Title Length

Yann Bramoullé[a], Lorenzo Ductor[b,*]

[a]*Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, Centre de la Vieille Charité, Marseille, France*
[b]*Middlesex University London, The Burroughs, London, United Kingdom*

Abstract

We document strong and robust negative correlations between the length of the title of an economics article and different measures of scientific quality. Analyzing all articles published between 1970 and 2011 and referenced in *EconLit*, we find that articles with shorter titles tend to be published in better journals, to be more cited and to be more innovative. These correlations hold controlling for unobserved time-invariant and observed time-varying characteristics of teams of authors.

Keywords: Title length, Journal Quality, Citations, Novelty *JEL Classification Numbers:* A10, J24

*Corresponding author at: Middlesex University London, The Burroughs, London, United Kingdom, NW4 4BT; Email Address: lductor@gmail.com. (L.Ductor).

1. Introduction

We document a strong and robust negative correlation between the length of the title of an economics article and its scientific quality. Articles with shorter titles tend to be published in better journals. Controlling for journal quality, articles with shorter titles tend to receive more citations. To illustrate, papers published between 1970 and 2011 in economics journals had a title composed, on average, of 78 characters. Over the same period, the average title length of papers in the *American Economic Review* was 60 characters. Within the *AER*, the top 20 most cited articles had an average title length of 52 characters.

These correlations may have different explanations. Title length could have a causal impact on quality measures. For instance a short title could make a paper easier to memorize, which could directly increase its citations. Title length and quality could also have common determinants. Authors' research and writing skills, in particular, could affect both the paper's measured outcomes and the length of its title. Good researchers may systematically adopt shorter titles and publish well-cited papers in good journals. Other common determinants may be harder to measure. We notably expect more novel papers to have shorter titles. Seminal articles often have very short titles; consider "A Theory of Production" (Cobb & Douglas 1928), with 22 characters or "Economic Growth and Income Inequality" (Kuznets 1955) with 37 characters. Further studies on these topics had to adopt longer titles. Novelty is not easy to measure, however, and could be imperfectly correlated with journal quality or citation counts.

We provide an in-depth analysis of the relationship between title length and scientific quality. We examine all articles published between 1970 and 2011 and referenced in *EconLit*. Our initial sample contains more than 580,000 articles.[1] We consider three measures of article quality. We look at an impact factor of the journal in which the article is published and at the number of citations received by the article. Citations capture the impact of the article on subsequent research. Some recent studies indicate, however, that citations may be little correlated with the article's novelty, see Lee et al. (2015). We build on recent advances in bibliometric analysis and construct an index of novelty based on keyword atypicality, as in Boudreau et al. (2017) and Sreenivasan (2013). Keywords capture central aspects of research and more novel papers likely have atypical keywords or keyword combinations. These three measures - journal quality, citations and novelty - capture different dimensions of the scientific quality of an article. Our main objective is to better understand the relationship between title length and these different

---

[1]We provide descriptive statistics and Figures based on this initial sample in Section 2. In our regressions, the largest sample studied contains about 490,000 articles.

dimensions.

We estimate regressions of an article's quality measure on title length and on an expanding set of controls. We control, first, for the article's number of pages and for JEL code fixed effects, capturing systematic differences across fields. When studying citations and novelty, we control for journal fixed effects. We thus analyze how citations or novelty are related to title length controlling for journal unobserved heterogeneity. Importantly, we also control for characteristics of the authors. We include a full set of dummies at the team level.[2] These team of authors dummies capture time-invariant (observed and unobserved) characteristics of the teams, such as team ability and gender composition. We also control for observable time-varying team characteristics such as specialization, past team output and average past output of the team members. We expect these team dummies and time-varying characteristics to account reasonably well for the impact of authors' skills and ability. Finally, we include novelty as a control in the regressions on journal quality and citations to see whether the observed relationships between title length and these measures are mediated through novelty.

In all specifications, our main estimates are negative and statistically and quantitatively significant. Articles with shorter titles tend to be published in better journals. They tend to be more cited and to get higher novelty scores. Moreover, these tendencies are more pronounced in better journals. Controlling for team characteristics has a strong impact in the regressions on journal quality. The main estimate, which remains statistically and quantitatively significant, is approximately divided by three. This shows that better teams indeed have a strong systematic tendency to publish articles in better journals and with shorter titles. By contrast, the main estimates are essentially unchanged when controlling for team characteristics in regressions on citations or novelty. Articles with shorter titles tend to be more cited and more novel than articles with longer titles, controlling for authors' skills and abilities.

Moreover, including novelty in the regressions on journal quality and citations has essentially no impact on the estimates. This means that the observed relation between title length and citations is not explained by the fact that novel articles tend to have shorter titles and to be more cited. Together, these results show that title length correlates well with the overall scientific quality of a paper.

---

[2]Suppose, for instance, that author a has published single-authored articles, articles with author b, articles with authors b and c, and articles with author d. This leads to the inclusion of four different dummies in the regressions: $\delta_{\{a\}}$, $\delta_{\{a,b\}}$, $\delta_{\{a,b,c\}}$ and $\delta_{\{d\}}$ where, for instance, $\delta_{\{a,b\}} = 1$ for articles written by a and b and 0 otherwise.

Our analysis contributes to the scientific study of the academic process.[3] Two recent studies document a negative association between title length and citations on large samples, controlling for journal fixed effects. Letchford et al. (2015) look at highly cited articles in the Web of Science across all disciplines. Independently of our analysis, Gnewuch & Wohlrabe (2017) consider economics article published between 1980 and 2015 and referenced in the Web of Science.[4]

We build on these earlier studies and make two main contributions to the study of the relation between title length and scientific quality. First, we analyze two other dimensions of scientific quality than citations. We provide the first large-scale study of the relation between title length and the impact factor of the journal in which the article is published. Perhaps more importantly, we build an original index of novelty, based on keyword atypicality, and provide the first analysis of the link between novelty and title length. We show that title length is also strongly negatively correlated with these two measures. Second, we provide the first analysis controlling for time-invariant and observed time-varying characteristics of teams of authors. This allows us to account for main determinants of scientific quality, which were ignored in previous studies. We show that two thirds of the association between title length and journal quality is explained by team characteristics while surprisingly the observed relation between title length and citations is quantitatively robust to the inclusion of team characteristics in the regressions.[5] Overall, our results reveal a strong and robust negative relation between the length of the title of an article and its scientific quality.

The remainder of the paper is organized as follows. We describe the data in Section 2. We present our empirical specifications in Section 3. We study the relation between title length, journal quality, citations and novelty in Section 4 and conclude in Section 5.

2. The Data

We analyse two main datasets. We first use information on all articles published between 1970 and 2011 in journals listed in *EconLit*, the bibliography of economics journals compiled by

---

[3]In economics, see, for instance, Bramoullé et al. (2014), Ductor et al. (2014), Fafchamps, Goyal & van der Leij (2010).

[4]Researchers have been looking for correlates of articles' quality and impact, see, e.g., Haslam et al. (2008), Hartley et al. (1988), Smart & Bayer (1986), Webster et al. (2009). For instance, article length appears to be a good predictor of citations, across different periods and disciplines, see Card & DellaVigna (2013), Falagas et al. (2013), Fox et al. (2016). The negative relation with title length was noticed in some early studies on small samples, see Jacques & Sebire (2010), Jamali & Nikzad (2011) and Habibzadeh & Yadollahie (2010).

[5]We also control for field fixed effects, which are ignored in Letchford et al. (2015) and Genwuch & Wohlrabe (2017).

the *American Economic Association*. We exclude from the sample articles containing in their title the words "note", "comment", "preface", "remark", "reply" and "foreword".[6] This first dataset contains 580,055 articles published in 1,617 journals, and is thus composed of the entire body of academic economics publications over the period. We measure journal quality through a standard impact factor introduced by Kodrzycki & Yu (2006). The Kodrzycki & Yu index (hereafter, KY) uses an iterative process, excluding self-citations of the journal and weighting citations on the influence of the citing journal. As this index is not available for all the journals listed in *EconLit*, we use the predicted impact factor derived in Ductor et al. (2014) for journals not listed in KY.[7] Note that this measure is constant and does not vary over time. Computing a time-varying impact factor for the $1,617$ journals listed in EconLit would be quite challenging as most of these journals are relatively new and citations are not easily available for most of them. In addition, changes in journals impact factors appear to be quite small in economics, both in absolute term and relatively to other disciplines, see Althouse et al. (2009). As a robustness check, we consider alternative measures of journal quality. We notably analyze two rankings based on coarse categories, which should be unaffected by marginal changes in specific journals impact factors, see Supplementary Appendix A. We find that results are robust to the quality measure.

Since *EconLit* does not provide information on citations, we retrieved citations from the Web of Science (Thomson Reuters 2014). Due to the expensive and time-consuming nature of the retrieval process, we focus on journals appearing in the list of the Tinbergen Institute. This list contains 117 journals ranked in three categories: AA (5 journals), A (27 journals), and B (85 journals), see Appendix. It arguably contains most of the established journals of our discipline. The citation dataset includes information on 161,699 articles and the number of citations they received yearly until 2013. The most cited paper, "Prospect Theory: an Analysis of Decision under Risk" (Kahneman & Tversky 1979), received 8,571 citations, while 13% of the papers have no citations.

We measure the novelty of an article by building an index based on *atypicality of keyword combinations*, following Boudreau et al. (2017) and Sreenivasan (2013). Keywords capture central aspects of the research. This index measures the relative infrequency of pairs of keywords

---

[6]These atypical articles only represent 1.1% of the full sample and tend to have longer titles and to be less cited. Including them in the analysis leads to estimates of the relation between title length and scientific quality that are slightly greater in magnitude.

[7]To illustrate, the KY impact factor is 27.1 for the *American Economic Review*, 6.6 for the *Journal of Development Economics* and 0.6 for the *Pacific Economic Review*.

of the article, compared to previously published articles. Formally, let K_i denote the set of keywords of article i. Let $N_{t;k_1,k_2}$ denote the number of articles published in year t or before which contain the two keywords k_1 and k_2. Let N_t denote the overall number of pairs of keywords contained in articles published in year t or before. Define the probability of observing the pair of keywords $\{k_1, k_2\}$ before year t as

$$P_t(k_1, k_2) = \frac{N_{t;k_1,k_2}}{N_t}$$

Note that $-\ln(P_t)$ measures the atypicality of the pair of keywords $\{k_1, k_2\}$ in this period. Novelty of article i published in year t is then equal to the normalized average atypicality of its pairs of keywords. Formally,

$$Nov_i = -\frac{\sum_{k_1 \neq k_2 \in K_i} \ln(P_t(k_1, k_2))}{|K_i||K_i + 1| \ln(N_t)}$$

where $|K_i|$ is the number of keywords of article i. Note that $0 \leq -\ln(P_t) \leq \ln(N_t)$ and hence Nov_i varies between 0 and 1. An article's novelty gets close to its maximal value of 1 when all its pairs of keywords are extremely atypical.[8]

We measure the length of an article's title by simply counting the number of characters, including spaces and punctuations. Average title length in the Econlit sample is 78; some papers have very long titles.[9] Descriptive statistics give us a first idea of the link between title length and scientific quality. Over the whole period, average title length is 61 for AA journals, 66 for A journals, 70 for B journals and 81 for unranked journals. Figure 1 shows the distributions of title length across journal categories using the entire EconLit sample, 580,055 articles. We observe a clear ordering of the distributions by journal quality. Statistically, any notion that distributions are not ordered through First-Order Stochastic Dominance relations aligned with quality is rejected. Figure 2 shows the distributions of title length for articles in the first decile of the citation distribution, those in the last decile and all other articles using the TI sample, 161,699 articles. Again, these distributions are unambiguously ordered and articles with more citations appear to have shorter titles. Figure 3 shows the distribution of title length across different percentiles of the novelty distribution, for the sample of articles with at least one keyword, 464,835 articles. Clearly, the title length of articles in the top 10% of novelty have shorter titles than the rest of articles.

---

[8]For robustness, we also look at two alternative indices: one built from relative frequencies of keyword pairs compared to articles published in the same year and another based on the atypicality of individual keywords. The results are robust to these alternative measures of novelty, see Supplementary Appendix A.

[9]For instance, the following title has 203 characters: "Analysis of the carbon sequestration costs of afforestation and reforestation agroforestry practices and the use of cost curves to evaluate their potential for implementation of climate change mitigation", Torres et al. (2010).

Figure 4 shows the evolution of average title length over time and by journal quality using the EconLit sample from 1974 to 2011, 567,218 articles. For unranked and B ranked journals, we observe a clear increasing tendency. The average title length of articles published in these two categories has increased substantially over time (despite a slight temporary decrease around 1990). This increase is consistent with the large rise in the number of articles published over the period and with a general tendency towards more specialization. As authors develop finer and finer extensions of seminal works, they likely need to adopt longer titles to describe their research. Interestingly, however, A and AA ranked journals display different trends. For top field and top generalist journals, we also observe a similar increasing trend at the beginning of the period. However, a break occurs at some time in the 1980s for the two categories. Average title length then becomes slightly decreasing with time for articles published in A ranked journals and stays essentially invariant for articles published in AA ranked journals. A possible explanation behind these distinct tendencies could be the documented increase in competition over the period (Card and DellaVigna 2013).[10] As publication slots become more and more scarce in top journals authors likely increase the quality of their submitted papers, leading to shorter titles.

3. Empirical model

In this section, we present the main empirical specifications that we use to study the relationship between title length and article quality. For journal quality, we consider variants of the following econometric model:

$$
\begin{aligned}
log(q_i) \quad = \quad & \beta_0 + \beta_1 length_i + \beta_2 pages_i + \sum_{f=1}^{F} \lambda_f JEL_{i,f} \\
& + \beta_3 H_{r,t-1} + \beta_4 T_{r,t-1} + \beta_5 A_{r,t-1} + \delta_r + \mu_t + \epsilon_i,
\end{aligned} \tag{1}
$$

where i denotes an article, r the article's research team and t the year where this article is published. The research team of an article is the set of its author(s). For instance, $r = \{a\}$ for all articles single-authored by author a while $r = \{a, b\}$ for articles written by authors a and b. Note that authors are identified in the data on the basis of their first and last names. In the relatively infrequent cases where first names are not available, we apply the name disambiguation algorithm designed by Van der Leij (2006).[11]

---

[10]Previously documented patterns consistent with increased competition include an increase in the number of submissions to the top 5 (Card and DellaVigna 2013), in number of coauthors (Ductor 2015), in papers' length (Card and Della Vigna 2014) and in turnaround time, (Ellison 2002).

[11]We also apply this algorithm to account for middle initials, see the Appendix of Chapter 2 in Van der Leij (2006) for details.

The left-hand side variable, $log(q_i)$, denotes the log of the impact factor of the journal in which article i is published. On the right hand side, $length_i$ is the length of the title of the article, $pages_i$ is number of pages and $JEL_{i,f}$ denotes the proportion of the JEL codes of the articles in category f.[12] Next, the variables $H_{r,t-1}$, $T_{r,t-1}$ and $A_{r,t-1}$ represent time-varying characteristics of the article's research team. $H_{r,t-1}$ is the degree of specialization of research team r, computed over all publications by research team r published in year $t-1$ or before.[13] To measure specialization, we use the Herfindahl index as in Ductor (2015). Formally, let $n_{t-1,f}$ denote the number of articles written by the team in field f and published in year $t-1$ or before. Let $n_{t-1} = \sum_{f=1}^{F} n_{t-1,f}$ denote the overall number of articles published by the team until $t-1$. Articles with multiple JEL codes are divided and assigned proportionally to each field. Then,

$$H_{r,t-1} = \sum_{f=1}^{F} \left(\frac{n_{t-1,f}}{n_{t-1}} \right)^2 .$$

This measure is maximum and equal to 1 when all the team's papers are written in a unique field. A lower value indicates a lower degree of specialization.

$T_{r,t-1}$ is the past output of the research team, computed as the overall number of papers, weighted by journal quality, published by research team r until $t-1$. $A_{r,t-1}$ captures the average past output of the different authors in the team, computed as the overall number of papers, weighted by journal quality, published by all authors of the research team. Note that these two measures coincide in some specific cases, for instance when r has a single author who has never written with someone else or when two authors write all their papers together. They generally differ, however. $T_{r,t-1}$ measures the track record of the team itself while $A_{r,t-1}$ captures the track record of the teams members. Thus, a team composed of two senior researchers who start collaborating will have a high $A_{r,t-1}$ and a low $T_{r,t-1}$.

We also include research team dummies δ_r. For instance, if authors a and b write three articles i, j and k together, the dummy variable $\delta_{\{a,b\}}$ is equal to 1 for articles i, j and k and to 0 for all other articles. These dummies control for time-invariant characteristics of the research teams, such as innate abilities or ethnic and gender composition. We also present results without team dummies, but with dummies for the number of authors of the article. Finally, we include year of publication dummies, μ_t, for each year from 1970 to 2011. These year dummies account for any possible time trends in journal quality. In some specifications, we further interact year

---

[12]For instance if an article has JEL codes $B2$, $C1$ and $C5$, then $JEL_{i,B} = 1/3$, $JEL_{i,C} = 2/3$ and $JEL_{i,f} = 0$ if $f \notin \{B,C\}$.

[13]By assumption, $H_{r,t-1} = T_{r,t-1} = 0$ if research team r has not published any article in year $t-1$ or before.

dummies, or a linear time trend, with title length, to detect differences in the effect of title length across time.

To study citations and novelty, we augment the previous model with journal fixed effects. For instance, let c_i denote the number of citations gathered by article i and j denote the journal in which the article is published. We base our citation regressions on the following model:

$$
\begin{aligned}
log(c_i + 1) \; = \; & \beta_0 + \beta_1 length_i + \beta_2 pages_i + \sum_{f=1}^{F} \lambda_f JEL_{i,f} \\
& + \beta_3 H_{r,t-1} + \beta_4 T_{r,t-1} + \beta_5 A_{r,t-1} + \delta_r + \nu_j + \mu_t + \epsilon_i
\end{aligned} \tag{2}
$$

where journal fixed effects, ν_j, now capture systematic differences in citation patterns across journals. We adopt a similar model to analyze an article's novelty, Nov_i. The log plus one transformation of the number of citations mitigates the effect of highly cited papers, see Ductor et al. (2015).[14]

4. Results

We now present our main results, analyzing the relation between title length and the three different quality measures in turn. Table 1 presents estimates of variants of equation (1), regressing the logarithm of the impact factor of the journal in which the article is published over the length of the article's title and expanding sets of controls.[15] Column (1) presents results of estimations without controls. We focus in our regressions on articles published in journals whose extended KY impact factor can be computed and with JEL code information. About 13.7% of the papers in the original sample are published in journals whose impact factor cannot be computed and a further 2.8% does not have information on JEL codes. Our main regression sample thus contains information on $488,649$ articles. The raw estimate is -0.0064. To get a sense of the magnitude, this estimate implies that a switch from the first decile (39) to the last decile (121) of the title length distribution is associated with a 52% increase in the journal's impact factor. In the estimations underlying Column (2), we control for the article's number of pages, dummies for different numbers of authors, JEL code dummies, and year dummies. Controlling for these factors has little impact on the estimate.

We then include research team dummies and three time-varying characteristics of the research teams (degree of specialization, past output jointly published by the team and average

---

[14]Our results are robust to looking at the number of citations directly and to count data models such as poisson and negative binomial, see Supplementary Appendix A.

[15]Our results are robust to regressing journal quality over powers of the article's title length, see Supplementary Appendix A.

past output of all the authors in the research team). The team dummies control for heterogeneity at the level of the team of authors. Note that identification now relies on articles published by teams with at least two publications. To detect potential selection effects, we report in Column (3) results of estimation on this subsample without team fixed effects and characteristics. The estimate is essentially unchanged. We then report results of estimations including team fixed effects in Column (4). The estimate is roughly divided by 3. Therefore, two thirds of the original effect is due to composition: better teams tend to publish papers in better journals and with shorter titles. The remaining third captures the fact that, even controlling for time-invariant team heterogeneity, papers with shorter titles tend to be published in better journals. Note that the effect is still quantitatively significant. A switch from first to last decile of the title length distribution is now associated with a 16% increase in journal impact factor.

Adding time-varying team characteristics, in Column (5), leaves the estimate associated with title length essentially unaffected. These results reveal some interesting effects, however. Specialization is positively associated with journal quality, which is consistent with the existence of positive returns to specialization in the publication process. Authors' past output is also positively associated with journal quality, which could reflect a positive impact of experience and past success on publications. Controlling for authors' past output, however, team past output is negatively associated with journal quality. Authors with diverse sets of coauthors thus tend to publish in better journals than authors who always publish together.

Finally, we investigate whether the observed negative relationship between title length and journal quality can be explained by novelty, as more novel articles may have shorter titles and be published in better journals. To compute the novelty score, we need to further restrict the sample to articles for which keywords are available. In Column (6), we report results from the regressions underlying Column (5) on this new subsample. The main estimate is little affected. We then add the novelty score to the controls and report results in Column (7). The estimate remains unchanged. Therefore, the negative association between title length and journal quality is not mediated through novelty.

We next turn to the relation between title length and citations. Following equation (2), we regress the logarithm of the number of citations of the article in 2011 plus one against the length of the article's title and controls. Year dummies control for the fact that citations accumulate over time.[16] Column (1) of Table 2 presents results of regressions including the same controls

---

[16]Our results are robust to alternative ways to account for this accumulation process, such as looking at citations per year, number of citations during the first five years after publication, or the percentile in the citation distribution, see Supplementary Appendix A.

as in Column (2) of Table (1). The estimate is -0.0017. A switch from first to last decile of the title length distribution is then associated with a 14% increase in the number of citations.

In the next set of regressions, we include journal dummies. Results are reported in Column (2) of Table 2. The estimate decreases in magnitude and is now -0.0008. Comparing Columns (1) and (2) we see that an important fraction of the association between title length and citations is explained by time invariant characteristics specific to the journal where the article is published. We then include research team dummies and time varying team characteristics, we report results in Column (4) and (5). Interestingly the estimate, which is still negative and statistically significant, is now larger in magnitude.[17] Therefore, articles with shorter titles tend to be more cited, even controlling for journal time-invariant characteristics and for team time-invariant and time-varying observed characteristics. To illustrate, consider a research team publishing two articles in the *American Economic Review* in 2001, one with 39 characters in its title and the other with 122 characters. The article with the shorter title is predicted to receive 15.3 citations in 2013 while the one with the longer title would receive 13.6 citations.[18]

Controlling for team characteristics therefore has a very different effect depending on whether we look at journal quality or citations. A large part of the negative association between title length and journal quality is due to the fact that teams with better ability and, for instance, more experienced authors publish articles in better journals and with shorter titles. This conforms to our initial expectation since team characteristics are primary determinants of an article's outcomes and content. Surprisingly, however, controlling for team characteristics has little impact on the association between title length and citations.

We also find it quite interesting to contrast result reported in Column (5) of Table 2 with those reported in Column (5) of Table 1. Specialization is now negatively associated with citations, in contrast to its positive association with journal quality. The articles published by more specialized teams therefore receive less citations, even though they tend to be published in better journals. This could reflect the fact that they tend to reach a narrower audience. By contrast, estimates of past team output and past output of team members have similar signs. Articles published by authors with better track records thus tend to receive more citations and to be published in better journals, consistently with the Matthew effect (Merton, 1968). And articles published by authors who publish exclusively with the same coauthors tend to receive

---

[17]As shown by results reported in Column (3), part of this increase is due to a composition effect. The association between title length and citations is stronger on the subset of articles published by research teams with at least two publications.

[18]To get these predictions, the other control factors are evaluated at the mean.

less citations and to be published in lower-ranked journals.

Finally, we include novelty as a control and report results in Column (7). The main estimate is unchanged. The negative association between title length and citations is therefore not explained by the fact that more novel papers tend to have shorter titles and to be more cited. Surprisingly, the relationship between novelty and citations is negative and statistically significant. Controlling for title length, journal time invariant characteristics and team characteristics, more novel articles therefore tend to be less cited. These findings complement recent evidence on novelty in research: Boudreau et al. (2016) show that novel research grant proposals are associated with lower evaluations by committees while Lee et al. (2015) show in a different context that novel articles indeed appear to accumulate citations at a lower pace.

We further look at how the relation between title length and citations varies by journal quality. We first include the journal's impact factor and its interaction with title length in a citation regression. The results are reported in Column (1) of Table 3. The negative coefficient of the interaction term shows that the relation between title length and citations is higher for journals with higher impact factor. We then estimate separate citation regressions for each journal category. The results are reported in Columns (2), (3) and (4) of Table 3. The effect is twice as large in A and AA journals as in B journals. The relation between title length and citations is therefore stronger in better journals.

Third, we study the association between title length and the novelty of the research article measured using the atypicality of keywords pairs. The results presented in table 4 reveal a negative robust association between title length and novelty. The estimate is equal to -0.0003 in the specification that control for team characteristics, see Column (5), and changes very little across specifications. Quantitatively, a switch from the first decile to the last decile of the title length distribution is associated with a 0.0247 increase in novelty, corresponding to a $+8.3\%$ increase compared to average novelty, equal to 0.31. Articles with shorter titles therefore tend to score higher on the novelty index, even when controlling for journal fixed effects and team characteristics.[19]

Fourth, we analyze whether the relation between title length and scientific quality varied over time. We modify our preferred empirical specifications (including journal and team fixed effects and team time-varying characteristics) in two ways. We first include linear time trends (rather than year fixed effects) and linear time trends interacted with title length in the regressions.

---

[19]In line with expectations, team specialization is negatively associated with novelty while team members past output is positively associated with novelty.

Columns (1), (2) and (3) of Table 4 report results of these regressions on the three different quality measures studied: journal quality, citations and novelty. We then relax the linearity assumption and include year dummies as well as year dummies interacted with title length in the regressions. Results from this more flexible specification are reported in Column (4), (5) and (6) of Table 4.

On journal quality, results from the specification with linear trends indicate that the relation between title length and journal quality has decreased in magnitude over time (Column (1)). However, this apparent decrease is not robust to relaxing linearity. In Column (4), we see that the coefficients of the interacted terms are jointly statistically insignificant. On citations, results across the two specifications are consistent and indicate that the relation between title length and citations has not changed over time. Results from the two specifications are also consistent for novelty and indicate that the relation between title length and novelty has increased in magnitude over time. Moreover, this increase seems concentrated in the last decade of the period studied (from 2000 to 2011).[20]

Finally, we investigate whether title length and its assocation with scientific quality differ between field and general purpose journals.[21] Intuitively, articles published in field journals may be more specialized and may require longer titles to convey their intentions to the readers. To study this, we first classify journals in specialist and generalist, see Apppendix. We find that, indeed, articles published in specialist journals have longer titles, on average, (70.1 characters) than those published in generalist journals (64.2 characters). We then introduce an interaction term between a generalist dummy and title length to our main empirical specifications. We report results in Table 6. We find that the relation between title length and citations or novelty is not statistically different between the two kinds of journals, see Columns (2) and (3). By contrast, the relation between title length and journal quality appears to differ between field and general purpose journals. In particular, we do not detect a significant relation between title length and journal quality for specialist journals, once we control for team characteristics.

5. Conclusions

Analyzing all articles published between 1970 and 2011 and referenced in *EconLit*, we document a strong and robust negative correlation between the length of the title of an article

---

[20]We also analyze how the relation between title length and research quality evolve over time across journal quality, by running for each outcome separate regressions including year dummies and year dummies interacted with title length for the three journal categories of the Tinbergen Institute. In each of these nine regressions, the interaction terms are jointly insignificant, indicating an absence of time trends by journal category. Results are available upon request

[21]We thank an anonymous referee for this suggestion.

and three measures of scientific quality: the impact factor of the journal in which the article is published, the number of citations it receives and an index of novelty based on keyword combination atypicality. These correlations hold conditioning on key common determinants such as year dummies and field shares, journal dummies (for citations and novelty), team of authors dummies and time-varying characteristics of authors' teams such as specialization and track record.

Our results admit two possible explanations. On one hand, title length could have a causal impact on journal quality or citations. A short title could make an article easier to memorize, affecting citations and, possibly, editorial decisions. Identifying such causal impacts empirically would be challenging. It would require finding exogenous sources of variation in title length, independent of the content of the article itself. On the other hand, title length could proxy for the true, unobserved qualities of the article. Articles with a strong potential to influence subsequent research could thus both generate more citations and have shorter titles. Similarly, articles with truly novel content could both tend to have shorter titles and atypical keywords pairs. Note that scientific quality is, in general, hard to define, hard to measure and surely multidimensional. Thus, we find it quite remarkable that title length, which is a very simple and easy to measure feature of a research article, displays such strong and robust correlations with different measures of scientific quality.

Bibliography

[1] Althouse, Benjamin M., West, Jevin D., Bergstrom, Carl T., and Bergstrom, Theodore, 2009. Differences in impact factor across fields and over time. *Journal of the Association for Information Science and Technology*, 60(1), 27-34.

[2] Arrow, Kenneth J., Bernheim, B. Douglas, Feldstein, Martin S.,McFadden, Daniel L., Poterba, James M., and Robert M. Solow, 2011. 100 years of the *American Economic Review*: the Top 20 Articles. *American Economic Review* 101(1): 1–8.

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

[4] Bramoullé, Yann, Currarini, Sergio, Jackson, Matthew O., Pin, Paolo, and Robert W. Brian, 2012. Homophily and Long-Run Integration in Social Networks. *Journal of Economic Theory* 147(5): 1754–1786.

[5] Card, David, and Stefano DellaVigna, 2013. Nine facts about top journals in economics. *Journal of Economic Literature*, 51(1), 144–161.

[6] Card, David, and Stefano DellaVigna, 2014. Page limits on economics articles: Evidence from two journals. *The Journal of Economic Perspectives*, 28(3), 149–167.

[7] Cobb, Charles W., and Paul H. Douglas, 1928. A theory of production. *The American Economic Review*, 18(1), 139–165.

[8] Ductor, Lorenzo, 2015. Does Coauthorship Lead to Higher Academic Productivity? *Oxford Bulletin of Economics and Statistics*, 77(3), 385–407.

[9] Ductor, Lorenzo, Fafchamps, Marcel, Goyal, Sanjeev, and Marco J. van der Leij, 2014. Social networks and research output. *Review of Economics and Statistics*, 96(5), 936–948.

[10] Ellison, Gleen, 2002. The Slowdown of the Economics Publishing Process. *Journal of Political Economy*, 110(5), 947-993.

[11] Falagas, Matthew E., Zarkali, Angeliki, Karageorgopoulos, Drosos E., Bardakas, Vangelis, and Michael N. Mavros, 2013. The impact of article length on the number of future citations: a bibliometric analysis of general medicine journals. *PLoS One*, 8(2), p.e49476.

[12] Fafchamps, Marcel, Leij, Marco J., and Sanjeev Goyal, 2010. Matching and network effects. *Journal of the European Economic Association*, 8(1), 203–231.

[13] Fox, Charles. W., Paine, C.E. Timothy, and Boris Sauterey, 2016. Citations increase with manuscript length, author number, and references cited in ecology journals. *Ecology and Evolution*, 6(21), 7717–7726.

[14] Gnewuch, Mathias and Klaus Wohlrabe, 2017. Title characteristics and citations in economics. *Scientometrics*. doi:10.1007/s11192-016-2216-7

[15] Habibzadeh, Farrokh, and Mahboobeh Yadollahie, 2010. Are shorter article titles more attractive for citations? Crosssectional study of 22 scientific journals. *Croatian Medical Journal*, 51(2), 165–170.

[16] Hartley, James, Trueman, Myfanwy, and A. Jack Meadows, 1988. Readability and prestige in scientific journals. *Journal of Information Science*, 14(2), 69–75.

[17] Haslam, Nick, Ban, Lauren, Kaufmann, Leah, Loughnan, Stephen, Peters, Kim, Whelan, Jennifer, and Sam Wilson, 2008. What makes an article influential? Predicting impact in social and personality psychology. *Scientometrics*, 76(1), 169–185.

[18] Jamali, Hamid R., and Mahsa Nikzad, 2011. Article title type and its relation with the number of downloads and citations. *Scientometrics*, 88, 653–661.

[19] Jacques, Thomas S., and Neil J. Sebire, 2009. The impact of article titles on citation hits: an analysis of general and specialist medical journals. *Journal of the Royal Society of Medicine Short Reports* 1, 1–5.

[20] Kahneman, Daniel, and Amos Tversky, 1979. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the econometric society*, 47(2),263–291.

[21] Kodrzycki, Yolanda K., and Pingkang Yu, 2006. New approaches to ranking economics journals. *The BE Journal of Economic Analysis & Policy*, 5(1).

[22] Kuznets, Simon, 1955. Economic growth and income inequality. *The American Economic Review*, 45(1), 1–28.

[23] Lee, Yun N., Walsh, John P., and Jian Wang, 2015. Creativity in scientific teams: Unpacking novelty and impact. *Research Policy*, 44(3), 684–697.

[24] Letchford, Adrian, Moat, Helen S., and Tobias Preis, 2015. The advantage of short paper titles. *Royal Society open science*, 2(8), 150–266.

[25] Merton, Robert K, 1968. The Matthew effect in science. *Science*, 159(3810), 56-63.

[26] Torres, Arturo B., Marchant, Rob, Lovett, Jon C., Smart, James C.R., and Richard Tipper, 2010. Analysis of the carbon sequestration costs of afforestation and reforestation agroforestry practices and the use of cost curves to evaluate their potential for implementation of climate change mitigation. *Ecological Economics*, 69(3), 469-477.

[27] Smart, John C., and Alan E. Bayer, 1986. Author collaboration and impact: a note on citation rates of single and multiple authored articles. *Scientometrics*, 10(5–6), 297–305.

[28] Sreenivasan, Sameet, 2013. Quantitative analysis of the evolution of novelty in cinema through crowdsourced keywords. *Scientific reports*, 3.

[29] Van der Leij, Marco J., 2006. The economics of networks: Theory and empirics. The Tinbergen Institute Research Series, 384. Thela Thesis, Amsterdam.

[30] Webster, Gregory D., Jonason, Peter K., and Tatiana O. Schember, 2009. Hot topics and popular papers in evolutionary psychology: Analyses of title words and citation counts in evolution and human behavior, 1979–2008. *Evolutionary Psychology*, 7(3), 348–362.

6. Appendix: List of journals

The journals in the sample are the economics, econometrics and finance journals that appear on the journal list of the Tinbergen Institute. This list was used to evaluate the research output of three leading economics departments in the Netherlands. It has also been used by Goyal et al.(2006) and Fafchamps et al. (2010).

Top 5 / AA-ranked journals: American Economic Review, Econometrica, Journal of Political Economy, Quarterly Journal of Economics, Review of Economic Studies

A-ranked journals: Accounting Review, Bell Journal of Economics, Econometric Theory, Economic Journal, European Economic Review, Games and Economic Behavior, International Economic Review, Journal of Accounting and Economics, Journal of Business and Economic Statistics, Journal of Econometrics, Journal of Economic Literature, Journal of Economic Perspectives, Journal of Economic Theory, Journal of Environmental Economics and Management, Journal of Finance, Journal of Financial Economics, Journal of Health Economics, Journal of Human Resources, Journal of International Economics, Journal of Labor Economics, Journal of Marketing Research, Journal of Monetary Economics, Journal of Public Economics, RAND Journal of Economics, Review of Economics and Statistics, Review of Financial Studies, World Bank Economic Review

B-ranked journals: American Journal of Agricultural Economics, Applied Economics, Bulletin of the Institute of Economics and Statistics (Oxford-University), Cambridge Journal of Economics, Contemporary Economic Policy, De Economist, Ecological Economics, Economica, Economic Development and Cultural Change, Economic Geography, Economic History Review, Economic History Review (Second Series), Economic Inquiry, Economic Policy: A European Forum, Economic Record, Economics and Philosophy, Economics Letters, Economic Theory, Energy Economics, Environmental and Resource Economics, Environment and Planning A, Explorations in Economic History, Financial Management, Health Economics, IMF Economic Review, Industrial and Labor Relations Review, Insurance: Mathematics and Economics, International Journal of Forecasting, International Journal of Game Theory, International Journal of Industrial Organization, International Review of Law and Economics, International Tax and Public Finance, Journal of Accounting Research, Journal of Applied Econometrics, Journal of Banking and Finance, Journal of Business, Journal of Comparative Economics, Journal of Development Economics, Journal of Economic Behavior and Organization, Journal of Economic Dynamics and Control, Journal of Economic History, Journal of Economic Issues, Journal of Economic Psychology, Journal of Economics and Management Strategy, Journal of Evolution-

ary Economics, Journal of Financial and Quantitative Analysis, Journal of Financial Intermediation, Journal of Forecasting, Journal of Industrial Economics, Journal of Institutional and Theoretical Economics, Journal of International Money and Finance, Journal of Law, Economics, and Organization, Journal of Law and Economics, Journal of Macroeconomics, Journal of Mathematical Economics, Journal of Money, Credit, and Banking, Journal of Population Economics, Journal of Post Keynesian Economics, Journal of Risk and Uncertainty, Journal of Transport Economics and Policy, Journal of Urban Economics, Kyklos , Land Economics, Macroeconomic Dynamics, Mathematical Finance, Marketing Science, National Tax Journal, Oxford Bulletin of Economics and Statistics, Oxford Review of Economic Policy, Public Choice, Regional Science and Urban Economics, Resource and Energy Economics, Resources and Energy, Review of Income and Wealth, Scandinavian Journal of Economics, Scottish Journal of Political Economy, Small Business Economics, Social Choice and Welfare, Southern Economic Journal, Theory and Decision, Transportation Research: Part B: Methodological, Review of World Economics, Western Economic Journal, World Development, World Economy.

Generalist journals from the Tinbergen Institute: American Economic Review, Journal of Politica Economy, Quarterly Journal of Economics, Review of Economic Studies, Econometrica, Economic Journal, European Economic Review, International Economic Review, Journal of Economic Literature, Journal of-Economic Perspectives, Review of Economics and Statistics, World Bank Economic Review, Oxford Bulletin of Economics and Statistics, Cambridge Journal of Economics, Contemporary Economic Policy, De Economist, Economic Inquiry, Economic Record, Economica, Journal of Applied Econometrics, Economics Letters, Journal of Economic Behavior and Organization, Journal of Economic Issues, Kyklos, Oxford Review of Economic Policy, Scandinavian Journal of Economics, Southern Economic Journal, Swedish Journal of Economics and Western-Economic-Journal.

Figure 1: Distribution of title length by journal quality
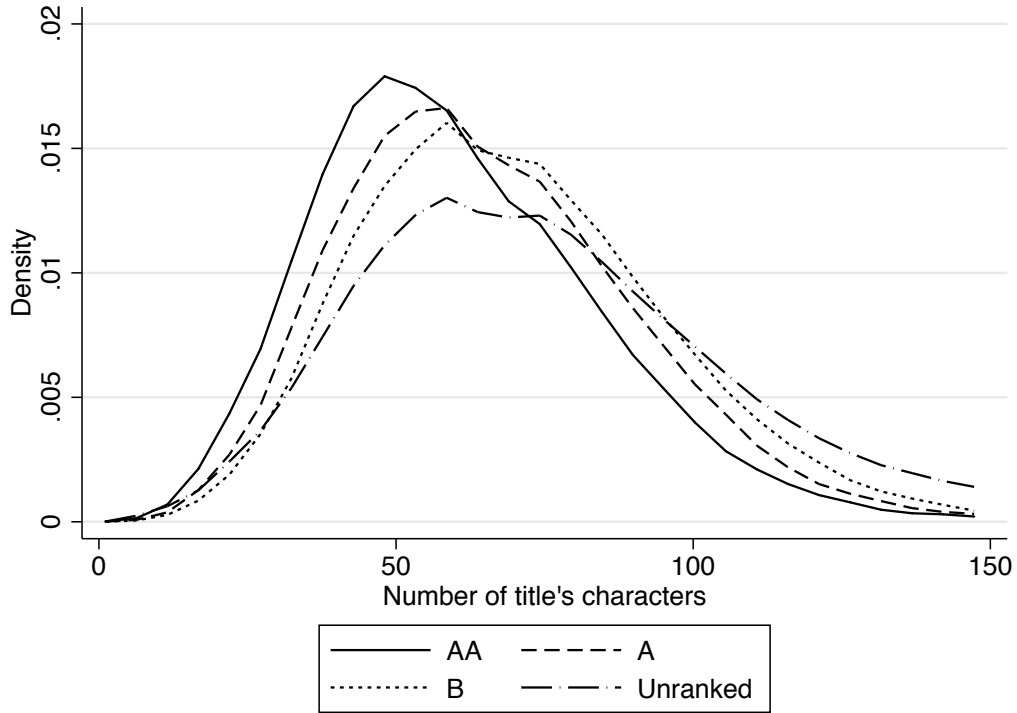
Notes: The sample consists of all regular articles published between 1970 and 2011 and referenced in *EconLit*, 580,055 articles. We consider the Epanechnikov kernel.

Figure 2: Distribution of title length by citations
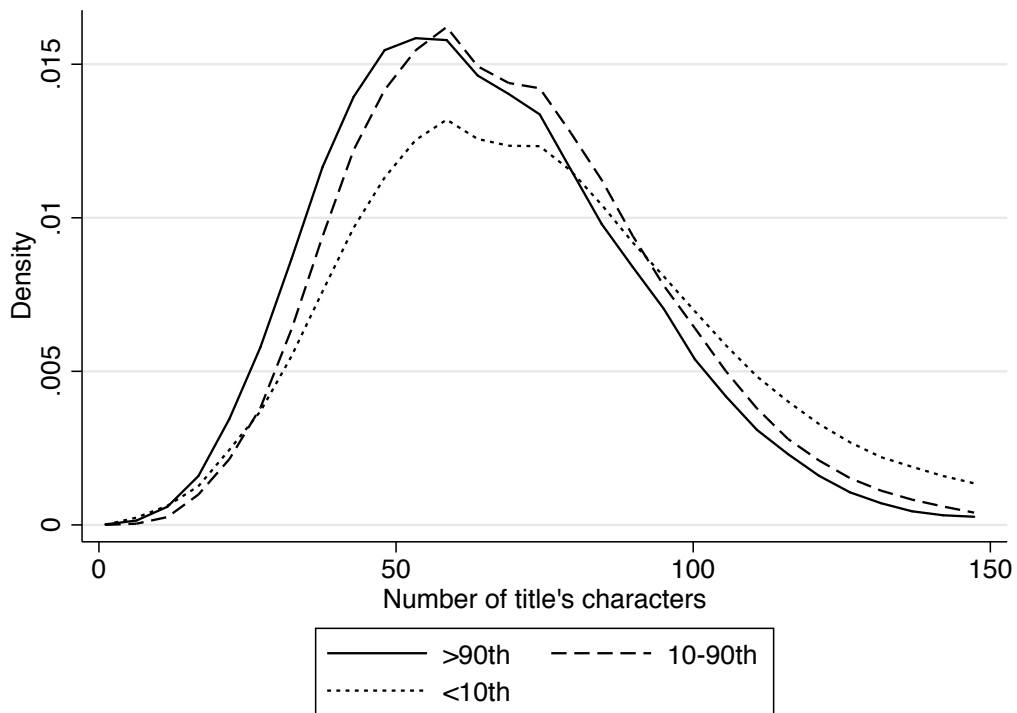
Notes: The sample consists of all regular articles published between 1970 and 2011 in a journal listed in the Tinbergen Institute list, 161,699 articles. We consider the Epanechnikov kernel.

Figure 3: Distribution of title length by novelty

Notes: The sample consists of all regular articles published between 1970 and 2011 in *EconLit* with at least one keyword, 464,835 articles. We consider the Epanechnikov kernel.

Figure 4: Average title length over time and by journal quality

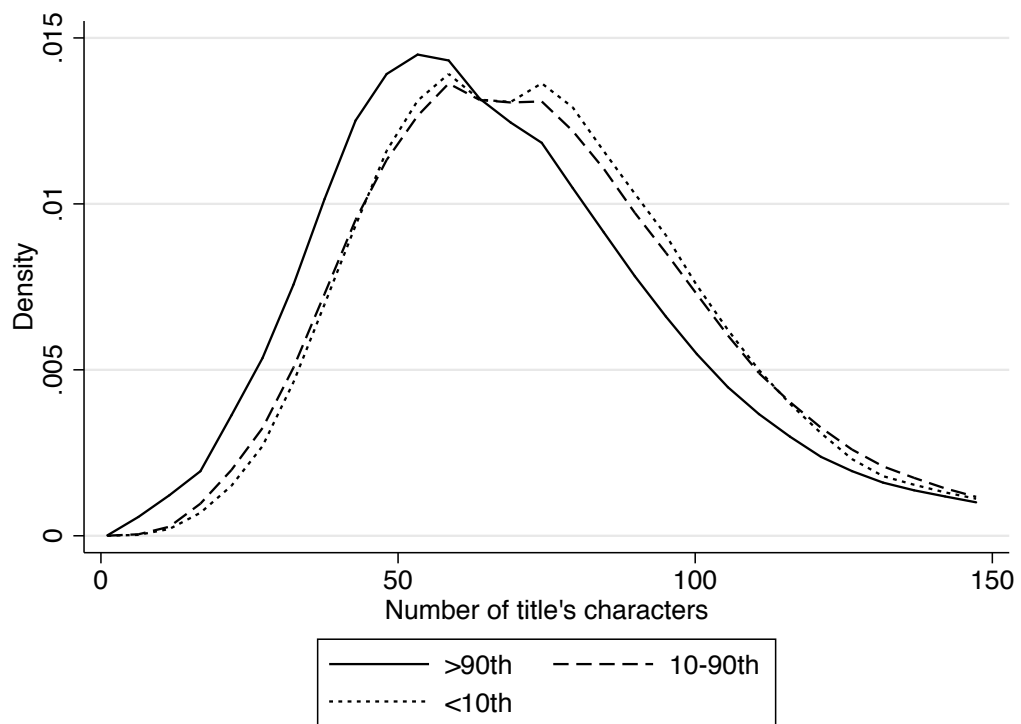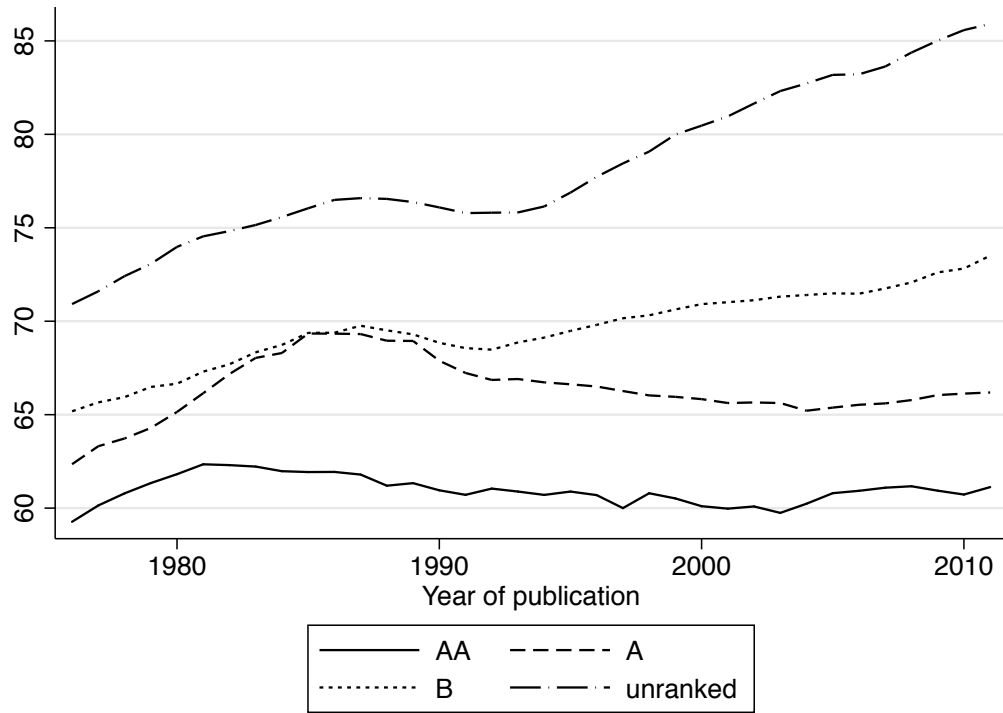*Notes*: Five year moving averages of title length. The sample consists of all regular articles published between 1974 and 2011 and referenced in *EconLit*, 567,218 articles. We consider the Epanechnikov kernel.

Table 1: Title length and journal quality

Variables/Samples:	(1) All	(2) All	(3) TFE	(4) TFE	(5) TFE	(6) TFE	(7) TFE
Title length	-0.0064***	-0.0060***	-0.0062***	-0.0019***	-0.0018***	-0.0020***	-0.0020***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Pages		0.0149***	0.0147***	0.0036***	0.0037***	0.0039***	0.0038***
		(0.0005)	(0.0008)	(0.0003)	(0.0003)	(0.0004)	(0.0004)
Number of authors = 2		0.2895***	0.4103***				
		(0.0069)	(0.0112)				
Number of authors = 3		0.3349***	0.4867***				
		(0.0082)	(0.0195)				
Number of authors = 4		0.2532***	0.4828***				
		(0.0153)	(0.0704)				
Team Specialization					0.1422***	0.1170***	0.1167***
					(0.0100)	(0.0127)	(0.0127)
Past output team					-0.1889***	-0.1868***	-0.1867***
					(0.0041)	(0.0047)	(0.0047)
Avg. past output authors					0.0205***	0.0291***	0.0291***
					(0.0051)	(0.0060)	(0.0060)
Novelty							-0.0152
							(0.0110)
Observations	488,649	488,649	279,239	279,239	279,239	216,494	216,494
R-squared	0.0274	0.1306	0.1371	0.6276	0.6359	0.6397	0.6397
Year dummies	NO	YES	YES	YES	YES	YES	YES
Fields shares	NO	YES	YES	YES	YES	YES	YES
Team dummies	NO	NO	NO	YES	YES	YES	YES

Notes: In columns 1 to 7, we estimate the relationship between the impact factor of the journal where the article is published and the article's title length, the dependent variable is in log. In columns 1 and 2 (All) we consider the full sample of articles. In columns 3, 4 and 5 (TFE) the sample is restricted to articles published by a research team (and sole authors) with at least two publications. In columns 6 and 7 the sample is restricted to articles published by a research team (and sole authors) with at least two publications and articles with at least one keyword. *Pages* is the number of pages of the article; *Number of authors==2, 3, >3* are dummy variables if the number of authors publishing the article is 2, 3, or more than 3, respectively; *Team specialization* is a herfindhal index obtained using the shares of past publications in different fields in economics, as defined by the first digit of the JEL codes; *Past output team* is the number of papers adjusted by quality that the research team has published together in the past; *Avg. past output authors* is the average number of papers adjusted by quality published by the authors of the research team in the past; *Novelty* is an index of the novelty of the article obtained measuring the atypicality of the keywords of the article. All the regressions use clustered standard errors at the team level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2: Title length and citations

Variables/Samples:	(1) All	(2) All	(3) TFE	(4) TFE	(5) TFE	(6) TFE	(7) TFE
Title length	-0.0017***	-0.0008***	-0.0013***	-0.0015***	-0.0014***	-0.0014***	-0.0014***
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Pages	0.0224***	0.0164***	0.0269***	0.0217***	0.0217***	0.0232***	0.0231***
	(0.0017)	(0.0016)	(0.0029)	(0.0030)	(0.0030)	(0.0027)	(0.0027)
Number of authors = 2,	0.2769***	0.2270***					
	(0.0099)	(0.0084)					
Number of authors = 3	0.3943***	0.3170***					
	(0.0121)	(0.0107)					
Number of authors >3	0.5548***	0.4223***					
	(0.0234)	(0.0213)					
Team Specialization					-0.0774***	-0.0934***	-0.0940***
					(0.0239)	(0.0295)	(0.0295)
Past output team					-0.0507***	-0.0551***	-0.0551***
					(0.0044)	(0.0051)	(0.0051)
Avg. past output authors					0.0230***	0.0321***	0.0322***
					(0.0063)	(0.0073)	(0.0073)
Novelty							-0.0600***
							(0.0210)
Observations	161,699	161,699	87,728	87,728	87,728	66,823	66,823
R-squared	0.1502	0.2858	0.2847	0.5848	0.5859	0.5991	0.5992
Year dummies	YES	YES	YES	YES	YES	YES	YES
Fields shares	YES	YES	YES	YES	YES	YES	YES
Journals dummies	NO	YES	YES	YES	YES	YES	YES
Team dummies	NO	NO	NO	YES	YES	YES	YES

Notes: In columns 1 to 7, we estimate the relationship between cumulative citations from year of publication to 2013 and the article's title length, the dependent variable is in $log(x + 1)$. In columns 1 and 2 (All) we consider the full sample of articles. In columns 3, 4 and 5 (TFE) the sample is restricted to articles published by a research team (and sole authors) with at least two publications. In columns 6 and 7 the sample is restricted to articles published by a research team (and sole authors) with at least two publications and articles with at least one keyword. *Pages* is the number of pages of the article; *Number of authors==2, 3, >3* are dummy variables if the number of authors publishing the article is 2, 3, or more than 3, respectively; *Team specialization* is a herfindhal index obtained using the shares of past publications in different fields in economics, as defined by the first digit of the JEL codes; *Past output team* is the number of papers adjusted by quality that the research team has published together in the past; *Avg. past output authors* is the average number of papers adjusted by quality published by the authors of the research team in the past; *Novelty* is an index of the novelty of the article obtained measuring the atypicality of the keywords of the article. All the regressions use clustered standard errors at the team level. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Title length and citations by journal quality

	(1) TFE	(2) AA	(3) A	(4) B
Title length	-0.0006**	-0.0023**	-0.0026***	-0.0013***
	(0.0003)	(0.0009)	(0.0005)	(0.0002)
Journal Quality IF	0.2540***			
	(0.0098)			
Title length*Journal quality IF	-0.0004***			
	(0.0001)			
Pages	0.0230***	0.0388***	0.0105***	0.0381***
	(0.0028)	(0.0085)	(0.0032)	(0.0013)
Team Specialization	-0.0668***	0.0656	-0.1314**	-0.0795***
	(0.0242)	(0.1282)	(0.0623)	(0.0283)
Past output team	-0.0584***	-0.0726***	-0.0511***	-0.0568***
	(0.0047)	(0.0158)	(0.0088)	(0.0077)
Avg. past output authors	0.0236***	0.0067	0.0328**	0.0364***
	(0.0066)	(0.0231)	(0.0137)	(0.0094)
Observations	87,728	7,431	18,998	46,343
R-squared	0.5639	0.5776	0.5844	0.5883
Year dummies	YES	YES	YES	YES
Fields shares	YES	YES	YES	YES
Journals dummies	NO	YES	YES	YES
Team dummies	YES	YES	YES	YES

Notes: In column 1, the results are obtained using team of authors with more than one observation. Column 2, 3 and 4 report the result using the sample of articles published in AA-ranked journals, A-ranked journals and B-ranked journals, respectively. The dependent variable is in $log(x+1)$, Journal quality IF is in log. *Pages* is the number of pages of the article; *Number of authors==2, 3, >3* are dummy variables if the number of authors publishing the article is 2, 3, or more than 3, respectively; *Team specialization* is a herfindhal index obtained using the shares of past publications in different fields in economics, as defined by the first digit of the JEL codes; *Past output team* is the number of papers adjusted by quality that the research team has published together in the past; *Avg. past output authors* is the average number of papers adjusted by quality published by the authors of the research team in the past. All the regressions use clustered standard errors at the team level. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Title length and novelty

Variables/Samples:	(1) All	(2) All	(3) TFE	(4) TFE	(5) TFE
Title length	-0.0003***	-0.0004***	-0.0004***	-0.0003***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Pages	-0.0014***	-0.0012***	-0.0012***	-0.0010***	-0.0010***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Number of authors = 2,	-0.0080***	-0.0049***			
	(0.0010)	(0.0009)			
Number of authors = 3	-0.0057***	-0.0028**			
	(0.0013)	(0.0013)			
Number of authors >3	-0.0078***	-0.0078***			
	(0.0028)	(0.0028)			
Team Specialization					-0.0235***
					(0.0030)
Past output team					-0.0001
					(0.0008)
Avg. past output authors					0.0028***
					(0.0011)
Observations	464,815	464,614	246,824	246,824	246,824
R-squared	0.2985	0.3300	0.3299	0.5567	0.5569
Year dummies	YES	YES	YES	YES	YES
Fields shares	YES	YES	YES	YES	YES
Journals dummies	NO	YES	YES	YES	YES
Team dummies	NO	NO	NO	YES	YES

Notes: In columns 1 to 5, we estimate the relationship between the novelty of the articles obtained as keywords atypicality and the article's title length. In columns 1 and 3 (All) we consider the full sample of articles with keywords. In columns 3, 4 and 5 (TFE) the sample is restricted to articles published by a research team (and sole authors) with at least two publications. *Pages* is the number of pages of the article; *Number of authors==2, 3, >3* are dummy variables if the number of authors publishing the article is 2, 3, or more than 3, respectively; *Team specialization* is a herfindhal index obtained using the shares of past publications in different fields in economics, as defined by the first digit of the JEL codes; *Team experience* is the number of papers that the research team has published together; *Past output team* is the number of papers adjusted by quality that the research team has published together in the past; *Avg. past output authors* is the average number of papers adjusted by quality published by the authors of the research team in the past. All the regressions use clustered standard errors at the team level. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Titles length and research quality over time

VARIABLES	(1) Journal Quality	(2) Citations	(3) Novelty	(4) Journal Quality	(5) Citations	(6) Novelty
Title length	-0.0036***	-0.0027***	0.0001	-0.0031***	-0.0039	0.0001
Team Specialization	0.1148***	-0.0865***	-0.0165***	0.1091***	-0.0994***	-0.0230***
Linear trend	-0.0116***	-0.0206***	-0.0022***			
Title length * linear trend	0.0001***	0.0000	-0.0000***			
Year 1972*title length				0.0000	0.0000	0.0000
Year 1973*title length				0.0009	0.0017	0.0002
Year 1974*title length				0.0003	-0.0002	-0.0005
Year 1975*title length				-0.0012	0.0038	0.0003
Year 1976*title length				-0.0021*	0.0028	-0.0002
Year 1977*title length				0.0010	-0.0000	0.0000
Year 1978*title length				0.0007	0.0049	-0.0004
Year 1979*title length				0.0006	0.0006	-0.0001
Year 1980*title length				-0.0003	-0.0002	-0.0001
Year 1981*title length				0.0007	0.0019	0.0001
Year 1982*title length				0.0003	-0.0013	-0.0003
Year 1983*title length				0.0008	-0.0000	-0.0002
Year 1984*title length				0.0011	0.0001	-0.0003
Year 1985*title length				0.0005	-0.0002	-0.0003
Year 1986*title length				0.0001	0.0012	0.0002
Year 1987*title length				0.0010	0.0016	-0.0002
Year 1988*title length				0.0002	0.0019	0.0002
Year 1989*title length				0.0004	0.0011	-0.0001
Year 1990*title length				0.0015	0.0026	0.0001
Year 1991*title length				0.0009	0.0026	-0.0003
Year 1992*title length				0.0010	0.0024	-0.0003
Year 1993*title length				0.0007	0.0013	-0.0002
Year 1994*title length				0.0019**	0.0025	-0.0003
Year 1995*title length				0.0010	0.0024	-0.0003
Year 1996*title length				0.0012	0.0031	-0.0002
Year 1997*title length				0.0008	0.0041	-0.0001
Year 1998*title length				0.0009	0.0029	-0.0001
Year 1999*title length				0.0009	0.0023	-0.0002
Year 2000*title length				0.0014	0.0015	-0.0008***
Year 2001*title length				0.0013	0.0026	-0.0008***
Year 2002*title length				0.0011	0.0040	-0.0004
Year 2003*title length				0.0015	0.0033	-0.0006**
Year 2004*title length				0.0023**	0.0047	-0.0006**
Year 2005*title length				0.0027***	0.0028	-0.0007**
Year 2006*title length				0.0014	0.0033	-0.0003
Year 2007*title length				0.0019*	0.0026	-0.0004
Year 2008*title length				0.0020**	0.0029	-0.0004
Year 2009*title length				0.0021**	0.0040	-0.0005*
Year 2010*title length				0.0018*	0.0016	-0.0004
Year 2011*title length				0.0012	0.0030	-0.0006**
Past output team	-0.2510***	-0.0876***	0.0025**	-0.2558***	-0.0683***	0.0009
Avg. past output authors	0.0809***	0.0602***	0.0017	0.0825***	0.0438***	0.0022**
Pages	0.0038***	0.0242***	-0.0007***	0.0038***	0.0231***	-0.0011***
Novelty	0.0544***	-0.1631***		-0.0131	-0.0513**	
F-test interaction terms (p-value)				0.5780	0.3804	0.0057
Observations	216,494	66,823	246,824	216,494	66,823	246,824
R-squared	0.6419	0.5811	0.4123	0.6438	0.5996	0.5574
Year dummies	NO	NO	NO	YES	YES	YES
Journal dummies	NO	YES	YES	NO	YES	YES
Fields shares	YES	YES	YES	YES	YES	YES

Notes: The sample is restricted to articles published by a research team (and sole authors) with at least two publications and articles with at least one keyword. X is a dummy variable equal to 1 if the article was published in year X, the baseline year is 1971. *Pages* is the number of pages of the article; *Number of authors==2, 3, >3* are dummy variables if the number of authors publishing the article is 2, 3, or more than 3, respectively; *Team specialization* is a herfindhal index obtained using the shares of past publications in different fields in economics, as defined by the first digit of the JEL codes; *Past output team* is the number of papers adjusted by quality that the research team has published together in the past; *Avg. past output authors* is the average number of papers adjusted by quality published by the authors of the research team in the past; *Novelty* is an index of the novelty of the article obtained measuring the atypicality of the keywords of the article. The F-test interaction terms (p-value) is a test on the joint significance on the interacted terms between the year dummies and title length. Standard errors are not reported for the sake of brevity. All the regressions use cluster standard errors at the research team level*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Title length and citations: generalist vs specialist journals

VARIABLES	(1) TFE Journal Quality	(2) TFE Novelty	(3) TFE Citations
Title length	0.0000	-0.0003***	-0.0014***
	(0.0002)	(0.0001)	(0.0003)
Generalist	0.6436***	–	–
	(0.0298)		
Title length*Generalist	-0.0013***	-0.0000	-0.0007
	(0.0004)	(0.0001)	(0.0004)
Pages	0.0104***	-0.0014***	0.0231***
	(0.0008)	(0.0003)	(0.0026)
Team Specialization	0.1042***	-0.0114*	-0.0959***
	(0.0269)	(0.0067)	(0.0293)
Past output team	-0.0831***	0.0015	-0.0564***
	(0.0043)	(0.0012)	(0.0053)
Avg. past output authors	0.0124*	0.0001	0.0334***
	(0.0064)	(0.0016)	(0.0073)
Novelty	-0.0058		-0.0522**
	(0.0179)		(0.0207)
Observations	66,823	66,829	66,829
R-squared	0.7110	0.5831	0.5990
Year dummies	YES	YES	YES
Fields shares	YES	YES	YES
Journals FE	NO	YES	YES
Team dummies	YES	YES	YES

Notes: The sample is restricted to articles published by a research team (and sole authors) with at least two publications (TFE). The benchmark category is specialist journals. In columns 1 the dependent variable is the journal quality impact factor in logs. In column 2 the dependent variable is citations in $log(x+1)$. *Pages* is the number of pages of the article; *Number of authors==2, 3, >3* are dummy variables if the number of authors publishing the article is 2, 3, or more than 3, respectively; *Team specialization* is a herfindhal index obtained using the shares of past publications in different fields in economics, as defined by the first digit of the JEL codes; *Past output team* is the number of papers adjusted by quality that the research team has published together in the past; *Avg. past output authors* is the average number of papers adjusted by quality published by the authors of the research team in the past; *Novelty* is an index of the novelty of the article obtained measuring the atypicality of the keywords of the article.The F-test interaction terms (p-value) is a test on the joint significance on the interacted terms between the year dummies and title length. All the regressions use clustered standard errors at the research team level*** $p<0.01$, ** $p<0.05$, * $p<0.1$