# International Collaboration and Counting Inflation in the Assessment of National Research Productivity

**Mu-hsuan Huang**
Dept. of Library & Information Science
National Taiwan University
1, Sec. 4, Roosevelt Rd., Taipei 116, Taiwan
Email: mhhuang@ntu.edu.tw

**Chi-Shiou Lin**
Dept. of Library & Information Science
National Taiwan University
1, Sec. 4, Roosevelt Rd., Taipei 116, Taiwan
Email: chishioulin@ntu.edu.tw

## ABSTRACT

In this paper we describe how different accounting procedures affected the counting of scientific paper numbers at the country level and the country ranks based on paper production quantity in physics. Using 1989-2008 citation data, we also report the counting inflation ratio between different accounting procedures. We found that, in general, different accounting procedures yielded relatively similar and stable rankings. But for certain clusters of countries, the normal count procedure tended to favor the more advanced Western countries. In contrast, the newly developed countries received more credit in the adjusted and straight count procedures.

## Keywords
Research productivity, counting inflation, bibliometrics

## INTRODUCTION
The counting of scientific papers numbers is fundamental to the assessment of research productivity. However, in the age of increasing scientific collaboration across national borders, it becomes conceptually and methodologically challenging to conduct counting for an internationally collaborated paper to show each country's respective contribution. Pravdić & Oluić-Vukovic (1986, 1991) summarized four basic accounting procedures for assessing scientific productivity:

- Normal count: equal credit is given to all contributors; one full unit is assigned to each author or each country involved in a scientific paper.

- Adjusted count (fractional authorship): each multi-authored paper is divided by the number of all authors; productivity is expressed as a score obtained by summing only the corresponding parts of each paper.

- Straight count: only the first author or whose country receives full credit of a multi-authored paper.

- Modified straight count: each paper as a unit is allocated to the more productive author rather than the

first author; the identification of authors receiving credits is via sophisticated computation of the authors' collaboration links (Pravdić & Oluić-Vukovic, 1986).

None of the accounting procedures alone was ideal for assessing research productivity. Pravdić & Oluić-Vukovic (1986) suggested that a "dual approach" combining the normal count and modified straight count might be a better way for assessing research productivity. They further pointed out that assessment at individual and national levels faced different problems.

In this paper, we are concerned with another problem resulted from the use of different accounting procedures, namely, the counting inflation of research productivity at the national level. Although none of the aforementioned accounting procedures is perfect for research productivity assessment, the more convenient methods like normal count, adjusted count, and straight count continue to be used in large scale evaluation programs due to the availability and size of data and the ease of computation. When normal count is used, each country will unavoidably receive more paper counts than in other methods. The changes in paper numbers may consequently affect country rankings and obscure observations of development trends and directions in international scientific knowledge production and competition.

We are interested in whether different accounting procedures might affect the comparative rankings of national research performance in a real assessment program and to what extent paper numbers may inflate when different accounting procedures are used. This short paper reports results in research productivity. Our future work continues to examine the consequences of procedure selection in assessing other aspects, e.g., research impact. Our dataset is the complete 1989-2008 citation data of the physics journals from Thomson Reuter's ISI Web of Science (WoS). This large and comprehensive dataset allows us to depict a highly accurate picture of how accounting procedures affected country ranks and the extent of counting inflation in a scientific field characteristic of heavy and intensifying international collaboration.

## METHODOLOGY

We used the citation data from ISI WoS for the analysis. In October, 2008, 336 journals were listed under the category of physics in Essential Science Indicators (ESI). We analyzed the citation data of the past twenty years of the physics journals (1989-2008). Within the time frame, a total of 1,445,273 papers with authors from 165 countries were published in those 336 journals.

Our research team wrote a program to automatically parse the WoS data. For each citation entry, we calculated the author number based on the names recorded in the author field (AU). International collaboration papers were defined as co-authored papers with authors from different countries. Because the WoS data lacked information on each author's nationality, we used the nationalities of the authors' institutions instead. We determined whether a paper was internationally collaborated based on the author address field (C1), which theoretically included all authors' institution addresses including the first author's, and the corresponding author address field (RP), which listed only the corresponding author's institution address. It should be noted that 41,390 entries (2.86% of the total entries) lacked data in the C1 and RP fields and were purged from our ranking and counting inflation analyses.

We employed the following five accounting procedures to calculate the numbers of each country's international collaboration papers.

A. *All authors count* (normal count): regardless of the order of authorship, if the collaborating authors' institutions recorded in C1 are located in different countries, then each country is considered as having produced one paper. For example, if a paper has five authors from three institutions located in two different countries, then the two countries each receive one paper count.

B. *First author counts* (straight count): only the first institution address listed in C1 is counted. In our dataset, 250,958 entries (17.36% of the total) lacked the C1 field, and we used RP instead.

C. *Corresponding author counts* (straight count): only the institution address in RP is counted. 40,697 entries (2.82%) lack the RP field, and we used the first address in C1 instead.

D. *Divided share (I)* (adjusted count): all the institutions listed in C1 were used to count the national paper production; each institution received equal share. For example, if a paper has five authors from three institutions, two of which are located in country X and one in country Y, then X receives 2/3 paper count, and Y receives 1/3. Please note that 2.82% of our dataset lacked the C1 field, and RP was used instead. Because the RP field records only one institution address, these entries unavoidably were counted as being produced by

one single country, and thus misjudgment may have occurred in the calculation.

E. *Divided share (II)* (adjusted count): regardless of the numbers of collaborating institutions in C1, only the nationalities of the institutions are considered. For example, if a paper has five authors from three institutions which are located in two different countries, then each country receives 1/2 paper count.

## FINDINGS

### The Physics Papers in 1989-2008

The 1,445,273 physics papers published in 1989-2008 were produced by 6,658,522 authors. Each paper had 4.61 authors. Among them, 1,189,863 (82.33% of the total) were co-authored papers. The total number of authors of the collaboration papers was 6,403,112. That is, each collaborated paper had 5.38 authors in average.

International collaboration papers were of particular importance in assessing national research productivity. 329,447 (22.79%) of the entire papers were of international co-authorship judged by authors' institutional affiliations. The total number of authors of the international collaboration papers was 3,135,587. Each international collaboration paper had 9.52 authors in average. One can see that, although international collaboration papers only accounted for 22.79% of the physics papers, the huge difference between the paper number and author number and the high average author number per paper have anticipated noticeable counting inflation.

### Country Ranks by the Five Accounting Procedures

Here we report the rank changes in the top 30 countries by any of the five accounting procedures. As shown in Table 1, slight variation existed when different accounting procedures were applied, but the distribution of country ranks was rather similar in each method. United States ranked first in all of the five rankings; its paper production was approximately 2.3-2.4 times of the second country in quantity (i.e., Germany or Japan, depending on the procedures used). At the other end, Singapore and Hungary could both fall out of the top 30 when a certain procedure was applied. Hungary made it into the top 30 only in procedure A. In contrast, Singapore was able to hit the top 28 in all procedures but A.

Some countries' ranks never changed no matter which accounting procedure was applied, i.e., United States (1), Russia (5), France (6), U.K. (7), Italy (8), India (9), Poland (13), and Czech (25). For the other countries of varied ranks, a closer examination revealed some interesting patterns. First, one can identify several clusters of countries with adjacent ranks. Within each cluster, country ranks varied by procedure, but ranks were interchangeable only within the same cluster. For instance, in the first cluster of Germany, Japan, and China, one can see three different orders of ranks: 2-3-4 (A); 4-2-3 (B; D); 3-2-4(C; E). It should be noted that Japan was ranked as the 2nd in all procedures but

A. For the ranking variations in the other clusters, see Table 1. The two larger clusters are located at the lower half of the table, i.e., the cluster of Switzerland to Taiwan; and the cluster of Denmark to Singapore. An interesting observation of each country's rank changes due to procedures used is that the largest difference of ranks did not exceed 3 even in clusters composed of more countries, e.g., Taiwan ranked 15th the highest and 18th the lowest; Argentina ranked between 26th and 29th, and Singapore between 28th and 31st.

Another interesting observation is that, when procedure A was applied, the Western countries usually ranked higher than their peer countries within the same clusters. In contrast, when procedures B-D were applied, those countries of East Asia and those that can be described as "newly industrializing economies" (Central Intelligence Agency, n.d.) or "emergent markets" (CME Group Index Services, 2010) were ranked higher than their Western peers – for example, Japan and China as opposed to Germany (the 1st cluster); South Korea as opposed to Canada and Spain (the 2nd cluster); Brazil and Taiwan as opposed to Switzerland, Netherlands, and Australia (the 3rd cluster); Ukraine and Mexico as opposed to Belgium and Austria (the 4th cluster); Argentina and Singapore as opposed to their cluster peers (the 5th cluster). The more dramatic rank changes can observed in the following countries: South Korea rose from 12 (procedure A) to 10 (procedures B-D); Taiwan rose from 18 (A) to 15 (B; D); Argentina from 29 (A) to 26 (B; D); Singapore from 31 (A) to 28 (B-D). In contrast, Germany, Switzerland, and Switzerland could drop two ranks; Hungary dropped three in all four procedures.

Finally, comparing the results of the five different procedures, one can see that the results from procedures B-D were rather consistent in the direction of rank change. That is, when any of the procedures other than A was applied, a country's country rank could rise, remain the same, or drop. But the direction in terms of rising or dropping was all the same. Curiously, procedures B versus D, as well as procedures C versus E, yielded nearly identical rankings in the top 30 countries except Denmark and Argentina. This curious high similarity awaits further investigation. Given the similarity of the procedures B-D rankings, it seems safe to conclude that, among various accounting procedures, only procedure A made a larger difference to paper number counting.

## Counting Inflation
We further used the paper numbers from procedure A as the basis for calculating the ratio of counting inflation. We divided each country's paper number from procedure A by the numbers from procedures B-D respectively. Results showed that counting inflation ranged from as low as 1.11 (China in procedure B) to as high as 1.84 (Hungary in procedure C) (see Table 1). For the convenience of observation, we used different background colors for ratio

values within four equal sections between 1.00 and 2.00 (1.00-1.25; 1.26-1.50; 1.51-1.75; 1.76 and above).

One can see that, except the U.S., all other countries with inflation ratio lower than 1.25 were Asian countries, i.e., Japan, China, India, South Korea, Taiwan, and Singapore. Only Singapore had inflation ratio higher than 1.25 in two accounting procedures. The low inflation suggests that those countries possibly had not involved as much as the other countries in international collaboration.

All the Western countries within the top 10 countries including Russia had inflation ratio ranging between 1.26-1.50. Other Western countries in top 20-30 could have higher inflation. Switzerland, Denmark, Austria, and Hungary had the highest inflation as compared to other countries. The higher inflation ratio indicates that those countries, when participating in international collaboration, more often served supporting or facilitating roles rather than as the lead investigator.

It makes sense to see the degree of inflation echoed the rank rise and rank drop within each cluster of countries. Countries with lower inflation ratio all rose in ranks when procedures B-D were applied. In contrast, countries with relatively higher ratio in each cluster dropped in ranks when procedures B-D were used for paper counts. The wider gaps of counting inflation can be observed in the 3rd cluster (i.e., Switzerland vs. Taiwan) and the 7th cluster (i.e., Denmark and Hungary vs. Singapore).

## DISCUSSION & CONCLUSION
A major contribution of this project is that, compared to studies using sample data (e.g., Glanzel, 2002; Golnabi & Mahdieh, 2006; Kao, 2009), our large and relatively complete citation data of the physics discipline can yield more robust understandings of what really happened in the international scientific research arena. As one can see, country ranks were not affected greatly by accounting procedures. This suggests that all of the five accounting procedures are capable of yielding reliable comparisons of national research productivity.

Variations of country ranks within clusters showed that procedure A (normal count) generally favored the more advanced Western countries. The newly developed countries were accredited more in the straight count and adjusted count procedures. The differences between procedure A and the others may have occurred for two possible reasons. Firstly, institutions in the newly developed countries may have assumed leadership roles in international collaboration papers. But it was less discernible in the normal count approach in which every participating country received equal credit. Alternatively, it is likely that the rank differences were due to less participation of the newly developed countries in international collaboration. Our next step is to examine the extent of international and intra-national collaboration of each country to understand the cause.

In assessing research productivity, paper quantity is only one basic measure. Evaluations of research performance should incorporate both quantity and quality measures. The next logical step is to examine how the papers were cited and whether different accounting procedures also affected country-by-country citation counts. Through comparisons of quantity and quality measures based on authentic data, we will be able to identify the effects of accounting procedures in bibliometric analyses.

## REFERENCES

Central Intelligence Agency. (n.d.). *The world factbook: Appendix B: International organizations and groups*. Retrieved May 31, 2010, from: https://www.cia.gov/library/publications/the-world-factbook/appendix/appendix-b.html#N

CME Group Index Services. (2010). *Dow Jones total stock market indexes*. Retrieved May 31, 2010, from: http://www.djindexes.com/mdsidx/downloads/brochure_info/Dow_Jones_Total_Stock_Market_Indexes_Brochure.pdf

Glanzel, W. (2002). Coauthorship patterns and trends in the science (1980-1998): A bibliometric study with implications for database indexing and search strategies. *Library Trends*, *50*(3), 461-473.

Golnabi, H., & Mahdieh, M. H. (2006). Trend of laser research developments in global level. *Optics and Laser Technology*, *38*(2), 122-131.

Kao, C. (2009). The authorship and country spread of Operation Research journals. *Scientometrics*, *78*(3), 397-407.

Pravdić, N., & Oluić -Vukovic, V. (1991). Distribution of scientific productivity: Ambiguities in the assignment of author rank. *Scientometrics*, 20(1), 131-144.

Pravdić, N., & Oluić-Vuković, V. (1986). Dual approach to multiple authorship in the study of collaboration/ scientific output relationship. *Scientometrics*, *10*(5), 259-280.

.

| Country | 1989-2008 Paper Numbers by Different Accounting Procedures | | | | | Country Rank by Paper Count * | | | | | Ratio of Counting Inflation ** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | A | B | C | D | E | A/B | A/C | A/D | A/E |
| USA | 381,902 | 310,987 | 310,154 | 314,185 | 311,657 | 1 | 1 | 1 | 1 | 1 | 1.23 | 1.23 | 1.22 | 1.23 |
| Germany | 159,718 | 109,210 | 110,667 | 108,125 | 110,444 | 2 | 4 | 3 | 4 | 3 | 1.46 | 1.44 | 1.48 | 1.45 |
| Japan | 157,485 | 135,041 | 135,061 | 135,970 | 136,439 | 3 | 2 | 2 | 2 | 2 | 1.17 | 1.17 | 1.16 | 1.15 |
| China | 122,418 | 110,189 | 109,655 | 108,352 | 108,074 | 4 | 3 | 4 | 3 | 4 | 1.11 | 1.12 | 1.13 | 1.13 |
| Russia | 118,943 | 89,044 | 87,291 | 90,045 | 90,672 | 5 | 5 | 5 | 5 | 5 | 1.34 | 1.36 | 1.32 | 1.31 |
| France | 115,471 | 78,270 | 78,124 | 79,640 | 80,075 | 6 | 6 | 6 | 6 | 6 | 1.48 | 1.48 | 1.45 | 1.44 |
| UK | 102,943 | 72,637 | 72,739 | 72,717 | 74,502 | 7 | 7 | 7 | 7 | 7 | 1.42 | 1.42 | 1.42 | 1.38 |
| Italy | 75,820 | 54,118 | 53,560 | 54,321 | 51,812 | 8 | 8 | 8 | 8 | 8 | 1.40 | 1.42 | 1.40 | 1.46 |
| India | 47,917 | 41,250 | 41,109 | 40,879 | 40,934 | 9 | 9 | 9 | 9 | 9 | 1.16 | 1.17 | 1.17 | 1.17 |
| Canada | 43,892 | 30,170 | 30,393 | 30,364 | 31,080 | 10 | 11 | 11 | 11 | 11 | 1.45 | 1.44 | 1.45 | 1.41 |
| Spain | 41,946 | 28,853 | 29,198 | 27,958 | 28,327 | 11 | 12 | 12 | 12 | 12 | 1.45 | 1.44 | 1.50 | 1.48 |
| South Korea | 40,379 | 32,984 | 33,179 | 32,768 | 33,109 | 12 | 10 | 10 | 10 | 10 | 1.22 | 1.22 | 1.23 | 1.22 |
| Poland | 37,931 | 24,803 | 24,884 | 25,031 | 25,670 | 13 | 13 | 13 | 13 | 13 | 1.53 | 1.52 | 1.52 | 1.48 |
| Switzerland | 35,055 | 19,528 | 20,141 | 19,145 | 20,321 | 14 | 16 | 15 | 16 | 15 | 1.80 | 1.74 | 1.83 | 1.73 |
| Brazil | 28,183 | 21,462 | 21,595 | 20,972 | 21,185 | 15 | 14 | 14 | 14 | 14 | 1.31 | 1.31 | 1.34 | 1.33 |
| Netherlands | 27,560 | 17,560 | 17,661 | 17,687 | 18,209 | 16 | 17 | 17 | 17 | 17 | 1.57 | 1.56 | 1.56 | 1.51 |
| Australia | 23,651 | 17,390 | 17,494 | 17,261 | 17,515 | 17 | 18 | 18 | 18 | 18 | 1.36 | 1.35 | 1.37 | 1.35 |
| Taiwan | 23,622 | 19,955 | 20,127 | 19,753 | 19,930 | 18 | 15 | 16 | 15 | 16 | 1.18 | 1.17 | 1.20 | 1.19 |
| Sweden | 23,528 | 14,457 | 14,663 | 14,324 | 14,709 | 19 | 20 | 20 | 20 | 20 | 1.63 | 1.60 | 1.64 | 1.60 |
| Israel | 21,976 | 15,182 | 15,360 | 14,910 | 15,343 | 20 | 19 | 19 | 19 | 19 | 1.45 | 1.43 | 1.47 | 1.43 |
| Belgium | 18,511 | 11,426 | 11,615 | 11,096 | 11,539 | 21 | 22 | 22 | 22 | 22 | 1.62 | 1.59 | 1.67 | 1.60 |
| Ukraine | 17,785 | 12,523 | 12,381 | 12,550 | 12,823 | 22 | 21 | 21 | 21 | 21 | 1.42 | 1.44 | 1.42 | 1.39 |
| Austria | 13,562 | 7,869 | 7,993 | 7,711 | 8,058 | 23 | 24 | 24 | 24 | 24 | 1.72 | 1.70 | 1.76 | 1.68 |
| Mexico | 13,165 | 9,712 | 9,784 | 9,550 | 9,683 | 24 | 23 | 23 | 23 | 23 | 1.36 | 1.35 | 1.38 | 1.36 |
| Czech | 12,570 | 7,699 | 7,723 | 7,661 | 7,846 | 25 | 25 | 25 | 25 | 25 | 1.63 | 1.63 | 1.64 | 1.60 |
| Denmark | 12,560 | 6,986 | 7,023 | 7,144 | 7,032 | 26 | 27 | 26 | 26 | 27 | 1.80 | 1.79 | 1.76 | 1.79 |
| Finland | 10,565 | 6,504 | 6,541 | 6,313 | 6,446 | 27 | 30 | 30 | 30 | 30 | 1.62 | 1.62 | 1.67 | 1.64 |
| Greece | 10,415 | 6,611 | 6,713 | 6,447 | 6,568 | 28 | 29 | 29 | 29 | 29 | 1.58 | 1.55 | 1.62 | 1.59 |
| Argentina | 9,744 | 7,046 | 6,976 | 7,069 | 7,407 | 29 | 26 | 27 | 27 | 26 | 1.38 | 1.40 | 1.38 | 1.32 |
| Hungary | 8,940 | 4,873 | 4,849 | 4,965 | 5,129 | 30 | 32 | 32 | 32 | 32 | 1.83 | 1.84 | 1.80 | 1.74 |
| Singapore | 8,390 | 6,811 | 6,873 | 6,600 | 6,653 | 31 | 28 | 28 | 28 | 28 | 1.23 | 1.22 | 1.27 | 1.26 |

Note. * Red color indicates rank drop from procedure A to other procedures; blue color indicates rank rise. ** Background colors indicate the range of counting inflation ratio (1.00-1.25; 1.26-1.50; 1.51-1.75; 1.76 and above).

**Table 1. Paper counts, country ranks, and counting inflation**