TRAMES, 2013, **17**(**67/62**), 3, 199–214

FACTORS AFFECTING BIBLIOMETRIC INDICATORS OF SCIENTIFIC QUALITY

Jüri Allik

University of Tartu and Estonian Academy of Sciences

Abstract. The High Quality Science Index (HQSI) was constructed on the basis of the release of the *Essential Science Indicators* (Thomson Reuters) for the period from January 1, 2002 to August 31, 2012. The HQSI was computed for a country or territory as a sum of normalised scores of the mean impact (citations per paper) and the percentage of papers that reach the top-1% citation ranking. Expectedly, countries or territories that are producing larger Gross National Income per capita and allocate higher percentage of the produced economic wealth for the research and development (R&D) were more likely to achieve prominence in the scientific publications. The size of the country and its population were not important factors to excel in scientific research. Since economic and socio-demographic factors only partly predicted the quality of science in a given country or territory, there is considerable space for historical and science policy factors that could affect the quality of science in a given country. Several countries being in almost identical starting positions twenty years ago have developed on completely different trajectories dependent on policies and decisions made by their policy makers. Possibilities of how to improve reliability of measures of scientific quality have been discussed.

Keywords: bibliometric analysis, high quality science, impact factor, percentage of highly cited papers, country self-citation bias, mediocrity index

DOI: 10.3176/tr.2013.3.01

1. Introduction

The number of papers published in international peer-reviewed journals still serves as an indicator of scientific performance for countries with whom the authors of these papers are affiliated (King 2004, May 1997). For example, it was observed that China's growing share in the total number of articles published globally is now second only to the scientific world leader, the United States (Leydesdorff and Wagner 2009a, 2009b, The Royal Society 2011). A report issued by the Royal Society also noticed that Iran is the fastest growing country in terms of numbers of scientific publications in the world, growing from just 736 in 1996 to 13,238 in 2008 (The Royal Society 2011). At the same time Russia's annual growth in the number of publications has been minimal remaining approximately on the same absolute level during the last two decades (Adams and King 2010). However, the increase of the total number of publications is not an automatic guarantee of the similar increase in the quality of publications. Indeed, as it was observed by Eugene Garfield, out of about 38 million published papers indexed in the *Web of Science* (*WoS*, Thomson Reuters) approximately half were not cited at all (Garfield 2005). Probably for this reason the *Essential Science Indicators* (ESI, Thomson Reuters) identifies the 'essential core' of journals selecting the top 50% of journals by total citations in each of the 22 disciplines. But even if the increase of published papers in this 'essential core' is not accompanied by even more rapid increase of citations, it is a very problematic sign of progress. Understandably, policy makers and researchers themselves are increasingly more concerned about identification of indicators of high quality science (Moed 2005, Wagner and Leydesdorff 2012, van Leeuwen, Visser, Moed, Nederhof, and van Raan 2003). The average impact (citations per item) of all papers published by some country is certainly a more meaningful indicator of the scientific quality than a mere number of published papers. For example, it was shown that conventional bibliometric indicators, such as the number of citations, correlate with the number of Nobel Prize achievements in several advanced countries with similar research abilities (Rodriguez-Navarro 2011). Nevertheless, the analysis of the average impact should be supplemented by a careful inspection of the number of potential indicators such as highly cited papers or percentage of papers that are never cited (Wagner and Leydesdorff 2012, van Leeuwen et al. 2003).

It is not a well-guarded secret that high quality scientific research is mainly a privilege of rich nations. It is quite obvious that only these countries which are able to produce a sufficient economic wealth are also able to allocate a substantial amount of resources to scientific research and maintain a sufficient number of highly competent researchers. However, besides money high quality science depends on many other factors such as historical tradition, efficient educational system, healthy population, to say nothing about governmental policies that regulate research and development in a given country.

The three Baltic states, Estonia, Latvia and Lithuania, serve as a good example of how science could develop according to remarkably different paths (Allik 2008). In 1991 all these three countries, just liberated from the Soviet occupation, produced approximately 300 papers each in the journals indexed by the *WoS*. Twenty years later in 2011, Estonian scientists co-authored 1,863, Latvian 952, and Lithuanian 2,829 papers in journals, proceedings, and books indexed in *WoS*. Although the three countries had very similar political and economic histories they have remarkably different paths towards scientific distinction. In spite of identical starting positions, integration into the European Union, and joining NATO, scientific productivity of these three countries nowadays differs more than three times. Not only overall scientific productivity has developed along different trajectories in these neighbouring countries, but also the ability to promote and sustain scientific quality. Latvia was able to maintain the impact of its papers in spite of a relatively modest increase in productivity. Lithuania increased productivity almost 10 times during the last 20 years but largely at the expense of the impact of published papers. Out of the three Baltic states only Estonia managed to do both, to increase substantially the number of publication along with their average impact (Allik 2008).

The main goal of this paper is to look for the demographic and socio-economic factors that are relevant for high quality in scientific research. This goal is hopefully achieved by constructing a composite indicator of scientific excellence which combines the mean impact of published papers with the fraction of papers which achieve the status of highly cited papers. After instituting this index of scientific excellence the next step is to search for the best set of demographic and socio-economic variables which are associated with low or high position on this index.

2. Methods

Essential Science Indicators (*ESI*, Thomson Reuters) has been updated as of November 1, 2012 to cover a 10-year plus 8-month period, from January 1, 2002 to August 31, 2012. *ESI* operates on data received from *WoS* in which the whole science, except the humanities, is divided into 22 fields. In order to enter *ESI* there are different thresholds established for countries, institutions, individual scientists, single papers, and journals. About 5,000 journals out of the 10,000 are analysed which represent the top 50% by discipline and total citations over the 10-year and 8 months period. Journals are assigned uniquely to only one discipline (with the exception of multidisciplinary journals), but each country appears on average in about 13 disciplines. For countries of territories 148 are selected out of about 200 representing the top 50% by discipline and total citations over the 10-year period. Both *WoS* and *ESI* have their own tradition how to classify countries or territories. For example, Great Britain is represented by its four constituents England, Northern Ireland, Scotland, and Wales. The Bermudas are also listed as an independent entity, not as a British overseas territory. Although Taiwan is represented as an independent territorial unit, Hong Kong is absorbed into the general statistics of the People's Republic of China. Some political categories such as Yugoslavia which ceased to exist are still present. Since there is no fair method how to redistribute Yugoslavia's record between successor states these records were dismissed from the analysis.

Individual scientists, institutions and papers are included if they are among the top 1% most cited scientists, institutions, and papers respectively. The entrance thresholds are very different for different fields. For example, for a physicist to join the club of the top-1% most cited researchers it was necessary to collect at least 2,384 citations for publications appeared during the last 10 years and 8 months. The life for computer scientists was much easier because they needed to collect 'only' 186 citations to papers published during the last decade. Since older papers have more chances to be cited all citation thresholds are normalized not only relative to each field but also relative to time cohorts.

There are huge differences in the overall scientific productivity. Expectedly the most scientifically productive country was the U.S. whose scientists published 3,250,380 papers which were cited 51,546,380 times (15.9 per paper). The smallest contribution was made by Vatican scholars who had 94 papers which were cited 862 times (9.17 times per papers). In order to avoid a mixture of relevant players with small and often erratic participants it was decided to analyse only these countries whose scientists were able to publish 4,000 or more papers during the 10-years plus 8-month period. There were 82 countries or territories out of 148 whose scientists co-authored 4,000 or more papers during this period. Thus, there were 65 countries or territories that were not able to meet this 4,000-paper criterion. Among them were 8 countries producing between 3,000 and 4,000 papers during the observation period which were left out from the analysis: Sri Lanka (3,701), Republic of Georgia (3,686), Costa Rica (3,654), Oman (3,466), Luxembourg (3,400), Azerbaijan (3,219), and Ghana (3,125). Although the 4,000-paper threshold was established rather arbitrarily the included 82 countries produce more than 99% of all papers listed in the ESI.

From a large list of potentially important demographic and socio-economic variables I selected those which are typically regarded relevant for the human development: life expectancy on birth, schooling years, the size of population, median age, fertility and the Gross National Income per capita which were extracted from the latest Human Development Report (Human Development Report 2011). The research and development (R&D) expenditures (% of GDP) was taken from the World Bank statistics (http://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS). Some missing data were supplemented by data from other sources (http://www.nationmaster.com/graph/eco_res_and_dev_exp_of_gdp-economy-research-development-expenditure-gdp). The latest available data are reported. Nevertheless, there were 4 countries with missing R&D expenditure values.

3. Results

In Table 1 the main bibliometric, demographic, and economic indicators are shown for 82 countries/territories that were able to publish more than 4,000 papers during the 10-years and 8-months period. All countries/territories are ranked according to the impact (citations per paper) of published articles. According to this ranking, scientists working in Switzerland, Iceland, and Denmark were on the top of publishing papers with high citation. Papers co-authored by researchers working in Saudi Arabia, Algeria, and Serbia were on the bottom of the ranking. On average, Swiss papers were cited 16.83 times which is over 5 times more often than the mean impact of Serbian papers (3.13).

However, the mean number of citations per paper is not the only and perhaps the best indicator of the quality of scientific research. It is imaginable that researchers of some country produce a large number of mediocrely cited papers but are inept to publish outstanding papers which have exceptionally high citation rate. Therefore, as another indicator of scientific excellence I computed for each country the percentage of papers which have reached the top 1% citation rate in any of 22 science fields. Thus, unlike the mean number of citations this indicator is normalized relative to the citation baseline specific to each field. Expectedly, these two indicators (citations per paper and the percentage of highly cited papers) were highly correlated ($r = .83$, $p < .0001$). Since the mean impact of papers and the percentage of highly cited papers were correlated, it is possible to combine these two indicators into a composite measure the High Quality Science Index (HQSI). Before summation the number of citations per paper and the percentage of highly cited papers were separately transformed according to the following formula: $(X\text{-}M)/SD$ where *M* is the mean value of all *X*-s either citations per paper or the percentage of highly cited papers and *SD* is their standard deviation. Thus, to compute this index the both components were first normalized (the mean of the transformed values equals to zero and standard deviation equals to one) and then summed together (see the 6th column in Table 1).

It is anticipated that rich countries or territories have more resources, both physical and human, to produce high quality papers. Figure 1 demonstrates the relationship between the Gross National Income per capita (GNI, Human Development Report 2011) and HQSI. Although linear correlation between these two indices was reasonably high ($r = .65$, $p < .001$), there are obvious deviations from the linear regression. Without any doubt, Kuwait has rather modest results in science compared with the accumulated economic wealth. However, there is another group on countries whose science performance exceeds all expectations based on their economic indicators. Three African countries (Kenya, Tanzania, and Uganda) obviously benefited from cooperation with the leading countries in the world, especially with the United States, mainly studying maladies such as malaria and HIV that have inflicted populations and environment of these countries. Unlike other countries that form this exceptional group, Armenia is distinguished by a high quality of physics and space science. If I excluded three obvious deviants (Kuwait, Armenia, and Peru) the correlation between HQSI and GNI increased to the respectable .75.

Although the percentage of highly cited papers and the mean impact are substantially correlated, their difference can be used for the computation of what could be called Mediocrity Index. A scientific production of a country or territory which produces unexpectedly small number of highly influential papers compared with the mean impact of its published papers could be called ordinary or moderate quality, neither good nor bad. Aubtracting from the normalized score of the mean impact the normalized score of the percentage of highly cited papers, the highest scores on the Mediocrity Index was seen in Japan, Finland, and Sweden. The difference of Sweden and Finland from the Netherlands, Denmark and Switzerland

Table 1. Main bibliometric, demographic, and economic indicators for countries/territories publishing more than 4,000 papers

	Papers	Citations	Impact	Number of Highly Cited Papers	Missing Fields	High Quality Science Index	Life-expectancy	Schooling years	Population (millions)	Fertility	Median Age	Gross National Income	R&D (% of GDP)
SWITZERLAND	200 720	3 378 814	16.83	4 865	0	2.35	82.3	11.0	7.7	1.5	41.4	39 924	3.00
ICELAND	6 117	98 749	16.14	168	1	2.54	81.8	10.4	0.3	2.1	34.8	29 354	2.64
DENMARK	107 456	1 706 424	15.88	2 195	0	1.87	78.8	11.4	5.6	1.9	40.6	34 347	3.02
USA	3 250 380	51 546 027	15.86	59 513	0	1.68	78.5	12.4	313.1	2.1	36.9	43 017	2.79
NETHERLANDS	276 918	4 376 263	15.80	5 560	0	1.83	80.7	11.6	16.7	1.8	40.7	36 402	1.84
SCOTLAND	116 797	1 844 294	15.79	2 309	0	1.80	80.2	9.3	5.2	1.9	37.0	33 296	1.82
ENGLAND	747 867	11 306 492	15.12	13 768	0	1.58	80.2	9.3	53.0	1.9	39.8	33 296	1.82
SWEDEN	191 190	2 849 453	14.90	3 100	0	1.35	81.4	11.7	9.4	1.9	40.7	35 837	3.62
BELGIUM	152 108	2 137 230	14.05	2 745	0	1.40	80.0	10.9	10.8	1.8	41.2	33 357	1.96
FINLAND	95 314	1 300 687	13.65	1 336	0	0.98	80.0	10.3	5.4	1.9	42.0	32 438	3.84
GERMANY	845 448	11 340 845	13.41	12 697	0	1.04	80.4	12.2	82.2	1.5	44.3	34 854	2.82
CANADA	493 736	6 588 762	13.34	7 531	0	1.05	81.0	12.1	34.3	1.7	39.9	35 166	1.95
AUSTRIA	104 856	1 379 259	13.15	1 712	0	1.12	80.9	10.8	8.4	1.4	41.8	35 719	2.75
WALES	39 656	508 225	12.82	612	0	0.99	80.2	9.3	3.1	1.9	39.8	33 296	1.82
NORWAY	80 781	1 031 223	12.77	1 260	0	1.00	81.1	12.6	4.9	2.0	38.7	47 557	1.80
ISRAEL	117 799	1 497 439	12.71	1 609	0	0.82	81.6	11.9	7.6	2.9	30.1	25 849	4.27
FRANCE	603 328	7 556 253	12.52	8 125	0	0.77	81.5	10.6	63.1	2.0	39.9	30 462	2.23
AUSTRALIA	338 947	4 098 351	12.09	4 924	0	0.81	81.9	12.0	22.6	2.0	36.9	34 431	2.35
ITALY	470 648	5 689 120	12.09	5 906	0	0.63	81.9	10.1	60.8	1.5	43.2	26 484	1.27
NORTH IRELAND	19 471	234 682	12.05	235	0	0.58	80.2	9.3	1.8	1.9	39.8	33 296	1.82
IRELAND	52 419	622 442	11.87	832	0	0.90	80.6	11.6	4.5	2.1	34.7	29 322	1.77
NEW ZEALAND	63 682	707 805	11.11	858	0	0.58	80.7	12.5	4.4	2.1	36.6	23 737	1.17
PERU	4 858	53 260	10.96	92	1	1.05	74.0	8.7	29.4	2.4	25.6	8 389	0.10
KENYA	8 523	91 090	10.69	110	6	0.47	57.1	7.0	41.6	4.0	18.5	1 492	0.42
SPAIN	382 658	4 079 695	10.66	4 438	0	0.35	81.4	10.4	46.5	1.5	40.1	26 508	1.38

	Papers	Citations	Impact	Number of Highly Cited Papers	Missing Fields	High Quality Science Index	Life-expectancy	Schooling years	Population (millions)	Fertility	Median Age	Gross National Income	R&D (% of GDP)
JAPAN	803 857	8 390 864	10.44	6 121	0	-0.04	83.4	11.6	126.5	1.4	44.7	32 295	3.45
TANZANIA	4 305	44 809	10.41	58	8	0.48	58.2	5.1	46.2	5.5	17.5	1 328	0.43
SINGAPORE	75 883	786 658	10.37	1 172	1	0.65	81.1	8.8	5.2	1.4	37.6	52 569	2.66
HUNGARY	53 908	557 321	10.34	618	0	0.29	74.4	11.1	10.0	1.4	39.8	16 581	1.15
UGANDA	4 168	41 144	9.87	57	11	0.42	54.1	4.7	34.5	3.1	15.7	1 124	0.41
ESTONIA	10 231	99 834	9.76	154	0	0.53	74.8	12.0	1.3	1.7	39.7	16 799	1.44
PORTUGAL	73 918	691 573	9.36	794	0	0.09	79.5	7.7	10.7	1.3	41.0	20 573	1.66
PHILIPPINES	6 599	61 598	9.33	106	3	0.56	68.7	8.9	94.9	3.1	22.2	3 478	0.11
GREECE	92 420	854 244	9.24	897	0	-0.02	79.9	10.1	11.4	1.5	41.4	23 747	0.58
URUGUAY	5 390	49 394	9.16	45	3	-0.16	77.0	8.5	3.4	2.0	33.7	13 242	0.66
CHILE	39 335	347 398	8.83	367	0	-0.11	79.1	9.7	17.3	1.8	32.1	13 329	0.39
SOUTH AFRICA	61 564	534 352	8.68	730	0	0.09	52.8	8.5	50.5	2.4	24.9	9 469	0.93
ARGENTINA	63 824	534 486	8.37	479	0	-0.34	75.9	9.3	40.8	2.2	30.4	14 527	0.52
CZECH REPUBLIC	75 905	627 048	8.26	739	0	-0.16	77.7	12.3	10.5	1.5	39.4	21 405	1.53
THAILAND	38 864	310 728	8.00	276	0	-0.43	74.1	6.6	69.5	1.5	34.2	7 694	0.21
INDONESIA	7 773	61 525	7.92	81	2	-0.14	69.4	5.8	242.3	2.4	27.8	3 716	0.08
ARMENIA	5 042	36 972	7.33	122	16	1.02	74.2	10.8	3.1	1.7	32.1	5 188	0.27
SOUTH KOREA	325 403	2 385 604	7.33	2 293	0	-0.53	80.6	11.6	48.4	1.4	37.9	28 230	3.36
TAIWAN	202 868	1 485 933	7.32	1 331	0	-0.57	73.5	7.5	23.2	1.1	38.1	7 476	2.58
SLOVENIA	26 989	195 294	7.24	247	0	-0.35	79.3	11.6	2.0	1.5	41.7	24 914	1.86
MEXICO	83 984	604 675	7.20	574	0	-0.56	77.0	8.5	114.8	2.2	26.6	13 245	0.37
VENEZUELA	12 021	86 548	7.20	66	1	-0.69	74.4	7.6	29.4	2.4	26.1	10 656	0.40
CYPRUS	4 716	32 840	6.96	71	6	0.14	79.6	9.8	1.1	1.5	34.2	24 841	0.46
VIETNAM	8 811	60 854	6.91	74	2	-0.46	75.2	5.5	85.8	1.9	28.2	2 805	0.19
LEBANON	6 059	41 441	6.84	45	3	-0.56	72.6	7.9	4.3	1.8	29.1	13 076	n.a.
LATVIA	4 131	28 238	6.84	29	4	-0.60	73.3	11.5	2.2	1.5	40.2	14 293	0.46

	Papers	Citations	Impact	Number of Highly Cited Papers	Missing Fields	High Quality Science Index	Life-expectancy	Schooling years	Population (millions)	Fertility	Median Age	Gross National Income	R&D (% of GDP)
POLAND	170 212	1 159 713	6.81	1 142	0	-0.63	76.1	10.0	38.3	1.4	38.0	17 451	0.68
BULGARIA	20 371	138 122	6.78	126	0	-0.68	73.4	10.6	7.4	1.6	41.6	11 412	0.53
SLOVAKIA	26 068	174 555	6.70	203	0	-0.55	75.4	11.6	5.5	1.4	36.9	19 998	0.48
BANGLADESH	7 659	51 124	6.68	46	4	-0.71	68.9	4.8	150.5	3.2	24.2	1 529	n.a.
CAMEROON	4 245	28 298	6.67	15	5	-0.94	51.6	5.9	20.0	4.3	19.3	2 031	n.a.
COLOMBIA	16 632	108 842	6.54	204	0	-0.17	73.7	7.3	46.9	2.3	26.8	8 315	0.16
PEOPLES R CHINA	1 022 597	6 653 426	6.51	7 951	0	-0.58	73.5	7.5	1347.6	1.6	34.5	7 476	1.47
BRAZIL	244 250	1 581 093	6.47	1 232	0	-0.83	73.5	7.2	196.7	1.8	29.1	10 162	1.08
CUBA	7 615	47 476	6.23	35	4	-0.90	79.1	9.9	11.3	1.5	38.4	5 416	0.49
ETHIOPIA	4 177	25 724	6.16	20	8	-0.89	59.3	1.5	84.7	3.9	18.7	971	0.17
INDIA	334 660	2 059 832	6.15	1 606	0	-0.89	65.4	4.4	1241.5	2.5	25.1	3 468	0.76
CROATIA	24 973	147 839	5.92	207	0	-0.61	76.6	9.8	4.4	1.5	41.5	15 729	0.83
KUWAIT	6 188	34 967	5.65	21	5	-1.09	74.6	6.1	2.8	2.4	28.2	47 926	0.11
U ARAB EMIRATES	7 634	42 976	5.63	37	0	-0.96	76.5	9.3	7.9	1.7	30.1	5 993	n.a.
TURKEY	175 432	976 847	5.57	1 066	0	-0.86	74.0	6.5	73.6	2.0	28.3	12 246	0.85
EGYPT	42 453	223 793	5.27	157	0	-1.12	73.2	6.4	82.5	2.6	24.4	5 269	0.21
MOROCCO	12 107	62 976	5.20	66	1	-0.97	72.2	4.4	32.3	2.2	26.3	4 196	0.60
LITHUANIA	14 153	73 109	5.17	131	1	-0.63	72.2	10.9	3.3	1.5	39.3	16 234	0.84
RUSSIA	275 325	1 371 065	4.98	1 220	0	-1.09	68.8	9.8	142.8	1.5	37.9	14 561	1.25
JORDAN	8 574	42 071	4.91	50	6	-0.97	73.4	8.6	6.3	2.9	20.7	5 300	0.42
BYELARUS	10 453	47 822	4.57	92	10	-0.75	70.3	9.3	9.6	1.5	38.3	13 439	0.64
NIGERIA	13 557	61 245	4.52	42	2	-1.27	51.9	5.0	162.5	5.4	18.5	2 069	0.22
PAKISTAN	26 065	117 629	4.51	233	0	-0.75	65.4	4.9	176.7	3.2	21.7	2 550	0.46
ROMANIA	43 514	195 308	4.49	286	0	-0.97	74.0	10.4	21.4	1.5	38.5	11 046	0.48
IRAN	102 693	457 808	4.46	608	0	-1.03	73.0	7.3	74.8	1.6	27.1	10 164	0.79
TUNISIA	18 503	81 760	4.42	60	2	-1.28	74.5	6.5	10.6	1.9	28.9	7 281	1.10

	Papers	Citations	Impact	Number of Highly Cited Papers	Missing Fields	High Quality Science Index	Life-expectancy	Schooling years	Population (millions)	Fertility	Median Age	Gross National Income	R&D (% of GDP)
MALAYSIA	33 895	146 640	4.33	204	1	-1.04	74.2	9.5	28.9	2.6	26.0	13 685	0.63
UKRAINE	45 820	197 020	4.30	161	2	-1.27	68.5	11.3	45.2	1.5	39.3	6 175	0.86
SAUDI ARABIA	25 903	108 188	4.18	204	1	-0.89	73.9	7.8	28.1	2.6	25.9	23 274	0.08
ALGERIA	11 815	49 214	4.17	38	5	-1.31	73.1	7.0	36.0	2.1	26.2	7 658	0.07
SERBIA	19 983	62 530	3.13	205	3	-0.82	74.5	10.2	9.9	1.6	37.6	10 236	0.89

Notes: Impact = the mean number of citations per paper (Citations/Papers); Missing fields = the number of fields out of 22 in which the country or territory failed to pass 50% citation threshold; R&D = expenditures on the research and development as a percentage of GDP; latest available data; n.a. = not available.

in producing highly cited papers has been previously noted (Karlsson and Persson 2012). On the opposite pole of the scale were Cyprus, Armenia, and Serbia who published much more highly visible papers than it could be expected from the country's mean impact.

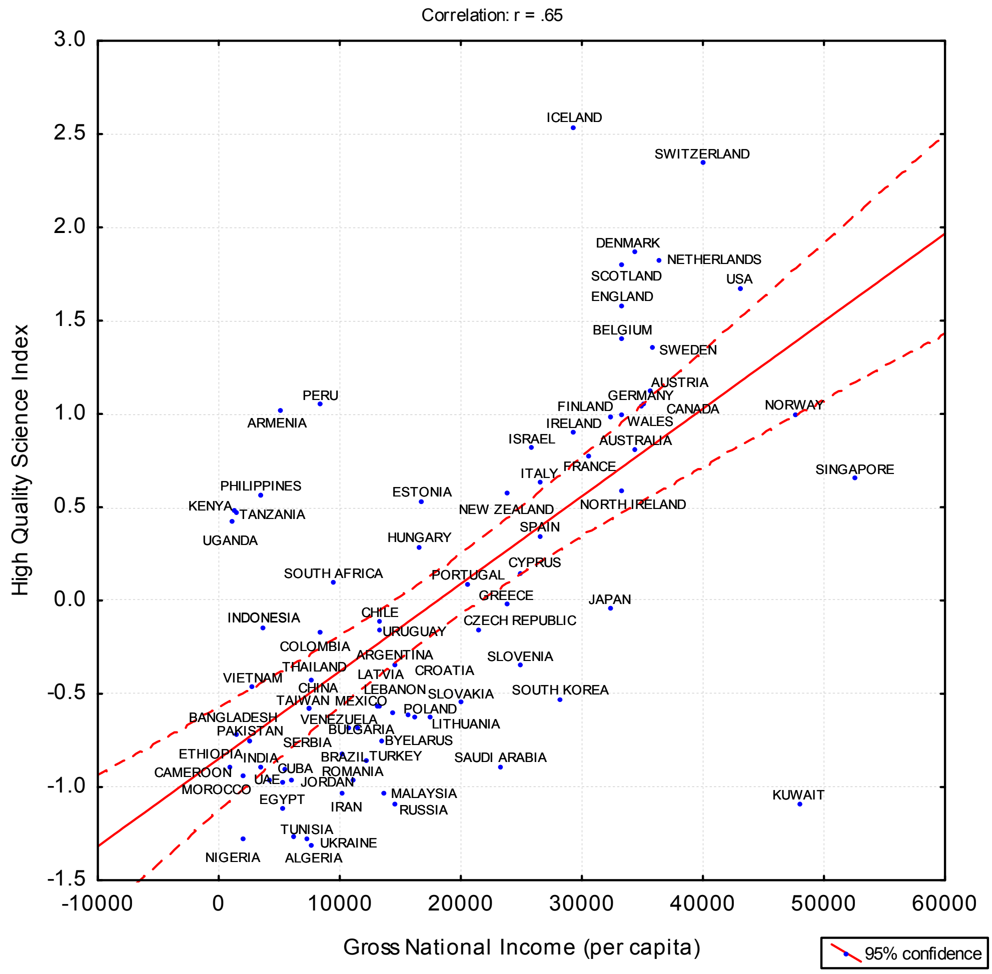

Figure 1. The relationship between the Gross National Income per capita and the High Quality Science Index.

Next I tried to predict standing on HQSI from demographic and socio-economic variables. Using forward step-wise multiple regressions it turned out that 6 relevant variables explained 54.3% of the HQSI variance: $R = .74$, $F(6,71) = 14.02$, $p < .001$. Expectedly, the largest contributions were made by the Gross National Income per capita ($\beta = 0.389$, $p = .005$) and the percentage of GDP that was spent on R&D (β = 0.284, p = .018). Thus, very rich nations spending also a considerable amount of the produced wealth on R&D are more likely to have high-quality papers. In addition, fertility (β = 0.311, p = .014), schooling years (β = 0.216, p = .075), the number of fields in which the threshold of the ESI was reached (β = 0.093, p = .090), and life-expectancy at birth (β = 0.161, p = .297) made their contribution in the explanation of the scientific excellence. It was expected that countries where people live and attend school longer have better chances to produce high quality science. It was more anticipated that countries with a lower fertility rate rather than those where the demographic revolution has not been accomplished yet are more successful in writing highly cited articles. In reality, fertility has a positive impact on scientific excellence when all other factors were controlled. Finally, it seems to be easier to achieve high quality of science selectively in some fields than in all 22 fields simultaneously.

4. Discussion

What seems to be obvious even without sophisticated computations is that the size of the country has a negligible effect on publishing highly visible and influential papers in the most prestigious scientific journals. A tiny Iceland with the population of 320,000 and spending 2.64% of its GDP for R&D is on the top of scientific excellence. It is also not surprising that the top of the HQSI is occupied by relatively small countries or territories such as Switzerland, Denmark, Netherlands, and Scotland who all achieved scientific excellence beyond what could be predicted on the basis of their economic indicators alone. Out of ten most populous countries only the USA is on the top of countries or territories ranked by their scientific excellence measured by bibliometric indicators. The rest like China, India, and Indonesia are all in the lower half of the scientific quality ranking. The former superpower Russia occupies 76th position out of 82 countries or territories being in terms of its scientific impact virtually on the same level with Nigeria and Pakistan.

It was certainly instructive to observe what has happened to the former Communist bloc countries. After the dissolution of the Soviet Union its successor, the Russian Federation has obviously failed to modernize it science which structure still reflects science of the 20th century, not science of the 21st century which is characteristically dominated by medical and life sciences (Adams and King 2010). The relative output of Russian scientists has remained basically on the same level but the impact of their papers has even decreased (Markusova, Jansz, Libkind, Libkind, and Varshavsky 2009). Surprisingly, Armenia outperformed other former Communist bloc countries in terms of the mean impact and proportion of highly cited papers. This seems to be achieved by concentrating material and human resources into only few research areas. Armenian scientists have passed the threshold of essential science only in 6 out of 22 research areas. Out of 122 Armenian papers that have achieved the status of highly cited papers

101 were co-authored by physicists. Many of these physics papers were written collectively in CERN by multinational groups of researchers where the number can exceed several thousand. If Armenia could be discarded as an artefact created by excellent performance of only one or two groups then two former Communist bloc achievers are Estonia and Hungary. They passed the threshold of essential science in all 22 categories, increased steadily the mean impact of their publications, and managed to publish a remarkable percentage of papers which were in the top-1% citations. Scientific performance of Estonia and Hungary is even more astonishing since their governments spend a very modest amount of their GDP on the research and development 1.44% and 1.15% respectively which is far less of what was promised in the Lisbon strategy's 3% of GDP. These two countries are followed by Czech Republic, Slovenia, and Croatia who performed in science slightly below of what could be expected from the level of their economic development. All other former Soviet bloc countries and especially former Soviet republics, except perhaps Armenia and Estonia, have underperformed concerning scientific excellence.

Although economic wealth is one of the major factors of the high quality science, many ranking positions can be explained by political decisions made in the past. For example, it is generally acknowledged that Latvia made a costly mistake by practically abolishing permanent science funding after regaining independence in 1991 replacing it almost totally with relatively insecure and insufficient temporary grant money (Allik 2003, Kristapsons, Martinson, and Dagyte 2003). In spite of the substantial brain-drain Latvia succeeded in maintaining a relatively high impact of its publications but only at the expense of a modest general productivity. On the other hand, Lithuania decided to increase the overall scientific productivity which was not accompanied with an equal concern about the quality of scientific publications. Dozens of local Lithuanian journals were established which publish not only in English but Lithuanian as well. These local journals have obviously less stringent criteria for the acceptance of scientific contributions especially when they come from the Lithuanian authors. When the Thomson Reuters announced on May 28, 2008 that nearly a thousand new regional journals have been added to the *WoS*, Lithuania managed to include 27 of their own local journals. As a consequence, the impact of Lithuanian scientific papers has dramatically dropped compared to all other countries in the world. One likely candidate for the decrease of the impact is a diminished motivation of Lithuanian researchers to publish their papers in the most prestigious and competitive international journals since it is much easier to publish in one of these local journals which impact is relatively low. As a consequence, the impact of Lithuanian papers has dramatically decreased dropping on the same level with Russia (Allik 2011). Unfortunately, there is no information of how many Lithuanian journals are in the top 5,000 journals selected by ESI.

Out of the three Baltic countries only Estonia was able to increase not only the volume of its publications but also their mean impact. There are several reasons for this accomplishment from which several are worth mentioning . From the beginning of regaining independence in 1991 all research money applications in Estonia were also written in English which allowed using foreign experts in evaluation, to say nothing about an invaluable practice for writing successful grant applications. Perhaps even more crucial was the fact that scientific assessment and decision-making was given to panels consisting of scientist who were mandated to make sovereign decisions that have been rarely reversed by non-scientific authorities. These factors in addition to a tight cooperation with researchers from the world's leading sciences, especially with Finland and Sweden, were essential in the growth of excellence in Estonian science (Allik 2008, 2011).

Although the EU share of total publications is greater than that of the U.S. in many separate disciplines as well as in all sciences as a whole, the U.S. almost always surpasses the EU when it concerns the upper tail of citation distributions, that is the number of highly cited papers (Albarran, Crespo, Ortuno, and Ruiz-Castillo 2010, Bornmann and Leydesdorff 2013, Leydesdorff and Wagner 2009a). In the light of these data it is surprising that the U.S. occupies only the 6th position in the ranking of high quality science. This is even less expected because the U.S. authors are subjected to one of the highest known country self-citation bias (Allik 2013a, 2013b, Jaffe 2011). Since there seems to be a pervasive tendency that researchers residing in the U.S. are more likely to cite an article by U.S. authors rather than by non-U.S. authors, the mean impact and highly cited articles scores for the U.S. may be inflated. Of course, many other countries, such as China and Iran, demonstrate similar country self-citation bias (Jaffe 2011). However, when it comes to deciding to cite or not to cite a previously published paper, the U.S. authors give more than others a preference to papers of those colleagues whose research questions, used methods, and proposed theories they know best (Allik 2013b).

There are many ways how measures of scientific excellence can be improved. Although several new measures of scientific excellence have been proposed it is obvious that only a combination of several of them could achieve sufficient reliability (Wagner and Leydesdorff 2012, van Leeuwen et al. 2003). Recent years have witnessed a considerable increase in the number of papers with authors in excess of 50 and it is not unusual to have reports whose author counts exceed 1,000 or even more. Since many of these multiauthorship papers concern hot data produced by large international collaborative projects such as Large Hadron Collider, Laser Interferometer Gravitational Wave Observatory, or Antiretroviral Therapy Cohort Collaboration even smaller stakeholders share an equal amount of recognition with the principal investigators of these large projects. Since the number citations is equally distributed between all co-authors small countries can potentially have an easy profit which was actually earned by others. However, it seems to be a considerable waste of time investing energy into inventing algorithms which could differentiate individual contributions of authors or countries that are behind them (cf. Põder 2010). Nevertheless, it would be instructive to know the ratio between highly cited papers written in cooperation with authors from many countries and papers authored exclusively by authors working in the same country.

However, the largest potential for improving measures of scientific excellence is by differentiating fields of science (Aksnes, Schneider, and Gunnarsson 2012, Leydesdorff and Bornmann 2011, Leydesdorff and Opthof 2010). For example, the average citation rate for papers published in the period 2002-2012 is quite different for various fields. The average citation rate papers published in molecular biology and genetics was 23.49 and in immunology 21.10. On the other hand, papers in mathematics and computer sciences were cited with average rate 3.53 and 4.07 citations respectively. This means that a country or territory whose strength is in mathematics and computer science has at least five times less chances to influence ranking on the scale of science quality than some other country who has invested in the development of molecular biology, genetics, and immunology. A telling example is Tanzania. Out of 58 highly cited papers which were co-authored by Tanzanians 28 were published in journals classified in the clinical medicine category. Since the average citation rate in clinical medicine is relatively high (12.62 citation per paper) it was more advantageous for Tanzania to develop its strength in this field rather than in agriculture (7.24 citations/papers) or economics or business (6.49 citations/papers). Fortunately, subfield normalization seems to change the global country ranking only marginally (Aksnes et al. 2012, Herranz and Ruiz-Castillo 2012). The proposed HQSI is a half-step in the desired direction since one of its components – the percentage of papers which have reached the top 1% citation rate in one of 22 science fields – takes differences between fields into account.

Address:
Jüri Allik
Department of Psychology
University of Tartu
Näituse 2
Tartu 50409, Estonia
E-mail: juri.allik@ut.ee

References

Adams, J. and C. King (2010) *Global research report: Russia. Research and collaboration in the new geography of science*. Leeds: Evidence.

Aksnes, Dag W., Jesper W. Schneider, and Magnus Gunnarsson (2012) "Ranking national research systems by citation indicators. A comparative analysis using whole and fractionalised counting methods". *Journal of Informetrics* 6, 1, 36–43. doi: 10.1016/j.joi.2011.08.002

Albarran, Pedro, Juan A.Crespo, Ignacio Ortuno, and Javier Ruiz-Castillo (2010) "A comparison of the scientific performance of the US and the European union at the turn of the 21st century". *Scientometrics* 85, 1, 329–344. doi: 10.1007/s11192-010-0223-7

Allik, J. (2003) "The quality of science in Estonia, Latvia, and Lithuania after the first decade of independence". *Trames* 7, 1, 40–52.

Allik, J. (2008) "Quality of Estonian science estimated through bibliometric indicators (1997–2007)". *Proceedings of the Estonian Academy of Sciences* 57, 255–264.

Allik, J. (2011) "Estonian science estimated through bibliometric indicators". In J. Engelbrecht, ed. *Research in Estonia: present and future*, 456–469. Tallinn: Estonian Academy of Sciences.

Allik, J. (2013a) "Bibliometric analysis of the journal of cross-cultural psychology during the first ten years of the new millennium". *Journal of Cross-Cultural Psychology* 44, 4, 657–667. doi: 10.1177/0022022112461941

Allik, J. (2013b) "Personality psychology in the first decade of the new millennium: a bibliometric portrait". *European Journal of Personality* 27, 1, 5–14. doi: 10.1002/per.1843

Bornmann, Lutz, and Loet Leydesdorff (2013) "Macro-indicators of citation impacts of six prolific countries: In Cites data and the statistical significance of trends". *Plos One* 8, 2. doi: 10.1371/journal.pone.0056768

Garfield, E. (2005) "The agony and the ecstasy – the history and meaning of the journal impact factor". Paper presented at the International Congress on Peer Review and Biomedical Publication, Chicago, September 16, 2005.

Herranz, Neus, and Javier Ruiz-Castillo (2012) "Sub-field normalization in the multiplicative case: average-based citation indicators". *Journal of Informetrics* 6, 4, 543–556. doi: 10.1016/j.joi.2012.02.006

Human Development Report (2011*) Human development report 2011: sustainability and equity: a better future for all*. New York: United Nations Development Programme.

Jaffe, K. (2011) "Do countries with lower self-citation rates produce higher impact papers? Or, does humility pay?". *Interciencia* 36, 9, 694–698.

Karlsson, S., and O. Persson (2012) *The Swedish production of highly cited papers*. (Vetenskapsrådet lilla rapportserie, 5.) Stockholm: Vetenskapsrådet.

King, D. A. (2004) "The scientific impact of nations". *Nature* 430, 6997, 311–316. doi: 10.1038/430311a

Kristapsons, J., H. Martinson, and I. Dagyte (2003) *Baltic R&D systems in transitions: experiences and future prospects*. Riga: Zinatne.

Leydesdorff, Loet, and Lutz Bornmann (2011) "How fractional counting of citations affects the impact factor: normalization in terms of differences in citation potentials among fields of science". *Journal of the American Society for Information Science and Technology* 62, 2, 217–229. doi: 10.1002/asi.21450

Leydesdorff, Loet, and Tobias Opthof (2010) "Normalization at the field level: fractional counting of citations". *Journal of Informetrics* 4, 4, 644–646. doi: 10.1016/j.joi.2010.05.003

Leydesdorff, Loet, and Wagner, Caroline. (2009a) Is the United States losing ground in science? A global perspective on the world science system. Scientometrics, 78(1), 23–36. doi: 10.1007/s11192-008-1830-4

Leydesdorff, Loet, and Caroline Wagner (2009b) "Macro-level indicators of the relations between research funding and research output". *Journal of Informetrics* 3, 4, 353–362. doi: 10.1016/j.joi.2009.05.005

Markusova, V. A., M. Jansz, A. N. Libkind, I. Libkind, and A. Varshavsky (2009) "Trends in Russian research output in post-Soviet era". *Scientometrics* 79, 2, 249–260. doi: 10.1007/s11192-009-0416-0

May, R. M. (1997) "The scientific wealth of nations". *Science* 275, 5301, 793–796. doi: 10.1126/science.275.5301.793

Moed, H. F. (2005) *Citation analysis in research evaluation*. Dordrecht: Springer.

Põder, E. (2010) "Let's correct that small mistake". *Journal of the American Society for Information Science and Technology* 61, 12, 2593–2594. doi: 10.1002/asi.21438

Rodriguez-Navarro, A. (2011) "Measuring research excellence number of Nobel Prize achievements versus conventional bibliometric indicators". *Journal of Documentation* 67, 4, 582–600. doi: 10.1108/00220411111145007

The Royal Society (2011) *Knowledge, networks and nations: global scientific collaboration in the 21st century*. London: The Royal Society.

Wagner, Caroline S. and, Loet Leydesdorff (2012) "An integrated impact indicator: a new definition of 'impact' with policy relevance". *Research Evaluation* 21, 3, 183–188. doi: 10.1093/reseval/rvs012

van Leeuwen, T. N., M. S. Visser, H. F. Moed, T. J. Nederhof, and A. F. J. van Raan (2003) "Holy Grail of science policy: exploring and combining bibliometric tools in search of scientific excellence". *Scientometrics* 57, 2, 257–280. doi: 10.1023/A:1024141819302