1 Introduction

Comoving with the ever growing importance attached to academic journals in most fields, a discernible trend of scholarly effort has emerged, aimed at providing reliable accounts of journal "quality". Several reasons could be advanced as contributing motivations for this trend. The perceived quality of one's published work is assessed mostly on the basis of the journals containing one's articles. This evaluation has been very important for quite some time now in determining key decisions such as one's salary, tenure, and promotion. Inter-department evaluations and comparisons of research output are also becoming increasingly common worldwide. For instance, departments at UK universities undergo an evaluation of their research performance every 4 to 5 years, upon which a good part of their funding is contingent. Similar measures are currently being discussed in a number of European countries due to a perception that the university system lacks adequate incentives to motivate performance. The need for reliable quality indicators for the professional journals is thus well-established, for all parties concerned, individuals as well as institutions.[1] Furthermore, it is highly desirable that this information be updated on a regular basis, in view of the extremely rapid increase of the number of economics journals over the last decade.

An important feature of a journal quality index is the objectivity of the criteria upon which it is based. In the abscence of objective measures, the evaluation process is subject to political manipulation and other natural limitations (such as individuals' professional biases), and can thus produce

[1] A separate literature dealing broadly with these and related issues has developed over the years. A partial list (far from exhaustive) of such studies can be given: Frey and Eichenberger (1993), Kirman and Dahl (1994), Stigler, Stigler and Friedland (1995), Conroy and Dusansky (1995), Dusansky and Vernon (1998), Kalaitzidakis, Mamuneas and Stengos (1999, 2001).

rather divergent results.[2]

The first objective criterion used in ranking economics journals was proposed by Liebowitz and Palmer (1984). It is based on counting citations from the SSCI made in a large sample of journals in 1980 to articles published in the five-year period 1975-79, with the key feature of their measure being based on discriminating across citations according to the quality of the journal that issued them. Since the latter quality was precisely what the authors were purporting to assess, there was a need for a recursive procedure. In the first iteration, the total citations received by a journal was used to weight the value of each citation from that journal. The numbers thus obtained were then used in the next iteration as the new weights assigned to each citation, and so on, ad infinitum. They also provide an initial ranking based only on total citations, and two more rankings based on impact-adjusted citations per article and per character. Their pioneering work was updated by Laband and Piette (1994), using a larger sample containing new journals. Both studies derived impact-adjusted rankings that clearly conformed much better (though with a few exceptions discussed below) to a subjective perception of journal quality in the profession than the (first) simple citations ranking, thus suggesting that their criterion is appropriate.[3]

[2] A ranking used in Holland until 1997, assigning grades A down to E, featured *Games and Economic Behavior* as a D journal and *Journal of Economics* as a B journal. The current ranking (with numerical scores) corrected the above flaw among others, but produced some new surprises. In particular, it assigned the highest score of 9.4 to *The Economist*, 3.4 to *Quarterly Journal of Economics*, 1.01 to *Journal of Environmental Economics and Management* and .58 to *Journal of Economic Theory (*JET*)*!

[3] A weakness of the LP ranking is that it attributes a dismal performance to some top field journals such as *Journal of Development Economics* or *Journal of Environmental Economics and Management.* While this is a reflection of the lack of influence these fields have in economics overall, these journals would most likely perform better if the LP

Liebowitz and Palmer did not provide a study of the foundations or of the general properties of their procedure. The purpose of the present note is to fill this gap. We show that their recursive process can be modelled as a Markov chain, and that the impact-adjusted citations measures can be obtained in terms of the limiting distribution of the Markov chain. This provides a simple and rigorous theoretical foundation for their procedure, with many useful conclusions. In particular, it establishes that the process is always well-defined, and will necessarily converge, regardless of the journal sample size or the specifics of the particular data set. Adapting the probabilistic interpretations of some concepts from the theory of Markov chains, the characterization allows a precise interpretation for, and a critical evaluation of, the measure. This assessment of the validity of the measure points to some weaknesses, but mostly justifies its appropriateness as a reflection of the influence levels of journals. Overall then, our results provide a theoretical foundation and a critical appraisal of the intuitive arguments of Liebowitz and Palmer (1984), leading to qualified support for the appropriateness of their proposed measure. In particular, our arguments suggest an explanation for the likely most objectionable outcome of the procedure: The unduly high ranking of two finance journals. Suggestions for some potentially important improvements are also given.

---

journal sample included all social science journals, due to the proximity of these fields to other subjects such as sociology, political science, ecology,...While these considerations are beyond the scope of this paper, they emerge to a limited extent as part of the discussions in Sections 4-5.

2 The Ideal Theoretical Framework

This section provides the theoretical foundations of the impact-adjusted citations measure, ignoring some important issues relating to the delineation of the relevant journal sample. These issues will be dealt with subsequently. Our general results are presented along with an interesting probabilistic interpretation of the impact measure. A summary of all relevant results from the theory of Markov chains is provided in Appendix B.

Consider an ideal situation where an n-journal sample can be unequivocally defined for a given discipline, say economics. Let C_{ij} be the number of citations from journal i to journal j, where $i, j \in \{1, 2, ..., n\}$. In order to be able to invoke results from the theory of Markov chains in a direct way, it is convenient to represent our data in proportion terms. Thus, let

$$c_{ij} \triangleq \frac{C_{ij}}{\sum_{j=1}^{n} C_{ij}} \ , \tag{1}$$

where c_{ij} is the proportion of journal i's total citations that are to journal j. The $n \times n$ matrix $C \triangleq [c_{ij}]$, clearly having each entry in the interval $[0, 1]$ and each row sum equal to 1, is a stochastic matrix. As c_{ij} may be interpreted as the probability that a randomly selected citation in journal i is to journal j, the matrix C may be viewed as a Markov chain. The state space is the set of all journals under consideration, and thus has cardinality n, and a transition from journal i to journal j reflects the fact that a citation contained in journal i, randomly drawn according to the probability distribution $(c_{i1}, c_{i2}, ..., c_{in})$, actually goes to journal j.

An important advantage of this representation is that it allows direct insightful analogies with the standard probabilistic interpretations of various maningful derived variables in the theory of (finite-state) Markov chains. Thus, for instance, the ij^{th} entry of the matrix $C^2 \triangleq C \times C$ represents the probability that a randomly selected citation in journal i goes to journal j in exactly two steps, or in other words, the probability that the citation is first made from journal i to any of the n journals, and then from the latter to journal j. Likewise, but more generally, the ij^{th} entry of the matrix C^t, denoted[4] $c_{ij}^{(t)}$, stands for the probability that a randomly selected citation in journal i goes to journal j in exactly t steps, for any $t = 1, 2, ...$Following Liebowitz and Palmer (1984), henceforth LP, we define normalized impact-adjusted value iterates for each of the n journals recursively as follows:

$$q_i^{(0)} = \sum_{k=1}^{n} c_{ki} \text{ and } q_i^{(t)} = \sum_{j=1}^{n} c_{ji} q_j^{(t-1)} \text{ for all } t = 1, 2, ... \tag{2}$$

Although the $q_i^{(t)\prime}s$ play only an intermediate role in defining journal quality, we provide a suitable interpretation for them below. We begin with a preliminary simple result that sheds some light on the meaning of the normalized Impact-Adjusted value iterates.

Lemma 1 *For each* $i = 1, 2, ..., n$, $q_i^{(t)} = \sum_{k=1}^{n} c_{ki}^{(t+1)} = i^{th}$ *column sum of the matrix* C^{t+1}.

Proof. We give the proof for $t = 1$ only, as the rest follows analogously (by induction on t).

$$q_i^{(1)} = \sum_{j=1}^{n} c_{ji} \left(\sum_{k=1}^{n} c_{kj} \right) = \sum_{j=1}^{n} \sum_{k=1}^{n} c_{ji} c_{kj} = \sum_{j=1}^{n} \sum_{k=1}^{n} c_{kj} c_{ji} = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} c_{kj} c_{ji} \right) = \sum_{k=1}^{n} c_{ki}^{(2)},$$

where all the middle equalities follow from standard properties of double summation, and the last one from matrix multiplication rules. ■

These impact-adjusted value iterates can thus be interpreted as follows. $q_i^{(0)}$ is simply the proportion of the total citations that is received by journal

[4]Note that while the t in C^t indeed stands for an exponent (in the sense of matrix multiplication), the t in $c_{ij}^{(t)}$ and $Q_i^{(t)}$ does not. Rather, it indicates the iteration stage, whence the parenthesis (see Appendix A.)

i, and hence the basis for Liebowitz and Palmer's initial ranking. Then, for each t, $q_i^{(t)}$ is the proportion of the total t-stage citations (as described in the previous subsection) that would accrue to journal i.

Our main result requires the following assumption (interpreted and discussed below):

(A1) *All pairs (i, j) communicate, i.e., for every i, j, there exist t such that $c_{ij}^{(t)} > 0$.*

Assumption (A1) is the so-called irreducibility property of Markov chains. In the context of the model at hand, Assumption (A1) simply says that while it may be true that some journal has not received any citations at all from some other journal, it will always be possible to find a finite sequence of journals starting from the former and ending with the latter, with each journal along the sequence having referenced the preceding one at least once. This is certainly a very plausible assumption here[5].

Proposition 2 *Assuming (A1), there exists a vector $c^* = (c_1^*, c_2^*, ..., c_n^*)$ such that as $t \to \infty$,*

(i) the matrix C^t converges (entry-wise) to the matrix C^* having identical rows, each equal to c^*, or in other words, $c_{ij}^{(t)} \to c_j^*$, with the rate of convergence being geometric, and

(ii) for each $i = 1, 2, ..., n$, the impact-adjusted value iterates $q_i^{(t)}$ converge to $q_i^* = nc_i^*$.

[5]Thus, for instance, while there may be no references in *Yale Law Journal* (YLJ) to *Journal of Mathematical Economics* (JME) for the period considered, it is rather likely one can find references in YLJ to *Journal of Law and Economics* and in the latter to *Bell Journal of Economics*, and then on to JME. On the other hand, Assumption (A1) does preclude the presence of noneconomics (say mathematics) journals that are referenced by some economics journals, but never themselves reference articles from economics journals. The presence of one such journal would imply the Markov chain C has an absorbing state.

Furthermore, the vector c^* is the unique invariant distribution of the Markov chain, i.e., the unique solution to the equation $x = Cx$, where x is an n-dimensional probability vector.

Proof. Part (i) follows directly from Theorem (A.1) (in Appendix). Part (ii) follows from Lemma 1 and the fact that the rows of C^* are all identical (from Part (i)). ■

The impact-adjusted citation value of journal i can now be formally defined in an asymptotic sense as q_i^*, or equivalently as c_i^* as the two differ only by the factor n, for all i.

Liebowitz and Palmer (1984) also defined two other measures of influence: impact-adjusted citations per article and per character. Since these are defined from the previous numbers by simple division, by the numbers of articles and of characters, respectively, appropriately modified versions of Proposition 2 will also apply to these cases.

Next, a discussion of the Proposition is provided, along with an interpretation of the c_i^*'s. It is certainly comforting to know that the inderlying process is always well-defined, and converges to the desired limit as long as Assumption (A1) holds, independently of the number of journals in the sample and of the distribution of the particular citation data at hand. Furthermore, the geometric rate of convergence is relevant for computational purposes, as one could truncate the recursive algorithm after a relatively small number of iterations and still obtain a good estimate of the desired values.

Although the problem at hand is deterministic, one can adapt the well-known probabilistic interpretations associated with various notions in Markov chain theory to the present set-up, and thereby derive useful insights. Consider the following fictitious experiment. Pick an initial journal according to the uniform distribution (i.e. with probability $1/n$ for each journal.) Draw a citation from this journal (for the relevant period) according to the distribution given by the corresponding row of C. Repeat this process starting from the journal that received the latter citation, and so on, ad infinitum...Then, c_i^* may succinctly be interpreted as the long-run proportion of time the citation process is at Journal i, for all i. In other words, c_i^* is a measure of the steady-state frequency with which Journal i is invoked by all the journals in the sample. From this standpoint, c_i^* emerges as a very appropriate measure of the relative influence of journal i, as reflected in all the journals in the sample.

If the stochastic vector π is the probability distribution for picking a journal this period, then $\pi^{'} = C\pi$ gives the probability distribution of the next journal (or chosen citation). Since c^* (uniquely) satisfies $c^* = Cc^*$, it may also be seen as a measure of the relative recurrence of the different journals in the citation process. If each journal i is selected with probability c_i^* this period, it will emerge again with probability c_i^* next period.

These interpretations constitute a precise confirmation of LP's original intuition in devising this measure as an attempt to incorporate the important notion that citations from higher-quality journals should be assigned higher weights than citations from lesser journals, with journal quality itself to be endogenously determined from the data. It is certainly of critical importance for future use of this method to know that it is well-defined and will always converge to the desired values, independently of the specifics of the data set.

3 The Liebowitz-Palmer Framework

The actual LP framework differs from our ideal model in some important respects. The data used is from the Social Science Citation Index (or SSCI), which provides information on the total number of citations from about 4300 journals, covering all of the social sciences and some natural sciences. The selection of 107 economics journals actually ranked by LP was suggested by the fact that these journals were featured in recent issues of the *Journal of Economic Literature* (or JEL.) This list indeed contains virtually all academic economics journals, in addition to a *selection* of finance, law and statistics journals.

LP deals with citation totals as opposed to proportions. Their procedure is defined as follows. In the first step, each citation from any JEL-suggested is weighted according to the total number of citations received from *all SSCI journals* by the citing journal. Thus[6]

$$Q_i^{(1)} = \sum_{j=1}^{n} C_{ji} \left(\sum_{k=1}^{m} C_{kj} \right), i = 1, 2, ..., n. \tag{3}$$

where C_{ij} is the number of citations from journal i to journal j, m is the number of SSCI journals (i.e. 4300), n is the number of JEL-suggested journals (i.e. 107), and $Q_i^{(1)}$ is the first iteration of the impact-adjusted value for journal i. The numbers $Q_1^{(1)}, Q_2^{(1)}, ..., Q_n^{(1)}$ thus generated are then used as weights for the next iteration in this process, and so on, ad infinitum. In other words, the t^{th} iteration ($t = 2, 3, ...,$) is

$$Q_i^{(t)} = \sum_{j=1}^{n} C_{ji} Q_j^{(t-1)} \text{ for } i = 2, 3, ..., n. \tag{4}$$

Observe that the 0^{th} iteration, $Q_i^0 = \sum_{k=1}^{m} C_{kj}$ is based on all 4300 SSCI journals, while all the subsequent iterations are based only on the 107 economics journals. Clearly, Q_i^0 is exactly the total citations received by journal i from

[6]We have kept LP's notation except that we have interchanged rows and columns to conform to standard notation for Markov chains, with the transition matrix having row sums (not column sums) equal to 1.

all the 4300 SSCI journals, and the Q_i^0's constitute the first ranking in LP (given in the middle column of their Table 1, for articles published during 1975-79). As noted by LP, this ranking reflects a measure of unweighted influence accross all the social sciences. It is thus not surprising that two law journals appear at #5 and #12, and a statistics journal at # 4. These three journals probably garnered most of their cites from noneconomics journals, and do indeed represent influencial sources in their respective fields. With a view to arrive at a ranking of economics journals alone, LP devised the above recursive scheme where the citations from journals outside the JEL-suggested list of "economics" journals were counted only in the first step.

3.1 A Critical Assessment of the LP Measure

Relative to the ideal framework of Section 2, the LP procedure incorporates two nonstandard steps in adapting the available data from the SSCI to the goals of their work. We now provide a description of these two steps and an analysis of the likely effects they have had on the LP measure. We argue that, in view of the desired goal of an economics ranking only, one of these two steps was favorable to the journals that are on the fringes of economics, but within the JEL-suggested list. These are the *selected* finance, law and statistics journals. These will be referred to as the fringe journals below, while the rest of the JEL-suggested journals will simply be the economics journals.

The first of these steps is the hybrid definition of the $Q_i^{(t)}$'s, with a clear difference between the first iteration and all subsequent ones (see equations above). We will refer to this step as the truncation effect. Fortunately, using our Markov chain characterization, we know that this step is of no consequence to the final measure. Indeed, this follows directly from the ergodic nature of the Markov chain (i.e. the fact that $x_{t+1} = Cx_t$ converges to the unique steady-state distribution *for any initial state* x_0, see Appendix)[7].

To consider the second nonstandard step, it is instructive to paraphrase the LP procedure in the language of Markov chains. Let the initial data from the SSCI be represented in proportion terms as in Section 2 by an $m \times m$ matrix C, with the journals ordered w.l.o.g. in such a way that the first n journals are the JEL-suggested journals. Take its $n \times n$ submatrix to the North-West, call it $\widehat{C}$. As this is a sub-stochastic matrix (row sums are less than 1), divide each entry by the corresponding row sum to convert it into a stochastic matrix[8], call it $\widehat{C}_s$. Then find the invariant distribution of $\widehat{C}_s$: This gives the LP impact-adjusted values for the n journals. We will refer to this step as the normalization effect (of unit row sums in going from $\widehat{C}$ to $\widehat{C}_s$). For each row $i = 1, 2, ..., n$, this amounts to a redistribution of the cites from economics journal i to all noneconomics journals, towards all the economics journals, in a way that preserves the latters' original proportions of citations from i.

The number of cites to non-JEL journals is clearly higher for the fringe journals than for the economics journals, because only a selection of the for-

[7]The truncation effect does introduce a one-time bias that favors the fringe journals since the column sums corresponding to these journals increase by larger amounts, when taken over all SSCI journals instead of over only JEL-suggested journals, than those corresponding to economics journals. In other words, the amount of citations received from SSCI journals that are not JEL-suggested is probably much higher for a fringe journal than for an economics journal. This is simply because all economics journals are on the JEL-suggested list, while only a selection of (say) the law journals are. The latter naturally get a lot of cites from other law journals that were not JEL-suggested.

[8]Recall that Proposition 2 relies crucially on the matrix being a stochastic matrix. Otherwise, if the row sums are larger (resp. smaller) than one, powers of the matrix will generally converge to ∞ (resp. 0). It is not entirely clear what normalization LP used: see their Footnote 7 on this point.

mer is in the JEL list. This implies that the normalization effect is more significant for the fringe journals than for the economics journals, in the sense that more "mass" gets redistributed. Furthermore, another important difference between the economics and the fringe (say, finance) journals is that the latter cite each other much more than they cite economics journals. Thus, setting w.l.o.g. *Journal of Financial Economics* (JFE) and *Journal of Finance* (JF) as the first two journals on our ordered list, we know that c_{11}, c_{12}, c_{21}, and c_{22} are relatively large numbers to start with. After the sizable proportional normalization, these four numbers will get even larger. Hence these two entries will get a lot of weight in the invariant distribution[9], i.e. high impact-adjusted values. This effect probably contributes in a substantive way to the surprisingly high performance of JFE and JF[10]: See the illustrative example in Appendix A.

Similar arguments apply to the statistics and the law journals. While these did not rank as highly as the finance journals did, they peformed better than they probably should when viewed as economics journals. Indeed, *Journal of Legal Studies* impact-ranked higher than *European Economic Review* and *Yale Law Review* higher than *Journal of Industrial Economics.* In fact, relative to one another, the fringe journals were probably ranked by the LP procedure just about as one would expect them to be, *as economics journals.* Thus the rankings simply reflect the fact that finance journals are

[9]As an extreme case, if {1,2} were an absorbing class (i.e. if $c_{11} + c_{12} = c_{21} + c_{22} = 1$), one invariant distribution will have Journals 1 and 2 sharing the total impact-adjusted value, with all other journals getting 0. While this would clearly violate Assumption **(A1),** a slight redistribution of mass would not.

[10]In the follow-up ranking by Laband and Piette (1994), the JFE, a specialized journal, even ranked first in impact-adjusted citations per article, which we argue below along with LP, is the more reliable measure. Recall that both rankings are based on periods before the current boom in the finance literature.

closer to economics than are law or statistics journals[11]. The three fringe fields were treated in similar terms by the LP procedure: The JEL list included only a few of the journals from each of the three disciplines, so that the two effects discussed above probably led to distortions of comparable magnitude.

3.2 Amending the LP Procedure

The selection by the JEL of the particular noneconomics journals used by LP was motivated by their proximity to economics, relative to other journals in the same discipline. Unfortunately, this otherwise reasonable choice seems to have biased the measure via the mechanism described above. The analysis presented so far suggests a number of improvements that might be useful in a future journal ranking. The key issue here being the *definition of the relevant sample of journals*, three alternative ways of performing an impact-adjusted ranking with minor distortions from the available SSCI data emerge.

The first is to deal with the economics journals only, from beginning to end. This would amount to following the Ideal Theoretical Framework of Section 2 with the economics journals only, i.e. without the selected fringe journals (from finance, law and statistics.)

The second procedure is to follow the Ideal Theoretical Framework of Section 2 with the economics journals, *and all* the journals from any other social science one wishes to include, e.g. finance, law and/or statistics[12]. This

[11] Furthermore, assuming that (say) law journals quote law and economics journals more than the other way around, the presence of some law journals favored the law and economics journals, relative (say) to the economic history journals, in view of the abscence of any history journals.

[12] In selecting a subset of these related disciplines to include, one must always bear in mind the potentially important point raised in Footnote 10.

recommendation follows the principle that if one includes *some* journals from a field related to economics, then one must include *all* the journals from that field, in order to prevent the two distortive effects described above from being as influential as they seem to have been.

The third alternative would be to apply our ideal measure to all the SSCI journals, and *at the very end* of the procedure, discard the subset of journals one is not interested in. The latter subset could then be selected with quite some flexibility; It may include any noneconomics journals, a subset of the finance journals only, etc...This procedure would then measure the overall influence across all the social sciences of economics journals (and any other social science journals), relative to one another.

The actual selection of the most appropriate procedure among the above three would naturally depend on the ultimate aims of the ranking. A standard economics department would probably find the first method most appropriate, possibly the second as well. On the other hand, a business school would probably prefer to invoke a ranking based on the third method, in case comparisons across various fields are desired.

4 Assessing the Ideal Impact-Adjusted Measure

In this section, an attempt is made to critically evaluate the impact-adjusted citation measure in the ideal world of Section 2, i.e. abstracting away from the sample definition problems of Section 3. The essential nature of the measure has already been discussed in Section 2. Here, we simply expand on some specific points that deserve particular attention, and expose some weaknesses. This discussion makes critical use of the insights from our characterization of the LP measure as the invariant distribution of a Markov chain.

4.1 Total vs per-Article Impact-Adjusted Measures

Here, we provide some arguments confirming LP's suggestion that a per article measure is more appropriate than a total measure for capturing a quality notion for journals. Suppose two journals A and B cite all other journals in the same proportions. Then, invoking our characterization, it is easy to see that a merger of the two journals would yield a new journal whose LP total impact ranking in an $(n-1)$-journal sample is the sum of the rankings of A and B in the original n-journal sample[13]. This is clearly a very undesirable property of the total-impact measure, as the new journal should be regarded as having the average quality of its constituent parts. The per-article measure clearly satisfies the last property, and is thus immune to mergers or break-ups. Likewise, the per-character measure also has this property. This is a crucial invariance property for a measure of journal "quality".

A related question of interest is whether the two LP measures favor journals in large fields over those in small fields of economics. The following illustration sheds light on this issue: Consider a three-journal world with A as a general-interest journal, and B and C as (essentially unrelated) field journals such that C's field is twice as large as B's field (i.e. C contains twice as many articles as B). Let the citation proportions matrix be given by

[13]Analogous remarks can be made for the reverse operation of a break-up of a journal. This happened in mathematics when *Annals of Probability and Statistics* was broken up to form *Annals of Probability* and *Annals of Statistics.*

	A	B	C
A	2/3	1/9	2/9
B	.4	.6	0
C	.4	0	.6

Thus A is equally likely to quote a single article from B and C, and each article from B or C attaches the same importance to general-interest articles and (for simplicity) none to the other field. The total-impact measure can be computed roughly as $(.55, .15, .30)$, thus indicating C is twice as influential as B. Yet, the data at hand can be naturally interpreted as pointing to an equal influence level for the two journals.

Since C contains twice as many articles as B, the LP per-article measure would correct the above discrepancy. Thus this provides another perspective pointing towards the superiority of the per-article measure. Note that the above argument is robust to many aspects of the specification of the citations matrix, provided the described essential points are maintained.

On the other hand, it is worthwhile to point out that the size of the field is probably still relevant to the LP measures, in that a good correlation can be expected between the weight of the different fields in the composition of articles in the general-interest journals and the size of those fields. Thus, using our illustration, articles in A would be more likely to quote a single article in C than in B, something not reflected in our example above.

4.2 Some Pros and Cons of the LP Measures

Here, we simply provide a brief description of some of the strengths and weaknesses of the "ideal" LP measures (without explicitly distinguishing between the two.)

4.2.1 Effects of the Citation Horizon

Both LP and Laband and Piette (1994) considered citations made in a given base year to articles published in the preceding five-year period. This certainly captures the immediate impact of the article. Yet, we all know that some classic papers in economics continue to be intensely cited twenty years after publication, and even beyond. In fact, it would be reasonable to attribute a high quality mark to such proof of ”durability”. A natural way of amending the procedure in this direction is simply to add a few more years to the citation horizon, and consider (say) the 10th year and the 15th year following publication, in addition to the five-year period. These two extras years should probably be weighted even higher than the other years, to reflect the value of long-lasting influence (in other words, one might use a ”discount” factor larger than 1.) Such an amendment would also serve to partially factor out fad effects in economic research, and reward journals publishing more fundamental research work as opposed to work motivated largely by current trends or policy debates[14].

4.2.2 Sensitivity to Citation Traditions

The LP measure has the drawback of being sensitive to citing traditions in the different fields of economics, a relatively broad discipline. This feature can be illustrated within the setting of a three-journal world again. Suppose

[14] In this regard, it is worthwhile to point out that the LP measures would have given a dismal score to Cournot’s 1937 work for about one hundred and thirty years, only to award him with a massive score in subsequent decades. While this example is indeed extreme, milder versions of the same phenomenon are quite easy to find. An example is a pioneering article by Ruff in JET 1969 on R&D cooperation, that waited until the topic blossomed in the 90s with more than 100 articles published to date, to go from no quotes to plenty of them.

A is the only general-interest journal, and it contains half of its articles in the field of Journal B and half in the field of Journal C. Each of the former articles makes 15 citations, two-thirds to A and one-third to B. Each of the later articles makes 30 citations, also two-thirds to A and one-third to B. Each field journal has on average two-thirds of its cites to A and one-third to itself. This gives rise to the matrix

	A	B	C
A	2/3	1/9	2/9
B	2/3	1/3	0
C	2/3	0	1/3

whose invariant distribution is $(2/3, 1/9, 2/9)$. Hence, a double propensity to cite in its field leads to a double impact for Journal C, relative to Journal B! This is clearly a major theoretical shortcoming of the measure. Its practical magnitude clearly depends on the actual relative differences in citation lengths across different fields. A quick check of Journal Citation Reports indicates that articles in *Journal of Labor Economics* have on average 31 references, against 21 for the *Journal of Economic Theory*. If such differences prove significant, it would be appropriate to introduce a corrective normalization.

A related positive aspect of the LP measure is that if any subset of the journals experience citation inflation (a term coined by LP) of a *uniform* sort, modelled by multiplying all of the citations of each of those journals by a constant (depending on the journal), the LP measure will be unaffected, since the citation proportions matrix remains the same.

5 Conclusion and Recommendations

This paper has provided a simple theoretical foundation, based on the theory of Markov chains, for the objective procedure for journal ranking devised by LP. An elementary characterization is derived that substantially clarifies the meaning and scope of the procedure, and provides an intuitive interpretation for it. A detailed and illustrative discussion of various strengths and weaknesses of the measure is also made possible by this characterization. In particular, the excessively high rankings of the finance journals is explained as a consequence of some data-handling steps performed by LP. Several suggestions are made to correct the shortcomings of the LP study that may prove useful in much-needed current rankings. The most important issue raised centers around the appropriate definition of a *comprehensive* and *coherent* sample of journals, given the available SSCI data. Overall, an amended version of the LP procedure would emerge as a highly appropriate measure[15].

Last but not least, our simple analysis provides simple and neat answers to skeptics whose typical first question is, understandably, "but, how do we know that the procedure will always converge, and if so, to anything meaningful?", and might offer convincing evidence against those who prefer no measure at all to a somewhat imperfect measure. Yet, it remains inevitable that with any systematic evaluation system in place, the academic research process will eventually respond in strategic ways, thereby reducing some of

[15]One amendment of interest adopted in the latest study on both journal and department rankings, is to drop self-citations for each journal, to correct for an inherent bias of the citations process. Our characterization stilll holds in this case, the only change being that the diagonal mass in each row now gets redistributed uniformly to the other elements of the same row (here, we are assuming that each journal would "communicate" with itself in two or more steps.)

the concomitant benefits[16].

6 Appendix A

We provide an illustration of the normalization effect to get a feel for the distortion it entails in favor of (say) the finance journals. Consider a 6-journal sample comprising 2 finance journals (F_1 and F_2), and 4 economics journals of which 2 are general-interest (EG_1 and EG_2) and the other 2 are specialized (ES_1 and ES_2). With the order on journals being ($F_1, EG_1, EG_2, ES_1, ES_2$, F_2), let the citation matrix C be given by

$$C = \begin{bmatrix} .5 & .05 & .05 & 0 & 0 & .4 \\ .05 & .4 & .4 & .1 & .05 & 0 \\ .05 & .4 & .4 & .05 & .1 & 0 \\ 0 & .35 & .35 & .25 & .05 & 0 \\ .05 & .35 & .35 & .2 & .05 & 0 \\ .4 & .05 & .05 & .05 & .05 & .4 \end{bmatrix}$$

The citation proportions have clearly been chosen so as to reflect the nature of the journals at hand in a realistic way. The impact-adjusted values are then numerically given by the invariant distribution of the matrix C

$$q_C = (.144\,76, .308\,45, .308\,45,\ .08363\,7, .05818\,3, .09652\,7).$$

Next, we discard the finance journal F_2 from the sample (last row and column), and reallocate mass so as to make the resulting $5x5$ matrix stochastic. We obtain:

[16] As in physics, the mere fact of observing a system may well uncontrollably change its behavior.

$$E=\begin{bmatrix} .833\,33 & .0833 & .0833 & 0 & 0 \\ .05 & .4 & .4 & .1 & .05 \\ .05 & .4 & .4 & .05 & .1 \\ 0 & .35 & .35 & .25 & .05 \\ .05 & .35 & .35 & .2 & .05 \end{bmatrix}$$

for which the invariant distribution is

$$q_E=(.\,219\,29,.\,323\,06,.\,323\,06,.07\,901\,1,.05\,503\,1).$$

Thus, due to normalization, the finance journal gained $(.\,219\,29-.144\,76)/.144\,76=51.5\%$ while each of the two general-interest journals only gained $(.\,323\,06-.308\,45)/.308\,45=4.7\%$!

7 Appendix B

We present here a brief summary of the main properties of stationary (finite-state) Markov chains used in the text. A complete treatment can be found in Kemeny and Snell (1960).

An $n\times n$ matrix $[p_{ij}]$ is called a stochastic matrix if $p_{ij}\geq 0$ for all i,j and the sum of the elements across each row $\sum_{j=1}^{n}p_{ij}=1$. In other words, each row is a probability vector. With $S=\{1,2,...,n\}$ being a description of the set of states, the pair (S,P) is a stationary (finite-state) Markov chain describing the following dynamic (Markov) process: If the current state is $i\in S$, the next state will be j with probability p_{ij}, independently of all previous values, i.e., for any i,j,k,l in S,

$$P(x_{t+1}=j/x_t=i,x_{t-1}=k,...,x_0=l)=P(x_{t+1}=j/x_t=i)=p_{ij}$$

If the current state is i, the probability that in t stages the state is j is $p_{ij}^{(t)} = P(x_t = j/x_0 = i)$. It can be shown that the matrix $\left[p_{ij}^{(t)}\right]$ equals P^t, the t^{th} power[17] of P.

Two states i and j are said to communicate if each can be reached from the other with positive probability (in finite time), i.e., if there exist integers t_1 and t_2 such that $p_{ij}^{(t_1)} > 0$ and $p_{ij}^{(t_2)} > 0$. If all possible pairs of states in S comminicate in a stationary Markov chain, it is said to be irreducible. This property excludes absorbing states (a state i is absorbing if $p_{ii} = 1$) as well as transient states (a state i is transient if $p_{ii}^{(t)} \longrightarrow 0$ as $t \longrightarrow \infty$).

Theorem A.1: *For every irreducible stationary Markov chain P, there exists an $n \times n$ matrix P^* with identical rows such that P^t converges (entry-wise) to P^* as $t \to \infty$. Furthermore, the equation $Px = x$, with x being an n-dimensional probability vector has q as its unique solution, $q = (q_1, q_2, ..., q_n)$ being the common row of P^*.*

In other words, the theorem says that $p_{ij}^{(t)} \to q_j$ as $t \to \infty$, for all $i \in S$. We conclude with some standard interpretations. First, in view of this result, each q_i may be viewed as the long-run proportion of time the system spends in state i. Observe that this proportion of time (q_i) is independent of the initial state $x_0 \in S$, for all i. Second, since $Pq = q$, it follows that if the initial state of the process is selected according to the probability distribution q, the next state will also be distributed according to q, and so on, ad infinitum.

[17] $p_{ij}^{(t)}$ is not the same as the t^{th} power of p_{ij}, whence the parenthesis stressing that (t) is only a time index.

References

[1] Conroy, M. and R. Dusansky (1995), The Productivity of Economics Departments in the US: Publications in the Core Journals, *Journal of Economic Literature*, 33, 1966-71.

[2] Dusansky, R. and C. Vernon (1998), Rankings of U.S. Economics Departments, *Journal of Economic Perspectives*, 12(1), 157-70.

[3] Frey, B. and Eichenberger, R. (1993). American and European Economics and Economists. *Journal of Economic Perspectives*, vol. 7, pp. 185-93.

[4] Kalaitzidakis, P., T. Mamuneas and T. Stengos (1999), European Economics: An Analysis Based on Publications in the core Journals, *European Economic Review*, 43, 1150-1168.

[5] Kalaitzidakis, P., T. Mamuneas and T. Stengos (2001), Ranking of Academic Journals and Institutions in Economics, *European Economic Review*, to appear.

[6] Kemeny, J. and Snell L. (1960). *Finite Markov Chains*. Van Nostrand, Princeton, N.J.

[7] Kirman, A. and M. Dahl (1994), Economic Research in Europe, *European Economic Review*, 38, 505-522.

[8] Laband, D. and Piette, M. (1994). The Relative Impacts of Economics Journals, *Journal of Economic Literature*, vol. 32, 640-66.

[9] Liebowitz, S. and J. Palmer (1984). Assessing the Relative Impacts of Economics Journals, *Journal of Economic Literature*, vol. 22, 77-88.

[10] Stigler, G., S. Stigler and C. Friedland (1995), The Journals of Economics, *Journal of Political Economy,* 103, 331-59.