ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Wohlrabe, Klaus; Bürgi, Constantin

Working Paper

What Is the Benefit from Publishing a Working Paper in a Journal in Terms of Citations? Evidence from Economics

CESifo Working Paper, No. 8925

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Wohlrabe, Klaus; Bürgi, Constantin (2021) : What Is the Benefit from Publishing a Working Paper in a Journal in Terms of Citations? Evidence from Economics, CESifo Working Paper, No. 8925, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/235295

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

CESifo WORKING PAPERS

8925
2021

March 2021

What Is the Benefit from Publishing a Working Paper in a Journal in Terms of Citations? Evidence from Economics

Klaus Wohlrabe, Constantin Bürgi

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8925*

What Is the Benefit from Publishing a Working Paper in a Journal in Terms of Citations? Evidence from Economics

Abstract

Many papers in economics that are published in peer reviewed journals are initially released in widely circulated working paper series. This raises the question about the benefit of publishing in a peer-reviewed journal in terms of citations. Specifically, we address the question: To what extent does the stamp of approval obtained by publishing in a peer-reviewed journal lead to more subsequent citations for papers that are already available in working paper series? Our data set comprises about 28.000 working papers from four major working paper series in economics. Using panel data methods, we show that the publication in a peer reviewed journal results in around twice the number of yearly citations relative to working papers that never get published in a journal. Our results hold in several robustness checks.

JEL-Codes: A100, A120.

Keywords: peer review, citations, RePEc, working paper, preprint.

Klaus Wohlrabe
ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5
Germany – 81679 Munich
wohlrabe@ifo.de

Constantin Bürgi
St. Mary's College of Maryland
47645 College Drive
USA – St. Mary's City, MD, 20686
crburgi@smcm.edu

February 26, 2021

1 Introduction

One of the most important metrics to assess the importance of a research article, its impact or value is its citation count. The citation count across articles is in turn used in rankings of academic journals, authors, departments, universities or as an input for tenure and promotion decisions (e.g. Segalla (2008) or Seiler and Wohlrabe (2012)). Due to this, the expected citation count is one of the factors that determines researchers' decisions of how to disseminate their research, be it as a published article or not.[1]

Traditionally, publishing an article in a peer-reviewed journal has been an effective way to get one's work cited. For example, it increases the outreach of the article, makes it available online (for a fee or free as open access), and it provides a stamp of approval by peers. However, many of these benefits can readily be obtained by omitting the peer review process and publishing a paper in a working paper series or as a preprint. Indeed, there are many working papers that received a substantial number of citations even if they never successfully completed the peer review process to become journal articles.

In addition, the literature has identified several limitations of the peer review process that do not apply to working papers. Bornmann (2011) provides an excellent description. For example, Laband and Piette (1994), Hodgson and Rothman (1999) and Ductor and Visser (2020) suggest that the editor and his or her relationship to the authors can have a large impact on the chances of a paper getting published in a high tier journal. In addition, working papers are available without delay while the time from submission to publication in peer-reviewed journals can take up to multiple years and has increased over time (e.g. Ellison (2002)). The almost immediate availability of working papers allows a timelier discourse of the findings and potentially more citations.[2]

While there are some shortcomings in the peer review process that is required to publish in a journal, this process also has some key advantages that can impact the number of citations

[1] See Spiewanowski and Talavera (2021) for a recent study of publication behaviour for UK-based economists.

[2] Brown and Zimmermann (2017) provide a detailed discussion on this issue. Sarabipour et al. (2019) outline the value of preprints for early stage researchers.

a paper gets. As there is no formal peer review before preprints or working papers become available online, they can potentially include errors, key omissions, flaws or caveats. The peer review process thus acts as a quality control of research papers and can foster trust in the science system. Additional potential benefits include reaching a new and potentially wider audience and an improved readability due to a more standardized formatting.

Because there are both factors which suggest that working papers should get more citations (e.g. timeliness) and factors which suggest that journal articles should get more citations (e.g. potential flaws in working papers), it is not obvious whether publishing in a journal increases the citation count. If publishing an article in a peer-reviewed journal has little or no impact on the citations a paper receives, this would put into question some of the motivations for publishing an article as well as the importance of working papers relative to published articles.

In order to measure the impact of publishing in a journal, we first denote the collection of benefits that could lead to additional citations the stamp of approval effect from publishing in a peer-reviewed journal. This name is chosen, as at least the editor and the peer reviewers need to give their approval for publication. We then choose to measure this benefit in terms of citations by comparing working paper articles that have been published either in a peer reviewed journal or not. Our data set comprises about 28,000 papers published in four major economics working paper series. The bibliometric data builds upon the RePEc website (Research Papers in Economics, www.repec.org) and the citation data were retrieved from CitEc, which is closely related to RePEc.[3] We estimate the stamp of approval effect on the yearly citation count.

We focus on the potential increase in citations due to publishing in a journal article but there are additional benefits like prestige or tenure and promotion which can be important reasons to go through the publishing process as well. However, as citations tend to play a role for these benefits as well, our findings can have some relevance for those benefits as well.

[3] An comprehensive overview provides Zimmermann (2013). RePEc data has been used in bibliometric analysis by Rath and Wohlrabe (2016), García-Suaza et al. (2020) or Wohlrabe and Gralka (2020), among others.

Specifically, our results suggest that the stamp of approval effect leads to a doubling in the number of yearly citations for journal articles and hence it makes sense to attribute a higher value to journal publications than working papers.

This paper also complements the literature on the value of preprints and working papers. Specifically, it has been shown extensively that journal articles which are available as preprints and working papers (and open access) as well are cited more often (e.g. Sarabipour et al. (2019), Fraser et al. (2020), Fu and Hughey (2019), or Wohlrabe and Bürgi (2021)). We show here that the reverse is true also and hence the stamp of approval is important as well.

The remainder of the paper is structured as follows: the next section describes the data used followed by the panel regressions. The following section offers robustness checks, and the final section concludes.

2 Data

We build upon the data set of Baumann and Wohlrabe (2020a) who also provide more background information. The four working paper series considered stand out with respect to prestige and influence in the field of economics. They belong to the most cited and downloaded working paper series on RePEc.[4] The four series are published by networks of economists. Submitting papers to the series is only allowed to members of the corresponding networks and joining the network is only possible by invitation. Once an author is a member of the specific network he or she is free to submit any working paper. With this procedure, the networks want to assure a specific level of quality of the submitted papers as invitations are only issued to established or promising researchers.

The original data set comprises 28,877 working papers from between 2000 and 2012. For our analysis we exclude those working papers that have been published as a chapter in a book (1,120). This leaves us with 27,757 papers. The majority of papers has been published in the NBER working paper series (10,364), closely followed by CEPR (6,699) and IZA (6,904).

[4]Citations: https://ideas.repec.org/top/top.wpseries.simple.html; Downloads: https://ideas.repec.org/top/top.wpseries.download10.html

The least number of papers were issued in the CESifo series (3,790). In Table 1, we show how many papers of each series have been published in a refereed journal. In total and for each series this share is approximately 50%. These numbers differ a little bit from the estimates provided by Baumann and Wohlrabe (2020a) who report a share of about 66% based on a random sample. In the robustness section we address this issue in more detail.

Our citation data comes from CitEc (http://citec.repec.org/). This website provides the citation data for the RePEc network.[5] Each working paper and journal article has a unique identifier. We used this identifier to extract citations on a yearly basis for working papers and the journal articles separately. However, as the working paper citations contain the journal citations and vice versa, it is not possible to obtain accurate citation counts for the two versions separately. Due to this restriction, we use the maximum citations per year from either version as the citation count for our dependent variable.[6] As Baumann and Wohlrabe (2020a) have shown, many papers have been published in several working paper series simultaneously. CitEc also consolidates citations across versions, i.e. it assigns citations received by one working paper version also to the other versions. Due to time delays in the consolidation process at CitEc, the citation numbers are not always identical across series. We therefore take always the maximum citation count across working papers.

In Table 1 we provide the descriptive statistics for the citation data. Besides the full sample, we also distinguish whether an article has been published in a journal or not. The average citation count across all working papers is 43. This number is higher for papers in the NBER series (61) and smaller for the IZA (25) and CESifo series (21). The citation distribution is quite dispersed as the standard deviation exceeds the mean considerably. This is also confirmed in the left graph of Figure 1. The most cited paper published in a journal is the one by Melitz (2003) which appeared both in the NBER and the CEPR working paper series. The article by Pesaran (2004) is the most cited paper (904) that never

[5] We thank Jose Manuel Barrueco for help with the citation data. The citations were retrieved in February 2020.

[6] Instead of the maximum, we also ran the estimations the minimum and our results remained qualitatively the same as shown in Table 6.

appeared in a journal. It was published both in the IZA and the CESifo working paper series. When comparing articles in our two subgroups, we see a clear difference between the citation counts. Articles published in a journal received 70 citations on average whereas the average citation counts for unpublished papers only amounts to 15. In the lower panel of Table 1, we report the p-value of a two-sided non-parametric Wilcoxon-Mann-Whitney test. The null hypothesis of an equal median of citations is clearly rejected both for the full sample and the four series.[7] The results reported in Table 1 are supported by the kernel estimates of the citation distribution depicted in the left graph of Figure 1.[8] They show that most of the papers with 10 or fewer citations have not been published in a journal. Additionally, journal articles in our sample exhibit more mass compared to the control group across almost the entire citation distribution.

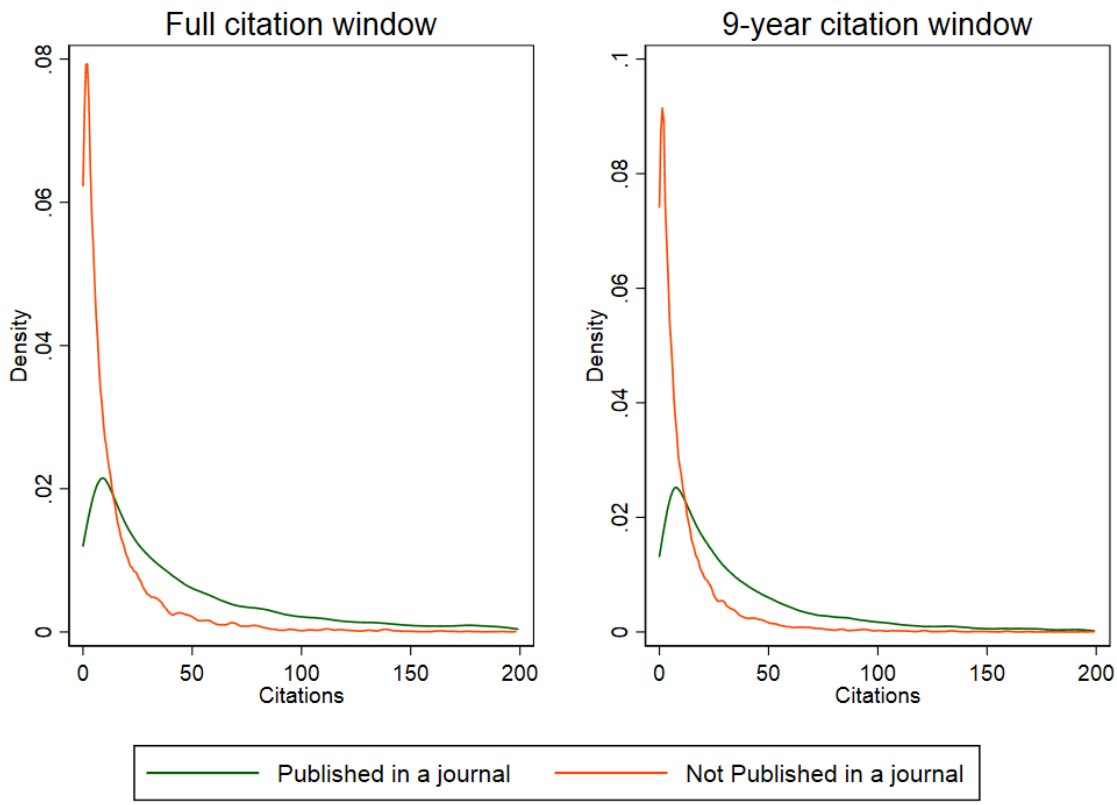
[7] We obtain the same results in case of two-sided t-test.

[8] In order to increase readability, we capped the citations at 200.

Table 1: Descriptive statistics for citation count

	Full citation window						9-year citation window				
	N	Mean	Standard deviation	Minimum	Maximum	t-Test p-value	Mean	Standard deviation	Minimum	Maximum	Wilcoxon-Mann-Whitney test p-value
Total	27,757	43	109	0	5,150		29	54	0	1,693	
CESifo	3,790	21	58	0	1,296		15	30	0	462	
NBER	10,364	61	139	0	5,150		39	67	0	1,693	
CEPR	6,699	46	114	0	5,150		30	54	0	1,198	
IZA	6,904	25	56	0	1,368		19	36	0	889	
Published in a journal											
Total	14,212	70	144	0	5,150		44	69	0	1,693	
CESifo	1,896	34	76	0	1,296		24	39	0	462	
NBER	5,542	96	179	0	5,150		58	84	0	1,693	
CEPR	3,510	74	149	0	5,150		46	67	0	1,198	
IZA	3,264	42	74	0	1,368		31	48	0	889	
Not published in a journal											
Total	13,545	15	33	0	904	0.000	12	21	0	330	0.000
CESifo	1,894	8	25	0	904	0.000	7	13	0	215	0.000
NBER	4,822	21	40	0	861	0.000	16	26	0	330	0.000
CEPR	3,189	15	32	0	861	0.000	12	20	0	313	0.000
IZA	3,640	11	24	0	904	0.000	9	15	0	330	0.000

Figure 1: Kernel estimates for citations distributions



The comparisons so far might be biased as the publication date of working papers ranges from 2000 to 2012. So, more recent papers had less time to accumulate citations. Therefore, we repeat the analysis using a constant citation window of nine years starting in the publication year of the working paper. As Table 1 and Figure 1 show that the results stay qualitatively the same, papers that are eventually published in a journal gather more citations. As expected, the citation counts are smaller due to the smaller citation window.

These simple mean comparisons leave out many of the factors which might drive citation counts of papers and do not necessarily reflect a precise estimate. See Bornmann and Daniel (2008) and Tahamtan and Bornmann (2019) for a literature overview. We address these issues in the next section.

3 Methodology and main results

3.1 Methodology

In order to test the hypothesis that a journal publication leads to more citations, we first create a dummy variable that takes a value of 0 if the paper was only available as a working paper in a specific year and 1 otherwise (i.e. if it was published in a journal). Because this dummy takes value 1 in all years for papers that become journal articles in the first year, the stamp of approval effect cannot be separated from paper fixed effects and we exclude them from our analysis. This holds for 1,878 articles. Baumann and Wohlrabe (2020b) report more on the differences between the publication dates of working papers and the corresponding journal articles.

Due to the panel nature of our data set, we omit the controls used in the literature as they are either captured by the paper fixed effects (e.g. the number of authors, the working paper series or the number of working papers) or captured by time fixed effects (e.g. age of the papers).[9] In addition, the panel structure with fixed effects also controls for many unobservable characteristics specific to a paper that do not change over time and can influence the outcomes. For example, only the better papers might eventually be published in a journal and this quality difference is captured by paper fixed effects (FE).[10] In a first step we run a pooled regression and a random effects (RE) specification. The likelihood ratio test clearly rejects the pooled model in favor of the RE model (p-value = 0.000). Similarly, the Hausman test rejects the RE model in favor of the FE model (p-value = 0.000). We then estimate the following equation as our main specification:

$$citpy_{it} = \beta journal_{it} + \nu_i + \mu_t + \varepsilon_{it} \tag{1}$$

where *citpy* are the citations per year for each article and *journal* is a dummy that takes

[9] This setup also controls for more sophisticated age structures as in Anauati et al. (2016) and is broadly in line with the structure in Fraser et al. (2020) or Fu and Hughey (2019).

[10] Obviously, unobservable characteristics that change over time are not captured by the paper and time fixed effect like for example a quality change between the working paper and the published paper. We address this issue in Section 3.3.

value one for the years in which the journal article is available. ν_i are the paper fixed effects and μ_t the time fixed effects. These include also dummy variables representing each year in our data sets. This captures general citations trends. Negative binomial panel regressions with fixed effects would be natural choice for count data. However, these are not feasible (e.g. see Wooldridge (1999)). Due to this limitation, we use Poisson regressions with both time and paper fixed effects instead. We run the regression both across all papers and for each working paper series individually and cluster the standard errors at the paper level. We report incident rate ratios (the exponential of the coefficients) and the effect is multiplicative. This means for example that an incident rate ratio of 2 implies a doubling of the number of yearly citations and 1 means no change in citations.

3.2 Main results

Table 2 reports the estimates both for the full sample and for each working paper series separately. The coefficient of the journal dummy is large and highly significant across all regressions. The estimates imply that the stamp of approval effect of publishing an article in a peer-reviewed journal results in more than a doubling of the yearly citation count relative to only making the article available as a working paper. This effect is not only statistically significant but also substantial in its size. Looking at the working paper series, the journal effect is larger for the CESifo and IZA working papers. For these, the publication in a journal leads to an increase in citations by a factor of around 2.5 while it is a touch smaller for NBER and CEPR working papers with an increase in citations by a factor of 2.3.

3.3 Initial robustness checks

A working paper might appear in multiple working paper series. This might cause the paper to reach a wider audience before becoming a journal article and hence the benefit from publishing it becomes smaller. To test this, we restrict our sample to papers that have been released only in one working paper series. The result in column 1 of Table 3 show that papers that are only in one working paper series indeed benefit more from being published. We can

Table 2: Publication impact using panel regression

	Dependent variable:				
	Citations				
	(1)	(2)	(3)	(4)	(5)
	Total	CESifo	NBER	CEPR	IZA
Years in Journal	2.382***	2.579***	2.392***	2.326***	2.442***
	(0.039)	(0.106)	(0.065)	(0.073)	(0.066)
Year FE	Yes	Yes	Yes	Yes	Yes
Paper FE	Yes	Yes	Yes	Yes	Yes
Observations	369,375	48,008	144,625	92,352	84,390
ID	25,879	3,374	9,977	6,242	6,286
Log pseudo-likelihood	-569,319	-53,888	-259,504	-145,724	-109398

This table shows the coefficient of interest based on panel fixed effects Poisson regressions. The reported coefficients are incident rate ratios (the exponential of the coefficients). For example, the coefficient of 2.382 in the first column implies that articles published in journals receive 2.382 times as many yearly citations as working papers. Paper-level clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

look at this argument also from a different perspective. Being released in only one working paper series and then either published in a journal or not might be an unfair comparison. The article finally published in a journal has an additional outlet and therefore a higher visibility. In order to account for that issue, we restrict our sample to articles published in a journal with only one working paper and compare them to articles available in two working paper series but not in a journal. Column 2 of Table 3 shows that the effect becomes a touch smaller but remains highly significant. In column 3 we report the outcome of a comparison between articles that appeared in two working paper series and in a journal with those that appeared in three working paper series but not in a journal. Again, publishing in a journal leads to a significantly higher citation count.

An additional concern is the issue of the time lag between working paper and journal publication. It could be the case that papers that take longer to get published in a journal

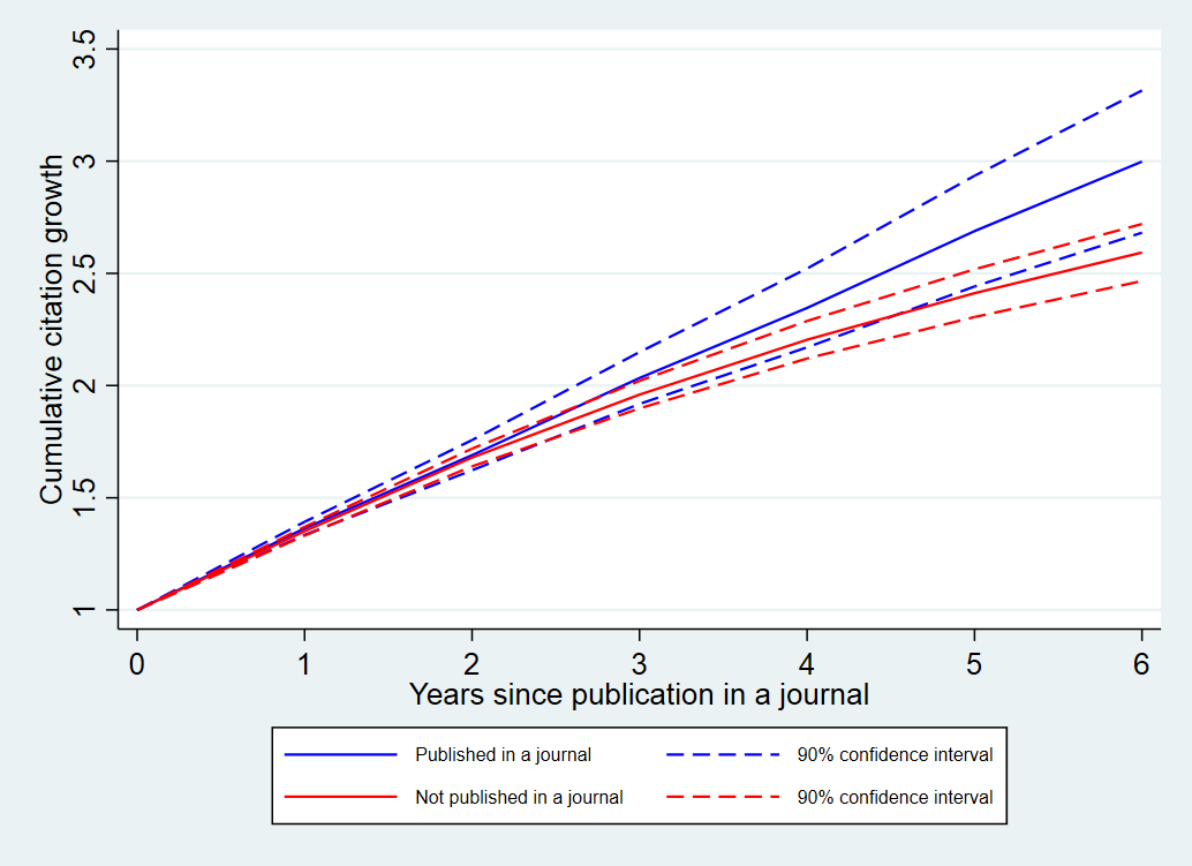
have a large change in quality. For example, they might go through more review rounds and have more substantial revisions during the submission process. As a consequence, they might receive a larger benefit from publishing. As already noted, papers with identical publication years in both outlets are excluded by construction hence we took only those papers where the time lag is one year. This limits the revisions possible and hence the potential quality change of the paper. In the fourth column of Table 3 we show that the journal effect is larger for the papers which only had limited time for quality changes. This suggests that the (unobserved) quality change is not the main driver of our results.

Table 3: Initial robustness checks

	(1) appeared in only in one WP series	(2) WP + Journal vs. 2 WP	(3) 2 WP + Journal vs. 3 WP	(4) Journal publication appeared one year later
Journal	2.944***	2.342***	2.429***	4.641***
	(0.081)	(0.060)	(0.060)	(0.141)
Observations	162,576	129,434	74,772	220,447
Number of ID	11,172	22,682	18,752	15,378
Year FE	YES	YES	YES	YES
Paper FE	YES	YES	YES	YES
Log pseudo-likelihood	-213,323	-181,548	-95,208	-278,385

This table shows panel Poisson regressions results. The first column only includes papers that appeared in only one working paper series. In the other columns we compare working papers that were later published in a journal with working paper only appeared in different number of series. All coefficients are incident rate ratios (the exponential of the coefficients). Paper-level clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Figure 2: Development for cumulative citation count for a matched sample



This figure shows the cumulative citations relative to the first year of the paper on the vertical axis for 151 matched papers that have at least 20 cumulative citations in the first three years. The blue line are papers that have been published in a journal in the third year and the red line are the matched papers that were not published. The years are on the horizontal axis.

Our next robustness check concerns a potential endogeneity issue. Specifically, one could imagine that only better papers get published and worse papers remain working papers. An editor could see that a working paper has already accumulated a substantial number of citations and might have a positive bias towards these papers. This could cause our coefficients to capture the paper quality rather than the stamp of approval effect. In order to test this, we reduce our sample to papers that cumulatively received at least 20 citations in the first three years and in addition only keep published papers that were published in the

third year of becoming available as working paper. These papers are then matched to similar working papers that have never appeared in a journal. Therefore, they have been issued in the same year and in the same working paper series. Furthermore, the cumulative citation ratio between two papers must have been between 0.9 and 1.1. Based on these conditions we were able to match 151 published papers. For each of the first seven years, we calculate the cumulative citations relative to the first year and average this number across the two groups. Figure 2 shows the resulting line for the published articles in blue and the line for the working papers in red plus the 90% confidence intervals. While both the published and unpublished papers have similar citations in the first three years, the published articles receive more citations once they get published. This suggests that there is indeed a stamp of approval effect, as the papers in both groups should have a similar quality based on the citations received in the first three years, but published papers receive more citations in later years.

4 Additional robustness checks

4.1 Accounting for non-tracked articles

So far, we assumed that we were able to find all published articles corresponding to working papers. However, Baumann and Wohlrabe (2020a) showed that the actual share of working papers in our sample which are finally published in a journal is about 66% using a random sub-sample (instead of the 50% we found). In this section, we want to make sure that our results are not driven by our potential inability to track all published papers. In order to test this, we repeat the regression in Table 2 but treat a random 30%, 20%, 10%, 5% and 1% of the published papers as if they had zero citations in every year. Specifically, we run a simulation with 1000 repetitions for each threshold where we randomly assign zero citations to the specified share of published articles. This simulation is equivalent to adding a similar percentage of the least cited unpublished working papers randomly to the published journals. Replacing 10% of the citations with zeros is equivalent to not having found around

1,400 journal publications. Replacing the citations with zeros are also a clear worst-case scenario, as it is likely that at least some of the published articles we could not find have more than zero citations.

Table 4: Robustness: Accounting for non-tracked articles using simulations

	(1)	(2)	(3)	(4)	(5)
	30%	20%	10%	5%	1%
Mean	1.761	1.970	2.176	2.279	2.361
High	1.791	2.000	2.197	2.295	2.369
Low	1.726	1.944	2.155	2.264	2.354
Largest SE	0.005	0.005	0.004	0.004	0.004
Significant	1000	1000	1000	1000	1000

This table shows the distribution of coefficients for 1000 replications and various thresholds. Aside from setting x% of published articles have their citations set equal to 0, the estimations correspond to column 1 of Table 2.

As Table 4 shows, all simulations have a positive and significant coefficient and the coefficient becomes smaller the more journal articles are changed to have 0 citations. The mean coefficient of the 1000 simulations is also reasonably close to the highest and lowest coefficient in the simulation, even if the difference is several standard errors. With 30% missed articles, the stamp of approval effect is smaller but still corresponds to a 75% increase in citations for articles that are published in a journal relative to the ones that remain working papers.

4.2 Skewness of the citation distribution

It is well-known that the citation distribution is skewed (e.g. see Seiler and Wohlrabe (2014)). Our sample includes several papers that have very large numbers of citations which could potentially drive our results. Also, there is a number of papers that have never been cited, which could also impact our results. In order to check whether the papers with many citations drive our results, we repeat the regressions in Table 2 column 1 but exclude papers that have at least 1000, 500 and 100 citations from our sample. The results are presented in the first

three columns of Table 5.

Table 5: Robustness: Accounting for skewed citation distribution

	Dependent variable:				
	Citations				
	(1)	(2)	(3)	(4)	(5)
Excluding	1000	500	100	0	0 and 100
Journal	2.430***	2.483***	2.600***	2.382***	2.601***
	(0.032)	(0.031)	(0.030)	(0.039)	(0.030)
Observations	368,423	365,518	325,915	369,375	325,543
Number of ID	25,825	25,657	23,097	25,879	23,073
Year FE	YES	YES	YES	YES	YES
Paper FE	YES	YES	YES	YES	YES
Log pseudo-likelihood	-557,282	-569,319	-398,725	-569,319	-398,725

This table shows panel Poisson regressions where in each regression articles with at least X citations are excluded. In column (4) papers with 0 citations were removed before the estimation. In column (5), both papers with at least 100 citations and papers with 0 citations were removed before the estimation. All coefficients are incident rate ratios (the exponential of the coefficients). Paper-level clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Compared with the results where we did not exclude these papers, the impact of publishing a paper are broadly unchanged or there is a larger stamp of approval effect. This result is also robust to omitting the papers with zero citations as shown in column (4) and to omitting papers with both 0 citations in addition to the ones with at least 100 citations shown in the last column in Table 5.

4.3 Further robustness checks

Finally, we consider two additional issues. First, our data has more observations for older papers than for newer ones as our sample stops in 2020 for all papers. This could potentially influence our results as this causes older papers to have a higher weight in the panel. In order to assess the impact of this, we repeat the regressions but only include the first nine years for each paper (a nine-year citation window). The results for this regression are shown in

the first column of Table 6. While the coefficient declines to a 90% increase, it is still highly significant.

Table 6: Further robustness checks

	(1) constant time window	(2) Minimum citations
Journal	1.907***	2.382***
	(0.026)	(0.039)
Number of ID	25,630	25,879
Year FE	YES	YES
Paper FE	YES	YES
Log pseudo-likelihood	-351,239	-569,319

This table shows panel Poisson regressions results. The first column only uses the first nine years of data for each paper, and the second one uses the smaller of the two citation sources (working paper and journal citations). All coefficients are incident rate ratios (the exponential of the coefficients). Paper-level clustered standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Second and as mentioned in the methodology section, we have multiple sources for citations. CitEc reports citations for the working paper and for the journal article separately, even if both citation numbers contain citations for either. So far, we used the maximum of the two numbers as the journal citations in the analysis. In column 2 of Table 6 we report the regression result for the smaller of the two numbers. The coefficient and the significance are virtually unchanged when compared to the ones obtained in Table 2.

5 Conclusion

This paper showed that the yearly citation count of working papers which are subsequently published in a journal more than doubles relative to unpublished working papers. This increase in citations is substantial, controlling for a number of observable and unobservable variables that are time or paper specific in a panel regression. There is thus a substantial

stamp of approval effect when publishing in journals.

While we were able to identify what we call the stamp of approval effect, we were not able to identify what component of publishing an article causes this effect. There are many possible channels, which could cause this increase in citations. For example, the peer review process that approves and improves the paper or the additional outreach of a journal could both be potential causes that increase the citation count and we found evidence for both of these channels. This paper also focused on citations only. There are other potential benefits beyond citations to both the researcher and the field as a whole, which we did not address here.

References

Anauati, V., Galiani, S., and Gálvez, R. H. (2016). Quantifying the life cycle of scholarly articles across fields of economic research. *Economic Inquiry*, 54(2):1339–1355.

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

Baumann, A. and Wohlrabe, K. (2020b). Where have all the working papers gone? evidence from four major economics working paper series. CESifo Working Paper 8328.

Bornmann, L. (2011). Scientific peer review. *Annual review of information science and technology*, 45(1):197–245.

Bornmann, L. and Daniel, H.-D. (2008). What do citation counts measure? a review of studies on citing behavior. *Journal of Documentation*, 64(1):45–80.

Brown, A. J. and Zimmermann, K. F. (2017). Three decades of publishing research in population economics. *Journal of Population Economics*, 30(1):11–27.

Ductor, L. and Visser, B. (2020). Why are connections to editorial board members of economics journals valuable? Technical report, Department of Economic Theory and Economic History of the University of Granada.

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

Fraser, N., Momeni, F., Mayr, P., and Peters, I. (2020). The effect of biorxiv preprints on citations and altmetrics. *Quantative Science Studies*, 1(2):618–638.

Fu, D. Y. and Hughey, J. J. (2019). Releasing a preprint is associated with more attention and citations for the peer-reviewed article. *Elife*, 8:e52646.

García-Suaza, A., Otero, J., and Winkelmann, R. (2020). Predicting early career productivity of phd economists: Does advisor-match matter? *Scientometrics*, 122(1):429–449.

Hodgson, G. M. and Rothman, H. (1999). The editors and authors of economics journals: a case of institutional oligopoly? *Economic journal*, 109(453):165–186.

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

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Pesaran, H. (2004). General diagnostic tests for cross-sectional dependence in panels. CESifo Working Paper 1229.

Rath, K. and Wohlrabe, K. (2016). Recent trends in co-authorship in economics: Evidence from RePEc. *Applied Economics Letters*, 23(12):897–902.

Sarabipour, S., Debat, H. J., Emmott, E., Burgess, S. J., Schwessinger, B., and Hensel, Z. (2019). On the value of preprints: An early career researcher perspective. *PLoS Biology*, 17(2):e3000151.

Segalla, M. (2008). Publishing in the right place or publishing the right thing: journal targeting and citations' strategies for promotion and tenure committees. *European Journal of International Management*, 2(2):122–127.

Seiler, C. and Wohlrabe, K. (2012). Ranking economists on the basis of many indicators: An alternative approach using repec data. *Journal of Informetrics*, 6(3):389–402.

Seiler, C. and Wohlrabe, K. (2014). How robust are journal rankings based on the impact factor? evidence from the economic sciences. *Journal of Informetrics*, 8(4):904–911.

Spiewanowski, P. and Talavera, O. (2021). Journal rankings and publication strategy. *Scientometrics*, Forthcoming.

Tahamtan, I. and Bornmann, L. (2019). What do citation counts measure? an updated review of studies on citations in scientific documents published between 2006 and 2018. *Scientometrics*, 121(3):1635–1684.

Wohlrabe, K. and Bürgi, C. (2021). Do working papers increase journal citations? evidence from the top 5 journals in economics. *Applied Economics Letters*, forthcoming.

Wohlrabe, K. and Gralka, S. (2020). Using archetypoid analysis to classify institutions and faculties of economics. *Scientometrics*, 123:159–179.

Wooldridge, J. M. (1999). Distribution-free estimation of some nonlinear panel data models. *Journal of Econometrics*, 90(1):77–97.

Zimmermann, C. (2013). Academic rankings with repec. *Econometrics*, 1(3):249–280.