How Important is Editorial Gatekeeping? Evidence from Top Biomedical Journals*

Joshua L. Krieger[†] Kyle R. Myers[‡] Ariel D. Stern[§]

January 9, 2023

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.

JEL Codes: I23, J24, O31, O34

*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. ADS acknowledges support from the Kauffman Foundation. All errors are our own.

†Harvard Business School. Email: jkrieger@hbs.edu

‡Harvard Business School. Email: kmyers@hbs.edu

§Harvard Business School. Email: astern@hbs.edu

"...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.*

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 and 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 and 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 and 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

[1] Another 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. (2020) and references therein.

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 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, non-trivial changes in the composition of an editorial board would not meaningfully alter journals' scientific content.

Our paper 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 paper 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 and 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 editorial 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 paper, and consider the broader implications of our findings for future work in the Discussion.

1 Background, Data, and Summary Statistics

1.1 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

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

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 non-scientific editorial positions from our analyses (e.g., copyeditors).

1.2 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 A.1. 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 is 7 years (s.d.=7 years), and, because of the skewness in tenure lengths, the average active editor is in their sixth year of tenure in our data.

1.3 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 and 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.

1.4 Categorizing Topics with Medical Subject Headings (MeSH)

In order 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

[3] Our 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.

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

Diabetes Mellitus and *Adrenal Gland Diseases* are sub-terms 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 paper 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, 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 3.4, 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.

2 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^P_{mjt}, to the same share measure based on editors' publications prior to their start at the journal, S^E_{mjt}[5]:

$$S^P_{mjt} = \alpha + S^E_{mjt}\beta + \epsilon_{mjt} , \tag{1}$$

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, $\widehat{\beta}$, 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

[5]We 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.

[6]The partial-R^2 for a given independent variable x is defined as $\frac{R^2 - R^2_{-x}}{1-R^2_{-x}}$, where R^2 corresponds to the saturated model and R^2_{-x} corresponds to the model where x is excluded.

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^E_{mjt} 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 and Sampat (2015) or competitive pressures as in Hill and Stein (2021)).

Panel (c) of Table 1 reports the distributions of these share measures (S^P, S^E). Clearly, there are a large number zeros 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^E_{mjt} 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 D.6).

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.

2.1 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, 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). 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

[7]As 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.

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).

2.2 Main Regression Model

In order 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:

$$S^P_{mjt} = \underbrace{\gamma_{mj}}_{\substack{\text{journal's}\\ \text{mission}}} + \underbrace{\sigma_{mt}}_{\substack{\text{topic}\\ \text{markets}}} + \underbrace{S^E_{mjt}\beta}_{\substack{\text{editors' scientific}\\ \text{homophily}}} + \epsilon_{mjt} \,. \tag{2}$$

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.

2.3 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]

2.4 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

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

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 C.3). 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.

3 Results

3.1 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.

[9] In the demand model described in Appendix B, we formalize these local shocks and the assumption about non-negative selection.

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.

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^E_{mjt}) than there is for journals' publications (S^P_{mjt}).[10]

[10]The coefficient of variation of these non-zero distributions is 25% larger for S^E_{mjt} than S^P_{mjt}.

3.2 Alternative Specifications & 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 in order 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 non-research (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 D.4). Lastly, 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)" 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 C.3 illustrates how the (small) effects we estimate could instead be plausibly explained by unobservable time varying trends.

3.3 Proxying for Quality with Forward Citations

Given that we identify a non-zero 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 D.9 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.

3.4 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 and 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 D.10 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.

3.5 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 of these observed similarities, which range from about 0.25 up to nearly 1 (mean= 0.75, s.d.=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% (s.d.=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.

4 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. Thus, 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.

Tables & Figures

Table 1: Summary Statistics

Panel (a): Sample Journal List

Am Rev Respir Dis	*Anesth Analg*	*Anesthesiology*
Ann Intern Med	*Ann Surg*	*Br J Anaesth*
Circ Res	*Circulation*	*Hum Brain Mapp*
J Am Coll Cardiol	*J Nucl Med*	*JAMA*
N Eng J Med	*Radiology*	*Stem Cells*

Panel (b): Journal & Editor Statistics (Journal-Year obs.)

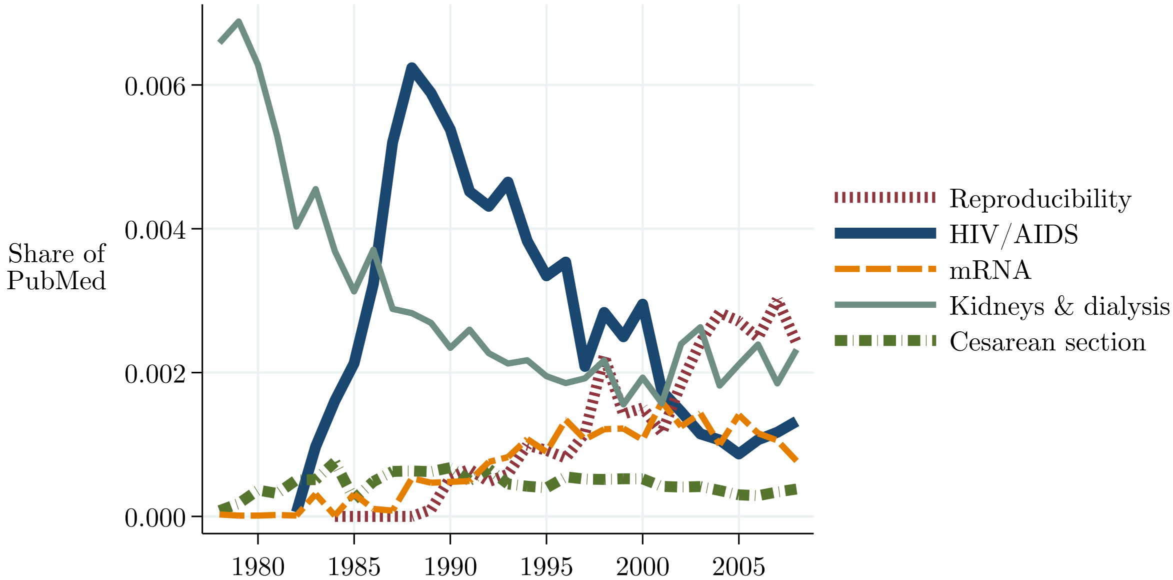
	Obs.	Mean	S.D.	Min.	$p25$	$p50$	$p75$	Max
Year	336	1996.9	6.9	1985	1991	1997	2003	2008
Editors								
Num. Matched Eds.	336	57.8	51.8	7	25	38	76	278
Eds'. Current Tenure	336	5.70	2.51	0.00	4.11	5.19	6.59	14.88
Pubs. per Ed.	336	72.4	27.5	33.0	49.2	66.7	95.5	151.1
Share of MeSH Tree	336	0.302	0.098	0.097	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	S.D.	Min.	$p25$	$p50$	$p75$	Max
Journals' Pubs.								
$\mathbf{1}\{S^P_{mjt}\}$	1,954,148	0.154						
S^P_{mjt} 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.								
$\mathbf{1}\{S^E_{mjt}\}$	1,954,148	0.318						
S^E_{mjt} if >0	620,759	$5.41e^{-4}$	$2.71e^{-3}$	$9.35e^{-7}$	$3.38e^{-5}$	$1.01e^{-4}$	$3.31e^{-4}$	0.107

Notes: Panel (a) lists the fifteen 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 non-zero 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 non-zero (if >0), which are approximately log-normal.

Figure 1: Example MeSH Topic Trends



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

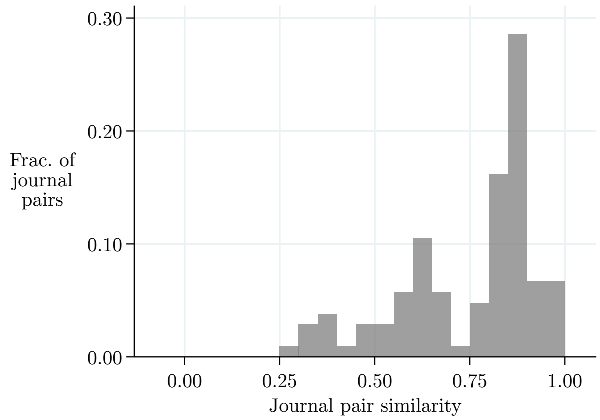
Table 2: Main Results

	(1)	(2)	(3)	(4)
Panel (a): D.V. = S^P_{mjt}				
S^E_{mjt}	1.044	0.245	0.867	0.230
	(0.054)	(0.038)	(0.038)	(0.029)
Elasticity at means	1.043	0.245	0.867	0.230
	(0.041)	(0.038)	(0.038)	(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. = $\mathbf{1}\{S^P_{mjt} > 0\}$				
$\mathbf{1}\{S^E_{mjt} > 0\}$	0.353	0.065	0.129	0.021
	(0.005)	(0.002)	(0.002)	(0.001)
Elasticity at means	0.723	0.133	0.265	0.044
	(0.004)	(0.004)	(0.004)	(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. = $log(S^P_{mjt})$				
$log(S^E_{mjt})$	0.530	0.089	0.466	0.046
	(0.016)	(0.007)	(0.007)	(0.004)
Elasticity at means	0.530	0.089	0.466	0.046
	(0.016)	(0.007)	(0.007)	(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

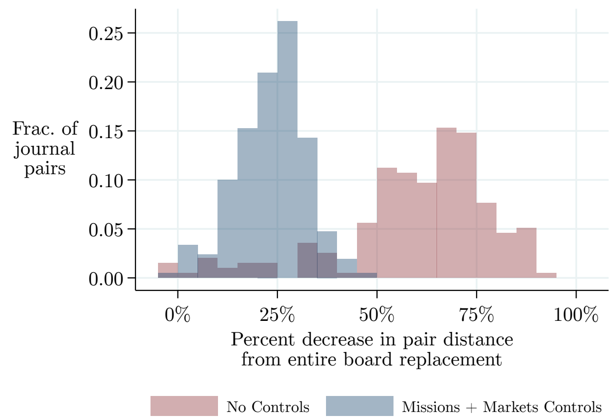
Notes: Standard errors are in parentheses and are clustered at the MeSH (m) level. Elasticities are reported at the means of the independent and dependent variables, with standard errors calculated via the delta method. The bottom two rows apply to all panels and indicate which columns are based on specifications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

Figure 2: Journal Similarity, Observed and After Simulated "Takeovers"

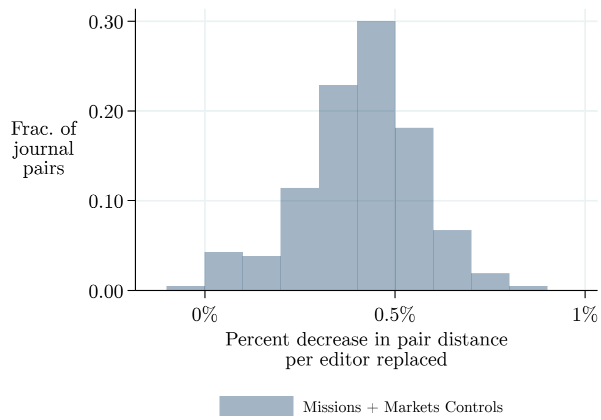
(a) Observed Journal-Journal Similarity Scores



(b) Decrease in Distance Post-Takeover, Total Board Replacement



(c) Decrease in Distance Post-Takeover, Per Editor Replaced



Notes: 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.

References

Amrein, K., Langmann, A., Fahrleitner-Pammer, A., Pieber, T. R., and Zollner-Schwetz, I. (2011). Women underrepresented on editorial boards of 60 major medical journals. *Gender Medicine*, 8(6):378–387.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics*. Princeton university press.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262.

Boudreau, K. J., Guinan, E. C., Lakhani, K. R., and Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, 62(10):2765–2783.

Brogaard, J., Engelberg, J., and Parsons, C. A. (2014). Networks and productivity: Causal evidence from editor rotations. *Journal of Financial Economics*, 111(1):251–270.

Bryan, K. A. and Ozcan, Y. (2021). The impact of open access mandates on invention. *Review of Economics and Statistics*, Forthcoming.

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

Carrell, S., Figlio, D., and Lusher, L. (2020). Clubs and networks in economic reviewing. *Mimeo*.

Crane, D. (1967). The gatekeepers of science: Some factors affecting the selection of articles for scientific journals. *American Sociologist*, 2(4):195–201.

Dasgupta, P. and David, P. A. (1994). Toward a new economics of science. *Research Policy*, 23(5):487–521.

Einav, L., Finkelstein, A., and Polyakova, M. (2018). Private provision of social insurance: drug-specific price elasticities and cost sharing in medicare part d. *American Economic Journal: Economic Policy*, 10(3):122–53.

Espey, J. A. and Espey, M. (2004). Turning on the lights: A meta-analysis of residential electricity demand elasticities. *Journal of Agricultural and Applied Economics*, 36(1):65–81.

Espey, M. (1998). Gasoline demand revisited: an international meta-analysis of elasticities. *Energy economics*, 20(3):273–295.

Gentzkow, M. and Shapiro, J. M. (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1):35–71.

Ginther, D. K., Basner, J., Jensen, U., Schnell, J., Kington, R., and Schaffer, W. T. (2018). Publications as predictors of racial and ethnic differences in NIH research awards. *PloS One*, 13(11):e0205929.

Glasser, T. L., Allen, D. S., and Blanks, S. E. (1989). The influence of chain ownership on news play: A case study. *Journalism Quarterly*, 66(3):607–614.

Hegde, D. and Sampat, B. (2015). Can private money buy public science? disease group lobbying and federal funding for biomedical research. *Management Science*, 61(10):2281–2298.

Hill, R. and Stein, C. (2021). Race to the bottom: Competition and quality in science. *Mimeo*.

Laband, D. N. and Piette, M. J. (1994). Favoritism versus search for good papers: Empirical evidence regarding the behavior of journal editors. *Journal of Political Economy*, 102(1):194–203.

Lerchenmueller, M. J. and Sorenson, O. (2016). Author disambiguation in pubmed: evidence on the precision and recall of author-ity among nih-funded scientists. *PLoS One*, 11(7):e0158731.

Li, D. (2017). Expertise versus bias in evaluation: Evidence from the NIH. *American Economic Journal: Applied Economics*, 9(2):60–92.

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., and Leibowitz, A. (1987). Health insurance and the demand for medical care: evidence from a randomized experiment. *The American economic review*, pages 251–277.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

Myers, K. (2020). The elasticity of science. *American Economic Journal: Applied Economics*, 12(4):103–134.

Oster, E. (2020). Health recommendations and selection in health behaviors. *American Economic Review: Insights*, 2(2):143–60.

Stephan, P. E. (1996). The economics of science. *Journal of Economic Literature*, 34(3):1199–1235.

Torvik, V. I. and Smalheiser, N. R. (2009). Author name disambiguation in medline. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 3(3):1–29.

Wang, J., Veugelers, R., and Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8):1416–1436.

Way, S. F., Morgan, A. C., Clauset, A., and Larremore, D. B. (2017). The misleading narrative of the canonical faculty productivity trajectory. *Proceedings of the National Academy of Sciences*, 114(44):E9216–E9223.

Zuckerman, H. and Merton, R. K. (1971). Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system. *Minerva*, 9(1):66–100.

Appendix

A Data Construction and Additional Summary Statistics

A.1 Journal Data

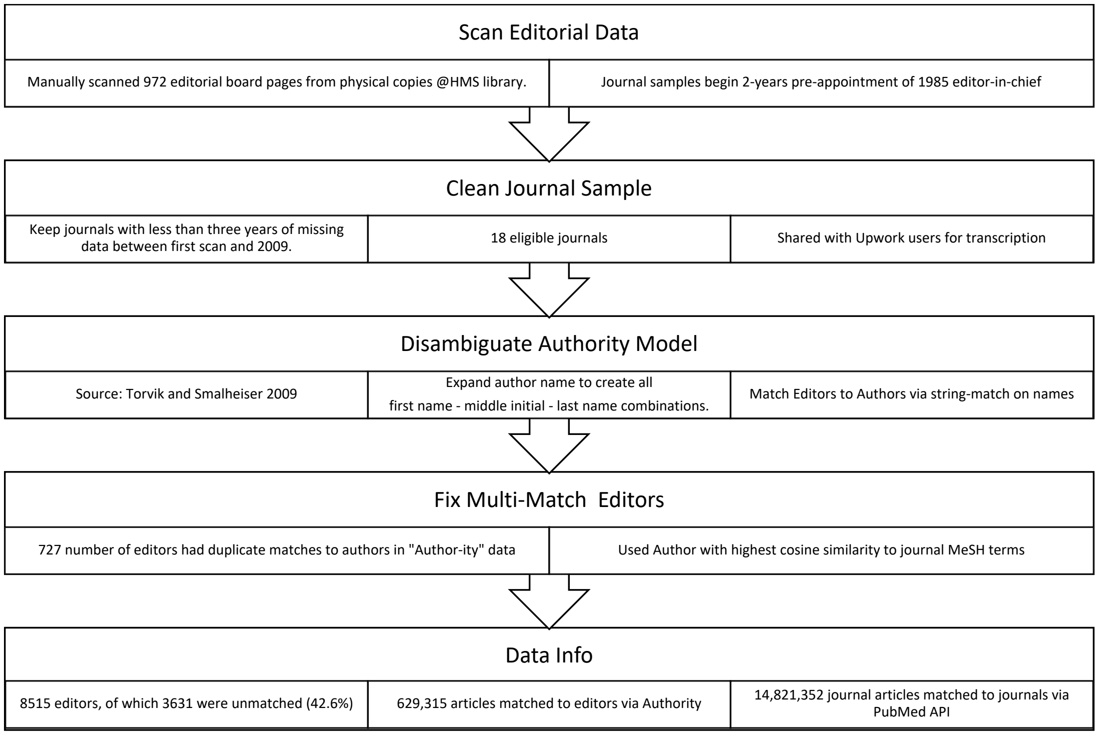
Figure A.1 outlines our process. To construct our set of journals, we follow related work by Amrein et al. (2011) and select the five top-rated journals from general medicine and the specialties of anesthesiology, critical care, radiology, surgery, and hematology. We then restrict the sample to journals where we were able to obtain physical copies (which are necessary to observe the masthead for the editorial staff information) prior to 1995 and less than three years of hard copies were missing. To the extent possible, we obtained one physical copy per year for each journal where the earliest year pursued was set by a rule-of-thumb to be two years prior to the appointment of whomever was editor-in-chief of the journal in 1985. From scans of these physical copies, we outsourced the manual transcription of the editorial board members.

Table A.1: Journal Summary Statistics

Journal	Earliest Year	Latest Year	Pub.-Matched Num. Eds.		Editors' Pubs. per Ed.		Journals' Pubs per Year	
Am Rev Respir Dis	1978	2008	62.3	(25.7)	75.8	(13.1)	522.6	(124.0)
Anesth Analg	1976	2008	21.7	(21)	44.7	(14.3)	535.1	(225.0)
Anesthesiology	1975	2008	22.9	(14.8)	40.7	(6.69)	433.5	(104.0)
Ann Intern Med	1969	2008	25.1	(8.19)	58.8	(16.2)	481.6	(76.4)
Ann Surg	1972	2008	24.6	(15.7)	120.1	(15.3)	231.1	(33.8)
Br J Anaesth	1981	2008	17.4	(4.64)	43.8	(4.98)	338.5	(70.2)
Circ Res	1980	2008	82.2	(38.4)	84.1	(11.1)	301.8	(86.5)
Circulation	1981	2008	160.4	(80.6)	106.7	(3.98)	889.8	(369.0)
Hum Brain Mapp	1998	2008	37.2	(3.95)	54.2	(8.94)	76.7	(27.5)
J Am Coll Cardiol	1983	2008	102.5	(51.4)	103.9	(19.5)	586.2	(137.0)
J Nucl Med	1975	2008	45.5	(26.4)	56.7	(14.8)	325.9	(81.6)
JAMA	1980	2008	28.0	(3.56)	60.8	(10.6)	1,000.1	(105.0)
N Eng J Med	1976	2008	19.0	(12.5)	77.7	(18.7)	1,132.0	(232)
Radiology	1964	2008	49.3	(37.1)	39.3	(10.1)	593.9	(133.0)
Stem Cells	1996	2008	54.4	(30.5)	111.6	(17.2)	150.7	(113.0)

Notes: The rightmost six columns report the mean and s.d. in parentheses for the three variables listed in the heading. "Pub.-Matched" counts only editors that we are successfully able to merge to an entry in the Author-ity (Torvik and Smalheiser 2009) disambiguated publication database.

Figure A.1: Flowchart of Data Construction



Notes: Outlines the data construction process.

A.2 Publication Data and MeSH Categorizations

MeSH terms are assigned at the article level. However, by combining sets of articles by journal(s) and author, we can measure the relative prevalence of MeSH topics for broader corpuses of work. As a result, our key regression variables are *shares* of articles containing a given topic within a journal-year or editors own publications. Aggregating MeSH topics poses two main challenges. First, any given PubMed article can be associated with multiple MeSH terms. Some articles have dozens of assigned MeSH topics while others only have a handful. To address this issue, we weight each topic by the inverse of total topics per article–such that topics get larger weight assigned when the article has fewer total topics.

The second issue in aggregating MeSH topics' relative frequencies is that the MeSH tree's "branches" vary in length and specificity.[1] For example, one article might be tagged with the general (three digit) topic of "Digestive System Diseases," while another article is assigned the (15 digit) topic of "Somatostatinoma," a particular type of pancreatic cancer, which is a sub-branch of the larger "Digestive System Diseases" topic. In handling these

[1]Within the MeSH system, the levels of depth of a given sub-topic is denoted by the number of digits associated with its identifier.

types of comparisons, we don't want to throw away valuable detail within MeSH branches, but we also want to make sure we are not overstating differences between topics within the same larger branches.

Our solution, loosely speaking, is to calculate the "distance down a branch" each MeSH term is. For instance, if a certain branch of the MeSH tree contains a total of 3 successive terms, we'd denote the first term as being 1/3 specificity, the second being 2/3, and the last being 3/3. Then, we can make a choice about how aggregated (less specificity; shorter distances down the branch) or disaggregated (more specificity; longer distances down the branch) we'd like the data to be. In our preferred specifications, we aggregate the MeSH tree to 50% of all branches. As shown below, we obtain very similar estimates when disaggregating more or less.

B Motivating Demand Model

The following presents two interconnected demand models of how scientists choose content to publish in journals (which generates variation in our dependent variable), and how they choose to fill editorial positions (which generates variation in our focal independent variable). Besides motivating our regressions, the purpose of this exercise is to formalize our argument as to why our estimate of the scientific homophily effect is likely an upper bound of the true effect.

To preface, there are two major departures from reality in the demand models that follow. First, we model the joint utility to both non-editors and editors from publishing content and filling editorial positions, because we cannot observe rejected paper submissions or unsuccessful bids to become an editor. That is, we ignore the facts that a given article is only ever submitted to a very few number of journals, and only a very few number of (non-editor) individuals are considered for editorships. Thus, the homophily effect we estimate will be driven by a combination of the choices made by non-editors and editors.[2]

Second, we ignore the fact that publication choices are made at the level of indivisible papers and editorship choices are made at the level of indivisible humans. Instead, we model choices at the level of the scientific (MeSH) topics that describe the content of papers and, based on the papers they've written, the specialties of scientists. In other words, our model describes a world where portions of an article could be published in different journals and only portions of a scientists' expertise could be awarded an editorship. This abstraction dramatically simplifies our empirical approach and directly motivates our regressions with a specific set of fixed effects. These fixed effects remove variation in the data due to the forces described earlier – namely, journals' missions and market-wide topic trends – which, if not accounted for, may yield incorrect estimates of the homophily effect.[3]

[2]For instance, how much of the homophily effect is driven by authors choices about which journals to submit papers to (due to their beliefs about the homophily effect), and how much is due to editors choices amongst the set of papers they actually receive? We cannot separate these effects, which may be an interesting area for future research. However, our goal here is to understand the magnitude of this aggregate effect, because if it is small (as our results indicate), it suggests that allocating resources to understanding or addressing this particular homophily effect may not be worthwhile.

[3]Unlike traditional discrete choice models of demand (e.g., Berry 1994), our units of observation are actually based on features of the (true) product (i.e., topics of these papers). Practically speaking, in the colloquial terms of "wide" and "long" data, where traditional demand models are wide – products have multiple features described by multiple data columns – our model is long – where papers exist in multiple data rows.

B.1 Publication choices

The joint utility to authors and editors considering each article a (the consumers) from publishing their topical content m (the products) in a given journal-year jt (the markets) depends on features of the journal and the topics of the paper. The outside good is not publishing any content that could have possibly fit within the journal's pages.

We follow Berry's (1994) approach to modeling a discrete choice and write the joint indirect utility from publishing article a, which contains topics m, in journal j at year t as:

$$u_{amjt} = \delta_{mjt} + \epsilon_{amjt}\,, \tag{1}$$

where δ_{mjt} is the mean indirect utility term common to all mjt, and ϵ_{amjt} is an i.i.d. idiosyncratic term with a Type 1 extreme value distribution.[4]

The mean indirect utility from publication depends on journal-time specific quality thresholds and/or space constraints (α_{jt}), the journal-topic specific match due to the time invariant missions of journals (γ_{mj}), aggregate market-wide trends in the supply of or demand for topics over time (σ_{mt}), "local" supply and/or demand shocks at the journal-topic-time level ξ_{mjt}, and the journal-specific expertise and backgrounds of current editors, summarized by the parameter δ'_{mjt}:

$$\delta_{mjt} = \alpha_{jt} + \gamma_{mj} + \sigma_{mt} + \delta'_{mjt}\beta + \xi_{mjt}\,. \tag{2}$$

For example, a more negative α_{jt} would indicate that it is harder to publish in jt compared to other journals at that time; a more positive γ_{mj} would indicate that it is always easier to publish in journal j when the article is related to topic m.

We will construct a measure of δ'_{mjt} with data, and β is the focal homophily parameter that governs how much editors' research backgrounds influence the composition of the publications in their journals (i.e., via gatekeeping). Our estimate of β will capture the net effects of editors possibly having stronger preferences for the topics that they themselves study, as well as the possibility that editors evaluate work more closely aligned with their

[4]This commonly used assumption yields a closed form solution for choice probabilities and facilitates the linear regressions below. The independence of these terms is implausible since the ϵ_{amjt} values associated with each article will very likely be correlated. Furthermore, it also assumes the independence of irrelevant alternatives, which can be problematic when constructing elasticity matrices for substitution patterns. We are not very concerned with these issues given the objective of our analyses – inferring the magnitude of any homophily effect. The first issue likely implies that our counterfactual exercise will allow more flexibility in the content publishing process (since, per our model, an editor could technically publish a portion of a paper), which may yield larger changes due to homophily than would be seen in reality, continuing our interpretation of our results as upper bounds. The second issue is likely not relevant since we do not investigate any true counterfactuals related to the introduction of new products.

own specialties more critically (i.e., Boudreau et al. 2016; Li 2017). The local-shock term (ξ_{mjt}) is unobservable, and we cannot empirically account for it – it is what will lead us to interpret our estimate of β as an upper bound as we outline below.

B.2 Editorship choices

Next, we specify a utility model of editors choosing their replacements. The joint utility to potential-editors and current editors i (the consumers) from having the potential-editor with expertise on topic m (the products) become an editor at journal j in year t (the markets) depends on features of the journal and the experience of the individual, as proxied by the topics of the papers they have written. Assume for simplicity that each individual can only be in one position at a time, and the outside good is defined as not filling a potential editor slot. Since we're again modeling choices at the topic-level, markets here are indexed by journal-time (jt) pairs and we have another joint indirect utility function:

$$u'_{imjt} = \delta'_{mjt} + \epsilon'_{mjt}, \tag{3}$$

where again, δ'_{mjt} is the mean indirect utility term common to all mjt, and ϵ_{amjt} is an i.i.d. idiosyncratic term with a Type 1 extreme value distribution. Similarly as we do above, we decompose the mean utility term into an additive series of jt-, mj-, mt-, and mjt-specific effects:

$$\delta'_{mjt} = \alpha'_{jt} + \gamma'_{mj} + \sigma'_{mt} + \xi'_{mjt} \, . \tag{4}$$

Again, we allow the effects of journal size, journal scope, and topic trends to influence the value of a scientist with a particular specialty becoming an editor.

In practice, editorships are filled for terms that often last many years (e.g., in our data the average tenure is about sixteen years). Thus, potential and current editors must use their beliefs (about γ'_{mj}) and forecasts (about $\alpha'_{jt}, \sigma'_{mt}$, and ξ'_{mjt}) when making editorship decisions. With this in mind, we view the four right-hand-side parameters in Equation (4) as the beliefs/forecasts about the corresponding parameters in Equation (2) (i.e., ξ'_{mjt} is the forecast of ξ_{mjt}).

It will be useful to formalize the process relating the realized and forecasted local shocks (ξ_{mjt}, ξ'_{mjt}), so we assume:

$$\xi'_{mjt} = \delta\xi_{mjt} + \mu_{mjt} \, , \tag{5}$$

where the δ parameter captures the accuracy of forecasts, and μ_{mjt} captures "noise" in the specialties of an editorial board that arises when forecasts of local demand shocks are not perfectly accurate.

The key assumption that leads us to interpret our estimate of the homophily effect (β) is to assume

$$\delta \geq 0\,. \tag{6}$$

That is, we assume there is a non-negative correlation between forecasted and realized local demand shocks for papers and editors that specialize in specific topics at specific journals at a specific time. Next, we connect the two choice models above and make use of this assumption to motivate our main regression model.

B.3 Empirical model

Based on Equations (1–2), and following Berry (1994), we can define mean utility and observed publication market shares of each topics s_{mjt} and some outside good s_{0jt}[5]:

$$log\big(s_{mjt}/s_{0jt}\big) = \delta_{mjt} = \alpha_{jt} + \gamma_{mj} + \sigma_{mt} + \delta'_{mjt}\beta + \xi_{mjt} \tag{7}$$

Likewise, based on Equations (3–4), we can define mean utility and observed editorial market shares of each topic s'_{mjt} and the outside good s'_{0jt} as:

$$log\big(s'_{mjt}/s'_{0jt}\big) = \delta'_{mjt} = \alpha'_{jt} + \gamma'_{mj} + \sigma'_{mt} + \xi'_{mjt} \tag{8}$$

Substituting δ'_{mjt} from Equation (8) into Equation (7) and making use of Equation (5), the local-shock forecasting assumption, yields:

$$\begin{aligned} log\big(s_{mjt}/s_{0jt}\big) &= \alpha_{jt} + \gamma_{mj} + \sigma_{mt} + log\big(s'_{mjt}/s'_{0jt}\big)\beta + \xi_{mjt} \\ &= (\alpha_{jt} + \beta\alpha'_{jt}) + (\gamma_{mj} + \beta\gamma'_{mj}) + (\sigma_{mt} + \beta\sigma'_{mt}) + \beta\big(\delta\xi_{mjt} + \mu_{mjt}\big) + \xi_{mjt}\;. \end{aligned} \tag{9}$$

First, note that we can ignore both the α_{jt} parameters and the denominators of the logged share ratios, the outside good shares. Both of these are irrelevant because they are all constant within jt, and the only two other data-based variables in the model (s_{mjt}, s'_{mjt}) are both scaled at the jt-level (since they are each m's shares of in-journal publications or editors' publications – the variation is only within-jt). Thus, we can rewrite that first line of Equation (9) to be:

$$log(s_{mjt}) = \gamma_{mj} + \sigma_{mt} + log(s'_{mjt})\beta + \xi_{mjt}\;, \tag{10}$$

[5]As shown below in Eq. 10, the specific definition of the outside good is irrelevant as it will drop out of the estimating equation. In theory, the outside good reflects the value of *not* publishing any more content, just as in a traditional product demand model the outside good often reflects the value of the no-purchase decision.

which mirrors our main regression specification, Equation (2), except for the log transformation. The last line of Equation (9) shows how our empirical estimate of β will depend on the (unknowable) value of δ, which describes the accuracy of current editors' forecasts of the local shocks ξ_{mjt} when choosing future editors.

In one boundary case, if these local shocks are forecasted perfectly and $\delta = 1$, then $\xi_{mjt} = \xi'_{mjt}$ and $\mu_{mjt} = 0 \,\forall\, mjt$. In this case, our estimate of β would equal one (and the parameters would have no standard errors) because the only variation in $log\big(s'_{mjt}/s'_{0mt}\big)$ (conditional on the fixed effects) will be due to ξ_{mjt} – there would be no residuals in the regression. This would also lead to an upward bias in the within-R^2 associated with the gatekeeping effect.

In the other boundary case, if forecasts have no predictive value and $\delta = 0$, then ξ_{mjt} and ξ'_{mjt} are orthogonal. In this case, our estimate of β would represent the true effect since the only residual variation in $log\big(s'_{mjt}/s'_{0mt}\big)$ (conditional on the fixed effects) is due only to the "noise" of μ_{mjt}.

To summarize, we cannot observe local shocks ξ_{mjt}[6], which describe journal-time-specific shocks to the supply of and/or demand for research on a topic. The impact of this on our estimate of the gatekeeper effect (β) will depend on the accuracy of forecasts about these local shocks (e.g., do current editors (care to) accurately forecast the demand for a certain type of research and choose a new editor who has expertise in that topic?). As these forecasts becoming more accurate and important ($\delta \to 1$), our estimate of β will be biased upwards and we will overstate the amount of variation in publications due to gatekeeping. Conversely, as they become less important ($\delta \to 0$), our estimate of β will approach the true value. Thus, under these assumptions, our regressions identify a plausible upper bound of the amount of homophily created by editorial gatekeeping.

[6] Nor can we condition out this variation with fixed effects since that would eliminate the variation in editorial specialties across journals and time.

C Alternative Empirical Approach: Editor-Journal Pairs

The empirical model and data-frame at the core of our analyses, which is centered on MeSH-journal-time (mjt) observations, is useful because it maps directly to a model of editors' utility (see Section B) and allows us to clearly illustrate the roles of missions and markets relative to that of homophily. However, a limitation of this approach is that it aggregates all editors into a single mass, and cannot shed light on the role individual editors play in this setting.

As an alternative approach, we construct an editor-journal-time (ijt) dataset that focuses on a measure of the scientific similarity between the editor's own publications and those of each journal in our sample. Recasting the data in this way allows us to ask whether the articles published during each editor's tenure are more or less similar to their own work before or after their tenure. Formally, we estimate the following regression:

$$s_{ijt} = \beta \mathbf{1}\{\text{active}_{ijt}\} + \gamma_{ij} + \alpha_{jt} + \sigma_{it} + \epsilon_{ijt} \ , \tag{11}$$

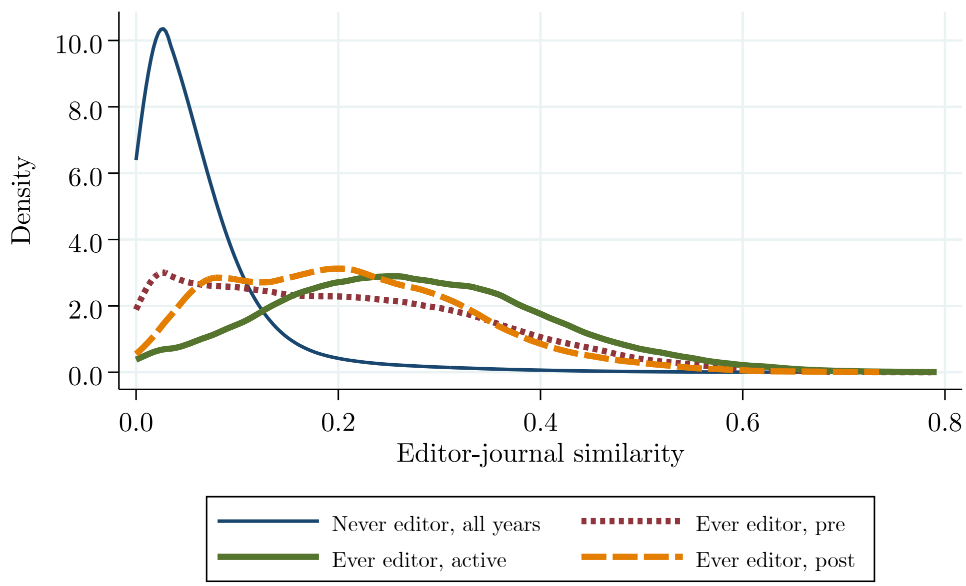
where s_{ijt} is the cosine similarity between each editor's prior publications (up to and including the current year) and each journals current-year publications. The cosine similarities are constructed using the MeSH terms of both sets of publications, such that $s_{ijt} \in [0, 1]$. $\mathbf{1}\{\text{active}_{ijt}\}$ is an indicator that editor i is on the board of journal j in year t, and β is the focal parameter of interest indicating the magnitude of similarity shifts during tenure. The γ_{ij} fixed effects condition out persistent differences in editor-journal similarities, which are analogous to our "mission" controls in the preferred specification. The α_{jt} and σ_{it} fixed effects condition out time-varying differences across editors and journals, which are analogous to our "markets" controls in the preferred specification.

We construct the underlying ijt dataset for all editor-journal pairs in our data, restricting it to editor-years in which the editors are actively publishing (i.e., between the first and last years of any publications) as well as to journal-years for which we have data on editorial boards.

Figure C.2 plots the raw similarity score distributions separately based on different criteria about whether an editor is ever on the board of a particular journal and, if so, whether the observation is from a pre-, active-, or post-tenure year.

Clearly, scientists who ever become editors at a particular journal are substantially aligned with the science that tends to be published in that journal. Furthermore, the possibility of a (small) treatment effect of active editorial status is apparent, as indicated by the right-ward shift in the "Ever editor, active" distribution relative to the distributions of "Ever editors" pre and post tenure.

Figure C.2: Editor-Journal Cosine Similarity Distributions



Notes: Plots the cosine similarity between all possible editor-journal pairs, separately depending on the status of the editor-journal pair.

Table C.2 reports the results of estimating Equation 11, phasing in the fixed effects. In the loosest specification without any controls for mission- or market-wide forces (Col. 1), we estimate that being an active editor leads to a 0.2 increase in the cosine similarity between the editor's prior publications and the current publications in the journal. Compared to the mean similarity in the full dataset (0.06), this is a sizable effect, which also accounts for a non-trivial amount of variation in the data (the partial-R^2 is approximately 0.07).

However, just as in our main results, once we account for stable differences across editor-journal pairs (i.e., "markets"), the effect of editorship appears to be dramatically smaller. Columns 2–3 indicate that the shift in similarity once becoming an editor is likely more on the scale of 0.03, with this status now explaining very little variation in the data (the partial-R^2 is approximately 0.009). As we saw in the MeSH-journal-year analyses in the preferred specifications, scientists that become editors at particular journals are persistently aligned with those journals, and not accounting for this persistence can lead one to drastically overstate the influence of an editor's prior work on the content published during their editorship.

This dataframe permits an even more in-depth event study of how the editor-journal similarity evolves before, during, and after the editors tenure. To investigate the evolution of

Table C.2: Editor-Journal Cosine Regression Results

	(1)	(2)	(3)
$\mathbf{1}\{\text{active}_{ijt}\}$	0.206***	0.039***	0.034***
	(0.003)	(0.002)	(0.001)
Obs., ijt	2,020,080	2,020,080	2,020,080
Total R^2	0.076	0.811	0.879
partial-R^2, ij-missions		0.796	0.807
partial-R^2, it, jt-markets			0.384
partial-R^2, sci. homophily	0.070	0.008	0.009
Incl. γ_{ij}		Y	Y
Incl. $\alpha_{jt} + \sigma_{it}$			Y

Notes: Reports the results from estimating Equation 11. Columns 1–2 both include year (t) fixed effects to account for any aggregate trends.

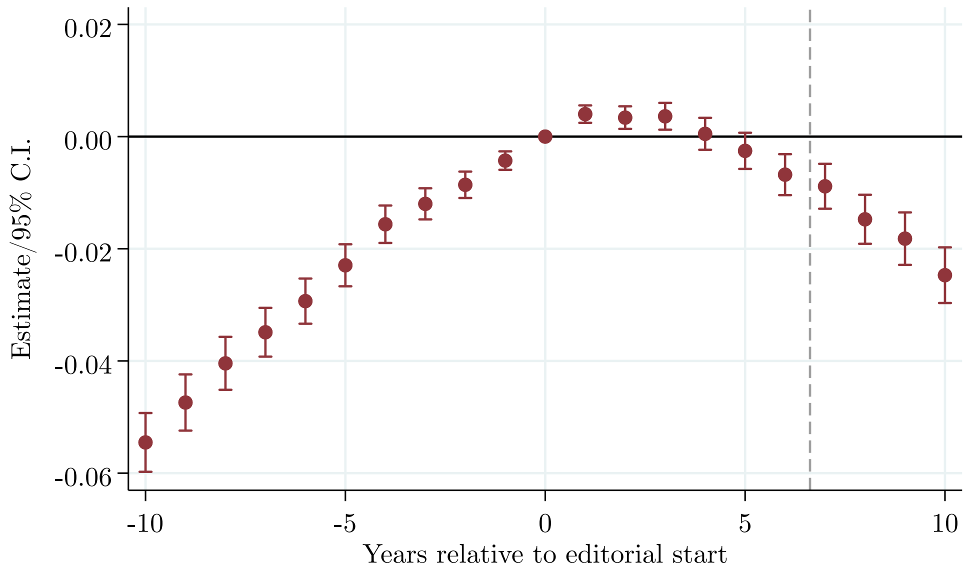
similarity between an (eventual) editor and the journal they join, we estimate the following regression:

$$s_{ijt} = \alpha + \sum_{\tau=-10}^{10} \beta_\tau \mathbf{1}\{\text{years since start}_{ijt}\} + \epsilon_{ijt} \ , \tag{12}$$

where, for ease of interpretation, we use only ij pairs that ever result in i becoming an editor at journal j.

Figure C.3 plots the estimates of the β_τ parameters, which does not illustrate any discrete change in editor-journal similarity at the start of one's tenure, but rather shows a very smooth process whereby editors and journals appear to become more and more similar as the start of a tenure approaches. There is a statistically significant increase once a tenure begins. But overall, the trend appears much more consistent with a process where selection effects (i.e., missions and markets) are much more important drivers of variation in the data relative to any treatment effects (i.e., via scientific homophily).

Figure C.3: Editor-Journal Cosine Similarity Event Study



Notes: Plots the β_τ estimates and 95% confidence intervals per Eq. 12, where the first year of an editor's tenure is the reference point omitted from the vector of indicators. The vertical dashed line is at the average time when an editor's tenure ends.

D Additional Results

D.1 Alternative Specifications

Alternative Missions and Markets Controls

In Table D.3, we replicate our preferred specifications using data-based proxies (instead of fixed effects) to condition out the mission- and market-based variation. For an alternative approach to conditioning out the mission-specific variation, we include one-year lagged values of the dependent variable within the MeSH-journal panels.[7] For an alternative topic-market control, we include the jackknifed average MeSH shares from all journals *except* the focal journal in a given year. Using these alternative approaches yields elasticities all within the bounds of our main results reported in Table 2.

Alternative Risk Set Definitions

As shown in Table 1 Panel (b), there are a large fraction of zeros in our two key MeSH share variables. Given this, one reasonable concern is that these observations may not be "true zeros," in the sense that they may not reflect MeSH-journal pairs where there is a non-zero possibility that either an editor will have published on the topic in the past, or the journal will have published on the topic in the present. In other words, our decision to include all combinations of MeSH terms and journals in the data might cause spurious results.

To investigate this possibility, Table D.4 replicates our main results using four alternative sub-samples where we place stricter and stricter thresholds on which MeSH-journal pairs are included in the analyses. In all cases, we obtain virtually identical estimates to our main results.

Alternative Topic Specificities

We noted in Section 1.4 that using the MeSH hierarchy to index scientific topics requires making a decision about how much (not) to aggregate the topic tree hierarchy. This decision amounts to deciding how specific the MeSH terms will be in the analyses. Our main results are based on, what we term, a "50% aggregation" of the hierarchy in the sense that it is halfway between the most specific and most aggregated we could make our MeSH terms.

In Table D.5, we report results from our preferred specifications using alternative levels of specificity spanning from 10% (very aggregated) to 100% (using all unique MeSH terms).

[7]As noted by Angrist and Pischke (2008), results from the panel fixed effects (i.e., the main specifications using mj fixed effects) and lagged dependent variables should bound the true estimate.

There is some variation in our point estimates, but it is not substantial. In linear and binary models, the point estimates increase with the specificity of the MeSH terms in the analyses. In log-transformed models, the point estimates decrease with specificity of terms. Still, in all cases the elasticity estimates are relatively similar and so we do not infer any meaningful heterogeneity from these results.

Alternative Editorial Board Aggregation

In our main specifications, we construct the focal independent variable, S^E, by effectively treating editorial boards as one giant individual in the sense that we aggregate all publications by all editors together when calculating S^E. This approach implicitly puts more weight on editors with larger publication records since the topics they studied prior to their tenure will, mechanically, appear more frequently. An alternative approach would be to give each individual editor equal weight by first averaging MeSH shares within individuals, and then across all individuals. Table D.6 reports the results from this alternative approach and we find highly similar results. That is to say, editors appear to have equal influence regardless of their prior publication output levels.

Table D.3: Alternative Mission and Market Controls

	Linear-Linear		Binary-Binary		log-log	
	(1)	(2)	(3)	(4)	(5)	(6)
S^E_{mjt}	0.124***	0.099***	0.079***	0.087***	0.218***	0.190***
	(0.016)	(0.011)	(0.001)	(0.001)	(0.004)	(0.005)
Elasticity at means	0.124	0.099	0.163	0.180	0.218	0.190
Total R^2	0.959	0.955	0.504	0.452	0.742	0.650
partial-R^2, mj-missions	0.698	0.723	0.145	0.169	0.353	0.416
partial-R^2, mt-markets	0.101	0.010	0.156	0.069	0.264	0.000
partial-R^2, sci. homoph.	0.029	0.023	0.010	0.015	0.062	0.065
Obs., mjt	1,862,633	1,862,633	1,862,633	1,862,633	156,359	159,900
Missions, control type	Lag D.V.	Lag D.V.	Lag D.V.	Lag D.V.	Lag D.V.	Lag D.V.
Markets, control type	F.E.	Jackknife avg.	F.E.	Jackknife avg.	F.E.	Jackknife avg.

Notes: Reports estimates of the main regression using either linear (cols. 1-2), binary (cols. 3-4), or log (cols. 5-6) transformations of both the dependent and independent variables. Standard errors are in parentheses and are clustered at the MeSH (m) level. The bottom two rows indicate the specification of the mission (mj-level) and market (mt-level) controls. "Lag D.V." indicates that a one-year lag of the dependent variable is included to control for differences in journal missions, "F.E." indicates that fixed effects are used, and "Jackknife avg." indicates that a same-year journal-jackknifed average of the dependent variable is used to control for differences in topic markets.

Table D.4: Alternative MeSH-Journal Samples

	Full Sample (1)	Ever In Journal's (2)	Ever In Editor's (3)	Ever In Both (4)	In Both Abv. $p10$ (5)
Panel (a): D.V. = S^P_{mjt}					
S^E_{mjt}	0.230***	0.231***	0.230***	0.231***	0.231***
	(0.029)	(0.030)	(0.029)	(0.030)	(0.030)
Elasticity at means	0.230	0.227	0.232	0.229	0.229
Total R^2	0.952	0.952	0.952	0.952	0.952
partial-R^2, sci. homoph.	0.024	0.024	0.024	0.024	0.024
Obs., mjt	1,953,818	877,486	960,556	731,153	648,725
Panel (b): D.V. = $\mathbf{1}\{S^P_{mjt}\}$					
$\mathbf{1}\{S^E_{mjt}\}$	0.021***	0.023***	0.013***	0.019***	0.020***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
Elasticity at means	0.044	0.039	0.028	0.035	0.036
Total R^2	0.611	0.530	0.574	0.529	0.519
partial-R^2, sci. homoph.	0.001	0.001	0.001	0.001	0.001
Obs., mjt	1,953,818	877,486	960,556	731,153	648,725
Panel (c): D.V. = $log(S^P_{mjt})$					
$log(S^E_{mjt})$	0.046***	0.046***	0.046***	0.046***	0.046***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Elasticity at means	0.046	0.046	0.046	0.046	0.046
Total R^2	0.803	0.803	0.803	0.803	0.802
partial-R^2, sci. homoph.	0.001	0.001	0.001	0.001	0.001
Obs., mjt	224,224	224,224	224,224	224,224	223,543

Notes: Standard errors are in parentheses and are clustered at the MeSH (m) level. All specifications include MeSH-journal (mj) and MeSH-year (mt) fixed effects. The column headers indicate which pairs of MeSH topics and journals (mj) are included in the regression based on whether the MeSH term ever appears in the journal's and/or the editors' publications. The last column includes only MeSH-journal pairs where the average appearance of the MeSH term in both the journal's and editors' publications is above the 10th percentile of the non-zero distribution.

Table D.5: Alternative MeSH Hierarchy Aggregations

	Level of MeSH Specificity				
	10%	25%	50%	75%	100%
	(1)	(2)	(3)	(4)	(5)
Panel (a): D.V. = S^P_{mjt}					
S^E_{mjt}	0.211***	0.228***	0.230***	0.234***	0.246***
	(0.044)	(0.038)	(0.029)	(0.032)	(0.032)
Elasticity at means	0.211	0.228	0.230	0.234	0.246
Total R^2	0.967	0.961	0.952	0.949	0.949
partial-R^2, mj-missions (γ_{mj})	0.671	0.642	0.646	0.639	0.632
partial-R^2, mt-markets (σ_{mt})	0.237	0.205	0.155	0.155	0.157
partial-R^2, sci. homoph. (S^E_{mjt})	0.024	0.025	0.024	0.023	0.023
Obs., mjt	38,819	355,675	1,953,818	4,937,169	8,317,602
Panel (b): D.V. = $\mathbf{1}\{S^P_{mjt}\}$					
$\mathbf{1}\{S^E_{mjt}\}$	0.001	0.023***	0.021***	0.020***	0.019***
	(0.017)	(0.003)	(0.001)	(0.001)	(0.001)
Elasticity at means	0.001	0.036	0.044	0.049	0.051
Total R^2	0.737	0.733	0.611	0.543	0.521
partial-R^2, mj-missions (γ_{mj})	0.342	0.334	0.323	0.308	0.304
partial-R^2, mt-markets (σ_{mt})	0.134	0.120	0.115	0.114	0.112
partial-R^2, sci. homoph. (S^E_{mjt})	0.000	0.001	0.001	0.001	0.001
Obs., mjt	38,819	355,675	1,953,818	4,937,169	8,317,602
Panel (c): D.V. = $log(S^P_{mjt})$					
$log(S^E_{mjt})$	0.084***	0.069***	0.046***	0.034***	0.030***
	(0.019)	(0.008)	(0.004)	(0.004)	(0.004)
Elasticity at means	0.084	0.069	0.046	0.034	0.030
Total R^2	0.915	0.864	0.803	0.818	0.823
partial-R^2, mj-missions (γ_{mj})	0.492	0.495	0.499	0.518	0.525
partial-R^2, mt-markets (σ_{mt})	0.166	0.214	0.274	0.329	0.356
partial-R^2, sci. homoph. (S^E_{mjt})	0.004	0.002	0.001	0.001	0.000
Obs., mjt	31,910	96,142	224,224	242,541	242,724

Notes: Reports estimates of the main regression using alternative levels of specificity when aggregating the MeSH hierarchy – smaller specificity values indicate the tree has been aggregated to broader topics. A specificity of 100% indicates that no aggregation is performed, and the 50% specificity is the specification used in all other tables and figures. Standard errors are in parentheses and are clustered at the MeSH (m) level. All regressions include both journal mission and topic market fixed effects.

Table D.6: Alternative Aggregation of Editorial Boards

	Linear-Linear		Binary-Binary		log-log	
	Pub.-wgt.	Person-wgt.	Pub.-wgt.	Person-wgt.	Pub.-wgt.	Person-wgt.
	(1)	(2)	(3)	(4)	(5)	(6)
S^E_{mjt}	0.230***	0.232***	0.021***	0.021***	0.046***	0.040***
	(0.029)	(0.032)	(0.001)	(0.001)	(0.004)	(0.004)
Elasticity at means	0.230	0.232	0.044	0.044	0.046	0.040
Total R^2	0.952	0.952	0.611	0.611	0.803	0.803
partial-R^2, mj-missions	0.646	0.630	0.323	0.323	0.499	0.509
partial-R^2, mt-markets	0.155	0.153	0.115	0.115	0.274	0.274
partial-R^2, sci. homoph.	0.024	0.025	0.001	0.001	0.001	0.001
Obs., mjt	1,953,818	1,953,818	1,953,818	1,953,818	224,224	224,224

Notes: Reports estimates of the main regression using either linear (cols. 1-2), binary (cols. 3-4), or log (cols. 5-6) transformations of both the dependent and independent variables. Standard errors are in parentheses and are clustered at the MeSH (m) level. Cols. 1, 3, and 5 reflect the main specifications where the MeSH-shares of editorial boards are aggregated at the publication-level. Cols. 2, 4, and 6 reflect an alternative specification where the MeSH-shares of editorial boards are first aggregated at the publication-level within each person, and then again at the person level.

D.2 Heterogeneity

Effects by Publication Types

Our results thus far are based on MeSH share variables (S^P and S^E) that are derived from *all types* of publications in journals and written by authors. However, the PubMed data does allow us to make one major distinction in terms of the nature of these publications in that we can separate "research articles," as in the standard, peer-reviewed article that comprises most of these journals, from other article types such as editorials, comments, and letters. These latter publications are rarely peer-reviewed, but they do represent an opportunity for us to see if the effect we identify is driven more by editors' scientific expertise on certain topics *per se* (which we expect to be more represented in research articles) or instead by their preferences or beliefs about which topics are germane (which we expect to be relatively more represented in their non-research publications).

In Table D.7, we report results from separating both journals' and editors' publications into research versus non-research types. Overall, it appears that the main effect we observe is driven by editors' research publications, and that the effect is relatively similar on both the research and non-research publications in their journals. Had the scientific homophily effect been driven by editors non-research publications, concern might be warranted regarding these preferences. But given that it appears the effect is driven almost entirely by the research publications, we see this as yet another piece of supporting evidence that the scientific homophily effect is likely not a first-order cause for concern.

Effects over Time and by Topic Age

Figure D.4 reports the estimates we obtain when we allow the scientific homophily parameter to vary by year. Regardless of the different sample inclusion criteria we use, we find evidence that, if anything, the magnitude of the effect is declining over time to the point of not reaching statistical significance. We lack a strategy for shedding light on why the effect may be declining over time.

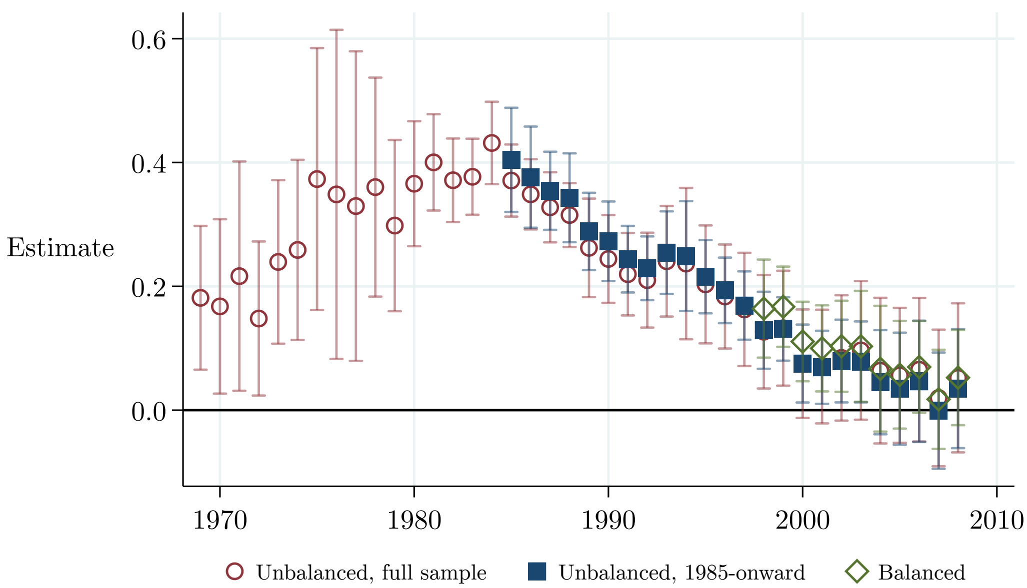
Another temporal dimension along which one might expect to observe heterogenous effects is the "age" of the topic at hand – are editors more or less likely to tilt journals' towards their research when the science is particularly new (or old)? We investigate this question by proxying for a topic's age 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, reported in Table D.8, are not very conclusive and generally do not reveal any consistent patterns. Our approach is limited by our data, and is crude at best. Still, the fact that that no clear pattern emerges suggests that there are likely no first-order differences along this dimension.

Table D.7: Effects by Type of Publication: Research Articles vs. Other (e.g., Editorials, Letters)

	S^P_{mjt}, all	S^P_{mjt}, articles	S^P_{mjt}, other
	(1)	(2)	(3)
S^E_{mjt}, articles	0.214***	0.199***	0.259***
	(0.028)	(0.026)	(0.053)
S^E_{mjt}, other	0.018	0.021***	0.023*
	(0.011)	(0.008)	(0.013)
Elasticity at means, articles	0.214	0.199	0.271
Elasticity at means, other	0.018	0.020	0.024
partial-R^2, sci. homoph., articles	0.019	0.017	0.003
partial-R^2, sci. homoph., other	0.001	0.001	0.000
Obs., mjt	1,953,818	1,953,818	1,953,818

Notes: Standard errors are in parentheses and are clustered at the MeSH (m) level.

Figure D.4: Year-specific Scientific Homophily Effect Estimates



Notes: Plots the point estimates and 95% confidence intervals from estimating the main regression specification (incl. journal-mission and topic-market fixed effects), but allowing the scientific homophily effect to vary by year. The "Unbalanced, full sample" includes all journal-years for which we obtained data. The "Unbalanced, 1985-onward" is the sample used in all other regressions. The "Balanced" sample includes all years from 1998 onward, since this is the earliest year we could obtain publication-matched editorial data for all in-sample journals.

Table D.8: Effects by Age of MeSH

	Per min. (1)	Per mean (2)	Per max (3)
S^E_{mjt} × Young	0.215***	0.294**	-0.046
	(0.028)	(0.126)	(0.663)
S^E_{mjt} × Middle	0.328*	0.215***	0.395
	(0.197)	(0.048)	(0.299)
S^E_{mjt} × Old	0.218***	0.235***	0.216***
	(0.029)	(0.032)	(0.021)
Total R^2	0.952	0.952	0.952
Obs., mjt	1,953,818	1,953,818	1,953,818

Notes: Standard errors are in parentheses and are clustered at the MeSH (m) level. Each column corresponds to an alternative statistic used to summarize the age of an aggregated MeSH term based on the year each unique (non-aggregated) MeSH term within an aggregation was introduced into the hierarchy. "Young," "Middle," and "Old" refer to roughly tercile-based splits of each statistic.

Table D.9: Citation-weighted Results

	Linear-Linear		Binary-Binary		log-log	
	Raw (1)	Cite-wgt. (2)	Raw (3)	Cite-wgt. (4)	Raw (5)	Cite-wgt. (6)
S^E_{mjt}	0.230***	0.114***	0.021***	0.016***	0.046***	0.037***
	(0.029)	(0.030)	(0.001)	(0.001)	(0.004)	(0.008)
Elasticity at means	0.230	0.129	0.044	0.038	0.046	0.037
Total R^2	0.952	0.898	0.611	0.588	0.803	0.691
partial-R^2, mj-missions	0.646	0.568	0.323	0.318	0.499	0.364
partial-R^2, mt-markets	0.155	0.113	0.115	0.112	0.274	0.261
partial-R^2, sci. homoph.	0.024	0.003	0.001	0.000	0.001	0.000
Obs., mjt	1,953,818	1,953,818	1,953,818	1,953,818	224,224	192,930

Notes: Reports estimates of the main regression using either linear (cols. 1-2), binary (cols. 3-4), or log (cols. 5-6) transformations of both the dependent and independent variables. Standard errors are in parentheses and are clustered at the MeSH (m) level. Cols. 1, 3, and 5 reflect the main specifications where the MeSH-shares are unweighted. Cols. 2, 4, and 6 reflect an alternative specification where the MeSH-shares of publications are weighted by the number of forward-citations-per-year a publication receives.

Table D.10: Main Results using Affiliation Terms

	(1)	(2)	(3)	(4)
Panel (a): D.V. = S^P_{ajt}				
S^E_{ajt}	0.707***	0.102***	0.478***	0.052***
	(0.057)	(0.030)	(0.057)	(0.005)
Elasticity at means	0.707	0.102	0.478	0.052
Total R^2	0.754	0.967	0.832	0.977
partial-R^2, aj-missions		0.867		0.862
partial-R^2, at-markets			0.315	0.288
partial-R^2, sci. homophily	0.754	0.044	0.391	0.014
Obs., ajt	5,681,088	5,681,088	5,681,088	5,681,088
Panel (b): D.V. = $\mathbf{1}\{S^P_{ajt} > 0\}$				
$\mathbf{1}\{S^E_{ajt} > 0\}$	0.637***	0.064***	0.245***	0.020***
	(0.002)	(0.001)	(0.003)	(0.001)
Elasticity at means	0.096	0.010	0.037	0.003
Total R^2	0.087	0.551	0.409	0.602
partial-R^2, aj-missions		0.508		0.315
partial-R^2, at-markets			0.353	0.100
partial-R^2, sci. homophily	0.087	0.001	0.015	0.000
Obs., ajt	5,681,088	5,681,088	5,681,088	5,681,088
Panel (c): D.V. = $log(S^P_{ajt})$				
$log(S^E_{ajt})$	0.708***	0.075***	0.362***	0.050***
	(0.012)	(0.003)	(0.008)	(0.003)
Elasticity at means	0.708	0.075	0.362	0.050
Total R^2	0.435	0.888	0.786	0.929
partial-R^2, aj-missions		0.803		0.665
partial-R^2, at-markets			0.621	0.356
partial-R^2, sci. homophily	0.435	0.006	0.198	0.003
Obs., ajt	214,197	214,197	214,197	214,197
Incl. γ_{aj}		Y		Y
Incl. σ_{at}			Y	Y

Notes: Standard errors are in parentheses and are clustered at the affiliation term (a) level. The bottom two rows apply to all panels and indicate which columns are based on specifications that include affiliation-journal (aj) and/or affiliation-year (at) fixed effects.