

Reliability of Scopus subject classification of journals and its impact on bibliometric research

Nikolaj Lazić

Faculty of Humanities and
Social Sciences

Stjepan Mateljan

Institute for Social Research in Zagreb

Maja Jokić

Institute for Social Research in Zagreb

Funded by





Outline...

- introduction
- database sample
- fields of science
 - Scopus classification
 - OECD classification
 - Croatian classification



- methodology
 - analysis of journals and articles
- results
- conclusion
- future work



Introduction



- this research was conducted within the project “Research activity, collaboration and orientation and social sciences in Croatia and other post-socialist European countries – RACOSS”, financed by Croatian Science Foundation
- the first indicator in the analysis of reliability of bibliographic records taken from the database Scopus, was the classification of the journals. We found Scopus database as more relevant in the field of social sciences than WoS, based on our experience, as well as other researchers’ experiences



Introduction



- bibliometric or scientometric research could be used to evaluate scientific work (and is constantly pushed as something that should be done to measure scientific work in Croatia)
 - problem with fundamental scientific research that is sometimes not recognized
 - “follow the herd” in science research
- will Scopus subject categories influence processing and use of bibliographic data in bibliometric or scientometric research?
 - will it influence policy?



Database sample



- scientific papers from the field of social sciences
 - authors, or at least one author, were scientists from 15 post-socialist European countries (Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia, Bosnia and Herzegovina, Former Yugoslav Republic of Macedonia, Montenegro and Serbia)
 - time frame: 1996-2013.
- recall
 - **104837** documents (article, article in press, review, conference paper, editorial, note, letter, and short survey, book, book chapters and erratum)
 - **83059** scientific papers



Fields of science



Scopus social sciences classification

- social sciences
- business management and accounting
- economics, econometrics and finance
- psychology

OECD social sciences

- psychology
- economics and business
- educational sciences
- sociology
- law
- political science
- social and economic geography
- media and communication
- other social sciences

Croatian social science classification

- economics
- law
- political science
- information and communication sciences
- sociology
- psychology
- pedagogy
- special education
- logopedy
- kinesiology
- demography
- social work
- security studies
- interdisciplinary social sciences



Methodology



- we analysed journals from the obtained sample - we selected journals that mainly deal with the issues of natural or technical sciences, or issues of the area of humanities, with a negligible small share of social sciences
- some journals in the obtained set that almost entirely deal with humanities
 - there could be some justification since at the time we retrieved Scopus data (October 2015), there was a default search limit option for only four major areas: Health sciences, Life sciences, Physical sciences and Social sciences. The category of Social Sciences included Humanities and Arts
 - some fields of Social sciences and Humanities differ significantly in scientific communication, they use different publishing media, publishing dynamics are different, citation behaviour is different, and the results of bibliometric analyses are different, we decided to conduct a detailed analysis of the aim and the scope of each journal included



Methodology



- we've selected **4896** journals with initial set of **83059** papers
- determine the similarity or difference between the Scopus classification of social sciences and the official Classification of social sciences in Croatia
 - big differences between classification systems
- analyse in which group does some journal belong
 - sorting journals according to fields
 - analyse scientific papers and journals



Methodology



- reclassification was done using journal names and using keywords in the journal titles to determine the field (for cca 1500 journals) and looking at journal web pages to determine the field (for remaining 600 journals)
 - some keywords:
 - "^nature\$", "biology", "geograph", "mountain", "geograf", "human.*evolution", "chemistr", "ecolog", "statistic", "environmenta", "natura.*histor", "astrono"
 - "enginee", "computer", "operations", "applied sciences", "architect", "geodet", "tehnicki", "hydropow", "applied.*analy"
 - "med[ie]cine", "health", "psychia", "alzhei", "traum.*stre", "neuropsy", "neurolo", "substan.*misus", "eatin.*disor", "anesth"
 - "agricult", "poljoprivr", "garden", "sericult", "agrarian", "landscape", "forestr"
 - "bogosl", "antropol", "anthropolog", "arch[ae]olo", "prahisto", "english", "iberic", "sefarad", "byzant", "mu[sz]je*k"
 - "^science\$", "scient.*americ", "academy.*sciences", "metadata", "scien.*letter", "reasoning", "royal soci"

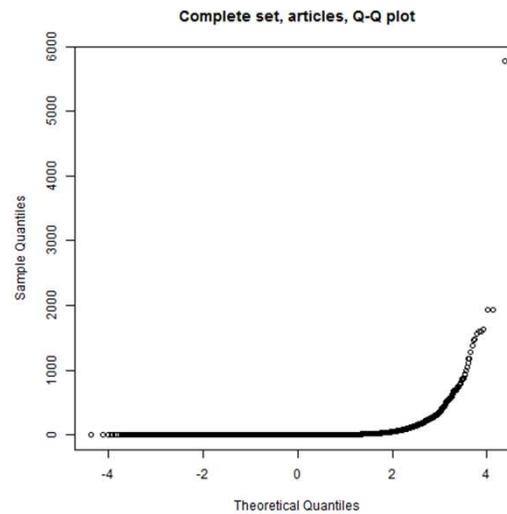


Results



- results from complete set of articles

min	1stQ	median	mean	3rdQ	max
0	0	1	6.686	4	5771





Results



- problem with Scopus's classification of social sciences in our sample is the inclusion of a journals that deals almost exclusively with the issues of natural sciences or biomedicine. As an example, we get journals: Nature, Science, Proceedings of the National Academy of Sciences of the United States of America, Fuzzy Sets and Systems
- 1799 or 2.16% papers with at least one of the co-authors from at least one of the 15 countries from our sample
- this is total of 200,694 citations, which is 36.1% of the total number of citations obtained from papers published in 4896 journals



Results

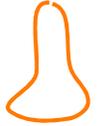


- reclassification of these **4896** journals
 - **2744** (56%) primarily deal with the issues of social sciences according to the mentioned Croatian classification of social sciences
 - remaining **2152** journals (44%) did not belong to social sciences and they were reclassified to Croatian classification of science into: natural sciences, engineering and technology, medical and health sciences, agricultural science and biotechnology, humanities

	min	1stQ	median	mean	3rdQ	max
soc. sc.	0	0	1	5.237	4	1592
non s.s.	0	0	1	7.891	3	5771



Results



- Distribution of the number of journals, papers and citations in a set of journals classified by the relevant experts in the field of social sciences

Subject field	N journals	N papers (1996-2013)	N citations (1996-2015)	citation/paper
Economics	695	11991	48360	4.03
Education	295	2935	9785	3.33
Information and communication sciences	163	2239	19077	8.52
Law	89	885	1088	1.23
Political sciences	158	1224	5098	0.41
Psychology	289	4045	48409	108.78
Sociology	56	1472	3446	2.34
Social work	27	524	1295	2.49
Multidisciplinary social sciences	559	5662	24495	4.32
Multidisciplinary social science & humanities	155	2222	3968	1.78
Multidisciplinary social science & other science fields (medicine, natural sciences, technology etc.)	258	4508	32458	7.2



Results



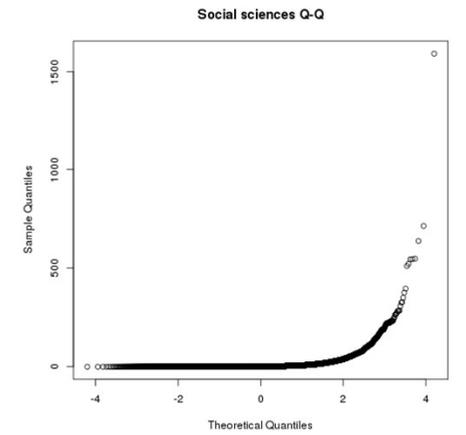
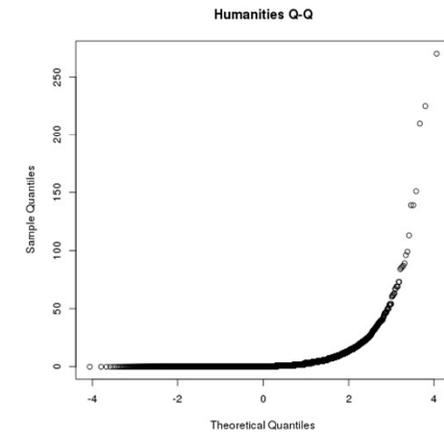
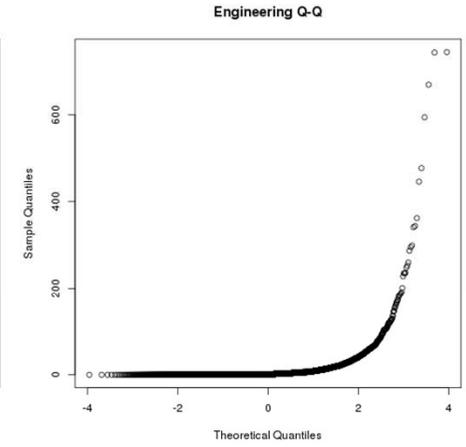
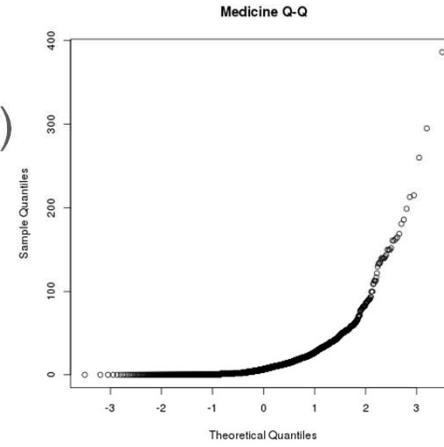
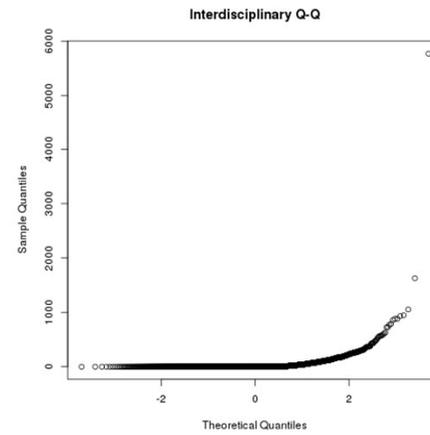
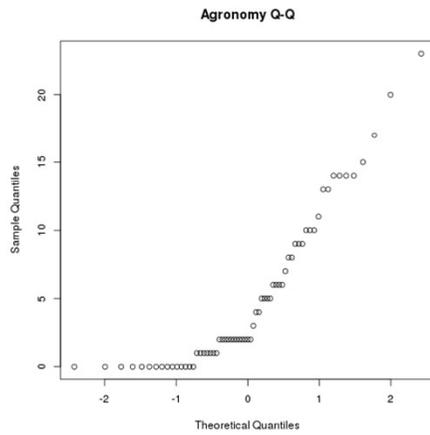
- additional analysis of non social sciences articles

	min	1stQ	median	mean	3rdQ	max
humanities	0	0	0	1.703	1	270
medicine	0	1	7	14.85	18	386
engineering	0	0	1	5.706	5	744
natural sc.	0	0	1	18.04	5	1937
agricultural	0	1	2	5.138	9	23
interdisciplinary	0	0	1	5.237	4	1592



Results

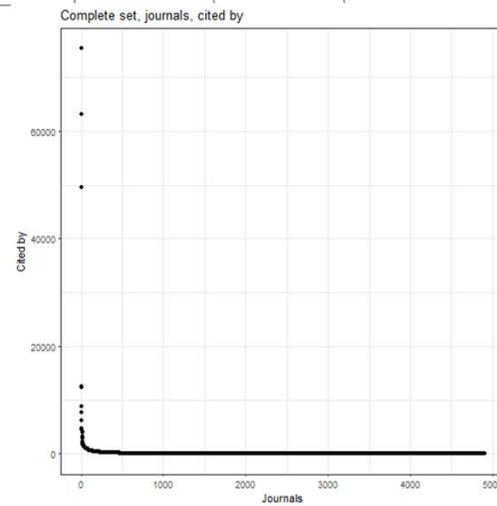
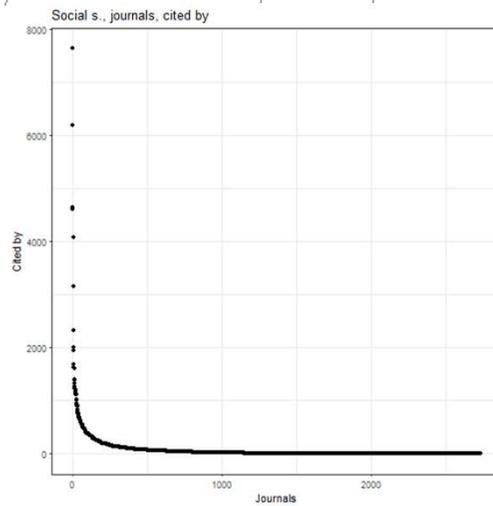
- similar distribution (Pareto like density)
(quantile-quantile graph)





Results - number of citations per journal

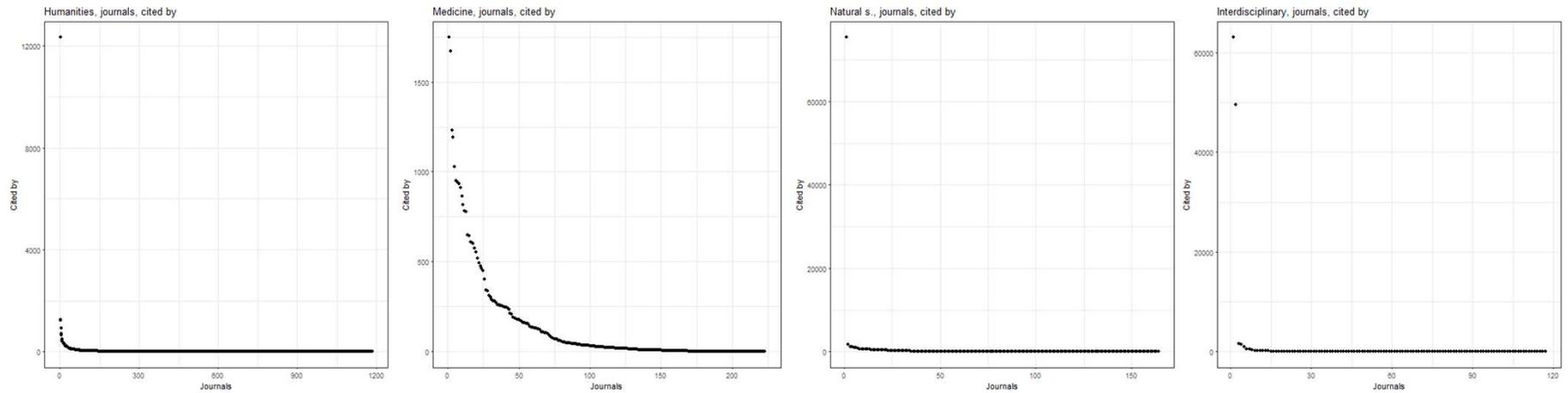
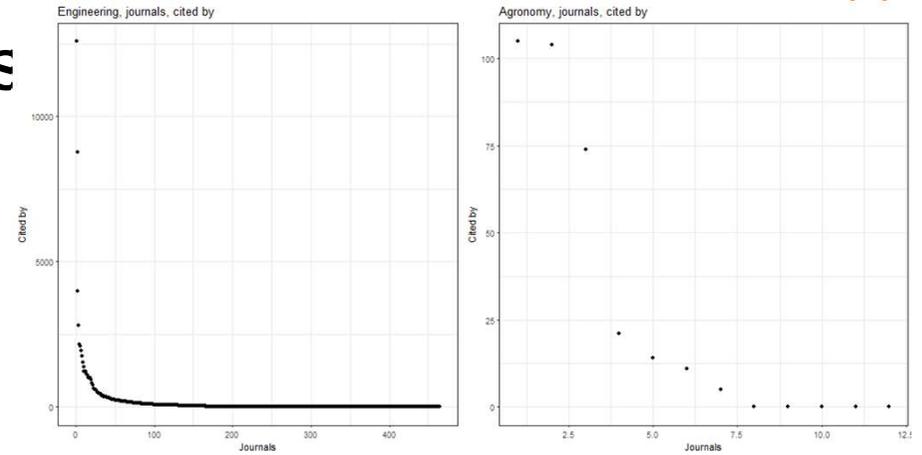
	min	1stQ	median	mean	3rdQ	max
soc. sc.	0	2	10	72.13	46	7639
complete set	0	1	7	113.3	37	75450





Results - number of citations

- number of citations per journal for non soc. sci.





Results - number of citations



- number of citations per journal in our dataset

	min	1stQ	median	mean	3rdQ	max
humanities	0	0	1	29.03	8	12343
medicine	0	4	24	144.6	154	1749
engineering	0	1	10	161.43	70.25	12595
natural sc.	0	4	19.5	581.1	97	75447
agricultural	0	0	8	27.83	34.25	105
interdisciplinary	0	1	6	1032	32	63118
social sc.	0	2	10	72.12	46	7630



Conclusion



- be careful when deriving bibliographic data from the Scopus database or any other database for bibliometric research
- most of the bibliometric researches are made for the needs of policy making, then the conclusions based on such researches should not be made without checking the database response



Future work



- visual representation of keywords space for easier identification of the field
- linking to UDC or some other kind of classification?



Thank you!

