

HOW IMPORTANT IS EDITORIAL GATEKEEPING? EVIDENCE FROM TOP BIOMEDICAL JOURNALS

Joshua L. Krieger, Kyle R. Myers, and Ariel D. Stern*

Abstract—We examine editors' influence on the scientific content of academic journals by unpacking the role of three major forces: journals' stated missions, the aggregate supply of and demand for specific topics, and scientific homophily via editorial gatekeeping. In a sample of top biomedical journals, we find the first two forces explain the vast majority of variation in published content. The upper bound of the homophily effect is statistically significant but practically much less important. Marginal changes to the composition of editorial boards do not meaningfully impact journals' content in the short run. However, we cannot rule out persistent or pervasive frictions in the publication process.

"...journal editors say that they are not tally clerks and that decisions to publish are theirs, not the reviewers'..."

— Altman, L.K. (2006). For Science's Gatekeepers, a Credibility Gap.

The New York Times.

I. Introduction

BECAUSE publications are the prime currency in science, editors of academic journals have the potential to influence the allocation of research grants (Ginther et al., 2018), career trajectories (Way et al., 2017), downstream inventions (Bryan & Ozcan, 2021), and consumer behavior (Oster, 2020). Against this backdrop, it has long been noted that premier scientists may generate distortions in the diffusion of knowledge because of their own experiences, preferences, and/or beliefs (Merton, 1973; Dasgupta & David, 1994; Stephan, 1996). However, in the case of journal editors, these individuals are still subject to institutional and market forces such that the potential magnitude of any distortions due to "gatekeeping" remains unclear.

Typically, the concern is that gatekeepers can create inefficient or inequitable levels of homophily, conferring greater benefits to those with similar preferences, beliefs, or connections. We focus on this concern by identifying the extent to which editors induce *scientific homophily* at a sample of

top biomedical journals. We test for scientific homophily using two alternative measures. The first, our more direct but narrow measure, is based on a comprehensive vocabulary of topics that are assigned to all publications in our sample by the National Library of Medicine. The second, our broader but more indirect measure, is based on authors' affiliation bylines in their publications. Using these measures, we test how the distribution of science published in a journal covaries over time with the distribution of science pursued by the journals' editors before they became editors.¹

A long history of empirical work has found evidence suggestive of homophily at academic journals (e.g., Crane, 1967; Zuckerman & Merton, 1971; Brogaard et al., 2014). But this work focuses mostly on social connections, not scientific content, and it has been unclear how much this homophily reflects (i) the fact that journals have specific missions that reflect a persistent preference for certain topics, (ii) aggregate trends in the supply of and demand for certain topics, or (iii) editorial gatekeeping *per se*.

Our empirical analyses focus on disentangling these three forces—which we refer to as "missions" (journals' stable revealed preferences), "markets" (aggregate supply and demand shifts), and scientific homophily—by evaluating how the natural churn of editors at fifteen top biomedical journals affects the content published in those journals. We focus on variation in editors' idiosyncratic research backgrounds and identify a plausible upper bound of how much editors may steer journals towards the topics that they themselves have studied. We examine publishing in the biomedical sciences for two reasons. First, biomedical research is an especially important sector of the economy, and one in which the precious few slots in top journals for featuring research projects may have significant downstream impact. Second, this setting allows us to use validated topic modeling tools, which we supplement with new hand-collected data on editor tenures.

Our findings align with prior research in that we estimate a statistically significant homophily effect using both of our measures of scientific content. As the share of editorial experience in a topic increases, so too does the share of the journals' publications on that same topic. We cannot separate how much of this effect is due to the supply of submissions (by authors) versus the selection thereof (by editors). Furthermore, we observe "pre-trends" where journals appear to bring scientists on board as editors just as the content of

Received for publication November 8, 2021. Revision accepted for publication January 9, 2023. Editor: Pierre Azoulay.

*Krieger, Myers, and Stern: Harvard Business School.

The authors are grateful to Bryce Turner and Mats Terwiesch for excellent research assistance. John-Paul Ferguson, Michael Ransom, Carolyn Stein, Brad Wible and participants at the HBS Health Care Initiative Faculty Research Seminar, McGill Organizational Behavior Seminar, the Western Economic Association annual meeting, and the NBER Science of Science Funding meeting provided helpful comments. The authors are grateful for research support from the Harvard Business School Division of Faculty Research and Development. A.D.S. acknowledges support from the Kauffman Foundation. All errors are our own.

A supplemental appendix is available online at https://doi.org/10.1162/rest_a_01340.

¹Another concern is related to the distribution of opportunities across genders and socioeconomic status. We lack the data to investigate these features; for work focusing on the economics discipline, see Laband and Piette (1994), Brogaard et al. (2014), Card et al. (2020), or Carrell et al. (2024) and references therein.

the journal is moving towards the topics those scientists specialize in. Thus, we view our estimate as an upper bound of the net effects of these forces, which is relevant for predicting the effect of policies that affect editorial turnover (e.g., tenure limits). We believe this upper bound is informative because, despite the prevalence of narratives asserting that “certain editors only like certain types of papers,” there is little prior empirical evidence on the practical magnitude of such an effect.

Still, while we can reject a null hypothesis of no scientific homophily, the role of such gate-keeping is dwarfed by the importance of the mission- and market-based forces. The homophily effect typically explains less than three percent of the variation in published content that is not already explained by missions and markets. Conversely, missions and markets often explain anywhere from forty to eighty percent of the variation in published content that is not explained by scientific homophily. This pattern holds through a battery of robustness tests and alternative specifications. Additionally, we find that the change in article composition induced by the homophily channel leads to slightly fewer citations to the journal, which suggests editors trade off content for impact (as proxied by citations).

To better understand the practical implications of our point estimates, we perform a series of simulations we refer to as “editorial takeovers.” We simulate the replacement of editors at one journal with those from another, and use our estimates of the homophily effect to predict how much closer (in terms of topics published) these editor replacements and the ensuing associated gatekeeping might bring two journals. This exercise indicates that, even within our diverse sample, nontrivial changes in the composition of an editorial board would not meaningfully alter journals’ scientific content.

Our article is most closely related to Li (2017), which analyzes how the composition of National Institutes of Health (NIH) grant review panels influences the direction and quality of their funding decisions. Li (2017) shows that greater scientific homophily between the evaluator and a project can lead a scientist to be both better informed and more biased about the quality of the project. Our article complements this work by illustrating the aggregate effect of such forces in the market for publications. Despite being in a very different setting, our analysis is also related to work in media markets that investigates how much of the slant in newspapers is due to ownership’s preferences versus other factors such as readers’ demand (e.g., Glasser et al., 1989; Gentzkow & Shapiro, 2010). In a similar vein as Gentzkow and Shapiro (2010), our results show that the vast majority of the variation in slant is driven by the pull of the readership as opposed to owners pushing their own agendas.

Overall, our analyses lead us to conclude that editors do favor articles from researchers in their scientific club—those who share their research interests and affiliations. However, the magnitude of the homophily effects we identify suggest that a relatively small portion of a journal’s pages are captured by articles steered into that journal due to edito-

rial idiosyncrasies. Further, our analyses show that failing to account for the role of missions and markets would significantly overstate these homophily effects. We cannot test whether or not certain topics go unpublished because *all* editors “dislike” those topics; our results do not rule out persistent or pervasive frictions in the peer-review or editor selection. We highlight this and other limitations of our approach throughout the article, and we consider the broader implications of our findings for future work in the Discussion.

II. Background, Data, and Summary Statistics

A. Biomedical Research Publishing and Editors

The publication process in biomedical sciences follows most of the contemporary norms for peer-reviewed academic work.² In general, new editors are selected by existing editors. The *New England Journal of Medicine (NEJM)*, which is included in our sample of journals and is one of the oldest and most prestigious biomedical journals, provides a helpful summary of editor selection and their duties:

“...editors are chosen for their expertise in major areas of medicine. Associate editors play central roles in managing the peer review process and in decisions to accept or decline manuscripts for publication in *NEJM*. In addition to their work for *NEJM*, they also hold full-time positions at academic medical centers.”

Editorial responsibilities can roughly be categorized into three parts: selection of which submissions to advance to peer-review; selection of reviewers; and publication decisions given reviews. All of these choices are mechanisms through which editors can influence the content of their journals. As we describe below, we focus only on editors at these journals who have previously published papers themselves; this necessarily excludes any nonscientific editorial positions from our analyses (e.g., copyeditors).

B. Sample of Top Biomedical Journals

Appendix A describes how we constructed our sample of fifteen top biomedical journals,³ which span a range of general and specialized disciplines and are listed in panel a of table 1 and described in further detail in table A1. On average, we have about 30 years of data per journal. Since our data collection efforts centered on 1985 as the earliest year, we use that year as the earliest date for all of our main analyses. The average tenure length for all editors in our sample

²Article review times are often on the order of weeks to a few months. For more, see www.nejm.org/media-center/publication-process.

³Our sample deliberately excludes journals with professional, full-time editors (e.g., *Nature*, *Science*) because of their rarity and the challenges of clearly measuring those editors’ scientific expertise as they have highly varied publication records.

TABLE 1.—SUMMARY STATISTICS

Panel (a): Sample journal list								
<i>Am. Rev. Respir. Dis.</i>								
<i>Ann. Intern. Med.</i>								
<i>Circ. Res.</i>								
<i>J. Am. Coll. Cardiol.</i>								
<i>N. Eng. J. Med.</i>								
<i>Anesth. Analg.</i>								
<i>Ann. Surg.</i>								
<i>Circulation</i>								
<i>J. Nucl. Med.</i>								
<i>Radiology</i>								
<i>Anesthesiology</i>								
<i>Br. J. Anaesth.</i>								
<i>Hum. Brain Mapp.</i>								
<i>JAMA</i>								
<i>Stem Cells</i>								
Panel (b): Journal & editor statistics (Journal-Year obs.)								
	Obs.	Mean	SD	Min.	p25	p50	p75	Max
Year	336	1996.9	6.9	1985	1991	1997	2003	2008
Editors								
Num. Matched Eds.	336	57.8	51.8	7	25	38	76	278
Eds'. Current Tenure	336	5.70	2.51	0.00	4.11	5.19	6.59	14.88
Pubs. per Ed.	336	72.4	27.5	33.0	49.2	66.7	95.5	151.1
Share of MeSH Tree	336	0.302	0.098	0.097	0.224	0.306	0.364	0.552
Journals' Pubs.								
Pubs.	336	563.7	323.7	36	336	500	717	1,476
Share of MeSH Tree	336	0.147	0.052	0.018	0.110	0.144	0.182	0.294
Panel (c): MeSH share statistics (MeSH-Journal-Year obs.)								
	Obs.	Mean	SD	Min.	p25	p50	p75	Max
Journals' Pubs.								
$\mathbf{1}\{S_{mjt}^P\}$	1,954,148	0.154						
S_{mjt}^P if > 0	302,652	$1.11e^{-3}$	$4.40e^{-3}$	$1.74e^{-5}$	$1.55e^{-4}$	$3.22e^{-4}$	$7.81e^{-4}$	0.157
Editors' Pubs.								
$\mathbf{1}\{S_{mjt}^E\}$	1,954,148	0.318						
S_{mjt}^E if > 0	620,759	$5.41e^{-4}$	$2.71e^{-3}$	$9.35e^{-7}$	$3.38e^{-5}$	$1.01e^{-4}$	$3.31e^{-4}$	0.107

Panel a lists the 15 in-sample journals per their ISO 4 abbreviations. Panel b reports summary statistics for the editors and journals in our data. Editors' Publications per Editor reports the average number of publications per editor at the journal, counting only publications prior to the editors start. Share of MeSH Tree reports the fraction of the entire MeSH tree used in our main analyses, which is comprised of approximately 6,100 terms, with nonzero publications per year (i.e., a value of 0.1 indicates that, on average, the publications in a journal span 10% of the MeSH tree each year). Panel c reports statistics for the MeSH-journal-year (mjt) share variables, including the fraction that are greater than zero ($\mathbf{1}\{\}$) and the distribution of shares conditional on being nonzero (if > 0), which are approximately log-normal.

is 7 years (SD = 7 years), and, because of the skewness in tenure lengths, the average active editor is in their sixth year of tenure in our data.

C. Publication and Author Data

Our primary source of biomedical publication data is the "Author-ity" dataset developed by Torvik and Smalheiser (2009). This dataset offers a disambiguated publication history for virtually all authors with publications in the National Library of Medicine's PubMed Database until 2008, and it has been shown to have very high precision and recall (Lerchenmueller & Sorenson, 2016). We supplemented this with additional publication data sourced directly from PubMed as needed. Summary statistics of editors we are able to match to their publication record are displayed in panel b of table 1.

D. Categorizing Topics with Medical Subject Headings (MeSH)

To index topic space—the scientific content of publications—we make use of the National Library of Medicine's Medical Subject Heading (MeSH) system. MeSH is "a controlled and hierarchically organized vocabulary," and perhaps the most widely used classification system for biomedical research.⁴ All articles in PubMed are

assigned MeSH terms. Importantly, these assignments are performed by the National Library of Medicine's algorithms and staff, not authors.

The MeSH system's specificity and hierarchical organization makes it a valuable tool. It presents topics as discrete, mutually exclusive, nested components. For example, both *Diabetes Mellitus* and *Adrenal Gland Diseases* are subterms under the broader category of *Endocrine System Diseases*. MeSH terms also capture the study's scientific approach. For example, many articles are tagged with MeSH terms related to the organisms, methodologies, clinical activities, and interventions involved.

Panel b of table 1 reports the share of the MeSH tree that appears in the publications (1) previously written by editors and (2) in journals. On average, each journal spans roughly 15% of the tree, whereas the editors at these journals have published on a wider swath of topics, covering about 30% on average. Our focus in this article is testing whether or not it is the case that editors steer their journals towards the portions of the MeSH tree that they themselves have previously focused on.

An important caveat to our topic modeling approach is that we are treating papers as purely a combination of MeSH terms—if two papers have the same MeSH terms then they are "equal" for our purposes. This prevents us from speaking to whether or not editors may have preferences over other features of papers conditional on their MeSH content (e.g., risk, novelty, clarity, implications, etc.). In short,

⁴For more, see www.nlm.nih.gov/mesh/meshhome.html.

MeSH provides a useful tool for indexing the “Background” and “Methods” sections of papers, but it is much less useful for investigating the “Conclusions” of a study. In section IVD, we report additional results where we use the authors’ affiliation bylines as an alternative proxy for the scientific community and institutional origins of papers, which yields very similar results.

III. Empirical Framework

To begin, consider a simple test for the presence of scientific homophily that relates the share of publications on MeSH topic m at journal j in year t , S_{mjt}^P , to the same share measure based on editors’ publications prior to their start at the journal, S_{mjt}^E ⁵:

$$S_{mjt}^P = \alpha + S_{mjt}^E \beta + \epsilon_{mjt}, \quad (1)$$

where β is the focal parameter describing the scientific homophily effect—the extent to which editors are more likely to publish articles related to the topics they have previously studied. Once we introduce further controls and assumptions (described below), we take the statistical significance of our estimate, $\hat{\beta}$, as evidence that gatekeeping exists.

Additionally, we will rely on R^2 and partial- R^2 statistics to understand how much variation in content within and across journals can be attributed to different forces.⁶ Beyond that, we use simulations to better convey the practical magnitude of any homophily effect. For the purposes of describing our empirical framework, we focus on our direct measure of scientific content based on MeSH terms. The same logic holds when using our more indirect measure of science based on the words that appear in authors’ affiliation bylines.

We refer to S_{mjt}^E as capturing editors’ “experience” or “background” to reflect pure content of this measure—it says nothing about the quality of editors’ prior publications. This variable will thus reflect some combination of editors’ skills, expertise, their preferences, or any other force that previously influenced the direction of their research (e.g., funding opportunities as in Myers, 2020 and Hegde & Sampat, 2015 or competitive pressures as in Hill & Stein, forthcoming).

Panel c of table 1 reports the distributions of these share measures (S^P , S^E). Clearly, a large number of zeros are in our data. On average each year, journals publish articles related to about 15% of the more than 6,000 MeSH terms we use in our preferred data construction. Likewise, editor’s publications only cover about 30% MeSH topics at each journal in a year. Given this sparsity, we estimate multiple versions of equation (1) using linear, binary, and log

transformations of the share variables to investigate intensive and/or extensive margin effects.

Our preferred construction of S_{mjt}^E is based on the aggregated publication records of all current editors prior to the start of their tenure. This approach implicitly puts more weight on the publication records of individuals with more publications. As reported in the appendix, we obtain similar results if we calculate the topic shares first at the individual-level and then take the average of those values, which gives each editor equal weight (see appendix figure D6).

In appendix C, we provide an alternative approach that recasts our data from this MeSH-focused level to one at the individual editor level. This alternative approach focuses on cosine similarity scores and how they vary across editor-journal pairs over time. The results from this approach are very consistent with the main specifications reported here—the role of missions and markets dwarfs any effect that arises from scientific homophily.

A. Journal Missions and Topic Markets

Our goal is for the focal parameter in equation (1), β , to capture only the scientific homophily effect that describes how editors affect published content for idiosyncratic reasons related to their own background. This requires an understanding of what generates variation in the composition of editors across and within journals since we lack any systematic sources of exogenous variation in editors. In particular, two important forces are in play: journal’s missions to focus on particular topics, and time-varying, aggregate (i.e., across all journals) trends in the supply of and demand for research on certain topics.

Journals’ missions. All journals in our sample, and virtually all academic journals, have “missions” of varying scope. For example, consider the following excerpts from the latest mission statements or mastheads of a few journals in our sample:

Anesthesia & Analgesia: “articles on the latest advances in drugs, preoperative preparation, patient monitoring, pain management, pathophysiology, and many other timely topics”

Stem Cells: “laboratory investigations of stem cells and the translation of their clinical aspects”

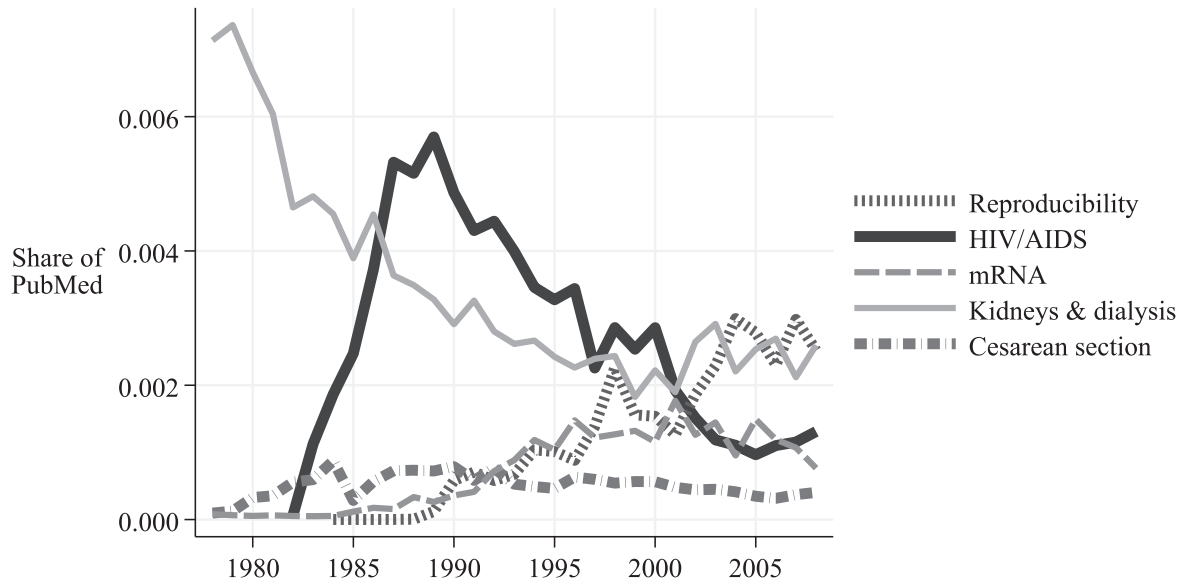
Annals of Surgery: “contributions to the advancement of surgical science and practice”

Clearly, given their titles, these journals vary in terms of the topics of articles they publish. Notably, these missions are extremely persistent over time. Consider, for example,

⁵We focus on editors’ publications prior to their start on a board to prevent feedback effects whereby editors’ publications might be influenced by the work they evaluate.

⁶The partial- R^2 for a given independent variable x is defined as $(R^2 - R_{-x}^2)/(1 - R_{-x}^2)$, where R^2 corresponds to the saturated model and R_{-x}^2 corresponds to the model where x is excluded.

FIGURE 1.—EXAMPLE MESH TOPIC TRENDS



Plots the annual share of the total PubMed publication record from 1978 to 2008 that is related to five example topics.

the *Annals of Surgery*'s missions from about 35 and 135 years ago:

Annals of Surgery, 1985: "articles in the field of surgery... devoted to the surgical sciences"

Annals of Surgery, 1885: "monthly review of surgical science and practice"

Of course the *Annals of Surgery* is, and has always been, focused on publishing research on surgery. This may seem trivial, but it has implications for our empirical approach. When current editors are choosing their replacements, we should expect them to choose experts on the same topics they themselves are experts on, and for these editors, in turn, to choose to publish new papers on similar topics as before. We would not classify this persistence as any kind of editor-specific gatekeeping *per se*. Empirically speaking, this means that our model should account for the unique, persistent preference that each journal (j) has for each MeSH topic (m) given its mission.⁷

Topic markets. The notion of "topic markets"—the aggregate supply of and demand for research on certain topics—is also relevant. Over time, the costs of studying any given topic will fluctuate as science progresses or stalls. Likewise, the benefits of new knowledge on a topic will fluctuate with the preferences of the consumers of these papers (e.g., clinicians, other academics, funders, policy-makers).

⁷As shown below, we use a fixed effects approach to capture variation due to these missions rather than attempting to infer anything directly from the mission statements. The examples here illustrate what the data reveals: journals have highly persistent preferences for certain topics.

As shown in figure 1, which plots the PubMed-wide trends for a few select MeSH topics, these trends can vary widely in any year and over time.

Thus, when editors choose their replacements, we should expect them to form expectations about future supply and demand and choose experts on topics that they anticipate will be germane in the near future. This implies that we could observe a correlation between editors' backgrounds and publications in their journal simply because of trends in topic markets. Empirically speaking, this means that our model should account for aggregate changes/shocks that are common to all journals, but unique to MeSH topics (m) and time-varying (t).

B. Main Regression Model

To account for the role of journals' missions and topic markets, we modify equation (1) to include two vectors of fixed effects at the MeSH-journal (mj) level to account for missions and at the MeSH-year (mt) level to account for markets. This yields our main regression model:

$$S_{mjt}^P = \underbrace{\gamma_{mj}}_{\text{journal's mission}} + \underbrace{\sigma_{mt}}_{\text{topic markets}} + \underbrace{S_{mjt}^E \beta}_{\text{editors' scientific homophily}} + \epsilon_{mjt}. \quad (2)$$

Even with this rich set of fixed effects, there is still variation in editor composition, largely due to the term limits associated with most editorial positions (as well as any idiosyncratic events that lead to churn in editors). In our sample, researchers move in and out of these positions fairly frequently; roughly 18% of editors are new to a journal in a given year.

C. Connection to a Model of Demand for Editors and Content

While relatively intuitive, equation (2) does not obviously reflect a model of the choices that generate our data. However, in appendix B we follow the discrete choice theory of Berry (1994) closely and show that equation (2) approximates a stylized logit model of the aggregate demand for publications and new editors by current editors. In this model, we consider two “products” that current editors demand: (i) the MeSH topics of new papers to be published, and (ii) the MeSH topics of prior papers written by researchers who might be chosen as new editors. It is these two choices that generate variation in our independent (S^E) and dependent (S^P) variables. We relate the mean utility of these two products to each other, while incorporating the same mission- and market-based forces, and arrive at an empirical model that is analogous to equation (2).⁸

D. Estimating the Upper Bound of the Scientific Homophily Effect

The fixed effects included in equation (2) likely eliminate most of the endogenous variation in our data. However, we are concerned that there still may be unobservable “local” shocks—unique at the *mjt*-level observation—that codetermine both publication outcomes and editor selection. For instance, consider the large surge in AIDS-related research illustrated in figure 1. While our MeSH-year fixed effects (σ_{mt}) will remove this aggregate trend from the data, it surely must have been the case that this surge was more relevant for certain journals (e.g., the general interest, clinically focused *Journal of the American Medical Association*) compared to others (e.g., *Anesthesiology*) and in turn we would expect these journals to have responded differentially to this event.

We have no way of accounting for these sort of local shocks, but, presumably, editors do. Thus, we assume that existing editors tend to choose new editors who, if anything, have more experience in topics where such shocks are more positive (e.g., due to an upward surge in supply or demand for the topic by the journal). If correct, this implies a non-negative correlation between S^E and ϵ , even conditional on our fixed effects.⁹

In appendix C, we perform an editor-level analysis and find clear evidence of pre-trends whereby the similarity between a (future editor) scientist and a journal increase markedly up to the point at which they become editors (see figure C3). This is consistent with a positive correlation between S^E and ϵ and could bias our estimate of β upwards; for this reason, we view our estimates of the scientific homophily effect as upper bounds.

⁸This demand model specifically motivates a regression using log transformations of the share variables, which we report in many results tables.

⁹In the demand model described in appendix B, we formalize these local shocks and the assumption about nonnegative selection.

IV. Results

A. Main Results

Table 2 displays the main results, with three panels that correspond to regressions based on (a) linear, (b) binary, and (c) log transformations of the focal share variables. For all panels, column 1 reports the results from regressions with no fixed effects, columns 2–3 include the mission and market fixed effects (respectively), and column 4 includes both vectors of fixed effects.

In all three models, the naive regressions shown in column 1 suggest elasticities between 0.5 and 1.0 and editors’ research backgrounds alone can explain a large portion of the content published in their journals; the R^2 values range from 0.2 to 0.8. As a starting point, these magnitudes suggest that editor’s preferences for scientific homophily may shape the content of their journals.

However, the journal-mission and topic-market fixed effects each independently absorb a large amount of variation in the data as seen in columns 2–3. In all models, both the point estimate and partial- R^2 of the scientific homophily effect decrease significantly when either of the controls are included.

In our preferred specifications, shown in column 4, we obtain elasticity estimates of roughly 0.230 with the linear model, 0.044 with the binary model focusing on extensive margins, and 0.046 with the log model focusing on intensive margins. Our estimated confidence intervals of these elasticities span from a low of 0.038 (with the log model) to a high of 0.287 (with the linear model).

The lower rows of each panel in table 2 report a set of R^2 statistics for each regression. Focusing on the partial- R^2 corresponding to the scientific homophily effect, we observe a pattern similar to what happens with the point estimates of the coefficients. With both sets of fixed effects are introduced (column 4), the partial- R^2 decreases by almost two orders of magnitude to 0.024 for the linear outcome measure (panel A) and approximately 0.001 for the binary and log outcome measures (panels B and C). Conversely, missions and markets appear to be very good at explaining what is published in these journals—the partial- R^2 of these factors are in the range of 0.12 to 0.65. And because a significant amount of the variation in editors’ own research occurs along these same dimensions, not accounting for missions or markets appears to severely overstate the implied importance of scientific homophily due to gatekeeping.

Interpreting magnitudes. While the magnitudes associated with the scientific homophily effect reported above are statistically significant, they also appear to be rather small in our preferred specifications—certainly relative to the naive regressions with no controls. Recall that we hypothesize these magnitudes reflect an upper bound of the true effect. However, it is still difficult to gauge any practical importance because we do not have a benchmark for what is “large” versus “small” in this context.

TABLE 2.—MAIN RESULTS

	(1)	(2)	(3)	(4)
Panel (a): D.V. = S_{mjt}^P				
S_{mjt}^E	1.044 (0.054)	0.245 (0.038)	0.867 (0.038)	0.230 (0.029)
Elasticity at means	1.043 (0.041)	0.245 (0.038)	0.867 (0.038)	0.230 (0.029)
Total R^2	0.822	0.943	0.864	0.952
partial- R^2 , mj -missions		0.680		0.646
partial- R^2 , mt -markets			0.238	0.155
partial- R^2 , sci. homophily	0.822	0.029	0.457	0.024
Obs., mjt	1,953,818	1,953,818	1,953,818	1,953,818
Panel (b): D.V. = $\mathbf{1}\{S_{mjt}^P > 0\}$				
$\mathbf{1}\{S_{mjt}^E > 0\}$	0.353 (0.005)	0.065 (0.002)	0.129 (0.002)	0.021 (0.001)
Elasticity at means	0.723 (0.004)	0.133 (0.004)	0.265 (0.004)	0.044 (0.003)
Total R^2	0.206	0.556	0.420	0.611
partial- R^2 , mj -missions		0.441		0.323
partial- R^2 , mt -markets			0.269	0.115
partial- R^2 , sci. homophily	0.206	0.004	0.023	0.001
Obs., mjt	1,953,818	1,953,818	1,953,818	1,953,818
Panel (c): D.V. = $\log(S_{mjt}^P)$				
$\log(S_{mjt}^E)$	0.530 (0.016)	0.089 (0.007)	0.466 (0.007)	0.046 (0.004)
Elasticity at means	0.530 (0.016)	0.089 (0.007)	0.466 (0.007)	0.046 (0.004)
Total R^2	0.398	0.725	0.601	0.803
partial- R^2 , mj -missions		0.543		0.499
partial- R^2 , mt -markets			0.338	0.274
partial- R^2 , sci. homophily	0.398	0.006	0.241	0.001
Obs., mjt	224,224	224,224	224,224	224,224
Incl. γ_{mj}		Y		Y
Incl. σ_{mt}			Y	Y

Standard errors are in parentheses and are clustered at the MeSH (m) level. Elasticities are reported at the means of the independent and dependent variables, with standard errors calculated via the delta method. The bottom two rows apply to all panels and indicate which columns are based on specifications that include MeSH-journal (mj) and/or MeSH-year (mt) fixed effects.

One way to consider the importance of the scientific homophily effect is to focus on the corresponding partial- R^2 statistic. As noted above, editors' backgrounds explain only a tiny share of variation in published content. This small value could be because either there is a truly large scientific homophily effect, but the observed variation in editors' backgrounds is so small that the effect does not generate sizable variation in content, or the effect is in fact small. The data suggest the latter: as reported in the summary statistics of panel c table 1, there is actually more variation in MeSH shares for editors (S_{mjt}^E) than there is for journals' publications (S_{mjt}^P).¹⁰

B. Alternative Specifications and Heterogeneity

In appendix D, we present a number of additional results based on alternative specifications. We implement alternative controls for missions and markets, change our definitions of the risk sets (i.e., which journal-MeSH pairs are feasible), explore different levels of aggregation of the MeSH

hierarchy, and test alternative ways of aggregating editorial boards. The magnitudes and patterns documented in our main results persist across all specifications, suggesting that no single decision related to our data construction or estimation are leading to spurious results.

Appendix D also explores a number of dimensions of heterogeneity to further explore the nature of the homophily effect we identify. First, to get some sense as to whether the effect is driven more by an editors' "expertise" versus their "preferences" (although the two are surely correlated), we report results after separating publications into two groups that should reflect these two forces: research (i.e., peer-reviewed articles) and nonresearch (e.g., editorials, comments, letters) articles. The homophily effect we are identifying appears to be almost entirely driven by editors' research publications.

We also investigate the extent to which the homophily effect may be changing over time, and we find evidence that the magnitude of the effect is declining over time to the point of losing statistical significance (see figure D4). Last, motivated by "novelty bias" identified in prior research (Boudreau et al., 2016), we test whether editors appear to induce more or less homophily when focusing on "new(er)"

¹⁰The coefficient of variation of these nonzero distributions is 25% larger for S_{mjt}^E than S_{mjt}^P .

topics. We investigate this question by proxying for a topic's age based on the year it was introduced into the MeSH hierarchy, and allowing the homophily effect to be a function of this age. Our results yield no conclusive evidence that the homophily effect is stronger (or weaker) for newer (versus older) topics.

Appendix C describes an alternative approach to estimating the impact of scientific homophily by constructing an editor-level data frame. Rather than treating the editorial board as a single entity, we use editor-specific publication histories and a MeSH-based cosine measure to construct an editor-journal specific scientific similarity score. We then estimate how the similarity between editor-journal pairs changes during one's editorial tenure. The results corroborate our main findings. When scientists become editors, the similarity between their own research and what is published in their journal increases, but the magnitude of this effect and the amount of variation in the data explained by it is very small. Furthermore, appendix figure C3 illustrates how the (small) effects we estimate could instead be plausibly explained by unobservable time varying trends.

C. Proxying for Quality with Forward Citations

Given that we identify a nonzero scientific homophily effect, an obvious follow-on question is whether the content marginally steered into these journals and associated with this force is of differential quality relative to what would have been published in its absence.

As is customary, we proxy for publication impact using forward citations (scaled by the year of publication) and estimate citation-weighted versions of our main regressions. Table D9 reports these citation-weighted results alongside the unweighted regressions. In all cases, the citation-weighted effects are smaller than the unweighted effect, which indicates that this new content obtains fewer citations relative to the additional space in the journal it obtains. Under the admittedly generous assumption that editors are able to forecast these citations accurately, the magnitudes imply that editors are willing to accept about a 2%–5% decline in (expected) forward citations for every 10% increase in proximity to their own expertise. Yet whether this truly reflects any sort of welfare loss is unclear.

At first glance, a willingness to trade off content for impact appears contrary to Li (2017), the results of which suggest that articles published via the homophily channel should receive *more* forward citations due to editor expertise in those articles' subfields. We do not view these results as inconsistent for three reasons. First, our estimates incorporate both demand-side (editors choosing submissions) and supply-side effects (submissions from authors), whereas the supply is essentially fixed in Li (2017). Further, if authors believe they can capitalize on taste-based discrimination, they may employ an even "lower bar" in deciding whether or not their paper is suitable for submitting to a journal with a related editor. Second, we hypothesize that the

scale of variation in topic overlap in our data is substantially larger than in the data used by Li (2017). Third, the measure of homophily in Li (2017) relies on citations between reviewers and applicants, and therefore mixes both social and intellectual relatedness. By contrast, our approach separates topic/intellectual similarity from social/professional connections.

While forward citation counts provide an intuitive and simple proxy for scientific value, work such as Wang et al. (2017) has shown that novel science tends to have a delayed accumulation of citations. Thus, our finding could easily be rationalized by editors having more private information about research more closely related to their own expertise that will be valuable in the longer run.

D. Alternative Approach: Affiliation-Based Topic Modeling

As an alternative approach to indexing scientific topic space, we make use of the affiliation bylines in authors' publications. Thankfully, the "Author-ity" database (Torvik & Smalheiser, 2009) contains data on the top twenty most common keywords that appear in each authors' publications. These keywords reflect some combination connections that may be of geographic (e.g., state names), institutional (e.g., university names), and scientific (e.g., department names). Thus, they may reflect a more holistic index of the content of a given authors' publications and should capture many of the more nuanced ways in which two researchers may belong to the same scientific club.

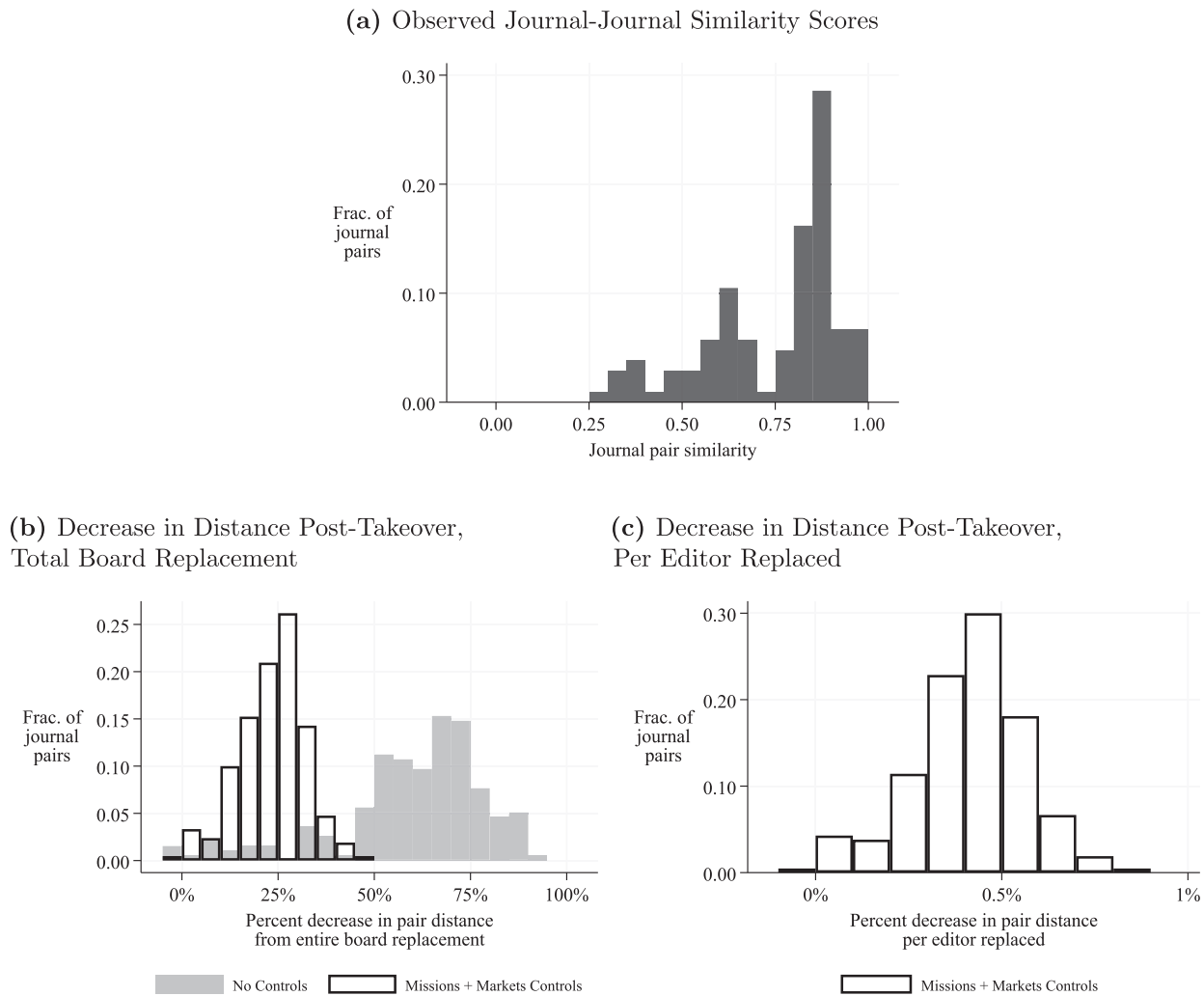
Table D10 recreates our main results table (table 2) using these affiliation terms. Interestingly, we obtain very similar results in all regards. The similarity in these results may reflect the fact that both this and the MeSH-based approach are capturing the same phenomena. But, given the large role of institutions in this affiliation-based approach, it may also be because the scientific and institutional homophily preferences of editors are of roughly the same size. Regardless, the results again suggest that the nature of homophily we are isolating is not a very large determinant of what gets published at these journals.

E. Understanding Magnitudes via Simulated Editorial Takeovers

To provide another perspective on the magnitude of our estimates, we perform a series of simulated editorial "takeovers" to explore how much new editors might alter the content of a journal. This exercise is meant only to provide an illustration of these magnitudes in a more tangible framing.

We begin this exercise by first estimating the scientific similarity of all possible pairs of journals in our data. We use the cosine similarity metric based on the MeSH terms, which yields a number ranging from zero to one, with larger values indicating higher similarity. Figure 2a plots the distribution

FIGURE 2.—JOURNAL SIMILARITY, OBSERVED AND AFTER SIMULATED “TAKEOVERS”



Panel a plots the distribution of all pairwise similarity scores between journals based on the cosine similarity of the average rate of MeSH term appearance in their respective publications. Panel b plots the decrease in the distance (the inverse of similarity) between two journals after replacing an entire editorial board and then only allowing the homophily effect (estimated from a regression including no controls or both markets and missions controls) to alter the content of the journal. Panel c reports the same changes as panel b, but scaled by the average number of editors at each journal to estimate the distance decrease per editor.

of these observed similarities, which range from about 0.25 up to nearly 1 (mean = 0.75, SD = 0.18). As a test of face validity, the two most similar journals in our sample are, reassuringly, *Anesthesiology* and *Anesthesia & Analgesia*.

We then iterate through each of the fifteen journals in our sample and (i) replace the entire editorial board at the other fourteen journals with the editors at the focal journal, (ii) estimate the change in the published content that would be expected given our estimates of the scientific homophily effect, and finally (iii) estimate the similarity between the (unchanged) focal journal and the fourteen other (changed) journals. We then calculate the change in the similarity post-takeover and scale these changes by the number of editors replaced to obtain effect sizes per editor. To be as generous as possible, we perform these simulations using the point estimates of the homophily effect from our linear model, panel a of table 2, which yielded one of the largest estimates across our specifications.

Figure 2b plots the results of these takeover simulations. We plot the percent decrease in the distance (the inverse of similarity) between two journals content post-takeover. Using the large estimates from the model with no controls yields an average decrease in the distance of the journals post-takeover of approximately 60% (SD = 20%), which is roughly equivalent to 0.9 standard deviation of the baseline differences across journals.

When using the homophily estimate from our preferred specification, the average decrease in distance post-takeover is only 24%, or 0.3 standard deviations of baseline differences. This is a relatively small change considering that the entire editorial board is replaced. In terms of a per-person effect—levels of editorial churn we see in practice—figure 2c plots the same changes, but scaled by the number of editors being replaced. Here, each editor is responsible for bringing the two journals closer together by only about 0.4% or 0.005 standard deviation of the baseline differences

across the journals. This exercise indicates that incremental changes in editors would not have a meaningful effect on the content of a journal's publications.

V. Discussion

Overall, our results point to a scientific homophily effect that is practically small and declining over time. The role of this effect is dwarfed by the importance of journals' missions and the aggregate topic-specific trends despite the appearance of gatekeeping in the unconditional correlations.

The journals in our sample publish research that can be neatly characterized as either experimental (e.g., randomized clinical trials) or descriptive (e.g., reporting of disease prevalence). This suggests there may be fewer subjective dimensions for evaluation in this sample compared to, for example, the social sciences where quasi-experimental methods with more subjective criteria are prevalent. Replicating our analyses in other fields could prove useful for understanding how we evaluate and disseminate science across disciplines.

An important limitation of our approach is that we cannot investigate any persistent or pervasive biases since our empirical model subsumes these effects into our "missions" and "markets" controls. Thus, we cannot speak to circumstances where editors across all journals may not be willing to publish a particular idea or ideas from particular scientists. Similarly, we cannot speak to biases against novelty or controversy since our methodology is not suited for differentiating between papers on the same topics that draw different conclusions.

We believe future work studying the academic publication process should focus on better understanding persistent biases over long periods of time (i.e., understanding what drives variation in γ_{mj} across journals) as well as any pervasive biases related to particular topics (i.e., understanding what drives variation in σ_{mt} across time). Evaluating these long-run, large-scale issues will require unique data and creative research designs, but our results suggest that short-run concerns pertaining to scientific homophily might not be very important.

REFERENCES

- Berry, S. T., "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics* 25:2 (1994), 242–262. 10.2307/2555829
- Boudreau, K. J., E. C. Guinan, K. R. Lakhani, and C. Riedl, "Looking across and Looking Beyond the Knowledge Frontier: Intellectual Distance, Novelty, and Resource Allocation in Science," *Management Science* 62:10 (2016), 2765–2783. 10.1287/mnsc.2015.2285
- Brogaard, J., J. Engelberg, and C. A. Parsons, "Networks and Productivity: Causal Evidence from Editor Rotations," *Journal of Financial Economics* 111:1 (2014), 251–270. 10.1016/j.jfineco.2013.10.006
- Bryan, K. A., and Y. Ozcan, "The Impact of Open Access Mandates on Invention," this REVIEW 103:5 (2021), 954–967.
- Card, D., S. DellaVigna, P. Funk, and N. Iriberry, "Are Referees and Editors in Economics Gender Neutral?" *Quarterly Journal of Economics* 135:1 (2020), 269–327. 10.1093/qje/qjz035
- Carrell, S., D. Figlio, and L. Lusher, "Clubs and Networks in Economic Reviewing," *Journal of Political Economy* 132:9 (2024), 2999–3024.
- Crane, D., "The Gatekeepers of Science: Some Factors Affecting the Selection of Articles for Scientific Journals," *American Sociologist* 2:4 (1967), 195–201.
- Dasgupta, P., and P. A. David, "Toward a New Economics of Science," *Research Policy* 23:5 (1994), 487–521. 10.1016/0048-7333(94)01002-1
- Gentzkow, M., and J. M. Shapiro, "What Drives Media Slant? Evidence from US Daily Newspapers," *Econometrica* 78:1 (2010), 35–71. 10.3982/ECTA7195
- Ginther, D. K., J. Basner, U. Jensen, J. Schnell, R. Kington, and W. T. Schaffer, "Publications as Predictors of Racial and Ethnic Differences in NIH Research Awards," *PLoS One* 13:11 (2018), e0205929. 10.1371/journal.pone.0205929
- Glasser, T. L., D. S. Allen, and S. E. Blanks, "The Influence of Chain Ownership on News Play: A Case Study," *Journalism Quarterly* 66:3 (1989), 607–614. 10.1177/107769908906600311
- Hegde, D., and B. Sampat, "Can Private Money Buy Public Science? Disease Group Lobbying and Federal Funding for Biomedical Research," *Management Science* 61:10 (2015), 2281–2298. 10.1287/mnsc.2014.2107
- Hill, R., and C. Stein, "Race to the Bottom: Competition and Quality in Science," *The Quarterly Journal of Economics* (forthcoming).
- Laband, D. N., and M. J. Piette, "Favoritism versus Search for Good Papers: Empirical Evidence Regarding the Behavior of Journal Editors," *Journal of Political Economy* 102:1 (1994), 194–203. 10.1086/261927
- Lerchenmueller, M. J., and O. Sorenson, "Author Disambiguation in PubMed: Evidence on the Precision and Recall of Author-ity among NIH-Funded Scientists," *PLoS One* 11:7 (2016), e0158731. 10.1371/journal.pone.0158731
- Li, D., "Expertise versus Bias in Evaluation: Evidence from the NIH," *American Economic Journal: Applied Economics* 9:2 (2017), 60–92. 10.1257/app.20150421
- Merton, R. K., *The Sociology of Science: Theoretical and Empirical Investigations* (Chicago: University of Chicago Press, 1973).
- Myers, K., "The Elasticity of Science," *American Economic Journal: Applied Economics* 12:4 (2020), 103–134. 10.1257/app.20180518
- Oster, E., "Health Recommendations and Selection in Health Behaviors," *American Economic Review: Insights* 2:2 (2020), 143–160. 10.1257/aeri.20190355
- Stephan, P. E., "The Economics of Science," *Journal of Economic Literature* 34:3 (1996), 1199–1235.
- Torvik, V. I., and N. R. Smalheiser, "Author Name Disambiguation in MEDLINE," *ACM Transactions on Knowledge Discovery from Data (TKDD)* 3:3 (2009), 1–29. 10.1145/1552303.1552304
- Wang, J., R. Veugelers, and P. Stephan, "Bias against Novelty in Science: A Cautionary Tale for Users of Bibliometric Indicators," *Research Policy* 46:8 (2017), 1416–1436. 10.1016/j.respol.2017.06.006
- Way, S. F., A. C. Morgan, A. Clauset, and D. B. Larremore, "The Misleading Narrative of the Canonical Faculty Productivity Trajectory," *Proceedings of the National Academy of Sciences* 114:44 (2017), E9216–E9223. 10.1073/pnas.1702121114
- Zuckerman, H., and R. K. Merton, "Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System," *Minerva* 9:1 (1971), 66–100. 10.1007/BF01553188