Robust Ranking of Journal Quality:
An Application to Economics*

Chia-Lin Chang
Department of Applied Economics
Department of Finance
National Chung Hsing University
Taiwan

Esfandiar Maasoumi
Department of Economics
Emory University

Michael McAleer
Department of Quantitative Finance
National Tsing Hua University
Taiwan
and
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute
The Netherlands
and
Department of Quantitative Economics
Complutense University of Madrid

Revised: October 2013

* The authors are grateful to two referees for helpful comments and suggestions. For financial support, the first author wishes to thank the National Science Council, Taiwan, and the third author wishes to acknowledge the Australian Research Council and the National Science Council, Taiwan.
Abstract

The paper focuses on the robustness of rankings of academic journal quality and research impact in general, and in Economics, in particular, based on the widely-used Thomson Reuters ISI Web of Science citations database (ISI). The paper analyses 299 leading international journals in Economics using quantifiable Research Assessment Measures (RAMs), and highlights the similarities and differences in various RAMs, which are based on alternative transformations of citations and influence. All existing RAMs to date have been static, so two new dynamic RAMs are developed to capture changes in impact factor over time and escalating journal self citations. Alternative RAMs may be calculated annually or updated daily to determine When, Where and How (frequently) published papers are cited (see Chang et al. (2011a, b, c)). The RAMs are grouped in four distinct classes that include impact factor, mean citations and non-citations, journal policy, number of high quality papers, and journal influence and article influence. These classes include the most widely used RAMs, namely the classic 2-year impact factor including journal self citations (2YIF), 2-year impact factor excluding journal self citations (2YIF*), 5-year impact factor including journal self citations (5YIF), Eigenfactor (or Journal Influence), Article Influence, h-index, and PI-BETA (Papers Ignored - By Even The Authors). As all existing RAMs to date have been static, two new dynamic RAMs are developed to capture changes in impact factor over time (5YD2 = 5YIF/2YIF) and Escalating Self Citations. We highlight robust rankings based on the harmonic mean of the ranks of RAMs across the 4 classes. It is shown that emphasizing the 2-year impact factor of a journal, which partly answers the question as to When published papers are cited, to the exclusion of other informative RAMs, which answer Where and How (frequently) published papers are cited, can lead to a distorted evaluation of journal quality, impact and influence relative to the harmonic mean of the ranks.

Keywords: Research assessment measures, Impact factor, IFI, C3PO, PI-BETA, STAR, Eigenfactor, Article Influence, h-index, 5YD2, ESC, harmonic mean of the ranks, economics, journal rankings.

JEL Classifications: C18, C81, Y10.
1. Introduction

The perceived quality of academic journals is routinely based on untested expert assessments of journal impact and influence, the number of high quality papers, journal policy, and quantitative or qualitative information about a journal, as well as quantifiable bibliometric Research Assessment Measures (RAMs). In this context, the leading database for generating RAMs to evaluate the research performance of individual researchers and the quality of academic journals is the Thomson Reuters ISI Web of Science (2011) database (hereafter ISI), where most RAMs are based on alternative transformations of citations and influence data. All existing RAMs to date have been static, so two new dynamic RAMs are developed to capture changes in impact factor over time and escalating journal self citations.

Although there are important caveats regarding the methodology and data collection methods underlying any database (see, for example, Seglen (1997) and Chang et al. (2011a, b, c, d) for caveats regarding ISI), the ISI citations database is the oldest and most prestigious source of RAMs, and undoubtedly the benchmark against which other general databases, such as SciVerse Scopus, Google Scholar and Microsoft Academic Search, social science open access repositories, such as the Social Science Research Network (SSRN), and discipline-specific databases, such as Research Papers in Economics (RePEc), are compared. Journal publishers promote the ISI impact factor (see below) of their journals and, if their journals do not yet have an impact factor, publicize the fact that their journals have either been selected for coverage in ISI or have applied for inclusion in ISI.

Various RAMs have been used to compare journals in a wide range of ISI disciplines, such as the 40 leading journals in Economics and the leading 10 journals in each of Management, Finance and Marketing (Chang et al. (2011a)), the leading 6 journals in each of 20 disciplines in the Sciences (Chang et al (2011b)), the leading journals in a sub-discipline of Economics, namely Econometrics, and Statistics (Chang et al. (2011c)), and the leading 26 journals in Neuroscience (Chang et al. (2011d)). As not all of the leading journals in the ISI discipline of Economics have yet been analysed in terms of citations, quality and impact, one of the primary aims of this paper is to undertake such a rankings analysis.

When impact factors and other RAMs-based citations data are used without appropriate care, misleading, unintended inferences may be drawn. Seglen (1997) cautioned against using
impact factors of journals to evaluate scientific research. Nevertheless, as quantified metrics, citations are necessary for evaluating the impact and visibility of high quality and significant scientific research output. Embracing journal citations as a valid measure of scientific research output, Hirsch (2005) suggested a widely-used measure, the h-index, for quantifying an individual researcher’s scientific research output. Although citations data are used more widely as a measure of research productivity in the sciences than in the social sciences, the h-index is now widely used to evaluate both the research output of individual researchers and to quantify the number of highly-cited publications in academic journals in both the sciences and social sciences.

The perceived research performance of individual researchers is a key issue in hiring, tenure and promotion decisions. The perceived quality of academic journals has long been used as a suitable proxy for quality, especially for less established scholars, and especially in the social sciences, as leading journals tend to publish significant scientific research output. Evaluations of individual researchers, institutions and journals have been undertaken over an extended period (see, for example, Neary et al. (2003), and the other papers in the same issue of the Journal of the European Economic Association). Emphasis on the appropriate weights attached to journals has been analysed, with the suggestions that it can be important, especially in ranking individuals, though not necessarily institutions. Kalaitzidakis et al. (2003) determine weights for the thirty leading research journals, while Axarloglou and Theoharakis (2003) analyse the survey results of journal quality perceptions of 2,103 AEA economists worldwide. The overall impression given by these interesting papers is that consensus is difficult to reach in terms of ranking the leading journals in economics.

The convention in the sciences and social sciences is such that the acceptance of a paper for journal publication is based on the expertise of a few editors and referees. Although the number varies considerably across disciplines, acceptance of a paper for journal publication undeniably relies on a handful of decision makers, who determine the explicit rejection rate of a journal before publication. As editors and referees are not immune from making type 1 and type 2 errors regarding the latent quality and likely future impact of submitted papers, the rejection of a paper by a journal is not necessarily a correct decision, just as acceptance of a paper for publication is not a guarantee that it will have future impact and influence.
In comparison with the rejection rate of a journal before publication, there is an equally important implicit rejection rate after publication. Rather than relying on a small number of editors and reviewers, the rejection rate after publication relies on the worldwide scientific community. As argued in Chang et al. (2011c), the proportion of published papers that is ignored by the profession, and possibly by the authors themselves, is an important impact performance measure after publication.

The paper is also concerned with highlighting the upsurge in journal self citations in recent years. It would seem useful to present RAMs that capture such an escalation of journal self citations over time, and also to mitigate such an effect. One new dynamic RAM addresses the different speeds at which citations are accrued in the sciences and social sciences, and a second new dynamic RAM captures the escalation of journal self citations over time.

The RAMs may be classified according to four distinct classes, namely Class 1: “impact factor, mean citations and non-citations”, Class 2: “journal policy”, Class 3: “number of high quality papers”, and Class 4: “journal influence and article influence”. It is shown that emphasizing the 2-year impact factor of a journal to the exclusion of other informative RAMs can lead to a distorted evaluation of journal quality, impact and influence relative to the harmonic mean of the ranks of 13 existing and 2 new dynamic RAMs across the 4 classes. Together with the arithmetic and geometric means, the harmonic mean is one of the three Pythagorean means, and is defined as the reciprocal of the arithmetic mean of the reciprocals.

This paper examines the importance of RAMs as viable rankings criteria in Economics, and attempts to answer some important questions raised in Chang et al. (2011a, b, c), namely When, Where and How (frequently) are published papers cited in leading journals in a discipline. In this paper, we evaluate the usefulness of 15 RAMs for 299 leading journals, and suggest a robust rankings method of alternative RAMs using the harmonic mean of the ranks. The rankings based on any single RAM, such as the h-index or the 2 year impact factors are placed in context, and may be seen as extremes since they are clearly subsumed by the harmonic mean of the ranks when all other RAMs are given zero weights, except the RAM in question.

The plan of the remainder of the paper is as follows. Section 2 presents some key RAMs using ISI data that may be calculated annually or updated daily, including the most widely
used RAM, namely the classic 2-year impact factor including journal self citations (2YIF), 2-year impact factor excluding journal self citations (2YIF*), 5-year impact factor including journal self citations (5YIF), Immediacy (or zero-year impact factor (0YIF)), Eigenfactor (or Journal Influence), Article Influence, C3PO (Citation Performance Per Paper Online), h-index, PI-BETA (Papers Ignored - By Even The Authors), 2-year Self-citation Threshold Approval Ratings (2Y-STAR), Historical Self-citation Threshold Approval Ratings (H-STAR), Impact Factor Inflation (IFI), and Cited Article Influence (CAI). Two new dynamic RAMs are developed, namely 5YD2 (5YIF Divided by 2YIF) and ESC (Escalating Self Citations). Section 3 discusses and analyses 15 RAMs for 299 leading journals in the ISI category of Economics, and provides a harmonic mean of the ranks as a robust rankings method of alternative RAMs. Section 4 summarizes the ranking outcomes and gives some practical suggestions as to how to rank journal quality and impact.

2. Research Assessment Measures (RAM)

A widely-used RAM database for evaluating journal impact and quality is the Thomson Reuters ISI Web of Science (2011). As discussed in a number of papers (for example, Chang et al. (2011a, b, c)), the RAMs are intended as descriptive statistics to capture journal impact and performance, and are not based on a mathematical model. Hence, in what follows, no optimization or estimation is required in calculating the alternative RAMs.

As the alternative RAMs that are provided in ISI and in several recent publications may not be widely known, this section provides a brief description and definition of 13 RAMs that may be calculated annually or updated daily to answer the questions as to When, and Where and How (frequently), published papers are cited (for further details, see Chang et al. (2011a, b, c)). Two new dynamic RAMs that are calculated annually, namely 5YD2 and ESC, are also suggested. The answers to When published papers are cited are based on the set \{2YIF, 2YIF*, 5YIF, Immediacy\}, and the answers to Where and How (frequently) published papers are cited are based on the set \{Eigenfactor, Article Influence, IFI, 5YD2, H-STAR, 2Y-STAR, ESC, C3PO, h-index, PI-BETA, CAI\}, as will be discussed below.

2.1 Annual RAM
With three exceptions, namely Eigenfactor, Article Influence and Cited Article Influence, existing RAMs are based on citations data and are reported separately for the sciences and social sciences. RAMs may be computed annually or updated daily. The annual RAMs given below are calculated for a Journal Citations Reports (JCR) calendar year, which is the year before the annual RAM are released. For example, the RAMs were released in late-June 2011 for the JCR calendar year 2010.

(1) 2-year impact factor including journal self citations (2YIF):
The classic 2-year impact factor including journal self citations (2YIF) of a journal is typically referred to as “the impact factor”, is calculated annually, and is defined as “Total citations in a year to papers published in a journal in the previous 2 years / Total papers published in a journal in the previous 2 years”. The choice of 2 years by ISI is arbitrary. It is widely held in the academic community, and certainly by the editors and publishers of journals, that a higher 2YIF is better than lower.

(2) 2-year impact factor excluding journal self citations (2YIF*):
ISI also reports a 2-year impact factor without journal self citations (that is, citations to a journal in which a citing paper is published), which is calculated annually. As this impact factor is not widely known or used, Chang et al. (2011c) refer to this RAM as 2YIF*. Although 2YIF* is rarely reported, a higher value would be preferred to lower.

(3) 5-year impact factor including journal self citations (5YIF):
The 5-year impact factor including journal self citations (5YIF) of a journal is calculated annually, and is defined as “Total citations in a year to papers published in a journal in the previous 5 years / Total papers published in a journal in the previous 5 years.” The choice of 5 years by ISI is arbitrary. Although 5YIF is not widely reported, a higher value would be preferred to lower.

(4) Immediacy, or zero-year impact factor including journal self citations (0YIF):
Immediacy is a zero-year impact factor including journal self citations (0YIF) of a journal, is calculated annually, and is defined as “Total citations to papers published in a journal in the same year / Total papers published in a journal in the same year.” The choice of the same year by ISI is arbitrary, but the nature of Immediacy makes it clear that a very short run
outcome is under consideration. Although Immediacy is rarely reported, a higher value would be preferred to lower.

(5) 5YIF Divided by 2YIF (5YD2):
As both 2YIF and 5YIF include journal self citations, if it is assumed that journal self citations are uniformly distributed over the 5-year period for calculating 5YIF, their ratio will eliminate the effect of journal self citations and capture the increase in the citation rate over time. In any event, the impact of journal self citations should be mitigated with the ratio of 5YIF to 2YIF. We define a new dynamic RAM as 5YD2 as “5YD2 = 5YIF / 2YIF”. In the natural, physical and medical sciences, where citations are observed with a frequency of weeks and months rather than years, it is typically the case that 5YIF < 2YIF (see Chang et al. (2011b, d)), whereas the reverse, 5YIF > 2YIF, seems to hold generally in the social sciences, where citations tend to increase gradually over time (see Chang et al. (2011a, c)). Thus, emphasizing the different speeds at which citations are accrued over time, a lower 5YD2 would be preferred to higher in the sciences, while a higher 5YD2 would be preferred to lower in the social sciences.

(6) Eigenfactor (or Journal Influence):
The Eigenfactor score (see Bergstrom (2007), Bergstrom and West (2008), Bergstrom, West and Wiseman (2008)) is calculated annually (see www.eigenfactor.org), and is defined as: “The Eigenfactor Score calculation is based on the number of times articles from the journal published in the past five years have been cited in the JCR year, but it also considers which journals have contributed these citations so that highly cited journals will influence the network more than lesser cited journals. References from one article in a journal to another article from the same journal are removed, so that Eigenfactor Scores are not influenced by journal self-citation.” The value of the threshold that separates ‘highly cited’ from ‘lesser cited’ journals, as well as how the former might ‘influence the network more’ than the latter, are based on the Eigenfactor score of the citing journal. Thus, Eigenfactor might usefully be interpreted as a weighted total citations score, or a “Journal Influence” measure. A higher Eigenfactor score would be preferred to lower.

(7) Article Influence (or Journal Influence per Article):
Article Influence (see Bergstrom (2007), Bergstrom and West (2008), Bergstrom, West and Wiseman (2008)) measures the relative importance of a journal’s citation influence on a per-
Despite the misleading suggestion of measuring “Article Influence”, as each journal has only a single “Article Influence” score, this RAM is actually a “Journal Influence per Article” score. Article Influence is a scaled Eigenfactor score, is calculated annually, is standardized to have a mean of one across all journals in the Thomson Reuters ISI database, and is defined as “Eigenfactor score divided by the fraction of all articles published by a journal.” A higher Article Influence would be preferred to lower.

(8) IFI:
The ratio of 2YIF to 2YIF* is intended to capture how journal self citations can inflate the impact factor of a journal, whether this is an unconscious self-promotion decision made independently by publishing authors or as an administrative decision undertaken by a journal’s editors and/or publishers. Chang et al. (2011a) define Impact Factor Inflation (IFI) as “IFI = 2YIF / 2YIF*”. The minimum value for IFI is 1, with any value above the minimum capturing the effect of journal self citations on the 2-year impact factor. A lower IFI would be preferred to higher.

(9) H-STAR:
ISI has implicitly recognized the inflation in journal self citations by calculating an impact factor that excludes self citations, and provides data on journal self citations, both historically (for the life of the journal) and for the preceding two years, in calculating 2YIF. Chang et al. (2011b) define the Self-citation Threshold Approval Rating (STAR) as the percentage difference between citations in other journals and journal self citations. If HS = historical journal self citations, then Historical STAR (H-STAR) is defined as “H-STAR = [(100-HS) - HS] = (100-2HS)”. If HS = 0 (minimum), 50 or 100 (maximum) percent, for example, H-STAR = 100, 0 and -100, respectively. A higher H-STAR would be preferred to lower.

(10) 2Y-STAR:
If 2YS = journal self citations over the preceding 2-year period, then the 2-Year STAR is defined as “2Y-STAR = [(100-2YS) – 2YS] = (100-2(2YS))”. If 2YS = 0 (minimum), 50 or 100 (maximum) percent, for example, 2Y-STAR = 100, 0 and -100, respectively. A higher 2Y-STAR would be preferred to lower.

(11) Escalating Self Citations (ESC):
As self citations for many journals in the sciences and social sciences have been increasing over time, it would seem useful to present a dynamic RAM that captures such an escalation over time. The difference $2YS – HS$ measures Escalating Self Citations in journals over the most recent 2 years relative to the historical period for calculating citations, which will differ across journals. We define a new dynamic RAM as “$ESC = 2YS – HS = (H\text{-STAR} – 2Y\text{-STAR}) / 2$”. Given the range of each of H\text{-STAR} and 2Y\text{-STAR} is (-100, 100), the range of ESC is also (-100, 100), with -100 denoting minimum, and 100 denoting maximum, escalation. A lower ESC would be preferred to higher.

2.2 Daily Updated RAM

Some RAMs are updated daily, and are reported for a given day in a calendar year rather than for a JCR year.

(12) C3PO:
ISI reports the mean number of citations for a journal, namely total citations up to a given day divided by the number of papers published in a journal up to the same day, as the “average” number of citations. In order to distinguish the mean from the median and mode, the C3PO of an ISI journal on any given day is defined by Chang et al. (2011a) as “C3PO (Citation Performance Per Paper Online) = Total citations to a journal / Total papers published in a journal.” A higher C3PO would be preferred to lower. [Note: C3PO should not be confused with C-3PO, the Star Wars android.]

(13) h-index:
The h-index (Hirsch, 2005)) was originally proposed to assess the scientific research productivity and citations impact of individual researchers. However, the h-index can also be calculated for journals, and should be interpreted as assessing the impact or influence of highly cited journal publications. The h-index of a journal on any given day is based on historically cited and citing papers, including journal self citations, and is defined as “$h\text{-index} = \text{number of published papers, where each has at least h citations}$.” The h-index differs from an impact factor in that the h-index measures the number of highly cited papers historically. A higher h-index would be preferred to lower.

(14) PI-BETA:
This RAM measures the proportion of papers in a journal that has never been cited, As such, PI-BETA is, in effect, a rejection rate of a journal **after** publication. Chang et al. (2011c) argue that lack of citations of a published paper, especially if it is not a recent publication, reflects on the quality of a journal by exposing: (i) what might be considered as incorrect decisions by the members of the editorial board of a journal; and (ii) the lost opportunities of papers that might have been cited had they not been rejected by the journal. Chang et al. (2011c) propose that a paper with zero citations in ISI journals can be measured by PI-BETA (= Papers Ignored (PI) - By Even The Authors (BETA)), which is calculated for an ISI journal on any given day as “Number of papers with zero citations in a journal / Total papers published in a journal.” As journals would typically prefer a higher proportion of published papers being cited rather than ignored, a lower PI-BETA would be preferred to higher.

**CAI:**
Article Influence is intended to measure the average influence of an article across the sciences and social sciences. As an article with zero citations typically does not have any (academic) influence, a more suitable measure of the influence of cited articles would seem to be Cited Article Influence (CAI). Chang et al. (2011b) define CAI as “CAI = (1 - PI-BETA)(Article Influence)”. If PI-BETA = 0, then CAI is equivalent to Article Influence; if PI-BETA = 1, then CAI = 0. As Article Influence is calculated annually and PI-BETA is updated daily, CAI may be updated daily. A higher CAI would be preferred to lower.

3. Analysis of RAM for 299 Leading Journals in Economics

As no single RAM captures adequately the quality, impact and influence of a journal, any general measure of journal quality and impact, such as a harmonic mean of the ranks as a robust rankings method of alternative RAMs (see, for example, Chang and McAleer (2013)), should depend on the following four distinct classes:

(i) **Class 1**: “impact factor, mean citations and non-citations” (2YIF, 2YIF*, 5YIF, Immediacy, C3PO, PI-BETA);
(ii) **Class 2**: “journal policy” (IFI, H-STAR, 2Y-STAR, 5YD2, ESC);
(iii) **Class 3**: “number of high quality papers” (h-index);
(iv) **Class 4**: “journal influence and article influence” (Eigenfactor, Article Influence, CAI).
As each of the four classes has equal weight in the calculation of the harmonic mean of the ranks, the h-index has the single highest weight of the 15 RAMs. For journals that have been included in ISI for less than five years, Class 1 does not include 5YIF, Class 2 does not include 5YD2, and Class 4 does not include Article Influence and CAI, in calculating the harmonic mean of the ranks of the RAMs. Class 3 includes only the h-index. When RAM data for only Eigenfactor are available, Class 4 would be a “journal influence” rather than “journal influence and article influence” class.

As PI-BETA in Class 1 ranks journals from low to high rather than high to low, 1 – PI-BETA would be used in calculating the harmonic mean of the ranks of the original RAMs in Class 1, as appropriate. In a similar vein, IFI and ESC in Class 2 also rank journals from low to high rather than high to low, so that 1/IFI and –ESC would be used in calculating the harmonic mean of the ranks of the original RAMs, as appropriate.

The harmonic mean of the ranks of Classes 1 and 2 are based on 5 and 4 RAMs, respectively, whereas the rankings according to h-index and Eigenfactor are the sole representatives in Classes 3 and 4, respectively. As Classes 1, 2, 3 and 4 have, respectively, 5, 4, 1 and 1 journals in calculating the harmonic mean of the ranks of the 4 classes, the weights for the RAMs in Classes 3 and 4 are the highest, followed by Classes 2 and 1, respectively. The harmonic mean of the ranks across the 4 distinct classes lead to a weighted harmonic mean of the ranks of the 11 RAMs.

The ISI category of Economics has one of the largest numbers of journals, at 304, of any discipline, and therefore has broad coverage, including the sub-disciplines and overlapping disciplines of, among others, accounting, agriculture, banking, derivatives, econometrics, economic history, economic theory, education, energy, environment, experiments, forecasting, futures, game theory, growth, health, history, industrial organization, innovation, insurance, international economics, labour, law, macroeconomics, mathematics, management, media, microeconomics, money, network, organisation, philosophy, policy, psychology, real estate, regional science, regulation, resources, risk, sociology, spatial analysis, statistics, strategy, taxation, technology, time series analysis, transportation, uncertainty, and welfare.
We compare the RAMs that are based on ISI citations data (see Tables 1-5). Only articles from the ISI Web of Science are included in the citations data, which were downloaded from ISI on 10 August 2011 for all journals. The ISI data set starts in 1899, so all data are from the inception of the respective journals, except for American Economic Review (from 1964), Value in Health (from 2006), Economic Journal (from 1957), American Journal of Agricultural Economics (from 1984), and Journal of Economic History (from 1962) (the numbers in parentheses are the first years in which the numbers of articles in the respective journals were below 10,000, which is the upper limit for which daily RAM (namely, h-index, C3PO, PI-BETA and CAI) are reported in ISI).

Some comments on the 304 journals in the ISI category of Economics are in order. Annual Review of Economics, Spanish Economic Review, Review of Agricultural Economics, and Investigaciones Economicas had blank (as distinct from zero) entries for Immediacy in the ISI dataset. Zero entries have been substituted rather than deleting these 4 journals from the rankings analysis as they have non-zero 2YIF. Inzinerine Ekonomika – Engineering Economics has a non-zero 2YIF but a zero entry for 2YIF*, so that IFI cannot be calculated. This journal has also been deleted from the dataset. Estudios de Economia has a zero 2YIF entry, while 3 journals, namely Applied Economic Perspectives and Policy, IMF Economic Review, and Series – Journal of the Spanish Economic Association, have blank 2YIF entries. As a non-zero 2YIF is required for ranking the journals, these 4 journals are deleted from the dataset. Of the remaining 299 journals listed in ISI in Table 1, 89 journals have been included in ISI for less than 5 years, so that the RAMs for 5YIF, Article Influence, CAI and 5YD2 are available for 210 journals.

In Table 1 we evaluate 15 RAMs for the 299 leading journals in Economics, which are ranked according to 2YIF. The means and ranges of 2YIF are, respectively, 1.036 and (0.003, 7.432), of 2YIF* are 0.889 and (0.001, 7.270), of 5YIF are 1.595 and (0.058, 8.076), and of Immediacy are 0.237 and (0, 3.467). These impact factors are generally consistent with the related areas of Business - Finance, Management, and Marketing (see Chang et al. (2011a)), but are typically lower than many disciplines in the sciences (see Chang et al. (2011b)). Two surprises in the top 10 journals based on 2YIF are Technological and Economic Development of Economy (at number 3) and Journal of Business Economics and Management (at number 7), both of which are co-published with Vilnius Gediminas Technical University, Lithuania. The Immediacy of Asian Economic Policy Review is extraordinarily high at 3.467, especially
relative to the mean value. In Table 1, the mean and range of 5YD2 are 1.380 and (0.686, 3.205), respectively, so that 5YIF is considerably higher than 2YIF, which is to be expected in Economics, which is a social sciences discipline as compared with many journals in the sciences. Developing Economics has a very high 5YD2 compared with the mean RAM value.

Journal self citations in Economics seem relatively high, with a mean IFI of 1.442 and a range of (1, 25.417), with 9 IFI scores in excess of 3, namely Economia Politica (at 25.417), Asian Journal of Technology Innovation (at 9.927), Pacific Economic Bulletin (at 6.706), Ekonomista (at 4.651), Economia Chilena (at 4.083), Journal of Banking & Finance (at 3.651), Amfiteatru Economic (at 3.636), Politicka Ekonomie (at 3.457), and Actual Problems of Economics (at 3). On average, the 299 leading journals in Economics have 2YIF that is inflated by a factor of 1.442 through journal self citations. It is also worth mentioning that 31 of the 299 journals have zero self citations.

The h-index has a mean of 27.244 and a range of (1, 215), with the three highest h-index values being 215, 210 and 197 for American Economic Review, Econometrica and Journal of Political Economy, respectively. There are 117 journals with an h-index that is less than 10, including 9 journals with an h-index of 1. The median h-index is 17, and the mode is 3. Many of the journals with low h-indexes have been included in ISI for less than five years.

In terms of mean citations, C3PO has a mean of 5.51 and a range of (0.01, 59.65), with significant contributions coming from the leading 3 journals, namely Journal of Financial Economics, Quarterly Journal of Economics, and Journal of Political Economy. The median C3PO is 2.46, and 28% of the 299 journals have C3PO values that are less than one. Eigenfactor has a mean of 0.005 and a range of (0, 0.101), with 2 journals, American Economic Review and Journal of Finance, clearly having the highest scores, and hence the greatest Journal Influence. Article Influence has a mean of 1.334 and a range of (0.012, 11.741), with 4 journals, Quarterly Journal of Economics, Journal of Political Economy, Econometrica, and Journal of Economic Literature, having the greatest journal influence. As Article Influence is standardized to have a mean of one across all journals in the Thomson Reuters ISI database, the mean article influence in Economics is greater than for the full list of journals in the ISI database. Cited Article Influence (CAI) has a mean of 0.925 and a range of (0, 10.309), with 3 journals, Quarterly Journal of Economics, Journal of Political
Economy, and Review of Economic Studies, having the greatest influence on the basis of cited journal articles.

H-STAR and 2Y-STAR for the 299 journals are not high, with a mean of 72.5 and a range of (-64, 100) for H-STAR, and a much lower mean of 63.9 and a wider range of (-92, 100) for 2Y-STAR. The H-STAR and 2Y-STAR means of 72 and 64 reflect journal self citations of 14% and 18%, respectively, historically and for the preceding two years. On average, journal self citations have increased over the preceding two years as compared with historical levels. The ESC mean is 4.3 and has a range of (-28, 45). On average, self citations are escalating, with 35 journals having no change in the preceding 2 years relative to historical levels, 69 journals decreasing in self citations, and 195 journals increasing in self citations. Overall, two-thirds of the ISI Economics journals have escalating self citations relative to historical levels.

The PI-BETA scores are illuminating. The mean is 0.492 and the median is 0.471 so that, on average, almost one of every 2 papers that are published in the leading 299 journals in Economics is not cited. The range of (0.054, 0.989) suggests that the journal with the highest percentage of cited papers, Oxford Review of Economic Policy, has one uncited paper for every 20 published papers, while the journal with the lowest percentage of cited papers, Ekonomista, has virtually no cited papers. Of the 299 Economics journals in Table 1, 16 journals have PI-BETA that exceeds 0.9, which means that more than 9 of every 10 published papers in these journals have zero citations. At the other end of the scale, 12 journals have PI-BETA that are less than 0.1, which means that a very high proportion of the papers published in these journals are cited. The PI-BETA values in Table 1 are typically much higher than many disciplines in the sciences (see Chang et al. (2011b)).

As 89 journals have been included in ISI for less than 5 years, and hence do not have corresponding RAMs for 5YIF, 5YD2, Article Influence and CAI, the simple correlations of 15 RAMs for the 210 leading journals in Economics are given in Table 2, while the simple correlations of 11 RAMs for the 299 leading journals are given in Table 3.

There are 6 and 1 RAM pairs for which the correlations exceed 0.9 (in absolute value) in Tables 2 and 3, respectively, and 10 and 3 RAM pairs in Tables 2 and 3, respectively, for which the correlations are in the range (0.8, 0.9), in absolute value. The correlations of 0.984
and 0.98 between 2YIF and 2YIF* in Tables 2 and 3, respectively, are extremely high, which suggests that the 2-year impact factors including and excluding self citations are very similar for leading journals in Economics. A similar comment applies to the very high correlations for the pairs (2YIF, 5YIF), (2YIF*, 5YIF) and (Article Influence, CAI) in Table 2. The 2 new RAMs, 5YD2 and ESC, are not highly correlated with each other or any other RAMs in tables 2 and 3, which suggests that they provide useful additional information about journal impact and influence.

One of the primary purposes of the paper is to determine if reliance on the classic 2-year impact factor of a journal, 2YIF, to the exclusion of the other RAMs can lead to a distorted evaluation of journal quality, impact and influence. In order to provide a robust rankings measure based on the 11 RAMs, 6 of which, namely 2YIF, 2YIF*, IFI, Immediacy, C3PO and PI-BETA, are based on ratios, the robust rankings of the 299 leading journals in Economics given in Table 4 are based on the harmonic mean of the ranks.

Although there are 5 RAMs in Class 1, namely 2YIF, 2YIF*, Immediacy, C3PO and PI-BETA, there are 48 journals with Immediacy values of zero. As the inclusion of Immediacy would restrict discrimination of the journals, the harmonic mean of the ranks for Class 1 is based on 2YIF, 2YIF*, C3PO and PI-BETA. Of the 4 RAMs in Class 2, namely IFI, H-STAR, 2Y-STAR and ESC, there are 31 IFI scores of 1, 16 H-STAR scores of 1, 31 2Y-STAR scores of 1, and 35 ESC scores of 70, the outcome being 10 journals ranked equal first according to the harmonic mean of the ranks. As a reasonably large number of journals seem to have displayed similar “journal policy” regarding self citations over the past 2 years, 5 years and historically, the RAMs in Class 2 are not able to discriminate among the leading journals in Economics, and hence will not be used in calculating the harmonic mean of the ranks. The harmonic mean of the ranks of the 11 RAMs of the 299 journals are, therefore, based on the harmonic mean of the harmonic means of the ranks of Class 1, h-index from Class 3, and Eigenfactor from Class 4.

The journals in Table 4 are ranked according to the harmonic mean of the ranks (given as Harmonic Mean). The number 1 ranked journal is American Economic Review, which has moved up 13 places (given in the last column as Difference = 2YIF ranking – Harmonic Mean ranking) from 14 according to 2YIF. In comparison with the rankings in Table 1 that are based on 2YIF, only 2 journals remain unchanged in Table 4, namely Journal of Finance
at number 5 and Economia Chilena at number 289. Many journals have had substantial shifts in rankings. The greatest improvement was 167 for Economics Letters (from 209 to 42), and the largest drop was 126 for Transformations in Business & Economics (from 53 to 179). There were 7 journals that improved their ranking by more than 100, and 7 journals that fell by more than 100 in the rankings.

Of the leading 10 journals according to 2YIF in Table 1, 6 journals remain in the top 10 according to the Harmonic Mean, namely Journal of Economic Literature (from 1 to 2), Quarterly Journal of Economics (from 2 to 3), Journal of Financial Economics (from 8 to 4), Journal of Finance (remaining at 5), Journal of Political Economy (from 6 to 7), and Review of Financial Studies (from 4 to 9). The 4 journals to have slipped out of the top 10 are Journal of Economic Perspectives (from 10 to 12), Technological and Economic Development of Economy (from 3 to 18), Brookings Papers on Economic Activity (from 9 to 45), and Journal of Business Economics and Management (from 7 to 57).

The use of the harmonic mean of the ranks may be seen as rewarding or penalizing widely-varying rankings across alternative RAMs. The harmonic mean of the ranks tends to reward journals with strong individual performances according to one or more RAMs, so that even one very strong performance can lead to a greatly improved ranking. There can be disagreement among the weights to be used, as well as about whether the harmonic, geometric or arithmetic means of the ranks might be the most appropriate Pythagorean mean of the ranks. The RAMs provided in Tables 1 and 4 allow alternative weights to be used for different journals, but concentration on 2YIF alone, with a zero weight for all other RAMs, would seem to be highly restrictive.

The results in Table 4 could also be used to rank journals in various sub-disciplines in economics, such as economic theory, econometrics, macroeconomics and financial economics, as well as journals of academic societies, such as various journals of the American Economic Association. Chang et al. (2011c) ranked the top 10 journals in econometrics using an earlier data set, and these could easily be updated using these results.

The simple ranking correlations of the 11 RAMs for the 299 leading journals in Economics, based on the rankings in Table 4, are given in Table 5. The correlations in Table 5 are not very close (in absolute value) to the correlations in Table 3 for the original RAM scores.
There are 7 RAM pairs for which the correlations exceed 0.9 (in absolute value), with the 2 highest correlations being for the pair (IFI, 2Y-STAR) at 0.998 and (2YIF, 2YIF*) at 0.97. There are also 5 RAM pairs for which the simple correlations are in the range (0.8, 0.9), in absolute value. The correlations of 0.998 and 0.97 for the pairs (IFI, 2Y-STAR) and (2YIF, 2YIF*) suggest that the rankings according to IFI and 2Y-STAR, as well as according to 2YIF and 2YIF*, would be virtually identical.

In Table 5, the 5 highest correlations with the Harmonic Mean are for C3PO (at 0.906), Eigenfactor (at 0.901), h-index (at 0.9), 2YIF* (at 0.864), and 2YIF (at 0.856), which suggests that the classic two-year impact factor including journal self citations is less highly correlated with the Harmonic Mean than are C3PO, Eigenfactor, h-index and the two-year impact factor excluding journal self citations. Thus, 2YIF would not seem to be the most appropriate or robust individual RAM to use if it were intended to capture the harmonic mean of the ranks. Indeed, using 2YIF as a single RAM to capture the quality of a journal would lead to a distorted evaluation of a journal’s impact and influence.

4. Concluding Remarks

The paper evaluated the ranking of academic journal quality and research impact using the Thomson Reuters ISI Web of Science (2011) citations database (hereafter ISI) for the Economics category. As all existing RAMs to date have been static, two new dynamic RAMs are developed to capture changes in impact factor over time and escalating journal self citations. This paper analysed the leading 299 journals in the ISI category of Economics using 15 quantifiable Research Assessment Measures (RAMs). The 15 RAMs that may be calculated annually or updated daily are used to answer the questions as to When, and Where and How (frequently), published papers are cited. The answers to When published papers are cited are based on the set \{2YIF, 2YIF*, 5YIF, Immediacy\}, and the answers to Where and How (frequently) published papers are cited are based on the set \{Eigenfactor, Article Influence, Cited Article Influence, IFI, 5YD2, H-STAR, 2Y-STAR, ESC, C3PO, h-index, PI-BETA\}.

The paper highlighted the similarities and differences in alternative RAMs, and showed that several RAMs were highly correlated so that they had little informative incremental value in
capturing the impact and performance of the highly-cited journals. Other RAMs were not highly correlated with each other, including the 2 new dynamic RAMs, namely 5YD2 and ESC, thereby providing additional information about journal impact and influence. The harmonic mean of the ranks of 11 RAMs were also presented for these 299 leading journals as a robust rankings method.

It was shown that emphasizing the 2-year impact factor of a journal, which partly answers the question as to When published papers are cited, to the exclusion of other informative RAMs, which answer Where and How (frequently) published papers are cited, could lead to a distorted evaluation of journal quality, impact and influence relative to the harmonic mean of the ranks of RAMs across distinct classes that include impact factor, mean citations and non-citations, journal policy, number of high quality papers, and journal influence and article influence.

The detailed RAMs provided in Tables 1 and 4 for the 299 leading journals in Economics permit robust rankings analyses of various sub-disciplines. Although Chang et al. (2011c) have analysed the leading journals in Econometrics and Statistics, a detailed analysis of the ranking of journals in various sub-disciplines in Economics is a topic for future research.
References


Chang, C.-L., M. McAleer and L. Oxley (2011b), What makes a great journal great in the sciences? Which came first, the chicken or the egg?, *Scientometrics*, 87(1), 17-40.


ISI Web of Science (2011), *Journal Citation Reports, Essential Science Indicators*, Thomson Reuters ISI.

