HOW SHOULD RESEARCH PERFORMANCE BE MEASURED? EVIDENCE FROM RANKINGS OF ACADEMIC ECONOMISTS[*]

Magnus Henrekson[†] and Daniel Waldenström[‡]

March 4, 2008

Abstract: Billions of euros are allocated every year to university research. Increased specialisation and international integration of research and researchers has sharply raised the need for comparisons of performance across fields, institutions and individual researchers. However, there is still no consensus regarding how such rankings should be conducted and what output measures are appropriate to use.

We rank all full professors in a particular discipline, economics, in one European nation using seven established, and some of them commonly used, measures of research performance. Our examination shows both that the rank order can vary greatly across measures, and that depending on the measure used the distribution of total research output is valued very differently. The renowned KMS measure in economics stands out among the measures analysed here. It exhibits the weakest correlation with the others used in our study.

We conclude by giving advice to funding councils and others assessing research quality on how to think about the use of both quantitative and qualitative measures of performance.

*Keywords***:** Impact of research, Ranking, Research productivity, Bibliometrics, Impact Factor.

*JEL***:** A11, A14, B41.

[†] Department of Economics, Stockholm School of Economics, P.O. Box 6501, SE-113 83 Stockholm; and Research Institute of Industrial Economics (IFN), P.O. Box 55665, SE-102 15 Stockholm, Sweden. Ph: +8-46-6654502. E-mail: Magnus.Henrekson@ifn.se. Web: www.ifn.se/mh.
[‡] Research Institute of Industrial Economics (IFN), P.O. Box 55665, SE-102 15 Sweden. Ph: +8-46-6654531. E-mail: Daniel.Waldenström@ifn.se. Web: www.ifn.se/danielw.

Introduction

The increased integration has made the total research market larger, but also more complex. In some disciplines, such as economics, there exists a core of general ideas and concepts that are widely acknowledged and used. At the same time, numerous global and specialised sub-disciplines have emerged and continue to emerge. Patterns are similar in other scientific disciplines. Enhanced cross-border cooperation, integration and exchange of ideas and research personnel have resulted in a sharply increased demand for objective and internationally comparable evaluations of institutions and individual researchers. This is particularly important for the pressing issue in the most advanced nations of finding appropriate criteria for the future allocation of funds to scientific research.

Given this demand, it is of crucial importance that the measures used really capture what is essential, namely quality-adjusted research output including the impact on the research of others. However, at present there exists no perfect measure, and as we will see it is highly unlikely that there will ever be a single measure capturing all relevant aspects. Moreover, available measures in economics and other disciplines have been constructed by individuals and organisations with their specific agenda and vested interests.[1] As a result, the relative valuation of a certain type of research is likely to depend on which measure is used. In the natural sciences, where there is a much longer tradition of evaluating and ranking research performance, the debate on the relative strengths and weaknesses of different measures already has a long history (e.g., Hecht et al. 1998, and Cameron 2005). In our view, it is imperative that this discussion is both broadened and deepened in economics as well.

What should an output measure capture? Citation based measures are most frequently used, but is it necessarily true that the most (least) cited research is also the best

---

[1] In economics considerable efforts have already been made to measure research output and productivity, and to rank individual researchers and institutions. First in the U.S. (e.g. Conroy and Dusansky 1995; Dusansky and Vernon 1998), and then in Europe (e.g., Kalaitzidakis et al. 1999, 2003; Combes and Linnemer 2003; Coupé 2003; Axarloglou and Theoharakis 2003; Lubrano et al. 2003; Tombazos 2005). Especially in Europe the Kalaitzidakis et al. (2003) (KMS) studies have raised controversy (Tombazos 2005), and in particular young economists seem to be greatly influenced both by their relative ranking of journals and the absolute metric they come up with (e.g., Oswald 2006), where a handful of top-ranked journals are given extreme weights relative to other frequently used measures.

1

(worst) research?[2] Can we assume that all important research results are published in refereed journals or should we also include monographs, book chapters and text-books? Is it sufficient to evaluate research based on in which journal an article is published or how many citations it gets?[3] How do we handle the fact that many more articles are published in some sub-disciplines and hence get more citations? How do we assess a researcher who has published one short article in a top-ranked journal relative to a researcher with several frequently cited articles in field journals of relatively low rank? How do we handle dynamic problems arising because the overall competition and journal ranking changes over time? Should we give weight to the impact outside academia, such as impact on policymaking or the policy debate?

The relative importance of these questions differs depending on the issue at hand. Typical decisions where guidance from quantitative measures can be used and is increasingly used are hiring, tenure and promotion decisions, and the allocation of research funds across individuals, research groups, departments, disciplines and universities. In this study we explore whether and to what extent the assessments depend on the measure used. This is done through a detailed analysis of the research performances of a reasonably homogeneous population of researchers, namely all full economics professors in one country, Sweden.

We document how the ranking and distribution of research performances are influenced by the use of different measures of research output. Specifically, we study seven commonly used measures that all capture essential aspects of research performance. Three of these (KMS, Impact Factor (IF), and a measure by Kodrzycki and Yu 2006 (KY)) are journal weight-based measures, i.e., specific citation-based journal rankings that when multiplied with an individual scholar's journal article production produce a score. Two of the measures are direct citation counts, defined as the number of citations of the individual researcher's five most cited works in two different citation databases (Social Sciences Citation Index (SSCI) and Google Scholar (GS)). The sixth measure is the so-called Hirsch- or *h*-index, which indicates the number of pa-

---

[2] There is research showing that famous scholars receive more citations (Coupé 2003), while articles settling academic debates or that provide important robustness checks tend to receive almost no citations at all (Mayer 2004; van Dalen and Klamer 2005).

[3] As Oswald (2007) shows, even in the top-ranked journals there are several articles that receive no or very few citations.

pers of a researcher that obtains at least an equal number of citations (in GS). The seventh measure is a "raw" output measure: the number of works in *EconLit*.

Our empirical analysis is based on a newly constructed database containing all documented international research publications by all currently active full professors in economics tenured at Swedish universities, as of winter of 2007.[4] The population consists of 93 professors, 87 men and 6 women, born between 1939 and 1968. Since we are studying professors in Sweden and not Swedish professors, there are foreign-born scholars in our sample and there are Swedes tenured at foreign universities who are not in our sample.[5]

Each of the seven measures analysed captures at least some relevant aspect of a researcher's output. A key finding is that the distribution of total research output is valued very differently across measures, and there is also large variation in the rank order of professors across measures.

1 Measures of research output

This section presents seven measures of research output that we use to assess the research performance of Swedish economics professors and how the distribution of performance across professors varies depending on the measure used.[6] The measures are:

Measures based on weighted journal publications:

1. Sum of KMS weighted journal articles (Kalaitzidakis et al. 2003)

2. Sum of IF weighted journal articles (Thomson Scientific 2003)

3. Sum of KY weighted journal articles (Kodrzycki and Yu 2006)

Measures based on citations to most cited works:

4. Sum of citations of five most cited works in the SSCI

---

[4] As professors in economics we include chaired and non-chaired professors (*befordringsprofessorer*) at economics departments at Swedish universities and colleges, as well as "professors in economics" active at other university departments (e.g., industrial dynamics, economics and management). A list of all professors including affiliation, birth year, year of Ph. D. and year of promotion to full professor (for the first time at a Swedish university) is available upon request.

[5] Casual observation suggests Swedes abroad tend to be above-average successful Swedish economists and their inclusion would thus increase the measured skewness in the distribution of individual performances.

[6] For earlier preliminary rankings and valuations of Swedish academic economists, see Henrekson, Magnusson and Waldenström (2006) and Sarafoglou (2006).

5. Sum of citations of the five most cited works in Google Scholar (GS)

6. Individual *h*-index (Hirsch 2005; Harzing 2007)

<u>Measures based on the number of international publications:</u>

7. Number of published works (articles, book chapters, books) in *EconLit*

These seven measures capture most of the relevant dimensions along which we normally think of when quantifying the volume and quality of individual research performance. In the first three measures, only journal articles are counted and the measures differ in the way they give relative merit to the journals in which these articles appear.

In the fourth and fifth measures, actual citations of the researcher's most influential journal articles (SSCI) or works in general (GS) are counted. Rather than summing all citations to a researcher's works or just the top citation we sum the citations to the five most cited works. Obtaining the sum for all works is tricky, in particular if the researcher has a common name, and just taking the most cited work may give undue weight to a one-time success article. This tendency is exacerbated by the so-called "halo effect", i.e., that very successful publications tend to be cited in excess of their scientific merit (Ursprung and Zimmer 2007). The halo effect is totally avoided in our sixth measure (individual *h*-index), which indicates the number of papers of a researcher that obtains at least an equal number of citations in GS. Lastly, the seventh measure includes all kinds of internationally published output in a researcher's performance, given that the publication is listed in *EconLit*.[7]

We have adjusted for co-authorship by weighting down publications with *n* co-authors by 1/*n*, which is in line with most previous work in this type of literature.[8]

---

[7] No doubt, there are a number of other measures used, such as the Laband and Piette (1994) measure which is still widely used, particularly in the U.S. Coupé (2003) uses several additional measures. However, the inclusion of further measures would not strengthen the general point made in this paper.

[8] For robustness purposes, we also weighted the co-authored publications by the square root of *n* and the results across measures are basically identical (although with slightly higher concentration of performances than when the weight 1/*n* is used). Further support for the view that the method for dealing with co-authorship is of marginal importance is provided by Coupé (2003). He refers to evidence suggesting that tenure decisions are only marginally influenced by the presence of co-authors.

1.1 Kalaitzidakis, Mamuneas and Stengos (KMS)

One of the most widely cited rankings of economics journals during recent years is that of Kalaitzidakis et al. (2003). This ranking received enormous attention, not least among European economists, after it was described by three recent Presidents of the European Economic Association (EEA) – John Mirlees, Peter Neary and Jean Tirole – as being "the most up-to-date set of objective journal weights available" (Neary et al. 2003, p. 1247).[9] These weights, which we call KMS (after the authors behind them), assign relative merits to each of the 159 journals in SSCI's "economics" category in the following way. First the total number of citations (as published in the 1998 issue of *Journal Citation Reports*, *JCR*) are collected for the ten years before 1998. Then the authors exclude within-journal citations ("self"-citations) and "older" citations, here defined as articles published before 1994 (i.e., older than four years). In order to remove any influence of journal size on the number of citations received, a weight relating each journal's annual number of pages with the average number of pages of all journals is computed. Finally, and most importantly, citations are weighted according to their "impact", meaning that citations from relatively well-cited journals are given more weight than citations coming from less cited journals. Altogether, KMS is a self-citation-, age-, size- and "impact"-adjusted journal weight based on actual citations at the end of the 1990s.

Substantial critique has been directed towards KMS and its sources. In particular, the iteration procedure used by the authors to account for the "impact" (of citing journals) allegedly boosts the weights primarily at the absolute top of the journal distribution, giving rise to extreme differences between journals. For example, a single article in the *American Economic Review* is valued more highly in KMS than ten articles in the *Journal of Financial Economics* or 25 articles in the *Journal of Law and Economics*, 60 articles in the *Journal of Health Economics*, or all the 400+ articles published in the *Journal of Evolutionary Economics* since its first issue in 1991.

---

[9] In 2000, the EEA invited bids for constructing journal and scholar rankings and five contributions were selected and published in a special issue of the *JEEA* in 2003. Based on these five contributions the current and past presidents and the president-elect of the EEA chose to emphasise KMS, although they also made clear that "there is no single way to carry out studies like these" (p. 1240). For further information, see the EEA website: http://www.eeassoc.org/default.asp?AId=37 (Oct. 23, 2007).

Another critique of KMS concerns its adjustment for journal size (number of pages), where Palacios-Huerta and Volij (2004) argue that the relevant level of analysis of scholarly work is articles, and that a more appropriate size measure would therefore be the number of articles published in a journal per year. Furthermore, KMS focuses only on the *JCR* "economics" category which implies focusing on a quite limited subset of journals. Among other things, this leads to the exclusion of most finance journals, where economists regularly publish even some of their best work. Perhaps even more importantly, the "impact" weights used in KMS are based on the premise that both cit*ing* and cit*ed* journals are included in the SSCI database, which precludes the scientific impact across SSCI's narrowly defined academic fields.

We compute individual "KMS points" for all Swedish economics professors by multiplying their journal articles by their respective KMS weights. Data on individual publications are collected from *EconLit* and sometimes from departmental and personal websites. It should be noted that *EconLit* covers roughly 1,000 journals against the 159 covered by KMS, which means that a large number of articles, in economics as well as in neighboring academic disciplines, receive zero KMS points. *Figure 1* displays the distribution of research performance of Swedish economics professors when their performance is measured using KMS scores. On the *x*-axis all professors are ranked from the most productive to the far left (having a KMS score of 961.2) to the least productive to the far right (with a KMS score of zero). A more detailed analysis follows in section 3, and we settle for now by noting the remarkable skewness of the distribution of performances across the scholars, with a sizeable proportion of them having a negligible or zero KMS score. This is striking since, after all, we restrict our analysis to full professors with on average 23 years as graduated researchers in academia.

[Figure 1: KMS scores]

1.2   Impact Factor (IF)

Even if the KMS measure is likely to be the most influential journal ranking in European economics, there is little doubt that the most widely used journal impact metric across all fields of science is the so-called *Impact Factor* (IF). IF is calculated by Thomson Scientific (which also runs SSCI) and is reported in *JCR*. It is defined as the

current year's citations to a journal's articles by articles in other journals during the two preceding years divided by the total number of articles published in the target journal during these two years.[10] We use IF weights for journals in the *JCR* "economics" category, which means that we restrict the number of weighted journals to 169. Unlike KMS, IF does not exclude within-journal citations and it weights all citations to journals equally while, as already stated, KMS attributes less weight to less-cited journals.

The effect of these differences seems to be sizeable. For example, the IF score for the *Journal of Health Economics* (1.778) is only marginally lower than the IF score of the *AER* (1.938), while the *AER* vastly outmeasures the *Journal of Health Economics* in KMS by a factor of 60. Generally, even for the journal ranked 150 according to IF fewer than 10 articles are required to obtain an IF score on par with one *AER*-article.

We calculate the IF score for Swedish economics professors in exactly the same way as we did for KMS, i.e., by multiplying their (co-authorship weighted) journal articles by the respective IF weights. *Figure 2* shows the distribution of IF scores across the population of Swedish economics professors, and as a point of reference we include the cumulative frequencies of both the IF and KMS scores. As is immediately apparent from the chart, the IF weighting scheme results in a much less skewed distribution.

[Figure 2: IF scores]

1.3   Kodrzycki and Yu (KY)

As already mentioned, the KMS measure is only applied to the journals in the SSCI "economics" category and in the calculation of journal impact only citations from these very same journals are used. Kodrzycki and Yu (2006) argue that both these restrictions are problematic to a social science discipline such as economics given that economists regularly interact in different ways with neighboring disciplines such as finance, law, political science, medicine, criminology, psychology and sociology (to

---

[10] For example, the IF of the *American Economic Review* in 2005 was calculated as all citations during 2005 of *AER* articles published in 2003 and 2004 divided by the total number of *AER* articles during 2003 and 2004 (resulting in an IF of 1.806).

7

mention only a few). As we noted above, the SSCI "economics" category also excludes some of the journals that economists actually cite the most, including the *Industrial and Labor Relations Review*, the *Journal of Finance* and the *Review of Financial Studies*. To mitigate some of this narrowness, Kodrzycki and Yu propose alternative weighting schemes in which they extend the number of cit*ed* journals (to 181) but, more importantly, extend the scope of cit*ing* journals to include all social science journals in SSCI.[11] As in the case of IF, the KY measures draw their citation information from the *JCR* issue of 2003. But unlike the short window of the past two years' citations counted in IF, KY use citations of journal articles published between 1996 and 2003. Recall that KMS used citations during 1994–1998 reported in the *JCR*-issue of 1998. Hence, to the extent that citation patterns change over time, which they do as persuasively shown by Kim et al. (2006), the degree of overlap between these set of weights is reduced. *Figure 3* shows the ranking of Swedish professors when KY weights are used. The KY performance distribution is fairly similar to the KMS distribution, which is interesting given the differences in samples of cited as well as citing journals.

[Figure 3: KY scores]

1.4 Some comparisons between the three journal weight measures

In order to get a sense of how the journal rankings in KMS, IF and KY correspond to each other, we list correlation coefficients and the number of overlapping journals for the weights in *Table 1*. The table makes clear that KMS stands out from the other measures. IF and the KY measure are closer to each other than they are to KMS, both in terms of correlations and in the number of journal overlaps.

[Table 1: Journal weight correlations]

It is also important to carefully assess the distributions of weights across the three measures since this has a first-order impact on how the actual research performances

---

[11] Kodrzycki and Yu (2006) also present weights when only citations from "economics" and "policy" oriented journals are counted. We have analysed the former and its results are almost identical to the one using all social science journals (available upon request). Furthermore, Kodrzycki and Yu suggest that a per-article (as opposed to per-page or per-character) weighting of journal size is the most relevant to make. This is in line with the reasoning of Palacios-Huerta and Volij (2004).

of scholars are distributed. *Table 2* presents distributional statistics for the three journal weights, and once again KMS stands out as being the most unevenly distributed of the journal weights analysed here (see also *Figures B1* and *B2* in Appendix B).[12] For example, the coefficient of variation of KMS is twice as large as the one of IF, but only slightly larger than that of KY. KMS values the 90th percentile journal more than 100 times higher than the 25th percentile journal, whereas KY values it 32–33 times higher and IF only 4.4 times higher.[13] The valuation of the 90th percentile journal relative to the median journal (the P90/P50-ratio) is three times larger for KMS than for KY and ten times larger than for IF. Focusing on the share of journal weights held by different population fractiles, the top 5 percent journals (P95–P100) holds roughly the same share in KMS and KY. This share is about double compared to IF. More strikingly, the lower half of the journals receives only 2.2 percent of the sum all weights in KMS, whereas it has twice as much in KY (4.5) and almost ten times as much in IF (20.6). Finally, the Gini coefficient of KMS is the highest of all measures.

[Table 2: Journal weight concentration outcomes]

1.5   Social Sciences Citation Index (SSCI)

A common critique leveled against measures like KMS and IF (and KY) is that they assign the same value to all articles appearing in a journal based on the adjusted total number of citations to that journal. By doing this, they disregard the fact that articles in a journal may have quite different impact. As shown by Oswald (2007), the number of citations that the eighteen articles in one *AER* issue in 1981 received over the following 25 years ranged from 401 to 0. In other words, the publication of an article in the *AER* (or any other top-ranked journal) does not guarantee that the article will be widely cited.[14]

---

[12] From *Figure B1* one can infer that the ten most highly ranked journals in KMS are attributed 50 percent of the entire journal score. The corresponding figure for IF is only 25 percent. It is also clear from the figure that the KMS and KY measure both give extremely high weights to the top-5 journals. The reader should also keep in mind that roughly 85 percent of all journals in *EconLit* are given zero weight by all four measures. Hence, roughly 1 percent of the *EconLit* journals are attributed 50 percent of the KMS score. *Figure B2* shows the corresponding picture, but in the form of Lorenz curves.

[13] We would have wished to compare also with the 10th percentile journal, but KMS gives it a zero weight and hence its P90/P10-ratio equals infinity.

[14] Laband and Tollison (2003) report that in their citation count over the subsequent five years of all articles published in 1996 in 91 journals 70 percent of the articles received one cite or less.

We include two measures that account for actual – rather than assumed – citations to the scholars' works. The first such measure is the sum of citations to the five most cited works of each professor as recorded in the SSCI. We choose to sum the number of the five most cited works as we deem this to strike a reasonable balance between counting all citations to all works (which may give rise to spurious results at the lower end) or only to the single most cited work (which would give too little credit to scholars with a number of well-cited works).

The SSCI citation database is probably the world's largest citation database and a widely used source for assessing the impact of individual researchers (see, e.g., Klein and Chiang 2004). But there are some important caveats with the SSCI citations as well. Only citations from journals in the SSCI (a minority of all existing journals) are recorded. While SSCI does not adjust for author self-citations, we have done this afterwards by help of our new database of professors. Moreover, until the mid-1990s only first authors received citations, which benefited people having names beginning with a letter early in the alphabet. We have also discovered several examples of misspellings of names that would obviously risk discriminating Swedes with last names containing non-Anglo Saxon letters such as "å", "ä" and "ö" (they are 13 percent in our sample).[15] *Figure 4* displays the distribution of researcher performance measured as citations in the SSCI. While skewed, it is still considerably more evenly distributed than the KMS measure.

[Figure 4: SSCI citations]

1.6 Google Scholar (GS)

Our second measure of actual citations is drawn from a less commonly used data source: the Internet database Google Scholar (GS).[16] GS and SSCI differ in several important respects. First, GS records citations coming from a much larger pool of publication types, including working papers, reports and books and academic journals

---

[15] This is a potential source of error, but we have done as much as possible to avoid this problem. Given our familiarity with most of the people in the sample, and the fact that we only include the five most cited articles, we judge that the problem is negligible.

[16] http://scholar.google.com.

that are available on the Internet.[17] Second, whereas SSCI only includes articles in SSCI journals among their cited works, GS allows any of its recorded publication items to be among an author's cited works. That this difference may be of importance is shown by the fact that of the ten most cited works in GS in our sample of Swedish professors, two were textbooks, one a monograph and one a book chapter, together representing 41 percent of all citations in the top-10 group. Equally interesting, of the six journal articles in the top-10, only two were also among the respective author's five most cited articles in SSCI. This is a strong indication of the importance of citations, and hence "outside" scientific impact, outside the SSCI realm. Third, GS claims to exclude self-citations although we have found examples where this is not true. For clarity reasons, we have thoroughly analysed each professor's citations in the GS database, notably ascertaining that we have the correct author and that there are five distinct cited works. *Figure 5* shows the distribution of research output when counted as the sum of citations of the five most cited works in the GS. Interestingly, the share of the top scholars is even greater than for the KMS measure, but it is also noteworthy that the share of the lower tail for GS is greater than for KMS.

[Figure 5: GS citations]

1.7   The individual *h*-index

The *h*-index suggested by Hirsh (2005) is an index that attempts to quantify the scientific productivity and impact of a scientist based on his/her most quoted papers. A scientist with an *h*-index of *x* has published *x* papers that have at least *x* citations. The *h*-index differs from the previous two citation measures in that it de-emphasises a few successful publications in favour of sustained productivity.[18]

The original *h*-index does not account for the number of co-authors in a paper. We therefore implement Harzing's (2007) alternative *individual h*-index that first normalises the number of citations for each paper by dividing the number of citations by the number of authors for that paper and then calculates the *h*-index of the normalised ci-

---

[17] According to Jacsó (2005), however, GS has problems covering publications from some of the large international publishing houses.
[18] It may do so too strongly. Two scientists may have the same *h*-index, say, *h* = 25, but one has 20 papers that have been cited more than 500 times and the other has none. Clearly, the output of the former is more valuable.

tation counts.[19] The individual *h*-index is based on GS citations using the software *Publish and Perish* by Harzing (2007).[20] As shown by *Figure 6*, the *h*-index distributes research performances across the professor population much more evenly than the other citation-based measures. While this was expected as the measure caps top citations, the order of magnitude of this difference is noteworthy.

[Figure 6: *h*-index]

1.8 Works in *EconLit*

Our final measure of research performance is a pure quantity measure: the number of publications (adjusted for co-authorship) listed in *EconLit*. While this measure includes most internationally published works such as journal articles, book chapters and monographs, we have removed working papers and reprints of already published works. Of course, a raw output measure does not capture several of the most desired dimensions of scholarly work such as adjustment for the quality of a publication. Nevertheless, this measure provides a useful point of reference. From *Figure 7* it is immediately clear that when only counting raw output the distribution of research performances is markedly more equal.

[Figure 7: Works in EconLit]

2 Comparing performances of scholars across measures

This section presents the second part of the empirical analysis. First, we compare the concentration of the distribution of performances in the seven measures. Then we examine the degree of overlap in the rankings of the measures. Finally we econometrically analyse how various background factors are associated with individual performance.

2.1 Distributional outcomes

Analysing the shape of the performance distribution among scholars is important. Given the fact that quantitative measures often underlie hiring and funding decisions,

---

[19] This approach more accurately accounts for co-authorship effects, and hence better approximates the per-author impact than the alternative individualised *h*-index of Batista et al. (2006) which merely divide the original *h*-index by dividing the mean number of researchers in the *h* publications.
[20] The calculations were made on September 15, 2007, using *Publish or Perish*, version 2.3.

the relative merit awarded to the top compared to the bottom may influence individual differences in, e.g., salaries or the size of research grants. Some degree of discrepancy in the skewness of performances across the measures might be acceptable since each measure focuses on slightly different aspects of scientific activity and impact.

In *Table 3*, we present several statistical metrics of concentration of research performance computed for the population of Swedish economics professors. The results indicate large differences between the analysed measures. Some of them give rise to distributions that are particularly skewed towards the top, while others turn out to be more tilted towards the middle. Most notably, the KMS weights stand out as generating the most skewed distribution, although the cite count in Google Scholar is in some respects as skewed as the KMS score. Specifically, the P90/P10- and P90/P25-ratios are by far the highest for KMS, but the P90/P50-ratio is the highest for GS; GS assigns almost no value to the research output of the median professor.

Looking at the share of total performance attributable to the four most productive professors (roughly the top 5 percent), this amounts almost to a quarter in KMS and almost one third in GS. The lower half of the population based on KMS represents only 7.4 percent of total performance, whereas the share is somewhat larger for KY (9.4 percent) and about double for IF and the two citation measures SSCI and GS, three times larger in Works (25.8 percent) and almost five times larger in the *h*-index (35.4 percent). The Gini coefficients indicate a similar pattern, with KMS producing the most skewed performance distribution and the *h*-index by far the most even distribution.

[Table 3: Concentration outcomes]

2.2 Degree of overlap between the rankings

Some degree of discrepancy in the skewness of the distributions of performances across the measures is both natural and desirable, since each measure focuses on slightly different aspects of scientific activity and impact. But it would still be a desirable property of a measure that it produces a rank order of professors that is roughly similar to the other measures. In particular, we would like the measures based on journal weights (KMS, IF, KY) to rank professors in roughly the same way as the

measures based on citation counts (SSCI, GS, *h*-index) do, since these two types of measures both capture important aspects of performance. This section examines whether this is in fact the case.

We present two sets of analyses. First, we calculate pairwise correlations using the entire population of professors using both Pearson correlations (*Table 4a*), which incorporate the absolute distances in performance, and Spearman rank correlations (*Table 4b*), which only captures the degree of rank order similarity. The results suggest a high degree of overlap between the measures. However, the overlap is substantially larger *within* the different types of measures (i.e., between the journal weight-measures KMS, KY and IF, on the one hand and the two citation measures, SSCI, GS and *h*-index, on the other hand) than *between* the two types of measures. Yet, KY and IF appear to be more correlated with SSCI, GS and the *h*-index than KMS. It is also noteworthy that the correlation of raw output, Works, with all the other measures is substantially smaller.

[Table 4a & b: Correlation of performances across measures]

Our second assessment focuses on the overlap in the rankings of the top-10 professors across the seven measures. While this concerns a much smaller group of scholars, these scholars represent between one quarter and more than half of the aggregate performance in the different measures. *Figure 8* displays the ranks in all measures for the ten professors that are top ranked according to KMS. The numbers are also provided in *Table 5*. There appears to be a fairly high degree of overlap between the journal weight measures (KMS, IF, KY), but considerably less so between KMS and the citation measures (SSCI, GS and *h*-index), as well as with the unweighted output measure (Works). For example, among the top ten scholars based on KMS between four and six are not ranked among the top-20 according to the citation measures. Strikingly, one of the top-10 based on KMS even ranks near the bottom of the distribution (as number 78 and 82) in terms of citations in SSCI and GS.

[Table 5: KMS top-10 in the different measures]
[Figure 8: KMS top-10 in the different measures]

2.3 What determines success?

In the above analysis we found that in terms of both the skewness research performances and the composition of individuals in the top of the rankings differed markedly between the seven measures. It is not clear, however, whether this considerable variation imply that there is no core set of individual characteristics that determine successful research performance or if the fact that the also measures themselves capture somewhat different aspects of scholarly impact still allows for such characteristics to play a consistent role. In order to address this issue, we turn to linear regression analysis. We use the log of the research output measures as the dependent variable, which is "explained" by a set of background variables drawn from our database of full professors. The variables are: (i) sex (for which we have no prior regarding the effect on research performance); (ii) affiliation at an established research university (expected positive effect);[21] (iii) age at graduation to Ph. D. degree (no prior); (iv) number of years from receiving the Ph. D. to being promoted to full professor (expected negative effect as it can be an indicator of research skill) and (v) the number of years as professor up until 2007 (expected positive effect).[22] The regression equation then looks as follows:[23]

$$\ln[1 + Research\ performance]_i = \alpha + \beta_1 \cdot Sex_i + \beta_2 \cdot Research\ university_i + \beta_3 \cdot Age\ at\ Ph.D._i + \beta_4 \cdot Time\ to\ professorship_i + \beta_5 \cdot Years\ as\ professor_i + u_i$$

Admittedly, this model provides a rather simple framework for analysing the determinants of research performance, since our dataset only allows cross-sectional analyses. The results should therefore be interpreted as conditional correlations rather than causal effects.[24] Reassuringly, the estimation results appear to be highly robust. As shown by the alternative regressions reported in *Table A1* and *A2* (Appendix A), varying the definition of the dependent variable (using logs without adding a one; not us-

---

[21] We regard the universities in Gothenburg, Lund, Stockholm, Umeå and Uppsala and the Stockholm School of Economics as established research universities in economics.

[22] For a similar analysis of tenure success of Swedish economists, see Tasiran et al. (1997).

[23] We add a one to the measures of research performance in order to be able to include also those professors whose scores are zero, which are otherwise not defined in log form. Appendix A shows the results when using a dependent variable a) in log without the added one and b) not in log form.

[24] In particular, unobservable variables are likely to be correlated with some of the included variables (e.g., being at a research university may be related to a number of personal characteristics that drive research performance) and causality could also be bi-directional in the case of affiliation at a research university.

ing logs; scaling by years as professor or years as Ph. D.) does not change the overall findings, and neither does the inclusion of additional independent variables (e.g., time as professor squared) or the exclusion of some of the independent variables.

Table 6 presents the main regression results and several interesting findings emerge. First, KMS and the *h*-index are the only measures where the sex variable is statistically significant, indicating better performance for male professors. In terms of the coefficient estimate, KMS stands out from all the other measures in being associated with large differences between the sexes in terms of research performance. Second, affiliation to a research university is associated with a significantly higher performance for almost all measures (for GS, *h*-index and Works the point estimates are statistically insignificant). Third, becoming a Ph. D. at a relatively young age is associated with a better performance in the subsequent career, which may reflect that more talented people require less time to become "licensed" researchers, or that it is in fact an advantage to become a Ph. D at an earlier age. The relationship seems to be somewhat stronger for the measures based on journal-weighted articles (KMS, IF, KY) than for the top citation counts (SSCI, GS), and especially compared to the raw output measure (Works). Fourth, the time needed to become professor is negatively associated with performance in all measures, although with some variation in the level of significance. Fifth, and finally, the number of years as professor up until 2007 has no significant impact on relative research performance, except in terms of number of works.

[Table 6: Regression results]

Overall, the measures provide many similar results with respect to the determinants of research performance, but there are also important differences between them. There are two measures that deviate from the others. In particular, the raw output measure, Works, is quite dissimilar and this deviation most likely reflects the fundamental distinction between assessing quantity and quality in research. The other measure that deviates in most respects is KMS. This is the only measure that is associated with a significant male-female difference, it greatly boosts the relative output of the research elites at established universities while failing to show the importance of the time re-

quired to advance to full professor, a measure which turned out to be particularly important for the metrics of scientific impact based on actual citations.

3   Concluding discussion

Increased specialisation and international integration of research and researchers have sharply raised the need for comparisons of performance across fields, institutions and individual researchers. However, there is still no consensus regarding how such rankings should be conducted and what output measures are appropriate to use in order for research funds to be efficiently allocated.

In this article we have analysed seven established and some of them commonly used measures of research performance. Based on these measures we have ranked all full professors in economics tenured at Swedish universities in early 2007. Each of the studied measures has its *pros* and *cons* and in most cases it is not obvious which measure is the most appropriate to use. Our examination shows both that the rank order varies greatly across measures, and that depending on the measure used the distribution of total research output is valued very differently.

One of our strongest results is that the renowned KMS measure stands out among all the measures analysed here. It gives rise to a particularly skewed distribution of performances among professors, where the professors at the very top are attributed a very large share of the total output, while the absolute contribution of the lower half of the population is negligible. It is also noteworthy that it correlates less strongly with our two citation measures than does IF and KY, and that it appears to disfavour women relative to the other measures.

Our results make clear that there is no single unequivocal catch-all measure that can be used. All seven measures provide relevant information about the performance of individual researchers, and no doubt there are additional aspects that may be important that are largely overlooked by all of these measures. For instance, only a small subset of all journals are included in the KMS, IF and KY measures, and most measures either ignore or give little weight to impact outside economics or on policymaking. Hence, while quantitative measures are essential for assessing research they can-

not fully substitute for careful reading and individual assessment of the works of individual researchers.[25]

In recent years, researchers have become increasingly aware of what is measured, and there is a strong tendency to do what is measured (Holmström and Milgrom 1991; Frey and Osterloh 2006). This tendency is reinforced if universities, departments and research councils use a certain metric when making decisions about hiring, promotion, and the allocation of funds (Holcombe 2004, Oswald 2007).

Relative ranking and quality-adjusted quantification of research output is no temporary fad. Instead it is likely to continue to gain in importance. As soon as a certain measure is widely used, researchers can be expected to adjust behaviour in order to maximise their output as defined by this measure. Therefore, the choice of measures is of great importance unless it turns out that ranking and the relative valuation of different researchers and departments are largely invariant with respect to an array of output measures. The evidence presented in this study speaks strongly against any presumption of that sort.

References

Axarloglou, Kostas and Vasilis Theoharakis (2003), "Diversity in Economics: An Analysis of Journal Quality Perceptions." *Journal of the European Economic Association* 1(6), 1402–1423

Batista, Pablo D., Mônica G. Campiteli, Osame Konouchi and Alexandre S. Martinez (2006), "Is It Possible to Compare Researchers with Different Scientific Interests?" *Scientometrics* 68(1), 179–189.

Cameron, Brian D. (2005), "Trends in the Usage of ISI Bibliometric Data: Uses, Abuses, and Implications." *Libraries and the Academy* 5(1), 105–125.

Combes, Pierre-Philippe and Laurent Linnemer (2003), "Where Are the Economists Who Publish?" *Journal of the European Economic Association* 1(6), 1250–1308.

Conroy, Michael E. and Richard Dusansky (1995), "The Productivity of Economic Departments in the US: Publications in the Core Journals." *Journal of Economic Literature* 33(4), 1966–1971.

Coupé, Tom (2003), "Revealed Performances: Worldwide Rankings of Economists and Economics Departments, 1990–2000." *Journal of the European Economic Association* 1(6), 1309–1345.

---

[25] Van Fleet et al. (2000) come to the same conclusions after having examined the use of journal rankings as explicit targets for researchers in management departments.

Dusansky, Richard and Clayton J. Vernon (1998), "Rankings of U.S. Economics Departments." *Journal of Economic Perspectives* 12(1), 157–170.

Frey, Bruno S. and Margit Osterloh (2006), Evaluations: Hidden Costs, Questionable Benefits, and Superior Alternatives." Working Paper, Institute for Empirical Research in Economics, University of Zürich.

Harzing, Anne-Wil (2007), "Publish or Perish." Online: http://www.harzing.com (October 4, 2007).

Hecht Fredrick, Barbara K. Hecht and Avery A. Sandberg (1998), "The Journal 'Impact Factor': A Misnamed, Misleading, Misused Measure." *Cancer Genet Cytogenet* 104(2), 77–81.

Henrekson, Magnus, Kristin Magnusson and Daniel Waldenström (2006), "Hur bör forskningsprestationer mätas?" *Ekonomisk Debatt* 34(3), 62–75.

Hirsch, Jorge E. (2005), "An Index to Quantify an Individual's Scientific Research Output." *Proccedings of the National Academy of Sciences* 102(46),16569–16572

Holcombe, Randall G. (2004), "The National Research Council Ranking of Research Universities: Its Impact on Research in Economics." *Econ Journal Watch* 1(3), 498–514.

Holmström, Bengt and Paul M. Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7(1), 24–52.

Jacsó, Peter (2005), "Google Scholar: the Pros and Cons." *Online Information Review* 29, 268–214.

Kalaitzidakis, Pantelis, Theofanis P. Mamuneas and Thanasis Stengos (1999), "European Economics: An Analysis Based on Publications in the Core Journals." *European Economic Review* 43(4–6), 1150–1168.

Kalaitzidakis, Pantelis, Theofanis P. Mamuneas and Thanasis Stengos (2003), "Rankings of Academic Journals and Institutions in Economics." *Journal of the European Economic Association* 1(6), 1346–1366.

Klein Dan B. and Eric Chiang (2004), "Citation Counts and SSCI in Personnel Decisions: A Survey of Economics Departments." *Econ Journal Watch* 1(1) 166–174.

Kodrzycki, Yolanda K. and Pingkang David Yu (2006), "New Approaches to Ranking Economics Journals." *Contributions to Economic Analysis & Policy* 5(1), article 24.

Kim, E. Han, Adair Morse and Luigi Zingales (2006), "What Has Mattered to Economics Since 1970?" *Journal of Economic Perspectives* 20(4), 189–202.

Laband, David N. and Michael J. Piette (1994), "The Relative Impacts of Economics Journals 1970–1990." *Journal of Economic Literature* 32(2), 640–666.

Laband, David N. and Robert D. Tollison (2003), "Dry Holes in Economic Research." *Kyklos* 56(2), 161–174.

Lubrano, Michel, Luc Baluwens, Alan Kirman and Camelia Protopopescu (2003), "Ranking Economics Departments in Europe: A Statistical Approach." *Journal of the European Economic Association* 1(6), 1367–1401.

Mayer, Tomas (2004), "Dry Holes in Economic Research: Comment." *Kyklos* 57(4), 621–626.

Neary, J. Peter, James A. Mirrlees and Jean Tirole (2003), "Evaluating Economics Research in Europe: An Introduction." *Journal of the European Economic Association* 1(6), 1239–1249.

Oswald, Andrew J. (2006), "Prestige Labels." *Royal Economic Society Newsletter*, issue No. 135, October.

Oswald, Andrew J. (2007), "An Examination of the Reliability of Prestigious Scholarly Journals: Evidence and Implications for Decision-Makers." *Economica* 74(293), 21–31.

Palacios-Huerta, Ignacio and Oscar Volij (2004), "The Measurement of Intellectual Influence." *Econometrica* 72(3), 963–977.

Sarafoglou, Nikias (2006), "Hur mäta produktivitet och hur produktiva är svenska professorer?" *Ekonomiska Samfundets Tidskrift* 59(2), 95–111.

Tasiran, Ali C., Ann Veiderpass and Bo Sandelin (1997), "Climbing Career Steps: Becoming a Full Professor of Economics." *Scandinavian Journal of Economics* 99(3), 471–484.

Thomson Scientific (2003), *Journal Citation Reports 2003. Social Sciences Edition,* Institute for Scientific Information, Philadelphia.

Tombazos, Christis G. (2005), "A Revisionist Perspective of European Research in Economics." *European Economic Review* 49(2), 251–277.

Ursprung, Heinrich, W. and Markus Zimmer (2007), "Who is the 'Platz-Hirsch' of the German Economics Profession." *Journal of Economics and Statistics* 227(2), 187–208.

Van Dalen, Hendrik P. and Arjo Klamer (2005), "Is Science A Case of Wasteful Competition?" *Kyklos* 58(3), 395–414.

Van Fleet, David D., Abagail McWilliams and Donald S. Siegel (2000), "A Theoretical and Empirical Analysis of Journal Rankings: The Case of Formal Lists." *Journal of Management* 26(5), 839–861.

Appendices

Appendix A: Robustness tables

Table A1: Regression results without adding 1 to the logged dependent variable

	KMS	IF	KY	SSCI	GS	*h*-index	Works
Sex (Male = 1)	1.76*	0.33	0.98	0.56	0.68	0.67**	0.31
	(0.76)	(0.51)	(0.67)	(0.57)	(0.47)	(0.20)	(0.17)
Research university	1.58**	0.69**	1.22**	0.63*	0.31	–0.04	0.07
	(0.36)	(0.24)	(0.31)	(0.30)	(0.21)	(0.14)	(0.16)
Age at Ph. D.	–0.10**	–0.07**	–0.11**	–0.01	–0.05	–0.04*	–0.02
	(0.04)	(0.02)	(0.03)	(0.04)	(0.03)	(0.02)	(0.02)
Time to professorship	–0.03	–0.05*	–0.04	–0.07**	–0.06**	–0.05**	–0.02
	(0.03)	(0.02)	(0.03)	(0.03)	(0.02)	(0.01)	(0.01)
Years as professor	0.02	0.01	0.02	0.01	0.00	-0.00	0.03**
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Constant	4.21*	3.05**	5.21**	3.27*	5.97**	3.35**	2.70**
	(1.91)	(1.12)	(1.58)	(1.56)	(1.32)	(0.65)	(0.81)
Observations	90	90	91	92	93	93	93
R–squared	0.41	0.34	0.43	0.20	0.23	0.41	0.30

*Note*: Robust standard errors in parentheses. * significant at 5%; ** significant at 1%.

Table A2: Regression results when the dependent variable is not logged

	KMS	IF	KY	SSCI	GS	*h*-index	Works
Sex (Male = 1)	59.16*	1.70	27.94	10.43	55.32	2.94**	4.10*
	(24.21)	(1.61)	(15.54)	(14.80)	(39.22)	(1.10)	(1.99)
Research university	73.22**	2.48**	35.65**	12.40	36.57	–0.20	1.20
	(19.52)	(0.84)	(10.92)	(11.24)	(29.16)	(0.82)	(1.93)
Age at Ph. D.	–8.27*	–0.44*	–5.57*	–2.33	–14.24*	–0.39**	–0.30
	(3.97)	(0.17)	(2.34)	(1.65)	(6.01)	(0.12)	(0.29)
Time to professorship	–1.74	–0.19*	–2.20	–2.00*	–7.61*	–0.30**	–0.27
	(2.16)	(0.08)	(1.28)	(0.85)	(3.42)	(0.07)	(0.16)
Years as professor	3.16	0.10	1.37	0.59	2.10	0.03	0.61**
	(1.78)	(0.08)	(1.02)	(0.72)	(3.21)	(0.06)	(0.16)
Constant	248.45	17.74*	200.81*	120.27	604.72*	21.10**	17.67
	(165.16)	(6.83)	(97.51)	(70.89)	(242.69)	(4.92)	(11.06)
Observations	93	93	93	93	93	93	93
R-squared	0.23	0.27	0.24	0.16	0.13	0.30	0.31

Note: Robust standard errors in parentheses. * significant at 5%; ** significant at 1%.

Appendix B: Two further illustrations of the distribution of weights

Figure B1: Cumulative shares of journal weights in the three measures
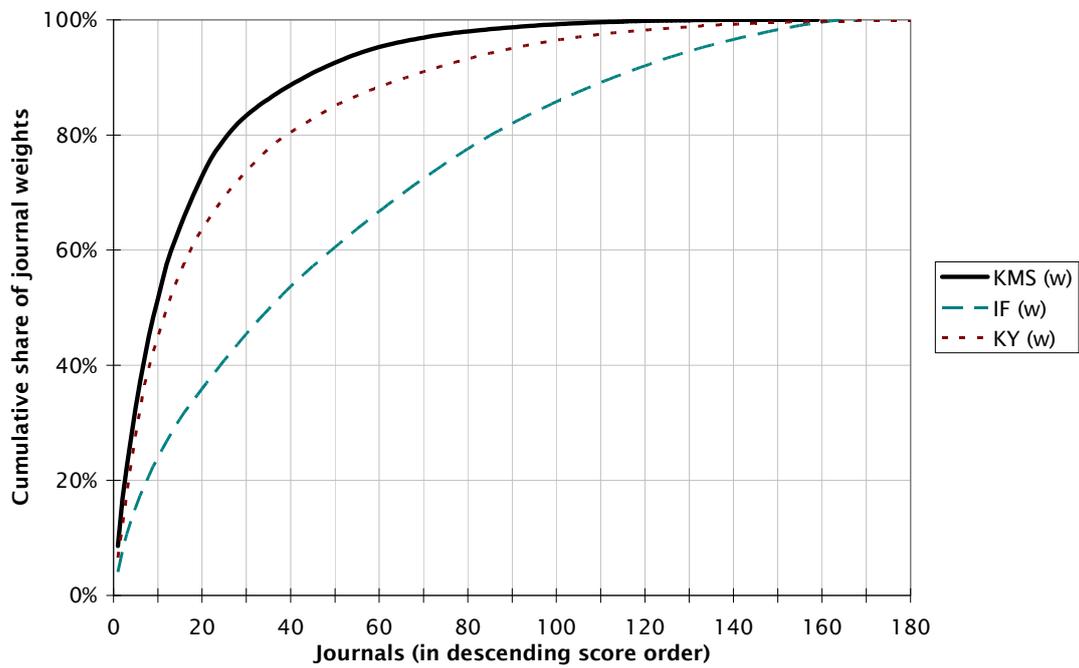
Note: See Table 1 for a description of the notation.

Figure B2: Lorenz curves for the three journal weight measures
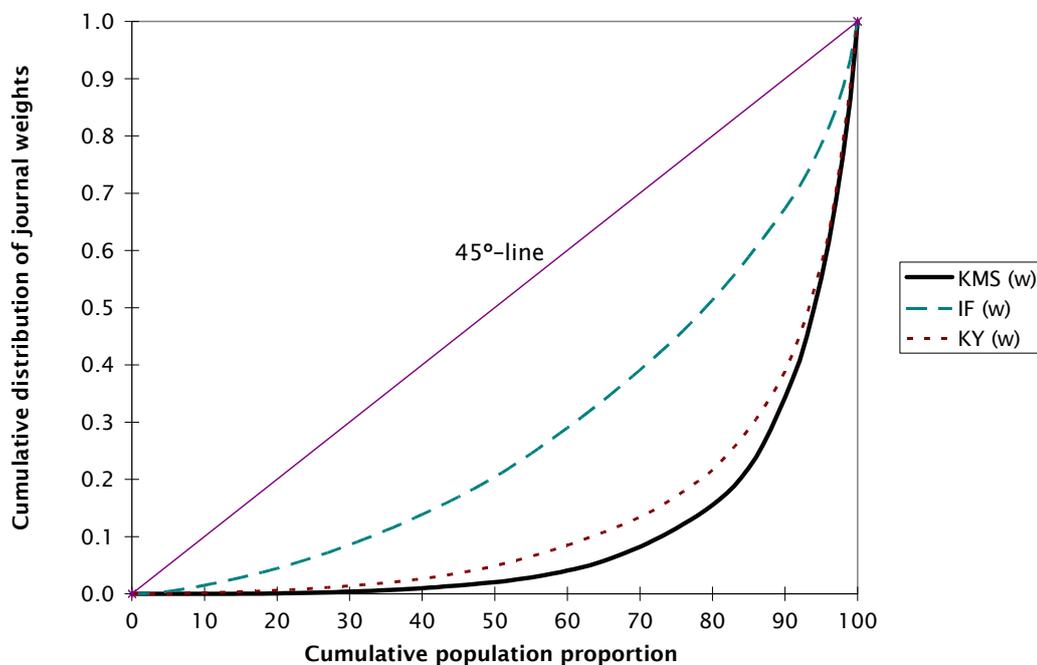
Note: See Table 1 for a description of the notation.

Table 1: Journal weight correlations

	KMS (w)	IF (w)	KY (w)
KMS (w)	1		
	(159)		
IF (w)	0.443**	1	
	(142)	(169)	
KY (w)	0.624**	0.776**	1
	(130)	(149)	(181)

Note: Correlations are pairwise Pearson correlations. Number of observations in parentheses. The "(w)" is added as a reminder that the table examines journal weights and not performance scores for the population of Swedish professors (which are analysed below). ** significant at 1%.

Table 2: Summary statistics and concentration estimates for journal weights of the three measures

Variable	KMS (w)	IF (w)	KY (w)
Max	100	5.243	100
Median	0.92	0.566	2.03
Min	0	0	0
Mean	7.32	0.77	7.38
Std. Dev.	16.05	0.75	14.76
C.V.	2.19	0.98	2.00
P90/P10	∞	7.1	116.9
P90/P25	102.0	4.4	33.1
P90/P50	24.4	2.4	8.6
P95–P100	41.2	20.7	42.3
P0–P50	2.2	20.6	4.5
Gini coefficient	0.791	0.440	0.738

Note: The "(w)" denotes journal weights (instead of performance scores). C.V. is the coefficient of variation (Std. Dev/Mean). P90/P10 is the ratio between the 90th percentile professor (P90) and the 10th percentile professor (P10), and analogously for the P90/P25 and P90/P50 ratios. P95–P100 is the share of points/citations/works attributable to the professors in the top 5% of the distribution and P0–P50 is the share attributable to the lower half.

Table 3: Summary statistics and concentration estimates for research performance according to the seven measures

Variable	KMS	IF	KY	SSCI	GS	*h*-index	Works
Metric	score	score	score	score	# citations	# citations	# works
Co-author weights	yes	yes	yes	yes	yes	yes	yes
Max	961.2	35.7	603.1	297.2	1438.0	28.0	50.7
Median	52.0	3.9	31.3	28.0	69.7	7.0	14.0
Min	0	0	0	0	4.5	1	2.0
Mean	103.7	5.6	61.6	46.7	154.4	7.8	16.3
Std. Dev.	143.8	5.8	83.0	50.0	240.3	4.7	10.9
C.V.	1.4	1.0	1.3	1.1	1.6	0.6	0.7
P90/P10	97.7	16.0	51.0	22.1	19.0	4.3	5.2
P90/P25	30.5	6.8	20.6	8.2	10.0	2.6	3.7
P90/P50	4.8	3.0	5.1	3.6	5.6	1.9	2.1
P95–P100	24.4	19.2	22.4	19.3	30.0	9.8	12.3
P0–P50	7.4	15.7	9.4	13.9	12.7	35.5	25.8
Gini coefficient	0.62	0.50	0.60	0.52	0.60	0.31	0.35

Note: See table 2 for a description of the statistical indicators.

Table 4: Correlations of individual performance across the seven measures

a) Pearson correlations

	KMS	IF	KY	SSCI	GS	*h*-index	Works
KMS	1						
IF	0.826	1					
KY	0.964	0.889	1				
SSCI	0.699	0.734	0.750	1			
GS	0.650	0.651	0.697	0.795	1		
h-index	0.537	0.684	0.613	0.703	0.711	1	
Works	0.428	0.618	0.443	0.429	0.393	0.603	1

Note: The number of observations is 93 in all cases. All coefficients are significant at the 1%-level.

b) Spearman rank correlations

	KMS	IF	KY	SSCI	GS	*h*-index	Works
KMS	1						
IF	0.861	1					
KY	0.953	0.923	1				
SSCI	0.539	0.682	0.638	1			
GS	0.504	0.623	0.612	0.848	1		
h-index	0.490	0.641	0.579	0.674	0.789	1	
Works	0.396	0.565	0.426	0.510	0.507	0.667	1

Note: The number of observations is 93 in all cases. All coefficients are significant at the 1%-level.

Table 5: Ranks in other measures of those ranked as top 10 according to KMS

KMS	IF	KY	SSCI	GS	h-index	Works
1	1	1	1	1	1	4
2	4	2	4	2	23	20
3	8	5	7	21	67	34
4	6	4	21	19	18	47
5	12	3	2	6	28	65
6	7	10	25	35	21	6
7	5	11	12	26	4	5
8	17	13	78	82	48	37
9	10	8	36	34	39	24
10	9	6	9	16	10	35

Table 6: Regression results

	KMS	IF	KY	SSCI	GS	h-index	Works
Sex (Male = 1)	1.83**	0.43	1.25	0.49	0.66	0.52**	0.29
	(0.67)	(0.37)	(0.65)	(0.52)	(0.45)	(0.16)	(0.16)
Research university	1.72**	0.56**	1.23**	0.69*	0.30	−0.04	0.06
	(0.35)	(0.16)	(0.30)	(0.27)	(0.21)	(0.11)	(0.14)
Age at Ph. D.	−0.08*	−0.05*	−0.09**	0.00	−0.05	−0.04**	−0.01
	(0.04)	(0.02)	(0.03)	(0.03)	(0.03)	(0.01)	(0.02)
Time to professorship	−0.04	−0.04**	−0.05*	−0.08**	−0.06**	−0.04**	−0.02
	(0.03)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Years as professor	0.01	0.01	0.01	0.00	0.00	−0.00	0.03**
	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Constant	3.68*	2.72**	4.64**	3.08*	5.98**	3.38**	2.76**
	(1.71)	(0.83)	(1.45)	(1.48)	(1.29)	(0.54)	(0.73)
Observations	93	93	93	93	93	93	93
R-squared	0.45	0.37	0.45	0.25	0.22	0.39	0.31

Note: Robust standard errors in parentheses. * significant at 5%; ** significant at 1%.

Figure 1: Ranking of Swedish economics professors according to KMS
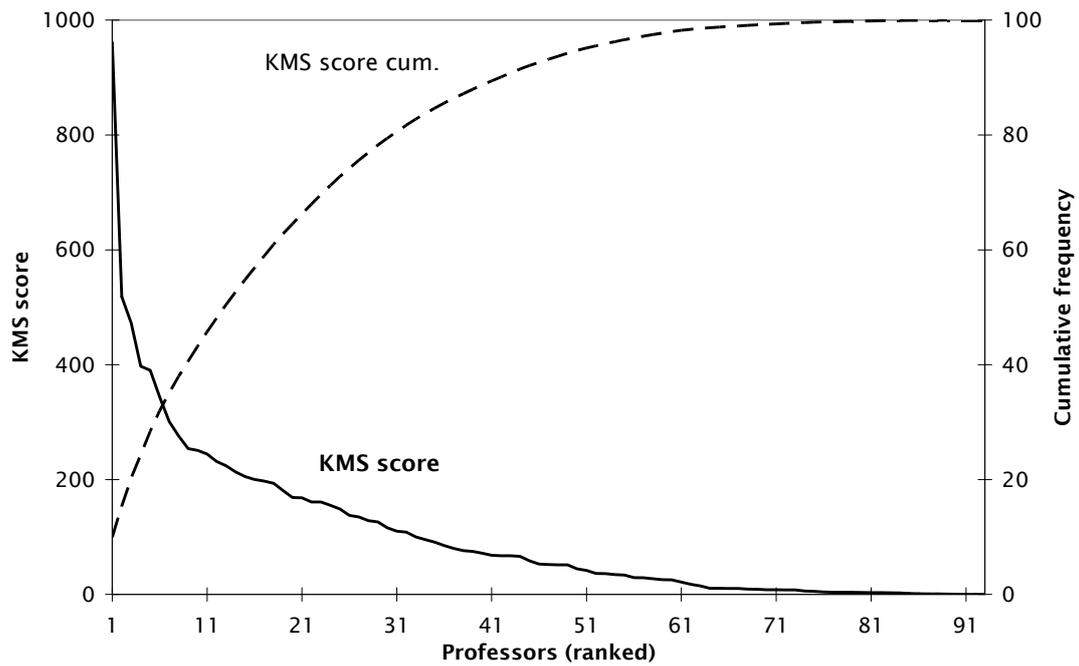
Note: This score is based on the KMS (w) set of journal weights. It awards a score of 100 to an article in the *AER*, which is the highest ranked journal. The journal score then falls rapidly: the 10th ranked journal has a score of 23 and the median journal has a score of 0.92. The lowest third of all journals ranked in KMS are awarded negligible scores (< 0.4).

Figure 2: Ranking of Swedish economics professors according to IF
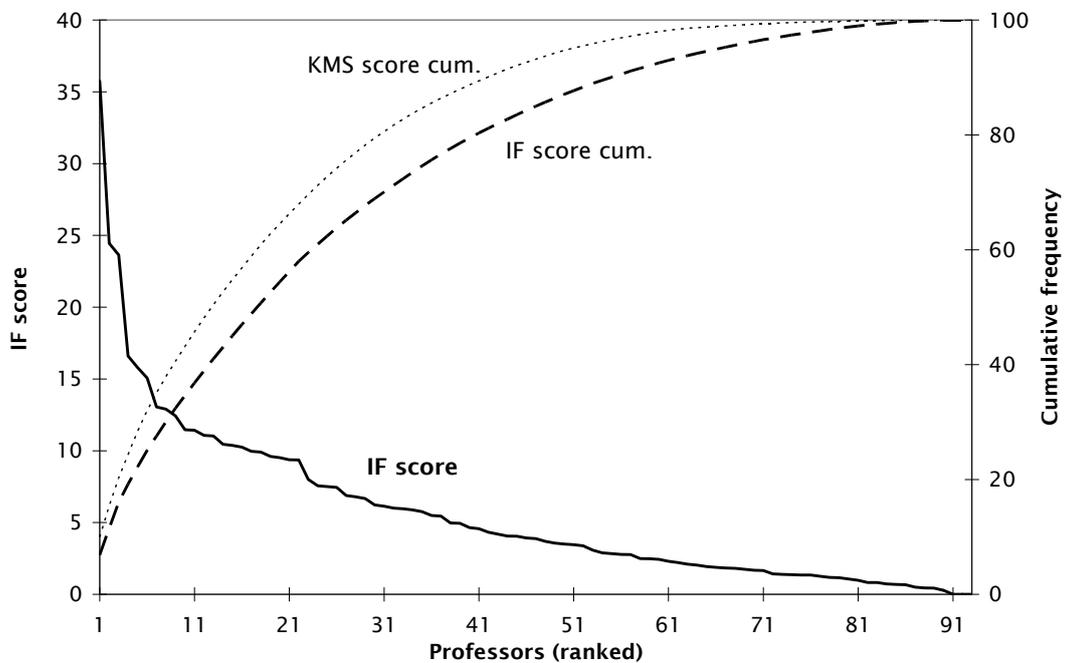
Figure 3: Ranking of Swedish economics professors according to KY
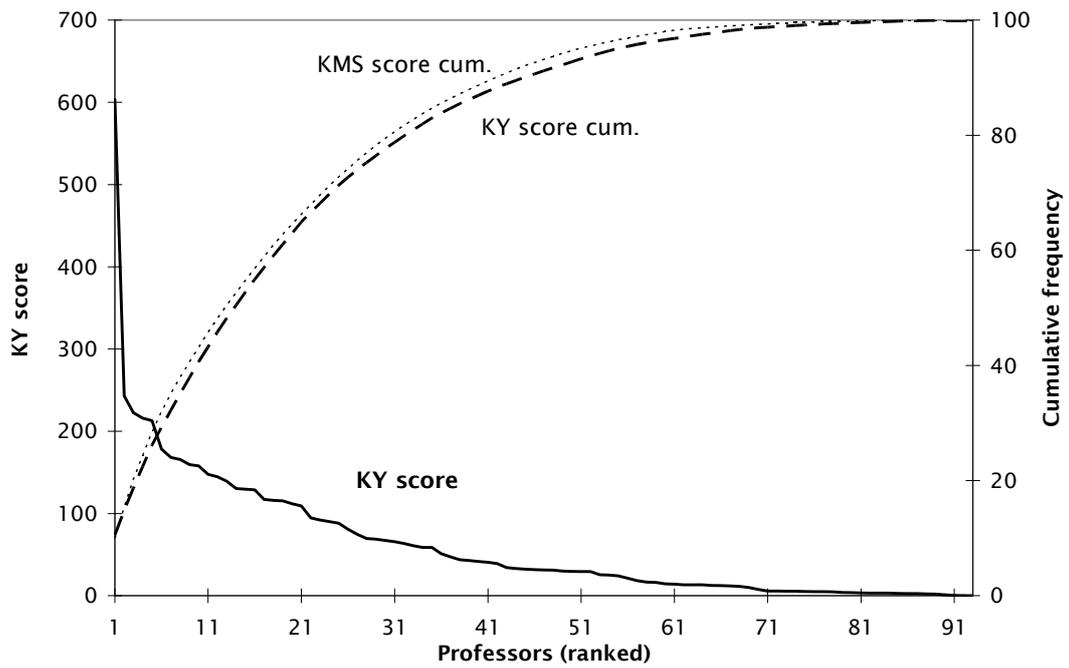
Figure 4: Ranking of Swedish economics professors according to SSCI
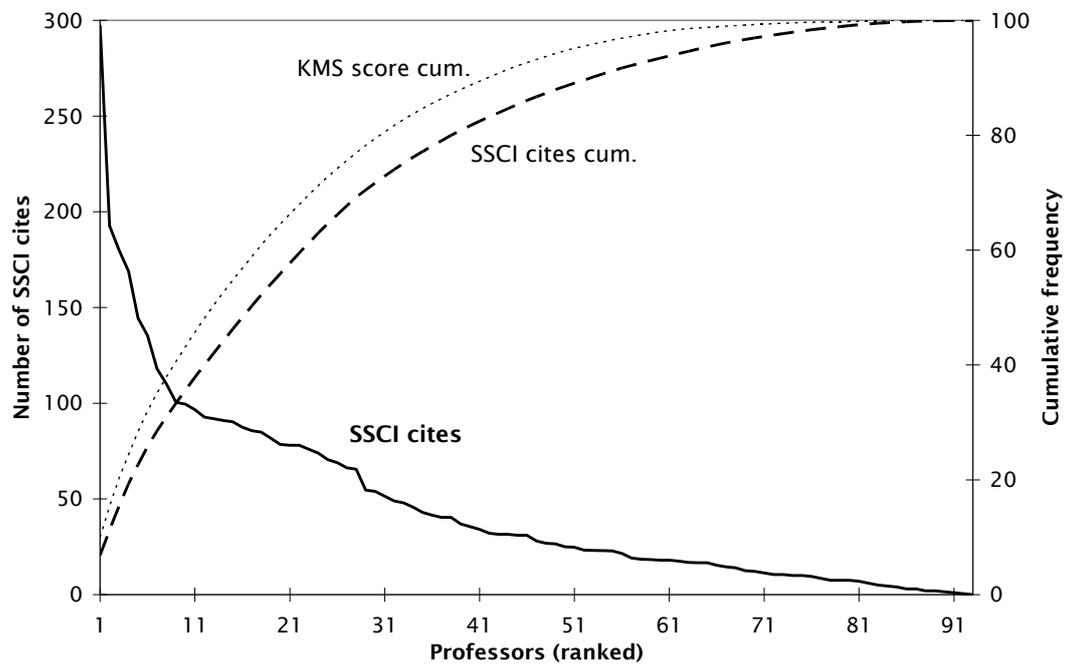
Figure 5: Ranking of Swedish economics professors according to GS
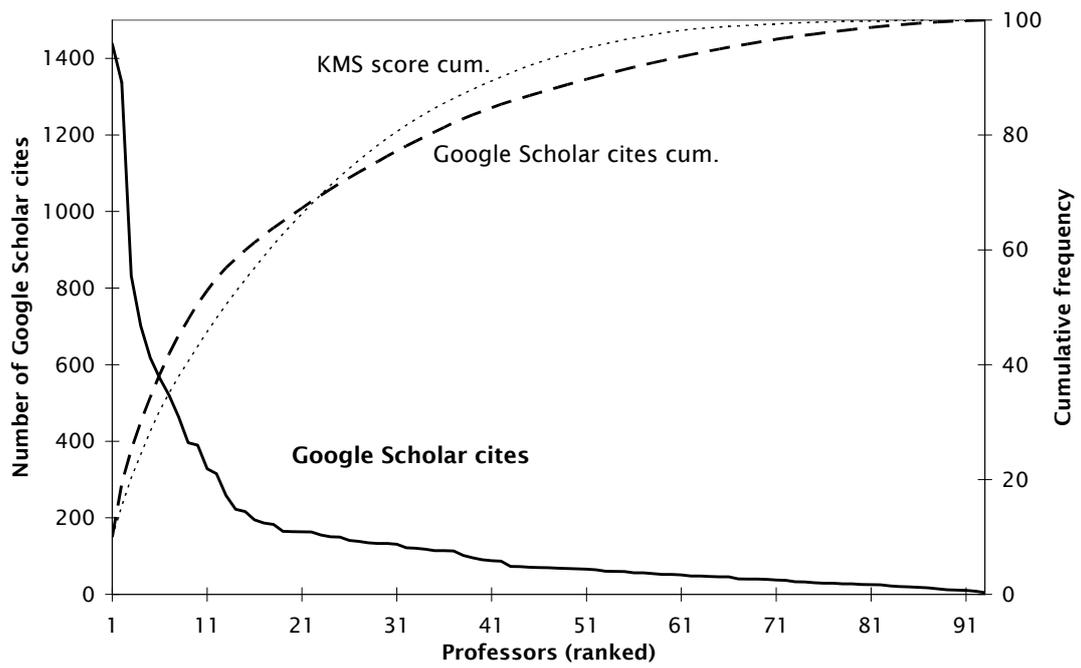
Figure 6: Ranking of Swedish economics professors according to individual *h*-index
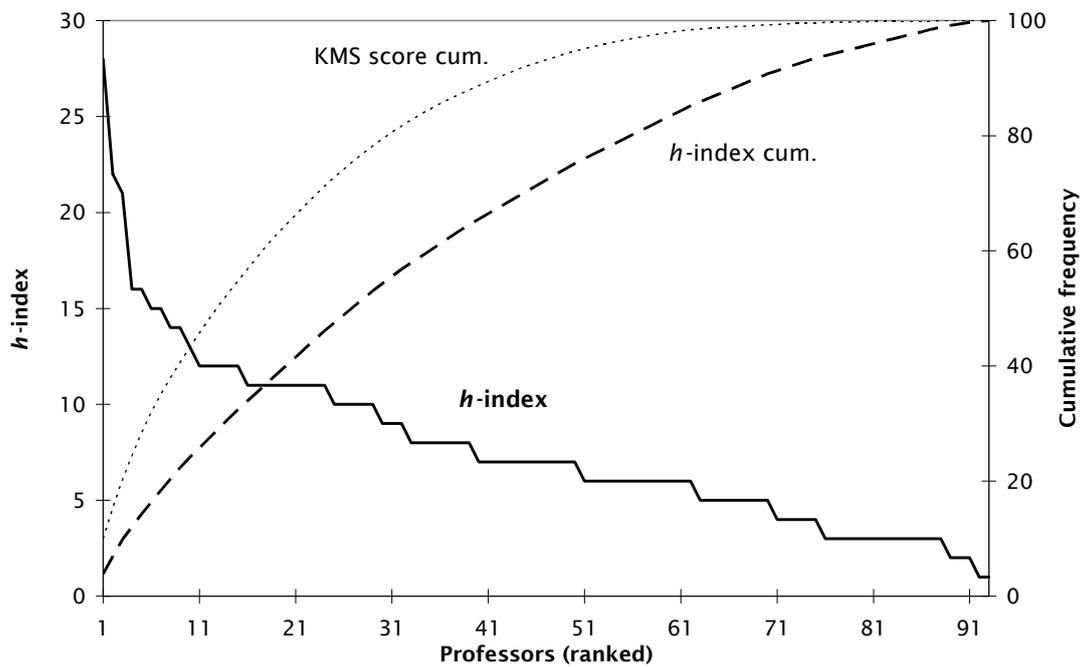
Figure 7: Ranking of Swedish economics professors according to Works
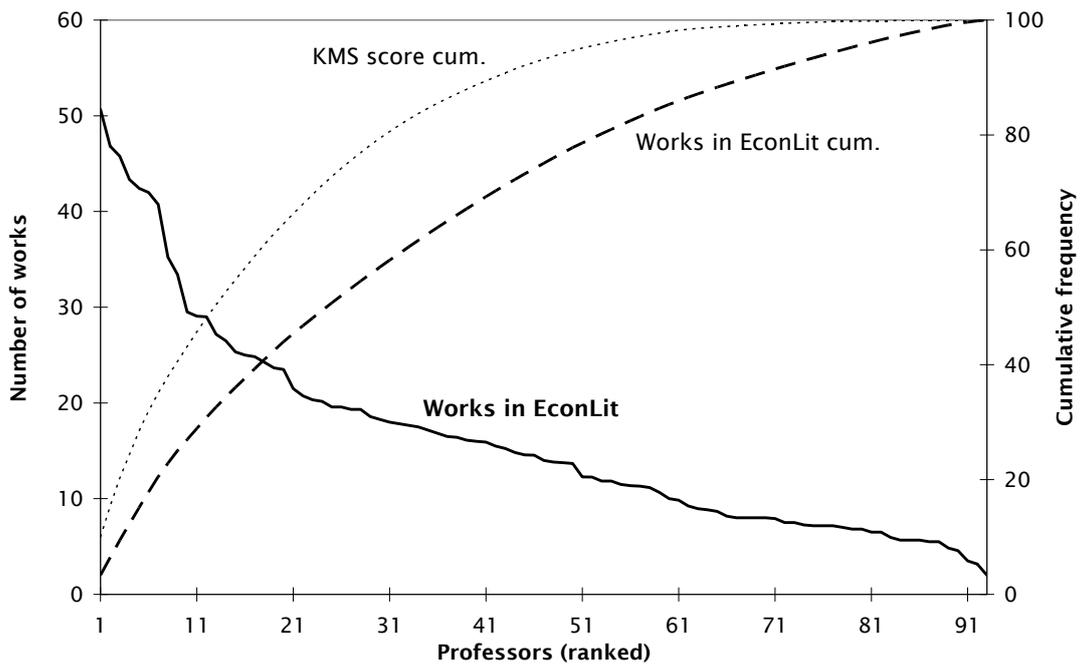
Figure 8: Ranking according to all seven measures of those ranked top-10 in KMS