Paper to be presented at the DRUID24 Conference
SKEMA, Université Côte d'Azur
June 13-15, 2024

Exploring the Role of Social Media in the Diffusion of Economic Research

Yotam Sofer Department of Strategy and Innovation ys.si@cbs.dk

Abstract

For more than a decade, social media has become a key channel for knowledge dissemination used by scientists as a whole, and economists in particular. However, its role in the diffusion of knowledge is understudied. This article investigates the role of social media visibility of working papers on diffusion outcomes. While previous studies focused on the diffusion of STEM research, this article explores the diffusion of economic research. To do so, a data set of all NBER working papers published between 2015-2018, covering their social media mentions, as well as bibliometric and altmetric indicators, is used. To estimate the causal effect of social media visibility, an instrumental variable approach, leveraging quasi-random variation in social media posting policy of the NBER's communication office, is employed. The results indicate heterogeneity in the role social media play in the diffusion of economic research. Increased social media visibility of working papers positively affects the likelihood and the extent to which research is diffused to public discourse (measured by blog and news mentions), within the first year from publication, as well as within the scientific community (measured by academic citations), four years post-publication. No effect on citations in policy documents was found. Lastly, the likelihood to publish a working paper in a peer reviewed journal is found to be unrelated to social media visibility of the working paper. The results of this paper provide evidence for the role social media play in the diffusion of economic knowledge.

Exploring the Role of Social Media in the Diffusion of Economic Research

March 2, 2024

Abstract

For more than a decade, social media has become a key channel for knowledge dissemination used by scientists as a whole, and economists in particular. However, its role in the diffusion of knowledge is understudied. This article investigates the role of social media visibility of working papers on diffusion outcomes. While previous studies focused on the diffusion of STEM research, this article explores the diffusion of economic research. To do so, a data set of all NBER working papers published between 2015-2018, covering their social media mentions, as well as bibliometric and altmetric indicators, is used. To estimate the causal effect of social media visibility, an instrumental variable approach, leveraging quasi-random variation in social media posting policy of the NBER's communication office, is employed. The results indicate heterogeneity in the role social media play in the diffusion of economic research. Increased social media visibility of working papers positively affects the likelihood and the extent to which research is diffused to public discourse (measured by blog and news mentions), within the first year from publication, as well as within the scientific community (measured by academic citations), four years post-publication. No effect on citations in policy documents was found. Lastly, the likelihood to publish a working paper in a peer reviewed journal is found to be unrelated to social media visibility of the working paper. The results of this paper provide evidence for the role social media play in the diffusion of economic knowledge.

Keywords: Knowledge Diffusion, Social Media, Science Communication

JEL Codes: O30, O33, O39

1 Introduction

The diffusion of scientific research is key to the promotion of economic development (Sorenson and Fleming, 2004). Previous studies have focused on knowledge diffusion through, for example, knowledge spillovers, academic engagement, technology transfers, and more (Jaffe et al., 1993; Perkmann et al., 2013, 2021; Arora et al., 2021; Hausman, 2022; Andrews, 2023). This study diverges from previous research, which has predominantly concentrated on investigating the diffusion of STEM research through patents, by delving into the diffusion patterns of economic research, encompassing its unique dissemination domains of public discourse and policy citations (Blinder and Krueger, 2004; Yin et al., 2022). At the core of the present paper is the examination of social media as a central diffusion channel, given its emergence over the past decade as a primary communication channel utilized by scientists overall, and economists specifically (Rowlands et al., 2011; Darling et al., 2013; Sugimoto et al., 2017; Giusta et al., 2021; Howoldt et al., 2023). Although a growing number of sociological and scientometric studies have explored the motives and scope of this phenomenon, causal evidence documenting the role social media plays in the diffusion of research is lacking (Tonia et al., 2016, 2020; Chan et al., 2023). Therefore, the main research question in this study is *How and to what extent does social media visibility affect the diffusion of economic research?*

The diffusion of scientific knowledge fosters economic development as it leads to technological development and productivity gains. Economists have stressed the key role of technological development in promoting economic growth (Solow, 1956; Arrow, 1962; Uzawa, 1965; Romer, 1990; Aghion and Howitt, 1992; Kogan et al., 2017; Romer, 1990; Aghion and Howitt, 1992). Moreover, scientific research has long been considered a fundamental driver of technological change, dating back to Adam Smith's "The Wealth of Nations" 1776[1]. However, the diffusion of knowledge is critical to fostering the relationship between scientific technological change and economic development (Sorenson and Fleming, 2004).

[1]Smith (1776) refers to scientists as "philosophers", see also in Pavitt (2013).

In recent decades, scholars have studied different aspects, channels, and trends related to the diffusion of scientific knowledge. This vast literature explored multiple diffusion mechanisms, such as knowledge spillovers and the role of geography (Jaffe et al., 1993; Jaffe and Trajtenberg, 1996; Bikard and Marx, 2020; Hausman, 2022; Andrews, 2023), training in teaching, which embeds scientific knowledge in students and researchers who later move to industry (Belenzon and Schankerman, 2013; Cantoni and Yuchtman, 2014; Biasi and Ma, 2022). In addition, others studied diffusion-supporting policies or institutions, such as the Bayh-Dole act or technology transfer offices (Henderson et al., 1998; Jensen and Thursby, 2001; Mowery and Ziedonis, 2002; Sampat, 2006; Hausman, 2022; Lerner et al., 2024). Deviating from the focus of the previous literature on STEM research, the study provides insights related to the diffusion of knowledge in the context of economic research.

This research adds to the literature on scientific knowledge diffusion by focusing on an understudied context, economic research. Previous studies explored diffusion patterns, channels, and consequences of STEM fields' research, arguably due to two primary reasons. First, and rather technical, is the availability of data, and a clear 'paper trail' form scientific publications to patents. Second, it is possible to quantify the monetary value of scientific knowledge when it is embedded in patents, or related products. Unlike scientific knowledge produced in STEM fields, economic research is only seldomly featured in patents, as the relevant diffusion domains differ from those of STEM research. Yin et al. (2022) reported that economic research is mainly used by the public through "news" [2] and "governament". Consistent with this insight, this research will consider public discourse (proxied by blog posts and mainstream media mentions) and well as mentions in policy documents, to be the relevant diffusion domains for economic research. Besides defining and investigating the relevant diffusion domains for economic research this paper also explores the role of an understudied diffusion channel, social media.

Studies in sociology and scientometric documented how social media platforms become

[2] Blinder and Krueger (2004) found that the main source of economic information for the greater public are mainstream media (network and cable TV, as well as newspapers).

a key communication channel used by scientists as a whole and economists in particular (Rowlands et al., 2011; Darling et al., 2013; Sugimoto et al., 2017; Giusta et al., 2021; Howoldt et al., 2023). While some of these studies explored the motives and scope of this phenomenon (Mohammadi et al., 2018; Klar et al., 2020; Chugh et al., 2021), causal evidence documenting the role social media play in the diffusion of knowledge is lacking (Tonia et al., 2016, 2020; Chan et al., 2023). Hence, this article investigates the impact of social media visibility of working papers on their diffusion outcomes. In addition, this study joins the exponentially growing economic literature investigating the societal effects of social media platforms (Aridor et al., 2024).

To answer the research question, a dataset of all NBER working papers published between 2015-2018, linked to their future publication outcomes, was constructed. Next, it was supplemented with bibliometric indicators (from OpenAlex), such as journal impact factor and academic citations, as well as altmetric indicators such as blog posts, traditional media, and policy mentions (from Altmetric.com). Importantly, detailed data about social media visibility was collected. Specifically, this paper is using Twitter as a setting (Twitter data from Altmetric.com, complemented using the Twitter API). Studying the causal effect of social media visibility on diffusion indicators, as previous studies noted, is challenging due to endogeneity concerns. That is, for example, due to the existence of unobserved drivers of both higher visibility on Twitter, and increased diffusion indicators, or selection concerns rooted in authors' choices to promote their research on social media.

To overcome these challenges and to estimate the causal effect of visibility on social media, an instrumental variable approach is used, capitalizing on the random and dispersed social media posting policy of the NBER's communication office. Two instrumental variables are introduced, the first is "News Pressure Index", indicating whether on the day the NBER tweeted about a given working paper, a major news event occurred, potentially leading to reduced attention to the NBER post. The second is the "Twitter Attention Index", which indicates if the NBER tweeted about a working paper at a time economists tend to use the platform, and therefore potentially increase the paper's visibility. I find that increased social media visibility of working papers positively affects the likelihood and rate of blog posts, traditional media, and academic citations. No effect was found on the number of policy citations or the likelihood of publishing in a peer-reviewed journal. Further, I explore temporal variation in the headline effects.

2 Empirical Setting and Data Collection

This paper relies on NBER working papers (NBER-WPs) as the primary data source. The data collection process began by obtaining the DOI of each NBER-WP and its corresponding published version. Bibliometric information was gathered from two sources: OpenAlex provided traditional indicators such as publication date, citation count, journal and authors' information; while Altmetrics.com supplied alternative indicators including social media mentions, blog posts, traditional media, and policy mentions. However, due to legal restrictions on Twitter data, Altmetrics.com could not share detailed Twitter mentions information. Therefore, information on Twitter mentions was collected using TwitterAPI. In addition to these data sources, this research project utilize news pressure index used by Foerderer and Schuetz (2022), and a large dataset of #EconTwitter records collected by Enrico Bergamini. This section will describe the empirical setting, the database used, and the data curation process.

2.1 The NBER Working Paper Series

The National Bureau of Economic Research, based in Cambridge, Massachusetts, is a nonpartisan private organization. Its mission is to "disseminate research findings to academics, public and private-sector decision-makers, and the public"[3]. To support this mission, NBER publishes more than 1,200 working papers[4] per year; these are manuscripts that have not

[3]See 'About the NBER' page: `https://www.nber.org/about-nber`
[4]In some scientific fields the term preprint is used synonymously.

yet completed a peer review process. To publish a paper with the NBER two key conditions must be met. First, working papers should not include policy recommendations and are only screened to ensure that (working papers are not screened for the quality of the research).

Second, only NBER affiliates can submit a paper to the Working Paper Series, which means that at least one author of a working paper must be an NBER affiliate. To become an affiliate, one must be primarily employed by a North American institution and must be nominated and elected in a highly selective process. With more than 1700 affiliates to date, including 44 Nobel Prize winners and 13 former chairs of the (US) president's council of economic advisors, the NBER is home to North America's scholarly elite in the field of economics. A recent NBER-WP, forthcoming in the Journal of Political Economy, suggested that the NBER operates as a 'club' (Carrell et al., 2022).

Nevertheless, choosing the NBER as a setting has some meaningful empirical benefits. First, limiting this investigation to a given field, economics, alleviates concerns about field-level heterogeneity in norms regarding the use of social media, as well as the use of working papers (or preprints) as a common method of early dissemination of research (Klebel et al., 2020). Second, the NBER WP series is well known and highly regarded, which potentially reduces the risk of scooping associated with the release of preprints Hill and Stein (2019).

Several technical features promoted the use of the NBER Working Paper series as a setting for this project. First, each NBER Working Paper is posted as a page on the NBER website in a standardized way. These pages contain information such as title of the WP, names of authors, publication date, Digital Object Identifier (DOI) as well as records of revisions the working paper went through. Most importantly, the NBER website links working papers with their published version, by providing its reference. Data were collected on all NBER working papers published between 2010 and 2022. Figure 1 presents a summary of publication outcomes of NBER working papers.

Figure 1, show several clear trends. First, the number of working papers published each year seems fairly stable, with a slight upward trend. Second, the decline in the share of published working papers confirms the prolonged publication process in economics. Third, the number of working papers published in 2020 appears to be a clear outlier, a result of the COVID-19 pandemic. Overall, the share of published papers ranges between 50% and 60%, with the peak year being 2013 (68%), from which the share of those published by one of the top five economic journals is about 9%-13%. In addition, between 4%-9% of the working papers are published as a book chapter. These figures are slightly higher than those found in previous investigation of the NBER WP series (Baumann and Wohlrabe, 2020).

Choosing the NEBR as an empirical setting poses some potential challenges to external validity. First, the focus on a defined field (economics) could raise concerns about the generalizability of the results to other, STEM or even SSH scientific fields. Second, as noted above, the NBER could also be characterized as an elitist club.

An analysis of the institutional and geographical background of NBER affiliates provides supporting evidence for this characterization. Almost half (45%) of all NBER affiliates are based in ten universities[5]. Furthermore, when inspecting the distributions of institutions awarding Ph.D. degrees to NBER affiliates, a stark 68% of all affiliates were trained in these same ten universities[6]. In addition, NBER affiliates are concentrated in a handful of states[7], as indicated by their primary university affiliation, 60.5% of all NBER affiliates are based in seven states [8]. The identification strategy of this paper will aim to reduce the challenge to external validity posed by these observations.

2.1.1 NBER-WP Series: Sub-Fields and NBER Programs

The data collected from the NBER website were then supplemented with two additional indicators. First, since each WP must have at least one NBER affiliate as an author, and

[5]Harvard (7.6%), university of Chicago (6.2%), Stanford University (5.3%), Berkeley (4.4%), MIT (4%), Yale (3.9%), Colombia (3.9%), University of Pennsylvania (3.6%), NYU (3.3%), and Northwestern (3.2%). See Figure 9 in Appendix A.

[6]See Figure 7 in Appendix A.

[7]Only 5% of affiliates have a primary university affiliation outside of the US.

[8]California (19%), Massachusetts (15.7%), Illinois (10.5%), Pennsylvania (5.2%), Connecticut (4.6%), North Carolina (3.2%), and Texas (2.3%). See Figure 9 in Appendix A.

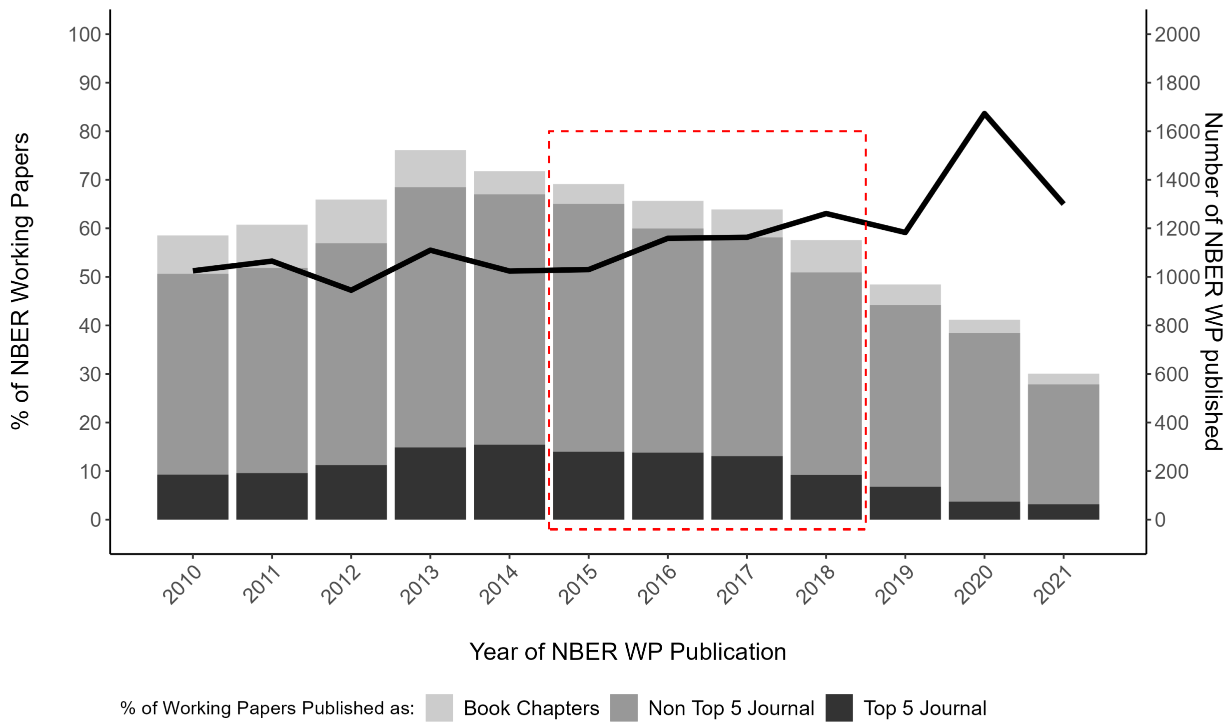


Figure 1: NBER Working Papers: publication outcomes and time trends, 2010-2022
Note: Data were collected from the NBER website. In dashed red rectangle: 2015-2018, the years used in this study.

each affiliate must be linked to at least one 'NBER program'[9], each WP could be linked to at least one NBER program. The database used for this research, linking working papers to NBER programs is publicly available online[10]. A second, and potentially a better proxy for the sub-field a WP is related to, are the WP's Journal of Economic Literature (JEL) classification codes[11]. At the time of submission to the NBER WP series, authors are asked to report JEL codes that best relate to the content of their paper. There is no limit to the number of JEL codes assigned to a WP.

Exploring the sub-field and topics distribution of NBER WPs used in this paper's, some patterns emerge. First, papers dealing with the broad categories of "Labor and Demographic Economics", "Health, Education, and Welfare", "Microeconomics", and "Macroeconomics

[9]The NBER, as of 2023, has 19 programs. Those are considered to be the organization's 'backbone'. Each of the programs is centered around a traditional economic research sub-field. For more information, visit: `https://www.nber.org/programs-projects/programs-working-groups`.

[10]See: `https://data.nber.org/nber-wp-logs/`.

[11]See: `https://www.aeaweb.org/econlit/jelCodes.php?view=jel`.

and Monetary Economics" are the most frequent in this sample, with more than a fifth of papers each. "General Economics and Teaching", "History of Economic Thought", and "Miscellaneous Categories" are the least frequent, with less than 1% of all papers[12]. Looking into the NBER working group assigned to each paper (based on internal NBER author affiliations), it is shown that papers related to "Labor Studies","Public Economics", and "Economic Fluctuations and Growth" are the largest groups, with more than 20% of all papers, and "Technical Working Papers" are the least frequent[13]. Next, using both the WP's and the linked journal publication's DOI, traditional bibliometric information, as well as alternative matrices, were collected from OpenAlex and `altmetric.com`, respectively.

2.2 Dependent and Independent Variables

To study the effects of social media on the diffusion of knowledge, several key inputs are required. First, indicators measuring the diffusion and impact of a scientific work must be collected. Second, to properly construct the experimental setting and account for time trends, timestamps are of critical importance. However, the NBER website does not provide this information, resulting in the need to obtain these data from other sources. Traditional bibliomentric data (e.g. citations, journal information) were collected from OpenAlex. OpenAlex is an open access repository of scientific publications, with almost triple the number of records found in the commercial Scopus and Web Of Science[14]. Next, `altmetric.com` was used to collect alternative matrices. `altmetric.com` collects mentions of scholarly work from a variety of sources, from social media (critical for this research), to traditional (mainstream) media, policy documents and blog posts.

The collection of data from OpenAlex and `altmetric.com` data followed a similar path. Using the data collected from the NBER platform, a query requesting data related to a list of DOIs, of both WP and published papers, was sent to these two repositories. Most papers

[12]See Table 8 in Appendix A.
[13]See Table 9 in Appendix A
[14]For more information, see: `https://openalex.org/about`.

(WP and published) were found in both repositories.

The key inputs to the empirical analysis of the research, obtained from `altmetric.com` are mentions of scholarly work on Twitter, this is due to two reasons. First, the independent variable of this study is the visibility on Twitter. Second, the empirical design applied in this article uses the timing of each NBER WP tweet by the NBER's Twitter account (@nberpubs)[15]. However, due to legal constraints, `altmetric.com` only provide mention and author identifications for Twitter mentions data, lacking some important metadata on each mention such as the timing of the tweet, it's content, and related matrices (e.g. like, retweet, comment counts). To overcome this deficiency and using the IDs provided, detailed data about each Twitter mention was collected directly from Twitter, using the Twitter API service.

2.2.1 Dependent Variables

Following the construction of the database as described above, a number of outcome variables were calculated. As the publication dates of the sample's WPs span across four years (2015-2018), it is important to make sure that the variables are adequately measured. For example, counting the accumulated number of citations, at the time of data collection[16], for a WP published in January 2015 relative to a comparable WP published in January 2018 would mechanically produce a higher citation count for the earlier WP, as it had more time to accumulate citations.

As a result, using the timestamp on all mentions' records collected from the NBER, `altmetric.com`, and OpenAlex, a count of mentions in blog posts, traditional media, policy documents, and forward citations by years from publication was calculated (up to 4 years after the WP publication). The definitions of what constitutes a blog post or a mainstream media mention are following altmetric.com, who track more than 15,000 different academic

[15] More on the empirical design in Section 3.

[16] Most data was collected during early 2023, but an additional round with Altmetric.com was conducted during late 2023.

and non-academic blogs, and about 5,000 news organizations worldwide (English and non-English speaking). When it comes to mentions in policy documents, altmetric.com defines policy documents as "any policy, guidance, or guidelines document from a governmental or non-governmental organisation"[17].

To provide the reader with a clearer understanding of these diffusion measures, Tables 10, 11 and 12 in appendix 5 provide a list of the top 20 sources of blog, media, and policy mentions related to this paper's sample. A common trend in all three indicators is that the mentions are fairly distributed across sources. Only four blog sources (National Affairs Online, himaginary, The Brookings Institution, and Marginal Revolution) account for more than 5% of all blog posts mentions. The corresponding numbers for news and policy mentions are zero and six, respectively.

In addition to diffusion outcomes, a set of variables related to the publication process of NBER working papers were calculated. The information regarding publication date of linked journal publication allows, in most cases, to produce a binary variable that gets a 1 if the WP was published within the first 4 years and a 0 if not. Additionally, using journal publication information, a similar binary variable, capturing publications in the top 5 economics journals[18] Finally, using publicly available data from the NBER platform, the number of revisions was calculated, that is, the number of times the authors sent a new version of a WP to the NEBR.

2.2.2 Independent Variable: Social Media Visibility

Using data obtained from Twitter and `altmetric.com`, a measure of the visibility of WPs on Twitter was calculated. To do so, the sum of WP mentions in tweets was calculated with their related 'likes count within the first week from an @nberpubs tweet. This sum is then log-transformed, to improve interpretability, and to reduce the skewness of the variable.

[17] See: https://help.altmetric.com/support/solutions/articles/6000235983-attention-sources-tracked-by-altmetric

[18] Top 5 journals are: The American Economic Review, Econometrica, the Journal of Political Economy, the Quarterly Journal of Economics, and the Review of Economic Studies (Heckman and Moktan, 2020).

Figure 3 presents the 'Twitter Visibility' measure for the sample of NBER WP used in this paper. Two clear trends are shown; first, an upward time trend is visible, in line with the overall expansion of Twitter's network size. Second, the spread of values is between 0[19] and 7, with no major outliers.

To offer a justification for the choice of a one-week window, Figure 2 presents the temporal accumulation of NBER WP tweets. Time 0 represents the posting time of the @nberpubs tweet. As shown in Figure 2, 34% of the mentions related to a NBER WP are tweeted before the @nberpubs tweet and therefore are not subject to the kind of quasi-randomised timing used for the empirical design. A possible explanation for this observation is that there is a delay between the time a WP is released on the NBER website and the time @nberpubs tweets about the WP. As WPs can be mentioned online by other Twitter users, including the authors themselves, others who follow the NBER WP series closely, such as those who are subscribed to the NBER weekly newsletter "New This Week" (Feenberg et al., 2017).

The figure also shows the short life cycle of NBER WPs on Twitter, 55% of all mentions are posted within a day following the @nberpubs tweet, and 64%, 74%, and 85% are posted within a week, a month, and a year, respectively. This short digital life cycle is consistent with previous studies on the visibility of academic work on social media (Shuai et al., 2012; Fang et al., 2021). Therefore, this evidence suggests that basing the visibility index on the mentions in the first week after the @nberpubs post can capture most of the mentions related to @nberpubs tweet. Section 3 provides a detailed discussion of the importance of the @nberpubs tweet for the empirical design of this article.

2.2.3 Instrumental Variables

As elaborated in detail in the next section (Section 3), the analysis addresses endogeneity concerns, which can lead to biased estimations of the impact of social media visibility on the diffusion of economic knowledge using an instrumental variables approach. Two instru-

[19] value of 0 means that the @nberpubs tweet is the only mention a WP received without any associated likes. This seems to be a relatively rare event in this sample.

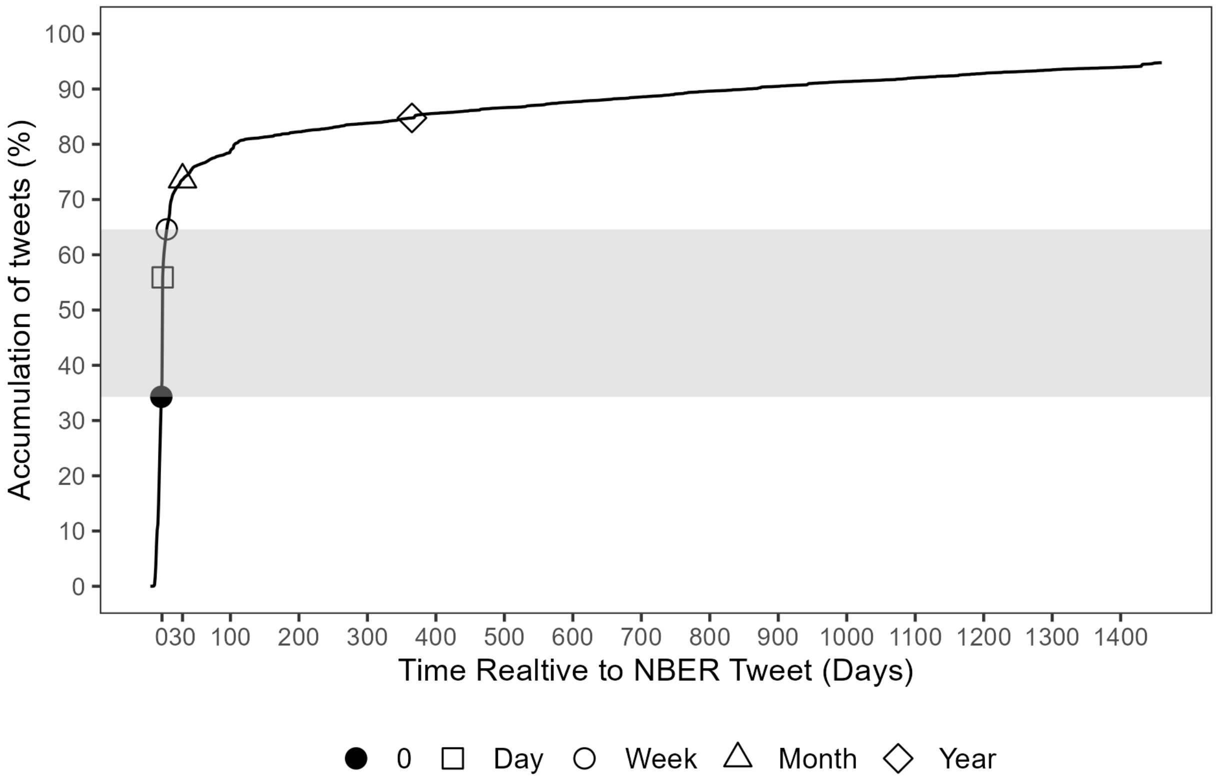


Figure 2: NBER WPs' Twitter mentions - temporal accumulation
Note: All tweets posting times were converted to Eastern Time Zone (EST), regardless of the location of the user posting it. The first week from the @nberpubs tweet about a working paper (time 0), is highlighted as it is used for variable construction.

ments are used, the #EconTwitter Attention Index and a news pressure measure. Next, a description of their data collection processes and how they were calculated is provided.

2.2.3.1 #EconTwitter Attention Index

Shortly prior to the restrictions on academics' use of the unlimited Twitter API, a data collection effort, led by Enrico Bergamini[20], was underway. More than 28 million tweets, posted by 2172 economists between 2007-2022 were collected. Bergamini's data collection strategy was simple, he downloaded, using the Twitter API, the entire tweeting activity of all economists listed in the *RePEc-registered Economists on Twitter* [21]. While his database

[20]PhD Student at the University of Turin, `https://www.enricobergamini.it/`

[21]The last available *RePEc-registered Economists on Twitter* list can be found here: `https://ideas.repec.org/i/etwitter.html`.

Figure 3: Likes weighted Twitter mentions (1st week after @nberpubs tweet) over time (log), 2015-2018

includes the content of these tweets, as well as user and tweet ID, it also registers the exact timing each tweet was posted. This feature of his database is key to the creation of the first instrument used in this paper, the 'Twitter Attention Index'.

The first step in calculating the 'Twitter Attention Index', is stratifying all observations in Bergamini's database into groups, by the year, the day of the week, and the hour of the day they were posted. Then, standardizing it to obtain the z-score for each hour of a weekday in a given year. As expressed in formula 1:

$$TwitterAttentionIndex_{ydh} = \frac{N_{ydh} - \bar{N}_{ydh}}{sd(N_{ydh})} \tag{1}$$

Where N_{ydh} is the number of tweets in year y day of the week d and hour of the day h. $\bar{N}_{ydh}$ is the mean of N_{dh} in year y and $sd(N_{dyh})$ is its standard deviation. Table 1 presents the standard deviations and means of N_{dh} over time.

The choice to design an annually standardized index is a result of some tension. First, calculating the index on a yearly basis accounts for possible changing Twitter usage patterns over time, as well as the exponential growth in Twitter's network size and increasing number of tweets at the relevant time period. However, it makes the assumption that a relatively busy Twitter feed in year y has the same meaning as that in year $t+4$. One might challenge this assumption, as the absolute number of economists and their tweets grows over time; see Table 1. Indeed, this concern would be easy to disregard if the growth in Twitter's network size and values of content shared on the platform, accompanied with an improvement to (1) the design of the Twitter algorithm, optimizing for user interests, and (2) the cognitive ability of users to filter content while using a social media platform. Although there is no clear evidence to support or refute this concern, the main specifications of this document use annual standardization. At the heart of the design choice is the need to account for time trends in the size of the network. Furthermore, the choice to standardize the index improves interpretability and accounts for possible volatility in the data (although Table 1 provides suggestive evidence that volatility is not a major issue in the data).

Table 1: #EconTwitter sample - descriptive statistics

	Year	# of Economists	# of Tweets	sd # of Tweets	mean # of Tweets
4	2015	1377	1929628	4203.593	11485.88
3	2016	1515	2293091	4748.350	13649.35
2	2017	1674	2693125	5731.857	16030.51
1	2018	1821	2768298	5873.930	16477.96
5	2015-2018	1895	9684142	5548.621	14410.93

Figure 4 presents a yearly heat map of the 'Rush Hour Index'. Few clear patterns emerge. First, economists tweet relatively less on weekends. Second, the morning hours of working days, particularly 9AM, are the busiest hours in terms of content posted by economists on Twitter. To a lesser extent, the early afternoon hours of 15-17 also exhibit increased posting activity[22]. Perhaps unsurprisingly, night hours experience less tweeting activity, especially around 23-01, which comprises the overlap of night time on both sides of the Atlantic Ocean.

[22]Previous studies found consistent time trends in social media engagement (Spasojevic et al., 2015).

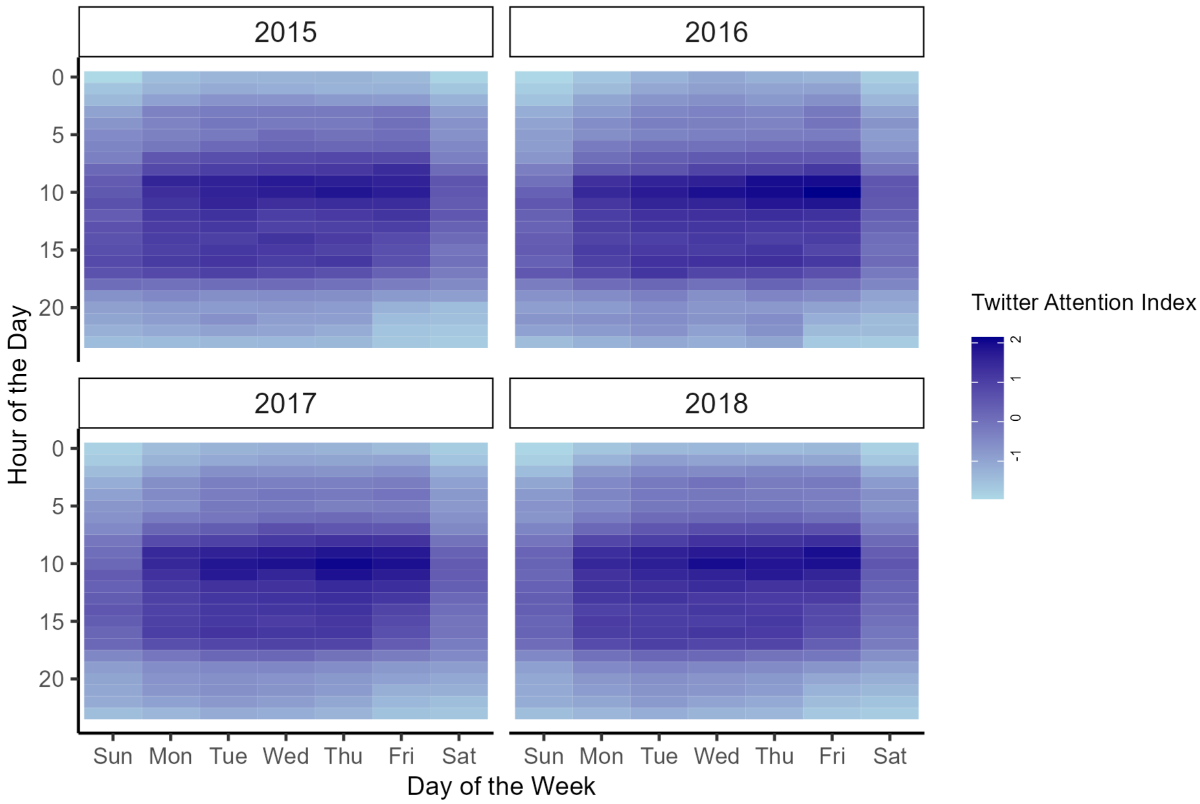


Figure 4: Twitter Attention Index - Yearly Heat Maps
Note: All tweets posting times were converted to Eastern Time Zone (EST), regardless of the location of the user posting it.

To conclude, Figure 4 clearly suggests that patterns of economists' tweeting timing are fairly stable over the years, even with the increase in network size and volume of tweets posted (shown in Table 1).

2.2.3.2 News Pressure

The second instrument used in this paper is capturing the intensity of the economic and business-related news in a given day. The measure of news pressure used in this paper is based on the *Wall Street Journal*'s (WSJ) front page "What's News" section. Every weekday, the WSJ collects the main stories of the day and lists them on the main page, under two categories 'world wide' and 'Business and finance'. The WSJ front page space occupied by the "What's News" section barely fluctuates overtime. Therefore, the number of news items in the 'Business and finance' section could serve as an indication for news pressure of high relevance to the economist, that is, how many news events are worthy of the readers' attention (according to the editors of the newspaper).

A previous study by Foerderer and Schuetz (2022) collected the number of items on the WSJ's "What's News" list between 2008-2018[23], and kindly agreed to share the data[24].

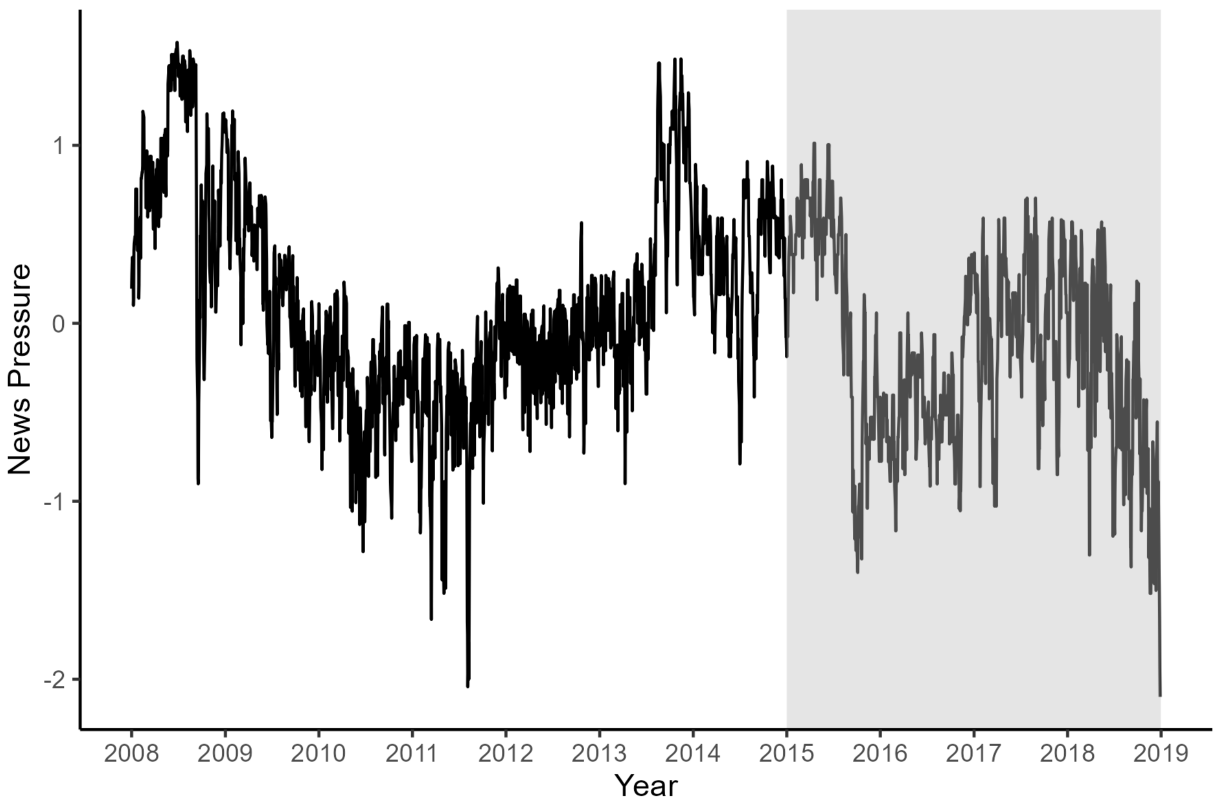


Figure 5: News Pressure Variable
Note: Date from Foerderer and Schuetz (2022). The highlighted period (2015-2019) is the one used in this analysis.

During the process of data curation, Foerderer and Schuetz (2022) noticed several shortcomings and, therefore, made corrections embedded in the final variable used in this paper. First, as a result of their specific research objective, they do not count news items related to data breaches[25] when calculating the news pressure variable. Second, the news pressure

[23]One should mention the growing use (in instrumental variables empirical designs, but not exclusively) of news pressure indicators in the economic and management literature (Eisensee and Strömberg, 2007; Manela, 2014; Jetter, 2017; Garz and Sörensen, 2017; Durante and Zhuravskaya, 2018; Peress and Schmidt, 2020).

[24]Specifically, Jens Foerderer was kind enough to permit the use of WSJ's "What's News" data for the benefit of this research project.

[25]As their study investigates firms' disclosure timing strategy when dealing with a data breach (Foerderer

variable provided by Foerderer and Schuetz (2022) is a seven-day moving average (an average of a given day, with a window of three days before and after). Third, the authors observed a major change in the layout of the “What’s News” section (occurred in August 2013), reducing it’s front page space from two columns to one. They address this change by standardizing the variable before and after the change. One major limitation of the data (in it’s current form), is the absence of weekends and holidays, as the WSJ only publish on weekdays, as well as a weekend addition on Saturdays, yet the Saturday addition’s “What’s News” section takes a different form than the weekdays’ one. The final ‘News Pressure’ variable is presented in Figure 5, where the period studied in this paper is highlighted. An important takeaway from Figure 5 is the presence of variation over time.

2.2.4 Final Sample and Descriptive Statistics

The final sample used for this study contains 2916 NBER WPs published between 2015-2018. Table 2 presents the summary statistics of the sample. This final sample, the observant reader would note, contains NBER WP that were (1) successfully linked to both `altmetric.com` and OpenAlex, (2) data about their @nberpubs tweet were available, and (3) received their @nberpub tweet on a weekday. In the first year since NBER WP is released, an average of 3 academic citations are registered (with a median of 1). However, as the standard deviation and maximum value suggest, there is substantial variation between WPs. When it comes to policy mentions, news stories, and blog posts, more than half of WPs are never mentioned, and the average number of mentions is 0.06, 0.54, and 0.44 respectively. About 50% of the WPs are published in the first four years after their release. 11% are published in the top 5 economic journals. This is lower than the shares presented in Figure 1, which is due to the censoring of the variable (Figure 1 does not condition on being published within the first 4 years).

Further, when it comes to visibility on Twitter, the table clearly shows that all NBER

and Schuetz, 2022).

WP receive at least one mention on Twitter (as each NBER WP is posted at least once on the platform, by the @nberpubs account). The average number of tweets and the 'likes' related to them is 28, with a median of 16. When applying a natural log on the likes weighted tweets indicator, the mean is about 3. As noted above, the minimum number of tweets and likes is 1, hence, by construction, when applying logarithmic scaled variable has a minimum of 0. The variation of both indicators indicated by the standard deviation of both of these variables is in line with that shown in Figure 3.

Lastly, Table 3 presents pairwise correlations between the different outcome variables, the Twitter visibility measure, and the proposed instrumental variables (discussed in the next section). It is clear that there is a correlation between the different outcomes, and between the outcomes and the main independent variable (likes weighted tweets). The two proposed instruments are correlated with the twitter visibility variable, although with different signs (positive correlation with 'Twitter Attention Index' and a negative one with 'News Pressure').

Table 2: Summary statistics - NBER working papers, 2015-2018

Statistic	N	Mean	St. Dev.	Min	Median	Max
# of academic citations (1 year)	2,916	3.123	7.306	0	1	162
# of policy citations (1 year)	2,916	0.066	0.403	0	0	8
# of News stories (1 year)	2,916	0.541	3.900	0	0	188
# of blog posts (1 year)	2,916	0.440	1.058	0	0	13
# of WP revisions (4 years)	2,916	0.434	0.894	0	0	9
Published (4 years)	2,656	0.498	0.500	0	0	1
Published in Top 5 (4 year)	2,656	0.114	0.318	0	0	1
Likes weighted tweets (1st week)	2,916	28.145	58.286	1	16	1,637
Likes weighted tweets (1st week) - log	2,916	2.813	0.913	0.000	2.773	7.401

Sample: NBER WP 2015-2018.

Table 3: Correlation matrix, 2015-2018

	# of academic citations	# of policy citations	# of News stories	# of blog posts	# of WP revisions	Published	Top 5	Likes weighted tweets (1st week)	Twitter Attention Index	News Pressure
# of academic citations	1.00									
# of policy citations	0.21***	1.00								
# of News stories	0.08***	0.06***	1.00							
# of blog posts	0.20***	0.25***	0.34***	1.00						
# of WP revisions	0.11***	0.02	0.07***	0.09***	1.00					
Published	0.08***	0.01	0.01	0.07***	0.04**	1.00				
Top 5	0.16***	0.00	0.05**	0.07***	0.15***	0.36***	1.00			
Likes weighted tweets (1st week)	0.16***	0.13***	0.11***	0.31***	0.04**	0.04*	0.07***	1.00		
Twitter Attention Index	0.05**	-0.01	0.07***	0.07***	0.03	0.00	0.01	0.07***	1.00	
News Pressure	0.01	-0.02	-0.05***	-0.05**	-0.03*	0.02	0.00	-0.12***	0.03*	1.00

*p<0.1; **p<0.05; ***p<0.01.

Note: Academic citations, policy documents, news mentions, and blog posts are measured in the first year since NBER WP publication, All outcome variables are measured for the first year since a NBER working paper publication date. Published and published in top 5 is measured in first four years. Sample: NBER WP 2015-2018.

3 Identification Strategy

Estimation of the effect of social media visibility on the diffusion of scientific knowledge, in a potential laboratory conditions, would be fairly simple (although unfeasible). The (social) scientist would randomly assign, to a carefully constructed sample of economists, different levels of visibility to working papers, and estimate how this 'treatment' affected an array of outcomes of interest using a simple OLS regression (applying it on this paper's setting):

$$Y_i = \beta_0 + \beta_1 TwitterVisbility_i + \varepsilon_i \tag{2}$$

Where $TwitterVisbility_i$ is a randomly assigned treatment on a NBER WP i. Y_i represents a set of outcome variables (such as the number of forward citations, the number of media mentions, etc.). β_1 is the coefficient of interest that measures the effect of increased visibility on Twitter on the relevant outcome. ε_i is an error term.

For many reasons (ethics, access to social media algorithm, to name a few), this experiment is not feasible, forcing the reliance on observational data. However, estimating a model similar to Model 2 would potentially result in a biased estimation of the effect. The main concern in this case would be endogeneity. The main sources of endogeneity concerns in this setting would be selection, omitted variable bias, and reverse causality. This section will detail how the setting chosen for this paper's empirical investigation, as well as the paper's identification strategy, addresses these endogeneity concerns.

First, a clear challenge when studying the effects of social media visibility on diffusion outcomes is the fact that knowledge, or papers, cannot tweet about themselves, it is rather a choice of the authors, readers, or related organizations, to post about them. This fact has empirical implications, as articles with any social media visibility might be systematically different than those without a social media presence. Choosing the NBER WP series addresses this concern by the fact that each NBER WP obtains at least one post on Twitter[26],

[26]The NBER communication office post about NBER WPs in all major social media platforms, including Facebook, LinkedIn and BlueSky.

regardless of the authors' preferences. The NBER communication office releases a standard tweet per paper; the tweet includes a short description of the paper[27], authors' name[28], a link to the full paper, and an image of the front page of the WP (containing the title and abstract of the paper). Importantly, the NBER communication strategy reduces concerns of selection, as it creates a setting in which NBER WPs 'have a life of their own' on social media, even if the authors themselves do not use social media.

Second, while this study examines, among others, the impact of social media visibility on diffusion to public discourse, it is plausible to consider that the relationship could be bidirectional. For example, news coverage on the insights provided by an NBER WP could boost visibility of this hypothetical paper. To address this reverse causality concern, the social media visibility variable is constructed to count only Twitter mentions and likes during the first week following the @nberpubs post. Additionally, the outcome variables used in the analysis exclude all mentions within that first week.

Third, one could think of multiple unobserved potential confounders, which affect both $Twitter\ Visbility_i$ and Y_i, leading to a biased estimation of β_1, due to an omitted variable bias. For example, working papers written by prominent researchers might attract more social media attention, as prominent researchers tend to have: larger network and increased social status, but also produce more timely and frontier research. Similar unobserved characteristics could also have a positive effect on impact measures (such as citations from other researchers, policymakers, blogs, or mainstream media). This endogeneity problem is present in most previous studies that report correlations between the visibility of research on social media and other impact measures. To overcome this obstacle, this paper capitalize on the quasi-random and disperse communication policy of the NEBR to apply an instrumental variables design.

[27]The short description is usually written by the authors, if not, by NBER communication officer.

[28]Authors who provide their Twitter handle to the NBER when submitting a WP are 'tagged' in the @nberpubs post.

3.0.1 @nberpubs Tweeting Policy

Critical to the empirical design of this research is the NBER communication strategy. Upon receiving a WP from an NBER affiliate, a serial number is assigned to the manuscript; this number is also integrated in the WP's DOI. The serial number is sequential and is determined by the order of papers that arrive at the NBER. For example, if affiliate A and affiliate B both send a paper to the NBER on a given day, at 13:00 and 13:01, respectively. The paper sent by affiliate A receives a serial number 1 while the paper of B is assigned number 2. Assuming that NBER affiliates do not have information on the timing of WP submission by their colleagues, the serial number assigned to each WP is quasi-random to the affiliate (Feenberg et al., 2017). Next, the manuscript is screened to ensure it does not contain policy recommendations, and once cleared, it is uploaded to the NEBR WP series website.

At this point, the paper is included in a pipeline of WPs, pending promotion on social media platforms. As noted above, the NBER releases a standardized post on multiple social media platforms. Since 2015, each week, an NBER communications officer prepares a schedule to ensure that the release of posts on social media is distributed throughout the hours of the week[29]. The order in this posting pipeline is determined by the serial number of the WP.

The NBER communication policy has several implications and promises related to this paper's identification strategy. First, it permits the use of the timing of @nberpubs posts about NBER WPs as random to the authors and the audience, and therefore as a generator of randomization in social media visibility. Second, as the even distribution of @nberpubs timing only begins in 2015, it determines the starting point of the sample. Therefore, to operationalize this randomization in the quest of alleviating endogeneity concerns, two instrumental variables, related to the timing of the @nberpubs tweet, are used.

[29]Before 2015 @nberpubs posted on most WP at the same hours of week.

3.1 Instrumental Variables

To apply the instrumental variables design, one should find at least one valid instrument, that is, an instrument that is exogenous and relevant. Importantly, one should make sure that the only possible causal path between the instrument and the outcome passes through the endogenous variable (Cunningham, 2021). Using the disperse and quasi-random timing of the @nberpubs tweet related to each NBER WP, two instruments are used: the '#EconTwitter Attention Index' and 'News Pressure' indicator. Moreover, two conditions must be met, the first, which could be empirically tested, is relevance, while the second, lacking a standardized statistical test, is validity.

To address the validity of these two instruments, the scholar must convince the reader that the relationship between the instruments and the outcome exists only through the endogenous variable. The quasi-experimental conditions used in this setting play an important role in this undertaking. It is hard to believe that neither instruments, the level of news pressure nor the observed average level of economists' Twitter activity affect the future outcomes of a NBER WP if not through the visibility of that WP on social media. Additionally, the instruments also meet the relevance criteria, as discussed below (in section 3.2). These two instruments are used in a first stage estimation, where the Twitter visibility is regressed on '#EconTwitter Attention Index' and 'News Pressure' variables:

$$TwitterVisbility_i = \beta_0 + \beta_1 AttentionIndex_t + \beta_2 NewsPressure_t + \varepsilon_i \tag{3}$$

Where $TwitterVisbility_i$ is the sum of tweets and likes related to an NBER working paper i in a time frame of 7 days after @nberpubs tweet is posted. $AttentionIndex_t$ is the attention index value assigned to NBER WP i, based on time t of @nberpubs account tweeted about it. Specifically, the value is assigned based on the hour of the week in a given year (e.g. Monday, 09:00-10:00 AM EST, 2016) and regardless of its calendar date. $NewsPressure_t$ represents the news pressure indicator for the calendar date the @nberpubs tweet is released.

β_1, β_2, are the correlations between $AttentionIndex_t$, and $NewsPressure_t$, respectively. β_0 denotes the intercept, and ε_i, is the error term.

When considering the two instruments carefully, a concern regarding serial correlation arises. Two NBER WP tweeted by the @nberpubs account on the same date and at the same hour (say, 1 March 2016, one at 09:15 and another at 09:20), will receive similar values of both instruments, leading to potential serial correlation. To address this issue, standard errors are clustered at the hourly level. That is, the time frame studied (1 January 2015 to 31 December 2018) is divided into hourly units, on which the standard errors are clustered.

Further, commonly seen as indicative of confirming the relevance of instruments is the F-statistic of the first-stage estimation. However, the use of clustered standard errors in the estimation of instrumental variables with one endogenous variable and two instruments warrants the use of 'Effective F-statistic' (James H Stock and Yogo, 2002; Olea and Pflueger, 2013; Lal et al., 2023)[30]. The critical value is similar to the convectional 'rule of thumb' of $F^{Effective} > 10$ (James H Stock and Yogo, 2002). Following the estimation of Model 3, the predicted values are calculated and integrated in the second stage estimation:

$$Y_i = \gamma_0 + \gamma_1 TwitterVisbility_i^{IV} + \eta_i \tag{4}$$

Where Y_i is paper i's focal outcome variable considered, and $TwitterVisbility_i^{IV}$ is the predicted value of $TwitterVisbility_i$, obtained from the first-stage model (Model 3). γ_0 and η_i are the intercept and error term, respectively. As noted above, standard errors are clustered at the hour level. γ_1 is the coefficient of interest.

3.2 First Stage

The results of the estimation of Model 3 (first-stage regression), are used to justify the use of the two proposed instruments and address the relevance condition. Table 4 shows the

[30]To obtain the value of 'Effective F-statistic', an R package "ivDiag" was used. See: https://cran.r-project.org/web/packages/ivDiag/ivDiag.pdf

results of the first-stage estimation. Model (1) and (2) estimate separately the relationship between the endogenous variable, Twitter visibility, and the instruments, Twitter attention index and news pressure, respectively. Model 3 estimate the first stage regression integrating both instruments, this is the model used in the final IV estimation. The sample used for the estimations presented in Table 4 is similar to the one used in most IV estimations[31].

The results presented in Table 4 indicate that both instruments correlate with the endogenous variable. This is in line with the correlations reported in table 3. As also indicated in Table 3, the two instruments are only marginally correlated with each other. This is reflected in model (3), since including both instruments in the same first-stage estimation seems to barely change their coefficients. The coefficient of the Twitter Attention Index instrument indicates a positive and statistically significant impact on the endogenous variable. This suggests that with more observed activity of economists on Twitter in a given weekly-hour on which the @nberpubs account releases a tweet, the more visibility the related NBER WP gets on Twitter. The coefficient of news pressure reports a negative and statistically significant impact of increased news pressure on the day of @nberpubs tweet on the visibility of a NBER WP. As one might expect, the discourse on social media is not detached from real-world events, so this result indicates that increased attention to news events might shift attention away from discourse about NBER WPs. Finally, Table 4 reports the effective F-statistic suggested by Olea and Pflueger (2013). Across all three models, the effective F-statistic is much above the conventional $F^{Effective} > 10$ rule of thumb (James H Stock and Yogo, 2002), suggesting that the instruments are relevant and 'strong'.

4 Results

This section will outline the main results of the paper. First, it will answer the main research question by investigating the effects of social media visibility on the diffusion of economic

[31]When estimating the impact on the likelihood to get published, several observation are dropped of the sample, as they are not successfully linked to a publication outcome, see Table 7. When running the first-stage for this smaller sample, similar results are found.

Table 4: First Stage Regressions (on Twitter likes weighted mentions (log))

	Dependent variable: Twitter Visibility (log)		
	Full Sample		
	(1)	(2)	(3)
Twitter Attention Index	0.113***		0.120***
	(0.029)		(0.029)
News pressure (WSJ)		−0.195***	−0.198***
		(0.030)	(0.030)
Intercept	2.691***	2.776***	2.647***
	(0.035)	(0.018)	(0.036)
Effective F - Statistic	14.75	40.03	27.25
Number of clusters	2822		2822
Observations	2,916	2,916	2,916
Adjusted R^2	0.004	0.013	0.018
Residual Std. Error	0.912 (df = 2914)	0.907 (df = 2914)	0.905 (df = 2913)

Note: Standard errors are clustered at the time-level (hour). *p<0.1; **p<0.05; ***p<0.01. Results for 4 years past WP publication date. Sample: NBER WP 2015-2018.

research. It will include an analysis of both the extensive and intensive margins. While the first section focuses on the relativity short term, that is, a year past NBER WP publication, the second subsection reports how these effects vary over time. The third subsection will dive into the effects of social media visibility on the process related to knowledge production.

4.1 Social Media and the Diffusion of Economic Knowledge

This subsection contains the main results of this paper. The results for the first year after NBER WP publication are reported in Tables 5 (for the extensive margin) and Table 6 (for the intensive margin). For each of the diffusion outcome variables [32], two estimations are reported, an IV estimation and an OLS estimation. These correspond to models 4 and 2, respectively. In addition, for the IV models, model diagnostics, effective F-statistic of the first-stage, as well as the p-values both the Wu-Hausman and Sargan tests (as the model includes more than one instrument) are reported.

[32] diffusion outcome variables are: mentions in blog posts, news stories, policy documents, and academic papers.

Table 5 reports the extensive margin of the effect. To do so, the count variable for each of the diffusion outcomes was converted to a binary variable, getting a 0 if no related mention was observed and 1 for at least one mention. This estimation illuminates whether social media visibility influences the likelihood of a NBER Working Paper being mentioned in a particular diffusion domain.

The results depicted in Table 5 demonstrate that social media plays a varied role in diffusing economic research across different domains. Columns (1) and (3) suggest that an increase of one standard deviation in the visibility on social media leads to an increase of 36% and 57% in the probability that a WP will be mentioned in blog posts and news stories, respectively. These positive effects are statistically significant at the 1% level. Furthermore, these effects appear to be higher than the OLS estimators presented in Models (2) and (4), suggesting that the OLS coefficients are downward biased. Indeed, the reported Wu-Hausman test's p-value confirms that endogeneity is present in the OLS estimation, potentially resulting in a biased OLS estimator. As per the reported Sargan test's p-value, the IV estimations do not violate the overidentification restriction (at a 5% level) for the blog posts estimation, but does for the news stories one.

Although the results suggest that social media play a role in the diffusion of economic knowledge to public discourse (as measured by blog posts and news stories), when it comes to the policy and scientific domains, a different image is discovered. The OLS estimates in columns (6) and (8) suggest that there are positive correlations between visibility on social media and the likelihood of being mentioned in policy documents and academic publications. However, the IV estimations provided in models (5) and (7) suggest that these correlations might be the result of the presence of upward bias. Although the Wu-Hausman test for Model (5) suggests that endogeneity might not be the source of this bias. Here, we should remind the reader that outcomes in Table 5 are measured for a one year past NBER WP publication window. As both policy citations and academic citations take time to accumulate, Subsection 4.2 further explores the dimension of time.

Table 5: IV and OLS regressions results - extensive margin

	Dependent variable:							
	Blog posts		News stories		Policy documents		Academic citations	
	instrumental variable	*OLS*	*instrumental variable*	*OLS*	*instrumental variable*	*OLS*	*instrumental variable*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Twitter Visibility (log)	0.360***	0.132***	0.569***	0.102***	0.018	0.025***	−0.049	0.087***
	(0.070)	(0.008)	(0.087)	(0.009)	(0.029)	(0.005)	(0.072)	(0.009)
Intercept	−0.758***	−0.116***	−1.343***	−0.029	−0.013	−0.030**	0.793***	0.412***
	(0.198)	(0.023)	(0.243)	(0.024)	(0.080)	(0.012)	(0.203)	(0.028)
Effective F-stat (1st Stage)	27.25		27.25		27.25		27.25	
Wu-Hausman (p-value)	0		0		0.83		0.05	
Sargan (p-value)	0.94		0		0.2		0.04	
Mean outcome variable	0.25	0.25	0.26	0.26	0.04	0.04	0.66	0.66
Number of clusters	2822		2822		2822		2822	
Observations	2,916	2,916	2,916	2,916	2,916	2,916	2,916	2,916

Note: Standard errors are clustered at the time-level (hour) *p<0.1; **p<0.05; ***p<0.01. Results for the first year past WP publication date. Sample: NBER WP 2015-2018.

Beyond the extensive margin, Table 6 provides the results of the intensive margin estimations. This investigation informs how social media increase the degree to which knowledge is diffused to focal domains. Overall, the results are consistent with those of the extensive margin. Columns (1) and (3) suggest that an increase of one standard deviation leads to an increase of 0.64 and 2.5 blog posts and news stories within the first year, respectively. Similarly to the extensive margin results in Table 5, the policy and academic citations IV coefficients, in columns (5) and (7), are not statistically significant, unlike the OLS estimators in, in columns (6) and (8) which show a positive and statistically significant coefficients. The Wu-Hausman test for models (5) and (7) accepts the null hypothesis that endogeneity is not a concern in this estimation.

4.2 Temporal Patterns

As noted above, the results in Subsection 4.1 confine the outcome variables to one year after NBER WP publication. However, for some of the outcomes, one year is a short time frame for knowledge to diffuse. This subsection explores this temporal dimension. To do so, Model 4 is estimated for a flow variable of each of the four diffusion outcomes, that is, the additional

Table 6: IV and OLS regressions results - intensive margin

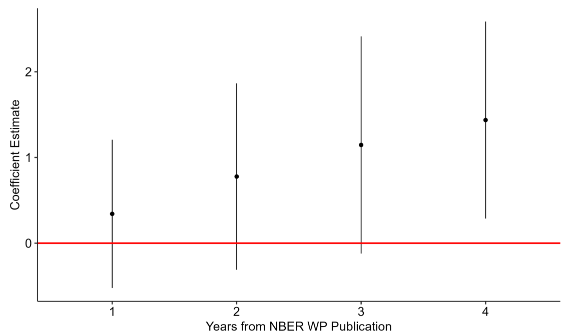
	Dependent variable:							
	Blog posts		News stories		Policy documents		Academic citations	
	instrumental variable	*OLS*	*instrumental variable*	*OLS*	*instrumental variable*	*OLS*	*instrumental variable*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Twitter Visibility (log)	0.641***	0.361***	2.500***	0.477***	0.049	0.057***	0.991	1.310***
	(0.149)	(0.030)	(0.941)	(0.083)	(0.056)	(0.013)	(1.100)	(0.200)
Intercept	−1.362***	−0.574***	−6.490**	−0.802***	−0.071	−0.094***	0.336	−0.560
	(0.418)	(0.073)	(2.586)	(0.179)	(0.155)	(0.032)	(3.078)	(0.503)
Effective F-stat (1st Stage)	27.25		27.25		27.25		27.25	
Wu-Hausman (p-value)	0.06		0		0.89		0.76	
Sargan (p-value)	0.06		0.1		0.24		0.02	
Mean outcome variable	0.44	0.44	0.54	0.54	0.07	0.07	3.12	3.12
Number of clusters	2822		2822		2822		2822	
Observations	2,916	2,916	2,916	2,916	2,916	2,916	2,916	2,916

Note: Standard errors are clustered at the time-level (hour). *p<0.1; **p<0.05; ***p<0.01. Results for the first year after the WP publication date. Sample: NBER WP 2015-2018.

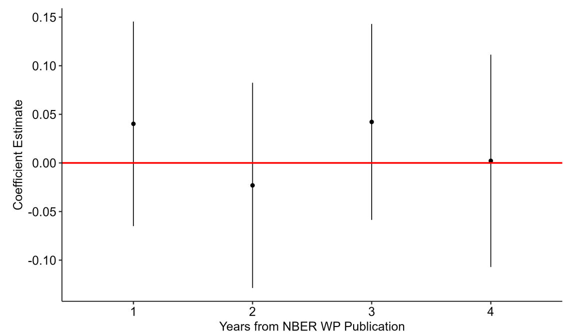
mentions in a given year since WP publication. The results are presented in figure 6 for up to five years after WP publication.

As reported in Subsection 4.1 a positive effect of social media visibility on diffusion of economic knowledge into public discourse is observed, while no effect is found for academic and policy citations. However, when it comes to academic citations, 6(a) shows that the choice of a one-year window seems to be behind the null effect, as it takes time for academic citations to accumulate. Figure 6(a) reports an increasing, with time, role of early social media visibility on the number of academic citations. By the fourth year after NBER WP publication, an increase of one standard deviation of Twitter visibility leads to an increase of 1.44 academic citations; this result is statistically significant at the 5% level.

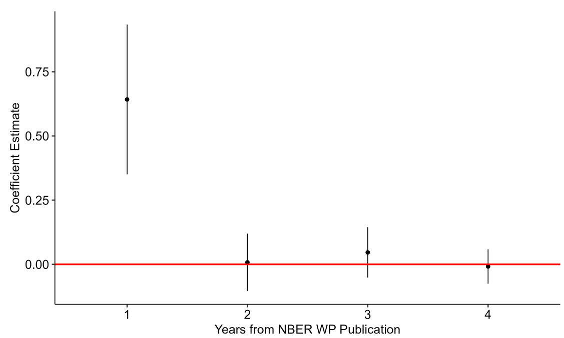
Unlike the result for academic citations, no effect was found on policy documents (see Figure 6(b)). The results for mentions in blog posts and mainstream media indicate that the effect of social media is concentrated in the first year since the publication of a WP.



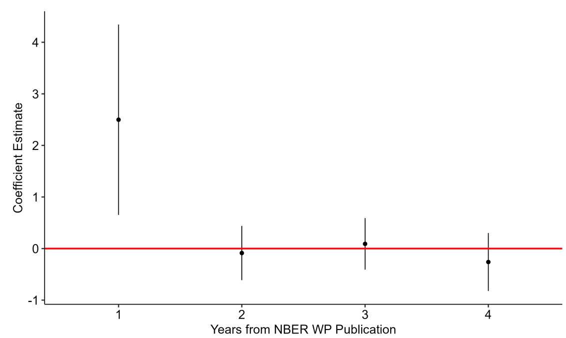
(a) Academic Citations



(b) Policy Document Mentions



(c) Blog Post Mentions



(d) Mainstream Media Mentions

Figure 6: Results by year

4.3 The Effect on Knowledge Production

Although the bulk of the research is centered around the diffusion of economic research, this subsection explores the role social media play in the publication outcomes of WP. It does so by estimating the effect of social media visibility on the likelihood of getting published in any peer-reviewed journals, in one of the top five economics journals, as well as the number of revisions made to the WP. The results are presented in Table 7. All outcomes are measured using a four-year post WP publication window.

For all three outcomes, the OLS and IV results diverge. For the publication outcomes, the OLS estimations presented in columns (4) and (6) indicate for a positive correlation between Twitter visibility in the week of @nberpubs tweet and the likelihood the publish, with a stronger correlation for the top five journals. However, the results of the IV estimation, in columns (3) and (5) indicate a null effect, suggesting that the OLS model might be upward biased (although the Wu-Hausman test does not reject the null hypothesis). Here one should remind the reader that the sample used for publication outcomes of WP is slightly smaller than the one used for all other estimation, as for some papers, no publication outcome was found. One implication is a reduction of the effective F-statistic, yet it is still well above the 10 threshold. Lastly, estimating the effect on the number of revisions reveals that an increase of one standard deviation of Twitter visibility leads to 0.32 revisions of a NBER WP.

Table 7: IV and OLS regressions results

	Dependent variable:					
	# WP revsions		Published		Top 5	
	instrumental variable	*OLS*	*instrumental variable*	*OLS*	*instrumental variable*	*OLS*
	(1)	(2)	(3)	(4)	(5)	(6)
Twitter Visibility (log)	0.323**	0.037	−0.063	0.020*	0.013	0.023***
	(0.136)	(0.136)	(0.077)	(0.011)	(0.050)	(0.007)
Intercept	−0.474	0.330	0.677***	0.441***	0.076	0.048***
	(0.382)	(0.382)	(0.218)	(0.032)	(0.141)	(0.019)
Effective F-stat (1st Stage)	27.25		23.07		23.07	
Wu-Hausman (p-value)	0.03		0.27		0.84	
Sargan (p-value)	0.64		0.69		0.52	
Mean outcome variable	0.43	0.43	0.5	0.5	0.11	0.11
Number of clusters	2822		2572		2572	
Observations	2,916	2,916	2,656	2,656	2,656	2,656

Note: Standard errors are clustered at the time-level (hour). *p<0.1; **p<0.05; ***p<0.01. Results for 4 years past WP publication date. Sample: NBER WP 2015-2018.

5 Concluding Discussion

Previous literature dealing with the diffusion of scientific knowledge centered around STEM fields and focused on its relevant diffusion channels, mechanisms, and diffusion domains. This paper contributes to existing knowledge in multiple ways. First, it investigates the diffusion of economic knowledge, departing from the STEM-centric stream of literature. Second, it explores the role of an increasingly central scientific communication channel, social media platforms. Third, to overcome endogeneity concerns and estimate causal effects, the paper, using a dataset of NBER WPs, applies instrumental variable identification strategy.

The results of this paper uncover a heterogeneous role that social media play in the diffusion of economic research. First, increased visibility of NBER WPs on social media leads to increased likelihood and extent to which NBER WPs are mentioned in public discourse (measured by blog posts and mainstream media mentions). This effect occurs within a year after the publication of an NBER WP. Second, social media visibility of WP also increases the number of academic citations, however, a longer time lag is reported, of more than 3 years. Third, no effect was found on diffusion into the domain of policy-making. Fourth, the papers that were more visible on social media did not show an increased propensity to get published in a peer reviewed journal.

The results of this paper demonstrate the role that social media play in the diffusion of economic knowledge. Social media facilitates fast dissemination of knowledge that informs the public discourse with frontier insights based on economic research. Working papers with increased social media visibility benefit from an increased number of citations, although this is only observed with a longer delay. However, the absence of an effect on diffusion to policy documents highlights the limitations of social media's role in disseminating knowledge, particularly in a critical domain aimed at maximizing societal returns from investment in economic research.

Additional analysis, exploring sub-field, authorship teams (composition in terms of: gender, seniority, and institutional prestige), will be conducted and added to this manuscript.

References

Aghion, P. and P. Howitt (1992). A model of growth through creative destruction. *Econometrica 60*(2), 323–351.

Andrews, M. J. (2023, May). How do institutions of higher education affect local invention? evidence from the establishment of us colleges. *American Economic Journal: Economic Policy 15*(2), 1–41.

Aridor, G., R. Jiménez-Durán, R. Levy, and L. Song (2024). The economics of social media.

Arora, A., S. Belenzon, and L. Sheer (2021, March). Knowledge spillovers and corporate investment in scientific research. *American Economic Review 111*(3), 871–98.

Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies 29*(3), 155–173.

Baumann, A. and K. Wohlrabe (2020). Where have all the working papers gone? evidence from four major economics working paper series. *Scientometrics 124*(3), 2433–2441.

Belenzon, S. and M. Schankerman (2013, 07). Spreading the Word: Geography, Policy, and Knowledge Spillovers. *The Review of Economics and Statistics 95*(3), 884–903.

Biasi, B. and S. Ma (2022, March). The education-innovation gap. Working Paper 29853.

Bikard, M. and M. Marx (2020). Bridging academia and industry: How geographic hubs connect university science and corporate technology. *Management Science 66*(8), 3425–3443.

Blinder, A. S. and A. B. Krueger (2004, September). What does the public know about economic policy, and how does it know it? Working Paper 10787, National Bureau of Economic Research.

Cantoni, D. and N. Yuchtman (2014, 03). Medieval Universities, Legal Institutions, and the Commercial Revolution *. *The Quarterly Journal of Economics 129*(2), 823–887.

Carrell, S. E., D. N. Figlio, and L. R. Lusher (2022). Clubs and networks in economics reviewing. Technical report, National Bureau of Economic Research.

Chan, H. F., A. S. Önder, S. Schweitzer, and B. Torgler (2023). Twitter and citations. *Economics Letters 231*, 111270.

Chugh, R., R. Grose, and S. A. Macht (2021). Social media usage by higher education academics: A scoping review of the literature. *Education and Information Technologies 26*, 983–999.

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale university press.

Darling, E. S., D. Shiffman, I. M. Côté, and J. A. Drew (2013). The role of twitter in the life cycle of a scientific publication.

Durante, R. and E. Zhuravskaya (2018). Attack when the world is not watching? us news and the israeli-palestinian conflict. *Journal of Political Economy 126*(3), 1085–1133.

Eisensee, T. and D. Strömberg (2007, 05). News Droughts, News Floods, and U. S. Disaster Relief*. *The Quarterly Journal of Economics 122*(2), 693–728.

Fang, Z., R. Costas, W. Tian, X. Wang, and P. Wouters (2021). How is science clicked on twitter? click metrics for bitly short links to scientific publications. *Journal of the Association for Information Science and Technology 72*(7), 918–932.

Feenberg, D., I. Ganguli, P. Gaulé, and J. Gruber (2017, 03). It's Good to Be First: Order Bias in Reading and Citing NBER Working Papers. *The Review of Economics and Statistics 99*(1), 32–39.

Foerderer, J. and S. W. Schuetz (2022). Data breach announcements and stock market reactions: A matter of timing? *Management Science 68*(10), 7298–7322.

Garz, M. and J. Sörensen (2017). Politicians under investigation: The news media's effect on the likelihood of resignation. *Journal of Public Economics 153*, 82–91.

Giusta, M. D., S. Jaworska, and D. V. Greetham (2021). Expert communication on twitter: Comparing economists' and scientists' social networks, topics and communicative styles. *Public Understanding of Science 30*(1), 75–90. PMID: 32933450.

Hausman, N. (2022, 07). University Innovation and Local Economic Growth. *The Review of Economics and Statistics 104*(4), 718–735.

Heckman, J. J. and S. Moktan (2020, June). Publishing and promotion in economics: The tyranny of the top five. *Journal of Economic Literature 58*(2), 419–70.

Henderson, R., A. B. Jaffe, and M. Trajtenberg (1998, 02). Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965–1988. *The Review of Economics and Statistics 80*(1), 119–127.

Hill, R. and C. Stein (2019). Scooped! estimating rewards for priority in science. *Journal of Political Economy*. Forthcoming.

Howoldt, D., H. Kroll, and P. Neuhäusler (2023). Relating popularity on twitter and linkedin to bibliometric indicators of visibility and interconnectedness: an analysis of 8512 applied researchers in germany. *Scientometrics*, 1–24.

Jaffe, A. and M. Trajtenberg (1996). Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. *Proceedings of the National Academy of Sciences 93*(23), 12671–12677.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993, 08). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations*. *The Quarterly Journal of Economics 108*(3), 577–598.

James H Stock, J. H. W. and M. Yogo (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics 20*(4), 518–529.

Jensen, R. and M. Thursby (2001, March). Proofs and prototypes for sale: The licensing of university inventions. *American Economic Review 91*(1), 240–259.

Jetter, M. (2017). The effect of media attention on terrorism. *Journal of Public Economics 153*, 32–48.

Klar, S., Y. Krupnikov, J. B. Ryan, K. Searles, and Y. Shmargad (2020, 04). Using social media to promote academic research: Identifying the benefits of twitter for sharing academic work. *PLOS ONE 15*(4), 1–15.

Klebel, T., S. Reichmann, J. Polka, G. McDowell, N. Penfold, S. Hindle, and T. Ross-Hellauer (2020, 10). Peer review and preprint policies are unclear at most major journals. *PLOS ONE 15*(10), 1–19.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017, 03). Technological Innovation, Resource Allocation, and Growth*. *The Quarterly Journal of Economics 132*(2), 665–712.

Lal, A., M. Lockhart, Y. Xu, and Z. Zu (2023). How much should we trust instrumental variable estimates in political science? practical advice based on over 60 replicated studies.

Lerner, J., H. J. Manley, C. Stein, and H. L. Williams (2024, January). The wandering scholars: Understanding the heterogeneity of university commercialization. Working Paper

32069, National Bureau of Economic Research.
Manela, A. (2014). The value of diffusing information. *Journal of Financial Economics 111*(1), 181–199.
Mohammadi, E., M. Thelwall, M. Kwasny, and K. L. Holmes (2018, 05). Academic information on twitter: A user survey. *PLOS ONE 13*(5), 1–18.
Mowery, D. C. and A. A. Ziedonis (2002). Academic patent quality and quantity before and after the bayh–dole act in the united states. *Research Policy 31*(3), 399–418.
Olea, J. L. M. and C. Pflueger (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics 31*(3), 358–369.
Pavitt, K. (2013). Academic research, technical change and government policy. *Companion Encyclopedia of Science in the Twentieth Century 143.*
Peress, J. and D. Schmidt (2020). Glued to the tv: Distracted noise traders and stock market liquidity. *The Journal of Finance 75*(2), 1083–1133.
Perkmann, M., R. Salandra, V. Tartari, M. McKelvey, and A. Hughes (2021). Academic engagement: A review of the literature 2011-2019. *Research Policy 50*(1), 104114.
Perkmann, M., V. Tartari, M. McKelvey, E. Autio, A. Broström, P. D'Este, R. Fini, A. Geuna, R. Grimaldi, A. Hughes, S. Krabel, M. Kitson, P. Llerena, F. Lissoni, A. Salter, and M. Sobrero (2013). Academic engagement and commercialisation: A review of the literature on university–industry relations. *Research Policy 42*(2), 423–442.
Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy 98*(5, Part 2), S71–S102.
Rowlands, I., D. Nicholas, B. Russell, N. Canty, and A. Watkinson (2011). Social media use in the research workflow. *Learned Publishing 24*(3), 183–195.
Sampat, B. N. (2006). Patenting and us academic research in the 20th century: The world before and after bayh-dole. *Research Policy 35*(6), 772–789. Property and the pursuit of knowledge: IPR issues affecting scientific research.
Shuai, X., A. Pepe, and J. Bollen (2012, 11). How the scientific community reacts to newly submitted preprints: Article downloads, twitter mentions, and citations. *PLOS ONE 7*(11), 1–8.
Smith, A. (1776). *The wealth of nations [1776].* ProQuest Ebook Central.
Solow, R. M. (1956, 02). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics 70*(1), 65–94.
Sorenson, O. and L. Fleming (2004). Science and the diffusion of knowledge. *Research Policy 33*(10), 1615–1634.
Spasojevic, N., Z. Li, A. Rao, and P. Bhattacharyya (2015). When-to-post on social networks. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, New York, NY, USA, pp. 2127–2136. Association for Computing Machinery.
Sugimoto, C. R., S. Work, V. Larivière, and S. Haustein (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology 68*(9), 2037–2062.
Tonia, T., H. Van Oyen, A. Berger, C. Schindler, and N. Künzli (2016). If i tweet will you cite? the effect of social media exposure of articles on downloads and citations. *International Journal of Public Health 61*(4), 513–520.
Tonia, T., H. Van Oyen, A. Berger, C. Schindler, and N. Künzli (2020). If i tweet will

you cite later? follow-up on the effect of social media exposure on article downloads and citations. *International Journal of Public Health 65*(9), 1797–1802.

Uzawa, H. (1965). Optimum Technical Change in An Aggregative Model of Economic Growth. *International Economic Review 6*(1), 18–31.

Yin, Y., Y. Dong, K. Wang, D. Wang, and B. F. Jones (2022). Public use and public funding of science. *Nature human behaviour 6*(10), 1344–1350.

Appendix A: Supplementary Materials

Table 8: NBER working papers, by JEL code, 2015-2018

JEL class	Category	# of papers	% of papers
A	General Economics and Teaching	24	0.81%
B	History of Economic Thought Methodology, and Heterodox Approaches	21	0.71%
C	Mathematical and Quantitative Methods	287	9.73%
D	Microeconomics	704	23.86%
E	Macroeconomics and Monetary Economics	653	22.14%
F	International Economics	421	14.27%
G	Financial Economics	543	18.41%
H	Public Economics	526	17.83%
I	Health, Education, and Welfare	708	24%
J	Labor and Demographic Economics	730	24.75%
K	Law and Economics	157	5.32%
L	Industrial Organization	329	11.15%
M	Business Administration and Business Economics • Marketing • Accounting • Personnel Economics	97	3.29%
N	Economic History	200	6.78%
O	Economic Development, Innovation, Technological Change, and Growth	507	17.19%
P	Political Economy and Comparative Economic Systems	69	2.34%
Q	Agricultural and Natural Resource Economics • Environmental and Ecological Economics	215	7.29%
R	Urban, Rural, Regional, Real Estate, and Transportation Economics	133	4.51%
Y	Miscellaneous Categories	2	0.07%
Z	Other Special Topics	56	1.9%

Note: Each paper could be assigned more than one JEL code. Based on data made public by the NBER.

Table 9: NBER working papers, by programs and program categories, 2015-2018

Program name	Program category	# of papers	% of papers
Asset Pricing	Finance	314	10.64%
Corporate Finance	Finance	292	9.9%
Economic Fluctuations and Growth	Macro/International	607	20.58%
Monetary Economics	Macro/International	327	11.08%
International Finance and Macroeconomics	Macro/International	314	10.64%
International Trade and Investment	Macro/International	212	7.19%
Labor Studies	Micro	699	23.69%
Public Economics	Micro	692	23.46%
Development Economics	Micro	378	12.81%
Health Economics	Micro	356	12.07%
Productivity, Innovation, and Entrepreneurship	Micro	338	11.46%
Health Care	Micro	284	9.63%
Industrial Organization	Micro	276	9.36%
Economics of Education	Micro	262	8.88%
Political Economics	Micro	259	8.78%
Children	Micro	242	8.2%
Environment and Energy Economics	Micro	225	7.63%
Economics of Aging	Micro	210	7.12%
Development of the American Economy	Micro	182	6.17%
Law and Economics	Micro	173	5.86%
Technical Working Papers	NA	50	1.69%

Note: Each paper could be assigned more than one program. Based on a public dataset, matching NBER WP to programs and program category (see Ben Davis's 'nberwp' repository, on github.)

Table 10: Top 20 Blog mentions‘ authors, 2015-2018

Blog author	n	%
National Affairs Online	221	12.29%
Himaginary	149	8.29%
The Brookings Institution	117	6.51%
Marginal Revolution	94	5.23%
Brad DeLong	79	4.39%
Real Time Economics	69	3.84%
Forum:Blog	46	2.56%
Nada es Gratis	40	2.22%
Noozilla Top	38	2.11%
Journalist's Resource	34	1.89%
Naked Capitalism	32	1.78%
World Bank Blogs	31	1.72%
Economist's View	25	1.39%
The Incidental Economist	25	1.39%
CityLab	19	1.06%
Liberty Street Economics	18	1%
Common Dreams	16	0.89%
Wonkblog	15	0.83%
Antitrust & Competition Policy Blog	14	0.78%
Freakonomics	14	0.78%
Total	1798	

Table 11: Top 20 policy mentions‘ authors, 2015-2018

Policy author	n	%
Urban Institute	112	14.74%
World Bank	74	9.74%
OECD	65	8.55%
Brookings Institute	45	5.92%
Center on Budget and Policy Priorities	45	5.92%
American Action Forum	35	4.61%
Economic Policy Institute	28	3.68%
The Heritage Foundation	24	3.16%
The Publications Office of the EU	20	2.63%
CPB Economic Policy Analysis	18	2.37%
UK Government	18	2.37%
Congressional Research Service	17	2.24%
The Inter-American Development Bank	14	1.84%
rijksoverheid.nl	14	1.84%
Center for Strategic and International Studies	13	1.71%
Food and Agriculture Organization of the UN	13	1.71%
World Health Organization	13	1.71%
Asian Development Bank	10	1.32%
Institute for Public Policy Research (IPPR)	9	1.18%
National Academies Press	9	1.18%
Total	760	

Table 12: Top 20 news mentions‘ authors, 2015-2018

Media author	n	%
Bloomberg	93	2.89%
Forbes	83	2.58%
MSN	74	2.3%
Yahoo! Finance USA	65	2.02%
New York Times	64	1.99%
Yahoo! News	58	1.8%
Brookings	57	1.77%
The Conversation	54	1.68%
Medium US	41	1.27%
Quartz	41	1.27%
Vox.com	37	1.15%
Phys.org	32	0.99%
Business Insider	29	0.9%
BullFax	28	0.87%
The Hill	28	0.87%
Medical Health News	27	0.84%
World Economic Forum	27	0.84%
The Guardian	24	0.75%
Foreign Affairs New Zealand	20	0.62%
Global Advisors	20	0.62%
Total	3220	NA

Note:)

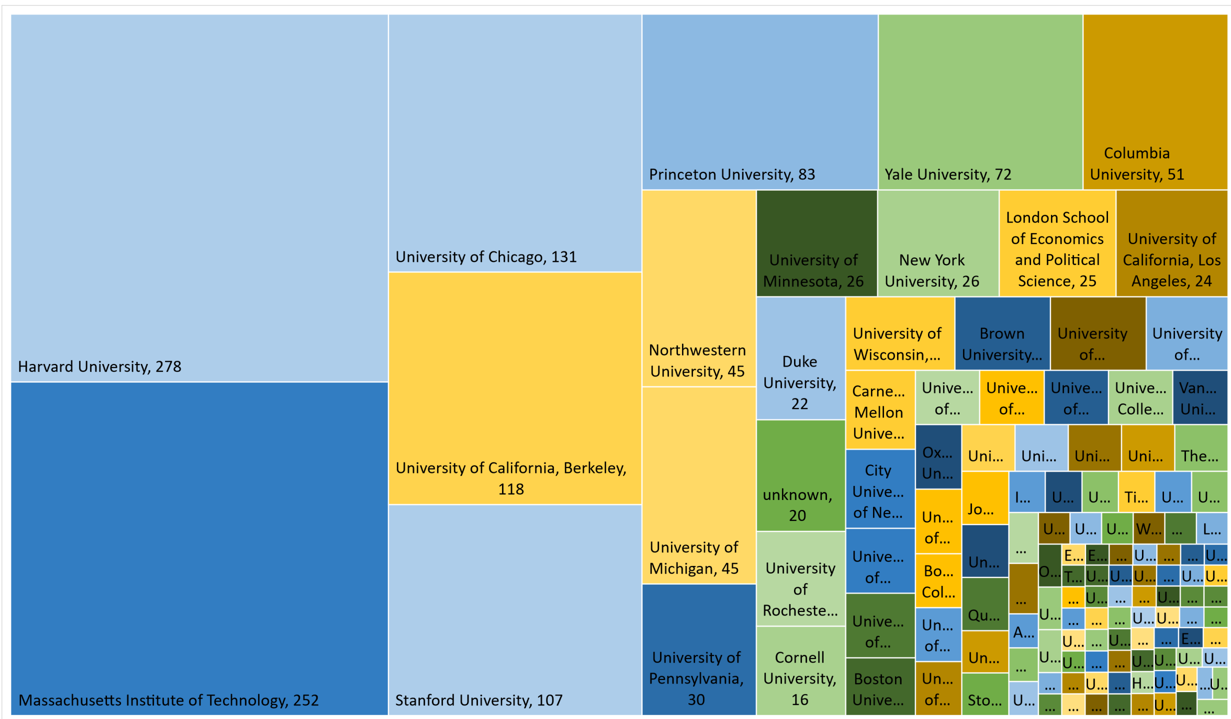


Figure 7: NBER affiliates: Ph.D.-awarding university
Note: Data for NEBR affiliates in December 2023 (1708 affiliates).

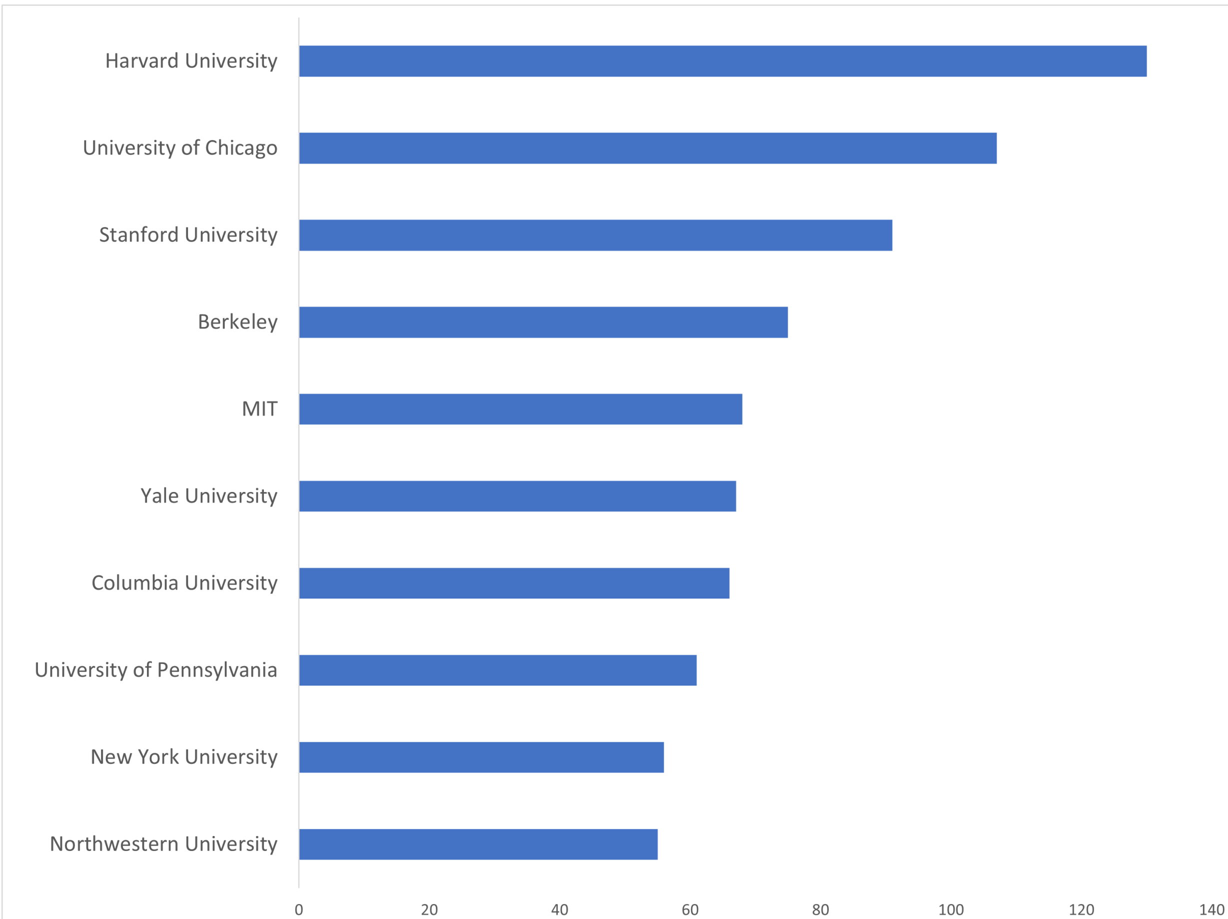


Figure 8: NBER affiliates: primary university affiliation 2023
Note: Data for NEBR affiliates in December 2023 (1708 affiliates).

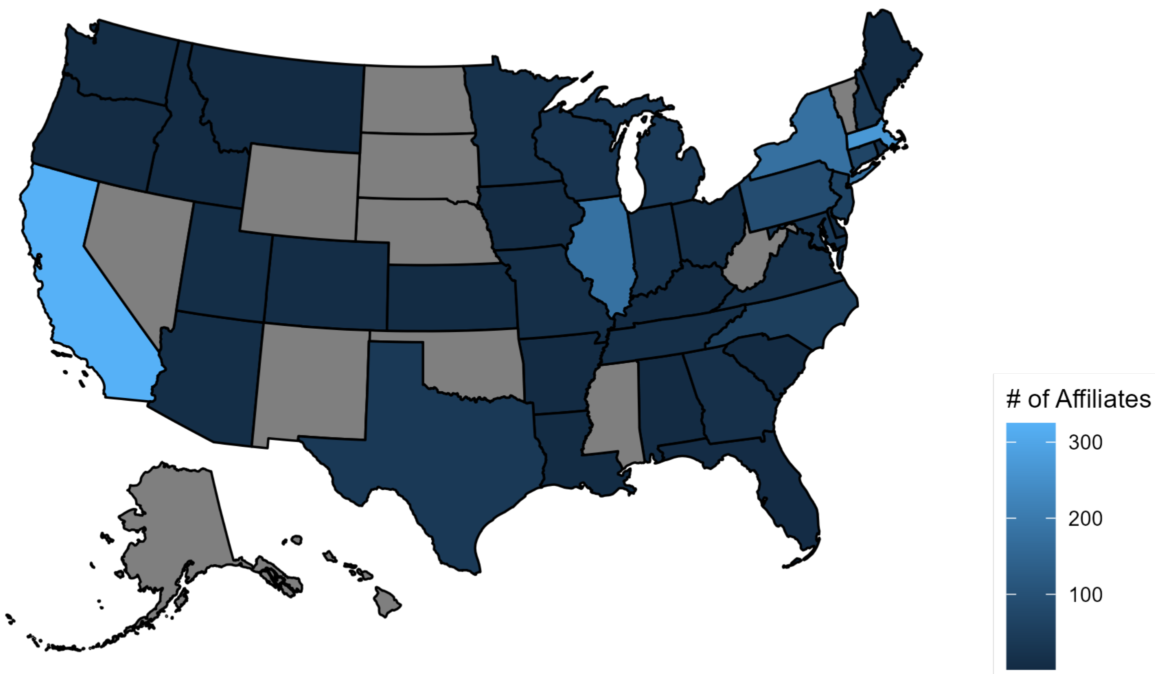


Figure 9: NBER affiliates: geographical distribution

Note: Data for NEBR affiliates in December 2023 (1708 affiliates). 5% of affiliates are based in a non-US university.