Uncertainty and the ranking of economics journals

Johan Lyhagen[1] · Per Ahlgren[1]

Abstract
Journal rankings often show significant changes compared to previous rankings. This gives rise to the question of how well estimated the rank of a journal is. In this contribution, we consider uncertainty in a ranking of economics journals. We use the invariant method of Pinski and Narin to rank the journals. We propose an uncertainty measure, which is based on a bootstrap approach. The measure is the average absolute change in rank, which we see as a reasonable uncertainty measure regarding rankings. We further calculate, based on the bootstrap method, 95% confidence interval for the observed values of the invariant method. We show that ranks of the highest, as well as the lowest, ranked journals are well estimated, while there is a high degree of uncertainty regarding the rank of many mid-ranked journals. The distribution of the underlying measure is useful for identifying groups of journals that are more or less of the same quality (from the point of view of the invariant measure). The journal with the highest observed value of the invariant measure, *Journal of Political Economy*, has the best performance and constitutes a singleton, whereas *Quarterly Journal of Economics* and *Econometrica* form the next group (there is a slight overlap between the two with respect to confidence intervals). The journals ranked between about 190–230 form another group in which there are no major quality differences between the journals, as the confidence intervals are overlapping.

Keywords Bootstrapping · Economic journals · Invariant method · Ranking · Uncertainty

Introduction

Nowadays, there is no lack of rankings of entities related to research. Examples of such entities are scientific journals and universities. For the latter, there are many ways to rank them. Academic Ranking of World Universities (more known as the Shanghai ranking), QS World University Ranking and THE World University Rankings are the most well-known university rankings. These rankings differ in indicators used and the weighting of the indicators, and their results have been shown to be only moderately correlated (Olcay and Bulu

✉ Johan Lyhagen
    johan.lyhagen@statistics.uu.se

    Per Ahlgren
    per.ahlgren@uadm.uu.se

[1] Department of Statistics, Uppsala University, Uppsala, Sweden

2017; Mussard and James 2018). CWTS Leiden Ranking, which uses an advanced bibliometric methodology and which do not rely on reputation measurement, avoids a composite indicator of university performance, and this constitutes a fundamental difference between this ranking and the three more well-known ones. University rankings have been exposed for various criticism, though. Regarding Academic Ranking of World Universities, Billaut et al. (2010) concluded that all criteria used by the designers of the ranking are only loosely connected with what they intended to capture. High volatility from year to year has been observed for THE World University Ranking (Bookstein et al. 2010), as well as high volatility in the ranking of universities after the 30th position for Academic Ranking of World Universities and THE World University Rankings (Saisana et al. 2011).

One main factor in the reputation of universities is the quality of the journals in which the researchers publish their work. Hence, there is also a great interest in ranking academic journals.

One can distinguish between expert-based and citation-based approaches to the ranking of journals. In some countries, for instance Norway, the allocation of university funding from the government partly depends on researcher driven quality classifications of journals (and of publishers) (Sivertsen 2010). Publications in journals assigned to the highest level yield a higher amount of funding. In this case, then, an expert-based approach to the ranking of journals is used. By contrast, in Sweden (the country where the authors of this paper are active), some universities allocate funding to their units (like departments) on the basis on a field normalized counterpart to the Journal Impact Factor (JIF), the latter a citation-based indicator that has been much used during a long time to evaluate academic journals (Garfield 2006). For each considered journal and publication year, a field normalized indicator value is calculated, operations that give rise to an (implicit) ranking of the journals. Then, for a given university unit, an average over the scores of the journals of the publications of the unit is obtained. In these Swedish cases, clearly, citation-based approaches to the ranking of journals are used. Research shows that there are differences between the expert-based and the citation-based approaches to journal rankings but the two approaches have been shown to be moderately or highly correlated (e.g. Serenko and Dohan 2011; Ahlgren and Waltman 2014; Haddawy et al. 2016; Saarela et al. 2016). However, Rigby et al. (2015), dealing with agricultural and environmental economics, did not find a significant correlation between journal scores based on expert assessments and the impact factors of the journals.

It is possible to combine expert-based and citation-based approaches to journal rankings. We have not found much earlier research on such hybrid approaches, however. Chen and Chen (2011) proposed a methodology for journal ranking that integrates experts' assessments with scores obtained by Google's PageRank (Brin and Page 1998), whereas Tsai et al. (2014) combined the two approaches by use of the Borda count approach (Reilly 2002).

There has been quite much research on the development of citation-based journal ranking approaches. One of the simplest approaches is to sum the number of citations from journals in the same field, for example as in Bush et al. (1974). This method has the obvious drawback that journals with a larger amount of papers will have more citations and, hence, will be ranked higher. The abovementioned JIF divides the number of citations received by a journal by the number of published papers in the journal such that a citation per paper indicator is obtained. The JIF can be attributed to Gross and Gross (1927) but Garfield (1955) also made important contributions (see the historical accounts of Archambault and Larivière (2009) and Garfield (2006)). Pinski and Narin (1976), in a seminal paper, proposed the first iterative algorithm, the invariant method,

for valuing citations. The invariant method is based on the idea that citations from higher-cited journals are more valuable than citations from lower-cited journals. Several other approaches, all based on citation weighting, were later developed with respect to journal ranking (e.g. Bergstrom 2007; Bollen et al. 2006; Giuffrida et al. 2019; Guerrero-Bote and Moya-Anegón 2012; Kalaitzidakis et al. 2011; Liebowitz and Palmer 1984). With regard to these examples, Liebowitz and Palmer (1984) used their proposed measures to rank economics journals. However, their measures favor journals that publish reviews. This is related to the fact that reviews are on average more frequently cited than non-review articles. Examples of rankings of economics journals using the invariant method include the rankings in Kodrzycki and Pingkang (2006) and Ritzberger (2008). For a somewhat more extensive account of the history of citation weighting, see Giuffrida et al. (2019).

It has been shown that also various citation-based approaches to journal rankings are correlated. For instance, Harzing and van der Wal (2009) and Garcia-Romero et al. (2016) found rank correlations in the range 0.67–0.91, whereas Bornmann et al. (2018) reported rank correlations in the range 0.89–0.99 for the journal indicators outlined in Google Scholar. Now, this seems to imply that the citation-based ranking approach would not be of any great importance. However, although the correlation is high, there can be a substantial difference in rank for individual journals. Liebowitz and Palmer (1984) observed substantial differences for individual journals when the used citation-based indicator was adjusted for journal size (in terms of total number of characters published by a journal): *Economic Journal* dropped from rank 7 to rank 23, *Oxford Economic Papers* from rank 12 to rank 42, and *Review of Economics and Statistics* from rank 5 to rank 16. Garcia-Romero et al. (2016) observed that *Journal of Economic Literature* would rank 2 using the Free Disposal Hull method (Deprins et al. 1984) but 13 using the KMS method (Kalaitzidakis et al. 2011). It was further observed that *Journal of Economic Growth* would rank 6 and 26, whereas *Economics and Politics* would rank 11 and 45 using the same two methods.

To our best knowledge, only a few papers have considered ranking uncertainty in economics. Lubrano et al. (2003) considered ranking uncertainty when ranking economics departments. Oswald (2007) noted that it is better to publish the best (in terms of citation counts) paper in a medium-ranked journal compared to publish the four worst papers in an elite journal. Wall (2009) noticed that the distribution of citations overlap between journals, implying a high degree of uncertainty in the ranking of journals. Halkos and Tzeremes (2011) used data envelopment analysis (DEA) to evaluate economics journals. They derived efficiency scores, and bootstrap confidence intervals of these scores, to measure the uncertainty. Stern (2013), who dealt with ranking of economics journals, took uncertainty into account for the (five-year) JIF in terms of confidence intervals. Significance testing was applied as well, and which journals that were significantly better than others in terms of the JIF was displayed. Hudson (2013) predicted the Keele journal classification by regression methods and measured the probability of correct classification. The robustness of rankings of economics journals was studied by Seiler and Wohlrabe (2014), where the rankings were based on the JIF, by considering the skewness of the distribution of citations to the articles in a journal. Horrace and Parmeter (2017) revisited the Stern study (2013) referred to above, but performed multivariate inference to control the overall error rate of the testing procedure. Confidence intervals around values of the $PP_{top\ 10\%}$ indicator, i.e. the share of papers that belong to the 10% most frequently cited papers in a certain subfield

and time period, was used by Bornmann and Wohlrabe (2019). The PP$_{top\ 10\%}$ results were used for assigning economics journals to four impact classes.[1]

We also mention the study by Bradshaw and Brook (2016), even if journals in economics were not taking into account in the study. These authors ranked journals based on mean ranks and with respect to five indicators. Uncertainty in mean rank for the considered journals was dealt with by the use of resampled uncertainty intervals around these ranks.

In this contribution, we consider uncertainty in a ranking of economics journals. The choice of economics as a field of study is motivated by the fact that economics is a field in which ranking of journals, departments and even individuals is common. We use the invariant method of Pinski and Narin (1976) to rank the journals as, theoretically, it is the most appealing method (see the next section). We propose an uncertainty measure, which is based on a bootstrap approach. The measure is the average absolute change in rank, which we see as a reasonable uncertainty measure regarding rankings. We further calculate, based on the bootstrap method, 95% confidence interval for the observed values of the invariant method.[2] We show that the confidence intervals adds information besides only using the rank. It is possible to identify groups of journals with practically equal performance, as well as identifying journals that substantially parts from their neighboring journals. Basically, if two confidence intervals overlap, this would indicate that the corresponding two journals are not significantly different. So far as we know, the current paper is the first to address uncertainty of journal rankings, where the rankings are based on the invariant method.

The remainder of the paper is structured as follows. The next section treats data and methods. The invariant method, as well as the bootstrap approach used, are described in the section. The results are given in the third section, whereas the results are discussed, and conclusions put forward, in the last section of the paper.

Data and methods

There are many methods when it comes to ranking journals. Each method has its advantages and disadvantages. In this section, we discuss the invariant method of Pinski and Narin (1976). The name origins from that the method is invariant to citation intensities but it also has some other interesting properties. Palacios-Huerta and Volij (2004) proved, using an axiomatic approach, that the invariant method satisfies four important properties: invariance to reference intensity (the average number references per paper with regard a journal), weak homogeneity, weak consistency and invariance to splitting of journals. The invariance to reference intensity means that each journal has one vote and the cited journals share this vote proportional to the proportion of citations they receive. Weak homogeneity says that the relative ranking between the two journals, where the two journals have the same number of publications and the same reference intensity, should be directly related to the ratio of how they cite each other. Weak consistency implies that if we consider a subset

---

[1] In Appendix 2 in Bornmann and Wohlrabe (2019), a detailed explanation of how the PP$_{top\ 10\%}$ indicator is calculated in their study is given.

[2] Instead of the bootstrap approach, an asymptotic approximation could have been used to derive the confidence intervals. However, there are two main advantages of the bootstrap compared to an asymptotic approach. First, the bootstrap most often yields intervals that are more reliable. Second, the bootstrap relies on fewer assumptions making it more robust.

of journals and redistribute the citations of the left out journals to the journals we keep according to how the left out journals cite, then the relative ranking should be the same as for the full set of journals. Invariance to splitting of journals means that if we split a journal to more than one journal with the same reference and citation profiles, then the relative ranking of the other journals will be the same. Palacios-Huerta and Volij (2004) proves that the invariant method is the only method that satisfies all four properties. A related method is the impact-adjusted measure of Liebowitz and Palmer (1984), which, as shown by Palacios-Huerta and Volij (2004), satisfies all properties except the invariance to reference intensity property.[3]

The journals are ranked according to the invariant measure. The observed rank of journal i is denoted by r_i. To evaluate the uncertainty, we used a bootstrap approach suggested by Andersson et al. (1998) in the context of ranking hospitals according to short-term mortality risk ratios for inpatients who had received treatment for myocardial infarction. Let c_{ij} denote the number of citations to journal i from journal j. Next, let $c_j = \sum_i c_{ij}$ be the sum of journal j's cited references that point to a journal included in our study. We took the number of references for journal j as given and randomized them according to probabilities defined as follows:

$$p_{ij} = \frac{c_{ij}}{c_j} \tag{1}$$

In other words, we took the number of references in journal j and randomly distributed them according to the proportion received citations for each journal. We then calculated values of the invariant measure for each journal and ranked the journals according to these values. This procedure was repeated B times, which yielded B rankings of the journals based on bootstrapping. Let r_{ib} be the new rank for journal i in replicate b, and $d_{ib} = |r_{ib} - r_i|$ the deviation of the new rank to the observed rank for i in replicate b. As a measure of variability, we calculated the average absolute change in rank for journal i, $\bar{d}_i$, as follows:

$$\bar{d}_i = \sum_{b=1}^{B} \frac{d_{ib}}{B} \tag{2}$$

We used 1,000 as the value of B. $\bar{d}_i$ can be interpreted as the expected random change in the ranking of the ith journal.

The average absolute change in rank over all journals, say C, can be interpreted as the expected random change in the ordering of the journals. C can be defined as follows:

$$C = \sum_{i=1}^{K} \frac{\bar{d}_i}{K} \tag{3}$$

where K is the number of considered journals.

The rankings are based on the invariant measure and their variability are informative. Hence, we also calculate bootstrapped-based 95% confidence intervals as a complement to the average absolute change in rank. We normalized the invariant measure such that the

---

[3] It is also interesting to note that Google's PageRank (Brin and Page 1998) is a version of the invariant method.

journal with the highest observed value of the measure has a value of 1, whereas the journal with the lowest observed value has a value of 0. A confidence interval was then formed for each journal using the bootstrapped values of the invariant measure. The additional information gained is that a journal can have a high average absolute change in rank due to that there are other journals with a similar value of the invariant measure, while another journal may have the same variability in the invariant measure but does not change rank as there are no other journals with similar values of the invariant measure.

The data was obtained from Journal Citation Reports (JCR), provided by Clarivate Analytics, and the publication period in the study is 1998–2013. We used all journals that belonged to the Web of Science subject category *Economics* in year 2013, 332 journals in total. Further, we only considered citations within the set of the 332 journals.

Results

The results are displayed in a number of tables and figures. To start with, we note that the value of C, i.e. of the average absolute change in rank over all journals, is 3.83 with a standard deviation of 3.61. This implies that, on average, the ranks are well estimated but that there are large individual variation in how well the ranks are estimated. In Table 1, the 60 highest ranked journals are shown. The journal names shown in the table, and in Table 2, are abbreviations of the full names. The average absolute change in rank ($\overline{d}_i$) and the 95% confidence intervals are also indicated. The highest ranked journals do not reveal much of surprises with *Journal of Political Economy*, *Quarterly Journal of Economics* and *Econometrica* in top. *Journal of Political Economy* is a clear winner and ranked first in all bootstrap replicates. *Quarterly Journal of Economics* is second in about 91% of the times and *Econometrica* 9% of the times, yielding only a small degree of uncertainty in the ranking between them. This uncertainty is confirmed by the 95% confidence intervals of the invariant measure that do overlap. For the remaining journals, there are some uncertainty in $\overline{d}_i$, but this uncertainty is too low to yield any substantial variability in rank. To exemplify the latter, the journal with the highest average absolute change in rank is *American Economic Journal: Macroeconomics*, which has rank 48 and a $\overline{d}$ value of about 2.3. This implies that a rank in the close neighborhood of 48 can be considered fairly certain.

In Table 2, the journals with the largest $\overline{d}$ values are shown. The 95% confidence intervals are all small and very close to 0, meaning that they are rather uninformative. *Baltic Journal of Economics*, with rank 249, is the journal with highest $\overline{d}$ value (about 30). At second place with an average rank change of nearly 26 positions is *Journal of Consumer Affairs* (rank 145). Then it is a rather big step down to the third journal, which is *Asian Economic Policy Review* (rank 217), with a $\overline{d}$ value of 15.5. Below that, there is a fair amount of journals with a $\overline{d}$ value in the interval 10–15. The journal with the 60th highest $\overline{d}$ value is *Spatial Economic Analysis* (rank 233; $\overline{d}$ is about 5.7).

In Figs. 1–3, the results for the average absolute change in rank is displayed, ordered by rank. In Fig. 1, the 100 highest ranked journals are shown. On the *y*-axis we have the rank $\pm\overline{d}$, i.e. the average absolute change in rank. We see that the variability for the first about almost 60 highest ranked journals are negligible, while amongst the top 100 there are a few with higher variability. Notably *Quantitative Economics* (ranked 68) with a $\overline{d}$ value of almost 11. This is also seen in the 95% confidence interval that is considerably larger than for the surrounding journals (Table 2; Fig. 4). Compared to the highest ranked journals, the mid-ranked (position 101–200) have higher variability in rank (Fig. 2). As noted

Table 1 Rank, average absolute change in rank ($\bar{d}$) and the 95% confidence interval for the normalized invariant measure for the 60 highest ranked journals

Rank	Journal	Avg abs change	Lower 95	Upper 95
1	J POLITE CON	0.000	1.000	1.000
2	Q J ECON	0.091	0.817	0.843
3	ECONOMETRICA	0.091	0.807	0.832
4	BROOKINGS PAP ECO AC	0.001	0.566	0.614
5	REV ECON STUD	0.001	0.539	0.559
6	J FINANC	0.000	0.480	0.504
7	J ECON LIT	0.000	0.422	0.445
8	AM ECON REV	0.000	0.292	0.300
9	RAND J ECON	0.000	0.242	0.257
10	J ECON THEORY	0.538	0.186	0.196
11	J MONETARY ECON	0.550	0.186	0.195
12	J ECON PERSPECT	0.482	0.181	0.192
13	J LABOR ECON	0.384	0.178	0.191
14	INT ECON REV	0.534	0.161	0.170
15	J ECON GROWTH	0.920	0.155	0.174
16	REV ECON STAT	0.573	0.159	0.166
17	J LAW ECON	0.560	0.152	0.165
18	J FINANC ECON	0.264	0.152	0.159
19	REV FINANC STUD	0.055	0.145	0.156
20	J HUM RESOUR	0.611	0.125	0.134
21	ECON J	0.527	0.125	0.132
22	J BUS ECON STAT	0.486	0.122	0.132
23	J ECONOMETRICS	0.000	0.113	0.120
24	J PUBLIC ECON	0.539	0.092	0.097
25	J LAW ECON ORGAN	0.884	0.089	0.099
26	J INT ECON	0.665	0.089	0.095
27	REV ECON DYNAM	0.911	0.087	0.096
28	J EUR ECON ASSOC	0.422	0.084	0.093
29	GAME ECON BEHAV	1.027	0.079	0.085
30	THEOR ECON	1.594	0.071	0.092
31	J APPL ECONOMET	0.984	0.076	0.086
32	J IND ECON	0.629	0.074	0.083
33	EUR ECON REV	0.289	0.073	0.078
34	J RISK UNCERTAINTY	0.622	0.063	0.072
35	ECONOMET THEOR	0.784	0.063	0.069
36	ECONOMICA	0.870	0.063	0.069
37	J ECON HIST	1.133	0.059	0.066
38	WORLD BANK ECON REV	1.110	0.059	0.066
39	EXP ECON	1.404	0.057	0.066
40	J FINANC QUANT ANAL	0.818	0.058	0.062
41	ECON POLICY	1.268	0.055	0.063
42	J MONEY CREDIT BANK	0.882	0.056	0.061
43	INT J GAME THEORY	0.888	0.054	0.061
44	J ACCOUNT ECON	0.317	0.050	0.055

Table 1 (continued)

Rank	Journal	Avg abs change	Lower 95	Upper 95
45	J MATH ECON	0.616	0.047	0.052
46	AM ECON J-APPL ECON	1.268	0.044	0.055
47	J DEV ECON	0.981	0.045	0.049
48	AM ECON J-MACROECON	2.337	0.039	0.053
49	ECONOMET REV	1.619	0.042	0.049
50	J URBAN ECON	1.164	0.043	0.047
51	J ECON MANAGE STRAT	1.577	0.042	0.048
52	OXFORD ECON PAP	1.176	0.042	0.046
53	ECON THEOR	1.332	0.038	0.041
54	ECON INQ	1.003	0.037	0.041
55	ANNU REV ECON	2.144	0.032	0.046
56	J FINANC ECONOMET	1.560	0.033	0.043
57	SCAND J ECON	0.891	0.035	0.039
58	J ENVIRON ECON MANAG	0.668	0.033	0.036
59	EXPLOR ECON HIST	1.360	0.030	0.036
60	CAN J ECON	1.519	0.030	0.034

above, *Journal of Consumer Affairs* at rank 145 has a $\bar{d}$ value of about 26. For the lowest ranked journals, the average mean absolute change in rank seems to be lower and getting lower with lower rank (Fig. 3). An exception is the most volatile journal of all, *Baltic Journal of Economics*. The general picture is that the rank of the highest and lowest journals are well estimated when using the invariant method, while for the mid-ranked journals there is considerably more uncertainty.

In Figs. 4–6, the values of the invariant measure and the corresponding bootstrapped 95% confidence intervals are indicated. Figure 4 shows that the highest ranked 10 journals are considerably better compared to the rest and that there are notably differences within the first 10 journals as well. The highest ranked journal, *Journal of Political Economy*, is normalized to have a value of one, while the second and third journals (*Quarterly Journal of Economics* and *Econometrica*) have values of about 0.8. The fourth and fifth journals have values around 0.55, while the sixth has a value of about 0.48 (see also Table 1). The confidence intervals are rather large, but as there are large differences in the observed values of the invariant measure there is very little overlap in intervals. The latter implies high certainty in rank. In Fig. 4, we have zoomed the confidence intervals for journals ranked 10–25. It can be seen that there are three groups within these journals. The first group consists of journals with ranks 10–13, the second of journals with ranks 14–19 and the third of journals with ranks 20–23. The journals at rank 24 and 25 seems to belong to the journals with higher rank. There are then subgroups of journals in which the journals are more or less equal. Notice that this information cannot be gained from observing the corresponding $\bar{d}$ values in Table 1.

For the mid-ranked journals (Fig. 5), the confidence intervals are wide in comparison to the value of the invariant measure, and there are no clear groups of journals. The same applies to a majority of the lowest ranked journals (Fig. 6): wide confidence intervals for journals up to rank about 290 and no clear groups of journals. For the journals above rank 290, the confidence intervals are more narrow implying that their rank is better estimated.

Table 2 Rank, average absolute change in rank ($\bar{d}$) and the 95% confidence interval for the normalized invariant measure for the 60 journals with the largest average absolute change in rank

Rank	Journal	Avg abs change	Lower 95	Upper 95
249	BALT J ECON	30.180	0.000	0.002
145	J CONSUM AFF	25.873	0.001	0.012
217	ASIAN ECON POLICY R	15.469	0.001	0.003
276	CEPAL REV	14.757	0.000	0.001
227	ASIAN ECON J	14.180	0.001	0.002
277	ACTA OECON	14.175	0.000	0.001
140	CLIOMETRICA	13.246	0.003	0.009
271	SINGAP ECON REV	13.118	0.000	0.001
205	ECON SOC REV	12.890	0.001	0.003
162	ANN ECON FINANC	12.514	0.002	0.007
187	HITOTSUB J ECON	12.404	0.001	0.004
243	KOREAN ECON REV	12.227	0.000	0.002
180	J APPL ECON	12.093	0.001	0.004
214	ASIAN ECON PAP	11.825	0.001	0.003
228	REV ECON POLIT	11.706	0.001	0.002
253	ASIA-PAC J ACCOUNT E	11.534	0.000	0.001
201	ECON J WATCH	11.501	0.001	0.003
281	ESTUD ECONOMIA	11.480	0.000	0.001
139	PORT ECON J	11.141	0.004	0.009
270	SERIES-J SPAN ECON	11.047	0.000	0.001
68	QUANT ECON	10.804	0.013	0.049
218	J ECON INTERACT COOR	10.430	0.001	0.002
267	ASIAN-PAC ECON LIT	10.263	0.000	0.001
258	CHINA WORLD ECON	9.737	0.000	0.001
194	AUST ECON HIST REV	9.588	0.002	0.003
260	AUST ECON REV	8.910	0.000	0.001
169	N AM J ECON FINANC	8.900	0.002	0.005
251	J MEDIA ECON	8.858	0.000	0.001
254	J BEHAV FINANC	8.566	0.000	0.001
273	REV HIST INDUST	8.421	0.000	0.001
236	EMERG MARK REV	8.335	0.001	0.002
136	J CULT ECON	8.309	0.004	0.008
210	J COMPET LAW ECON	8.078	0.001	0.002
157	REV DERIV RES	7.847	0.003	0.006
248	J INT FINANC MARK I	7.644	0.001	0.001
198	PAC ECON REV	7.634	0.002	0.003
215	WORLD TRADE REV	7.607	0.001	0.002
197	INT LABOUR REV	7.566	0.002	0.003
186	J PENSION ECON FINAN	7.527	0.002	0.003
199	RECH ECON LOUVAIN	7.468	0.002	0.003
200	GENEVA RISK INS REV	7.416	0.002	0.003
196	DEV ECON	7.371	0.002	0.003
158	INT J ECON THEORY	7.339	0.004	0.006
288	HIST ECON IDEAS	7.301	0.000	0.001

Table 2 (continued)

Rank	Journal	Avg abs change	Lower 95	Upper 95
229	POST-SOV AFF	7.297	0.001	0.002
262	ANNU REV RESOUR ECON	7.091	0.000	0.001
146	COMPUT ECON	7.086	0.004	0.007
153	GER ECON REV	7.076	0.004	0.006
160	ECON HUM BIOL	7.050	0.004	0.006
225	HACIENDA PUBLICA ESP	6.962	0.001	0.002
264	ECONOMICS-KIEL	6.709	0.000	0.001
141	ECON GOV	6.561	0.004	0.007
246	S AFR J ECON	6.550	0.001	0.001
192	ECONOMIST-NETHERLAND	6.285	0.002	0.003
240	EMPIRICA	6.171	0.001	0.001
211	B INDONES ECON STUD	5.886	0.001	0.002
304	ECON LABOUR RELAT RE	5.834	0.000	0.000
148	REV HIST ECON	5.739	0.004	0.006
261	EUROPE-ASIA STUD	5.731	0.000	0.001
223	SPAT ECON ANAL	5.720	0.001	0.002
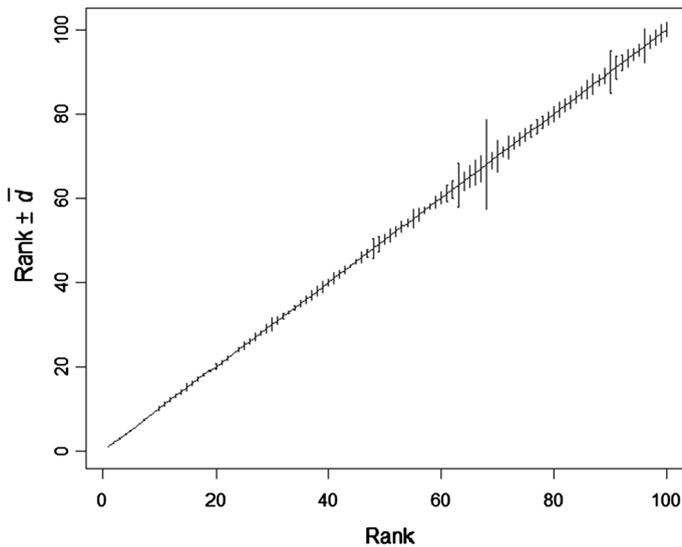
Fig. 1 The 100 highest ranked journals and their average absolute (bootstrapped) change in rank
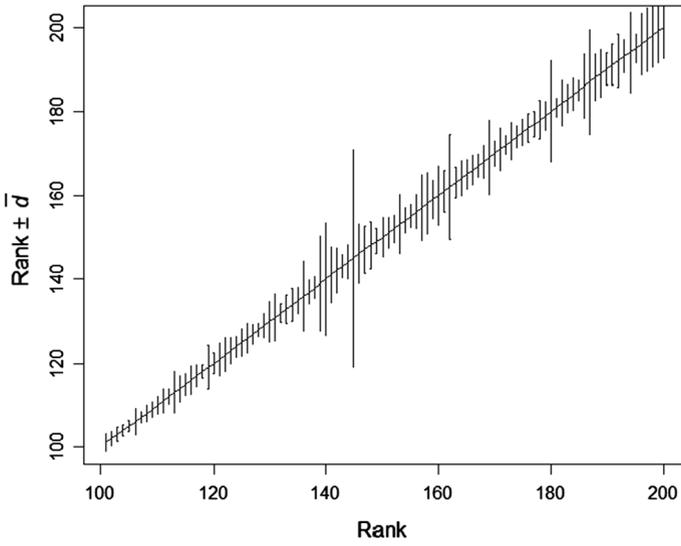
Fig. 2 The 100 mid-ranked (positions 101-200) journals and their average absolute (bootstrapped) change in rank
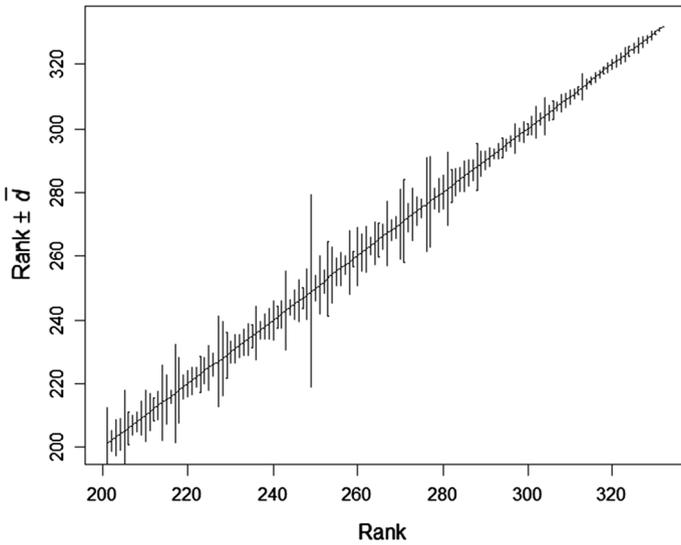
Fig. 3 The 132 lowest ranked journals (positions 201–332) and their average absolute (bootstrapped) change in rank
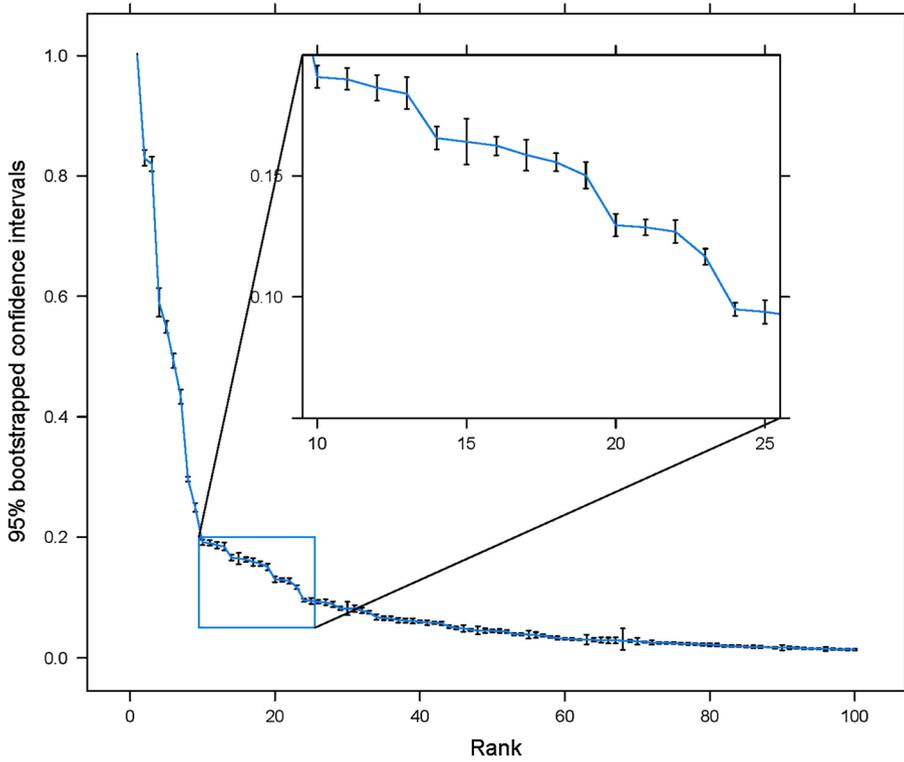
Fig. 4 The 100 highest ranked journals and bootstrapped 95% confidence intervals for the invariant measure, with zoom for journals at rank 10–25
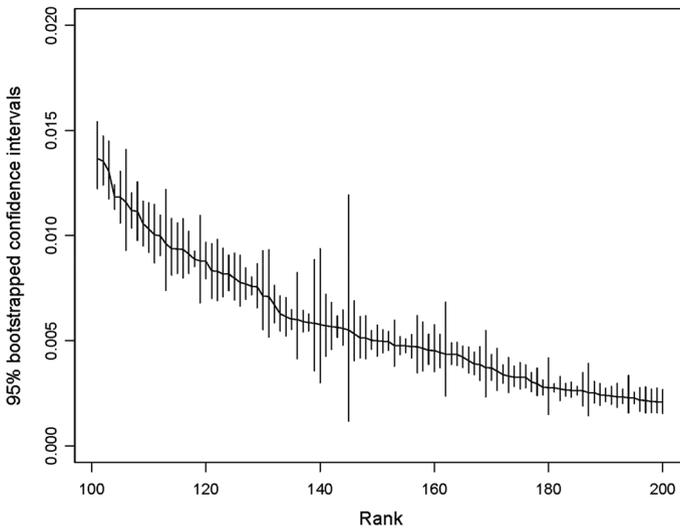

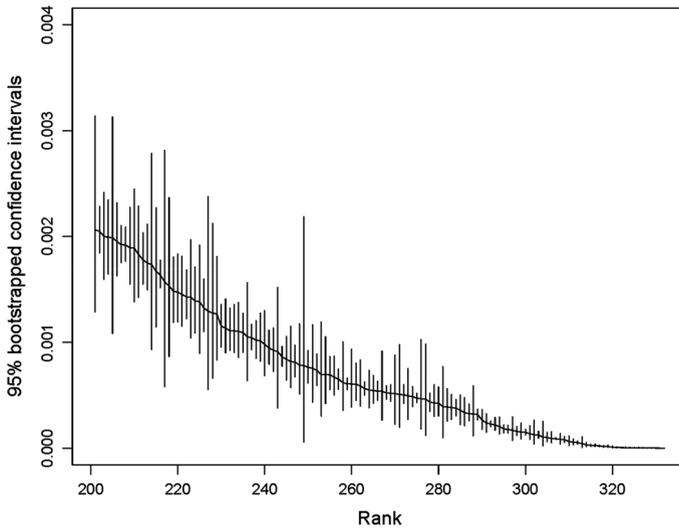**Fig. 5** The 100 mid-ranked journals and bootstrapped 95% confidence intervals for the invariant measure

Fig. 6 The 132 lowest ranked journals and bootstrapped 95% confidence intervals for the invariant measure

We can be quite certain that the lowest ranked journals actually belong to the group of low ranked journals.

Discussion and conclusions

Publishing in highly ranked scientific journals are not only important for the academic career of the individual researcher, but it is also important for academic institutions as better reputation makes it easier to attract good researchers. Moreover, it is not unusual that publication data is used to allocate financial resources to departments, where publications in higher ranked journals yields more resources. Regarding university rankings, briefly treated in the beginning of this paper, these are most often partly based on publication data. It is possible for higher ranked universities to have higher tuition fees compared to lower ranked ones.

Overall, there are many reasons for an interest in the ranking of scientific journals. The focus of this paper is the ranking of economic journals. Economics is a field in which ranking of journals, departments and even individuals is common. Our main contribution is to propose an uncertainty measure, the average absolute change in rank, which is based on a bootstrap approach. The bootstrap is a robust approach that assumes less than competing asymptotic approaches. We also consider the distribution of the underlying ranking measure, here the invariant measure of Pinski and Narin (1976), to gain additional useful information. We show that ranks of the highest, as well as the lowest, ranked journals are well estimated, while there is a high degree of uncertainty regarding the rank of many mid-ranked journals. Here, the uncertainty measure clearly shows that the highest and lowest ranked journals cannot be considered as outliers due to randomness as their ranks are well estimated. The distribution of the underlying measure is useful for identifying groups of journals that are more or less of the same quality (from the point of view of the invariant measure). The journal with the highest observed value of the invariant measure, *Journal*

of Political Economy, has the best performance and constitutes a singleton, whereas *Quarterly Journal of Economics* and *Econometrica* form the next group (there is a slight overlap between the two with respect to confidence intervals). The journals ranked between about 190 to 230 form another group in which there are no major quality differences between the journals, as the confidence intervals are overlapping (and the uncertainty of the ranks are high).

In this work, we have used the invariant method as the ranking method. Regarding future research, it would be interesting to study if ranking uncertainty is dependent on ranking method. It might be the case that there are ranking methods that are better than the invariant method from this perspective, i.e. methods associated with less ranking uncertainty compared to the invariant method.

References

Ahlgren, P., & Waltman, L. (2014). The correlation between citation-based and expert-based assessments of publication channels: SNIP and SJR vs. Norwegian quality assessments. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2014.09.010.

Andersson, J., Carling, K., & Mattson, S. (1998). Random ranking of hospitals is unsound. *CHANCE*. https://doi.org/10.1080/09332480.1998.10542106.

Archambault, É., & Larivière, V. (2009). History of the journal impact factor: Contingencies and consequences. *Scientometrics*. https://doi.org/10.1007/s11192-007-2036-x.

Bergstrom, C. T. (2007). Eigenfactor: Measuring the value and prestige of scholarly journals. *College & Research Libraries News*. https://doi.org/10.5860/crln.68.5.7804.

Billaut, J.-C., Bouyssou, D., & Vincke, P. (2010). Should you believe in the Shanghai ranking? *Scientometrics*. https://doi.org/10.1007/s11192-009-0115-x.

Bollen, J., Rodriguez, M. A., & van de Sompel, H. (2006). Journal status. *Scientometrics*. https://doi.org/10.1007/s11192-006-0176-z.

Bookstein, F. L., Seidler, H., Fieder, M., & Winckler, G. (2010). Too much noise in the times higher education rankings. *Scientometrics*. https://doi.org/10.1007/s11192-010-0189-5.

Bornmann, L., Butz, A., & Wohlrabe, K. (2018). What are the top five journals in economics? A new meta-ranking. *Applied Economics*. https://doi.org/10.1080/00036846.2017.1332753.

Bornmann, L., & Wohlrabe, K. (2019). Normalisation of citation impact in economics. *Scientometrics*. https://doi.org/10.1007/s11192-019-03140-w.

Bradshaw, C. J. A., & Brook, B. W. (2016). How to rank journals. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0149852.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems, 30*(1–2), 107–117.

Bush, W. C., Hamelman, P. W., & Staaf, R. J. (1974). A quality index for economic journals. *The Review of Economics and Statistics, 56*(1), 123–125.

Chen, Y.-L., & Chen, X.-H. (2011). An evolutionary PageRank approach for journal ranking with expert judgements. *Journal of Information Science*. https://doi.org/10.1177/0165551511402421.

Deprins, D., Simar, L., & Tulkens, H. (1984). Measuring labor efficiency in post offices. In M. Marchand, P. Pestieau, & H. Tulkens (Eds.), *The performance of public enterprises: concepts and measurement*. Amsterdam: North-Holland.

Garcia-Romero, A., Santin, D., & Sicilia, G. (2016). Another brick in the wall: a new ranking of academic journals in economics using fdh. *Scientometrics*. https://doi.org/10.1007/s11192-016-1843-3.

Garfield, E. (1955). Citation indexes for science. *Science, 122*(3159), 108–111.

Garfield, E. (2006). The history and meaning of the journal impact factor. *Journal of the American Medical Association*. https://doi.org/10.1001/jama.295.1.90.

Giuffrida, C., Abramo, G., & D'Angelo, C. A. (2019). Are all citations worth the same? Valuing citations by the value of the citing items. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2019.02.008.

Gross, P. L. K., & Gross, E. M. (1927). College libraries and chemical education. *Science, 66*(1713), 385–389.

Guerrero-Bote, V. P., & Moya-Anegón, F. (2012). A further step forward in measuring journals' scientific prestige: The SJR2 indicator. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2012.07.001.

Haddawy, P., Hassan, S.-U., Asghar, A., & Amin, S. (2016). A comprehensive examination of the relation of three citation-based journal metrics to expert judgment of journal quality. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2015.12.005.

Halkos, G. E., & Tzeremes, N. G. (2011). Measuring economic journals' citation efficiency: a data envelopment analysis approach. *Scientometrics*. https://doi.org/10.1007/s11192-011-0421-y.

Harzing, A.-W., & van der Wal, R. (2009). A google scholar h-index for journals: An alternative metric to measure journal impact in economics and business. *Journal of the American Society for Information Science and Technology*. https://doi.org/10.1002/asi.20953.

Horrace, W. C., & Parmeter, C. F. (2017). Accounting for multiplicity in inference o economics journal rankings. *Southern Economic Journal*. https://doi.org/10.1002/soej.12219.

Hudson, J. (2013). Ranking journals. *The Economic Journal, 123*(570), 202–222. https://doi.org/10.1111/ecoj.12064.

Kalaitzidakis, P., Mamuneas, T. P., & Stengos, T. (2011). An updated ranking of academic journals in economics. *Canadian Journal of Economics*. https://doi.org/10.1111/j.1540-5982.2011.01683.x.

Kodrzycki, Y. K., & Pingkang, Y. (2006). New Approaches to Ranking Economics Journals. *The B.E. Journal of Economic Analysis & Policy*. https://doi.org/10.1515/1538-0645.1520.

Liebowitz, S. J., & Palmer, J. P. (1984). Assessing the relative impacts of economics journals. *Journal of Economic Literature, 22*(1), 77–88.

Lubrano, M., Bauwens, L., Kirman, A., & Protopopescu, C. (2003). Ranking economics departments in Europe: A statistical approach. *Journal of the European Economic Association*. https://doi.org/10.1162/154247603322752575.

Mussard, M., & James, A. P. (2018). Engineering the global university rankings: gold standards, limitations and implications. *IEEE ACCESS*. https://doi.org/10.1109/ACCESS.2017.2789326.

Olcay, G. A., & Bulu, M. (2017). Is measuring the knowledge creation of universities possible?: A review of university rankings. *Technological Forecasting & Social Change*. https://doi.org/10.1016/j.techfore.2016.03.029.

Oswald, A. J. (2007). An examination of the reliability of prestigious scholarly journals: Evidence and implications for decision-makers. *Economica*. https://doi.org/10.1111/j.1468-0335.2006.00575.x.

Palacios-Huerta, I., & Volij, O. (2004). The measurement of intellectual influence. *Econometrica*. https://doi.org/10.1111/j.1468-0262.2004.00519.x.

Pinski, G., & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing & Management*. https://doi.org/10.1016/0306-4573(76)90048-0.

Reilly, B. (2002). Social choice in the South Seas: electoral innovation and the Borda count in the Pacific Island countries. *International Political Science Review*. https://doi.org/10.1177/0192512102023004002.

Rigby, D., Burton, M., & Lusk, J. L. (2015). Journals, preferences, and publishing in agricultural and environmental economics. *American Journal of Agricultural Economics*. https://doi.org/10.1093/ajae/aau102.

Ritzberger, K. (2008). A ranking of journals in economics and related fields. *German Economic Review*. https://doi.org/10.1111/j.1468-0475.2008.00447.x.

Saarela, M., Kärkkäinen, T., Lahtonen, T., & Rossi, T. (2016). Expert-based versus citation-based ranking of scholarly and scientific publication channels. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2016.03.004.

Saisana, M., d'Hombres, B., & Saltelli, A. (2011). Rickety numbers: Volatility of university rankings and policy implications. *Research Policy*. https://doi.org/10.1016/j.respol.2010.09.003.

Seiler, C., & Wohlrabe, K. (2014). How robust are journal rankings based on the impact factor? Evidence from the economic sciences. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2014.09.001.

Serenko, A., & Dohan, M. (2011). Comparing the expert survey and citation impact journal ranking methods: Example from the field of artificial intelligence. *Journal of Informetrics*. https://doi.org/10.1016/j.joi.2011.06.002.

Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. *ISSI Newsletter, 6*(1), 22–28.

Stern, D. I. (2013). Uncertainty measures for economics journal impact factors. *Journal of Economic Literature*. https://doi.org/10.1257/jel.51.1.173.

Tsai, C.-F., Hu, Y.-H., & Ke, S.-W. G. (2014). A Borda count approach to combine subjective and objective based MIS journal rankings. *Online Information Review*. https://doi.org/10.1108/OIR-11-2013-0253.

Wall, H. J. (2009). Don't get skewed over by journal rankings. *The B.E. Journal of Economic Analysis & Policy*. https://doi.org/10.2202/1935-1682.2280.