Department of Economics

WORKING PAPER

Lennart Ziegler

What is the Media Impact of Research in Economics?

Working Paper No. 2103

July 2021

What is the Media Impact of Research in Economics?*

Lennart Ziegler**

University of Vienna

July 2021

Abstract

Many research institutions aim to have a strong public impact but little evidence exists on the extent to which research findings reach a wider audience. Using a large sample of studies released in the working paper series of the National Bureau of Economic Research, I identify online coverage of research findings in 6 major news outlets. The analysis shows significant coverage rates in most newspapers in the first month after study release. Overall, about every 11th working paper is covered at least once during this period. I also find that media reporting is correlated with several author and study characteristics. While differences in coverage between most research areas are modest, empirical as well as US-focused studies receive substantially more attention. In particular, widely cited papers are covered more frequently, showing that academic success of studies serves as a strong predictor for wider public impact.

JEL Codes: A11, A14, L82

Keywords: Economics research, media coverage, public impact

*I am indebted to Dan Hamermesh who provided valuable comments and suggestions on an earlier draft of this paper.

**Address: University of Vienna - Department of Economics, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria. E-mail: lennart.ziegler@univie.ac.at.

1 Introduction

Researchers often seek to communicate their research findings to a broader audience beyond academic circles and contribute to the public debate. This is especially relevant at publicly funded institutions, which should generate knowledge that benefits society as a whole. In disciplines likes economics, research output can also inform policy makers regarding the implementation of public policies and it can help voters to make more educated decisions. Yet, little is known about the extent to which research findings reach a wider audience. While citations and publication records help to determine the academic success of researchers, there exist no comparable metrics to evaluate public outreach.

One important channel to disseminate knowledge is media coverage of research output. A discussion of new research findings in major news outlets allows to reach a much broader audience than academic publications. Furthermore, research journalists act as an important intermediary between academics and society. They process complex analyses into non-technical articles, which makes new findings also understandable for people without an academic background.

Media exposure and academic success do not need to be strongly correlated. While journalist may rely on journal rankings and citations to judge the quality of a study (if already available), interests and preferences may differ between academics and the broader public. Not all papers with a high citation count will necessarily have a similar public impact.

Despite the importance of media channels in spreading knowledge, little is known about the extent to which economics research is discussed in the media and how exposure differs between studies. This paper aims to fill this gap by examining the coverage of newly released working papers in 6 major US news outlets. Using a large sample of studies released in the working paper series of the National Bureau of Economic Research (NBER) between 2010 and 2019, I estimate the number of media citations and analyze how they differ by characteristics of studies and authors.

A key advantage of analyzing the coverage of working papers is that the immediate impact can easily be measured. Before studies are eventually published in academic journals, they are often circulated for the first time as working papers. Another important difference to journal articles is that recently published working papers are neither cited nor published yet and thus cannot affect the journalist's decision to cover the paper.

To obtain a measure of media exposure in the first weeks after the release of working papers, I match author names to the online content of 6 widely read international news outlets (*The New York Times*, *The Washington Post*, *The Wall Street Journal*, *The Financial Times*, *The Economist* and *CNN*). Estimation results show that the media outlets frequently cover research in economics. Overall, an average working paper is matched to about 0.16 news articles in the first 4 weeks after release. Every 11th study is covered in at least one of the news outlets during this period.

Differences in coverage rates by JEL classification of research area are, in most cases, relatively similar. Matching keywords in abstract texts, I find that empirical and US-focused studies receive particular media attention. Even though associations with the journal impact factor of eventually published papers are small, studies that are widely cited by the academic community get more widespread media coverage. I interpret this as evidence that academic success of a study serves as a strong predictor for wider public impact.

Next to these study differences, media attention also varies with author characteristics. Research strength of authors, as proxied by article views, is associated with higher coverage rates. Yet, this effect is smaller and insignificant when I exploit within-study differences of coverage between authors. Less experienced researchers and those affiliated to the top 5 percent institutions according to citations tend to get more coverage as well. These differences remain significant when I account for study fixed effects. For a given working paper, media articles are more likely to mention authors from research-strong institutions and those who have less experience in academia.

To my knowledge, this is the first study that quantifies the media exposure of economics research and examines differences by study and author characteristics. Most closely related to my analysis is an article by Hamermesh (2004), who provides media guidelines for economists and includes anecdotal evidence on what sort of studies are typically covered by media outlets. The paper highlights that interests of the public do not always align with academic interests. Studies that reveal important facts to a broader audience might sometimes be too specific to lay ground for further research. Furthermore, it is not clear how media attention affects the chances to be published in a top-ranked journal.

The methodological approach of this study bears similarity to a number of papers that focus on the determinants of academic success in the economics profession. Examining publication trends over 6 decades, Hamermesh (2013) shows that journals increasingly publish empirical studies with self-collected or experimental data. Card and DellaVigna (2020) analyze submissions to leading journals in economics, showing that referee and editor choices are strong predictors of study citations. A related study by Card et al. (2020) provides further evidence on gender discrimination in the publication process. The authors find that, conditional on study citations, female authors have lower chances to get a paper published in these journals.[1] From a broader perspective, my analysis also relates to the large literature in economics which examines how media channels influence politics and public policies.[2]

The rest of the paper is organized as follows. Section 2 describes the different data sources and provides descriptive statistics on working papers, authors and media coverage. The empir-

[1] Bransch and Kvasnicka (2017) find no evidence that these gender differences are driven by publishing boards consisting mainly of male editors.

[2] See Prat and Strömberg (2013) for a comprehensive review of recent studies in political economy.

ical framework is outlined in Section 3. Section 4 presents estimates for overall media coverage and shows how effects differ by study and author characteristics. Section 5 concludes.

2 Data

The estimation sample combines various data on media coverage, research papers and their authors. It is composed from different sources of publicly accessible data. Table A.1 in the appendix provides an overview with references to the respective sources and extraction dates.

2.1 Research output

The analysis focuses on all working papers which have been released by the National Bureau of Economic Research (NBER) between 2010 and 2019. A key focus of this non-profit organization is to disseminate research among academics, policy makers and professionals. The NBER working paper series is a prestigious and well-known source of recent research output. About 20 papers are released every week and assigned to different research programs, which cover all major fields in economics. Although access to NBER content is restricted, academic institutions as well newspapers and magazines can view and download the working papers. As such, the NBER publications are also frequently consulted by journalists.

To publish an NBER working paper, at least one author has to be affiliated to the organization as faculty research fellow or research associate. Despite this restriction, many coauthors are affiliated to non-US institutions and several studies focus on other countries than the US. For this reason, the working paper series can be seen as a source of international research, with impact also beyond the US.

While the analysis could be replicated with other working paper series, focusing on NBER working papers has several advantages. First, the organization is one of the largest and best-known platforms for research in economics. With around 1,000 releases per year, the large sample size simplifies the empirical analysis. Also, many journalists know the organization and are thus aware of recent research output. Second, the NBER programs cover all major areas of economics, which allows to examine differences in media coverage by subfield. Third, the exact date of release is available. As explained in the next section, this greatly simplifies matching papers to media reports and allows to estimate the immediate impact in the first weeks after release.

2.2 Study and author characteristics

To study heterogeneous effects, I merge additional data from different online sources. The NBER website provides for every working paper the date of release, JEL classification codes, a publication record (if any reported) and the abstract text. The widely-used JEL codes, developed by the Journal of Economic Literature, provide a detailed classification scheme for research articles in economics. Because of sample size restrictions, I focus on the 20 general categories.[3] While the JEL classification specifies the field of research in economics, the codes do not distinguish many relevant key characteristics of studies. To obtain additional information, I use word matches in the abstract text to proxy whether a study is empirical, theoretical, policy related and US focused.[4]

The NBER website also lists the researchers' name and affiliation. From the profile pages of authors, I obtain the year in which they released their first NBER publication. The years between release of a working paper and first NBER publication will serve as proxy for research experience in the subsequent analysis.

To determine the gender of authors, I follow a procedure similar to the one proposed by Card et al. (2020) and predict the gender based on first names. Using the R-package *gender,* I obtain the proportion of females and males with a given name from US social security records and from the Genderize.io database, which is compiled from user profiles on social networks. Because random measurement error in the explaining variable will downward bias regression coefficients, it is necessary to avoid mix-ups as much as possible. Therefore, I only classify the gender if the proportion of females or males is at least 90 percent. Overall, 92 percent of all names in the estimation sample can be identified as male or female. As a robustness check of the described procedure, I search for author matches in the RePEc ranking of the top 10 percent female economists and on a list of female members of the European Economic Association (EEA). Among the matched authors, only 2 percent of female economists are wrongly classified as males.

Additional information on research output and authors comes from the RePEc (Research Papers in Economics) platform. The RePEc service EconPapers provides a large collection of working papers and journal articles in economics. For each working paper, I obtain the total number of citations.[5] In addition, I count the number of additional working and discussion paper series in which the studies were released. Multiple releases can increase the visibility of a paper and potentially lead to more media coverage.

To evaluate the research strength of journals, institutions and authors, I collect data from the RePEc services IDEAS and CitEc, which provide various citation-based rankings. The CitEc journal ranking lists more than 600 journals along with their impact factor, which can be used as a measure of publication success. The IDEAS institution ranking contains the top 5 percent of institutions according to overall citations. Furthermore, I extract the number of article views for all authors who are registered on RePEc.

[3] Note that these groups are non-exclusive. In fact, most papers are classified as belonging to multiple categories.

[4] See Figure 4 for the definition of word pattern matches.

[5] If a working paper has already been published or if other working papers exist, I consider the total number of citations across all publications.

Table 1: Descriptive statistics - Working papers

	Mean	Std. Dev.	Min	Max	Obs
# authors	2.60	1.09	1	13	10,902
Single author	0.14	0.34	0	1	10,902
Publication					
Citations	25.93	56.36	0	1,493	10,902
Published (yet)	0.63	0.48	0	1	10,902
Top five journal	0.18	0.39	0	1	6,903
Journal impact factor	4.18	3.24	0.02	12.59	5,522
# other WP versions	0.80	1.23	0	11	10,902
JEL codes					
(J) Labor and Demographic Economics	0.24	0.43	0	1	10,902
(E) Macro-/Monetary Economics	0.23	0.42	0	1	10,902
(I) Health, Education, and Welfare	0.22	0.42	0	1	10,902
(D) Microeconomics	0.22	0.42	0	1	10,902
(G) Financial Economics	0.21	0.40	0	1	10,902
(H) Public Economics	0.17	0.38	0	1	10,902
(O) Econ. Develop., Techn. Change, and Growth	0.17	0.37	0	1	10,902
(F) International Economics	0.15	0.36	0	1	10,902
(C) Mathematical and Quantitative Methods	0.10	0.30	0	1	10,902
(L) Industrial Organization	0.11	0.31	0	1	10,902
Other (<10%)	0.28	0.45	0	1	10,902

NOTE: The sample includes NBER working papers released between 2010 and 2019. 'Top five journal' indicates publication in AER, Econometrica, JPE, QJE or REStud. Data on working paper versions, citations and journal impact factors were obtained in May/June 2021.

Table 1 and Table 2 provide summary statistics for all study and author characteristics. The estimation sample covers nearly 11,000 working papers. A study has, on average, 2-3 authors, and merely 14 percent of papers are single-authored. As of mid 2021, two third of all working papers have a publication record. This is, in part, driven by the fact that some papers have only been released recently. The publication rate ranges from 83 percent for papers from 2010 to 27 percent for papers released in 2019. Among articles with a publication record, about a fifth are published in one of the so-called top five journals, which are often regarded as the journals with the strongest academic impact. For 80 percent of the publications, I observe the journal impact factor in the CitEc ranking. Most of the remaining publications are studies that have been published in books or special editions for which impact factors are not available. The data from EconPapers reveal that several studies have also been released in other working paper series. On average, I observe 0.8 additional releases per study. The second half of Table 1 reports the (non-mutually exclusive) JEL general categories of working papers in the sample. The distribution shows that all major fields of economics are represented with substantial shares. The 10 largest groups each account for at least 10 percent of the studies. Because the remaining JEL codes are less frequently listed, I group them together for the empirical analysis.

As shown in Table 2, the estimation sample contains studies written by about 9,600 different authors. The majority publishes only one working paper during the sample period but there also exist some authors who publish very frequently in this series, yielding an average of almost 3 papers per author. Among the 92 percent of names that can be matched to a gender, a quarter of authors are female.[6] 56 percent of authors have a matching RePEc profile, which reports overall article views. Moreover, 59 percent of authors are affiliated to an institution that is among the top five percent in the IDEAS citation ranking.

Table 2: Descriptive statistics - Authors

	Mean	Std. Dev.	Min	Max	Obs
# WP (sample period)	2.94	4.32	1	75	9,634
Year first NBER WP/Pub	2,011.04	8.07	1,973	2,019	9,634
Female	0.26	0.44	0	1	8,840
Gender unknown	0.08	0.28	0	1	9,634
Author RePEc profile	0.56	0.50	0	1	9,634
Author RePEc article views	25,270.61	47,337.82	0	674,785	5,396
Insitution in top 5% cited	0.59	0.49	0	1	9,634
Institution citation score	112,548.26	119,804.41	6,645	470,596	5,729

Note: The sample includes authors of NBER working papers released between 2010 and 2019. Data on article views and citation scores were obtained in May/June 2021.

2.3 Media coverage

To measure media coverage, I collect data from the websites of four daily newspapers (The New York Times, The Washington Post, The Wall Street Journal, The Financial Times), one weekly magazine (The Economist) and one television channel (CNN).[7] I focus on these 6 sources because they all have substantial reach even beyond the US and report extensively on the economy and related topics. Being discussed in one of the outlets guarantees research papers to have a widespread audience. Even though this only represents a small share of potential media coverage, these sources thus serve as a reasonable proxy for media exposure.[8] The observed media content (reports, features, etc.) often overlaps with print content but

[6]Card et al. (2020) compute a share of about 18 percent female authors of publications in 53 different journals during a similar time period.

[7]The NBER website also lists news articles that cover the working papers. While these self-reported entries include a wider range of news outlets, they only cover a selective subset of media coverage. The number of entries substantially increases during the period of observation, which must be due to selective reporting. Moreover, it is possible that the selection of reported coverage is correlated with author and paper characteristics. Also, the entries do not readily allow to identify which authors of a study are mentioned in the media.

[8]The sample of newspapers can be further extended. Yet, not all websites qualify for web scraping in this context. The respective website has to have a reliable search engine which allows for whole word search and an accessible archive of published articles in the past years along with publication dates. These criteria exclude a number of potential news sources (e.g. Los Angeles Times, Chicago Tribune, FoxNews, CBS News, USA Today or Time Magazine).

also includes additional online-only content such as affiliated blog posts. To match media reports to papers, I use the websites' search engines to search for the names of all authors in the sample. Articles or blog entries that discuss a research paper typically mention the corresponding authors, which can be used to identify studies.[9] This simple procedure also allows to automatize the search process. Due to the large number of working papers and authors, a manual search for media coverage would hardly be feasible. A caveat of this approach is that I cannot distinguish the extent or type of coverage. While some articles provide an in-depth discussion of research findings, others may only make a short reference to working papers.

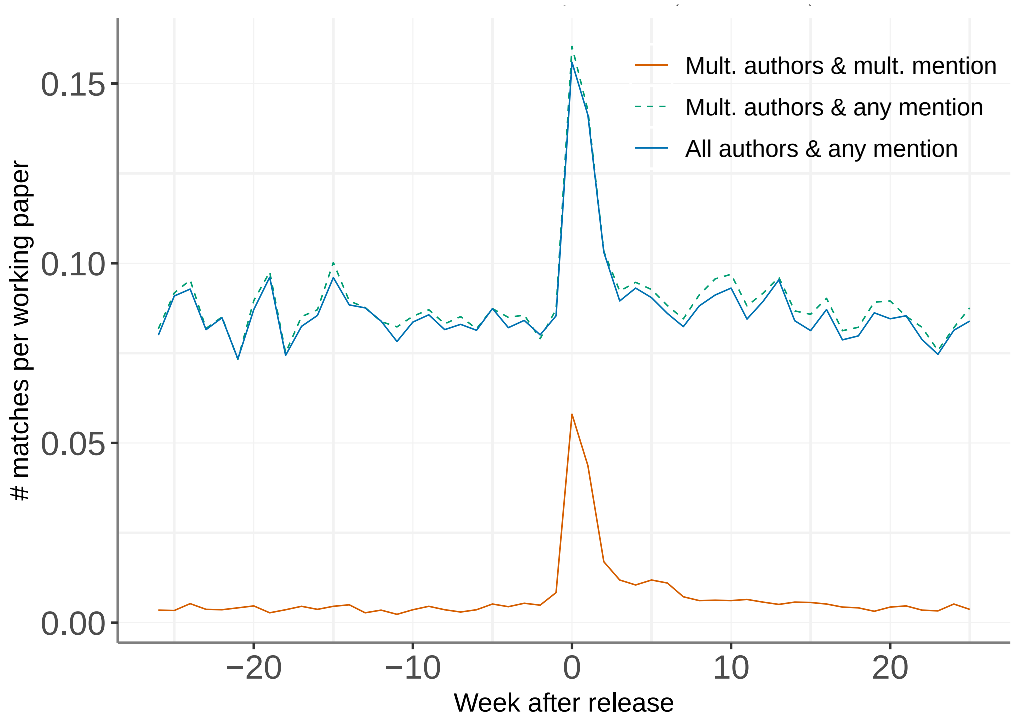
The main piece of information contained in the search results are the publishing dates of news articles. Not all observed name matches relate to news coverage of the respective studies. To identify coverage of working papers, I exploit the exact timing of release and compare matches in the first weeks after release to matches in the weeks before. Figure 1 plots the number of weekly matches per working paper by study release date, starting 6 months before and up to 6 months after release. The upper solid line shows the number of matches for the full sample. A news article is counted as match if it refers to at least one of the authors. In all weeks, I observe a large number of matches to working papers in my sample. Following the release of the paper, media hits increase sharply by more than 80 percent. After a few weeks, matches drop again to numbers similar to pre-release levels. While the large spike can arguably be attributed to coverage of the respective working papers, most of the matches, especially before the release dates, should be unrelated news articles.

When I only count matches if at least two authors of a study are referenced in a news article, the number of potential mismatches can be substantially reduced. By construction, this excludes single-authored working papers, which amount to 14 percent of the sample. The lower graph of Figure 1 shows that the corresponding number of matches is close to zero in most of the weeks. Almost all coverage happens in the first weeks after release. Compared to the unrestricted sample, the peak is somewhat less pronounced. This suggests that not all news articles provide full references when reporting on recently released working papers. To capture the full extent of media coverage, the subsequent analysis will focus on the full sample. Corresponding results for the restricted sample, which tend to be estimated with smaller standard errors due to a lower number of mismatches, are reported for comparison in the appendix.

For matches that are unrelated to coverage of a given working paper, a higher number of authors should increase the chances to observe a match. The dashed line plot of Figure 1 shows matches to multi-authored studies if news articles are also counted when only one author is

[9]For a small number of authors, I observe many mismatches because they share their name with other persons who are frequently mentioned in news sources. To reduce measurement error, I drop all working papers written by these authors, which corresponds to less than 1 percent of the sample.

mentioned. Compared to the full sample of papers, there are indeed somewhat more media matches. To account for this effect, I will control for the number of authors in the subsequent analysis.



3 Empirical framework

3.1 Measurement error

Using automated search of author names to identify news coverage poses several empirical challenges. As shown above, the procedure may wrongly attribute some news articles to research papers. In some cases, the mentioning of an author even coincides with the release of a working paper although the media coverage is unrelated to the respective study. This especially concerns economists who regularly write op-ed's or blogs. It is also possible that other unrelated persons with the same name are referenced instead. Moreover, media sources may report about a study without mentioning the full or correct name of authors. Both types of measurement error can bias the estimated treatment effect. In the following, I assume that, for study i in period t, the number of matches y^*_{it} is measured with error ϵ_{it}:

$$y_{it} = y^*_{it} + \epsilon_{it}$$

To identify coverage of working papers, I exploit the timing of release and compare matches in the first weeks after release to matches in the weeks before. The actual treatment effect on the *number of matches* is

$$\beta = E[y^*_{it}|D_{it} = 1] - E[y^*_{it}|D_{it} = 0],$$

where D_{it} indicates post-release periods. The estimated treatment effect is

$$E[\hat{\beta}] = E[y_{it}|D_{it} = 1] - E[y_{it}|D_{it} = 0] = \beta + E[\epsilon_{it}|D_{it} = 1] - E[\epsilon_{it}|D_{it} = 0].$$

The identifying assumption thus requires that the average error is the same before and after the release of a working paper:

$$E[\epsilon_{it}|D_{it} = 1] = E[\epsilon_{it}|D_{it} = 0]$$

This condition is likely satisfied when the mismeasurement can be attributed to over detection. Unless the author publishes multiple studies at the same time or suddenly becomes more present in the media, the extent of measurement error will not be affected. Also matches to other persons with the same name should not be correlated with the release date. However, if a media outlet covers a new working paper without mentioning the authors, this only leads to measurement error in post-release periods. As a result, the estimated treatment effect will be downward biased. Because it is common journalistic standard to reference at least one author when research studies are discussed in the media, I expect the resulting bias to be small.

Next to the number of matches, it is also informative to measure whether a study is covered at all in the media. Identifying this effect requires additional assumptions. The actual treatment effect on *having any match* is given by

$$\gamma = P[y^*_{it} > 0|D_{it} = 1] - P[y^*_{it} > 0|D_{it} = 0],$$

while the observed difference in the share of non-zero matches before and after release is

$$E[\hat{\gamma}] = P[y^*_{it} + \epsilon_{it} > 0|D_{it} = 1] - P[y^*_{it} + \epsilon_{it} > 0|D_{it} = 0].$$

Assuming that $\epsilon_{it} \geq 0$ and y^*_{it} and ϵ_{it} are independent, this can be expressed as

$$E[\hat{\gamma}] = \gamma \times (1 - P[\epsilon_{it} > 0|D_i = 1]).$$

To satisfy independence between y^*_{it} and ϵ_{it}, errors caused by mix-ups with other persons of the same name should not be correlated with coverage of the working paper. Furthermore, authors of covered studies should not be more present in the media either. Under these conditions, the measurement error biases the actual impact towards zero as shown above.[10] Assuming

[10]The derivation is provided in the appendix.

that an error is equally likely before and after study release, the matching probability in the pre-release period can be used to correct the downward bias.

3.2 Event study

To estimate media exposure of research papers by week after release, I conduct an event study. The estimation equation is given by

$$y_{it} = \alpha_t + \sum_{w=-12}^{12} \beta_w D_{it}^w + u_{it}$$

where y_{it} denotes the number of matches for study i in week t. Variables D_{it}^w indicate the week relative to the date of release, where the omitted indicator is the week just before release. To account for potential timing effects, I also include release date indicators (α_t). This regression allows to quantify news coverage by week before and after the release of a research paper. Coefficients β_w capture the increase in media matches relative to the base period. Under the conditions specified above, this should remove the measurement error. Furthermore, the event study design allows to reveal potential trends in the pre-release period.

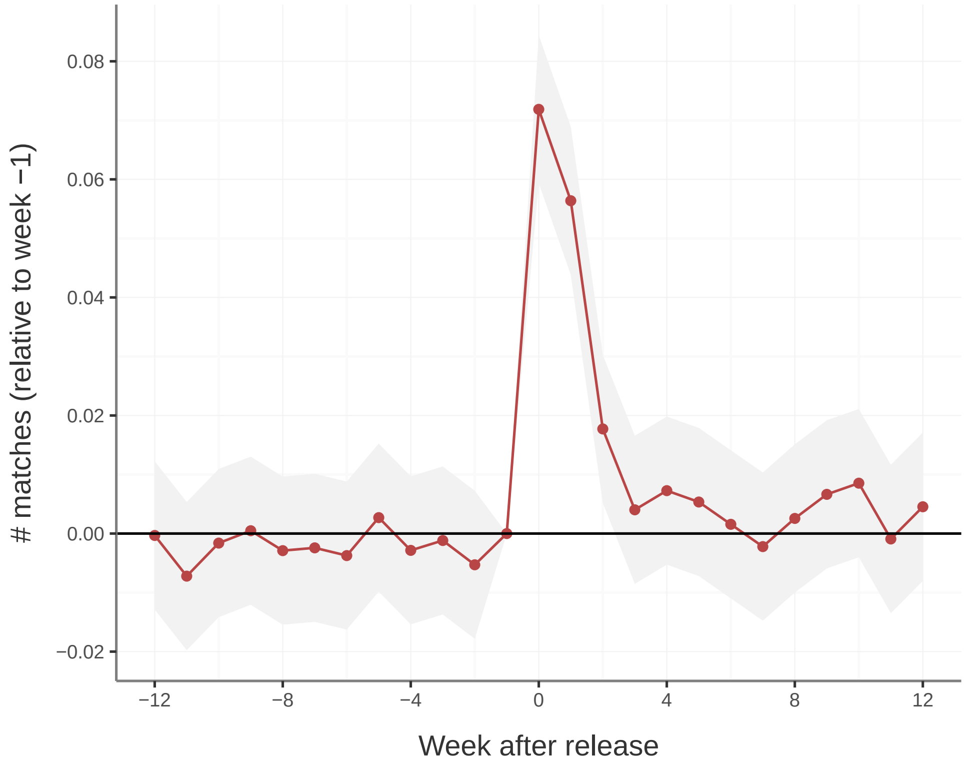
4 Results

4.1 Overall coverage

Figure 2 plots estimates of the event study, along with 95 percent confidence intervals, for the full sample of working papers. The graph shows a large spike in coverage immediately after release. In the first week, I estimate about 0.07 matches per study. Afterwards, coverage rates drop again but remain significant for another two weeks. In the subsequent weeks, coverage further decreases to estimates slightly above zero. Compared to the base period, coefficients are also close to zero in the weeks before the release. This suggests that researchers do not disseminate their research findings prior to the NBER release, which allows a clear identification of coverage around the release date.

To simplify the analysis in the remainder of this paper, I measure media coverage in the first 4 weeks after the release of working papers (treatment period) and compare it to matches in the preceding 4 weeks (control period). Table 3 reports the corresponding estimates for total coverage and coverage by media source. As shown in the upper panel, I observe nearly 0.16 media hits per paper. A lot of coverage stems from The Wall Street Journal, which accounts for half of the matches. I also find significant estimates for most of the remaining media outlets, ranging from about 0.003 to 0.024 matches per paper.

Figure 2: Event study graph



Note: The grey area indicates 95% confidence intervals.

The lower panel of Table 3 reports estimates for the likelihood to observe any match. As shown in the previous section, calculating the difference between treatment period and control period in the share of papers that have any match yields a downward biased estimate for the coverage likelihood. To correct for measurement error, I divide the difference by the share of papers that are not matched in the control period ($\frac{Share(y>0)_{D=1}-Share(y>0)_{D=0}}{1-Share(y>0)_{D=0}}$). The results show that approximately 9 percent of studies were covered in at least one of the 6 media sources. Again, I find the highest probability of coverage for The Wall Street Journal.

To examine changes in coverage over time, I compute average differences in media matches between treatment period and control period by month of release. Figure A.1 in the appendix shows that there is no clear trend in coverage over the sample period. I do not find evidence for seasonality in news coverage either. The estimates do not differ significantly by months, weeks or weekdays.

Table 3: Coverage estimates by sample

	All	NYT	WP	Economist	WSJ	FT	CNN
Outcome: Number of matches							
Coefficient	0.158	0.006	0.022	0.024	0.085	0.018	0.003
Standard error	(0.012)	(0.005)	(0.005)	(0.003)	(0.006)	(0.004)	(0.001)
Outcome: Any match							
Coefficient	0.087	0.008	0.020	0.019	0.058	0.009	0.003
Standard error	(0.005)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.001)

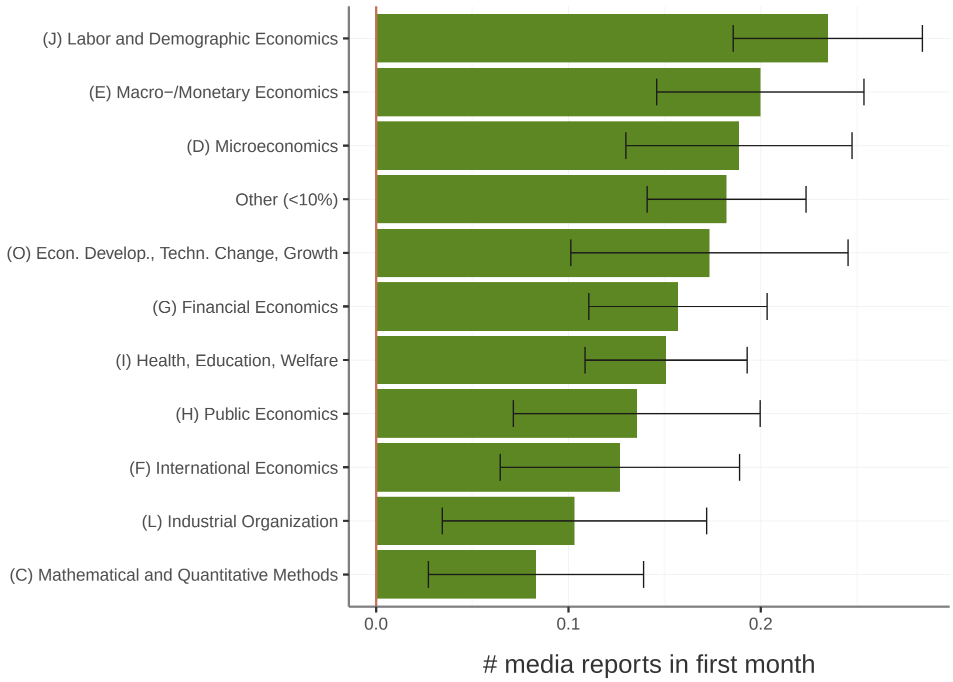
NOTE: The sample includes all papers (10,902). The upper panel reports differences in media matches between the first 4 weeks after the NBER release date (release period) and the 4 weeks before (control period). Standard errors reported in parentheses are clustered by paper. The lower panel shows differences in the share of papers with any match between release and control period divided by the share of papers without matches in the control period, where standard errors are approximated using the delta method.

Coverage rates estimated on the restricted sample, where news articles are only counted if they can be matched to at least two authors, are reported in the appendix (see Table A.2). For better comparison, the table also provides estimates for the sample of multi-authored papers, where news articles are also counted when only one author is mentioned. Compared to the results shown above, the estimate for the overall number of related news articles is about 30 percent lower in the restricted sample. Systematic measurement cannot explain this difference because the correlation between release date and measurement error would have to be unrealistically large. It is unlikely that authors are suddenly much more present in the media after the release of one study. The observed difference rather indicates that some news articles do not provide complete references. As a result, I observe a smaller number of matches in the restricted sample. Because of reduced measurement error, the corresponding t-statistics tend to be smaller as well.

4.2 Differences by study characteristics

Not all working papers receive the same degree of media attention. This section analyzes to what extent coverage differs by a number of observable study characteristics such as academic popularity and field of research. As discussed in the previous section, the estimated standard errors in this analysis are relatively large because many media matches in the full sample are unrelated to coverage of the working papers and have to be differenced out. The more restrictive definition, which requires that media articles mention at least two authors, misses out some related matches but allows for a more precise estimation of heterogenous effects. Tables and figures on heterogeneous effects for the full sample are presented in the main text while differences by study characteristics for the restricted sample are provided in the appendix.

Figure 3: Media reports by JEL classification

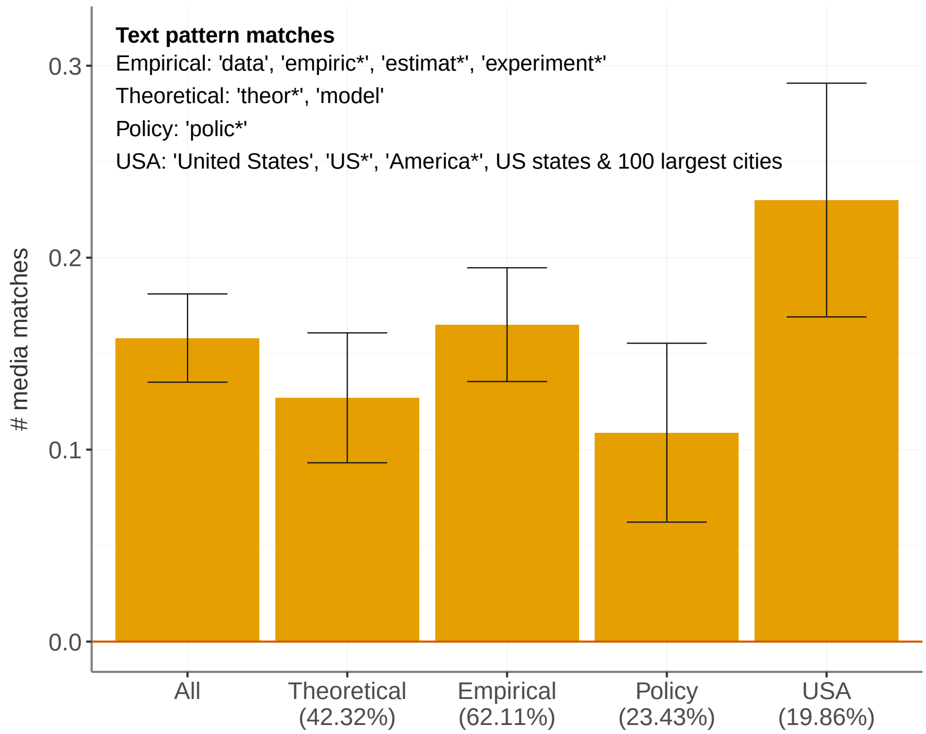


Note: The black lines indicate 95% confidence intervals.

Figure 3 plots coverage rates by general category of the JEL classification, where less frequent categories are again grouped together. For most of the categories, average media matches are close to the overall mean. The field that receives the most media exposure is *Labor and Demographic Economics*. Relative to the average, the category-specific coverage rate is about 50 percent higher. For papers classified under *Mathematical and Quantitive Methods*, the estimate is the smallest but still statistically significant. Technical studies may often be difficult to discuss in media outlets that are targeted at a wider audience without the required background knowledge.

Figure 4 shows differences in coverage by several additional study characteristics, which are proxied by word matches in abstract texts. The share of abstracts that could be matched to the respective keywords is reported in brackets. Compared to theoretical papers, studies with empirical content receive more media coverage. The estimated number of media matches is 30 percent higher. Analyses of data might be easier to break down for a non-academic audience. Readers may also be more interested in learning about empirical findings, which affect them more directly. Moreover, I find that working papers with policy references receive below-average media attention. This suggests that a discussion of more general topics is preferred over the analysis of specific policies. Figure 4 also shows that research with a focus on the United States receives much more media attention. Compared to studies without a US

Figure 4: Media reports by study characteristics



Note: The black lines indicate 95% confidence intervals.

reference, the estimated coverage rate is more than 60 percent higher. Although all 6 media outlets are internationally known, many readers are based in the United States. Journalists might thus be more likely to discuss a study when there exists a connection to the United States.

Next, I examine the relation between media coverage and academic success. Table 4 reports differences in coverage by citations of a study and, if the paper has already been published in a ranked journal, by journal impact factor and journal rank. To account for the fact that recently released papers are less likely to be published yet, all regressions include release date indicators. None of the estimated specifications reveal significant differences in coverage between published and unpublished studies.

As shown in the first column of Table 4, there exists a robust correlation between media reports and academic citations. A doubling of citations is associated with $0.055 \times ln(2) \approx 0.038$ additional media reports per study in the first weeks after release.[11]

[11] To also include papers without any references, I add one unit to every citation count before taking the logarithm. Card et al. (2020) use an inverse-hyperbolic-sine transformation, $asinh(cite) \equiv log(cite + (1 + cite^2)^{0.5})$, to accommodate observations with zero citations. Applying this transformation leads to similar coefficient estimates.

Table 4: Media reports by study characteristics

	(1)	(2)	(3)	(4)	(5)
WP released					
× ln(Citations+1)	0.055***			0.061***	0.060***
	(0.011)			(0.012)	(0.012)
× ln(Journal impact factor)		0.015		-0.012	
		(0.026)		(0.028)	
× ln(Journal rank)			-0.015		0.003
			(0.012)		(0.013)
× **I**(Ranked publication observed)		0.077	0.053	-0.011	0.040
		(0.133)	(0.048)	(0.139)	(0.048)

NOTE: The sample includes all papers (10,902). The coefficients measure media matches in the first 4 weeks after the NBER release date (release period) compared to the 4 weeks before (control period). All regressions control for levels of included characteristics, date of release, number of authors and JEL categories. Journal impact factors and ranks are set to the minimum if no (ranked) publication exists. Standard errors reported in parentheses are clustered by paper. **I**( ) denotes indicator functions. $*$ significant at 10% level, $**$ significant at 5% level, $***$ significant at 1% level.

When using the logarithm of the journal's impact factor or of the journal rank as proxy for academic success, estimated differences are small and statistically insignificant. For the restricted sample, I find very similar coefficients, which are significant at the 10 percent level (Table A.3). For a 100 percent increase in the journal's impact factor, measured media reports only increase by about 0.01. This shows that citations are a much stronger predictor for media coverage than journal strength.

Both measures of academic success, citations and journal rank, are highly correlated ($Corr[ln(\text{Journal rank}), ln(\text{Citations} + 1)] \approx -0.37$). Since citations are also observed after publication, appearing in a high-ranked journal might drive up the citation count. When I include both study citations and journal rank or journal impact factor in the regression, the coefficient on citations remains similar while differences by journal strength get even smaller. Results for the restricted sample show that the significant associations with journal rank and journal impact factor fade to zero once citations are included in the regression.

Given the lack of exogenous variation in coverage, it is unclear what drives the correlation between media exposure and academic success. Being discussed in the media might raise the chances of publication in a high-ranked journal. It is possible that editors prefer to publish papers with media coverage as public attention increases the visibility of the journal. Analogously, some researchers might cite a specific study because they have read a news article about it. It is also likely that academic and media interests align to a certain extent. Journalists and academics might have the same unobserved 'taste' for specific studies that are innovative or comprise novel findings. The estimation results discussed above suggest that the latter channel is more important. When controlling for citations, differences by journal rank fade to be significant. If media coverage increases the publication success of papers conditional on their academic value as proxied by citations, this association should persist.

Dissemination of research in multiple working paper series could substantially raise the visibility of a study in the media. To test this hypothesis, I estimate differences in coverage rates by the number of working paper versions listed on EconPapers. The regression results of Table 5 show that working papers released in multiple series tend to receive somewhat more media exposure but the effect is small and insignificant. One additional release is associated with an increase of about 0.01 media reports. The positive estimate is driven by studies with multiple other releases, which only concerns a small share of working papers in the sample. Less than 20 percent of studies have more than one additional release in other working paper series. The corresponding estimates for the restricted sample (Table A.4) are again more precisely estimated and confirm the findings for the full sample.

Table 5: Media reports by number of working paper versions

	(1)	(2)	(3)	(4)
WP released				
× # other WP versions	0.012	0.008		
	(0.012)	(0.012)		
× 1 other WP version			-0.015	-0.022
			(0.030)	(0.031)
× 2 other WP versions			0.057	0.049
			(0.041)	(0.042)
× 3+ other WP versions			0.024	0.011
			(0.044)	(0.046)
JEL Categories		✓		✓

NOTE: The sample includes all (10,902). The coefficients measure media matches in the first 4 weeks after the NBER release date (release period) compared to the 4 weeks before (control period). All regressions control for levels of included characteristics, date of release and number of authors. Standard errors reported in parentheses are clustered by paper. $*$ significant at 10% level, $**$ significant at 5% level, $***$ significant at 1% level.

4.3 Differences by author characteristics

Next to study characteristics, differences between authors may influence media exposure as well. Some researchers might be more likely to seek media attention than others. Visibility of an author may also be an important criterion for journalists to decide which study is worth covering.

Table 6 shows how media coverage differs by gender, experience, research strength and affiliation of authors. In contrast to the preceding analysis of study-level differences in coverage, media matches are now measured for every author-paper combination. This sample allows to examine whether studies are more likely to be covered if their authors possess specific characteristics. The analysis further shows which authors are mentioned when news articles do not reference all authors of a study. To capture heterogeneity with respect to the mentioning of specific authors, I estimate regressions that also include study indicators.

The results show no significant differences in media coverage between male and female authors. Compared to studies written by male economists, the difference is at most 0.01 and never statistically different from zero. Likewise, the results do not provide evidence that gender is a relevant factor when media reports only cite specific authors of a study. The point estimates remain similar when study fixed effects are included in the regression.

As measure of research strength, I use the authors' total number of article views on Econ-Papers, which yields a significant and robust coefficient. One standard-deviation increase in ln(article views) is associated with approximately 0.03 additional media matches in the regressions without study fixed-effects. Relative to the average number of media reports, this corresponds to an increase of about 20 percent. When study indicators are added to the regressions, the coefficients are only half as large and turn insignificant. This shows that the differential impact by articles views of authors can mostly be attributed to correlated study differences. When reporting on a specific working paper, media outlets are not significantly more likely to mention those researchers who are more visible in academia.[12]

Even if authors are individually less visible, journalists might be more likely to discuss their research if they are affiliated to institutions with a strong research background. The regression estimates show that authors affiliated to the top 5 percent institutions in the citation ranking receive significantly more coverage. This effect cannot be explained by study differences, suggesting that journalists rather cite those authors who are affiliated to a ranked university. Conditional on being in the ranking, the institutions' total citations are not associated with more coverage. The estimated coefficients are small and insignificant in both specifications.

Moreover, I estimate a negative correlation between media coverage and research experience of authors. For every ten years between study release and an author's first NBER publication, coverage decreases by about 0.035 reports per release when I control for author visibility. In the regressions with study fixed effects, the estimate decreases again but remains significant, showing that less experienced authors are more likely to be referenced in news articles about a working paper release.

[12] Authors without a RePEc profile, for whom article views are not available, show higher coverage rates. Yet, it is unclear how this effect should be interpreted as researchers who lack such a profile do not necessarily have less known research output.

As for the previous section, I repeat the heterogeneity analysis for the restricted sample of multi-authored papers, where news articles are only counted if they mention at least two authors. Because there exists little within-paper variation in media coverage of authors in this sample, the estimated specifications do not include study fixed effects. Table A.5 in the appendix reports the corresponding estimates. Compared to the results shown in Table 6, the coefficients are mostly smaller in magnitude but qualitatively similar. Due to fewer mismatches in this sample, the point estimates are again more precisely estimated. In contrast to the coefficients for the full sample, differences by the citation score of ranked institutions are marginally significant.

Table 6: Media reports by author characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
WP released						
× Gender - Female	-0.009	-0.008	-0.005	-0.007	-0.006	-0.008
	(0.012)	(0.013)	(0.012)	(0.013)	(0.012)	(0.013)
× Gender - Unknown	-0.009	-0.017	-0.005	-0.016	-0.005	-0.017
	(0.016)	(0.015)	(0.016)	(0.015)	(0.016)	(0.015)
× Experience (10^{-1}×years)	-0.012*	-0.017***	-0.034***	-0.025***	-0.035***	-0.027***
	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)	(0.008)
× ln(Author article views)			0.021***	0.008	0.019***	0.006
			(0.006)	(0.006)	(0.006)	(0.006)
× **I**(Author article views observed)			-0.188***	-0.071	-0.170***	-0.062
			(0.053)	(0.055)	(0.051)	(0.054)
× ln(School citation score)					0.005	-0.001
					(0.006)	(0.006)
× **I**(School citation score observed)					0.032**	0.037**
					(0.015)	(0.017)
Study FE		✓		✓		✓

NOTE: The sample includes all author-paper combinations of multi-authored papers (28,326). The coefficients measure media matches in the first 4 weeks after the NBER release date (release period) compared to the 4 weeks before (control period). All regressions control for levels of included characteristics, date of release, number of authors and JEL categories. Experience is proxied by years since first NBER publication. Article views and citation scores are set to the respective minimum if author/school data not available. **I**( ) denotes indicator functions. $*$ significant at 10% level, $**$ significant at 5% level, $***$ significant at 1% level.

5 Conclusion

A key challenge for researchers is to make their knowledge accessible to a broader audience outside academia. This paper provides evidence that top media outlets serve as an important intermediary between academia and the general public. Analyzing a large sample of working papers in economics released between 2010 and 2019, I find that new research findings are widely covered in leading newspapers. About 9 percent of studies are mentioned in at least one of 6 major news outlets in the first 4 weeks after release of a working paper.

While I observe media coverage, it is unclear how media outlets become aware of interesting studies. Journalists covering economics and related fields may consult publications of working paper series and decide what is worth reporting on. Sometimes researchers also directly seek contact to journalists to publicize new findings. Given the lack of empirical evidence on media coverage of research, it is difficult to judge which studies are of particular interest for the media. Although my estimates cannot be interpreted as causal differences, the findings show that several study and author characteristics serve as important predictors of media coverage. In particular, I find a strong correlation between media coverage and academic success as measured by citations. This indicates that many studies which are popular among academics are also of interest for the wider public. Moreover, I observe that media attention varies with author characteristics. When reporting on a working paper, journalists are more likely to mention those authors who have less experience and are affiliated to research-strong universities. Researchers with these characteristics might seek more media attention or could more likely be the corresponding author. Also, journalists might have a preference to cover or interview researchers with such a background.

A key question remains whether news coverage itself positively affects academic impact. As most of the media articles appear right after the initial release of a working paper, it is not possible to exploit differential timing of coverage. Future research may look at ways to exploit credible exogenous variation in media exposure.

References

Bransch, F. and Kvasnicka, M. (2017). Male gatekeepers gender bias in the publishing process? IZA Discussion Paper.

Card, D. and DellaVigna, S. (2020). What do editors maximize? Evidence from four economics journals. *Review of Economics and Statistics*, 102(1):195–217.

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.

Hamermesh, D. S. (2004). Maximizing the substance in the soundbite: A media guide for economists. *Journal of Economic Education*, 35(4):370–382.

Hamermesh, D. S. (2013). Six decades of top economics publishing: Who and how? *Journal of Economic Literature*, 51(1):162–72.

Prat, A. and Strömberg, D. (2013). The political economy of mass media. *Advances in economics and econometrics*, 2:135.

Appendix

Measurement error in the impact on having any match

Assumptions:

1. The error is non-negative ($\epsilon_{it} \geq 0$) and unrelated to the extent of coverage (y^*_{it} and ϵ_{it} are independent).

2. There exists no coverage before the release ($P[y^*_i > 0|D_i = 0] = 0$).

3. The error probability does not change with the release ($P[\epsilon_{it} = 0|D_{it} = 0] = P[\epsilon_{it} = 0|D_{it} = 1]$).

Derivation:

$$
\begin{aligned}
E[\hat{\gamma}] &= P[y^*_{it} + \epsilon_{it} > 0|D_{it} = 1] - P[y^*_{it} + \epsilon_{it} > 0|D_{it} = 0] \\
&\overset{(1)}{=} 1 - P[y^*_{it} = 0|D_{it} = 1]P[\epsilon_{it} = 0|D_{it} = 1] - (1 - P[y^*_{it} = 0|D_{it} = 0]P[\epsilon_{it} = 0|D_{it} = 0]) \\
&\overset{(2)}{=} \underbrace{P[y^*_{it} = 0|D_{it} = 0]}_{=1} P[\epsilon_{it} = 0|D_{it} = 0] - P[y^*_{it} = 0|D_{it} = 1]P[\epsilon_{it} = 0|D_{it} = 1] \\
&= P[\epsilon_{it} = 0|D_{it} = 0] - (1 - P[y^*_{it} > 0|D_{it} = 1])P[\epsilon_{it} = 0|D_{it} = 1] \\
&\overset{(3)}{=} \underbrace{P[\epsilon_{it} = 0|D_{it} = 0] - P[\epsilon_{it} = 0|D_{it} = 1]}_{=0} + P[y^*_{it} > 0|D_{it} = 1]P[\epsilon_{it} = 0|D_{it} = 1] \\
&\overset{(3)}{=} P[y^*_{it} > 0|D_{it} = 1]P[\epsilon_{it} = 0|D_{it} = 0] \\
&= \gamma \times P[\epsilon_{it} = 0|D_{it} = 0]
\end{aligned}
$$

Table A.1: Overview data sources

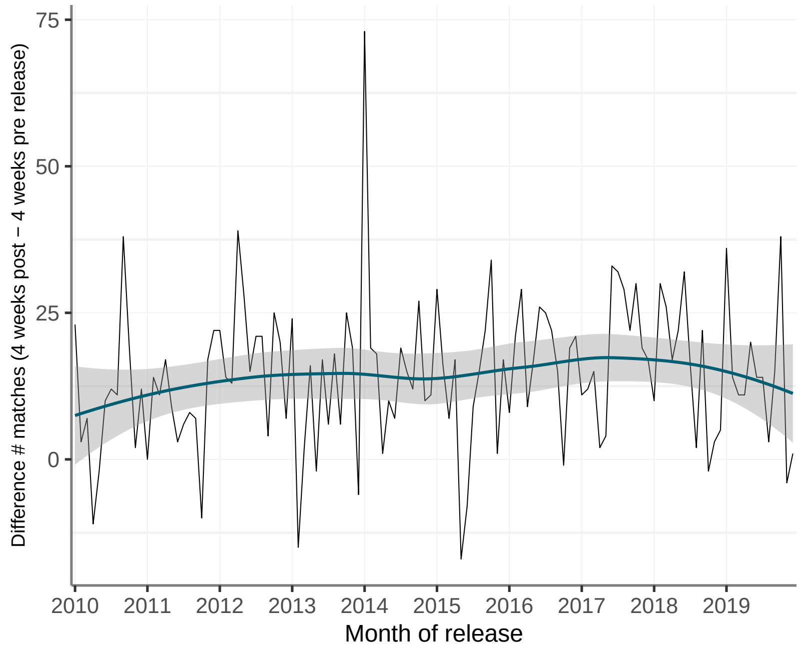
Type	Source	URL	Extraction date
NBER:			
Author profile	NBER website	www.nber.org	May 2021
Paper characteristics	NBER meta data	www2.nber.org/wp_metadata	
Media:			
New York Times	NYT website	www.nytimes.com	May 2021
Washington Post	WP website	www.washingtonpost.com	
Economist	Economist website	www.economist.com	
Wall Street Journal	WSJ website	www.wsj.com	
Financial Times	FT website	www.ft.com	
CNN	CNN website	www.cnn.com	
Additional information:			
Journal ranking 2020	CitEc website	https://citec.repec.org	May/June 2021
Institution ranking	IDEAS website	https://ideas.repec.org	
Author article views	LogEc website	https://logec.repec.org	
Working paper versions	EconPapers website	https://econpapers.repec.org	
Study citations			
R-package 'gender'	CRAN	https://cran.r-project.org	
Top female economists	IDEAS website	https://ideas.repec.org/top/top.women.html	
Female EEA members	EEA website	http://www.eeassoc.org	

Table A.2: Coverage estimates by sample

	All	NYT	WP	Economist	WSJ	FT	CNN
All matches - All papers (N=10,902)							
Coefficient	0.158	0.006	0.022	0.024	0.085	0.018	0.003
Standard error	(0.012)	(0.005)	(0.005)	(0.003)	(0.006)	(0.004)	(0.001)
All matches - Multiauthor papers (N=9,429)							
Coefficient	0.162	0.009	0.024	0.025	0.086	0.015	0.003
Standard error	(0.012)	(0.005)	(0.005)	(0.003)	(0.006)	(0.004)	(0.001)
At least two matches - Multiauthor papers (N=9,429)							
Coefficient	0.107	0.007	0.017	0.014	0.062	0.007	0.000
Standard error	(0.006)	(0.002)	(0.002)	(0.002)	(0.004)	(0.001)	(0.000)

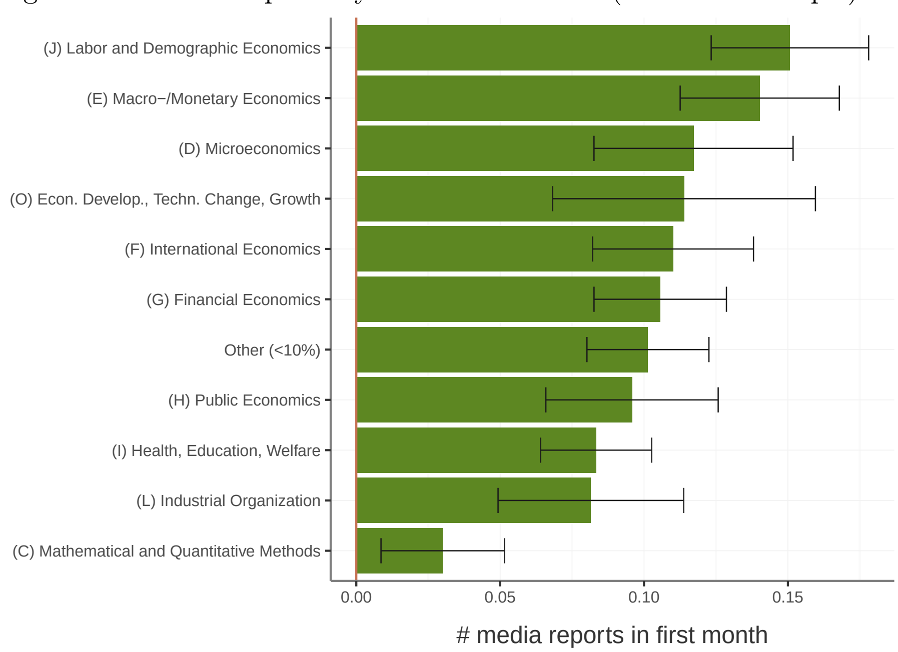
Note: The coefficients measure media matches in the first 4 weeks after the NBER release date (release period) compared to the 4 weeks before (control period). Standard errors reported in parentheses are clustered by paper.

Figure A.1: Name matches by month of working paper release



Note: The bold line shows smoothed values with 95% confidence intervals obtained from local regressions.

Figure A.2: Media reports by JEL classification (Restricted sample)



Note: The black lines indicate 95% confidence intervals.

Figure A.3: Media reports by study characteristics (Restricted sample)

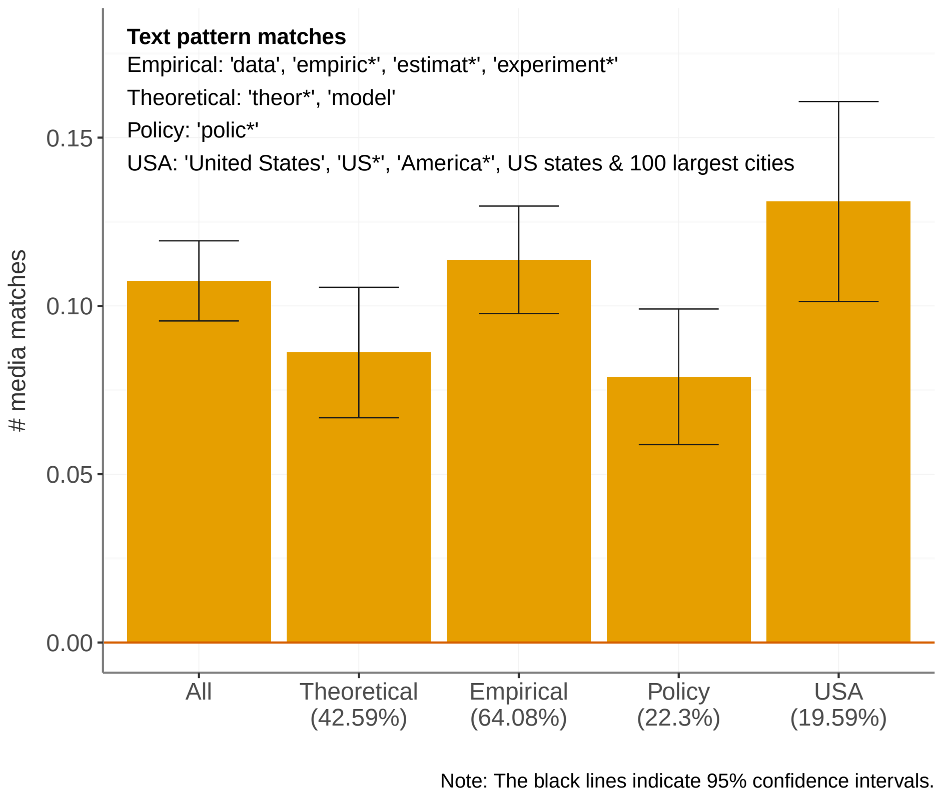


Table A.3: Media reports by study characteristics (Restricted sample)

	(1)	(2)	(3)	(4)	(5)
WP released					
× ln(Citations+1)	0.041***			0.042***	0.042***
	(0.006)			(0.006)	(0.006)
× ln(Journal impact factor)		0.017*		-0.001	
		(0.010)		(0.010)	
× ln(Journal rank)			-0.012*		0.000
			(0.006)		(0.006)
× **I**(Ranked publication observed)		0.067	0.019	0.003	0.009
		(0.048)	(0.020)	(0.048)	(0.020)

NOTE: The sample includes all papers (9,429). The coefficients measure media matches in the first 4 weeks after the NBER release date (release period) compared to the 4 weeks before (control period). All regressions control for levels of included characteristics, date of release, number of authors and JEL categories. Journal impact factors and ranks are set to the minimum if no (ranked) publication exists. Standard errors reported in parentheses are clustered by paper. **I**() denotes indicator functions. $*$ significant at 10% level, $**$ significant at 5% level, $***$ significant at 1% level.

Table A.4: Media reports by number of working paper versions (Restricted sample)

	(1)	(2)	(3)	(4)
WP released				
× # other WP versions	0.011**	0.006		
	(0.005)	(0.005)		
× 1 other WP version			0.027	0.019
			(0.018)	(0.018)
× 2 other WP versions			0.045**	0.034*
			(0.019)	(0.019)
× 3+ other WP versions			0.046**	0.031
			(0.021)	(0.021)
JEL Categories		✓		✓

NOTE: The sample includes all multi-authored papers (9,429). The coefficients measure media matches in the first 4 weeks after the NBER release date (release period) compared to the 4 weeks before (control period). All regressions control for levels of included characteristics, date of release and number of authors. Standard errors reported in parentheses are clustered by paper. $*$ significant at 10% level, $**$ significant at 5% level, $***$ significant at 1% level.

Table A.5: Author-level media reports by author characteristics (Restricted sample)

	(1)	(2)	(3)	(4)	(5)
WP released					
× Gender - Female	-0.007	-0.009	-0.005	-0.009	-0.006
	(0.007)	(0.008)	(0.008)	(0.008)	(0.008)
× Gender - Unknown	-0.001	-0.004	0.000	-0.001	0.001
	(0.012)	(0.012)	(0.011)	(0.012)	(0.011)
× Experience (10^{-1}×years)		-0.006	-0.023***	-0.011***	-0.025***
		(0.004)	(0.005)	(0.004)	(0.005)
× ln(Author article views+1)			0.016***		0.014***
			(0.004)		(0.004)
× **I**(Author article views observed)			-0.140***		-0.122***
			(0.039)		(0.037)
× ln(School citation score)				0.009**	0.008*
				(0.005)	(0.004)
× **I**(School citation score observed)				0.018*	0.017
				(0.010)	(0.010)

NOTE: The sample includes all author-paper combinations of multi-authored papers (26,853). The coefficients measure media matches in the first 4 weeks after the NBER release date (release period) compared to the 4 weeks before (control period). All regressions control for levels of included characteristics, date of release, number of authors and JEL categories. Experience is proxied by years since first NBER publication. Article views and citation scores are set to the respective minimum if author/school data not available. **I**( ) denotes indicator functions. $*$ significant at 10% level, $**$ significant at 5% level, $***$ significant at 1% level.