

An Examination of the Reliability of Prestigious Scholarly Journals: Evidence and Implications for Decision-Makers

By ANDREW J. OSWALD

University of Warwick

Final version received 17 September 2006.

Scientific-funding bodies are increasingly under pressure to use journal rankings to measure research quality. Hiring and promotion committees routinely hear an equivalent argument: 'this is important work because it is to be published in prestigious journal X'. But how persuasive is such an argument? This paper examines data on citations to articles published 25 years ago. It finds that it is better to write the best article published in an issue of a medium-quality journal such as the *OBES* than all four of the worst four articles published in an issue of an elite journal like the *AER*. Decision-makers need to understand this.

The results . . . will be expressed as quality profiles of research in each department submitted to the 2008 Research Assessment Exercise. They will determine the annual distribution of more than £8 billion for research in UK higher education institutions over a six-year period (www.rae.ac.uk).

The Government's firm presumption is that after . . . 2008 the system for assessing research quality . . . will be mainly metrics-based (www.hm-treasury.gov.uk/media/20E/EA/bud06_ch3_192).

INTRODUCTION

The United Kingdom is currently a useful test-bed for a worldwide problem: how to allocate resources to scientific research. Its forthcoming Research Assessment Exercise (RAE) will determine how much money should go to each department in more than 100 UK universities. To do this, a panel of experts will assess the quality of every department in every university. A small selection of each department's scholarly articles and books is to be given a quality rating by the appropriate panel. These will run from 4* down to 1*, where 4* corresponds to the highest world-class standard, and 1* corresponds to a national standard of research excellence. On such assessments will turn £8 billion of taxpayers' money. It appears that nations like Italy and Australia will soon follow the UK example and introduce a form of state-run university assessment exercise.

Partly because of the size of the undertaking, there will be pressure, if only covertly, on members of these peer review panels to use *journal labels* (e.g. X is a 4* journal, Y a 2* journal) in a mechanical way to decide on the quality of articles. Rumours of this, and guesstimates of the key list of journals, are currently circulating. Similar forces are discernible in other places. Seglen (1997), for instance, notes the rising use of journal prestige ratings as part of funding decisions in medical research. In the world of economics research, a

distinguished research institute in Amsterdam publishes a list of starred journals, ranked into categories of quality, to emphasize to its researchers that papers in certain journals should be viewed as of quality 'A' while others should be regarded as of quality 'B'.

It might seem natural that, despite UK government-backed formal disclaimers, the RAE expert panels should behave in this way. An obvious argument might go as follows. These papers have already been anonymously refereed, so the quality of a journal paper will be accurately captured by the prestige of the journal in which it has been published. Thanks to sources such as the ISI Web of Science database, journal standing can be judged fairly objectively by, for example, 'impact factors'. Journal labels and 'impact factors' are thereby becoming a kind of measuring rod, taking on a life of their own. This is not because of governments *per se*. Similar forces can be seen elsewhere. In universities around the globe, including the leading private US universities, hiring committees regularly listen to the argument that goes: 'this is important work because it is about to appear in prestigious journal X'.

But how persuasive are such arguments? Little research seems to have been directed at that question. As in most areas of life, prestige ratings in academia have their uses, and it is unlikely that any scholar would argue that labels are meaningless. Yet this does not mean that journal names are genuinely a sufficient statistic for quality.

This paper is an attempt to explore the reliability of prestige labels. It might be viewed as closely related to papers such as Laband and Tollison (2003) and newspaper articles such as Monastersky (2005), which emphasize that *where* a modern scientist publishes appears to be in some danger of becoming more important than what the author is actually saying. It is also potentially complementary to work such as Laband (1990), Oswald (1991), Laband and Piette (1994), Johnson (1997), Kalaitzidakis *et al.* (1999), Frey (2003), Seglen (1997), Coupe (2003) and Starbuck (2003, 2005), and is a small contribution to the field of scientometrics (van Dalen and Henkens 2005; Sussmuth *et al.* 2006). There is also a natural link to the work of information science researchers such as Oppenheim (1995), who have shown that in the United Kingdom the departmental rankings that emerge from the Research Assessment Exercise are closely correlated to those that would have emerged from a citations-based departmental ranking.

The paper collects data on the accumulated lifetime citations to papers published 25 years ago. It uses these to construct a simple test. The data come from issues of six economics journals of varying levels of reputation. These data show the expected ranking. However, and more interestingly, they also reveal that the best article in an issue of a good to medium-quality journal routinely goes on to have much more citations impact than a 'poor' article published in an issue of a more prestigious journal. This fact may not be known to all of the people who sit on funding councils, or perhaps even to many economists.

I. DATA COLLECTION AND ANALYSIS

Assume that after some decades the quality of a journal article is approximately known. Perhaps the most usual measure of impact is that

captured by the total number of citations the article has received (that is, the number of times the article has been cited in later researchers' bibliographies).

There is a considerable line of work that uses citations to assess intellectual output and productivity, and it has long been known that professorial salaries are correlated with researchers' lifetime citations, and that these citation counts are a good predictor of Nobel and other prizes. Moreover, better universities are led by more highly cited individuals; see e.g. Hamermesh *et al.* (1982), Laband (1990), Garfield and Welljams-Dorof (1992), Toutkoushian (1994), Moore *et al.* (1998), van Raan (1998), Thursby (2000), Bayers (2005) and Goodall (2006). As is also well known, citations are a noisy signal of quality—survey articles tend to garner citations more readily than regular papers; there may be some pro-US bias in citations; citation numbers are more open to manipulation than are publications figures; for some individuals self-citations can cause problems; and so on—but a common view is that citations are the most persuasive single measure of scholarly productivity. If the impact factors of journals become distorted over time, of course, as may happen if citations attract greater publicity and editors opportunistically try to manipulate their journals' citations totals, then the signal-to-noise ratio of citations may decline in the future.

For this paper, a selection of economics journals was made from the year 1981 (i.e. a quarter of a century ago, to allow a long lag for the 'true' quality of a journal paper to be revealed). The winter issue of the year was examined for the *American Economic Review*, *Econometrica*, the *Journal of Public Economics*, the *Economic Journal*, the *Journal of Industrial Economics* and the *Oxford Bulletin of Economics and Statistics*.

The *American Economic Review* (*AER*) and *Econometrica* are routinely viewed as two of the most prestigious journals in economics; in rankings they often appear near or at nos. 1 and 2 out of approximately 200 economics journals. The *Journal of Public Economics* (*JPubEcon*) and the *Economic Journal* (*EJ*) are usually viewed as good journals—routinely in the world's top 20. The *Journal of Industrial Economics* (*JIE*) and the *Oxford Bulletin of Economics and Statistics* (*OBES*) are typically rated still lower in prestige: they often appear around nos. 40–50 in journal rankings, and sometimes are very far below the position of the top journals in informally distributed 'rankings' for the UK's RAE.

At the time of writing, for example, the Web of Science total-citations rankings in the Economics category put the *AER* and *Econometrica* at nos. 1 and 2, the *EJ* at no. 9, *JPubEcon* at no. 16, *JIE* at no. 47 and the *OBES* at no. 51.

Data on total lifetime citations were collected on each article. The raw data are summarized in the Appendix. Table 1 lays out a summary of the data. As is known, the skewness of citation numbers implies that the mean values lie far above the median values. A small group of papers accounts for the majority of citations.

The remarkable variation in the number of times these journals' approximately ninety articles have been cited by other researchers is clear from the raw data. The single most cited paper is the famous theoretical analysis of trade unions by Ian McDonald and Robert Solow: published in the *American Economic Review*, this paper has garnered 401 citations to date. The

TABLE 1
DATA ON THE ACCUMULATED LIFETIME CITATIONS TO ARTICLES PUBLISHED IN 1981 IN SIX ECONOMIC JOURNALS

| | American Economic Review | Econometrica | Journal of Public Economics | Economic Journal | Journal of Industrial Economics | Oxford Bulletin of Economics and Statistics |
|--|--------------------------------|--------------|-----------------------------------|---------------------|---------------------------------------|---|
| Mean citations per article in that issue | 68 | 63 | 22 | 30 | 9 | 7 |
| Median citations per article in that issue | 23 | 22 | 9 | 11 | 3 | 2 |
| Combined citations to the 4 least-cited articles in that issue | 6 | 5 | 23 | 3 | 6 | 0 |
| Citations to the single most-cited article in that issue | 401 | 355 | 88 | 199 | 43 | 50 |

Notes: These are taken, for each journal, from the winter issue of the year 1981. The data measure the number of times over the ensuing 25 years that the articles have been cited by others. The source is the Web of Science's Social Sciences Citations Index in late March 2006. The data include short papers, partly because some of them are highly cited and partly because it was not possible to draw a dividing line between those and full papers, but exclude articles denoted 'Notes' and 'Comments', and also book reviews (where it was possible to assign these categories unambiguously).

next most influential article is the Hausman–Taylor econometric estimator paper published in *Econometrica*, which has been cited 355 times.

However, many of the papers attracted very few citations. For instance, over the 25 years, 15 have been cited either not at all or on only one occasion. Judged from the perspective of the time elapsed, it might be argued that these articles' contribution to intellectual output has been, and probably will continue to be, zero. In a sense (with the benefit of hindsight, needless to say), their publication might now be viewed as having been a mistake.

The mean lifetime citations across these six journals follow the broad pattern that might be expected. The prestige labels are, in a sense, valid: *AER* 68 citations; *Econometrica* 63 citations; *JPubEcon* 22; *EJ* 30; *JIE* 9; *OBES* 7. The top journals thus dominate. Similarly, median lifetime citations are: *AER* 23 citations; *Econometrica* 22; *JPubEcon* 9; *EJ* 11; *JIE* 3; *OBES* 2.

The variation in true quality, as measured by citations, is strikingly large. Because of this high variance, the less frequently cited articles in the top journals are easily bettered by good articles in less prestigious outlets. For instance, the fourth most cited article in the entire sample (199 citations) is that by Mansfield *et al.*, which appeared in the *Economic Journal*. As another example, in the *American Economic Review*, which is perhaps the most prestigious journal in the discipline, in its winter 1981 issue more than one-third of the articles had each, after 25 years, been cited fewer than 20 times. The very best papers in the other lower-quality journals had by then garnered far more citations in others' bibliographies—respectively, 88 (Sandmo in the *JPubEcon*), 199 (Mansfield *et al.* in the *EJ*), 43 (Teece in the *JIE*) and 50 (Sen in the *OBES*).

Consider, as a benchmark, the median number of citations. In the two top journals here it is approximately 22. A natural question is then: how many of the articles published in the other four journals turned out to exceed that level? These articles 'should', in principle, have appeared in the top journals. The answer is: approximately 16% of the articles. In the *JPubEcon* 1 out of 6 does; in the *EJ* 4 out of 15; in the *JIE* 2 articles out of 17; and in the *OBES* 1 out of 11.

One way to make this point more vividly is to take the mean value of citations among the four least-cited articles in each of the six journals. As shown in Table 1, those values are, respectively, 6, 5, 23, 3, 6 and 0 citation. Compared with the 'best' article published in the lesser journals, these are of the order of one-tenth as cited.

Ex post, therefore, labels cannot be relied upon to be free of significant error. It appears that the journal system often allocates high-quality papers to medium-quality journals, and vice versa. The data of this paper are consistent with the theoretical argument of Starbuck (2005), who points out, using simple statistical parameterizations, that an error-ridden system would generate empirical patterns of the sort documented here.

Although the implication of these data is that labels work too imperfectly to be taken as a sufficient statistic for the quality of an article, this does not automatically mean that peer reviewers can *ex ante* improve upon the journal labels. Perhaps the label is the best that can be done without waiting 25 years?

Nevertheless, simple evidence against such a view comes out of the raw data. There are signs that the journal editors did have an idea of which would

be the best papers in a particular issue of their journal. Judging by the way in which they arranged the order of publication, the editors turned out, *ex post*, to have what now looks like prior foresight. This can be seen informally by looking at the raw data. If we regress total citations, y , on publication order in the journal, x (that is, whether the paper appeared first, second, third, . . . , eighteenth), we get a more formal sense for the pattern. ('Notes' and 'Comments', it should perhaps be emphasized, were omitted from the data; the criterion was whether the papers had these words in their titles). Summarizing as regression lines:

$$\text{Econometrica citations} = 133.14 - 7.36 \text{ Order}$$

$$\text{AER citations} = 119.43 - 5.41 \text{ Order}$$

$$\text{EJ citations} = 66.68 - 4.57 \text{ Order}$$

$$\text{JPubEcon citations} = 58.93 - 10.60 \text{ Order}$$

$$\text{JIE citations} = 13.15 - 0.44 \text{ Order}$$

$$\text{OBES citations} = 19.42 - 2.05 \text{ Order}$$

Individually, the sample sizes here are too small to give well determined results (the six results vary in statistical significance from approximately the 5% significance level to approximately the 30% level); but, as can be checked by pooling the data with journal-dummy intercepts, as a group they are more persuasive. Hudson (2006), which was not available at the time the first draft of this paper was written, finds equivalent results on the statistically significant role of the order of journal papers within an econometric equation explaining citations. What editors know, and exactly how they know it, seems worth exploring in future research, because of the importance of peer review in the allocation of research funding in Western society. It is possible that this can be conveyed to the experts who sit on funding bodies.

It should be stressed that the point of this paper is not to argue that journal-quality rankings should, in all instances, be eschewed. When large samples of data are used, as in the assessment of a university department over a long period such as a decade or two, the error in judging articles in a mechanical way by simply assigning to each of them a journal-quality weight may turn out to be fairly small.

A referee has asked the related question: how then, in a practical way, should the RAE peer review actually operate? One approach might be to assume that the anticipated intellectual value, v , of an article can be approximated by

$$v = w(y)c + [1 - w(y)]e(j, r, i),$$

where w is a weight between zero and unity; y is the number of years since publication, $w(\cdot)$ is an increasing function; c is the flow of citations per unit of time since publication of the article; $e(\cdot, \cdot, \cdot)$ is a function that describes the *a priori* expected long-run number of citations to an article; j is the known modal or mean level of citations to all articles in the particular journal; r is the ordinal ranking of the article within the particular journal issue in which it appeared; and i is some quality factor, which might be termed academic 'instinct', assigned by an independent assessor or assessors, such as the considered view of a peer review panel. (In some disciplines it might not be appropriate to rely at all on the within-journal ordinal ranking factor, r .)

In the short-run, therefore, emphasis would be given to the modal or mean level of citations to the journal. The sufficient-statistic nature of a prestige label would (temporarily) be dominant in an assessment of the particular article's likely quality. However, after a number of years—in some scholarly disciplines only one or two years—considerable weight would be given to the actual number of citations acquired by the particular article. The quality of the actual article and the quality of the journal would gradually de-couple. In the long run virtually all the weight would be put upon the total citations acquired by an article, and the identity of the journal—the label itself—would cease to be important. This approach has a Bayesian spirit, of course. The exact way in which the weighting function $w(y)$ is decided would have to be determined by, among other things, the probability distribution of refereeing mistakes.

II. POTENTIAL OBJECTIONS

Several issues deserve consideration.

One objection is that the data-set used here is fairly small. However, examination of the Social Science Citations Index suggests that these characteristics are found repeatedly. The same kinds of patterns occur, for example, in the winter *American Economic Review* issues for the years 1982, 1983, 1984 and 1985. Looking at the 'worst' four articles in each issue, none of these articles has achieved ten citations after 25 years. Moreover, if we work through the 1980s issues of the *Oxford Bulletin of Economics and Statistics*, we find that the top article in each year attracts a considerable but variable number of citations: 23 citations in 1982 (Siddharthan–Lall), 57 in 1983 (Caves *et al.*); 24 in 1984 (Evans–Gulamani), 99 in 1985 (Nickell), 483 in 1986 (Granger), 54 in 1987 (Nickell), 79 in 1988 (Osborn *et al.*) and 48 in 1989 (Jackman *et al.*). While it might be useful for other reasons to extend the sample size, the paper's broad conclusions seem unlikely to change.

Second, could this phenomenon have disappeared over the last quarter-century, as journals may have improved their ability to sort papers into different quality classes? One way of checking this is to examine the winter issues of the *Oxford Bulletin of Economics and Statistics* and *American Economic Review* just five years ago. This test reproduces the principal conclusion. For example, the most cited paper in the December 2001 issue of *OBES* (by Lee and Strazicich) has to date acquired 12 citations; the four least cited papers in that issue of the *AER* sum to a total of 8 citations: Bianchi *et al.* 0, Deck 2, Bennett and LaManna 3, and Banerjee and Eckard 3.

A third objection might be that citations should be weighted by the importance of the journal *doing the citing*. Being mentioned in the *American Economic Review* is perhaps better, in some sense, than being cited in a less prestigious journal.

Opinions on this point differ. Some commentators argue that it is an undesirable form of double-counting. Another objection is that, because of their prominence, articles appearing in the most elite journals tend to be over-cited relative to their true worth. Perhaps the most sensible view is that it is only in the short run that a citation in a top journal matters more—because in the long run the issue is instead the stock of intellectual influence across the

discipline as measured by acquired citations in the year the article ceases to be mentioned. For the purposes of the present paper, however, the key point seems to be that the broad ideas are not going to be altered by weighting the citations totals. The reason is that the papers in *AER* and *Econometrica* that garner very few citations are not—it is straightforwardly checked—getting these citations disproportionately in the top journals.

Fourth, no adjustment has been made for the number of authors on each paper. (As Hudson 1996 shows, the proportion of single-authored economics papers steadily declines through time.) Nevertheless, as the focus here is on the impact of individual articles rather than the productivity of particular researchers or collaborations (as in Kalaitzidakis *et al.* 1999; Laband and Tollison 2000), it seems reasonable not to weight according to author numbers.

Fifth, it could be argued that self-citations are best removed from the data sample, as is often done in rankings of top economists and university departments. On balance, however, it seems appropriate not to do this here. It does not alter the conclusions of the paper (because self-citations are insignificant for important articles' total citations); and, for some of these highly influential researchers, there seems a logical case for leaving in own-mentions to those authors' important earlier papers.

III. CONCLUSIONS

This paper is a simple one. It provides evidence that it is dangerous to believe that a publication in famous journal X is more important than one published in medium-quality journal Y. This does not mean that young scholars ought to ignore top journals, nor that research funders should. Nevertheless, the publication system routinely pushes high-quality papers into medium-quality journals, and vice versa. Unless funding bodies are aware of this fact, they may make bad choices about how to allocate resources. It is likely that some senior scholars already understand the general point made in this paper, but young researchers and funding agencies may not.

By definition, scholarly articles in better journals go on, on average, to be more highly cited. This is not at issue. Importantly for decision-makers, however, there is a highly imperfect match between the quality of the journal and the lifetime citations of the individual articles. Approximately 16% of articles in the four less highly regarded journals studied here ended the period with more citations than the median citations of an article in one of the two elite journals, the *AER* and *Econometrica*. To make the point in a different way, if the criterion is intellectual impact measured by citations, in this sample it was better to publish the top article in an issue of the *Oxford Bulletin of Economics and Statistics* than to publish all of the bottom-four papers in an issue of the *American Economic Review*. If peer reviewers, and those who sit on funding panels, have expert insight allowing them to judge quality, then the results in this paper suggest that there is a case for them to do so. They should not rely simply on mechanical rules based on journal labels.

It might be objected that such reviewers may have no extra information that would allow them to rank journal papers (beyond the prestige of the journal itself). This possibility deserves to be taken seriously and needs further

study. Nevertheless, one counter-argument is to look at the citation levels of the journal papers by the order in which the paper appeared in the journal issue. The early-position papers, such as the Cooley–Leroy and Rosen papers in the 1981 *AER*, are more highly cited than articles lower down the order of appearance. This suggests that editors had some ability to forecast which would turn out, 25 years later, to be the best papers. The individuals on peer-review panels may be able to do the same.

Finally, although it must be left to future work, there are interesting dynamic effects to be considered. The fundamental difficulty here is that there can be a large discrepancy between the ‘market’ forecast of the stream of citations from an article and the actual time-pattern of realized citations. As this paper shows, a journal-prestige metric is noisy. Therefore funding agencies and university employers in promotion cases might, in principle, be able to improve the efficiency of the reward structure by developing some sort of *ex post* settling-up mechanism. Such a mechanism would reward *ex post* accumulated citations on a paper rather than merely the *ex ante* mean citation rate of the publishing journal.

Whether we will see these retrospective reward structures in scientific research in future is an open, but interesting, question.

APPENDIX: RAW DATA ON THE TOTAL CITATIONS TO EACH 1981 ARTICLE (IN THE ORDER THEY APPEARED IN THE JOURNAL ISSUE)

American Economic Review

Cooley–Leroy 118
 Rosen 123
 Kohn 23
 Howe–Roemer 8
 McDonald–Solow 401
 Hendershott 16
 Spulber 19
 Bresnahan 156
 Azariadis 16
 Jonung 23
 Startz 3
 Darity 3
 Caves *et al.* 147
 Akerlof–Main 45
 Walker 0
 Mussa 70
 Conybeare 0
 Boland 53

Econometrica

Malinvaud 28
 Hausman–Taylor 355
 Mundlak–Yahav 1
 Nickell 258

Economic Journal

Harris–Purvis 12
 Malcomson 44
 Bingswanger 77
 Dervis *et al.* 7
 Mansfield *et al.* 199
 Hughes–McCormick 54
 Spencer 4
 Von Ungernsternberg 15
 Skott 0
 Chiplin 6
 Hughes *et al.* 0
 Shah–Desai 11
 Masuda–Newman 3
 Formby *et al.* 20
 Shea 0

Journal of Industrial Economics

Williams–Laumas 13
 Lynn 2
 Aaranovitch–Sawyer 3
 Levine–Aaronovitch 7
 Teece 43
 Thompson 21
 Dnes 2

| | |
|------------------------------------|--|
| Geweke 40 | Feinberg 2 |
| Godfrey 21 | White 3 |
| Anderson 17 | Smith 23 |
| Bourguignon 11 | Likierman 0 |
| Harris–Raviv 97 | Hirschey–Pappas 2 |
| Edlefsen 21 | Highton–Webb 3 |
| Deaton–Muellbauer 32 | Lamm 15 |
| Pollak–Wales 142 | Bartlett 6 |
| Balk 1 | Baye 3 |
| Helpman 7 | Link-Long 7 |
| King 23 | |
| Nakamura–Nakamura 80 | <i>Oxford Bulletin of Economics and Statistics</i> |
| Bell 2 | Sen 50 |
| Rob 1 | Banerjee 8 |
| | Boltho 0 |
| | Stromback 0 |
| <i>Journal of Public Economics</i> | Winters 0 |
| Sandmo 88 | Mayhew–Rosewell 5 |
| Courant–Rubinfeld 9 | Lye–Silbertson 1 |
| Hey–Mavromaras 9 | Metwally–Tamaschke 2 |
| Weymark 5 | Tsegaye 0 |
| Bennett 0 | Brundell <i>et al.</i> 9 |
| Berglas 20 | King 3 |

ACKNOWLEDGMENTS

For their helpful comments, I thank Danny Blanchflower, Gordon D. A. Brown, David De Meza, Amanda Goodall, David Laband, Steve Nickell, Mark Stewart, Radu Vranceanu, Ian Walker, Ken Wallis, Michael Waterson and an anonymous referee and the editor of this Journal. I would like to record that my thoughts on retrospective reward structures have been heavily influenced by the excellent suggestions of, and electronic conversations with, David N. Laband, and that Michael Waterson's ideas first suggested to me that the order of articles within a journal might be correlated with their eventual citations.

REFERENCES

- BAYERS, N. K. (2005). Using ISI data in the analysis of German national and institutional research output. *Scientometrics*, **62**, 155–63.
- COUPE, T. (2003). The price is right: an analysis of best-paper prizes. Unpublished paper, National University of Kyiv-Mohyla, Ukraine.
- FREY, B. S. (2003). Publishing as prostitution? Choosing between one's own ideas and academic success. *Public Choice*, **116**, 205–23.
- GARFIELD, E. and WELLJAMS-DOROF, A. (1992). Of Nobel class: a citation perspective on high impact research authors. *Theoretical Medicine*, **13**, 117–35.
- GOODALL, A. H. (2006). Should research universities be led by top researchers, and are they? A citations analysis. *Journal of Documentation*, **62**, 388–411.
- HAMERMESH, D. S., JOHNSON, G. E. and WEISBROD, A. (1982). Scholarship, citations and salaries: economic rewards in economics. *Southern Economic Journal*, **49**, 472–81.
- HUDSON, J. (1996). Trends in multi-authored papers in economics. *Journal of Economic Perspectives*, **10**, 153–8.
- (2006). Be known by the company you keep: citations—quality or chance? Unpublished paper, University of Bath.
- JOHNSON, D. (1997). Getting noticed in economics: the determinants of academic citations. *The American Economist*, **41**, 43–52.

- KALAITZIDAKIS, P., MAMUNEAS, T. P. and STENGOS, T. (1999). European economics: an analysis based on publications in the core journals. *European Economic Review*, **43**, 1150–68.
- LABAND, D. N. (1990). Is there value-added from the review process in economics? Preliminary evidence from authors. *Quarterly Journal of Economics*, **105**, 341–52.
- and PIETTE, M. J. (1994). The relative impacts of economics journals, 1970–1990. *Journal of Economic Literature*, **32**, 640–66.
- and TOLLISON, R. D. (2000). Intellectual collaboration. *Journal of Political Economy*, **108**, 632–62.
- and ——— (2003). Dry holes in economic research. *Kyklos*, **56**, 161–73.
- MONASTERSKY, R. (2005). The number that is devouring science. *Chronicle of Higher Education*, 14 October.
- MOORE, W. J., NEWMAN, R. J. and TURNBULL, G. K. (1998). Do academic salaries decline with seniority? *Journal of Labor Economics*, **16**, 352–66.
- OPPENHEIM, C. (1995). The correlation between citation counts and the 1992 Research Assessment Exercise ratings for British library and information science university departments. *Journal of Documentation*, **51**, 18–27.
- OSWALD, A. J. (1991). Progress and microeconomic data. *Economic Journal*, **101**, 75–81.
- SEGLÉN, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, **314**, 497.
- STARBUCK, W. H. (2003). Turning lemons into lemonade: where is the value in peer reviews? *Journal of Management Inquiry*, **12**, 344–51.
- (2005). How much better are the most prestigious journals? The statistics of academic publication. *Organization Science*, **16**, 180–200.
- SUSSMUTH, B., STEININGER, M. and GHIO, S. (2006). Towards a European economics of economics: monitoring a decade of top research and providing some explanation. *Scientometrics*, **66**, 579–612.
- THURSBY, J. G. (2000). What do we say about ourselves and what does it mean? Yet another look at economics department research. *Journal of Economic Literature*, **38**, 383–404.
- TOUTKOUSHIAN, R. K. (1994). Using citations to measure sex discrimination in faculty salaries. *Review of Higher Education*, **18**, 61–82.
- VAN DALEN, H. P. and HENKENS, K. (2005). Signals in science: on the importance of signaling in gaining attention in science. *Scientometrics*, **64**, 209–33.
- VAN RAAN, A. F. J. (1998). Assessing the social sciences: the use of advanced bibliometric methods as a necessary complement to peer review. *Research Evaluation*, **7**, 2–6.