Standing on the shoulder of giants: The aspect of free-riding in RePEc rankings

Wohlrabe, Klaus and Meyer, Justus

21 March 2017

Standing on the shoulder of giants: The aspect of free-riding in RePEc rankings

Abstract

RePEc rankings have become a well-established source of information about actual and perceived academic performance of institutions, academic fields and their authors. One essential ingriedient are the impact factors calculated in RePEc which differ from the standard ones. RePEc reports the ratio of the cumulative citations of *all* articles of a journal and the number of listed items. The continuously updated RePEc impact factors account for the whole journal and citation history. This approach give rise to a potential free-riding of authors who profit from journal ranking established in the past. In this paper we demonstrate how the rankings of economists change if one calculates yearly impact factors. The distribution of gains and losses is most pronounced among middle-field ranked authors while the top group shows relative persistence.

Justus Meyer
University of Munich

Klaus Wohlrabe*
Ifo Institute – Leibniz-Institute
for Economic Research
at the University of Munich e.V.
Poschingerstr. 5
81679 Munich, Germany
Phone: +49(0)89/9224-1229
wohlrabe@ifo.de

* Corresponding author.

JEL Code: A12, A14
Keywords: impact factors; rankings; free-riding; RePEc

1 Introduction

RePEc rankings are a well-established source of information about actual and perceived academic performance of institutions, academic fields and their authors. Thereby the different rankings provide a significant mean for evaluating research quality by a single article and author publications based over time. An important feature of the RePEc rankings is their signaling effect. In particular, the authors rank has the potential to reflect the influence of a single author on the fields knowledge accumulation. The relevance of an author's position in the RePEc ranking also stems from the fact that individual publishing performance plays a crucial role in job-market decisions. The number and quality of publications is a key determinant who wins in the race for tenure and promotion. Furthermore, author rankings matter for reputation. As Haucap and Muck (2015) document, public rankings directly influence quality perceptions among economists about journals. Complementary to existing networks an author ranking can also be a crucial determinant for evaluating the quality of potential coauthors. A critical question, therefore, is how specific methodological characteristics shape the ranking outcomes.

Until now robustness checks for rankings in economics have mainly focused on the journal level or regional studies. For example, Costa Vieira (2004) analyses variability of rankings over time and presents different ways of evaluating a journal's impact. More recently Seiler and Wohlrabe (2014) find journal rankings in economics to be surprisingly robust.[1] New studies by Franses (2014) and Anderson and Tressler (forthcomming) review regional rankings. Franses (2014), for example, finds persistence in the top group of Danish economists but decreasing differences over time. This paper contributes to the literature by pointing at a particular shortcoming of the RePEc ranking procedure. The currently applied method significantly biases the perception of authors influence on the process of knowledge creation in economics. We find that there is a potential of free-riding in the impact-factor adjusted author ranking. The ongoing procedure of measuring an author's overall ranking is to adjust backwards for the publishing journal's quality. RePEc currently adjusts by weighing an authors publication by

[1] For a survey on the role of impact factors see Archambault and Larivière (2009).

the publishing journal's current impact factor. Impact factors, however, change over time as documented by Althouse *et al.* (2008). The reweighting of old articles by current impact factors raises the chance of involuntary free-riding. Authors benefit through previous publications in journals that climb the quality ladder over time. Instead, we propose to weight journal papers according to the impact factor during the year an article was published. We can show that the conventional system creates winners and losers compared to the alternative of a yearly impact factor weighting.

The remainder is structured as follows. First, we describe the methodology briefly behind the current RePEc ranking and introduce the data used. The subsequent section provides evidence on the free-riding effect in current author rankings. Finally, we conclude.

2 Data and impact factors in RePEc

We extracted our data from RePEc (Research Papers in Economics, www.repec.org). In economics, RePEc has become an essential source for the spread of knowledge and ranking of individual authors and academic institutions. RePEc is based on the "active participation principle", i.e. that authors, institutions and publishers have to register and to provide information to the network. This approach has the main advantage that an explicit assignment of works and citations to authors and articles is possible. Indeed, the RePEc story has become a success, with more than 45,000 registered authors with listed works and 2,500 journals in economic sciences worldwide as of January 2017. Using a unique identifier, we downloaded all meta-information for more than 1,000,000 journal articles listed in RePEc. This includes the title, the journal, number of authors and citations. Additionally, we restricted ourselves to data from 1980 to 2012, as for later years not all information for all journals were available. There are only few journals listed before 1980 and these also partly incomplete. We also excluded articles with apparently misclassified bibliometric information. All data were downloaded on 01/19/2015.[2] Finally, we have data for 800,577 journal articles published in 1,790

[2]This is the same database which has also been used in Rath and Wohlrabe (2016a), Rath and Wohlrabe (2016b) and Sommer and Wohlrabe (2017).

journals. This allows us to replicate the RePEc ranking for 33,214 authors. These authors are registered in RePEc and have claimed authorship of at least one journal article.

Based on all available bibliographic information within the network, RePEc calculates every month more than 30 different bibliometric indicators for registered authors and institutions. Table 1 provides an overview of these measures. There are five main categories: number of (published) works, citations, citation indices, citing authors, journal pages, and RePEc access statistics. Each of these main categories can be combined with different weighting schemes: simple or recursive impact factors, number of authors and combination of them. For the category 'distinct number of works' different version of a paper are counted only once. Published work is only counted if, first the publisher provides the meta data to RePEc and second, the author assigns this work to his/her account. Table 1 reveals that there is a focus on citations both directly and indirectly. In 14 rankings citations are count with quality and time adjustments. The indirect channel are the different impact factors. Zimmermann (2013) provides a detailed account on the methodology of RePEc.

The most well-known yearly impact factors are provided by Web of Science (WoS) from Thomson Scientific in its Journal Citation Report (JCR). Although they are criticized for a number of reasons, see Glänzel and Moed (2002) for an overview, they still provide a glimpse of the quality of a journal. Focusing on the economic sciences, the JCR impact factors have two major drawbacks: First, the average time for a journal article from publication to peak in citations is not always two years. Furthermore, the publication process in economics is rather slow compared to natural sciences, see Ellison (2002), which leads to the fact that the impact factors are rather small. Second, the impact factors from JCR is restricted to a specific journal list. The subsection 'economics' lists only 334 journals for the JCR 2015. Thus, many citations from other economic journals are potentially missing.[3] RePEc accounts for these two issues: First, citations of articles from the whole journal history available in the network are included. Second, RePEc considers citations from all indexed series. Based on this, impact factors for all listed series are available (journals, working papers and book series). Although impact factors

[3]See Nederhof (2006) for the issue of coverage in the Social Science Citation Index (SSCI), which contains the economics category as a subgroup, for the social sciences.

Table 1: Bibliometric measures in RePEc

		Without weighting	Simple Impact Factor	Recursive Impact Factor	Number of Authors	Number of Authors + Simple Impact Factor	Number of Authors + Recursive Impact Factor
Works	Overall	X					
	Distinct	X	X	X	X	X	X
Citations	Overall	X	X	X	X	X	X
	Discounted by citation year	X	X	X	X	X	X
Citing Authors	Overall	X					
	Weighted by authors rank	X					
Journal Pages		X	X	X	X	X	X
Access via RePEc	Abstract Views	X			X		
	Downloads	X			X		
Indices	h-Index	X					
	Wu-Index	X					
	Euclidian Index	X					
NEP Cites		X					
Strength of Students		X					
Betweenness		X					
Closeness		X					

in RePEc are also restricted to citations from listed series, this list is much larger compared to the economics subcategory in the JCR. Another difference between the standard and the RePEc impact factor is the exclusion of 'self-citations' to prevent 'self-inflation'. Finally, the JCR impact factors are only updated once a year, whereas in RePEc updates on a regular basis. In addition to the standard impact factor, RePEc provides also a recursive impact factor. It gives citations from journals with higher impact larger weights than citations from low-impact journals. In economics this method goes back to Liebowitz and Palmer (1984).

The other side of the coin of the RePEc impact factors is the aspect of potential positive and negative free-riding. The impact factor does not reflect both the history of the journal as well as the citation patterns across time.

3 Economists rankings based on different impact factors

Now we want to investigate how the definition of the impact factor affects the rankings in RePEc. We resemble the ranking 'distinct number of works weighted by simple impact factor' in RePEc. The impact factor is calculated (i) based on *all* listed items for each journal and (ii) based on *yearly* listed items. The latter one reflects both the journal quality at a given point in time and the citation behaviour at that time.

Differences in the placing of single authors can be dramatic as the sample moments in Table 2 illustrate. The mean number of rank changes in the overall sample is around three positions while the sub-samples show a much larger dynamic. Individual outliers cause the extreme values in the overall sample. An exclusion of those does not significantly change the results.[4] In particular, the average change for the top 100 is -12 places, for the top 1000 is -40 places. Among the top 1000 authors ordered by the initial RePEc ranking, the biggest gain is an improvement of 510 places while another author drops by 1600 places. Restricting the sample on the top 100 one finds range from improving by 57 places to a decline of almost 200. The increase of dispersion between overall and yearly based rankings is illustrated in Figure 1. The top group shows relative persistence. On the contrary in the middle of the rank distribution variability is much higher.

At the same time, we find a considerable overlap between both ranking procedures as measured by a Spearman rank correlation coefficient of 0.978.[5] As a matter of the sample size, however, already small deviations can have a significant impact on a substantial number of authors and their individual ranking. The increase in variability can also be seen from Figure 2. It visualises the increasing dispersion of differences in the two ranking methods by overall rank for the subsample of the top 1000 authors. Also here one sees variation in places increasing in a lower original RePEc rank. The more distant an author is from the top, the larger are potential gains and losses. An appealing explanation for this finding would be that variance in impact factors for top journals is lower than for lower ranked journals. Naturally,

[4] Excluding the most extreme 1% (5%) of observations causes a change in the overall sample mean to 20.54 (59.03).

[5] Again results are robust. If the most extreme 1% (5%) of observations were excluded this results in a change of the Spearman rank correlation coefficient to 0.982 (0.989).

economists who tend to be in the top group will publish relatively more in top journals which make them less affected. Another reason could be that those authors publish in the same journals. Hence they are affected by changes in a similar way than their peers at the top. Especially at the lower end of the ranking, publication patterns might be more diverse. As a result, overall variability at the top would be lower compared to the rest.

Table 3 provides a more detailed view of the top 20 economists according to the original RePEc ranking methodology. An adjustment of the ranking method causes partly significant but often rather minor changes. Among the first ranks changes are small. Both ranking methods list Andrei Shleifer, Peter C. B. Phillips and Jean Tirole as the three most influential economists according to our ranking. Some authors loose like Lawrence Summers who even drops out of the top ten. Others climb the ladder such as Robert Barro who gains nine places. The main changes take place in the middle field. That the proposed ranking procedure does not cause dramatic changes in the top group fits well in the picture as drawn from the overall sample. As seen dramatic changes in the ranking are more pronounced in the middle field of the overall ranking.

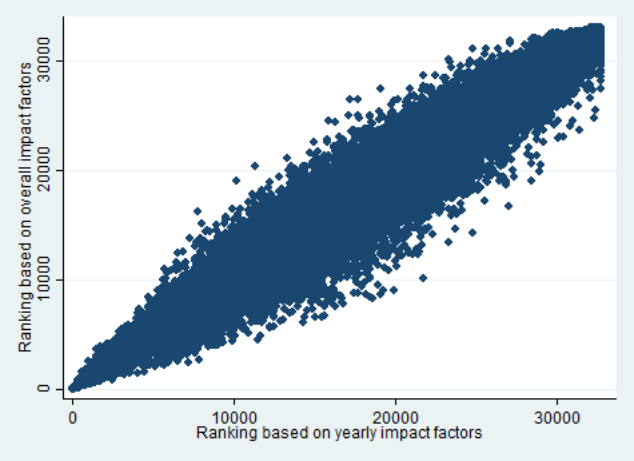
We repeated the whole analysis by adjusting for the number of authors and using the number of journal pages. We obtained quantitatively and qualitatively very similar results.

Our descriptive evidence confirms the hypothesis of a significant potential for free-riding in the current RePEc author ranking. Free-riding can be both intended and unintended. Past work can unjustifiedly feature a higher weight in the authors record if the publishing journal improves its position over time. By that, it can have a significant impact on the perceived position of an author in the field. Thus, authors benefit ex-post from having published in a journal that significantly strengthened its quality over time. Besides, an author who correctly anticipates an uprising journal can profit from the increase in its impact factor over time.

Table 2: Summary of Losses and Gains in Ranking Positions

Sample	Mean	Minimum	Maximum
Overall	2.71	-11604	9254
Top 1000	-40.60	-1600	510
Top 100	-12.21	-194	57

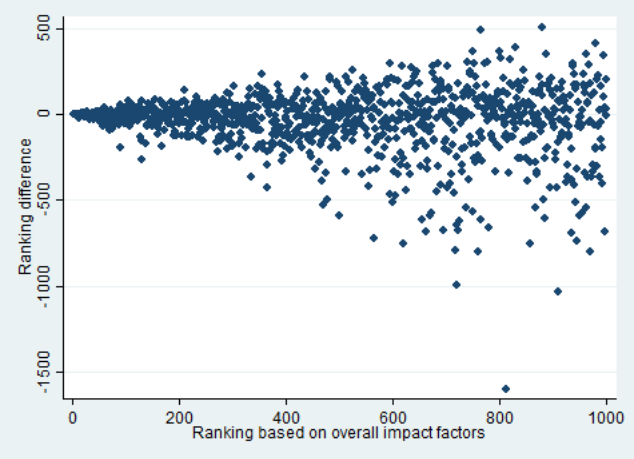
Figure 1: Ranking comparison using overall and yearly impact factors



4 Conclusion

Having collected data for 33,214 authors and their journal publications from the RePEc archive we have been able to demonstrate that methodology of the impact factor matters for author rankings. The current procedure gives rise to an opportunity of free-riding. We showed that significant losses and gains are possible if one adjusts the impact factor weighting by the date an article was published. The distribution of gains and losses is most pronounced among middle-field ranked authors while the top group shows relative persistence.

Figure 2: Rank differences relative to ranking using overall IF for the top 1000



References

Althouse, B. M., West, J. D., Bergstrom, C. T. and Bergstrom, T. (2008). Differences in impact factor across fields and over time. *Journal of the American Society for Information Science and Technology*, **60** (1), 27–34.

Anderson, D. L. and Tressler, J. (forthcomming). Researcher rank stability across alternative output measurement schemes in the context of a time limited research evaluation: the new zealand case. *Applied Economics*.

Archambault, É. and Larivière, V. (2009). History of the journal impact factor: Contingencies and consequences. *Scientometrics*, **79** (3), 635–649.

Costa Vieira, P. C. (2004). Statistical variability of top ranking economics journals impact. *Applied Economics Letters*, **11** (15), 945–948.

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

Franses, P. H. (2014). Trends in three decades of rankings of dutch economists. *Scientometrics*, **98** (2), 1257–1268.

Table 3: Top 20 Authors based on the ranking using the overall impact factor

Economist	Rank - Overall IF	Rank - Yearly IF	Difference
Andrei Shleifer	1	1	0
Peter C. B. Phillips	2	3	-1
Jean Tirole	3	2	1
Daron Acemoglu	4	5	-1
Joseph Stiglitz	5	6	-1
James J. Heckman	6	4	2
Lawrence H. Summers	7	12	-5
Donald W. K. Andrews	8	10	-2
John List	9	26	-17
Elhanan Helpman	10	20	-10
John J. Siegfried	11	34	-23
Mark Rosenzweig	12	22	-10
Jean-Jacques Laffont	13	8	5
Olivier Blanchard	14	13	1
RenÃ© M. Stulz	15	24	-9
Richard Blundell	16	16	0
Timothy Besley	17	9	8
David E. Card	18	17	1
John Y. Campbell	19	14	5
Robert J. Barro	20	11	9

Glänzel, W. and Moed, H. (2002). Journal impact measures in bibliometric research. *Scientometrics*, **53** (2), 171–193.

Haucap, J. and Muck, J. (2015). What drives the relevance and reputation of economics journals? an update from a survey among economists. *Scientometrics*, **103** (3), 849–877.

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

Nederhof, A. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics*, **66** (1), 81–100.

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

— and — (2016b). Trends in economics publications represented by jel categories between 2007 and 2013. *Applied Economics Letters*, **23** (9), 660–663.

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

Sommer, V. and Wohlrabe, K. (2017). Citations, journal ranking and multiple authorships reconsidered: evidence from almost one million articles. *Applied Economics Letters*, **forthcoming**.

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