Ruiz-Castillo, J. (2016), Research output indicators are not productivity indicators. Journal of Informetrics, v. 10, n. 2, pp. 661-663
Available in: https://doi.org/10.1016/j.joi.2016.04.004

Research output indicators are not productivity indicators

1. Introduction

I welcome the opportunity of offering a few comments on the Abramo and D'Angelo (2016) contribution—AA hereafter. To simplify matters, I will restrict myself to the issues in a single scientific field. I will also assume that all publications have a single author, overlooking the problem of assigning responsibility in the case of co-authored publications. I will discuss production functions, scientific productivity indices, and research output indicators

2. Production functions

Everybody would agree with AA's position that the natural setting for analyzing research performance is the notion of a production function where new knowledge output is a function of labor, capital, and possibly other inputs (Abramo & D'angelo, 2014). The problem, however, is twofold. First, even if we agree for the purpose of this note that publications and citations provide adequate measures of research output, measuring labor, capital, and other inputs is not easy. Second, estimation of production functions is a difficult econometric problem (Aguirregabiria, 2009; Ackerberg, Caves, & Frazer, 2015). Notwithstanding, conclusions reached in specific contexts may be useful elsewhere. For example, consider the research gap between Italy and an advanced scientific country, say the U.S. Casual observation indicates that the inferiority of Italian research can be partly explained by lower resources and the dominance of endogamic versus meritocratic criteria for the hiring and promotion of researchers. Any appropriate productivity comparison between these two countries would offer important lessons for Spain and other Mediterranean countries with similarly limited resources and poor governance as does Italy.

3. Productivity indices

Next, consider simply the possibility of collecting data for publications, citations, *and* labor inputs for some type of research units worldwide. As AA recognize "*We are aware that many countries do not have exhaustive databases of the composition of their university faculties, and that the disambiguation of author names on national scale remains a difficult task*." To this list we should add the disambiguation of research units' names.[1] Furthermore (i) the assignment of researchers to fields (ii) the estimation of equivalent full-time researchers for research unit members with and without teaching responsibilities, and (iii) the distinction between active researchers that publish at least one article in a given period of time and inactive researchers without any publications at all, remain formidable problems. Be that as it may, assume that we have information on the number of authors or active researchers N_i, the number of publications P_i, and the number of citations C_i for a set of I research units indexed by $i = 1,\ldots, I$. Assume that $P_i > P_j$ and/or $C_i > C_j$ for two research units. Can we conclude that unit i is superior to unit j in any interesting sense? Perhaps we can, for some purposes. However, for many other purposes it can be argued that using the total number of publications or the total number of citations as a performance indicator confounds *size* and *merit*. Without information on capital and other inputs, it seems acceptable to identify size with number of authors. As far as merit is concerned, I suggest distinguishing between three notions. Firstly, we could rank units in terms of P_i/N_i, $i = 1,\ldots, I$, a size-independent indicator of publication output. Secondly, we could rank research units in terms of C_i/N_i, $i = 1,\ldots, I$, a size-independent indicator of citation impact.[2] Third, for any unit of size N_k we could draw a large enough number R of random distributions of size N_k among all publications in the field, compute the total number of citations C^r for each $r = 1,\ldots,$

[1] Recall that only SCImago has a world institutions ranking (http://www.scimagoir.com). The Leiden Ranking produced by the CWTS only includes 750 universities (http://www.leidenranking.com).

[2] Although one can define other interesting options, for simplicity I would only consider average-based size-independent indicators.

R, and define unit k's citation merit m_k in terms of the percentile reached by C_k in the distribution $\{C^r, r = 1,..., R\}$. Then we could rank units in terms of m_k, $k = 1,..., I$.[3]

I believe that AA would accept the first two rankings – and perhaps the third – as relevant. However, recall that we have purposely overlooked a crucial dimension of the problem, namely, the capital equipment. Suppose, for example, that for two research units we have $C_i/N_i = C_j/N_j$ according to the second notion of merit, but unit i has old and small research facilities while unit j has brand new and ample research facilities. In this case, we would agree with AA that unit i is more efficient than unit j in the citation impact research front. The problem is that we will have to conclude that both units perform equally well because we do not have data on research facilities.[4]

4. Research output indicators

Finally, consider the current typical situation in which we only have information on P_i and C_i, $i = 1,..., I$. Of course, we can always rank research units in terms of the size dependent indicators P_i and C_i. For some purposes, this exercise could be useful, although this is not always the case. For example, consider the notion of the "European Paradox" – popularized in the *First European Report on Science and Technology Indicators* (EC, 1994) – according to which Europe plays a leading world role in terms of scientific excellence, measured in terms of the number of publications P, but lacks the entrepreneurial capacity of the U.S. to transform it into innovation, growth, and jobs. Apparently, the problem lies – not in the EU's scientific performance – but elsewhere. In his influential contribution, King (2004) states that "*the EU now matches the United States in the physical sciences, engineering and mathematics, although still lags in the life sciences*". The trouble is that King's statement refers to the share of total citations C, which is a mere consequence of the European superiority in the volume of publications. Albarrán, Crespo, Ortuño, and Ruiz-Castillo (2010) compare the publication shares of the U.S. and the EU at every percentile of the world citation distribution in each of 22 broad fields. It is found that – except for Agricultural Sciences – the U.S. always surpasses the EU when it counts, namely, at the upper tail of citation distributions. In other contributions, using different types of *citation impact per publication indicators* – namely, the type of indicators AA critize – it is established that the European Paradox masks a truly "European Drama": judging from citation impact per publication, the dominance of the U.S. over the EU is almost universal at different aggregation levels (Albarrán, Ortuño, & Ruiz-Castillo, 2011a; Herranz & Ruiz-Castillo, 2013).

Of course, indicators of citation impact per publication are not size-independent or productivity indicators in the sense of point 2. They are merely research output indicators. Nevertheless, they have served the purpose of discrediting the so-called European Paradox, re-directing the concern about the U.S./Europe research gap towards differences in resources and governance at both sides of the Atlantic—a useful purpose.

As AA emphasize, using indicators of citation impact per publication has obvious problems. In their example, unit i has $P_i = 100$, and $C_i = 1{,}000$, while unit j has $P_j = 200$, and $C_j = 1{,}500$. Since both units are assumed to have the same number of authors, we can take $N_i = N_j = 10$. In terms of mean citation per publication, $C_i/P_i = 10$ and $C_j/P_j = 7.5$, so that unit i appears to perform better than unit j. However, in terms of the first two notions of merit in point 2 we have: $P_i/N_i = 10 < P_j/N_j = 20$, and $C_i/N_i = 100 < C_j/N_j = 150$, indicating that unit j is more productive than unit i. In this case, the ranking according to the total number of publications, $P_i = 100 < P_j = 200$, and the total number of citations, $C_i = 1{,}000 < C_j = 1{,}500$, provide the right answer. However, this does not establish the superiority of the latter over the mean citation per publication: it is easy to construct an example in which the total number of publications and the total number of citations provide the wrong answer, while the mean citation per publication provides the right answer. It is also easy to construct an example in which the three indicators provide the wrong answer.[5]

Of course, when we rank units according to indicators of citation impact per publication we approximate a unit's size in terms of N_i by the size of its citation distribution, namely, the number of publications P_i. This is the sense in which most of the literature, and hence the title of AA's contribution, uses the term "size-independent" indicator.[6] Ranking two units according to size-independent indicators in this sense provide the wrong answer when P_i/P_j is very different from N_i/N_j—an unavoidable problem when we overlook the number of authors. Conceptually, this is the same difficulty that we encountered in point 2 when we "forgot" to take into account differences in capital inputs.

Naturally, as the last two examples have shown, using size dependent citation impact indicators in this sense does not solve the problem at all. Consequently, in the absence of information on inputs, it is convenient to have research units' rankings according to size-dependent and size-independent indicators where size is equal to the number of publications.[7] Depending on the issue at hand, different users can refer to one or the other ranking type. In my own case, when choosing

[3] This is the notion of merit introduced in Crespo, Ortuño, and Ruiz-Castillo (2012) when unit size is identified with number of publications. Note that unit k will have the same merit in terms of the percentile reached by C_k/N_k in the distribution $\{C/N_k\ r = 1,..., R\}$.

[4] This is, of course, an elementary point (see *inter alia* Abramo and D'angelo, 2015).

[5] First example: unit i has $N_i = 18$, $P_i = 12$, and $Ci = 100$, while unit j has $N_j = 8$, $P_j = 8$, and $C_j = 80$. Second example: unit i has $N_i = 25$, $P_i = 9$, and $C_i = 100$, while unit j has $N_j = 8$, $P_j = 8$, and $C_j = 80$.

[6] See *inter alia* my own publications Albarrán, Ortuño, and Ruiz-Castillo (2011b), and Perianes-Rodríguez and Ruiz-Castillo (2015).

[7] This is exactly what is done, for example, in the Leiden Ranking of universities. The SCImago ranking only ranks institutions according to a size-independent indicator of excellence. However, the availability of information on the number of publications would allow the user to construct a size dependent ranking.

an indicator of research output I would tend to use an indicator of citation impact per publication rather than an indicator of total citation impact.

5. Conclusions

In my opinion, AA should be complemented for refreshing our memories about the following:

(i) The production function approach provides a rich set up for the analysis of scientific research as an economic activity.
(ii) Having information regarding the number of authors per research unit makes valuable productivity comparisons possible by means of size-independent indicators.
(iii) In the meanwhile, when a unit's size can only be approximated by its number of publications, it should be clear that size-dependent and size-independent indicators of citation impact are merely indicators of research output. Consequently, they should not be interpreted as productivity indicators.

This is an important but elementary contribution to what Kuhn (1970) describes as *normal science*. AA's grandiose claim about a shift to a new research paradigm while the rest of us hold onto the current research paradigm, defending vested interests, etc., might be interpreted as manifestations of a certain Latin rhetorical style. Finally, even in the absence of information on inputs many of us will keep using research output indicators.

Acknowledgements

Conversations with Joan Crespo and Ignacio Ortuño are gratefully acknowledged. All remaining shortcomings and opinions are the author's sole responsibility.

References

Abramo, G., & D'angelo, C. A. (2014). How do you define and measure research productivity? *Scientometrics*, *101*, 1129–1144.
Abramo, G., & D'Angelo, C. A. (2016). A farewell to the MNCS and like size-independent indicators. *Journal of Informetrics*, *10*(2), 646–651.
Abramo, G., & D'angelo, C. A. (2015). Ranking research institutions by the number of highly-cited articles per scientist. *Journal of Informetrics*, *9*, 915–923.
Ackerberg, D. A., Caves, K., & Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, *83*, 2411–2451.
Aguirregabiria, V. (2009). *Econometric issues and methods in the estimation of production functions MPRA Paper No. 15973*. http://mpra.ub.uni-muenchen.de/15973
Albarrán, P., Crespo, J., Ortuño, I., & Ruiz-Castillo, J. (2010). A comparison of the scientific performance of the U.S. and Europe at the turn of the 21st Century. *Scientometrics*, *85*, 329–344.
Albarrán, P., Ortuño, I., & Ruiz-Castillo, J. (2011a). High- and low-impact citation measures: empirical applications. *Journal of Informetrics*, *5*, 122–145.
Albarrán, P., Ortuño, I., & Ruiz-Castillo, J. (2011b). High- and low-impact citation measures: technical results. *Journal of Informetrics*, *5*, 48–63.
Crespo, J. A., Ortuño, I., & Ruiz-Castillo, J. (2012). The citation merit of scientific publications. *PLoS One*, *7*(2012), e49156.
EC (1994). *First European Report on Science and Technology Indicators*, Directorate-General XII, Science, Research, and Development. Luxembourg: Office for Official Publications of the European Community.
Herranz, N., & Ruiz-Castillo, J. (2013). The end of the 'European Paradox'. *Scientometrics*, *95*, 453–464.
King, D. (2004). The scientific impact of nations. *Nature*, *430*, 311–316.
Kuhn, T. (1970). *The structure of scientific revolutions* (2nd. ed). Chicago: University of Chicago Press.
Perianes-Rodríguez, A., & Ruiz-Castillo, J. A. (2015). Comparison of two ways of evaluating research units working in different scientific fields. *Scientometrics*, *106*, 539–561.

Javier Ruiz-Castillo
Departamento de Economía, Universidad Carlos III, Spain
E-mail address: jrc@eco.uc3m.es