

# Scale-Independent Bibliometric Indicators

J. Sylvan Katz

*Science Policy Research Unit  
University of Sussex, Brighton, England*

Van Raan (this issue) makes an excellent case for using bibliometric data to measure some central aspects of scientific research and to construct indicators of groups: research groups, university departments, and institutes. He claims that, next to peer review, these indicators are indispensable for evaluating research and can be used in parallel with peer review processes. By way of an example, van Raan provides a table containing nine indicators for a German medical research institute. Two of these indicators—articles ( $P$ ) and citations ( $C$ )—are established proxy measures for the size of a group and the impact of its published research (Katz & Hicks, 1997). The ratio between citations and articles ( $CPP$ ) and the ratios between  $CPP$  and the mean Journal Citation Score and between the field-based world average and the Germany-specific world average—which are uniquely defined  $CPP$  reference values—are used to construct a set of indicators that van Raan suggests can be used to assess international research performance.

This commentary focuses solely on the use of bibliometric indicators to compare international research performance. It addresses the fundamental question of whether  $CPP$  or measures like  $CPP$  can be used to accurately compare the performance of groups of different sizes.

## SCALING RELATIONS

A scaling relation exists between two entities,  $x$  and  $y$ , if they are correlated by a power law given by the equation  $y = kx^n$ , where  $n$  is the scaling factor and  $k$  is a constant. There is evidence to suggest that  $C$  and  $P$  have a scaling relation when

---

Requests for reprints should be sent to J. Sylvan Katz, SPRU, the Freeman Centre, University of Sussex, Brighton, East Sussex, BN1 9QE. United Kingdom. E-mail: j.s.katz@sussex.ac.uk

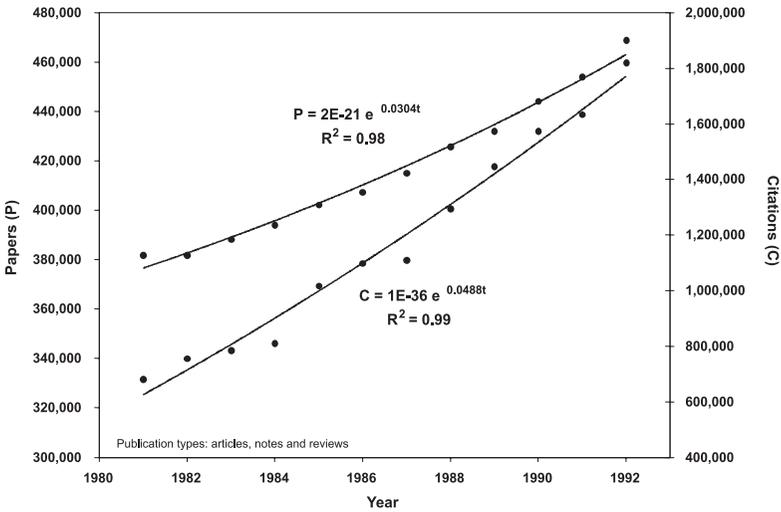


FIGURE 1 Exponential growth of citations and articles in the Science Citation Index.

measured (a) over time and (b) at a point in time. Each point will be considered separately.

It is known that the scientific literature tends to grow exponentially (Price, 1963; Wyatt, 1991). For example, Figure 1 uses data provided by the Institute for Scientific Information (ISI) from the 1981–1994 Science Citation Index (SCI).<sup>1</sup> It plots the growth from 1981 to 1992 of the number of articles indexed in the SCI—and the citations of these articles, counted using a three-year citation window. The data were not corrected for self-citations.<sup>2</sup> The graphs demonstrate that the data for  $P$  and  $C$  are accurately approximated by exponential growth curves.

It can be shown that any pair of coupled exponential processes will exhibit a power-law correlation where the scaling factor is given by the ratio of the exponents of the exponential processes (Katz, 2003). Assume we are given any two exponential processes  $x = am^{pt}$  and  $y = bm^{qt}$ . Let  $y = sx^n$ , then  $bm^{qt} = s(am^{pt})^n$  or  $b/s(a)^n = m^{(pn - q)t}$ . Because  $m^{(pn - q)t}$  is a time-dependent variable and it cannot be equal to  $b/s(a)^n$ , a constant, unless  $pn - q = 0$ , therefore,  $n = q/p$  and  $s = b/a^{q/p}$ . In other words, any pair of coupled exponential processes will exhibit a power-law relation with exponent  $n$  and intercept  $s$  that is predictable from the exponents and in-

<sup>1</sup>These data were purchased from ISI for a Natural Environment Research Council environmental research evaluation.

<sup>2</sup>It is unlikely that correcting for self-citations will affect the power-law relation between  $C$  and  $P$  shown in Figure 2; however, it might change the values of  $n$  and  $k$ .

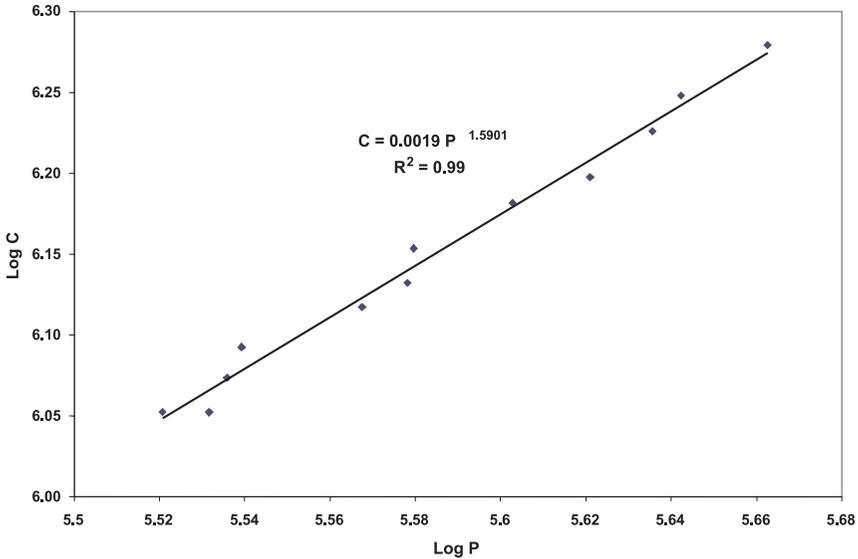


FIGURE 2 Scaling relation between citations and articles in the Science Citation Index.

tercepts of the individual exponential processes.<sup>3</sup> This relation holds even if the two processes are delayed in time with respect to each other.

The growth rates of  $P$  and  $C$  are exponential processes coupled in time. The preceding proof suggests that, using the regression information taken from Figure 1, we can predict that the scaling factor for the power-law relation between  $C$  and  $P$  should be  $0.0488/0.0304 = 1.61$ . As shown in Figure 2 the measured value was  $1.59 \pm 0.06$ , which is within 2% of the predicted values. The scaling factor tells us that, over the time interval when the number of  $P$  in the SCI doubled, the number of  $C$  counted using a 3-year citation<sup>4</sup> was expected to triple ( $2^{1.59} = 3.01$ ).

A variety of scaling relations exist between  $C$  and  $P$  at points in time (Katz, 1999, 2000). For example, when measured annually across ISI subject fields, the value of the scaling factor ranged from 1.24 to 1.28, with an average of  $1.27 \pm 0.03$ . The same analyses for a variety of national science systems and United Kingdom (UK) sectors found that  $n$  ranged from  $1.16 \pm 0.05$  to  $1.34 \pm 0.04$ . A scaling relation was also found across institutions within major UK sectors.  $C$

<sup>3</sup>The scaling factor,  $n$ , can also be used as a measure of the relative growth rate of  $x$  and  $y$ . When  $n = 1$ , the relative growth rates of  $x$  and  $y$  are the same; when  $n > 1$ , then  $y$  is growing faster than  $x$ ; and when  $n < 1$ , then  $x$  is growing faster than  $y$ .

<sup>4</sup>Counting citations using other citation window sizes may affect the sign and magnitude of the scaling factor.

and  $P$  probably scale across collections of research groups or university departments as well as institutes.

## SCALE-INDEPENDENT INDICATORS

A scaling relation can be used to construct a class of indicators, called scale-independent indicators (Katz, 2000), that are normalized for the effects of scaling. For example, assume that a collection of groups exhibits a scaling relation between  $C$  and  $P$  for a given characteristic at some point in time. The power-law regression between  $C$  and  $P$  describes the statistically average scaling relation among the members of the collection. The number of citations,  $C_p$ , that a group is expected to receive, given the nature of the collection, is given by  $C_p = kP^n$ , where  $k$  and  $n$  are given by the power-law regression and  $P$  is the size of the group. A relative citation index (RCI) can be calculated that compares a group's predicted citation count to its actual citation count, and is given by  $RCI = C/C_p$  (Katz, 2000). RCI can be directly compared across a collection of groups of differing sizes, whereas  $CPP$  should probably be compared only across groups of similar size or a collection of groups for which  $n = 1$ . When  $n = 1$ , the rank order of the members in the collection when ranked by  $CPP$  or RCI is the same. It can differ remarkably when  $n \neq 1$ .

One might argue that, when the scaling factor is close to 1.0, its effects can simply be ignored. Let us assume that  $n = 1.05$  and that we wish to compare two groups where one group published an order of magnitude more articles than another group. Given the scaling factor, we would expect the larger publisher to receive 11.2 times ( $10^{1.05}$ ) more citations than the smaller one. In other words, the larger group would be expected to receive 12% more citations than if we were to assume  $n = 1.0$ . A small scaling factor can have a large effect.

## CONCLUSION

In summary, scaling relations between  $C$  and  $P$  can exist over time and at points in time across research fields and institutes. They probably exist across research groups and departments, too. Conventional  $CPP$  measures can be supplemented with scale-independent measures that are corrected for the scaling correlation between  $C$  and  $P$ . Scale-independent indicators might increase our confidence in the indicators that we are using to evaluate research performance.

## REFERENCES

- Katz, J. S. (1999). The self-similar science system. *Research Policy*, 28, 501–517.  
 Katz, J. S. (2000). Scale independent indicators and research assessment. *Science and Public Policy*, 27, 23–36.

Katz, J. S. (2003). *The European Innovation System: A scale independent model*. Available at <http://www.sussex.ac.uk/Users/sylvank/plm>

Katz, J. S., & Hicks, D. M. (1997). *Bibliometric indicators for national systems of innovation*. Retrieved December 20, 2004, from <http://www.sussex.ac.uk/users/sylvank/pubs/Idea12.pdf>

Price, D. J. de S. (1963). *Little science, big science*. New York: Columbia University Press.

Wyatt, J. (1991). Use and sources of medical knowledge. *Lancet*, 338, 1368–1373.