

Summary

Academic establishments and funding agencies around the world are increasingly interested in assessing the quality of academic output. Most judgments about research are based on perceived quality of the publications. This study examined the factors associated with scholarly impact from a select macro and micro perspectives. The analyses was intended to provide evidence base and policy lead.

Objectives

The objectives of the study were to -

1. Examine the S&T publication patterns and their subject-wise distribution for major S&T publishing countries.
2. Analyze the overall citation patterns and their subject-wise trends.
3. Capture and analyze the comparative data on bibliometric and non-bibliometric variables for the identified countries.
4. Identify cases of publication impact and the factors associated with them from different countries.
5. Develop evidence based macro understanding for impact and what can be the learning from the international experience for Indian science.

Data and Methods

The data for the analysis were sourced from Scopus citation database and ScimagoJr.com.

The entire set of journal output and the associated citation related data were analysed to learn the functional relationship between citation variables and a set of other bibliographic variables.

Country-wise citation related analyses were narrowed down to 32 countries making up the top 90% of the scholarly literature output. This was done to make the analyses viable without leaving out any major contributor to research output. India is part of these countries.

Two micro analyses were carried out - (a) to understand the citations in the context of international collaborations using 2018 data pertaining to Immunology for select set of countries, including India; (b) to evaluate the validity of tail-end citations. This analysis used 2014-2015 data relating to Economics pertaining to India and the UK. The study also enquired into publisher interest in citations.

Extent of scholarly literature

Scholarly literature as indexed by Scopus for the period 1996-2018 includes output from 239 countries. Despite the vast representation, publications of 32 countries make up the top 90% of the output in Scopus. China and the US figure as the top two countries on CitableDocs. The contributions of Australia, Belgium, Canada, Denmark, Finland, Israel, Netherlands, Norway, Singapore, Sweden, and Switzerland have higher than average citedness per document.

Analysis of citations

The variables identified for this analyses were the extent of Citable Docs; TotalCites; TotalRefs; RefPerDoc; International Collaboration; Scimago Journal Ranking (SJR); CitesPerDoc - both Journal-wise and Country-wise. Both CitableDocs and CitesPerDocs were restricted to three-year period 2016-2018.

Total Citations

CitableDocs uniquely explained 78.1% of the observed variance in the Regression model (as derived from the part correlation), followed by SJR, which accounted for 20.1%. Articles with international

collaboration on its own accounted for 0.8% of the total, and RefPerDoc -0.7% of the accounted variance.

The same set of four predictor variables were regressed against TotalCites for articles in each of the 27 subject categories identified by Scopus. The results indicate that the four chosen independent variables explain a high degree of variance ranging from 97.7% ($R^2 = .977$) for Agricultural and Biological Sciences to 27.8% ($R^2 = .278$) for Economics. SJR makes a substantial and statistically significant difference in all the subjects in accruing total citations. The results on the whole reveal that citations accrued to articles in journals in different subjects depend mostly on the number of CitableDocs published by the journals indexed in Scopus. Higher the number, greater the citations accrued.

Actual and estimated citations to Indian journals

Total citations accrued to 499 Indian journals during 2016-2018 period was 93,380. As per the estimates based on regression equation derived in the multivariate model this should have been 154,803 citations. There was a shortfall of 61,423 going by the larger trends in Scopus indexed journals.

CitesPerDoc

It is not the total citations, but the higher citation impact of what is published, considering countries and their respective scholarly academic base comes in varying size.

The results of this analysis indicate that high impact for CitableDocs is a function of the Publication appearing in journals with high SJR more than the other variables. The Regression analyses for CitesPerDoc for 27 subjects show statistically significant results. Total Variance accounted for vary from a low of 7.6% (for Biochemistry, Genetics & Molecular biology) to a high of 96.8% (Energy).

Actual and estimated CitesPerDoc for Indian journals

Regression equation derived in the multivariate model indicate that the mean CitesPerDoc for Indian journals was 0.650 against the estimated value of 0.720. Scholarly contributions in our journals are less frequently cited that the estimated figures based on the world trends.

Analysis of citations at country level

Multivariate linear regression with the predictor variables - Researchers in R&D per capita; Total expenditure in R&D per capita (\$); University Education Index, GDP (PPP) per capita, Citable Documents with Total Citations as criterion variable returned R^2 of .991. Beta values significant in the context were University Education Index ($\beta .231^{**}$), GDP (PPP) (per capita) ($\beta .052^*$), and CitableDocs ($\beta .820^{**}$)

The results indicate that

- If we want higher citation figures against the country we can rely on publishing more citable documents, focus on university education standards, and economic development as reflected in GDP(PPP).
- However, this does not result in higher CitesPerDoc. It is the higher R&D Expenditure that matters more along with better economic development as reflected in GDP (PPP).
- Higher number of CitableDocs does not result in higher CitesPerDoc.

Estimated and actual citations to Indian contributions

Indian scholarly contributions have appeared both in Indian and foreign journals. During 2016-2018 our contributions had accrued 1,939,535 total citations against the estimate of 2,471,399 based on the Regression equation. We accrued 531,864 citations less. Our Cites PerDoc was only 4.33 for the period

as against the estimate of 4.68 based on Regression equation for CitesPerDoc, a shortfall of 0.35 per citable documents.

Are journals of some countries better than the others in citation yield?

This analysis explored possible mean difference in Total Cites among the journals of the countries making up the top 90% of the total citable documents in Scopus, using Anova statistic.

Total Citations

The main Anova inclusive of all the journals showed significant F Ratios indicating that there exists a statistically significant difference in Total Cites yield among country-wise grouping of journals. The analysis also indicate that in 13 of the 27 country-wise subject grouping of the journals a significant statistical difference is noticeable.

Indian journals accrue significantly lesser mean citations compared to those of Netherlands, Switzerland, the US, the UK as a whole, and so also in specific subjects, namely Engineering, Environmental Science, Medicine, Pharmacology, Toxicology, and Pharmaceutical, Social Sciences, and Veterinary Sciences.

Citesperdoc

Similar analysis carried out on variable CitesPerDoc for country-wise journal groupings and also their subject-wise groupings show a more complex pattern. Pecking order of journals for higher CitesPerDoc is those published in Netherlands, the US, the UK, Switzerland, Germany, and the others

Analyses of data for countries with high citation impact

This analysis explored the international collaboration, collaboration advantage, and desirability of publication in journals of higher SJR with publication and citation data pertaining to immunology. Four countries - Denmark, Netherlands, Sweden, and Switzerland - which have shown high citedness were considered for the analysis. Incidentally, these countries are also high on international collaboration in Citable Docs. Data pertaining to India was used for comparison.

Total Indian publications on immunology is more than twice that of Netherlands and Switzerland, and thrice as much as Sweden and almost four folds that of Denmark. However, the citation intensity for Denmark, Switzerland, Sweden, and Netherlands is twice or more than that of Indian publications. Despite the variation in number of total publications, extent of total authorial involvement remains more or less the same for India and Netherlands because of international collaboration. The number is not far behind for Switzerland. Mean number of authors for Indian contributions was 4.73 as against 10.72 for those of Denmark, and 9 and a fraction above for the other three.

Approximately 20% of Indian contributions had international collaboration compared to 77% for Denmark; 77% for Switzerland; 71% for Sweden and 75% for Netherlands.

In our international collaborative research projects we have been in lead in more than half the cases. The tabulated citation data shows that when Indian researchers were in the lead, almost for 25% of those collaborative publications the citation yield was 0 and so it goes. The other four countries had substantially greater proportion of CitableDocs with 10 or more citations. The distribution points to factors other than publication quality in play in citation yield.

The analysis was taken one step further to understand whether the publications of the five countries yield overlapping citations when they are published in journals of similar SJR category. Indian publications get significantly less citation yield compared to the other four countries individually. Indian research publications, despite being in the 'same company' seem to get significantly different citation impact. It is not where you publish, who you are seem to matter for citation yield.

Journal clustering on Publishers and citation distribution in Scopus

Top three publishers - Elsevier, Taylor & Francis, and Springer - own 4,769 (19.37%) of the total journals in Scopus. Each of these business houses published over 1,000 plus indexed journals. The next three publishers, in descending order of titles indexed - Sage, Wiley, and Blackwell - owned Journal titles ranging from 500 to 999 making up 9.14% of those indexed in Scopus. These six top publishers cumulatively publish **36.04%** of the total citable docs, and these have accounted for 42.29% of the citations accrued during 2016-2018 period. Elsevier alone made up 8.59% of the total journals, 17.86% of citable publications, and 24.92% of the citations.

At the other end of the journal distribution were those publishing one journal. There are 5877 of those making up 23.88% of the total, contributing 15.12% of the articles to Scopus, and got a citation yield of 5.20% of the total.

The distribution is explicit in being skewed towards a few top publishers both in terms of journal ownership, total CitableDocs, and TotalCites.

CitesPerDoc depends most on SJR. SJR is prestige of the journals. This prestige is a construct of the scholars themselves. The chances of an average journal (included in Scopus) from a developing country being in the top of this 'prestige' heap is not feasible and so higher CitesPerDoc from the CitableDocs in the journals from developing country like India is very less probable. Among the 4,533 Scopus indexed journals which have SJR one or higher, Brazil has 4, China 25, India 3, Mexico 1, and Turkey 0. These are the countries which are in the middle and lower income bracket. To contrast this are the journals from the US in this category are (1773) 28.55%; the UK (1502) 27.39%; Switzerland (120) 23.30%; Netherlands (719) 34.72%; Germany (256) 15.88%; Denmark (4) 11.11%. We have to understand that these journals are the product of the local research culture, and this culture sets the standards for science.

The analysis of data on immunology in the study also holds this out clearly. The chances of contributions with international collaboration getting cited is less for Indian publications falling in this category. The analysis also shows that even when Indian contributions appear in the journals of broadly overlapping SJR categories the citation accrual is significantly less. This trend supports the social constructivist argument of citations, as against the normative theory.

If Indian scientific contributions need such a recognition we need to work hard in multiple fronts of economy, as CitesPerDoc is a function of higher GDP(PPP) per capita, as also greater expenditure on R&D. It is only then we can enter the elite club. Without that even publications in higher SJR would run short on this.

The relevant question in the context is whether we can make a good scholarly contribution in the context of low investment or from the less developed countries? Though it is immensely possible, chances of them getting cited is low as the analysis from the Scopus data suggest.

Given the lukewarm acceptance of Indian scholarly publications as demonstrated in the analyses, the appropriate question could be how does it serve our science and what the society looks from such pursuits. The analysis shows that there is a pecking order among the journals and author affiliations. The source and also the contributions of some countries are preferred in citation terms than the others. This suggests the play out of social constructivist view of knowledge growth and citation practice. In that scheme of things both Indian citable documents and Indian journals do not figure prominently. It is so across the subjects. In such a context perhaps promotion of wider local science base and generating locally relevant knowledge needs emphasis, apart from engaging with

contemporary science and technology in general. Nonetheless, this has to be approached strategically with a longer term perspective. As we are so engaged we have to acquire the best practices in scholarly journal management, among others.

We may take a careful look at the evaluation practices like Sistema Nacional de Investigadores (National System of Researchers) adopted by Mexico as the country's main instrument for stimulating competitive research in science and technology. SNI is a cornerstone of the higher education system in Mexico, and is authorized to rank both research and researchers.

000