Leadership in Scholarship: Editors' Influence on the Profession's Narrative*

Ali Sina Önder Sergey V. Popov Sascha Schweitzer

December 19, 2019

Abstract

Academic journals disseminate new knowledge, and therefore can influence the direction and composition of ongoing research by choosing what to publish. We study the influence of editors and coeditors of the *American Economic Review (AER)* on the topic structure of papers published in the *AER* between 1976 and 2013 using a textual analysis of manuscripts. We compare *AER*'s topic structure to that of the other top general interest journals. The appointment of new *AER* editors, while accompanied by a minor comovement of *AER* topics towards topics of editor's post-appointment publications, is not an indicator of editor's personal taste in topics, but rather indicates the desire of those who appoint editors to premediate trends in other Top 5 journals.

JEL classification: A11, A14, O3

Keywords: Academia; Knowledge Dissemination; Influence; Journals; Editors; Publications; Latent Dirichlet Allocation; Topic Analysis.

*Önder: University of Portsmouth, ali.onder@port.ac.uk. Popov: Cardiff Business School, PopovS@cardiff.ac.uk. Schweitzer: University of Bayreuth, Sascha.Schweitzer@uni-bayreuth.de. We thank all our readers and audiences of seminars in Belfast, Cardiff, UIB, Bayreuth, and UCD, especially Marian Moszoro, Joshua Angrist, Rafel Crespí Cladera, Stefan Seifert, Ivan Pastine, Dan Hamermesh, Stas Kolenikov, and Yanos Zylberberg. The usual disclaimer applies.

1 Introduction

Publishing in top economics journals is increasingly competitive (Hamermesh, 2013) and extremely rewarding (Attema et al., 2014). Short-term rewards, such as promotions and grant awards, are prone to depend not only on publication content, but also on the journal prominence and publication counts (Heckman and Moktan, 2018). This creates a tradeoff between publishing what one thinks is important and what one thinks is likely to be published.[1] A new editor taking office in an influential journal may motivate researchers who seek recognition to steer knowledge generation towards the topics preferred by this editor. How strongly is the topic structure of a journal driven by editors' preferences in their own research?

To answer this question, we study the appointment of editors and coeditors of the *American Economic Review* (*AER*) taking office between 1985 and 2011.[2] We employ a high-detail textual analysis on the full texts of individual articles to identify the topics that emerge in the *AER* and the other leading general interest journals.[3]. We analyze how topic frequencies in the published research of a newly appointed editor co-move with topic frequencies observed in the *AER* before and after that editor's appointment. The other Top 5 constitute our control group.

We establish that, from the beginning, editors appointed to the *AER* tend to be more aligned with the *AER* and diverge from the other Top 5 in their topic profiles. During their term at the *AER*, editors' newly published topics continue to be positively related to those in the *AER*. This relationship persists when the time window of our analysis is altered; the coefficients are also qualitatively robust to changes in topic counts. We remain agnostic about cause and effect: editors could be appointed to lead the way to develop a research profile

[1]Ruhm (2018) argues methodological requirements might avert scholars away from important topics.

[2]Editors and coeditors wield equal decision-making power in the *AER*, whereas associate editors do not. We thank Dan Hamermesh for pointing this out, and past editors of the *AER* for confirmation. In the rest of this paper, we refer to editors as well as coeditors as *editors*.

[3]Namely, the *Quarterly Journal of Economics* (*QJE*), the *Journal of Political Economy* (*JPE*), *Econometrica*, and the *Review of Economic Studies* (*REStud*). These journals, together with the *AER*, make up the top group of the journal ranking documented by Combes and Linnemer (2010). Moreover, these are the conventional Top 5 economics journals that most academic economists would agree with. (cf Heckman and Moktan, 2018). In what follows, we refer to the above four leading general interest journals (Top 5 excluding the *AER*) as the *other Top 5*.

that continues to distinguish the *AER* from the other Top 5, or the development of the *AER* profile could have already been anticipated in the submission decision of authors.

2 Literature Review

We contribute to the empirical literature on knowledge dissemination by showing that editors can affect the profession, not only through their professional networks and their ties (Brogaard et al., 2014, Card and DellaVigna, 2017, Colussi, 2018, Medoff, 2003), but also through their influence on the topics and the narrative structures that appear in journals.

In our preliminary analysis in Section 3.2, we investigate the dynamics of topics covered by papers published in the *AER*, and, using topics suggested by machine learning instead of JEL codes, we obtain patterns similar to those documented in Figure 7 of Card and DellaVigna (2013) and in Figure 2 of Angrist et al. (2017). While the JEL codes are quite generic, there is little clarity about their persistence: it is not clear, for instance, if a paper on job market signaling would be best categorized as a micro paper, a labor paper, or both, with 50-50 allocation; and whether the decision regarding the allocation of such a paper to JEL codes would be the same in the 1990s and in the 2010s. When new topics arise or old topics fade away, the pre-defined JEL classifications are hardly ever adapted accordingly. Thus, new topics may be disguised under either very generic or rather odd JEL codes. Over time, this can lead to the overcrowding of some classes and the depopulation of others. Even a reform of the classification system, such as the one in 1990, brings inconsistencies of its own that complicate the investigation of the continuous development of topics (Cherrier, 2017).

Our approach does not suffer from this problem. It continuously tracks changes in topics and terminology, with no sudden artificial breaks. As long as the terminology persists, topics are assigned in the same way. Glandon et al. (2018) avoid using JEL codes in their analysis and classify macroeconomic papers manually, because JEL codes cannot capture the many nuances of different research areas within macroeconomics. For instance, they document that DSGE methodology became more prominent, so what constitutes macroeconomics changes in time; meanwhile the proportion of macroeconomic papers, according to Angrist et al. (2017), remained the same.

An overview of the methodology and research applications of textual analysis is described in Gentzkow et al. (2017). Analysis of the similarity between different text data has been used in various settings. For instance, Li (2017) investigates the quality of NIH grant applications by using a similarity measure between texts of NIH grant applications and publications. It becomes possible to determine which publications are directly linked to a specific NIH grant. We use a similar text analysis that quantifies the vectors of topic frequencies of all publications in the *AER*, in the other Top 5 and in editors' own publications, in order to measure topic similarity.

In studying publication patterns, a methodology similar to ours was applied by Mela et al. (2013) and Huber et al. (2014) to marketing literature. While they show that editors throughout their tenure feature different mixes of topics, they do not speculate as to why the topics of the text corpus moved in a certain direction. In economics, Angrist et al. (2017) study the development of economic literature over time. While finding little evidence for change in the composition of economics fields, they demonstrate a greater propensity for publishing empirical literature. Their analysis does not extend to studying whether or not the frequencies of topics of the journal co-move with the topic frequencies of the editors' own work. Kosnik (2015) uses topical analysis to study the corpus of seven journals in economics[4] published between 1960 and 2010. While this study finds suggestive evidence that research in macroeconomics diminishes, complemented by an increase in research in the microfoundations of macroeconomics, it does not concern editors' appointment, and does not compare trends across different journals. Ambrosino et al. (2018) uses all economics journals in JStor, but does not inquire into the editor's influence. Kosnik (2018) asks whether or not JEL codes are informative, and applies textual analysis to papers that share the same JEL code (using about 10 topics per JEL code), but does not study the dynamics of topics.

[4]The usual Top 5 (as we use in this paper as well) plus the Journal of Economic Literature and Journal of Economic Perspectives, both of which are by invitation only and therefore have significantly different incentive structures in the author-editor relationship.

3 Data and Methodology

We study the corpus of texts in the *AER*, *QJE*, *JPE*, *REStud*, and *Econometrica*, and all articles written by *AER*'s editors between 1976 and 2013 which are available at the JStor. We obtain our data from ITHAKA, the owners of JStor, the digital online library, which provides word and n-gram counts of academic papers for researchers[5]. We compare trends in topic frequencies in articles published by newly appointed editors of the *AER* who took office between 1985 and 2011 against topic frequencies observed in articles published in the *AER* and also those published in the other Top 5.

A topic in our context is not necessarily the same as something considered a field or a subfield in Economics research. A topic can be a field, or an aspect of a field, and it can even be a certain style of narrative that features distinct patterns that is picked up by our textual analysis.

3.1 Topic Analysis

The methodology of the analysis is based on reducing the inherently high dimensionality of textual data. This approach shares some similarities with principal components analysis: words (or combinations of words, such as "sovereign debt") that occur together with other specific words (such as "default") in many texts are likely to carry the same narrative purpose.

We preprocess full texts of research articles in our data through several technical steps. In the first step, common words are removed (such as "a", "above", "across", etc; full list of stop words available on request). In the second step, words are stemmed in order to abstract them from their different grammatical forms. The stemming procedure follows the standard approach described by Porter (1980). Finally, common multiple-word collocations (such as "nited States of America") are replaced by tokens. For the tokenizing, we employ the Python package `textmining` (Peccei, 2010). All of these preprocessing steps were performed using a Python script that is available on request.

[5] Data are provided by ITHAKA for research purposes upon request via `http://dfr.jstor.org/`, accessed 1 June 2017.

After preprocessing the text data, the topic analysis was performed using Latent Dirichlet Allocation (LDA)[6] model. Each topic is a probability distribution over words that are encountered in the whole text corpus. For each manuscript, LDA returns a list of mixing proportions: each document is modelled as a mixture distribution over topics, and therefore over words, and different documents have different topic loadings. An advantage of this methodology is that it is not driven by hand-picked sets of words ("unsupervised"): topics are constructed to fit a model consisting of a mixture of distributions over words, subject to a pre-specified number of topics. Our ex-ante specification is based on 200 topics; results remain qualitatively similar if the number of topics is increased (in which case additional topics become more specific, potentially containing more uninformative artifacts) or decreased (which makes topics more general, potentially concealing changes in time). We used the UMass Amherst's Machine Learning for Language Toolkit (MALLET) (McCallum, 2002) to carry out the estimation.[7]

3.2 Trends in Topics of the AER

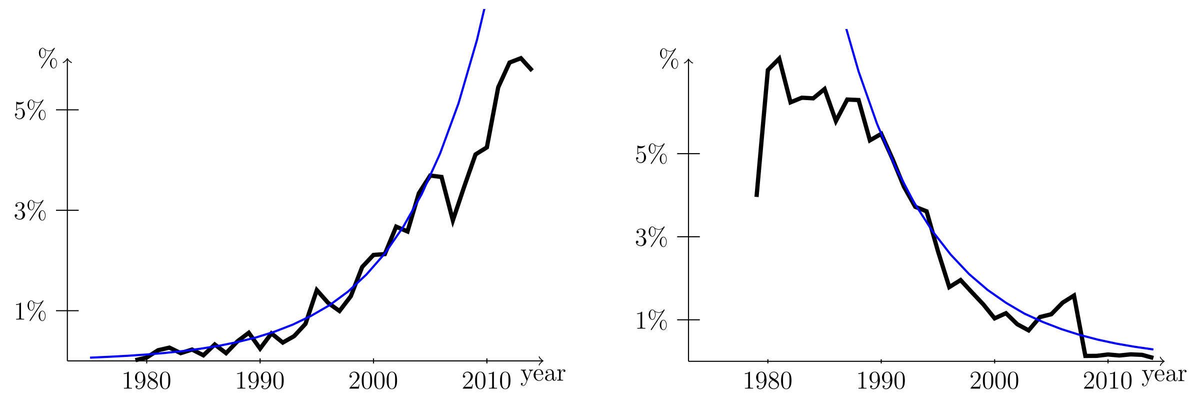
The topic analysis yields the topic frequencies in each article as well as the distribution of words in each topic. We omitted 5 topics that were clearly technical. The most popular topic overall constitutes around 4.5% of the corpus; 39 topics cover around 50% of the corpus.

Over time, trends may change: some topics can proliferate, while other topics may wither. To test for time trends in topics, we ran a time series regression for each topic[8], regressing a log of share of each topic on time and time-squared, with topic-specific coefficients. Then we conducted 195 F-tests to see whether the time trend was statistically significant, and kept the p-value of this test. Under the null hypothesis of no quadratic time trend across topics, the distribution of $p-$values should be close to uniform. In fact, it is not: the average $p-$value is somewhat less than 0.021, and 84.6% of topics have a $p-$value less than 0.01. A similar result is obtained if one attempts a panel regression with individual time trends:

[6]See Blei et al. (2003) for elaboration of the LDA machinery, and Ambrosino et al. (2018) on the interpretation of the topic loadings.

[7]Available at `https://mallet.cs.umass.edu`, accessed 1 June 2017.

[8]We used a four year window one year lag setting for this; similar results are obtainable for other settings. This allows us to use factor loadings from 1979 till 2014.



(a) Topic 42 (b) Topic 62

Figure 1: Topics change over time

the F-test value is 9836, which with degrees of freedom of 195×2 and 195×33 yields a numerically zero $p-$value. Implementing corrections (such as adjusting for non-normality, etc) could obviously increase the $p-$value.

Among individual topics, topic 42's linear slope coefficient is 0.1342. This topic includes stems such as

> `effect estim year result column us tabl control specif data sampl regress includ panel level coeffici fix-effect differ measur report`

and its share in *AER* publications increases in time, going from 0.14% of the text corpus in the late 1970s to 5.5% in the early 2010s. Meanwhile, topic 62's linear slope coefficient is -0.1229; it includes stems such as

> `tion re ing ment con ex vol de behavior com di iti paper exampl creas econom chang analysi eco-nomic robert`

and it accounts for 6.2% of the *AER* publications in late 1970s, but only for 0.1% of the text corpus in the late 2010. This does not necessarily mean that authors used the word `example` in 2010s less than they did before, it means that this characteristic accumulation of words tended to be part and parcel of a text more frequently before 2000 than afterwards. Both trends are plotted in Figure 1.

The nature of our topic data induces some of the trends: if there is a strong trend in one topic, there will be an opposite trend in the total loading of other topics, which is why it is hard to say which changes cause which other changes. We apply the Benjamini-Hochberg-Yekuteli algorithm[9] to choose a critical value to limit our false discovery rate from above by 1%, and still there are 165 topics that seem to exhibit a quadratic trend, and these topics cover about 89% of the corpus (if we just went with 1% significance, that would be 93% of the corpus). Therefore, it is safe to say that over 1979–2014 at least some changes in topics occurred in the papers covered by our corpus. Because our topics are narrower than the subfields of Economics, we detect some changes in the narrative that could not be captured by a coarser grouping methodology a lá Angrist et al. (2017).

3.3 Assigning Documents To Editors

We employ the topic frequencies of journals and editors based on three, four, and five year windows before and after an editor's tenure in our main analysis[10]. As already been pointed out by Ellison (2002) there are significant time lags between the crafting of a research paper and its actual publication. To accommodate publication lags, we compare results for one and two year lags. This means that with a three year window and one year lag, the editor appointed in 2000 is relevant for papers published in 2001, 2002, and 2003 (plus maybe additional years, but we deliberately do not include further years to study the effect of the appointment only); and we compare the topic loadings of these papers to topic loadings of papers published in 1998, 1999, and 2000.

The document sets and their notations are as follows: AER, $Top5$, and $Editor_i$ denote the AER, the other Top 5, and a specific editor i, respectively. $AER^c_{i,pre}$ and $AER^c_{i,post}$ denote the average frequency of topic c in articles published in the AER before and during tenure, respectively, of editor i in the AER. Similarly, $Top5^c_{i,pre}$ and $Top5^c_{i,post}$ denote the average frequency of topic c in articles published in the other Top 5 before and after the appointment, respectively, of editor i at the AER. The average frequency of topic c in articles written by editor i before and after her/his appointment at the AER is denoted by $Editor^c_{i,pre}$ and $Editor^c_{i,post}$, respectively. We take logarithms of all variables so that outliers are tamed and

[9]We use the conservative approach that allows for arbitrary dependence across outcomes of our tests, following Theorem 1.3 in Benjamini and Yekutieli (2001).

[10]A complete list of the AER's editors and coeditors covered in our analysis can be found in Table A.1 in the Appendix.

Table 1: Pairwise Correlations of Editors' and Journals' Topics using Four Year Window and One Year Lag

	$Editor^c_{i,post}$	$Editor^c_{i,pre}$	$AER^c_{i,post}$	$AER^c_{i,pre}$	$Top5^c_{i,post}$
$Editor^c_{i,pre}$	0.585***				
$AER^c_{i,post}$	0.661***	0.656***			
$AER^c_{i,pre}$	0.664***	0.658***	0.984***		
$Top5^c_{i,post}$	0.660***	0.645***	0.959***	0.955***	
$Top5^c_{i,pre}$	0.663***	0.653***	0.961***	0.959***	0.984***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

regression coefficients can be interpreted as respective elasticities. The difference between topic frequencies of the *AER* and the other Top 5 during the tenure of editor i is denoted $(AER - Top5)^c_{i,post}$.

3.4 Estimation

The unit of observation in our regression analysis is an editor-topic pair. When a four year window and a one year lag is used there are 2,925 editor-topic pairs. Table 1 shows the correlation coefficients of main variables we obtain from the textual analysis using a four year window and a one year lag.

We use OLS and two step LS (2SLS) estimations to investigate correlations between editors' and journals' topic frequencies. We regress topic frequencies observed in the *AER* and the other Top 5 during the tenure of an editor on her/his preference for topics and journals' topic frequencies which are observed prior to that editor's tenure. We not only control topic frequencies of the *AER* and the other Top 5 during editor i's tenure for editor's preferences but we control also for topic frequencies observed in the *AER* and the other Top 5 before editor i's tenure. Any discrepancy in topic frequencies of the *AER* and the other Top 5 may lead to a realignment in the next period, i.e. during editor i's tenure, independent of editor i's personal preferences. In particular we estimate:

$$AER^c_{i,post} = F_A(\textbf{Editor Preference}^c_i, AER^c_{i,pre}, Top5^c_{i,pre})$$

$$Top5^c_{i,post} = F_T(\textbf{Editor Preference}^c_i, AER^c_{i,pre}, Top5^c_{i,pre})$$

$$(AER - Top5)^c_{i,post} = H(\textbf{Editor Preference}^c_i, AER^c_{i,pre}, Top5^c_{i,pre})$$

where $\textbf{Editor Preference}^c_i$ is captured either by editor's topic frequencies prior to taking office or during her/his tenure at the *AER*.

Editors' topic frequencies during their tenure, however, might be influenced by topic frequencies observed in the *AER* or at the other Top 5 during that time. This poses the problem of endogeneity, and we use 2SLS to avoid this problem. That way we are able to isolate variations in topic frequencies of an editor's own research during her/his tenure to what can be explained by variations in topic frequencies observed before he/she has taken office at the *AER* either in her/his own research or in journal publications. In particular, we estimate

$$Editor^c_{i,post} = \beta_0 + \beta_1 Editor^c_{i,pre} + \beta_2 AER^c_{i,pre} + \beta_3 Top5^c_{i,pre} + \psi^c_i$$

and we obtain fitted values for editor i's topic frequencies during his/her tenure, denoted by $Editor^{c,fitted}_{i,post}$ which we refer to as the fitted topic frequency or the *fitted preference* of editor i. In the second stage, we use editor i's fitted preference as an independent variable in the estimation of topic frequencies in the *AER* and in the other Top 5 during editor i's tenure.

4 Results

We start with topic frequencies obtained from the textual analysis of a four year window with a one year lag. The list of editors included in this analysis is restricted to those who have been in office at least for the full length of the window and have sufficient text data for the textual analysis. For the rest of this paper, *post-tenure* refers to the time window (including lag) after the editor took office, and *pre-tenure* refers to that before they took office.

Estimated coefficients shown in Table A.2 reveal that editors' pre-tenure topic frequencies (for brevity, referred to as *topics*) are significantly and positively related to pre-tenure topics of the other Top 5, hence editors topics before their tenure at the AER strongly align with topics in the other Top 5. However, they have no significant relation to *AER*'s pre-tenure topics (column (1)). Editors' pre-tenure topics have also no statistically significant relation to *AER*'s or Top 5's post-tenure topics (columns (3) to (6)).

Table 2: Journals' Topics and Editor's Preference with Four Year Window and One Year Lag

	$Editor_i^c$		$AER_{i,post}^c$			$Top5_{i,post}^c$			$(AER$ - $Top5)_{i,post}^c$		
	(1)pre	(2)post	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$AER_{i,pre}^c$	-0.130	0.568**	0.197***	0.195***	0.193***	0.199**	0.197**	0.200**	-0.00198	-0.00185	-0.00714
	(0.225)	(0.181)	(0.0516)	(0.0514)	(0.0511)	(0.0728)	(0.0729)	(0.0728)	(0.0767)	(0.0769)	(0.0769)
$Top5_{i,pre}^c$	0.471**	0.347	0.276***	0.275***	0.274***	0.353***	0.350***	0.353***	-0.0762	-0.0749	-0.0794
	(0.175)	(0.210)	(0.0499)	(0.0495)	(0.0497)	(0.0626)	(0.0623)	(0.0627)	(0.0761)	(0.0763)	(0.0761)
$Editor_{i,pre}^c$		0.240***	0.00167			-0.000519			0.00219		
		(0.0203)	(0.00174)			(0.00157)			(0.00221)		
$Editor_{i,post}^c$				0.00429*			0.00506*			-0.000767	
				(0.00173)			(0.00210)			(0.00237)	
$Editor_{i,post}^{c,fitted}$					0.00695			-0.00216			0.00910
					(0.00724)			(0.00655)			(0.00921)
Topics FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2925	2925	2925	2925	2925	2925	2925	2925	2925	2925	2925
R^2	0.496	0.574	0.984	0.984	0.984	0.981	0.981	0.981	0.671	0.671	0.671

Standard errors in parentheses

$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Post-tenure topics of the *AER* as well as the other Top 5 are significantly and positively correlated with journals' (*AER* and the other Top 5) pre-tenure topics and with editors' post-tenure topics (columns (4) and (7)). Moreover, the coefficient for editors' post-tenure topics has a larger point estimate in case of the other Top 5 compared to its point estimate in case of the *AER*.

Editors are appointed to lead the way in which the research narrative unfolds in the corresponding journal. This is especially important when top journals are concerned. However, it is unclear whether editors lead the way by imposing their own pre-tenure preferences or whether they are affected by submissions or trends set by other major journals during their tenure. This endogeneity can be a problem when regressing topics of the *AER* on editors' topics during their tenure.

Editors' fitted preferences are obtained from regressing editors' post-tenure topics on editors' and journals' pre-tenure topics. Coefficient estimations from this regression are shown in column (2) of Table 2. Editors' post-tenure topics are positively related to their own and *AER*'s pre-tenure topics. We find no significant relation between editors' post-tenure and the other Top 5's pre-tenure topics. Editors' and the other Top 5's pre-tenure topics are highly correlated (column (1)) and the insignificant relation between editors' post-tenure topics and the other Top 5's pre-tenure topics shown in column (2) could mean that pre-tenure topics of the other Top 5 correlate with editors' post-tenure topics only to the extent they are contained in editors' pre-tenure topics. Fitted values for editors' post-tenure topics will this be a linear projection of the *AER*'s and editors' pre-tenure topics. As can be seen in columns (5) and (8), we obtain no significant partial correlation between editors' fitted preferences and the *AER*'s and the other Top 5's post-tenure topics, respectively. Neither journals' pre-tenure topics nor editors' pre- or post-tenure topics are significant in explaining the difference between post-tenure topics of the *AER* and the other Top 5 (columns (9) to (11).

We consider a longer publication lag and investigate a four year window with two year lag in Table 3. Editors' pre-tenure topics significantly correlate with pre-tenure topics of the *AER* as well as the other Top 5 (column (1)). Editors' post-tenure topics are significantly related to their pre-tenure topics and not journals' pre-tenure topics, as shown in column (2),

Table 3: Journals' Topics and Editor's Preference with Four Year Window and Two Year Lag

	$Editor^c_i$		$AER^c_{i,post}$			$Top5^c_{i,post}$			$(AER - Top5)^c_{i,post}$		
	(1)pre	(2)post	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$AER^c_{i,pre}$	0.394^{+}	0.271	0.238^{***}	0.238^{***}	0.237^{***}	0.176^{**}	0.175^{**}	0.173^{**}	0.0620	0.0631	0.0642
	(0.200)	(0.198)	(0.0525)	(0.0524)	(0.0525)	(0.0535)	(0.0535)	(0.0541)	(0.0637)	(0.0636)	(0.0640)
$Top5^c_{i,pre}$	0.422^{+}	-0.0580	0.225^{***}	0.225^{***}	0.225^{***}	0.333^{***}	0.334^{***}	0.333^{***}	-0.108	-0.108	-0.108
	(0.224)	(0.197)	(0.0481)	(0.0482)	(0.0482)	(0.0705)	(0.0705)	(0.0705)	(0.0758)	(0.0758)	(0.0758)
$Editor^c_{i,pre}$		0.235^{***}	0.00108			0.00298			-0.00191		
		(0.0273)	(0.00224)			(0.00289)			(0.00360)		
$Editor^c_{i,post}$				0.00142			0.00625^{**}			-0.00483^{+}	
				(0.00164)			(0.00216)			(0.00249)	
$Editor^{c,fitted}_{i,post}$					0.00458			0.0127			-0.00810
					(0.00954)			(0.0123)			(0.0153)
Topics FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2925	2925	2925	2925	2925	2925	2925	2925	2925	2925	2925
R^2	0.560	0.541	0.985	0.985	0.985	0.976	0.976	0.976	0.571	0.571	0.571

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

meaning that journals' pre-tenure topics are related to editors' post-tenure topics only to the extent they are embodied in editors' pre-tenure topics.

We find no statistically significant relation between editors' post-tenure topics and *AER*'s topics (columns (3) to (5)). Post-tenure topics of the other Top 5 are positively and significantly correlated with editors' post-tenure topics (column (7)), yet the significance is lost when we use fitted values for editors' post-tenure topics (column (8)). Furthermore, we find that the difference between journals' topic frequencies is also significantly related to editors' topics (column (10)) when a two year lag is considered.

Estimations using three and five year windows with lags of one and two years are shown in tables A.2, A.3, A.4, and A.5 in the Appendix. Five year window yields comparable results to those of the four year window. When the editor serving three years are included, then we obtain positive and significant relation between *AER*'s and editors' pre-tenure topics (Table A.3). Using a three year window with a one year lag we find negative and significant relation between editors' topics and topics of the other Top 5, as shown in Table A.2 columns (6) and (8). This is a hint that topics of the *AER* and the other Top 5 might be diverging due to editors' influence. The divergence of topics between *AER* and the other Top 5 is further confirmed when journals' differences in topic frequencies is regressed on editors' topics (columns (9) and (11)).

This finding can be interpreted as an alignment in topic preferences of the *AER* and its editors. Editors that remain in the office for three years and no more probably star on niche topics and are hired to make sure that the *AER* keeps up its line of publications that slightly diverge from existing topic trends in the other Top 5 journals. When editors who serve longer than three years are considered as it is the case for our four year and five year window analyses, we find that editors' topics align more with those of the other Top 5 than the *AER*. A possible reason is that these topics are broader topics with possibly larger impact. Hence editors who serve four years or longer contribute in such fields A possible interpretation might be that editors are hired to make sure that the *AER* keeps up its themes in line with existing topic trends so that their tenure at the *AER* serves the purpose that the *AER* does not miss out trending topics that have already taken off in the other Top 5.

5 Conclusion

We use textual analysis to quantify the topic frequency in the narrative of publications in the *AER* and ask if and how they align with the content of editors' individual publication portfolios. We find that topic frequencies that are observed in the *AER* align with those observed in editors' own publications while being an editor, but not much driven by editor's publications before becoming an editor. The size of the effect is quite small, amounting to a replacement of 1–3 *regular* papers in 1000 by a paper that is devoted only to the newly appointed editor's interests. Obviously, this could also mean that the papers submitted to the *AER* now have on average 0.1%–0.3% more irrelevant verbiage targeted at the new editor. This looks large; this is because most editors' work is not too far from what was getting published in the *AER* before their appointment, so 0.1-0.3% is the estimate of the appointment effect from below. However, for the natural reason of the secrecy covering author-editor relationships, we know neither the editors who were handling individual papers nor what was rejected by the very same editors. While the effect of the latter is unclear, the effect of the former clearly will make our coefficients biased towards zero. Our topic assignment is data-driven, not coming from a training dataset or heuristics, though either could have provided us with a better measure of topic dynamics; again, however, this would have biased the coefficients that we obtain towards zero. Heterogeneity in editors—some editors may be more prone to impose their own agenda, and some may be less—will add noise to our estimates, making our coefficients look statistically less significant, but will not alter the sign of the average effect.

We find that *AER*s topic frequencies align with those observed in editors' own publications while being an editor, which align with the topics of the other Top 5 before becoming an editor. Moreover, point estimates for editors topics when regressed on topics of other Top 5 are larger. Our favorite interpretation of these estimates is that editors are hired to make sure that the *AER* keeps up its line of publication topics in line with what is trending in the other Top 5 journals.

We provide estimates on multiple horizons because shorter horizons suffer less from the supply side issues (the academia can respond to an appointment by producing more papers in related fields), but longer horizons make sure that the new appointment had enough time to influence the publications. We cannot distinguish the decisions that the new editor makes from the decisions that other editors are making, either compensating for the new appointee's possible biases or embracing the new trends in the profession. Our data does not allow us to look inside of the black box of the editorship of the AER, but it does allow us to see that innovations in that black box does not seem to change the structure of the output beyond what was predictable from the deviation of the *AER* output from the rest of the Top 5.

References

Ambrosino, A., Cedrini, M., Davis, J. B., Fiori, S., Guerzoni, M., and Nuccio, M. (2018). What topic modeling could reveal about the evolution of economics. *Journal of Economic Methodology*, 25(4):329–348.

Angrist, J., Azoulay, P., Ellison, G., Hill, R., Lu, S. F., et al. (2017). Economic research evolves: Fields and styles. *American Economic Review*, 107(5):293–297.

Attema, A. E., Brouwer, W. B., and Van Exel, J. (2014). Your right arm for a publication in AER? *Economic Inquiry*, 52(1):495–502.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.*, 29(4):1165–1188.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022.

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.

Card, D. and DellaVigna, S. (2013). Nine facts about top journals in Economics. *Journal of Economic Literature*, 51(1):144–161.

Card, D. and DellaVigna, S. (2017). What do editors maximize? Evidence from four leading economics journals. Working Paper 23282, National Bureau of Economic Research.

Cherrier, B. (2017). Classifying economics: A history of the JEL codes. *Journal of Economic Literature*, 55(2):545–579.

Colussi, T. (2018). Social ties in academia: a friend is a treasure. *Review of Economics and Statistics*, 100(1):45–50.

Combes, P.-P. and Linnemer, L. (2010). Inferring missing citations: A quantitative multi-criteria ranking of all journals in economics. Technical report, mimeo.

Ellison, G. (2002). The slowdown of the economics publishing process. *Journal of Political Economy*, 110(5):947–993.

Gentzkow, M., Kelly, B. T., and Taddy, M. (2017). Text as data. Working Paper 23276, National Bureau of Economic Research.

Glandon, P., Kuttner, K. N., Mazumder, S., and Stroup, C. (2018). Macroeconomic research, present and past. Technical report, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3225417.

Hamermesh, D. S. (2013). Six decades of top economics publishing: Who and how? *Journal of Economic Literature*, 51(1):162–172.

Heckman, J. J. and Moktan, S. (2018). Publishing and promotion in economics: the tyranny of the top five. Technical report, National Bureau of Economic Research.

Huber, J., Kamakura, W., and Mela, C. F. (2014). A topical history of JMR. *Journal of Marketing Research*, 51(1):84–91.

Kosnik, L.-R. (2015). What have economists been doing for the last 50 years? a text analysis of published academic research from 1960-2010. *Economics: The Open-Access, Open-Assessment E-Journal*, 9:1–38.

Kosnik, L.-R. (2018). A survey of JEL codes: What do they mean and are they used consistently? *Journal of Economic Surveys*, 32(1):249–272.

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

McCallum, A. K. (2002). MALLET: A Machine Learning for Language Toolkit. http://mallet.cs.umass.edu.

Medoff, M. H. (2003). Editorial favoritism in Economics? *Southern Economic Journal*, 70(2):425–434.

Mela, C. F., Roos, J., and Deng, Y. (2013). A keyword history of Marketing Science. *Marketing Science*, 32(1):8–18.

Peccei, C. (2010). TextMining: Python text mining utilities. [Online; accessed Nov 17, 2017].

Porter, M. (1980). An algorithm for suffix stripping. *Program*, 14(3):130–137.

Ruhm, C. J. (2018). Shackling the identification police? Working Paper 25320, National Bureau of Economic Research.

Appendices

A A Model of Unbiased Change in Topics

To illustrate that the possible driving forces behind our findings do not require biased editing, we design a simple model of editor choice of which paper to publish. Assume there are two topics, indexed by $i \in \{1, 2\}$. Assume each paper can be either good (quality $q = 1$) or bad ($q = 0$), and the paper is good with probability π_i. Assume that at every period the representative editor obtains measure m_i of papers of topic i without knowing their true quality, and then for every paper with quality q of type i, the refereeing process (an interaction of editor's specialties, editor's networks, and the profession's supply of refereeing labour) provides a signal $q + \varepsilon$, where ε is distributed with the cdf $F_i(x)$.

Assume now that the editor selects papers based upon the threshold rule: if the signal is above $\bar{q}$, the paper is accepted, and the paper is rejected otherwise. This leads to the share of papers of topic 1 in the journal to be equal to:

$$\frac{m_1\left[(1-\pi_1)F(\bar{q}) + \pi_1 F(\bar{q}-1)\right]}{m_1\left[(1-\pi_1)F(\bar{q}) + \pi_1 F(\bar{q}-1)\right] + m_2\left[(1-\pi_2)F(\bar{q}) + \pi_2 F(\bar{q}-1)\right]}.$$

If there is a change in the proportion of topics published by the journal, does it have to be driven by the editor's leniency? No: it can be driven by the editor's specialization.

Result 1 *If the distribution of ϵ_i is uniform with support $[-b_i, b_i]$, $b_i > 1$, and $\bar{q} \in (0, 1)$, a marginal increase in b_i increases the proportion of published papers of topic i if $\pi_i < \bar{q}$, and increases otherwise.*

Proof. The probability that a paper of topic i of quality q will get published is:

$$P(q + \varepsilon_i > \bar{q}) = \frac{b_i - (\bar{q} - q)}{2b_i},$$

which leads to the calculation that the proportion of papers of topic i getting published is then:

$$(1-\pi_i)\overbrace{\frac{b_i-(\bar{q}-0)}{2b_i}}^{\text{bad paper is published}}+\pi_i\overbrace{\frac{b_i-(\bar{q}-1)}{2b_i}}^{\text{good paper is published}}=\frac{1}{2}+\frac{\pi_i-\bar{q}}{2b_i}.$$

Taking a derivative with respect to b_i, which is $-(\pi_i-\bar{q})/2b_i^2$, observe that it is negative when $\pi_i>\bar{q}$, and positive otherwise. The increase in the mass of papers of topic i getting accepted will lead to an increased proportion of papers of topic i in the journal. ■

This can be extended to a general setting, with general distributions, adjusting for the editor's choice of $\bar{q}$, having multiple thresholds $\bar{q}_i$ (for either the reason of bias, or a tradeoff between Type I and Type II errors, or both), introducing an endogenous decision of the topic choice or effort choice by the authors, having competing journals, etc. The purpose of this model is to illustrate that even under the simplest assumptions, a change in the refereeing process (an increase in one b_i and a decrease in another) can lead to a change in the composition of accepted papers, even if the editor applies the same acceptance rule to all papers.

Table A.1: List of Editors and Coeditors of the *AER* covered in our Analysis

Name	starting	ending	included when using a Window of Three Years	Four Years	Five Years
Editors : (1985 – 2011)					
Orley Ashenfelter	1985	2001	✓	✓	✓
Ben S. Bernanke	2001	2004	✓	✗	✗
Robert A. Moffitt	2004	2010	✓	✓	✓
Pinelopi K. Goldberg	2011	2016	**	**	**
Coeditors : (1985 – 2011)					
John B. Taylor	1985	1988	✓	✗	✗
Robert H. Haveman	1985	1991	✓	✓	✓
Hal R. Varian	1987	1989	✗	✗	✗
Bennett T. McCallum	1988	1991	✓	✗	✗
Paul R. Milgrom	1990	1993	✓	✗	✗
John Y. Campbell	1991	1993	✗	✗	✗
Roger H. Gordon	1991	1994	✓	✗	✗
Kenneth D. West	1993	1996	*	✗	✗
R. Preston McAfee	1993	2002	✓	✓	✓
Dennis N. Epple	1994	1999	*	✓	✓
Matthew D. Shapiro	1997	1999	✗	✗	✗
Valerie A. Ramey	1999	2002	*	✗	✗
Timothy J. Besley	1999	2004	✓	✓	✓
Orley Ashenfelter	2001	2002	**	**	**
David Card	2002	2004	✗	✗	✗
B. Douglas Bernheim	2002	2005	✓	✗	✗
Richard Rogerson	2003	2008	✓	✓	✓
Judith A. Chevalier	2004	2007	✓	✗	✗
Jeremy I. Bulow	2005	2008	✓	✗	✗
Vincent P. Crawford	2005	2009	✓	✓	✗
Mark Gertler	2005	2010	✓	✓	✓
Pinelopi K. Goldberg	2007	2010	✓	✓	✓
Alessandro Lizzeri	2008	2011	✓	✗	✗
Joel Sobel	2009	2010	✗	✗	✗
Dirk Krueger	2009	2011	✗	✗	✗
Larry Samuelson	2010	2016	✓	✓	✓
Martin Eichenbaum	2011	2014	✓	✓	✗
Andrzej Skrzypacz	2011	2014	✓	✗	✗
Marianne Bertrand	2011	2017	*	✓	✓
Hilary Hoynes	2011	2017	✓	✓	✓
Luigi Pistaferri	2011	2017	✓	✓	✓

(*)Editors who did not publish articles that meet our selection criteria for the duration of a window are not included in the analysis of that window.

(**)P.Goldberg and O.Ashenfelter have served as editor as well as coeditor. They enter our analysis only once at the starting date of either editorship or coeditorship whichever comes first.

Table A.2: Journals' Topics and Editor's Preference with Three Year Window and One Year Lag

	$Editor^c_i$		$AER^c_{i,post}$			$Top5^c_{i,post}$			$(AER - Top5)^c_{i,post}$		
	(1)pre	(2)post	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$AER^c_{i,pre}$	0.218	0.228	0.302***	0.303***	0.301***	0.210***	0.209***	0.213***	0.0928	0.0938	0.0885
	(0.205)	(0.178)	(0.0569)	(0.0570)	(0.0570)	(0.0405)	(0.0406)	(0.0405)	(0.0689)	(0.0688)	(0.0688)
$Top5^c_{i,pre}$	0.0400	0.444*	0.177***	0.176***	0.175***	0.328***	0.327***	0.334***	-0.151*	-0.151*	-0.159*
	(0.173)	(0.189)	(0.0462)	(0.0461)	(0.0465)	(0.0581)	(0.0582)	(0.0582)	(0.0658)	(0.0659)	(0.0662)
$Editor^c_{i,pre}$		0.202***	0.00100			-0.00282*			0.00383*		
		(0.0194)	(0.00122)			(0.00139)			(0.00167)		
$Editor^c_{i,post}$				0.000689			0.00121			-0.000520	
				(0.00143)			(0.00154)			(0.00175)	
$Editor^{c,fitted}_{i,post}$					0.00496			-0.0140*			0.0189*
					(0.00604)			(0.00686)			(0.00827)
Topics FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4485	4485	4485	4485	4485	4485	4485	4485	4485	4485	4485
R^2	0.480	0.477	0.981	0.981	0.981	0.978	0.978	0.978	0.554	0.553	0.554

Standard errors in parentheses

$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.3: Journals' Topics and Editor's Preference with Three Year Window and Two Year Lag

	$Editor_i^c$		$AER_{i,post}^c$			$Top5_{i,post}^c$			$(AER - Top5)_{i,post}^c$		
	(1)pre	(2)post	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$AER_{i,pre}^c$	0.682***	0.238	0.267***	0.266***	0.268***	0.256***	0.256***	0.255***	0.0113	0.00992	0.0130
	(0.186)	(0.172)	(0.0479)	(0.0480)	(0.0476)	(0.0534)	(0.0534)	(0.0534)	(0.0769)	(0.0771)	(0.0767)
$Top5_{i,pre}^c$	0.0352	0.362+	0.194***	0.194***	0.196***	0.206**	0.206**	0.205**	-0.0116	-0.0121	-0.00913
	(0.166)	(0.185)	(0.0567)	(0.0567)	(0.0568)	(0.0651)	(0.0652)	(0.0650)	(0.0878)	(0.0878)	(0.0878)
$Editor_{i,pre}^c$		0.199***	-0.000820			0.000563			-0.00138		
		(0.0177)	(0.00135)			(0.00140)			(0.00187)		
$Editor_{i,post}^c$				0.00163			0.000481			0.00115	
				(0.00118)			(0.00138)			(0.00160)	
$Editor_{i,post}^{c,fitted}$					-0.00413			0.00283			-0.00696
					(0.00681)			(0.00703)			(0.00939)
Topics FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4485	4485	4485	4485	4485	4485	4485	4485	4485	4485	4485
R^2	0.486	0.452	0.978	0.978	0.978	0.975	0.975	0.975	0.551	0.551	0.551

Standard errors in parentheses

$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.4: Journals' Topics and Editor's Preference with Five Year Window and One Year Lag

	$Editor^c_i$		$AER^c_{i,post}$			$Top5^c_{i,post}$			$(AER$ - $Top5)^c_{i,post}$		
	(1)pre	(2)post	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$AER^c_{i,pre}$	0.628*	-0.127	0.144**	0.144**	0.144**	0.0635	0.0638	0.0639	0.0803	0.0801	0.0801
	(0.246)	(0.239)	(0.0533)	(0.0534)	(0.0534)	(0.0801)	(0.0799)	(0.0799)	(0.0720)	(0.0717)	(0.0718)
$Top5^c_{i,pre}$	0.385+	0.452*	0.303***	0.301***	0.302***	0.380***	0.379***	0.379***	-0.0779	-0.0778	-0.0771
	(0.215)	(0.216)	(0.0623)	(0.0622)	(0.0628)	(0.0615)	(0.0613)	(0.0619)	(0.0761)	(0.0764)	(0.0752)
$Editor^c_{i,pre}$		0.304***	0.000321			0.000799			-0.000478		
		(0.0247)	(0.00199)			(0.00293)			(0.00271)		
$Editor^c_{i,post}$				0.00307			0.00352			-0.000453	
				(0.00188)			(0.00273)			(0.00280)	
$Editor^{c,fitted}_{i,post}$					0.00106			0.00263			-0.00157
					(0.00655)			(0.00962)			(0.00891)
Topics FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2535	2535	2535	2535	2535	2535	2535	2535	2535	2535	2535
R^2	0.598	0.598	0.986	0.986	0.986	0.980	0.980	0.980	0.700	0.700	0.700

Standard errors in parentheses

$^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.5: Journals' Topics and Editor's Preference with Five Year Window and Two Year Lag

	$Editor^c_i$		$AER^c_{i,post}$			$Top5^c_{i,post}$			$(AER$ - $Top5)^c_{i,post}$		
	(1)pre	(2)post	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$AER^c_{i,pre}$	0.226	0.156	0.259***	0.259***	0.257***	0.255*	0.253*	0.252*	0.00430	0.00614	0.00490
	(0.196)	(0.217)	(0.0694)	(0.0696)	(0.0695)	(0.0993)	(0.0997)	(0.0989)	(0.126)	(0.127)	(0.126)
$Top5^c_{i,pre}$	0.412^+	0.0612	0.0816	0.0828	0.0811	0.241**	0.241**	0.240**	-0.159	-0.158	-0.159
	(0.239)	(0.283)	(0.0673)	(0.0670)	(0.0673)	(0.0880)	(0.0870)	(0.0882)	(0.106)	(0.105)	(0.106)
$Editor^c_{i,pre}$		0.296***	0.00288			0.00402			-0.00114		
		(0.0307)	(0.00216)			(0.00310)			(0.00342)		
$Editor^c_{i,post}$				0.000380			0.00979**			-0.00941*	
				(0.00249)			(0.00302)			(0.00363)	
$Editor^{c,fitted}_{i,post}$					0.00972			0.0136			-0.00383
					(0.00730)			(0.0104)			(0.0115)
Topics FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2535	2535	2535	2535	2535	2535	2535	2535	2535	2535	2535
R^2	0.594	0.574	0.980	0.980	0.980	0.973	0.973	0.973	0.577	0.579	0.577

Standard errors in parentheses

$^+ p < 0.10$, $^* p < 0.05$, $^{**} p < 0.01$, $^{***} p < 0.001$