Ranking Journals Using Social Choice Theory Methods: A Novel Approach in Bibliometrics[1]

Fuad T. Aleskerov[2], Vladimir V. Pislyakov[3], Andrey N. Subochev[4]

Abstract
We use data on economic, management and political science journals to produce quantitative estimates of (in)consistency of evaluations based on seven popular bibliometric indicators (impact factor, 5-year impact factor, immediacy index, article influence score, h-index, SNIP and SJR). We propose a new approach to aggregating journal rankings: since rank aggregation is a multicriteria decision problem, ordinal ranking methods from social choice theory may solve it. We apply either a direct ranking method based on majority rule (the Copeland rule, the Markovian method) or a sorting procedure based on a tournament solution, such as the uncovered set and the minimal externally stable set. We demonstrate that aggregate rankings reduce the number of contradictions and represent the set of single-indicator-based rankings better than any of the seven rankings themselves.

Conference Topic
Methods and techniques

Introduction
After almost a century since Gross and Gross published their pioneering work (1927), ranking journals remains a problem. Introduction of the impact factor by Garfield and Sher (1963) ushered in the era of indicators. The emergence of the Scopus database and invention of the h-index (Hirsch 2005) reignited interest in developing various bibliometric measures. However, their growing multiplicity generates two questions:

(a) How do rankings based on different measures correlate with each other?
(b) How can we construct a “harmony” of rankings?

To answer the first question, we apply rank correlation analysis to rankings based on seven popular indicators. We find that all rankings positively correlate with each other, but there is a percentage of contradictions. We see no sufficient reason to presume that any indicator is somehow inferior to others. Therefore instead of trying to choose “the best” indicator, we suggest pooling the information contained in all rankings, even though this information is contradictory. For this purpose, we propose to use ordinal aggregation methods originated in social choice theory. To the best of our knowledge, these methods have never been used to rank journals. Rank correlation analysis confirm that aggregate rankings reduce the number of

[1] The study was financially supported through the Basic Research Program at the National Research University Higher School of Economics (HSE) and by the Russian Academic Excellence Project '5-100'.
[2] DeCAn Lab and Department of Mathematics, Faculty of Economic Sciences, National Research University Higher School of Economics, Moscow, Russia; Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia; alesk@hse.ru
[3] Library, National Research University Higher School of Economics, Moscow, pislyakov@hse.ru
[4] DeCAn Lab and Department of Mathematics, Faculty of Economic Sciences, National Research University Higher School of Economics, Moscow, asubochev@hse.ru

contradictions and represent the set of single-indicator-based rankings better than any of the seven rankings themselves.

Data

We consider three sets of journals representing three academic disciplines: economics, management and political science. Rankings are computed for each set separately. Sets of journals were taken from Journal Citation Reports database from Thomson Reuters, along with their IF, 5-year IF, immediacy index and AI indicators (all for JCR-2011 edition). SNIP and SJR metrics for 2011 were taken from Journal Metrics website powered by Scopus database; h-index for each journal was calculated manually by searching Web of Science database. To make h-index more definite, the exact publication and citation windows have been applied. Only papers appeared from 2007 to 2011 have been considered, and citations to them made during the same period, 2007–2011.

The selection of indicators contains all kinds of metrics. There are un-weighted as well as weighted (AI, SJR) measures. Indicators use different publication windows, from one (immediacy index) to five (5-year IF, AI) years. Moreover, they are taken from different databases. A choice of a database may significantly change the values of indicators even when they are based on the same methodology (Pislyakov 2009). Data sources and properties of metrics are summarized in Table 1.

Table 1. Indicators: sources and properties.

	Database	*Year*	*Publication window, years*	*Weighted*
2-year IF	WoS/JCR	2011	2	No
5-year IF	WoS/JCR	2011	5	No
immediacy index	WoS/JCR	2011	1	No
article influence	WoS/JCR	2011	5	Yes
h-index	WoS	2007–2011 (papers and citations)	5	No
SNIP	Scopus	2011	3	No
SJR	Scopus	2011	3	Yes

After exclusion of publications with missing values, the sets contain 212 economic journals, 93 management science journals and 99 political science journals.

Methods

We consider ranking of journals as a multicriteria decision problem. A classical solution is to apply some aggregation function, for instance a weighted sum, to alternative's criterial values and then rank alternatives by respective values of the function. However, this method has a fundamental deficiency related to its *cardinal* nature. To obtain meaningful results, one has to be sure that all aggregated indicators admit meaningful inter-indicator comparisons. In economics, this problem is known as the problem of interpersonal comparability of utilities. Bergson, Samuelson and Little built the so-called "new" welfare economics upon a postulate of incomparability of individual utilities. Arrow, the father of social choice, adopted this postulate and developed an *ordinal* approach to the aggregation problem (Arrow 1951). We propose to apply ordinal ranking methods from social choice since they are immune to incomparability problem and it is possible to frame any multicriteria decision problem as a social choice problem (Arrow & Raynaud 1986).

Basic notions

One of the main objectives of social choice theory is to determine what alternatives will be or should be chosen given a set of feasible alternatives and preferences of decision-makers (voters, experts). It is possible to transfer social choice methods to a multi-criteria setting if one treats a ranking based on a certain criterion as a representation of preferences of a certain voter. In our case, the set of rankings based on corresponding bibliometric indicators is treated as a profile of opinions of seven virtual experts.

Let A denote the set of feasible alternatives; let N denote a group of experts making a collective decision by vote. A decision is a choice of a subset from A. *Preferences* of a voter i, $i \in N$, are revealed through pairwise comparisons of alternatives and are modeled by a binary relation P_i on A, $P_i \subseteq A \times A$: if voter i prefers x to y, then the ordered pair (x, y) belongs to the relation P_i. If a voter is unable to compare two alternatives or thinks they are of equal value, it will be presumed that he is indifferent regarding the choice between them.

If chooser's preferences are known and a *choice rule* (a mapping of the set of binary relations on A onto the set of nonempty subsets of A) is given, then it is possible to determine what alternatives should be the result of her choice. Thus a social choice problem can be solved if one knows voters' preferences (experts' opinions), defines a binary relation μ, $\mu \subseteq A \times A$, that models social preferences (group's opinion), and determines a social choice rule $S(\mu, A)$: $\{\mu\} \rightarrow 2^A \backslash \varnothing$. Probably the most popular method to construct μ is to apply *the majority rule*: (x, y) belongs to μ if the number of those who think x is better than y is greater than the number of those who think y is better than x: $x\mu y \Leftrightarrow |N_1|>|N_2|$, where $N_1=\{i \in N | \ xP_iy\}$, $N_2=\{i \in N | \ yP_ix\}$. In this case, μ is called *the majority relation*.

The choice of this particular rule of aggregation is prescribed by the social choice theory since the majority rule, and this rule only, satisfies several important normative conditions (May 1952), such as *independence of irrelevant alternatives*, *Pareto-efficienc*y, *neutrality* (equal treatment of alternatives), and *anonymity* (equal treatment of voters).

The majority relation quite often happens not to be a ranking itself since it is generally nontransitive. That is, the majority relation may contain cycles. This result is known as *the Condorcet paradox* (Condorcet, 1785). In order to check if the majority relation is transitive or not and to evaluate how nontransitive it is, we calculate the number of 3-step μ-cycles, 4-step μ-cycles and 5-step μ-cycles for three sets of journals (Table 2).

Table 2. Numbers of 3-, 4- and 5-step μ-cycles for three sets of journals.

	3-step cycles	*4-step cycles*	*5-step cycles*
Economics	2446	22427	226103
Management	203	787	3254
Political Science	149	430	1344

As we see, the Condorcet paradox occurs in all three cases.

The Copeland rule

In order to bypass the nontransitivity problem, several methods have been proposed. Probably, the simplest one is the Copeland rule (Copeland, 1951). The idea behind it is the following: the greater the number of alternatives that are worse than a given one, the better this alternative is (the 2nd version of the Copeland rule); and it is determined through pairwise comparisons whether a given alternative is either better or worse than another one. Alternatively, it could be put that an alternative is good if the number of alternatives that are better is small (the 3rd version of the rule). Finally, one can subtract the number of alternatives that are more (socially) preferable than a given one from the number of alternatives less preferable and then rank alternatives by values of these differences (1st version of the rule). All three versions yield the same result when there are no ties. We used the second and the third versions of the Copeland rule.

A sorting procedure based on tournament solutions

In order to construct a ranking, we can use solutions to the problem of optimal social choice. A solution concept $S(\mu, A)$ is a choice rule that determines a set $B_{(1)}$ of those alternatives that are considered to be social optima: $B_{(1)}=S(\mu, A)$. Let us exclude them and repeat the sorting procedure for the subset $A\backslash B_{(1)}$. The set $B_{(2)}=S(\mu, A\backslash B_{(1)})=S(\mu, A\backslash S(\mu, A))$ contains second best choices, for they are worse than alternatives from $B_{(1)}$ and better than options from $A\backslash(B_{(1)}\cup B_{(2)})$). After a finite number of selections and exclusions, all alternatives from A will be separated by classes $B_{(k)}=S(\mu, A\backslash(B_{(k\text{-}1)}\cup B_{(k\text{-}2)}\cup...\cup B_{(2)}\cup B_{(1)}))$ according to their "quality", and these classes constitute a ranking.

We use two choice rules called *tournament solutions*: the uncovered set (Miller, 1980) and the externally stable set (von Neumann & Morgenstern, 1944; Aleskerov & Kurbanov 1999; Subochev, 2008; Aleskerov & Subochev, 2013). The former is based on the idea of choosing "strong" candidates; the latter chooses candidates from "strong" groups.

We say that an alternative *x covers* (meaning that it is definitively better than) an alternative y if x is (socially) preferred not only to y but also to all alternatives that are less preferable than y: $x\mu y \wedge \forall z\in A, y\mu z \Rightarrow x\mu z$. *The uncovered set UC* is comprised of all alternatives that are not covered by any other alternative.

The concept of a minimal externally stable set operationalizes the idea of a strong group of candidates. A set ES is *externally stable* if for any alternative x outside ES there exists an alternative y in ES that is more preferable (socially) than x: $\forall x\notin ES, \exists y: y\in ES \wedge y\mu x$. An externally stable set is *minimal* if none of its proper subsets is externally stable. An alternative is regarded to be optimal if it belongs to some minimal externally stable set; therefore, the solution is the union of all such sets and is denoted *MES*.

Both *UC* and *MES* are always nonempty and can be calculated through their matrix-vector representations given by Aleskerov and Subochev (2013).

The Markovian method

Finally, we apply a version of a ranking procedure called the Markovian method since it is based on an analysis of Markov chains that model stochastic moves from vertex to vertex via arcs of a digraph representing a binary relation μ. The earliest versions of this procedure were proposed by Daniels (1969) and Ushakov (1971). A similar method has been introduced in bibliometrics by Pinsky and Narin (1976). The detailed description of the procedure is given in (Aleskerov, Pislyakov & Subochev 2014).

The table with ranks of all journals in all rankings can be found in (Aleskerov, Pislyakov & Subochev 2014) as well.

Correlation analysis

To evaluate the (in)consistency of two rankings, we measure their correlation. In this paper, we use the Kendall rank correlation coefficient τ_b. Table 3 visualizes its values for all pairs of rankings, initial and aggregate. The corresponding numerical values of τ_b can be found in (Aleskerov, Pislyakov & Subochev 2014).

In all cases, ranking by values of the immediacy index demonstrates the lowest level of correlation with single-indicator-based rankings. This is possibly due to a very narrow publication window that this indicator is based on. In all cases, rankings based on the 5-year impact factor demonstrate the highest level of correlation among single-indicator-based rankings. In the previous study (Aleskerov et al. 2011), the most correlated ranking was one based on the classic impact factor, the 5-year impact being the second best. Systematic differences between rankings based on other indicators are not observed.

Direct observations of values of τ_b for pairs with an aggregate ranking confirm our previous results (Aleskerov et al. 2011). For each set of journals, all aggregate rankings correlate with any single-indicator-based ranking better than other single-indicator-based rankings do. The only exception is correlation of impact factor with 5-year IF, which is a bit higher than correlation with aggregate rankings. This is not true for 5-year IF, though. Formal comparisons based on majority rule (see (Aleskerov, Pislyakov & Subochev 2014) for details) confirm direct observations. In all cases, almost all aggregate ranking methods produce rankings that represent the set of single-indicator-based rankings better than any of these seven. Therefore replacing the set of seven single-indicator-based rankings with aggregate rankings is justified, the best method producing the most representative rankings being the third version of the Copeland rule.

Conclusion

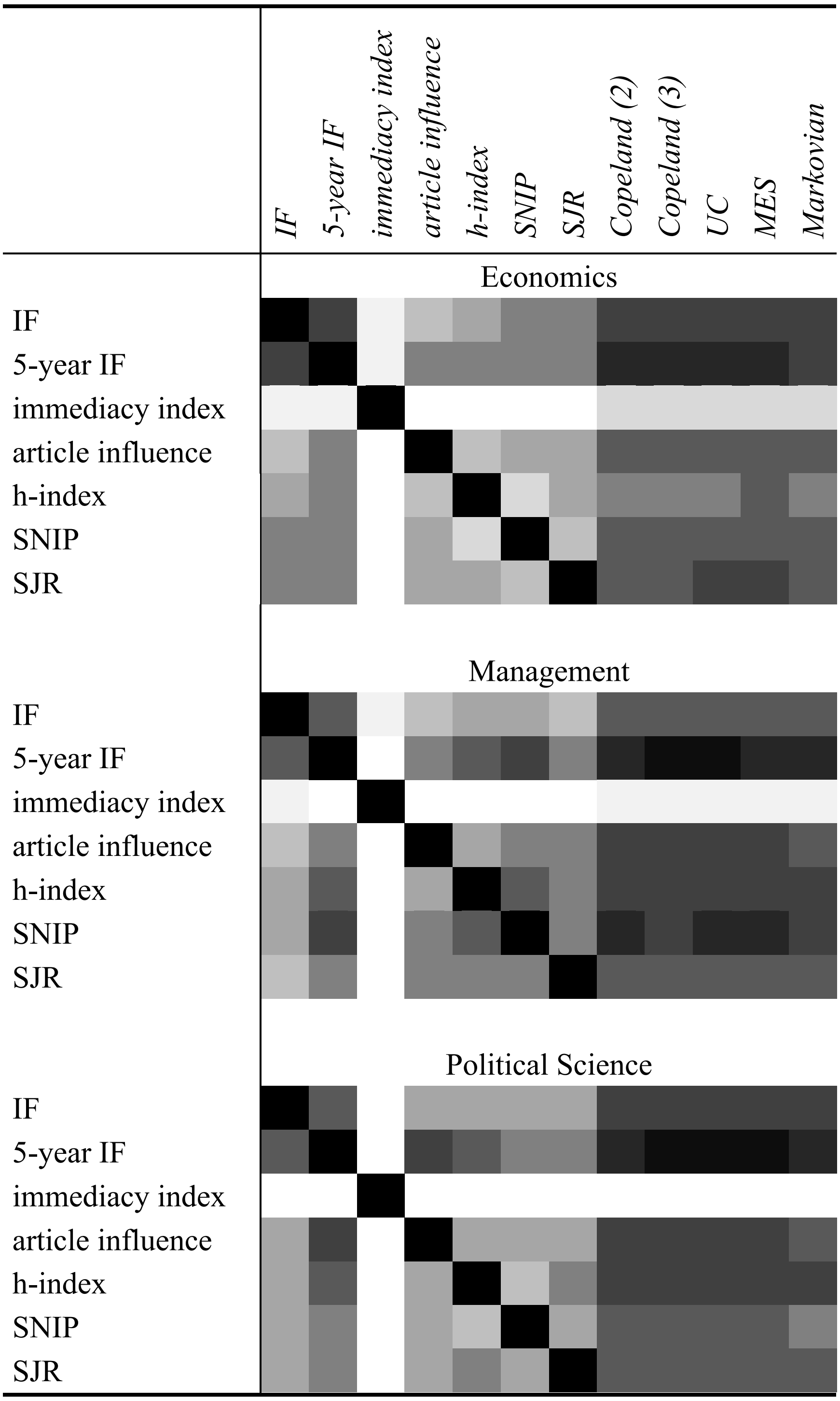
Measuring journal influence is a problem that has no clear-cut solution. Different approaches lead to different indicators, and each possesses its own justification. We took the values of seven popular bibliometric indicators as our data. The correlation analysis has shown that the 5-year impact factor is the best choice if one tries to represent seven single-indicator-based journal rankings by one of them. The least correlated are rankings based on the immediacy index. Other indicators are of more or less equal representativeness.

Despite the correlation of single-indicator-based rankings being high, there is a significant number of contradictions. We propose to minimize their number by replacing the set of rankings with an aggregate ranking. Aggregation can be performed in many ways. This report demonstrates the power of ordinal methods borrowed from social choice theory. This is a novel approach in bibliometrics. Ordinal procedures relieve a researcher from the burden of finding appropriate weights and theoretical justifications for arithmetic operations with aggregated variables. The correlation analysis has also shown that aggregate rankings reduce the number of contradictions and represent the set of single-indicator-based rankings better than any of the seven rankings themselves. Thus, aggregate rankings are more efficient instruments for the evaluation of journal influence.

Some of the aggregate rankings (produced by the Copeland rule and the Markovian method) are characterized by a high level of discrimination, and their shares of tied pairs are very small (less than 1%). For instance, the Markovian method discriminate almost all journals. Other rankings (those based on tournament solutions) are rough orderings, which could also be of value. One may even argue that these rough orderings, when many journals are regarded as equal to each other, better represent our intuitive judgments concerning journal influence.

Not all social choice ranking methods have been employed in this study. There are also other tournament solutions. The next logical step would be to widen both the arsenal of aggregation techniques and the set of empirical data.

Table 3. Kendall rank correlation coefficient τ_b visualized through a greyscale (the higher is the value , the darker is the cell; pure white corresponds to τ_b<0.5, pure black – to τ_b>0.95, the scale interval is 0.05).



References

Aleskerov, F. & Kurbanov, E. (1999). Degree of manipulability of social choice procedures. In A. Alkan, Ch.D. Aliprantis N.C. Yannelis (Eds.), *Current Trends in Economics: Theory and Applications* (pp. 13–27). N.Y.: Springer-Verlag.

Aleskerov, F.T., Pislyakov, V.V. & Subochev, A.N. (2014). *Ranking Journals in Economics, Management and Political Science by Social Choice Theory Methods. WP BRP 27/STI/2014.* Moscow: HSE.

Aleskerov, F.T., Pislyakov, V.V., Subochev, A.N. & Chistyakov, A.G. (2011). *Rankings of Management Science Journals Constructed by Methods from Social Choice Theory. Working paper WP7/2011/04.* Moscow: Higher School of Economics. (in Russian).

Aleskerov, F. & Subochev, A. (2013). Modeling optimal social choice: Matrix-vector representation of various solution concepts based on majority rule. *Journal of Global Optimization*, 56(2), 737–756.

Arrow, K.J. (1951). *Social Choice and Individual Values.* New York: Wiley.

Arrow, K.J. & Raynaud, H. (1986). *Social Choice and Multicriterion Decision-Making*. Cambridge (Mass.): MIT Press.

Condorcet, Marquis de. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix.* Paris: L'imprimerie royale.

Copeland, A.H. (1951). *A reasonable social welfare function, Seminar on Application of Mathematics to the Social Sciences, University of Michigan, Ann Arbor*. Mimeo.

Daniels, H.E. (1969). Round-robin tournament scores. *Biometrica*, 56(2), 295–299.

Garfield, E. & Sher, I.H. (1963). New Factors in the Evaluation of Scientific Literature through Citation Indexing. *American Documentation*, 14(3), 195–201.

Gross, P.L.K. & Gross, E.M. (1927). College libraries and chemical education. *Science*, 66(1713), 385–389.

Hirsch, J.E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569–16572.

Miller, N.R. (1980). A new solution set for tournaments and majority voting: Further graph-theoretical approaches to the theory of voting. *American Journal of Political Science*, 24(1), 68–96.

May, K.O. (1952). A set of independent necessary and sufficient conditions for simple majority decisions. Econometrica, 20(4), 680–684.

von Neumann, J. & Morgenstern O. (1944). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Pinski, G. & Narin, F. (1976). Citation influence for journal aggregates of scientific publications: Theory, with application to the literature of physics. *Information Processing and Management*, 12(5), 297–312.

Pislyakov, V. (2009). Comparing two "thermometers": Impact factors of 20 leading economic journals according to Journal Citation Reports and Scopus. *Scientometrics*, 79(3), 541–550.

Subochev, A. (2008). *Dominant, Weakly Stable, Uncovered Sets: Properties and Extensions. Working paper WP7/2008/03.* Moscow: SU – Higher School of Economics.

Ushakov, I.A. (1971). The problem of choosing the preferable object. *Izvestiya Akademii Nauk SSSR. Tekhnicheskaya Kibernetika*, 4, 3–7. (in Russian)