Authors' Publication Strategies and Citation Distributions in Journals

Andrey Lovakov[1] Vladimir Pislyakov[2]

[1] *lovakov@hse.ru*
Center for Institutional Studies, National Research University Higher School of Economics, Moscow (Russia)

[2] *pislyakov@hse.ru*
Library, National Research University Higher School of Economics, Moscow (Russia)

Abstract
Psychology is a discipline standing at the crossroads of hard and social sciences. Some of psychology journals are attributed to SCIE in Web of Science database while others to SSCI (and some to both). So it is especially interesting to study bibliometric characteristics of psychology journals. We study not the citedness itself (IF etc.) but the citation distribution across papers within psychology publications. This is, so to say, "indicators of the second order" which measure the digression of the citations received by individual papers from the journal's average. This also influences the publication strategies of the authors. Some journals guarantee to the author receiving of the mean number of citations while others have much more "All or Nothing" grade when any individual paper may have many cites or not have them at all. We also define four different types of psychology journals and explore their characteristics separately.

Conference Topic
Journals; Databases and electronic publications; Citation and co-citation analysis

Introduction
Citation distributions are generally characterized as "highly skewed", this means a non-Gaussian pattern and high concentration of citations on some items while others are low-cited or non-cited at all, see papers by Whitehand (1985), Seglen (1992) and many others. Still, the uneven distribution of citations has its own levels of inequality and may be measured by common indicators such as Gini index, which will be used in the present paper.

The Gini index is a measure of inequality of a distribution. In terms of citation distribution it is a comparison of the journal where all citations are received by one paper with the case when each journal's paper gets the same number of citations. The Gini index ranges from 0 to 1 where zero corresponds to perfect equality (each paper receives the same number of citations) and 1 corresponds to perfect inequality (only one paper has citations while others are not cited at all).

We are interested in how different journals gain their average citedness per paper. Some may get citedness from close to even distribution of citations, when each article has a high probability to receive an average number of cites. Other publications get their averages from a limited number of highly cited papers, while all others are not cited. The same research question was raised by Nuti et al. (2015) for cardiovascular journals. We choose psychology journals to investigate the issue. Psychology is chosen as a discipline standing at the crossroads of hard and social sciences, which makes it especially interesting for the analysis.

Gini index is used to measure the inequalities of citation distributions by Chi (2016) where the unevenness of book citations is compared with journals and the strong influence of uncited items on Gini coefficient is observed. Wren (2016) plainly states that "The higher Gini coefficients for bioinformatics journals suggest that development of novel bioinformatics resources may be somewhat 'hit or miss'. That is, some approaches become widely adopted and produce a disproportionate number of extremely highly cited papers, while most are not widely adopted <...>."

Gini coefficients for citations to a particular journal were also studied by, for example, Chatterjee, Ghosh & Chakrabarti (2016) and by Stegmann & Grohmann (2001). The latter paper analyzes journals on dermatology and explores the dependence of Gini index from journal impact factor. We address this question in the present paper too. Gini indices for leading multidisciplinary journals (*Science*, *Nature*, *PNAS* etc.) and for *PhysRev* series may be found in (Ghosh, Chattopadhyay & Chakrabarti, 2014).

Gini measure of inequality is also used in the other applied informetrics studies. For example Pislyakov (2008) applies this indicator to usage statistics of electronic journals. Leydesdorff & Rafols (2011) use it, along with other inequality indicators, to measure interdisciplinarity of a journal. They also agree that "The Gini coefficient <...> has the advantage of having been widely used in bibliometrics".

The new thing we want to introduce is the connection between the citation patterns in the journals and the choice the author makes when he/she chooses the appropriate journal to submit his/her manuscript. The choice is whether an author prefers to securely receive the average level of citations which is common to the journal (the case of publications with uniformly distributed citations across papers) or he/she wishes to risk and submit the paper to a journal where just several papers get extremely high level of citedness while others are being under-cited.

This choice marks different publication strategies of a scientist (whether chosen consciously or not). “All is safe” vs. “All or Nothing”.

Data and Methods

Data on 40 psychological journals from Social Sciences Citation Index (SSCI, Web of Science, Clarivate Analytics, ex-Thomson Reuters) were used (Table 1). We limit our analysis only to psychological journals, so there is no problem of cross-disciplinary comparison. However, we distinguish four types of journals in psychology, which, as we will further observe, demonstrate different patterns of citation inequality.

We have identified four types of psychological journals: multidisciplinary, sub-disciplinary, journals dedicated to a single particular topic, and methodological publications. In Web of Science there is another classification schema, with ‘Psychology: Experimental’, ‘Psychology: Developmental’ and others (8 different categories), but for the present research we need the classification ranging across all the sub-disciplines of psychology but marking the specific type of a journal.

Multidisciplinary group of journals publishes research on all spectrum of the science of psychology. For example, *Psychological Science* publishes papers on topics ranging from cognitive, social, developmental, and health psychology to behavioral neuroscience and biopsychology. Sub-disciplinary group includes journals which cover only one area of psychology. For example, *Journal of Personality and Social Psychology* publishes papers in areas of personality and social psychology while *Cognitive Psychology* publishes papers about attention, memory, language processing, perception, problem solving, and thinking. The group of journals dedicated to a single particular topic is focused on some special research area. For example, *Depression & Anxiety* focuses on studies related to different aspects of mood and anxiety disorders and related phenomena, and *Personal Relationships* publishes studies focused on attributes of individual partners in personal relationships, interactive relationship processes, and relationships in social contexts. Finally, methodological group includes journals which are

devoted to research design, methodology, measurement, quantitative and qualitative data analysis. For example, *Behavior Research Methods* publishes papers on methods, techniques, and instrumentation of research in experimental psychology, and *Psychological Methods* on methods for collecting, analyzing, understanding, and interpreting psychological data. The sample includes different journals in terms of impact factor and publisher.

Five-year journal impact-factors (2015) for each journal were extracted from Journal Citation Reports (JCR, Web of Science, Clarivate Analytics, ex-Thomson Reuters). Five Gini indexes were calculated for each journal. Every Gini index was based on citations of papers published in each single year, from 2010 to 2014 (only "articles" and "reviews" document types were taken into calculation). These five indexes were then aggregated using the arithmetic mean. The five-year 2015 impact is used as the same papers are assessed by it: citations of articles/reviews from 2010 to 2014 are included in the formulae of impact-factor.

For the data preparation, analysis and visualization we used R, a programming environment for statistical computing (R Core Team, 2016). The Gini indexes were calculated by the *reldist* package (Handcock, 2016). Additional packages were used: *dplyr* (Wickham & Francois, 2016), *ggplot2* (Wickham, 2009), *readxl* (Wickham, 2016).

Results and Discussion

Table 1 shows Gini index for 40 psychological journals chosen for our analysis. The Gini varies from 0.41 to 0.64 proving that there are different citation patterns, from close to "All or Nothing" as in *Psichothema* or *Behavior Research Methods* to "All is safe" as in *Psychology of Men and Masculinity*.

If we plot the dependence of Gini index from IF for all 40 analyzed journals, no clear regularity is observed (Figure 1). It seems that journal's citedness has no evident connection with the distribution of citations across its papers.

But when psychological publications are subdivided into classes by types of journals, the relationship of inequality of citation distribution to impact factor becomes more obvious. We may see it in Figure 2, where linear regression lines with 95% confidence intervals for predictions from a linear model are shown.

What is interesting, and what is the main point of our paper, this has a clear connection to a publication strategy of an author. For example, if you publish in sub-disciplinary journals, the more cited journal you choose, the more "safety" in receiving average citations you get (Pearson's $r = -0.87$, $p < 0.001$). But when an author submits his/her paper to a multidisciplinary journal, he/she may choose several ones *with the same IF* but with significantly different risk of receiving/non-receiving the average number of citations (so to say "All or Nothing" grade).

There is different level of inequality of citation distribution between groups of journals. Gini indexes of all journals from the methodological group are higher than 0.5, whereas Gini indexes of only a couple of journals from sub-disciplinary and special topic groups exceeds 0.5. It means that in general the inequality of citation distribution is higher in journals, which publish papers on methodology, methods, and tools of psychological research. These results are in accord with Wren (2016) that bioinformatics journals, which publish top papers, have higher Gini indexes. Publishing on methodology and research methods in psychology seems also to be 'hit or miss' in terms of citedness. The probable explanation may be that some methodological papers attract attention of many researchers who will use and cite them because these methods are well-

received by psychologists and are becoming widespread. While other papers are not so interesting and popular being rather complex or addressing too specific problems. It is in the methodological part of the science the inequality is therefore becomes more pronounced.

Table 1. Average Gini index and 5-year impact factor for the psychological journals.

Journal	*Average Gini index*	*5-year impact factor*
Multidisciplinary		
Psychological Science	0.46	6.29
British Journal of Psychology	0.52	3.26
Frontiers in Psychology	0.51	2.88
Acta Psychologica	0.51	2.41
Journal of Psychology	0.48	1.76
Scandinavian Journal of Psychology	0.52	1.55
International Journal of Psychology	0.56	1.42
Australian Journal of Psychology	0.45	1.26
Psicothema	0.64	1.23
Spanish Journal of Psychology	0.58	0.84
Sub-disciplinary		
Journal of Personality and Social Psychology	0.42	7.44
Child Development	0.47	5.81
Cognitive Psychology	0.46	5.49
Journal of Personality	0.47	4.18
Memory & Cognition	0.48	2.59
Personality and Individual Differences	0.48	2.42
Journal of Cross-Cultural Psychology	0.50	2.24
Journal of Economic Psychology	0.50	2.06
Small Group Research	0.48	1.32
Journal of Applied Social Psychology	0.53	1.31
Methodological		
Psychological Methods	0.63	9.46
Organizational Research Methods	0.55	6.51
Multivariate Behavioral Research	0.62	4.78
Behavior Research Methods	0.64	3.98
British Journal of Mathematical and Statistical Psychology	0.61	2.74
Journal of Mathematical Psychology	0.59	2.61
Psychometrika	0.54	2.58
Methodology. European Journal of Research Methods for the Behavioral and Social Sciences	0.56	2.17
Educational and Psychological Measurement	0.54	1.85
Journal of Educational Measurement	0.53	1.34
Special topic		
Depression and Anxiety	0.46	5.52
Emotion Review	0.49	4.12
International Journal of Eating Disorders	0.49	3.69
Autism	0.42	3.48
Psychology of Addictive Behaviors	0.43	3.42
Intelligence	0.48	3.29
Psychology of Men & Masculinity	0.41	3.12
Journal of Gambling Studies	0.43	2.98
Personal Relationships	0.47	1.57
Perception	0.56	1.13

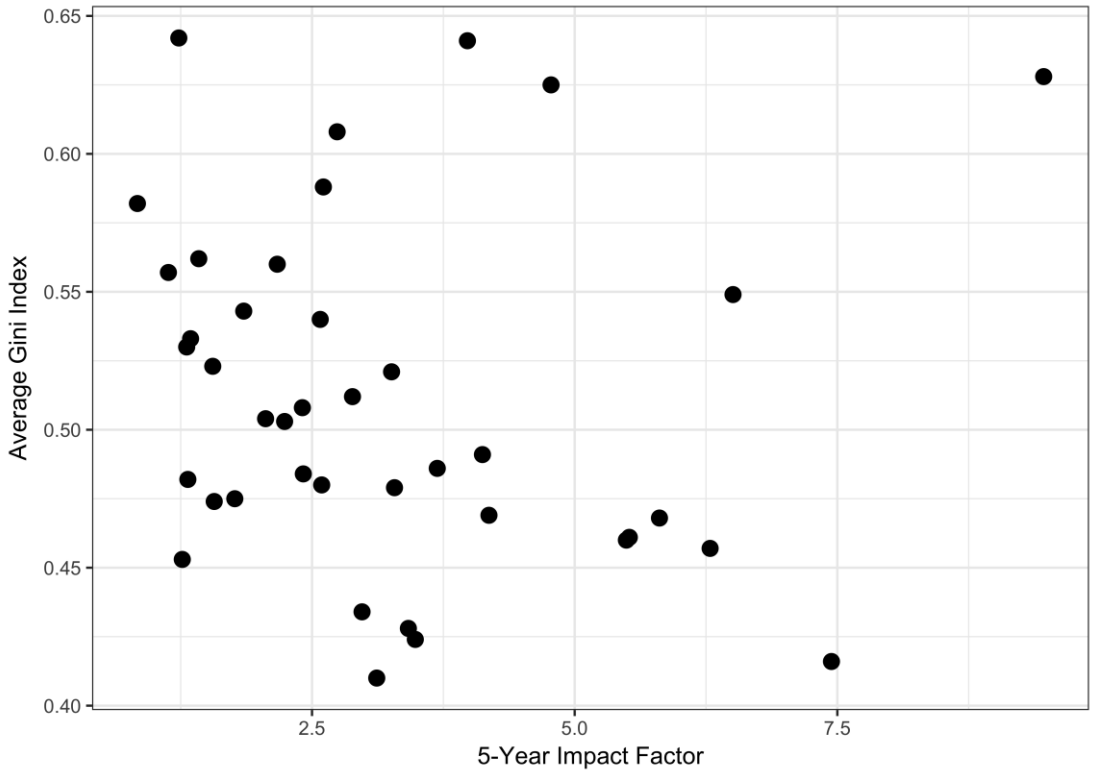

Figure 1. Relationship between average Gini index and 5-year impact factor (the whole sample of journals).

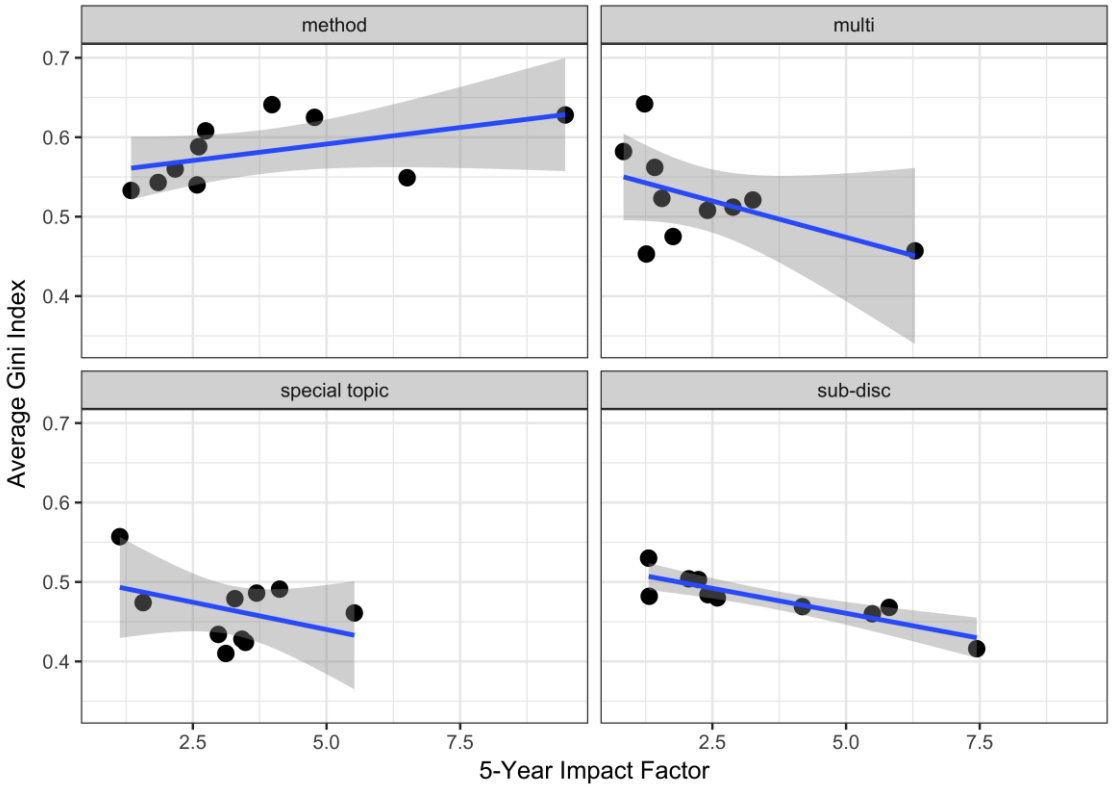

Figure 2. Relationship between average Gini index and 5-year impact factor by types of journals (linear regressions are shown, grey color displays 95% confidence interval for predictions from a linear model).

For instance, the most cited paper from methodological journals in our sample (860 total citations at the moment of data extraction) is an introduction to propensity score methods for observational studies (Austin, 2011) which are relevant not only for psychological research but also for all fields with observational (or nonrandomized) studies. And the third most cited paper (540 citations in total) demonstrates how to use Amazon's Mechanical Turk web site for conducting behavioral research (Mason & Suri, 2012) which is extremely useful tool for many researchers in psychology.

We may also remind the well-known '*Acta Crystallographica A* case' when only one methodological paper (Sheldrick, 2008) has changed the impact factor of the whole journal which has reached the top-3 overall (i.e. within *all* Web of Science) IF ranking. It was more than 20 times IF-rise for this journal. So, the methodological inequalities are common to the science and we may observe this in case of psychological publications too.

Conclusion

It is well known in bibliometrics that journal citedness (for example impact factors) should be compared only within the same scientific discipline. Moreover, there are differences between journal types (those which publish more review papers are generally cited more, for example). The present paper shows that the same approach should be used when we study not only average citedness but also patterns of digression of received by individual papers' citations from this average journal citedness. Psychology journals of different types have different Gini indices and different relationship between Gini and 5-year impact factor.

This clearly marks differences in publication strategies. Of course, when the author chooses the publication venue, at the first place stands the subject matter of the journal and its aptness to his/her research interests. Next, generally goes the international visibility/popularity of the journal which is often proxied by the citedness of the periodical.

But as we may see from our research, another characteristic of the journal also matters. It is so to say the grade of "All or Nothing". How strong the scientist may be confident in receiving average number of journal's citations to his/her individual paper. Some journals have more equal citation distribution across their papers than the others with the same IF.

Though often this choice is made not consciously by the authors, they may feel it from empiric impressions from different journals (if they have a scholarly career long enough). Anyhow, the authors should know and keep in mind these "differences of the second order" which may influence their publication strategies.

At last, we should underline that though the average of Gini for five years was taken for the present paper, the further research is needed to examine whether the regularities found are robust over time. Moreover, the effect most probably will vary across scientific disciplines which is also a fascinating topic of future analysis.

Acknowledgments

The paper was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE), and supported within the framework of a subsidy by the Russian Academic Excellence Project '5-100'.

References

Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research, 46*(3), 399–424. http://doi.org/10.1080/00273171.2011.568786

Chatterjee, A., Ghosh, A., & Chakrabarti, B. K. (2016). Universality of Citation Distributions for Academic Institutions and Journals. *PLOS ONE, 11*(1), e0146762. http://doi.org/10.1371/journal.pone.0146762

Chi, P.-S. (2016). Differing disciplinary citation concentration patterns of book and journal literature? *Journal of Informetrics, 10*(3), 814–829. http://doi.org/10.1016/j.joi.2016.05.005

Ghosh, A., Chattopadhyay, N., & Chakrabarti, B. K. (2014). Inequality in societies, academic institutions and science journals: Gini and k-indices. *Physica A: Statistical Mechanics and Its Applications, 410*, 30–34. http://doi.org/10.1016/j.physa.2014.05.026

Handcock, M. S. (2016). *Relative Distribution Methods*. Version 1.6-6. Project home page at http://www.stat.ucla.edu/~handcock/RelDist. URL: https://CRAN.R-project.org/package=reldist

Leydesdorff, L., & Rafols, I. (2011). Indicators of the Interdisciplinarity of Journals: Diversity, Centrality, and Citations. *Journal of Informetrics, 5*(1), 87-100. http://doi.org/10.1016/j.joi.2010.09.002

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*(1), 1–23. http://doi.org/10.3758/s13428-011-0124-6

Nuti, S. V., Ranasinghe, I., Murugiah, K., Shojaee, A., Li, S.-X., & Krumholz, H. M. (2015). Association Between Journal Citation Distribution and Impact Factor. *Journal of the American College of Cardiology, 65*(16), 1711–1712. http://doi.org/10.1016/j.jacc.2014.12.071

Pislyakov, V. (2008). *Informetric Modelling of the Process of Electronic Information Resources Usage*. PhD Thesis. Kazan. https://elib.hse.ru/science/papers/informetrics_full.pdf

R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Seglen, P. O. (1992). The skewness of science. Journal of the American Society for Information Science, 43(9), 628–638.

Sheldrick, G. (2008). A short history of SHELX. *Acta Crystallographica A, 64*(1), 112-122.

Stegmann, J., & Grohmann, G. (2001). Citation rates, knowledge export and international visibility of dermatology journals listed and not listed in the Journal Citation Reports. *Scientometrics, 50*(3), 483–502. http://doi.org/10.1023/A:1010562815788

Whitehand, J. W. R. (1985). Contributors to the recent development and influence of human geography: What citation analysis suggests. *Transactions - Institute of British Geographers, 10(2)*, 222–234.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.

Wickham, H. (2016). *readxl: Read Excel Files*. R package version 0.1.1. https://CRAN.R-project.org/package=readxl

Wickham, H. & Francois, R. (2016). *dplyr: A Grammar of Data Manipulation*. R package version 0.5.0. https://CRAN.R-project.org/package=dplyr

Wren, J. D. (2016). Bioinformatics programs are 31-fold over-represented among the highest impact scientific papers of the past two decades. *Bioinformatics, 32*(17), 2686–2691. http://doi.org/10.1093/bioinformatics/btw284