Journal of Informetrics (2017) 416–429

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Regular article

Ranking journals using social choice theory methods: A novel approach in bibliometrics

Andrey Subochev[a], Fuad Aleskerov[a,b], Vladimir Pislyakov[c,*]

[a] DeCAn Lab and Department of Mathematics, Faculty of Economic Sciences, National Research University Higher School of Economics, Moscow, Russia
[b] Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
[c] Library, National Research University Higher School of Economics, Moscow, Russia

ARTICLE INFO

Article history:
Received 23 October 2015
Accepted 8 March 2018

Keywords:
Social choice
Rank aggregation
Majority rule
Ordinal method
Copeland rule
Tournament solution
Externally stable
Multiple criteria
Impact factor
SNIP
SJR

ABSTRACT

We use data on economic, management and political science journals to produce quantitative estimates of (in)consistency of the evaluations based on six popular bibliometric indicators (impact factor, 5-year impact factor, immediacy index, article influence score, SNIP and SJR). We advocate a new approach to the aggregation of journal rankings. Since the rank aggregation is a multicriteria decision problem, ranking methods from social choice theory may solve it. We apply either a direct ranking method based on the 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 the aggregate rankings reduce the number of contradictions and represent the set of the single-indicator-based rankings better than any of the six rankings themselves.

1. Introduction

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

(a) How do the rankings based on different measures correlate with each other?
(b) What can a decision-maker do if there are several rankings but he/she needs just one?

Therefore, we start with the correlation analysis of the journal rankings. This has been done already in a number of comparative studies which were focused either on indicators from different databases (Archambault, Campbell, Gingras, & Larivière, 2009; Delgado & Repiso, 2013; Leydesdorff, 2009), or on citation, network and usage metrics (Bollen, Van de Sompel, Hagberg, & Chute, 2009). The reviews of Waltman (2016), Rousseau (2002) and Glänzel (2003) may serve as an

* Corresponding author.
E-mail addresses: asubochev@hse.ru (A. Subochev), alesk@hse.ru (F. Aleskerov), pislyakov@hse.ru (V. Pislyakov).

https://doi.org/10.1016/j.joi.2018.03.001
1751-1577/

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
SNIP	Scopus	2011	3	No
SJR	Scopus	2011	3	Yes

introduction to the vast literature on citation indicators. In agreement with the previous results, we find that all rankings correlate positively with each other. But our calculations also demonstrate that there is a non-negligible percentage of contradictions.

The multiplicity of contradicting evaluations is a problem for a decision-maker. To make decisions, there should be just one ranking. An obvious solution is to choose the best indicator. Unfortunately (for decision-makers), the academic discussion concerning relative advantages of various indicators has been inconclusive so far. Since there is no compelling reason to presume that one indicator is somehow inferior to the others, it is problematic to make the choice rationally.

The overarching goal of this paper is to relieve a prospective decision-maker from such a necessity. Instead of choosing the best indicator, a decision-maker may choose an appropriate aggregation procedure and use all rankings available. The theory of aggregation is a well-developed area, and, consequently, it allows one to make quite definite conclusions regarding the appropriateness of a choice.

To construct an aggregate ranking is to rank on a basis of multiple criteria. It is well known that there exists a formal analogy between the multicriteria decision-making and social choice (Arrow & Raynaud, 1986). Therefore, a decision-maker may consider the whole panoply of extensively studied and well-behaved social choice procedures. In this paper we propose to use *ordinal* aggregation methods based on the *majority rule*. To the best of our knowledge, none of them has ever been used to aggregate journal rankings.[1] The rank correlation analysis confirms that the aggregates thus obtained reduce the number of contradictions and represent the set of single-indicator-based rankings better than any member of the set.

2. Data

We consider three sets of journals representing three academic disciplines: economics, management, and political science. The rankings are calculated for each set separately. A journal is included if the following two criteria are both satisfied:

(a) both the Journal Citation Reports and the Scopus database classify the journal as either an economic, or management, or political science journal;
(b) values of all six bibliometric indicators are known.

After exclusion of journals with missing values, the sets contain 212 economic journals, 93 management science journals and 99 political science journals. Their lists are given in the extended preprint version of this paper (Aleskerov, Pislyakov, & Subochev, 2014).

Impact factor (IF), 5-year IF, immediacy index (II) and article influence score (AI) were taken from the Journal Citation Reports database (all for JCR-2011 edition). SNIP and SJR metrics for 2011 were taken from the Journal Metrics website powered by the Scopus database.

The main selection criteria for indicators were their popularity and diversity of data sources and methodologies. The latter is particularly important, since it is senseless to aggregate rankings if they are based on identical indicators. In order to capture the relatively vague concept of "journal influence/quality", it seems better to use measures which are as independent and dissimilar as it is possible.

The set of selected indicators contains all kinds of metrics. There are unweighted as well as weighted (AI, SJR) measures. The indicators use different publication windows from one (immediacy index) to five (5-year IF, AI) years. Moreover, they are taken from different databases, since a choice of a database may significantly change the values of indicators even when they are based on the same methodology (Pislyakov, 2009).

The selected indicators are well known and popular among the decision-makers and are calculated by the leaders of the citation database market, Clarivate Analytics (ex-Thomson Reuters IP) and Elsevier. They are easily available for large sets of journals, which is also important for decision-makers. The data sources and the properties of the indicators are summarized in Table 1.

[1] Cook, Raviv, and Richardson (2010) apply a non-majoritarian ordinal aggregation method called the Kemeny-Young rule. We make some comments concerning this rule in the Conclusion.

3. Correlation analysis of bibliometric-indicator-based rankings

To evaluate the (in)consistency of any two rankings, we measure their correlation. For this purpose, we use the Kendall rank correlation coefficient τ_b. The results for all the rankings are shown in Tables A.1 and A.2 (see Appendix A). The thick-bordered upper-left box in each section of Table A.1 contains values of τ_b for pairs of the single-indicator-based rankings. Table A.2 visualizes correlation coefficients on a greyscale.

In all cases, the correlation is positive. The rankings based on the immediacy index demonstrate the lowest level of correlation with the five others. If we exclude the immediacy index, all the values of τ_b are above 0.5, which means the correlation is quite strong. Even for the pairs containing an immediacy-index-based ranking, the values of τ_b do not fall below 0.356. In all cases, the rankings based on the 5-year IF demonstrate the highest correlation with the five others. In our previous study (Aleskerov, Pislyakov, & Subochev, 2011) where 82 management science journals were ranked using bibliometric data for the years 2008–2010, the most correlated ranking was one based on the classic IF, the 5-year-impact-based one being the second best. Systematic differences between the rankings based on other indicators are not observed. Despite the correlation being either moderate or high, the greatest value of τ_b for the initial indicators is 0.830, and all other values are below 0.8. This means the number of contradictions is non-negligible.[2]

4. Aggregation methods and their axiomatic analysis

For the reasons stated above, we prefer not to choose the best indicator and advocate an aggregation approach instead.

A classical solution is to apply some aggregation function, for instance a weighted sum, to alternatives' criterial values and then rank the alternatives by the 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 the 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. For a brief review of their works see Pattanaik (2008). Arrow, the father of social choice theory, 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 the incomparability problem.

4.1. 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 the ranking based on a certain criterion as the preferences of a certain voter. In our case, the set of rankings based on the corresponding bibliometric indicators is treated as a profile of opinions of six virtual experts. Since all the bibliometric indicators which we consider are numerical functions, we may regard them as utility functions representing corresponding rankings/preferences/opinions.

Let X denote the *general set* of alternatives. It is supposed that X is finite. Let A denote the *feasible set* of alternatives: $A \subseteq X \wedge A \neq \emptyset$. The feasible set is a variable. Let N denote a *society* (a group of people) making a collective decision. A *voter* $i \in N$ (a member of the society N) has preferences for the alternatives from X. It is supposed that individual preferences can be represented by a numerical function $u_i(x){:}X \to \mathbb{R}$, such that $u_i(x) > u_i(y)$ if and only if voter i strictly prefers x to y and $u_i(x) = u_i(y)$ if and only if i is indifferent regarding the choice between x and y. The value of $u_i(x)$ is interpreted as the *utility* of alternative x for voter i. The set $U = \{u_i(x) \mid i \in N\}$ of utility functions is called the *utility profile*.

A decision is either a choice of a subset S from A or a ranking R of the alternatives from A. In the former case, there exists a *social choice correspondence* $S(U, A)$ with arguments A and U and values in the set of subsets of A, which determines a set $B_{(1)}$ of social optima in A: $B_{(1)} = S(U, A) \subseteq A$. In the latter case, there exists a *social welfare functional* $R(U, A)$ with arguments A and U and values in the set of binary relations on A, $R(U, A) \subseteq A \times A$. Both social choice correspondences and social welfare functionals are also called *social choice rules*. It is presumed that all S and R depend on A and U only through restriction of U to A, $S(U, A) = S(U|_A)$, $R(U, A) = R(U|_A)$, i.e. social choices and rankings depend on utilities of available alternatives only. Further in the text, we let ourselves a slight abuse of notation and write simply U instead of $U|_A$ whenever this does not lead to confusion. The subrelation $P \subseteq R$, such that $(x, y) \in P \Leftrightarrow (x, y) \in R \wedge (y, x) \notin R$, is called the asymmetric part of R. Relations R and P represent weak and strict social preferences, correspondingly.

[2] A contradiction, or an inversion, is a situation in which a pair of alternatives is ranked differently in two rankings (a > b and b > a), tied pairs not being considered. Our calculations (Aleskerov et al., 2014) show that the percentage of inversions does not fall below 8% for all pairs of the six single-indicator-based rankings in all three cases, and the average value is 21% of inversions.

4.2. Axiomatic properties

Since the values of the bibliometric indicators are interpreted as the utilities of journals, our task is to choose an appropriate $R(U)$. To make a rational choice let us list the conditions it should satisfy.[3]

(C) *Completeness*: $\forall A \subseteq X, \forall x, y \in A, xRy \vee yRx$.

(T) *Transitivity*: $\forall A \subseteq X, \forall x, y, z \in A, xRy \wedge yRz \Rightarrow xRz$.

(N) *Neutrality*: the operations of permuting alternatives' names and aggregating utility functions commute, which means the rule treats all alternatives equally.

(SP) *Strong Pareto principle*: $\forall A \subseteq X, ((\forall x, y \in A, \forall i \in N, u_i(x) \geq u_i(y)) \Rightarrow xRy) \wedge ((\forall x, y \in A, ((\forall i \in N, u_i(x) \geq u_i(y)) \wedge (\exists j \in N: u_j(x) > u_j(y))) \Rightarrow xPy)$.

(A) *Anonymity*: any permutation of the functions in the utility profile leads to the same output, which means the rule treats all voters equally.

(UD) *Unrestricted domain*: the rule can be applied in all cases, that is, to any utility profile U.

(IIU) *Independence of irr elevant utilities*: $\forall A \subseteq X, R(U)|_A = R(U|_A)$

(O) *Ordinality*: if utility profiles U and U' are such that

$$\forall A \subseteq X,\ \forall x, y \in A,\ \forall i \in N,\ (u_i(x) > u_i(y) \Leftrightarrow u'_i(x) > u'_i(y)),\ \text{then } R(U|_A) = R(U'|_A).$$

It is not possible to satisfy all of the listed conditions (Arrow, 1951), therefore something has to be excluded. We cannot exclude (T) and (C), since we need a ranking. The relevance of (N), which ensures equal treatment of journals, and (SP) is straightforward, and they must be preserved as well. Anyway, dropping (SP) while preserving the rest of axioms does not make much sense: such a rule must be either dictatorial (one must choose one indicator and rank by its values) or antidictatorial (one must choose one indicator and rank by its values multiplied by −1, i.e. in the opposite direction) (Wilson, 1972).

That (A), which ensures equal treatment of indicators, should be kept is not evident, since indicators are not voters whose voting rights should necessarily be equal. But since we decided not to adjudicate claims concerning indicators' relative value, the Laplace principle applies. That is, the subjective probability of statements concerning values of different indicators being correct is constant. The anonymity follows.[4]

In our case, a utility profile is a table "journal/indicator value". A table is admissible if there is a citation matrix which generates the table. It is unclear whether the set of all admissible tables is restricted or not and how it is restricted if it is, therefore (UD) should be preserved.

Given (T), (C), (N), (SP), (A) and (UD), the independence of irrelevant utilities is incompatible with the ordinality, consequently, one must choose either (IIU) or (O). If we drop one of the two, the rest of the axioms can be satisfied.

Both axioms are important. (IIU) is a particularly attractive property of an aggregation procedure since it allows one to freely change the set of alternatives which are compared. Suppose we are interested in comparing only the objects from a particular group (e.g. the top-ranked journals) rather than all objects. We may either exclude the irrelevant alternatives and then apply a method of comparison to the rest; or compare all alternatives and then exclude the irrelevant ones. Consequently, we may obtain two different results. (IIU) solves this problem. Moreover, if the feasible set is changed, previous comparisons will remain valid.[5] Thus, (IIU) is a consistency requirement.

As it has been already stated, (O) is introduced in order to resolve the problem of incomparability of utilities of different persons. Roughly speaking, we may not know the utility substitution rates; consequently, if the utility of person i decreases, we are unable to keep the social welfare constant by increasing the utility of person j. Under (O), we do not presume that any quantities have any meaning, only order of numbers is meaningful, each voter has an ordinal scale of measurement, and these scales are incomparable. Thus, (O) is also called the axiom of ordinal noncomparability. It can be reformulated in such a way:

Ordinal Noncomparability: if utility profiles U and U' are such that $\forall i \in N, \exists g_i(u): \mathbb{R} \to \mathbb{R}$ – an arbitrary strictly increasing function, such that $\forall x \in X, u'_i(x) = g_i(u_i(x))$, then $\forall A \subseteq X, R(U|_A) = R(U'|_A)$. That is, $R(U)$ should be invariant under all monotonic transformations applied to utility functions separately.

The cardinal noncomparability is a weaker condition. It demands that $R(U)$ should be invariant only under any affine rescaling of utilities applied to utility functions separately.

(CN) *Cardinal Noncomparability*: if utility profiles U and U' are such that

$$\forall i \in N,\ \exists a_i, b_i \in \mathbb{R}:\ \forall x \in X,\ u'_i(x) = a_i\, u_i(x) + b_i,\ \text{then } \forall A \subseteq X,\ R(U|_A) = R(U'|_A).$$

[3] Most axioms from this list as well as the general statements about functionals satisfying a particular subset of the listed conditions are borrowed from Fleurbaey and Hammond (2004).

[4] It is also possible to adapt our framework to a case in which individual voters are not equal. The values attributed to their opinions can be reflected in weights of the voters or, equivalently, in the number of votes each voter is allowed to cast. The property of anonymity stays as the *anonymity of votes*: each vote has the same value no matter who casts it.

[5] A complete theoretical analysis of ranking aggregation procedures satisfying Arrowian IIA has been presented in Aleskerov (2002).

(CN) implies that it is possible to meaningfully compare differences of values of a given indicator. Under (CN), each voter has a cardinal scale of measurement without fixed unit and origin, but these scales remain incomparable. Unfortunately, moving from (O) to (CN) does not give one anything at all since (CN) is equivalent to (O) under (N), (IIU) and (UD) (d'Aspremont & Gevers, 1977). Thus to preserve (IIU) one must suppose some comparability of indicators. Unfortunately, we don't know how to make meaningful comparison of indicators' scales; moreover, we don't know if such comparisons are possible at all.

A rule satisfying (O) allows one to obtain aggregated evaluations of alternatives without recourse to arithmetic operations with different criteria, and consequently removes the problem of their theoretical justification. Though the use of ordinal methods involves some loss of information, we believe the information lost is either useless or irrelevant to a typical decision-maker interested in ranking journals.

Thus, it is (IIU) that should be dropped. This opens several possibilities. We advocate the use of the methods which are based on the majority rule. Our arguments are the following.

It is possible to satisfy all listed axioms if we drop (C). There is just one rule which does this. It is the Pareto rule defined by (SP) alone (Weymark, 1984). The obvious deficiency of this rule is that the number of ties in the Pareto dominance relation is high when the correlation of individual rankings is imperfect. This is because the Pareto rule makes insufficient use of the available information. Imposing an additional axiom of positive responsiveness can refine the Pareto dominance relation.

(PR) *Positive responsiveness*: if utility profiles U and U' are such that

$$\exists j \in N:\ (u_j(x) < u_i(y) \wedge u'_j(x) \geq u'_j(y)) \vee (u_j(x) = u_j(y) \wedge u'_j(x) > u'_j(y)) \text{ and}$$

$$\forall i \in N\backslash\{j\},\ u'_i(x) = u_i(x) \wedge u'_i(y) = u_i(y) \text{ and } xR(U)y \text{ and } yR(U)x \text{ then } xP(U')y \text{ for any pair } x \text{ and } y.$$

(PR) is the requirement to use all the information available under (N) and (A) to resolve ties in the Pareto dominance relation. May (1952) proved that the only aggregation rule which satisfies (UD), (N), (A) and (PR) is *the (simple) majority rule*: $x\,R\,y \Leftrightarrow |\{i \in N \mid u_i(x) > u_i(y)\}| \geq |\{i \in N \mid u_i(y) > u_i(x)\}|$. Obviously, this rule also satisfies (IIU), (O) and (SP). The latter means the Pareto dominance relation is a subrelation of *the majority relation*, which is the outcome of the majority rule. In addition, when there are only two alternatives in the feasible choice set, the majority rule trivially satisfies (T).

There are more arguments in favor of the majority rule. For instance, this rule satisfies the following conditions:

(CM) *Cardinal monotonicity*: if utility profiles U and U' are such that

$$\forall i \in N,\ u'_i(x) \geq u_i(x) \wedge u'_i(y) = u_i(y),\ \text{then } xR(U)y \Rightarrow xR(U')y.$$

Under (O), this property is equivalent to ordinal monotonicity, which is equivalent to Smith's monotonicity (Smith, 1973).

(OM) *Ordinal monotonicity*: if preference profiles (i.e. sets of individual preferences) Π and Π' are such that

$$\forall i \in N,\ xR_iy \Rightarrow xR_i'y \text{ and } xP_iy \Rightarrow xP_i'y,\ \text{where } R_i(P_i) \text{ are individual preferences of voter } i,\ \text{then } xR(U)y \Rightarrow xR(U')y.$$

(SCM) *Strict cardinal monotonicity*: if utility profiles U and U' are such that

$$\forall i \in N,\ u'_i(x) \geq u_i(x) \wedge u'_i(y) = u_i(y),\ \text{then } xR(U)y \Rightarrow xR(U')y \text{ and } xP(U)y \Rightarrow xP(U')y.$$

(CS) *Computational simplicity*: there exists a polynomial algorithm for computing $R(U)$.

Moreover, the majority relation minimizes the average number of contradictions with the rankings represented by the utility profile. Since the value of the Kendall rank correlation coefficient τ is just the renormalized number of contradictions between two rankings (the Kendall distance), the majority relation is the best correlate of the set of initial rankings.

Finally, if individual preferences are not subjective views but rather objective judgments concerning the state of affairs,[6] the Condorcet Jury Theorem can be applied. According to this theorem, results of majority rule-based binary comparisons are more likely to be true than any individual judgments concerning pairs of alternatives (Condorcet, 1785; Young, 1988).

Thus, if the majority relation is a ranking then it is clearly the best candidate for the aggregate. Unfortunately, the majority relation is generally nontransitive. For instance, the asymmetric part P of it may contain cycles. This result is known as *the Condorcet paradox* (Condorcet, 1785). To bypass the nontransitivity problem, it was proposed either to violate (UD) and consider only those profiles which yield only transitive relations, or to "mend" the nontransitive relation and, consequently, violate (IIU).

We have calculated the number of 3-step P-cycles, 4-step P-cycles and 5-step P-cycles for three sets of journals (Table 2) to check if the majority relation is transitive or not and to evaluate how nontransitive it is in our case.

As we see, the Condorcet paradox occurs in all three cases. Therefore, we are looking for a rule which is based on the majority relation and satisfies (T), (C), (N), (SP), (A), (UD), (O), (CM), (OM), (CS) plus an additional requirement of *the Condorcet consistency* (CC): the choice rule $R(U)$ must yield the majority relation, whenever the latter is a ranking.

[6] It is the case with all bibliometric indicators, since they are all proposed as measures of journal's influence, which is supposed to be an objective property.

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

	3-step cycles	4-step cycles	5-step cycles
Economics	167	822	3140
Management	19	36	57
Political Science	21	58	142

4.3. The Copeland rule

Probably, the simplest such method is the Copeland rule (Copeland, 1951). The idea behind it is the following: the greater the number of alternatives which 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. That is, we give 1 to any alternative for each win and 0 for each tie and each loss, add the numbers and rank alternatives by their respective score. Alternatively, it could be put that an option is good if the number of alternatives which are better is small (the 3rd version of the rule). That is, we give 1 to any alternative for each win and each tie and 0 for each loss. 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 the alternatives by the values of these differences (the 1st version of the rule). The latter is equivalent to giving 1 for each win, 0 for each loss and ½ for each tie. In fact, there is a continuum of versions of the Copland rule, each version being characterized by the value of a tie, which must necessarily be a number in the interval [0; 1], where a win is valued as 1 and a loss as 0. All versions yield the same ranking when there are no ties. In this paper, we used the second and the third versions of the Copeland rule, since they represent the borders of the interval and, consequently, are the most dissimilar versions of this rule.

The Copeland rule satisfies (T), (C), (N), (SP), (A), (UD), (O), (CM), (OM), (SCM), (CS) and (CC). This rule does not satisfy (IIU) and, consequently, the Arrowian axiom of Independence of Irrelevant Alternatives, which is a conjunction of (IIU) and (O).

Arrowian Independence of Irrelevant Alternatives (AIIA): Let R_i and R'_i denote the individual preference relation represented by $u_i(x)$ and $u'_i(x)$, respectively. Then $\forall A \subseteq X$, $\forall x, y \in A$,

$$(\forall i \in N,\ xR_iy \Leftrightarrow xR'_iy \wedge xP_iy \Leftrightarrow xP'_iy) \Rightarrow (xR(U|_A)y \Leftrightarrow xR(U'|_A)y \wedge xP(U|_A)y \Leftrightarrow xP(U'|_A)y).$$

(AIIA) states that the social ranking of x versus y depends only on ordinal binary comparisons of x with y. That is, if the values of indicators or the set of journals change so that this does not affect the position of journals x and y relative to each other in any indicator-based ranking, then the position of x relative to y in the aggregate ranking must not change.

Although the Copeland rule does not satisfy (AIIA), it satisfies the following weakening of (AIIA) (Rubinstein, 1980).

Weak Arrowian Independence of Irrelevant Alternatives (WAIIA): Suppose the feasible choice set A is fixed. Then $\forall x, y \in A$, $(\forall i \in N, \forall z \in A, (xR_iz \Leftrightarrow xR'_iz) \wedge (xP_iz \Leftrightarrow xP'_iz) \wedge (yR_iz \Leftrightarrow yR'_iz) \wedge (yP_iz \Leftrightarrow yP'_iz)) \Rightarrow ((xR(U|_A)y \Leftrightarrow xR(U'|_A)y \wedge (xP(U|_A)y \Leftrightarrow xP(U'|_A)y))$.

says that if the feasible choice set A does not change, the social ranking of x versus y depends only on ordinal binary comparisons of x and y with other alternatives from A. That is, if the set of journals *stays the same*, and the values of indicators change so that this does not affect the position of journals x and y relative to each other *and to any other journal* in any indicator-based ranking then the position of x relative to y in the aggregate ranking will not change.

4.4. A sorting procedure based on tournament solutions

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

In line with our general approach, we consider social choice rules S that depend on the utility profile U through the majority relation $P(U)$: $S = S(P(U), A) \subseteq A$. Furthermore, it is presumed that S depends on A and P only through restriction of P to A: $S = S(P|_A) \subseteq A$,[7] i.e. social choices are dependent on social preferences for available alternatives only. A *tournament solution* is a social choice correspondence $S(P)$ that has the following properties:

[7] The restriction $P|_A$ is defined unambiguously, since the majority rule satisfies both (IIU) and (O).

1. *Nonemptiness*: $\forall A \subseteq X \wedge A \neq \emptyset, \forall U, S(P|_A) \neq \emptyset$;
2. *Neutrality*: permutation of alternatives' names and social choice commute, i.e. the rule treats all alternatives equally;
3. *Condorcet consistency*: Let $w \in A$ be called *the Condorcet winner* if w is strictly preferred to any other alternative in A, $\forall y \in A$, wPy. If there exists the Condorcet winner w for $P|_A$ then $S(P|_A) = \{w\}$.

We use two tournament solutions: the uncovered set (Miller, 1980) and the minimal externally stable set (Aleskerov & Kurbanov, 1999; Aleskerov & Subochev, 2013; Subochev, 2008; von Neumann & Morgenstern, 1944; Wuffle, Feld, Owen, & Grofman, 1989) as examples. The former is based on the idea of choosing "strong" candidates; the latter chooses candidates from "strong" groups.

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

The concept of a minimal externally stable set operationalizes the idea of a strong group of candidates. A subset B of the feasible set A is *externally stable* if for any alternative x outside B there exists an alternative y in B that is more preferable (socially) than x: $\forall x \in A \backslash B, \exists y: y \in B \wedge yPx$. 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 in polynomial time through their matrix-vector representations given by Aleskerov and Subochev (2013). Therefore, the *UC* or *MES*-based sorting satisfies (CS).

Any sorting based on a tournament solution obviously satisfies (T), (C), (N), (A), (UD), (O) and (CC). It also satisfies the following weakening of (AIIA). Let us first compute the aggregate ranking and then exclude all journals ranked higher than an arbitrary rank r. Alternatively, let us first exclude all these journals from the feasible set and compute the aggregate ranking for the rest. Obviously, any sorting based on a choice correspondence will yield the same result in both cases.

If x Pareto-dominates y then x covers y in any A (Miller, 1980), consequently, the *UC*-based sorting also satisfies (SP). *MES* may select Pareto-inefficient options, therefore the *MES*-based sorting violates (SP) but it satisfies the

$$\textit{Weak Pareto principle}: \ \forall A \subseteq X, \ \forall x, \ y \in A, \ (\forall i \in N, \ u_i(x) \geq u_i(y)) \Rightarrow xRy. \tag{WP}$$

But *MES* satisfies the *generalized Nash Independence of Irrelevant Alternatives* (Subochev, 2017), while *UC* does not:

(GNIIA) : $\forall A \subseteq X, \ \forall B \subseteq X, \ S(A) \subseteq B \subseteq A \Rightarrow S(U, B) = S(U, A)$ (Nash, 1950).

Due to *MES* satisfying (GNIIA), the *MES*-based sorting yields consistent results not only when one excludes the alternatives from the top of the aggregate ranking but the bottom ones as well. Thus, though the *MES*-based sorting does not satisfy (IIU) and (AIIU), this rule satisfies a similar property, which can be called

(IICA) *the Independence of Irrelevant Classes of Alternatives*.

(IIU) says that *the utility profile-based* social ranking procedure must commute with the operation of restriction of the utility profile to the feasible set of alternatives. (AIIA) says that the *ordinal preference-based* social ranking procedure must commute with the operation of restriction of the preference profile (the set of binary relations R_i representing individual preferenses of members of the society N) to the feasible set of alternatives. (IICA) says that the ordinal preference-based social ranking procedure must commute with the operation of restriction of the preference profile to the feasible set of alternatives, *when all alternatives of the same aggregate rank are excluded*. That is, if we treat the indifference classes generated in A by $R(U)$ as new alternatives, then (IICA) means $R(U)$ satisfies an analogue of (AIIA) with respect to these «alternatives».

In our case, (IICA) seems to be more preferable than (WAIIA), since a researcher or a practitioner may wish or need to change the set of journals considered.

Both *UC* and *MES* are *P*-monotonic (Subochev, 2017), i.e. monotonic with respect to the collective preferences P, which means a top-ranked alternative x remains top-ranked if it moves up in the individual rankings, while the preference profile restricted to $A\backslash\{x\}$ remains the same. Since *MES* also satisfies (GNIIA), the *MES*-based selection procedure also satisfies (CM) and (OM), while *UC*-based sorting does not (Bouyssou, 2004). Both procedures satisfy neither (SCM), nor Smith's strong monotonicity axiom (Smith, 1973), but the latter condition is over-demanding, obviously.

4.5. The Markovian method

Finally, we apply a version of a ranking procedure called the Markovian method since it is based on the Markov chains that model stochastic moves from node to node via arcs of the digraph which represents a binary relation R. The earliest versions of this procedure were proposed by Daniels (1969) and Ushakov (1971). The version we use may be briefly described as follows.

First, let us sort the alternatives by the tournament solution called the weak top cycle *WTC* (Good, 1971; Schwartz, 1970, 1972, 1977; Smith, 1973; Ward, 1961). A subset *WTC* of the feasible set A is the weak top cycle of A if any alternative in *WTC* P-dominates any alternative outside *WTC*: $\forall x \in A\backslash WTC, y \in WTC \Rightarrow yPx$, and none of *WTC*'s proper subsets satisfies this property. After the sorting, we rank alternatives of the same sort in the following way. Let us consider the digraph G that represents the restriction of R to the set of alternatives $B_{(k)}$ of a given sort k. Let us consider the following Markov process of movement around G. Time is a sequence of moments $\{t_n\}$ indicated by a natural number n. At moment t_n, we are at some node x. With equal probability, we select a node from $B_{(k)}\backslash\{x\}$. Let it be y. If yRx then we move to y, otherwise we stay in x at t_{n+1}. The alternatives are ranked by the corresponding values $p(x) = \lim_{n\to\infty} p_n(x)$, where $p_n(x)$ is the probability that we are at node x at t_n. It can be proved that all $p(x)$ are independent of any $p_0(x)$, consequently, the initial conditions are irrelevant. The detailed description of this procedure is given by Aleskerov et al. (2014). Though in a different version, this method has been introduced in bibliometrics by Pinski and Narin (1976). Their version differs from the one described above in how the next contestant is selected. According to Pinski and Narin, if we are at node x at t_n another node y must be selected from the set $\{y \mid yRx\}$, consequently we are always obliged to move from x to y at t_{n+1}.

The version of the Markovian procedure used in this paper satisfies (T), (C), (N), (SP), (A), (UD), (O), (CM), (OM), (SCM), (CS), (CC) and does not satisfy (IIU), (WAIIA), (IICA). It should be noted that the Pinski-Narin version does not satisfy cardinal/ordinal monotonicity requirement (Kóczy & Strobel, 2007).

Table A.3 shows positions in the aggregate rankings for the ten journals with the highest IF. Complete rankings may be found in (Aleskerov et al., 2014).

5. The correlation analysis of the aggregate rankings

Let us consider the correlation of the aggregated rankings with each other and with the original six ones. The direct observation of the corresponding values of τ_b in Table A.1 and its visualization in Table A.2 confirm our supposition and our previous results (Aleskerov et al., 2011). For each set of journals, any single-indicator-based rating correlates better with any aggregate ranking than with any other single-indicator-based one. There are just two exceptions. In the case of management journals, the II-based ranking correlates with the IF-based ranking better than with the aggregates. In the case of political science journals, the AI-based ranking correlates with the 5-year IF-based ranking better than with the aggregates. This is not true for the IF and 5-year IF-based rankings, though. They correlate better with the aggregates than with any indicator-based rankings.

Let us employ the same idea of binary multicriteria comparisons to evaluate the rankings more formally. The problem of aggregation can be reformulated as a choice of a single object representing a given group of objects. In our case, we need to choose a ranking that serves as the best representative for the set of rankings based on the six bibliometric indicators. We have eleven candidates: the five aggregates and the six prime rankings themselves. If the six prime rankings were the preferences of some six voters, then we would expect that in a binary contest a voter would vote for a representative whose preferences are closer to his own. Let us again use the majority rule to determine the best representations. Let us say that ranking R_1 represents a given set of rankings better than ranking R_2 if R_1 is better correlated with the majority of rankings from this set than R_2. In our case, each ranking is characterized by the 6-tuple, its i-th component being the value of τ_b for this ranking and a corresponding single-indicator-based ranking. We compare these 6-tuples, compute the corresponding voting matrix[8] and the tournament matrix representing the majority relation on the set of the eleven rankings compared. We do this for each of the three journal sets separately.

Then we combine the results of all comparisons. For each pair of methods, we have $3 \times 6 = 18$ comparisons based on the proximity of the rankings obtained by these two methods to a single-indicator-based ranking with respect to a τ_b. For all three cases (sets of journals) and all six indicators, let us count how often method M_a "wins" over method M_b, $a = 1 \div 11$, $b = 1 \div 11$, that is, how often a ranking produced by method M_a happens to be closer to a single-indicator-based ranking than a ranking produced by method M_b. Let us say that method M_a performs generally better than method M_b, that is, M_a produces better representations of the sets of rankings based on six selected bibliometric indicators than M_b does, if M_a "wins" over M_b more often than M_b "wins" over M_a. That is, we calculate the sum of the three voting matrices obtained earlier and compute its corresponding tournament matrix.[9] We also compute the tournament matrix for the results of our previous study (Aleskerov et al., 2011).

If we apply the Copeland rule (version 2) to the tournament matrices obtained, we will get the five rankings of ranking methods. These rankings are presented in Table 3. The methods that produce rankings which better represent the group of the six single-indicator-based rankings are ranked higher. Please note that this is an *ex post* analysis.

The following observations can be made concerning the robustness of rankings with respect to the choice of the aggregation method. All rankings almost coincide with their majority relations.[10] Therefore, any neutral and Condorcet-consistent ranking method based on the majority rule will place the eleven rankings in almost the same order. In all five cases, any such method will place sorting by *UC* or *MES* and Copeland above Markov above 5-year impact above AI/IF/SJR/SNIP above Immediacy. Thus, we may conclude that the results of our comparisons are robust with respect to the choice of the aggregation method if it is a ranking procedure based on the majority rule.

[8] An entry v_{xy} of the voting matrix **V** contains the number of votes cast for candidate x in the binary contest where x's opponent was y.

[9] The scheme very much resembles the competition of nations during, say, the Olympic games. Methods are like nations. Rankings produced by methods are like sportsmen representing nations. Cases are like different sports (say, tennis, soccer and ping pong). "Wins" are points that sportsmen add to their national collection.

[10] In the majority relations, SJR is tied with AI for economic journals, AI is tied with IF for management journals, MES is tied with Copeland3 and SJR is tied with AI for political science journals, SNIP is tied with AI for all three sets combined.

Table 3
The Copeland ranking of rankings compared by Kendall's τ_b.

rank	Economics R_1	Management R_2	Political Science R_3	All 3 sets combined R_4	Previous results (2008) R_5
1	*UC*	*UC* / *MES*	*UC*	*UC*	*UC*
2	*MES*		*MES*	*MES*	*MES*
3	Copeland 2 / Copeland 3	Copeland 3	Copeland 2 / Copeland 3	Copeland 3	Copeland 3
4		Copeland 2		Copeland 2	Copeland 2
5	Markov	Markov	Markov	Markov	Markov
6	5-y. impact	5-y. impact	5-y. impact	5-y. impact	Impact
7	Impact	SNIP	Impact	Impact	5-y.impact
8	SJR	AI	SJR	AI / SJR	SJR
9	AI / SNIP	Impact / SJR	AI / SNIP		AI / SNIP
10				SNIP	
11	Immediacy	Immediacy	Immediacy	Immediacy	Immediacy

The formal comparison confirms the direct observations. In all five cases, the II-based ranking demonstrates the lowest level of correlation with the other single-indicator-based rankings. In all the cases except one related to the older data, the rankings based on the 5-year IF demonstrate the highest level of correlation among the single-indicator-based rankings. In the previous study, the most correlated ranking was one based on the classic IF, the 5-year impact being the second best. Systematic differences between rankings based on the other indicators are not observed.

In all five cases all aggregate rankings were placed above the initial six. This supports our assertion that the aggregation based on the majority rule produces rankings that represent the set of single-indicator-based rankings better than any of these six. Therefore replacing the set of six single-indicator-based rankings with aggregates is justified, at least for the datasets considered. Judging from Table 3, the best aggregation method seems to be the sorting by either *UC* or *MES*.

6. Conclusions

How to measure the influence of a journal is a problem that has no clear-cut solution. Different approaches lead to different indicators, and each has its own justification. We take the values of the six popular bibliometric indicators as our data. For the given datasets, the correlation analysis shows that the 5-year IF is the best choice if one tries to represent six 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.

Though the correlation of single-indicator-based rankings is high, there is a significant number of contradictions. Instead of choosing the best indicator, we consider the choice of an appropriate aggregation procedure. The aggregation can be performed in many ways. The framework of social welfare functionals borrowed from social choice theory allows us to consider and compare both ordinal and cardinal procedures. We discuss a set of relevant properties that an aggregation method should satisfy for ranking journals. Unfortunately, it is impossible to find a procedure that satisfies all of them. Ultimately, it is up for a decision-maker to determine what properties might be violated in a particular situation. In our case, the real choice is between just two axioms: the ordinal non-comparability (O) and the independence of irrelevant utilities (IIU). We argue in favor of violating (IIU) and preserving (O), though we recognize the importance of (IIU). As we have said already, we suggest the aggregation because we are (yet) unable to choose the best indicator rationally. Similarly, we suggest to preserve the axiom of ordinal non-comparability due to our inability to compare indicators' scales. When the scales of indicators are comparable, either cardinally or ordinally, (O) may and must be dropped, and cardinal or ordinal procedures satisfying (IIU) should be used instead. Until then, we are forced to choose those methods that do not presume any comparability of scales.

This paper demonstrates the power of ordinal methods based on the majority rule. This is a novel approach in bibliometrics. To the best of our knowledge, there was just one aggregation method applied to journal rankings (Cook et al., 2010) which satisfied ordinal non-comparability. This is the Kemeny-Young rule (Kemeny, 1959; Young & Levenglick, 1978) also known as "the minimal violation method" or "the maximal likelihood method" or "the median procedure". This rule prescribes one to use a linear order which is a median of a set of rankings with respect to the Kendall distance (the number of inversions or "violations").

Though the Kemeny-Young rule has its appeal, we argue that majority-rule-based methods on average perform better. Our arguments are the following. First, the Kemeny-Young rule violates the computational simplicity (CS) requirement, since the finding of a median linear order is an NP-complete problem. Second, the Kemeny-Young rule violates the axiom of unrestricted domain (UD), since the median may not be unique. For instance, there could be hundreds of median orders when the number of alternatives is 100. That is, in some instances this rule exacerbates the choice problem rather than solves it. At the same time, all the majority-rule-based methods we consider always produce a unique ranking. Finally, when we compare *only linear orders* as possible representatives of a set of rankings, the order which reduces the average number of contradictions to its minimum is, by definition, a Kemeny-Young median. But it is the majority relation that minimizes the number of contradictions when *all complete asymmetric binary relations* are taken into account. That is, the majority relation represents the set of rankings even better than the Kemeny-Young medians. The problem is that the majority relation may not be a ranking itself. But when the majority relation is a linear order then it is also a Kemeny-Young median, the one and only. Since the methods we consider are nothing more than some ways to "mend" the nontransitive majority relation, it is not surprising that the aggregate rankings, which they produce, reduce the number of contradictions and represent the set of single-indicator-based rankings better than any of the six rankings themselves, as the correlation analysis shows.

Naturally, a practitioner deals with a certain situation rather that with all possible situations. Then the following set of actions may be suggested, which combines both the majority rule and the Kemeny-Young method. First, compute the majority relation. If it is a ranking then use it. If it is not then try to compute the Kemeny-Young median. If you can compute it and it is unique then use the median. If you cannot compute it or the medians are many then compute one of the majority-rule-based rankings and use it.

Though all majority-rule-based rules violate the independence of irrelevant alternatives/utilities, some of them do satisfy some weaker versions of this property: the Copeland rule satisfies the weak Arrowian independence of irrelevant alternatives, while the sorting based on *MES* satisfies the independence of irrelevant classes of alternatives. The second property seems to be more preferable. The drawback of the *MES*-based sorting is that it satisfies the Pareto principle only weakly. There are other tournament solutions though (such as the minimal covering set, the bipartisan set) which, like *MES*, satisfy the Nash independence of irrelevant alternatives and pick up only Pareto-efficient alternatives, but these solutions are somewhat harder to compute.

For the given dataset, the sorting methods outperform all others, when they are compared by the Kendall τ_b.

As Table A.3 shows, the rankings produced by the Copeland rule and the Markovian method are characterized by a high level of journals' discrimination, and the share of tied pairs is very small (less than 1%). For instance, the Markovian method discriminates almost all journals. The sorting-based rankings are rough orderings. But one may even argue that they better represent our intuitive judgment concerning journal influence. Many journals may be considered as equal in their prominence, even though it is highly improbable that there is a pair of journals with identical values of all six indicators.

Author contributions

Fuad Aleskerov: Conceived and designed the analysis, Contributed analysis tools, Wrote the paper.

Vladimir Pislyakov: Conceived and designed the analysis, Collected the data, Wrote the paper.

Andrey Subochev: Conceived and designed the analysis, Contributed analysis tools, Performed the analysis, Wrote the paper.

Acknowledgements

The study was supported by the Basic Research Program at the National Research University Higher School of Economics (HSE) and by the Russian Academic Excellence Project '5-100'. We would like to thank Natalia Maksimova, the Director of the Higher School of Economics Library, for her support of our project and organizing institutional subscription to the bibliometric databases. We also thank Vyacheslav Yakuba, our colleague and friend from the Institute of Control Sciences (Russian Academy of Sciences, Moscow, Russia), for helping us with the computation of the Markovian rankings. The preliminary version of this paper was published as a preprint (Aleskerov et al., 2014) and was presented as a report at the Science and Technology Indicators (STI) Conference in 2016 (Aleskerov, Pislyakov, & Subochev, 2016).

Appendix A.

Table A.1
The Kendall rank correlation coefficient τ_b.

	impact factor	5-year impact factor	immediacy index	article influence	SNIP	SJR	Copeland (2)	Copeland (3)	*UC*	*MES*	Markovian
Economics											
impact factor		0.830	0.503	0.637	0.698	0.700	0.832	0.840	0.845	0.852	0.829
5-year impact factor	0.830		0.510	0.725	0.726	0.741	0.902	0.900	0.905	0.891	0.884
immediacy index	0.503	0.510		0.475	0.454	0.472	0.559	0.562	0.572	0.569	0.560
article influence	0.637	0.725	0.475		0.673	0.674	0.768	0.764	0.771	0.761	0.760
SNIP	0.698	0.726	0.454	0.673		0.638	0.773	0.773	0.780	0.779	0.771
SJR	0.700	0.741	0.472	0.674	0.638		0.785	0.782	0.790	0.785	0.775
Copeland rule (2 v.)	0.832	0.902	0.559	0.768	0.773	0.785		0.962	0.963	0.942	0.942
Copeland rule (3 v.)	0.840	0.900	0.562	0.764	0.773	0.782	0.962		0.968	0.951	0.963
sorting by *UC*	0.845	0.905	0.572	0.771	0.780	0.790	0.963	0.968		0.965	0.948
sorting by *MES*	0.852	0.891	0.569	0.761	0.779	0.785	0.942	0.951	0.965		0.938
Markovian method	0.829	0.884	0.560	0.760	0.771	0.775	0.942	0.963	0.948	0.938	
Management											
impact factor		0.790	0.520	0.641	0.679	0.626	0.809	0.800	0.808	0.799	0.797
5-year impact factor	0.790		0.475	0.743	0.798	0.702	0.904	0.897	0.901	0.904	0.881
immediacy index	0.520	0.475		0.456	0.399	0.391	0.510	0.512	0.516	0.511	0.511
article influence	0.641	0.743	0.456		0.695	0.728	0.803	0.809	0.815	0.811	0.799
SNIP	0.679	0.798	0.399	0.695		0.719	0.824	0.824	0.836	0.843	0.806
SJR	0.626	0.702	0.391	0.728	0.719		0.767	0.778	0.789	0.791	0.766
Copeland rule (2 v.)	0.809	0.904	0.510	0.803	0.824	0.767		0.966	0.972	0.964	0.938
Copeland rule (3 v.)	0.800	0.897	0.512	0.809	0.824	0.778	0.966		0.976	0.968	0.962
sorting by *UC*	0.808	0.901	0.516	0.815	0.836	0.789	0.972	0.976		0.978	0.952
sorting by *MES*	0.799	0.904	0.511	0.811	0.843	0.791	0.964	0.968	0.978		0.944
Markovian method	0.797	0.881	0.511	0.799	0.806	0.766	0.938	0.962	0.952	0.944	
Political science											
impact factor		0.773	0.422	0.671	0.653	0.673	0.812	0.806	0.810	0.815	0.795
5-year impact factor	0.773		0.374	0.835	0.705	0.717	0.896	0.901	0.905	0.896	0.871
immediacy index	0.422	0.374		0.356	0.372	0.398	0.440	0.439	0.446	0.451	0.451
article influence	0.671	0.835	0.356		0.671	0.653	0.820	0.823	0.829	0.830	0.789
SNIP	0.653	0.705	0.372	0.671		0.662	0.753	0.766	0.768	0.762	0.763
SJR	0.673	0.717	0.398	0.653	0.662		0.784	0.770	0.781	0.769	0.763
Copeland rule (2 v.)	0.812	0.896	0.440	0.820	0.753	0.784		0.958	0.963	0.939	0.933
Copeland rule (3 v.)	0.806	0.901	0.439	0.823	0.766	0.770	0.958		0.972	0.952	0.959
sorting by *UC*	0.810	0.905	0.446	0.829	0.768	0.781	0.963	0.972		0.968	0.942
sorting by *MES*	0.815	0.896	0.451	0.830	0.762	0.769	0.939	0.952	0.968		0.930
Markovian method	0.795	0.871	0.451	0.789	0.763	0.763	0.933	0.959	0.942	0.930	

Table A.2
Kendall rank correlation index τ_b visualized through a greyscale (the higher the value, the darker the cell; pure white corresponds to $\tau_b < 0.5$, pure black to $\tau_b > 0.95$; the scale interval is 0.05).

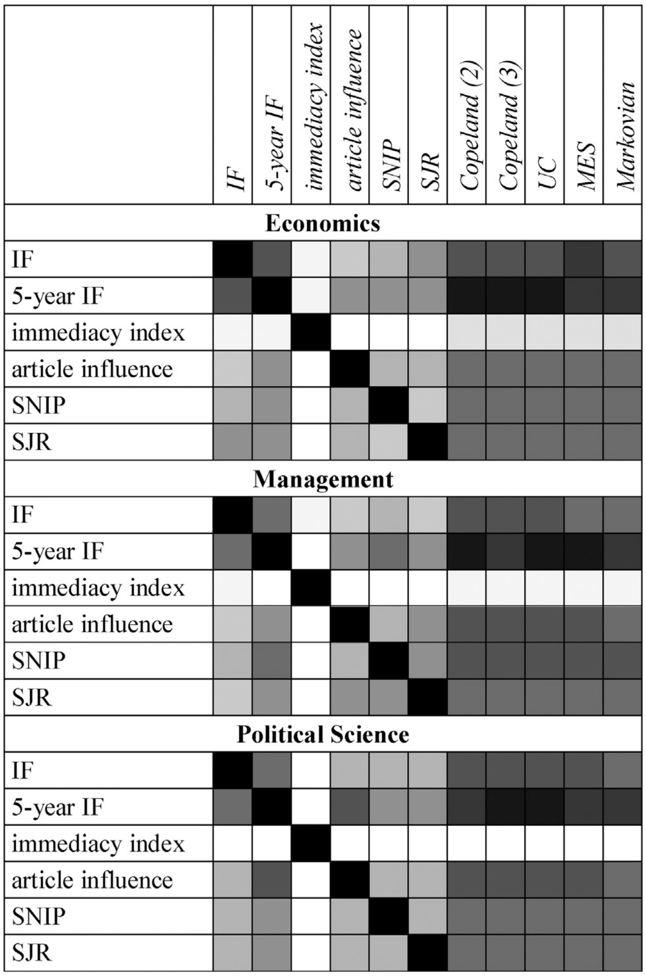

Table A.3
Aggregate ranks of the ten topmost (by impact factor) journals in economics, management and political science.

	IF	Copeland (2)	Copeland (3)	*UC*	*MES*	Markovian
	Economics[a]					
Total number of positions in a ranking	**200**	**136**	**139**	**44**	**46**	**207**
Journal of Economic Literature	1	1	1	1	1	1
Quarterly Journal of Economics	2	2	2	2	2	2
Review of Financial Studies	3	5	5	5	5	5
Journal of Finance	4	3	3	3	3	3
Journal of Economic Perspectives	5	4	4	4	4	4
Economic Geography	6	15	16	8	8	18
Journal of Financial Economics	7	5	5	5	5	5
Brookings Papers on Economic Activity	8	10	9	6	6	11
Journal of Accounting and Economics	9	9	7	6	6	10
Journal of Economic Geography	10	8	6	6	6	7
	Management					
Total number of positions in a ranking	**90**	**61**	**64**	**29**	**30**	**92**
Academy of Management Review	1	1	1	1	1	1
Academy of Management Journal	2	2	2	2	2	2
Academy of Management Learning and Education	3	16	14	8	9	18
Journal of Management	4	3	3	2	2	4
MIS Quarterly	5	3	3	2	2	3

Table A.3 (*Continued*)

	IF	Copeland (2)	Copeland (3)	*UC*	*MES*	Markovian
Journal of Operations Management	6	5	5	4	4	6
Organization Science	7	7	8	5	6	10
Journal of Applied Psychology	8	4	4	3	3	5
Journal of Management Studies	9	9	9	6	6	13
Administrative Science Quarterly	10	5	6	4	5	8
	Political Science					
Total number of positions in a ranking	**95**	**62**	**61**	**34**	**32**	**96**
American Political Science Review	1	1	2	1	1	2
American Journal of Political Science	2	2	2	2	1	2
Public Opinion Quarterly	3	4	3	3	2	4
Journal of Conflict Resolution	4	4	4	4	3	5
Political Analysis	5	1	1	1	1	1
Global Environmental Politics	6	13	12	7	8	19
Politics and Society	7	11	10	7	8	14
Political Geography	8	6	6	5	4	8
Journal of Peace Research	9	5	5	5	4	6
Policy Studies Journal	10	12	7	6	5	12

[a] Note: one journal, *Econometrica*, was not included into the set of economic journals by mistake. Since the majority rule satisfies Arrow's axiom of Independence of irrelevant alternatives (AIIA), our omission does not change the majority relation; consequently, the aggregate rankings should not change much after addition of *Econometrica* to the set.

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., & 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.

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., Pislyakov, V., & Subochev, A. (2014). *Ranking journals in economics, management and political science by social choice theory methods. WP BRP 27/STI/2014*. Moscow: HSE. Retrieved from https://www.hse.ru/data/2014/05/19/1322077930/27STI2014.pdf on 02 June 2017

Aleskerov, F. T., Pislyakov, V., & Subochev, A. (2016). Ranking journals using social choice theory methods: A novel approach in bibliometrics. In *Proceedings of the 21 international conference on science and technology indicators* (pp. 1266–1273).

Aleskerov, F. (2002). Categories of Arrowian voting schemes. In K. Arrow, A. Sen, & K. Suzumura (Eds.), *Handbook of social choice and welfare* (Vol. 1) (pp. 95–129). Amsterdam: Elsevier Science B.V. Handbook of economics series 19.

Archambault, É., Campbell, D., Gingras, Y., & Larivière, V. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology*, *60*(7), 1320–1326.

Arrow, K. J., & Raynaud, H. (1986). *Social choice and multicriterion decision-making*. Cambridge (Mass.): MIT Press.

Arrow, K. J. (1951). *Social choice and individual values*. New York: Wiley.

Bollen, J., Van de Sompel, H., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLoS One*, *4*(6).

Bouyssou, D. (2004). Monotonicity of ranking by choosing: A progress report. *Social Choice and Welfare*, *23*(2), 249–273. http://dx.doi.org/10.1007/s00355-003-0250-x

Condorcet, M. d. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: L'imprimerie royale. Retrieved from http://gallica.bnf.fr/ark:/12148/bpt6k417181 on 20 May 2017

Cook, W. D., Raviv, T., & Richardson, A. J. (2010). Aggregating incomplete lists of journal rankings: An application to academic accounting journals. *Accounting Perspectives*, *9*(3), 217–235.

Copeland, A. H. (1951). *A reasonable social welfare function, seminar on application of mathematics to the social sciences*. Ann Arbor, (Mimeo): University of Michigan.

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

d'Aspremont, C., & Gevers, L. (1977). Equity and the informational basis of collective choice. *Review of Economic Studies*, *44*, 199–209.

Delgado, E., & Repiso, R. (2013). The impact of scientific journals of communication: Comparing Google Scholar metrics, Web of Science and Scopus. *Comunicar*, *21*(41), 45–52.

Fleurbaey, M., & Hammond, P. (2004). *Interpersonally comparable utility*. pp. 1179–1285. *Handbook of utility theory* (Vol. 2) Boston: Springer.

Garfield, E., & Sher, I. H. (1963). New factors in the evaluation of scientific literature through citation indexing. *American Documentation*, *14*(3), 195–201.

Glänzel, W. (2003). *Bibliometrics as a research field: A course on theory and application of bibliometric indicators*. Course Handouts, Leuven.

Good, I. J. (1971). A note on Condorcet sets. *Public Choice*, *10*, 97–101.

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.

Kóczy, L. Á., & Strobel, M. (2007). Ranking of academic journals by tournament methods. In *22nd EEA meeting*.

Kemeny, J. (1959). Mathematics without numbers. *Daedalus*, *88*, 577–591.

Leydesdorff, L. (2009). How are new citation-based journal indicators adding to the bibliometric toolbox? *Journal of the American Society for Information Science and Technology*, *60*(7), 1327–1336.

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

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.

Nash, J. F. (1950). The bargaining problem. *Econometrica*, *18*(2), 155–162.

Pattanaik, P. K. (2008). Social welfare function. In S. N. Durlauf, & L. E. Blume (Eds.), *The new palgrave dictionary of economics. Second edition*. Palgrave Macmillan. Retrieved from https://doi.org/10.1057/9780230226203.1567 on 01 June 2017

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.

Rousseau, R. (2002). Journal evaluation: Technical and practical issues. *Library Trends*, *50*(3), 418–439.
Rubinstein, A. (1980). Ranking the participants in a tournament. *SIAM Journal on Applied Mathematics*, *38*(1), 108–111.
Schwartz, T. (1970). On the possibility of rational policy evaluation. *Theory and Decision*, *1*, 89–106.
Schwartz, T. (1972). Rationality and the myth of the maximum. *NoÛs*, *6*, 97–117.
Schwartz, T. (1977). Collective choice, separation of issues and vote trading. *The American Political Science Review*, *71*(3), 999–1010.
Smith, J. (1973). Aggregation of preferences with variable electorates. *Econometrica*, *41*(6), 1027–1041.
Subochev, A. (2008). *Dominant, weakly stable, uncovered sets: Properties and extensions. Working paper WP7/2008/03*. Moscow: SU – Higher School of Economics.
Subochev, A. (2017). Axiomatic analysis of the union of minimal externally stable sets concerning application and implementation of this novel social choice correspondence. In *Proceedings of the XVII April international academic conference on economic and social development: In 4 volumes. Volume 1* (pp. 111–120) (in Russian).
Ushakov, I. A. (1971). The problem of choosing the preferable object. In *Izvestiya Akademii Nauk SSSR. Tekhnicheskaya Kibernetika*. pp. 3–7. , 4, (in Russian).
von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.
Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, *10*(2), 365–391.
Ward, B. (1961). Majority rule and allocation. *Journal of Conflict Resolution*, *5*, 379–389.
Weymark, J. A. (1984). Arrow's theorem with social quasi-orderings. *Public Choice*, *42*(3), 235–246.
Wilson, R. B. (1972). Social choice without the Pareto principle. *Journal of Economic Theory*, *5*, 478–486.
Wuffle, A., Feld, S. L., Owen, G., & Grofman, B. (1989). Finagle's law and the finagle point, a new solution concept for two-candidate competition in spatial voting games without a core. *American Journal of Political Science*, *33*(2), 348–375.
Young, H. P., & Levenglick, A. (1978). A consistent extension of Condorcet's election principle. *SIAM Journal on Applied Mathematics*, *35*(2), 285–300.
Young, H. P. (1988). Condorcet's theory of voting. *American Political Science Review*, *82*(04), 1231–1244.