Citation Expectations: Are They Realized? Study of the Matthew Index for Russian Papers Published Abroad

Vladimir Pislyakov[1] and Ekaterina Dyachenko[2]

[1] *pislyakov@hse.ru*
Higher School of Economics, Library, 101000, Myasnitskaya, 20, Moscow (Russia)

[2] *dyachenkoel@hse.perm.ru*
Higher School of Economics, Perm Branch, Library, 614070, Studencheskaya, 38, Perm (Russia)

Abstract
We consider the "Matthew effect" in the citation process which leads to reallocation (or misallocation) of the citations received by scientific papers within the same journals. The case when such reallocation correlates with a country where an author works is investigated. Russian papers in chemistry and physics published abroad were examined. We found that in both disciplines in about 60% of journals Russian papers are cited less than an average one. However, if we consider each discipline as a whole, citedness of a Russian paper in physics will be on the average level, while chemistry publications receive about 16% citations less than one may expect from the citedness of the journals where they appear. Moreover, Russian chemistry papers mostly become undercited in the leading journals of the field. Characteristics of a "Matthew index" indicator and its significance for scientometric studies are also discussed.

Introduction
The term "Matthew effect" was introduced by Robert K. Merton (Merton, 1968). He described a psychosocial mechanism that led to misallocation of credit in the reward system of science. Papers written by eminent scholars (e. g. Nobel laureates) tend to get disproportionately great credit while relatively unknown scientists tend to get disproportionately little credit for contributions of the same quality. The name for the effect comes from the Gospel saying "for unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath" (Matthew 25:29). Later Merton (1988) developed his ideas further.
A group of German scholars (Bonitz, Bruckner & Scharnhorst, 1997; Bonitz, Bruckner & Scharnhorst, 1999; Bonitz & Scharnhorst, 2001) discovered and investigated a similar effect for citation of works written by scientists from different countries. They found that there are several countries that obtain more citations than it is supposed from the citedness of the journals where they publish their work. At the same time there are a plenty of other countries whose publications are cited less frequently than average papers in the same journals. German researchers also discovered that this regularity correlates with the impact factors of the journals where a country publishes its works. If a country has high "citation expectations" (based on the average citedness of the journals where it publishes), then it probably will get even more citations than expected. If a country has low citation expectations, its works generally will be even more undercited. This effect was called by Bonitz et al. "Matthew effect for countries" (Bonitz, Bruckner & Scharnhorst, 1997).
To measure the degree of Matthew effect an indicator named "Matthew index" was defined as the ratio of the difference between observed and expected numbers of citations to the expected number of citations. Here "expected number of citations" is calculated on the basis of the average number of cites per paper in a journal. Matthew index for any country may be calculated for a single journal (then it shows how papers of this country's authors are undercited or overcited in this particular journal), for a whole discipline or even for all its scientific publications.

Proceedings of ISSI 2009, pages 52–58.
Edited by B. Larsen and J. Leta.

We aim at investigating the Matthew index for Russian papers published abroad, i. e. in non-Russian journals. This will reveal how the publications of Russian scholars look against papers of their foreign colleagues in the context of the same journals. Bonitz (2002; 2005) emphasized that Matthew index measures "competitiveness" of a nation's scientific literature. It is undoubtedly true, especially if we keep in mind that Matthew effect for countries is in fact a *redistribution* of citations, and if one country gains positive Matthew index then some others should have negative one: it is *their* citations that are redistributed in favour of the winner(s).

For our research we chose two scientific fields which are considered traditionally strong in Russian science, chemistry and physics. According to Essential Science Indicators (ESI) ranking these fields have the highest number of papers and receive the greatest number of citations compared to all other Russian sectors of science. Our investigation should show how competitive Russian papers are in chemistry and physics when they are published in prestigious international journals.

Finally, recently not much attention has been paid to such an interesting scientometric aspect of a Matthew effect as "Matthew effect for countries", though its other manifestations were actively investigated, e. g. (Medoff, 2006; Morgan, Farkas & Hibel, 2008). Our contribution seeks to improve this situation.

Methodology

A Thomson Reuters database Science Citation Index Expanded hosted on the Web of Knowledge online platform was used to get citation data. Only documents of the type "Article" were taken into account. (Note that data were gathered in spring 2008, while in autumn 2008 some of the "Articles" transformed their document type in Web of Science to "Proceedings Paper", so that they both must be included now to reproduce our results, see (Thomson Reuters, 2009) for details.) A paper was attributed to Russia if country of at least one of the institutional affiliations of its authors was "Russia" (the so-called "whole counting" method; note that Bonitz et al. used first-author method).

Data were collected and Matthew indices were calculated for two disciplines, chemistry and physics. For subject classification of the journals ESI field classification was used. This system classifies all Web of Science journals into 22 broad fields, "Chemistry" and "Physics" are among them. Only one field is assigned to each journal, so two journal sets do not overlap. We chose two years of publication, 2003 and 1997. A 4-year citation period was chosen, including the year of publication: citations from 2003–2006 literature were counted for 2003 papers and from 1997–2000 literature for 1997 papers.

Table 1. Sources and papers included into analysis.

set	journals	papers, total	Russian papers	share of Russian papers, %
chemistry-2003	222	76359	1718	2.3
chemistry-1997	212	67488	1582	2.3
physics-2003	183	71498	4002	5.6
physics-1997	164	58854	3883	6.6

To investigate citedness of Russian papers published abroad, we excluded all Russian journals found in the ESI lists. As it is hard to determine "nationality" of the journal in this time of globalization of science, a formal criterion was used: all journals that published in the corresponding year more than 50% articles from Russia were excluded from further analysis. Journals that have not published a single Russian paper during that year were also excluded, as their citedness did not influence Matthew indices for Russia. Four final journal sets for the analysis are called “chemistry-2003”, “chemistry-1997”, “physics-2003” and “physics-1997”. They contain 222, 212, 183 and 164 sources respectively. We will also speak of “chemistry-2003 papers” (papers published in 2003 in chemistry-2003 journals) and so on. Quantitative characteristics of these sets are summarized in Table 1.

Results and discussion

Citation indicators

Figure 1 shows “gross” citation indicators (without descending to the journal level) for all 4 sets of papers. A sharp difference between relative standing of chemistry and physics papers published by Russian authors in foreign journals is obvious. While an average Russian paper in physics is cited on the world average level, a paper in chemistry obtains only about 70% of citations to the world’s average paper. These shares are almost constant for the years under consideration, though there is a slight positive movement from 1997 to 2003 in physics and slight negative shift in chemistry.

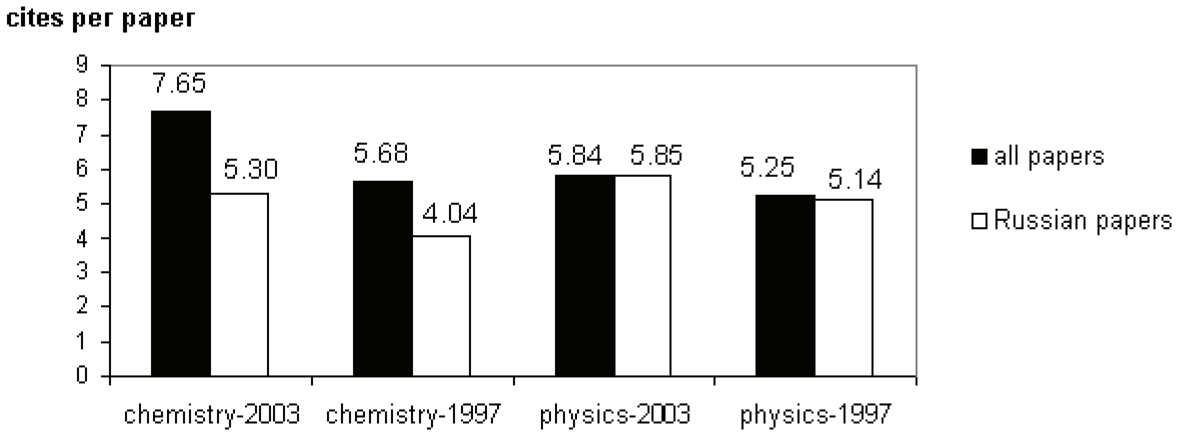

Figure 1. Cites per paper for 4 sets of papers.

Matthew indices for journals

Matthew index for Russian papers in a particular journal may be positive (Russian papers are cited more than journal’s average) or negative (they are cited less than journal’s average). Distribution of the numbers of journals with positive and negative indices for all 4 sets of sources is shown in Figure 2. For all cases there are more journals where Russian publications are undercited, generally about 60% of the sources have Matthew index less than zero, with the exception of chemistry-2003 set where this share amounts to 68%.

On the whole we found no significant correlation between Matthew index for Russian papers in a journal and (a) the average citedness of its papers (maximum absolute value of the Pearson correlation coefficient through all 4 sets is 0.20); (b) number of Russian papers in the journal (maximum is 0.10); (c) share of Russian papers in the journal (maximum is 0.08). In a certain sense this is in accordance with Bonitz (2002) who states that number of papers or citations, number of participating countries or impact factor hardly influence the “Matthew core journals” list (journals where papers are more likely to have higher absolute value of Matthew index).

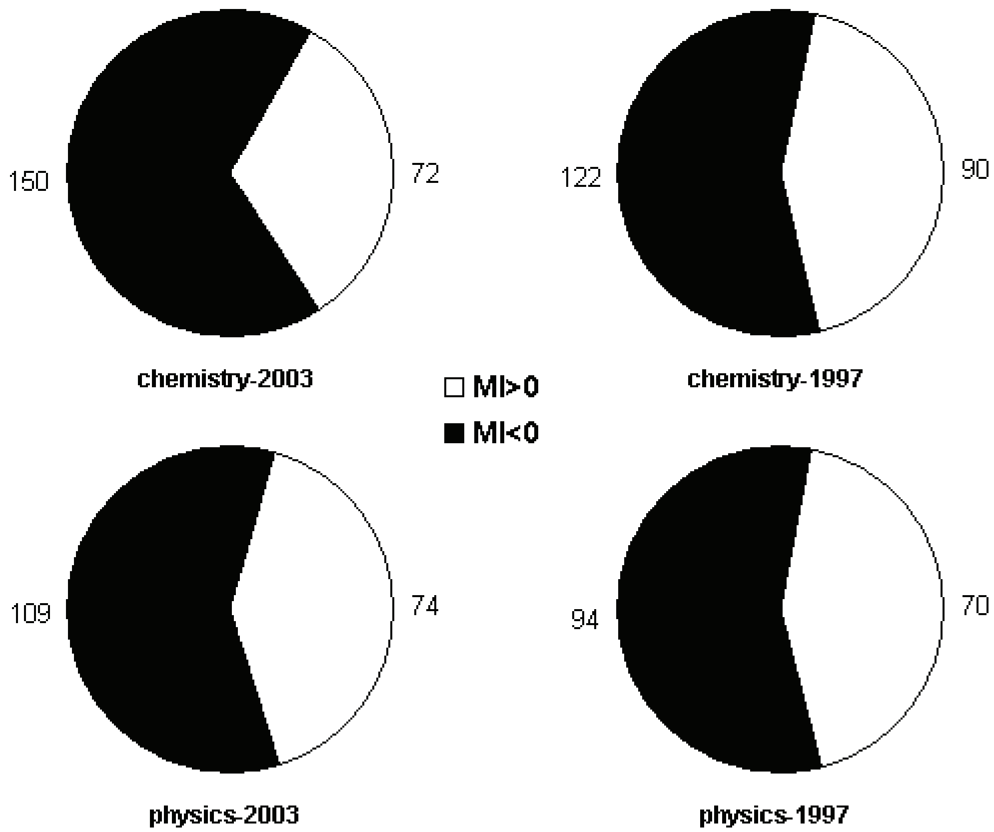

Figure 2. Matthew indices (MI) for Russian papers, by individual journals. The numbers next to the pie charts represent the number of journals with positive and negative values of the Matthew index, respectively.

Though overall correlation is not significant, it appeared useful to plot "relational charts" (Schubert & Braun, 1986) that show how citedness of Russian papers in a journal corresponds to average citedness of a journal's paper (Figure 3). The X-axis shows average citedness of a journal's paper and in fact reflects journal's "generalized impact factor" in a sense of Rousseau (1988) and Egghe (1988) concept (and counted for articles only). The Y-axis shows average citedness of Russian papers in the corresponding journal. If Russian publications are cited exactly as an average paper in a journal, this journal will lie on the $y=x$ line which is shown on all diagrams. More "observed" citations for Russian papers than "expected" level will move the journal above this line. If Russian papers are undercited, this will move the corresponding dots below $y=x$.

There is no clear regularity for dots scattering in Figure 3. Still, some valuable observations may be made. For both chemical sets we may see that in all journals with the highest impact factor (in a generalized sense mentioned above) Russian papers become undercited. The first 10 journals with most cited papers for chemistry-2003 and 14 journals for chemistry-1997 lie below the $y=x$ line. This means that in the most prestigious chemistry journals Russian papers are regularly cited below average level. The same is not true for physics publications. Among 10 most cited journals in physics-1997 set 3 have positive Matthew index for Russian papers. For physics-2003 this figure is even greater: in 8 out of 10 journals with the highest impact (including no. 1, it is *Physical Review Letters*) Russian articles are cited above average level. This distinction in getting credit in the most influential journals marks serious difference between levels of Russian literature in physics and chemistry.

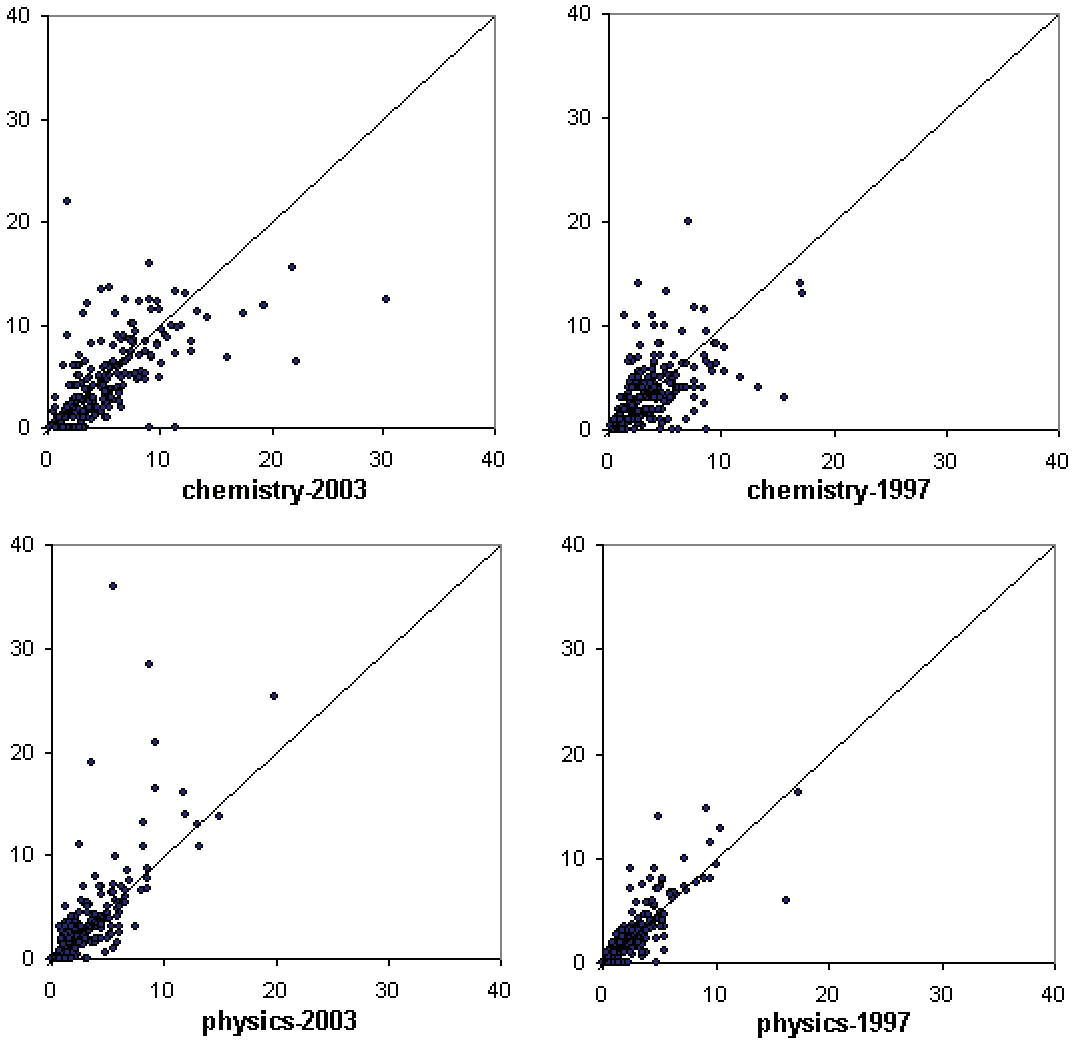

Figure 3. Relational charts for Russian papers.
X-axis — cites per average paper in a journal;
Y-axis — cites per average Russian paper in a journal.

Another interesting question is that of the stability of the Matthew index for a given journal through a period of time. On the basis of our statistics for two years of publication for two disciplines no stability was found. Journal sets chemistry-1997 and chemistry-2003 share 149 journals in common. Physics-1997 and physics-2003 have 133 overlapping journals. Pearson correlation coefficients for 1997 and 2003 Matthew indices for the same journals are 0.18 for chemistry and 0.11 for physics. It is very weak correlation which shows that there are no specific journals where Russian papers are systematically undercited and those where Russian papers always receive citations above average level. Additional argument for this is that index reverses its sign for 69 out of 149 chemistry journals and for 66 out of 133 physics ones. If Russian papers-1997 were undercited in these journals, then Russian papers-2003 were overcited or vice versa. Rather interestingly, if we limit the overlap of the sets to journals with significant number of Russian papers, the correlation between their Matthew indices will become somewhat stronger. Pearson coefficient for 1997 and 2003 Matthew indices of 14 chemistry journals with no less than 20 Russian papers equals 0.44, while for 33 such journals in physics it becomes 0.31. Still, even this correlation is not convincing.

Matthew indices for disciplines

If we combine data from all journals in each set, we will get Matthew indices for the whole disciplines. Table 2 contains these indices as well as the difference between observed and expected numbers of citations ("surplus").

Table 2. Matthew index for Russian papers, by disciplines.

set	"surplus" of citations	Matthew index
chemistry-2003	–1762	–16.2%
chemistry-1997	–723	–10.2%
physics-2003	+160	0.7%
physics-1997	+235	1.2%

We may conclude that Russian chemistry literature receives significantly less citations than it is expected on the basis of citedness of the journals where Russian papers appear. For 2003 articles Russian authors received 1762 citations less than average papers distributed among the same journals. On average Russian paper in chemistry gets 16% less citations.

Physics, on the contrary, shows almost exact average level of citedness. Its Matthew index is slightly above zero. The most interesting situation is with physics-1997 set. As we may infer from Figure 1, Russian papers in this set receive as a whole 2% citations below average level. However, Matthew index for them is positive. This demonstrates difference between "gross-" and "micro-level" (journal level) of citation analysis. On the journal level Russian papers were cited better than average papers in the same journal, but, most probably, distribution of papers by Russian authors skewed in 1997 towards sources with lower impact factor, so the overall citedness of Russian publications was below average level. This effect is closely linked with the concept of "relative publication strategy" (Vinkler, 1997).

It may be added that international collaboration strongly influences the Matthew index. Separate calculations for papers written *only* by Russian authors (no other country mentioned in all addresses of the authors' affiliations) discovered that their Matthew index falls to –32% for chemistry-2003 and to –31% for physics-2003. In many cases when Russian scientists publish their work abroad they need coauthors from other countries to have their papers cited on the average journal level. Perhaps one of the reasons why chemistry papers have lower Matthew index than physics ones is a greater share of the "pure Russian" publications in chemistry (41% against 29% in physics for 2003 sets).

Conclusion

We have studied visible traces of subtle processes in the universe of science communication that lead to reallocation of credit and redistribution of citations to scientific works. This results in non-zero "Matthew index" and both micro-structure and macro-structure of such a redistribution were considered. On the micro-level, considering separate journals, the effect was recorded as for Russian papers in chemistry so in physics. In about 60% of foreign journals Russian publications are cited less than an average paper.

As for the macro-view, on the discipline (field) level, Matthew index appeared to be negative for chemistry and slightly above zero for physics. This finding, together with the fact that Russian chemistry papers are undercited in all most prestigious journals, allowed us to detect a significant difference in the state of these sciences in Russia, or at least in the course of their internationalisation.

Matthew index proved to be a special, independent indicator that does not correlate with any other bibliometric characteristics of the studied sets. The intriguing question whether the observed undercitedness of some of the Russian literature is a pure psychosocial phenomenon, or it reflects real difference in value of this corpus of works, deserves more thorough examination in the future.

In closing it is worth mentioning that studies related to the Matthew effect in bibliometrics are not a merely theoretical activity, but can bring changes to the science policy procedures. For example, Bordons, Fernandez & Gomez (2002) recommend to use impact factors of journals where a paper is published and not the observed citations for scientometric analysis of "peripheral country's" research. This is to avoid Matthew-like effects and "sociological" component of the allocation of scientific reward. Anyhow, two main methods to measure a merit of a scientific paper are counting of the citations received and considering the impact factor of a journal that has published it. In scientometrics Matthew index stands in an important position, serving as a link between one method and another.

Acknowledgments

This research was supported by the Higher School of Economics Scientific Foundation (Grant 08-01-0050 "How papers of Russian scientists published abroad are cited? Study of the Matthew Effect"). Authors would like to thank Dr. Andrei V. Poletaev and Dr. Yuri Khripin for their valuable comments on the previous version of this paper.

References

Bonitz, M. (2002). Ranking of Nations and Heightened Competition in Matthew Core Journals: Two Faces of the Matthew Effect for Countries, *Library Trends*, 50, 440-460.

Bonitz, M. (2005). Ten years Matthew Effect for Countries, *Scientometrics*, 64, 375-379.

Bonitz, M., Bruckner, E. & Scharnhorst, A. (1997). Characteristics and Impact of the Matthew Effect for Countries, *Scientometrics*, 40, 407-422.

Bonitz, M., Bruckner, E. & Scharnhorst, A. (1999). The Matthew Index – Concentration Patterns and Matthew Core Journals, *Scientometrics*, 44, 361-378.

Bonitz, M. & Scharnhorst, A. (2001). Competition in Science and the Matthew Core Journals, *Scientometrics*, 51, 37-54.

Bordons, M., Fernandez, M.T. & Gomez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance in a peripheral country, *Scientometrics*, 53, 195-206.

Egghe, L. (1988). Mathematical relations between impact factors and average number of citations, *Information Processing & Management*, 24, 567-576.

Medoff, M.H. (2006). Evidence of a Harvard and Chicago Matthew effect, *Journal of Economic Methodology*, 13, 485-506.

Merton, R.K. (1968). The Matthew Effect in Science. *Science,* 3810, 56-63.

Merton, R.K. (1988). The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property, *ISIS*, 79, 606-623.

Morgan, P.L., Farkas, G. & Hibel, J. (2008). Matthew effects for whom? *Learning Disability Quarterly*, 31, 187-198.

Rousseau, R. (1988). Citation distribution of pure mathematics journals. In L. Egghe & R. Rousseau (Eds.), *Informetrics 87/88* (pp. 249-262). Amsterdam: Elsevier.

Schubert, A. & Braun, T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact, *Scientometrics*, 9, 281-291.

Thomson Reuters (2009). *Why the number of articles in Web of Science has gone down, and the number of proceedings papers has gone up*. Retrieved April 9, 2009 from: http://isiwebofknowledge.com/products_tools/multidisciplinary/webofscience/cpci/usingproceedings/

Vinkler, P. (1997). Relations of relative scientometric impact indicators. The relative publication strategy index, *Scientometrics*, 40, 163-169.