Article

A Correlation Analysis of Normalized Indicators of Citation

Dmitry M. Kochetkov [1],*

[1] Peoples' Friendship University of Russia; kochetkov_dm@rudn.university
* Correspondence: kochetkovdm@hotmail.com; Tel.: +7-925-865-64-45

Abstract: Recently, more and more countries are entering the global race for university competitiveness. On the one hand, global rankings are a convenient tool for quantitative analysis. On the other hand, their indicators are often difficult to quickly calculate, they often contradict each other. We thought about using widely available indicators for a quick analysis of the University's publication strategy. We opted for the normalized citation indicators available in SciVal analytical tool, i.e. Source Normalized Impact per Paper (SNIP) and Field-Weighted Citation Impact (FWCI). We have demonstrated the possibility of applying the correlation analysis to the impact indicators of a document and a journal on the sample of the social and humanitarian fields at Peoples' Friendship University of Russia. Particular attention was paid to the application of the results in practice.

Keywords: normalized indicators; correlation analysis; Source Normalized Impact per Paper; SNIP; Field-Weighted Citation Impact; FWCI

1. Introduction

Recently, programs for increasing the global competitiveness of universities funded by national and local governments have been launched almost all over the world. In Russia, the project called *5-100* [1] was launched in 2013. At present, 21 leading Russian universities are participating in the project; *Peoples' Friendship University of Russia* (PFUR, Russian "RUDN") is among them. The main goal of the Project is 5 Russian universities entering the TOP-100 of global university rankings. *QS World University Rankings* (QS) [2] and *Times Higher Education World University Rankings (THE)* [3] are among the most popular. Both rankings use the normalized citation indicator for the calculation. Without deepening into the calculation methodology, let's say that they differ significantly. The issues of normalizing citations for subject areas are discussed in the article by Waltmann & van Eck [4] and Waltman [5]. We have asked ourselves which metrics we can use for analysis within the framework of university research management. Accordingly, we decided to limit ourselves to the metrics available in widely distributed analytical packages that do not require large additional computations. Thus, our choice fell on *FWCI* and *SNIP*, which are available in SciVal from Elsevier.

Source Normalized Impact per Paper (SNIP) was introduced by Professor Henk Moed [6] at the *Centre for Science and Technology Studies (CTWS)*, University of Leiden. After receiving a dose of criticism, the indicator was revised [7]. He measured the impact of scientific publications through contextual citation weighting based on the total number of citations in the subject field using Scopus data. Or, as stated by CTWS, "SNIP is correct for differences in citation practices between scientific fields, allowing for more accurate between-fields comparisons of citation impact [8]." SNIP is calculated by dividing the number of citations per paper in the journal by the citation potential in the subject field. The study [9] confirmed that from various journal metrics, SNIP has the greatest correlation with expert ratings (for example, ERA [10]), possibly due to its normalized nature.

Field-Weighted Citation Impact (FWCI) is the ratio between the actual number of citations received by a publication set and the average number of citations received by all other similar publications.

46 The latter is referred to as the expected number of citations. Similar publications refer to the same
47 discipline of the same type and the same age. FWCI is a Snowball Metric [11]. FWCI is measured by
48 dividing the number of citations received by the publication by the average number of citations to
49 publications in the database published in the same year of the same type and within the same subject
50 category. When several publications are considered, the ratio between the actual and average
51 citations for each publication is first calculated. FWCI is calculated then as a mean value. Publications
52 can also be assigned to more than one subject category.

53 The use of quantitative indicators overcomes the subjectivity of peer review although it is
54 incorrect to rely solely on numbers [12]. Ideally, FWCI should be equal to SNIP. In reality, of course,
55 such a situation is unlikely, but it takes place as a direction for improvement of the publication
56 strategy of the university. We decided to focus on the problem area for our university (and for
57 Russian universities in general) in the social sciences and arts & humanities. It's no secret that Russian
58 universities are most often associated with the natural sciences. Besides, given the low citation
59 potential in these areas, only a few citations can affect the value of the normalized indicator rather
60 significantly. Our research question is how to use the SNIP and FWCI indicators together for the
61 purposes of research management and advancement in international rankings in terms of social
62 sciences and humanities. We also tried to elucidate the application of the results of this study in
63 practice.

2. Materials and Methods

65 The method of calculation of normalized citation indicators enables comparative analysis not
66 only with universities but also with university groups as the value is taken not as a sum, but as a
67 mean. Therefore, we decided to compare three objects: Project 5-100 in general, the Chinese *"League
68 of nine" C9* [13] and *Universities of Excellence* (Germany) [14]. The analysis was conducted on 6 subject
69 fields and areas (Arts & Humanities; Business, Management and Accounting; Decision Science;
70 Economics, Econometrics and Finance; Psychology; Social Sciences). We took only documents of the
71 "article" type for analysis due to the following:
72 1. Most often it is in journal articles that the approved original results of scientific research are
73   published;
74 2. It is very difficult to estimate the level of "non-journal" sources because of the lack of data.
75 First, we calculated the Pearson correlation coefficient for these subject fields and areas. The
76 correlation coefficient, like the covariance analysis, characterizes the degree to which the two
77 variables "change together". Unlike the covariance analysis, the correlation coefficient is scaled in
78 such a way that its value does not depend on the units in which the variables of the two
79 measurements are expressed. Any value of the correlation coefficient should be in the range from -1
80 to +1 inclusive. Correlation analysis elucidates whether data sets are associated in magnitude, i.e.
81 higher values of one data set are associated with larger values of the other set (positive correlation)
82 or on the contrary, low values of one set are associated with larger values of the other (negative
83 correlation), or data of two ranges are not related in any way (zero correlation).
84 Next, we constructed a dot diagram for publications in each of the subject areas based on FWCI
85 and SNIP indicators. Then we draw the lines through points 1 (global average citation potential) on
86 the abscissa and y-axis. Thus, we obtained a two-factor matrix, which gave us 4 quadrants (Figure 1).
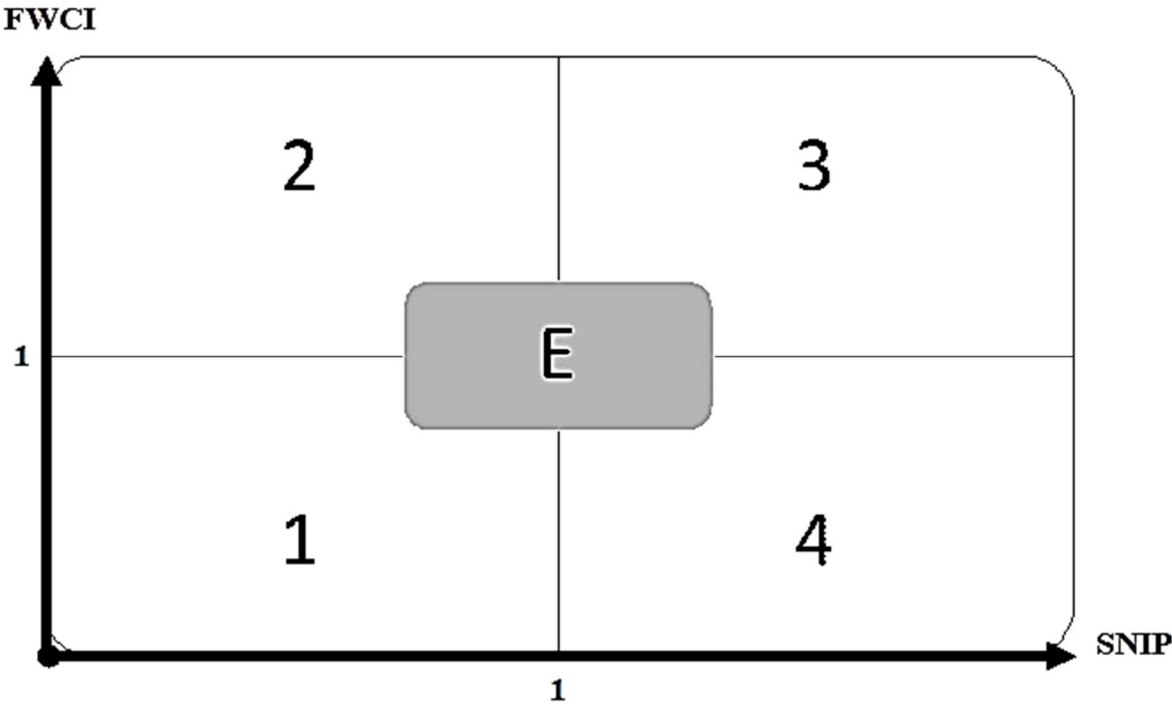
Figure 1. Two-factor matrix for the distribution of articles within the subject field/area.

Source: authors' own development.

Quadrant 1 "Everything is bad." The results of the university research are published in low-impact journals and receive low citation rates. The problem lies in the field of either relevance or importance of the research agenda as such, or in the fact that research in this field is still in its infancy (at the university).

Quadrant 2 "Rising Stars". The quality of publications in this field exceeds the level of the sources in which they are published. It's time for the university to make a shift towards the next level. Often such a situation is related with an underestimation by scientists of their own capabilities.

Quadrant 3 "Everything is fine." High-quality research results are published in high-quality sources. We must strive to ensure that most publications of the university are in this quadrant.

Quadrant 4 "Overestimation." University scientists overestimate the quality of their publications, which does not correspond to the level of sources in which they are published. Despite the review procedure, sometimes such situations arise because of the poor work of the reviewer or editor, or under the influence of the status of the researcher or the university itself. Most likely, the university should revise its research agenda, at least, partly.

The *point E* is the *balance*, the citation level of the document and the source coincide and at the same time correspond to the citation potential in the given subject field/area in a given period of time.

3. Results

Table 1 shows the results of the correlation analysis.

Table 1. Comparative correlation analysis of normalized citation of articles and normalized impact indicator of the source*.

Subject field/area	RUDN		Project 5-100		League C9		Universities of Excellence	
	N**	Corr***	N**	Corr***	N**	Corr***	N**	Corr***
Arts&Humanities	97	-0.01	2743	0.11	1788	0.12	4718	0.33

Business, Management and Accounting	83	0.19	1324	0.32	3763	0.35	2355	0.36
Decision Science	29	0.76	412	0.33	2697	0.28	1444	0.31
Economics, Econometrics and Finance	116	-0.04	2292	0.2	2503	0.39	2473	0.4
Psychology	15	0.63	616	0.36	1742	0.33	6126	0.4
Social Sciences	264	-0.02	5891	0.13	6231	0.4	8763	0.36

110   *Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

111   **Number of observations.

112   ***Correlation coefficient.

113   It became obvious from the table that Peoples' Friendship University of Russia is behind the
114   correlation coefficient from the compared groups of universities (including Project 5-100) in almost
115   all areas of social sciences and humanities. Especially critical situation is in arts & humanities,
116   economics and social sciences in general. High rates for Decision Science and Psychology do not seem
117   valid due to the extremely small sample size. In turn, the universities of Project 5-100 lag behind
118   Chinese and German colleagues. German Universities of Excellence are the leaders in virtually all
119   subject areas. Now let's analyze the situation in each subject area in more detail.

120   *3.1. Arts & Humanities*

121   The first attempt to build a point chart in the Arts & Humanities field was unsuccessful because
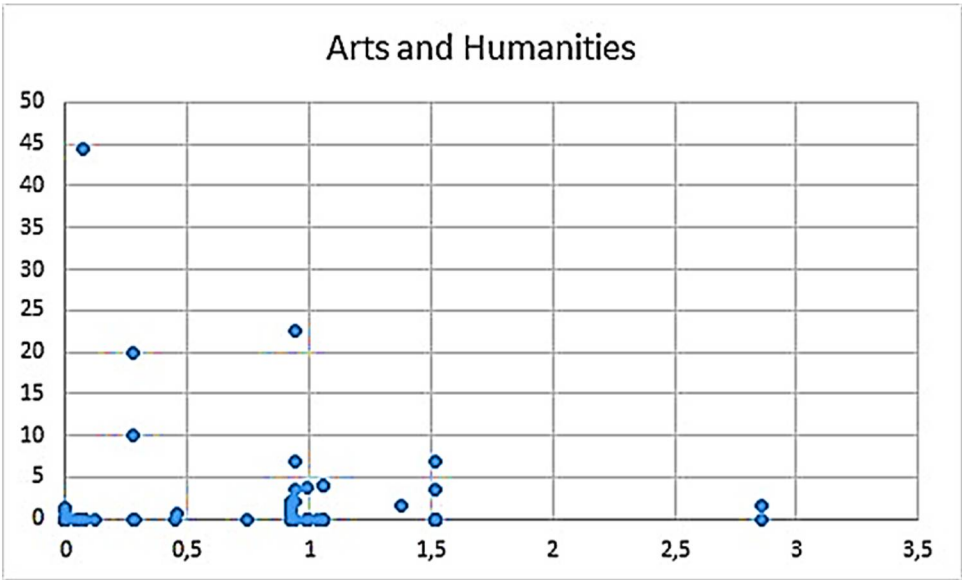122   of several "runouts" (e.g., one article has FWCI 44.39 - see Figure 2).

123

124   **Figure 2.** Distribution of publications in the subject category of *Arts & Humanities* by FWCI and SNIP.

125   Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

126   Therefore, we decided to remove 4 extreme FWCI values, eventually obtained a picture of the
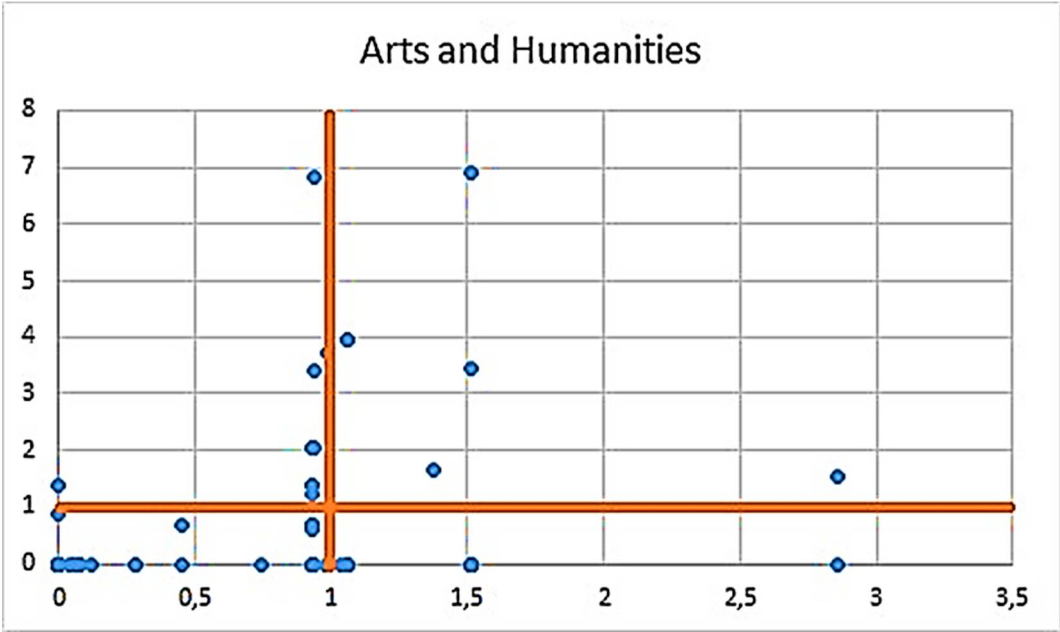127   distribution of publications (Figure 3).

128

129　**Figure 3.** Distribution of publications in the *Arts & Humanities* subject category by FWCI and SNIP
130　excluding extreme values.

131　Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

132　Most publications are in the first and second quadrants. Taking into account the removed
133　observations with the extreme value of FWCI, the main problem is seen in the discrepancy between
134　the selection of sources and the quality of published scientific results. Two out of the four removed
135　publications were published in the journal *Man in India*, another two in the *Pertanika Journal of Social*
136　*Sciences and Humanities*. The coverage of *Man in India* in *Scopus* was discontinued in 2017.

137　*3.2. Business, Management and Accounting*

138　In this subject area, there is the same problem as in the previous case: one of the publications has
139　FWCI 19.94 (by the way, it was also published in the *Pertanika Journal of Social Sciences and Humanities*).
140　In Figure 4 it came out that here, as, in the case of arts & humanities, most of the publications are in
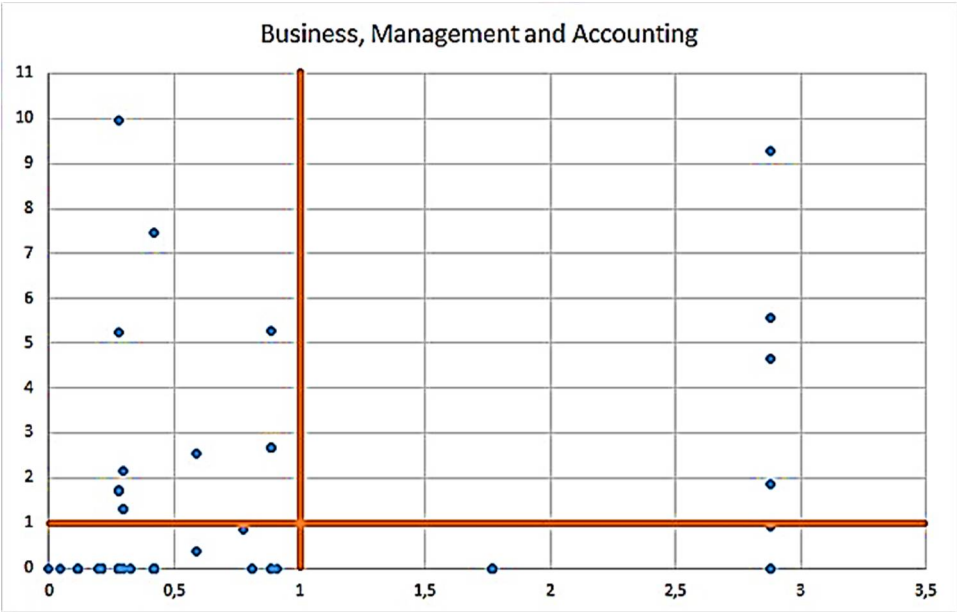141　the first and second quadrants with a big bias in the first.

142

143   **Figure 4.** Distribution of publications in the subject category of *Business, Management and Accounting*
144   by FWCI and SNIP without taking into account extreme values.

145   Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

146   Accordingly, in this case, we are dealing with the problems of relevance and significance of the
147   research itself.

148   *3.3. Decision Science*

149   As mentioned above, there are very few publications in this area, and almost all of them are in
150   the first quadrant (Figure 5).
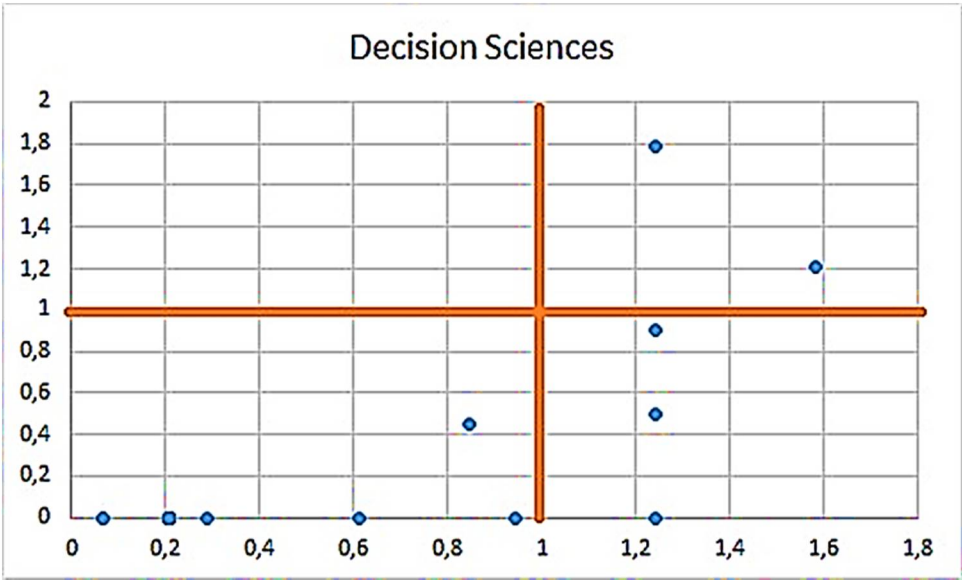
151

152   **Figure 5.** Distribution of publications in the subject category of *Decision Science* by FWCI and SNIP.

153   Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

154   Grounding on the results of the analysis, the development of research in this direction does not
155   seem very promising.

156   *3.4. Economics, Econometrics and Finance*

157   In this category there are 2 extreme values of FWCI, the first one is the same article as in the
158   category *Business, Management and Accounting* (and the second one in *Asian Journal of Social Science*,
159   which coverage in Scopus was discontinued in 2016). The distribution situation here is similar to Arts
160   & Humanities, i.e. publications are concentrated mainly in quadrants 1 and 2, while a lot of
161   qualitative publications are published in low-impact journals (Figure 6).
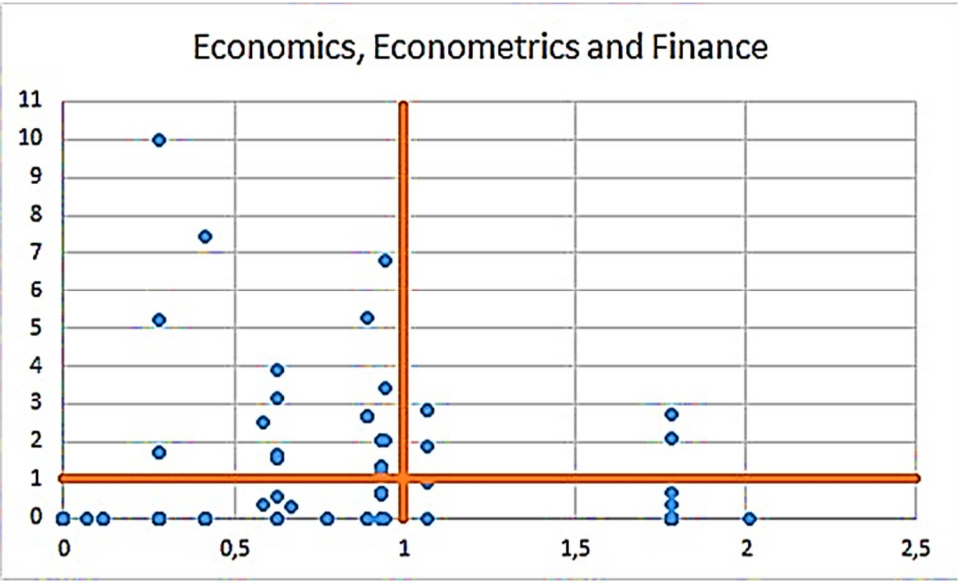
Figure 6. Distribution of publications in the subject category of *Economics, Econometrics and Finance* by FWCI and SNIP.

Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

3.5. Psychology

There are even fewer publications in this category than in Decision Science. Also, the absolute majority of documents are in the first quadrant (Figure 7).
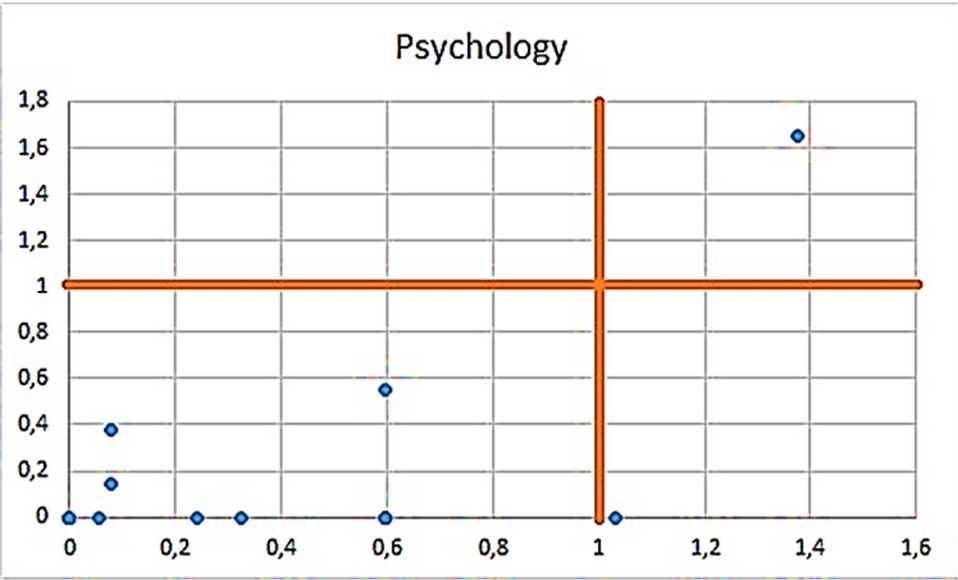

Figure 7. Distribution of publications in the subject category of *Psychology* by FWCI and SNIP.

Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

3.6. Social Sciences

Finally, in the social sciences, we deleted 5 publications with extreme FWCI values (Among them, the publications at the junction of the areas of Education and Mathematics look very interesting and promising). Among the remaining publications (in this subject field the number of publications

176 is the largest, 264) again the first quadrant prevails. Correlation between the level of the journal and
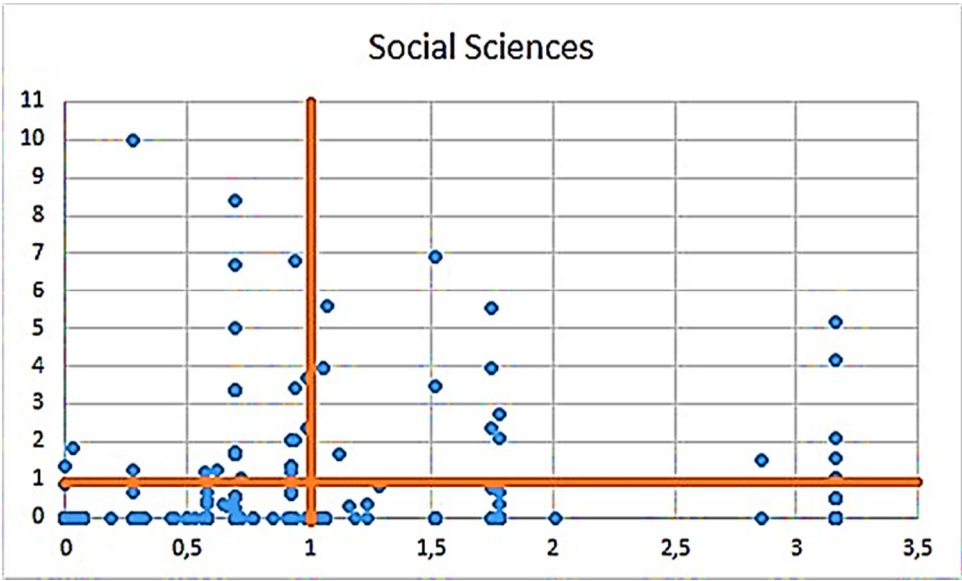177 articles' citation is missing (Figure 8).

178

Figure 8. Distribution of publications in the subject category of *Social Sciences* by FWCI and SNIP.

180 Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

4. Discussion and Conclusions

182   Grounding on the analysis, we can draw the following conclusions:
183   1.   With only two indicators available in SciVal, we can quickly analyze the University's
184 publication strategy in terms of subject areas.
185   2.   The correlation between the citation of publications and the level of the journal in Peoples'
186 Friendship University of Russia (RUDN) is much lower than that of comparable universities. One of
187 the possible reasons is a very high level of self-citation. Table 2 shows the FWCI values including and
188 without self-citations.

189   **Table 2.** Comparative analysis of normalized citation including and without self-citations*.

Subject area / Indicators	RUDN		5-100		C9		Universities of Excellence	
	FWCI **	FWCI 2 ***	FWCI **	FWCI 2 ***	FWCI **	FWCI 2 ***	FWCI **	FWCI 2 ***
Arts&Humanities	1.36	0.68	1.38	0.54	1.49	1.14	1.39	1.04
Business. Management and Accounting	1.12	0.63	1.11	0.59	1.47	1.17	1.37	1.06
Decision Science	0.15	0.07	1.08	0.46	1.32	1.01	1.34	0.94
Economics. Econometrics and Finance	1.24	0.65	1.02	0.41	1.33	1.08	1.44	1.17
Psychology	0.3	0.22	1.01	0.76	1.18	0.96	1.42	1.06
Social Sciences	1.22	0.62	1.09	0.49	1.26	0.96	1.52	1.16

190 *Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

191 **Field-Weighted Citation Impact including self-citations.

192 *** Field-Weighted Citation Impact excluding self-citations.

193 Comparative analysis showed that in 4 subject areas FWCI of publications including self-
194 citations is higher than one, i.e. corresponds to the global citation potential in a given field at a given
195 time period. A bad situation is observed only in the categories of Decision Science and Psychology.
196 But, when we removed self-citations, a fairly large gap with the C9 and Universities of Excellence
197 was revealed. The same problems are observed in other universities of the Project 5-100: the value of
198 indicators of normalized citation for RUDN is even higher than the average for the Project 5-100
199 (excluding two above-mentioned areas).
200 At the same time, we cannot say that the articles of Peoples' Friendship University of Russia or
201 other Russian universities receive fewer views than foreign comparators. If we take the normalized
202 *Field-Weighted Views Impact (FWVI)*, we see that the articles of the PFUR and 5-100 universities are
203 viewed much more often than foreign comparators with higher citation rates (Table 3)!

204 **Table 3.** Comparative analysis of the normalized indicator of publications views*.

Subject area / FWVI	RUDN	5-100	C9	Universities of Excellence
Arts&Humanities	2.46	2.51	1.13	1.05
Business. Management and Accounting	1.66	1.39	1.33	1.24
Decision Science	1.39	1.53	1.39	1.27
Economics. Econometrics and Finance	1.8	1.74	1.17	1.09
Psychology	0.79	1.67	1.21	1.24
Social Sciences	1.89	1.84	1.15	1.17

205 *Source: authors' own development based on the data of *SciVal, Elsevier B.V.*

206 Thus, the main problem is seen in the contents of the publications and the research activities of
207 the university.
208 3. There are very few publications in the subject areas of Decision Science and Psychology,
209 almost all of them are in the first quadrant. Accordingly, the development of these areas seems not
210 very promising.
211 4. In the subject fields/areas Arts & Humanities; Economics, Econometrics, and Finance; Social
212 Sciences quite a lot of research falls into the second quadrant, i.e. high-quality research results are
213 published in low-impact journals.
214 5. Practically in all subject fields/areas, there are articles in the fourth quadrant. Low citation
215 rates can be attributed both to the discrepancy between the quality of the publication and the level of
216 the journal and to the quality of academic English and metadata.

218 **Conflicts of Interest:** The author declares no conflict of interest.

219 **References**

220

221 1. The Ministry of Education and Science of the Russian Federation. Russian Academic Excellence Project
222 2018. Available online: https://5top100.ru/en/ (accessed July 31, 2018).
223 2. QS Quacquarelli Symonds Limited. TOPUNIVERSITIES 2018. Available online:
224 https://www.topuniversities.com/university-rankings (accessed July 31, 2018).
225 3. Times Higher Education. World University Rankings 2018. Available online:
226 https://www.timeshighereducation.com/world-university-rankings (accessed July 31, 2018).

227   4.   Waltman L, van Eck NJ. Source normalized indicators of citation impact: an overview of different
228   approaches and an empirical comparison. *Scientometrics* **2013**, 96, 699–716, DOI:10.1007/s11192-012-0913-4.
229   5.   Waltman L. A review of the literature on citation impact indicators. *J Informetr* **2016**, 10, 365–91,
230   DOI:10.1016/j.joi.2016.02.007.
231   6.   Moed HF. Measuring contextual citation impact of scientific journals. *J Informetr* **2010**, 4, 265–77,
232   DOI:10.1016/j.joi.2010.01.002.
233   7.   Waltman L, van Eck NJ, van Leeuwen TN, Visser MS. Some modifications to the SNIP journal impact
234   indicator. *J Informetr* **2013**, 7, 272–85. DOI:10.1016/j.joi.2012.11.011.
235   8.   Elsevier. Journal Metrics in Scopus: Source Normalized Impact per Paper (SNIP) 2018. Available online:
236   https://blog.scopus.com/posts/journal-metrics-in-scopus-source-normalized-impact-per-paper-snip
237   (accessed July 31, 2018).
238   9.   Haddawy P, Hassan S-U, Asghar A, Amin S. A comprehensive examination of the relation of three citation-
239   based journal metrics to expert judgment of journal quality. *J Informetr* **2016**, 10, 162–73,
240   DOI:10.1016/j.joi.2015.12.005.
241   10. Australian Research Council. Excellence in Research for Australia (ERA) 2018. Available online:
242   http://www.arc.gov.au/excellence-research-australia (accessed July 31, 2018).
243   11. Elsevier. Snowball Metrics 2018. Available online: https://www.snowballmetrics.com/metrics/ (accessed
244   July 31, 2018).
245   12. Hicks D, Wouters P, Waltman L, de Rijcke S, Rafols I. Bibliometrics: The Leiden Manifesto for research
246   metrics. *Nature* **2015**, 520, 429–31, DOI:10.1038/520429a.
247   13. C9   LEAGUE.   Your   University   Guide   n.d.   Available   online:
248   https://youruniversityguide.wordpress.com/east-asia/china/c9-league-九校联盟/ (accessed July 31, 2018).
249   14. Excellence   Initiative   (2005-2017).   DFG   2018.   Available   online:
250   http://www.dfg.de/en/research_funding/programmes/excellence_initiative/index.html (accessed July 31,
251   2018).