181-012-0604-2.

Smirnova, Inna, Daniel M. Romero, and Misha Teplitskiy. 2022. “Nudging Science Towards Fairer Evaluations: Evidence From Peer Review.”