DST-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy

Bibliometrics: Rules, applications and caveats

SU Library Week
13 May 2015

Johann Mouton (CREST and SciSTIP)
jm6@sun.ac.za

Some introductory observations (1)

- Bibliometrics is essentially a tool that can be applied for different purposes and within different contexts
- It has **descriptive** uses: to map scientific fields, to map collaborations between scientists, institutions and countries; to identify emerging fields and hot topics in science, to identify patterns and trends in the demographics of science over time, to describe patterns of publication across journals and also to identify trends in citation behaviour.

Some introductory observations (2)

- But – because of various developments (new accountability regimes, increasing internationalisation of science and continued constraints on resource allocation) – bibliometrics is increasingly used as a tool of assessment or evaluation (**evaluative** bibliometrics). Such applications are found when bibliometric indicators (citation impact scores, h-indices, journal impact factors) are used to assess research performance (from the individual to the national level).

Some introductory observations (3)

- So it is not surprising that bibliometrics is used by Deans to assess individual research performance, by Research managers to evaluate university performance and assess the strengths of institutes, centres and faculties with comparable entities elsewhere.
- Bibliometric indicators are used by national funding agencies to augment peer review, to make decisions about funding and rating.
- Bibliometric indicators are the mainstay of most university rankings (Shanghai, QS and Leiden U-multirank)

So what do library and information scientists want to use bibliometrics for?

Bibliometric indicators

RESEARCH DIMENSION	INDICATOR	LEVEL
Research output	Nr of full papers in accredited journals	University/ Campus Faculty/ College Department/ Centre
	Nr of article equivalents in accredited	
	Nr of article equivalents per most productive author	
	Nr of article equivalents per Journal	
	Nr of full papers per Journal	
	Total weighted research output (2005 – 2011)	University
	Total normed output (2005 – 201)	University
Shape of research production	Nr of article equivalents by Journal Index	University Faculty
	Nr of full papers by Journal Index	
	Nr of article equivalents by main field	University
	Nr of full papers by main field	University

Bibliometric indicators

RESEARCH DIMENSION	INDICATOR	LEVEL
Research demographics	Nr of article equivalents by gender	University Faculty
	Nr of article equivalents by race	
	Nr of article equivalents by age group	
	Nr of article equivalents by nationality	
Research productivity	Nr of papers per author	University Faculty/ Field
	Nr of article equivalents	
Research collaboration	Nr of Single Institution Papers (WoS)	University Field
	Nr of National Collaborative Papers (WoS)	
	Nr of International Collaborative Papers (WoS)	
	Research collaboration category and citation impact	

Bibliometric indicators

RESEARCH DIMENSION	INDICATOR	LEVEL
Citation Impact	MCS: Average number of citations of WoS papers	University Field
	MNJS: Average normalized citation score of the journals in which a unit has published (self-citations not included).	
	MNCS Average normalized number of citations of the publications of a unit (self-citations not included).	
	PPtop 10% Proportion publications of a unit belonging to the top 10% most frequently cited publications in their field (self-citations not included).	

Six rules of Evaluative Bibliometrics

Rule 1

Rule 1: Any measurement (and specifically when using indicators or metrics) is only as good as the underlying data source. If the data(base) or dataset that is used for a bibliometric measurement is full of errors and inaccuracies, incomplete or biased in some way, it is usually extremely difficult if not impossible to correct for such structural shortcomings. **The GIGO (garbage-in-garbage-out) rule**

Rule 2

Rule 2: Any form of scientific measurement (of which bibliometrics is an example) has to confront the problem of sampling error. Sample error is more prevalent where the sample of objects being measured (individual scientists or institutions) is very small or the objects being measured are not normally distributed. Small sample sizes and skewed distributions imply more careful analysis and interpretation of bibliometric results. This is increasingly relevant in evaluative bibliometrics where the performance of individual scientists are being assessed and the appropriateness of metrics of a single case (such as the *h*-index) are being questioned. **The rule of sampling error**

Rule 3

Rule 3: Not every aspect of science can be measured (quantified) or measured to the same degree. "Research" or "knowledge production" is a complex construct (because of disciplinary, epistemological, methodological, sociological and other differences) which makes it unlikely that this complexity can be captured in a single or few measures. The rule of partial indicators (or mono-operationism)

Rule 4

Rule 4: When confronted with the challenge of measuring the more ‘intangible” dimensions of research (such as “research quality” or “research collaboration”, there is a tendency to select metrics that are at best proxies of these properties. We then often tend to reduce the full meaning or content of that property to this one (proxy) measure. Examples: We reduce “”research collaboration” to “co-authorship of papers”. We reduce “research quality” to “citation impact”. In both these cases, the underlying problem is really that current bibliometric methods are simply inadequate (by themselves at least) to measure such properties adequately and need to be augmented by other evaluation methods. **The rule of non-reductionism**

Rule 5

Rule 5: Measurement at different levels of the research system have different implications and the results should not be confused. Metrics that are appropriate at one level (for example the Journal Impact factor at the journal level are usually inappropriate at lower levels of assessment (the journal paper level or the paper author). *The rule of the ecological fallacy*

Rule 6

Rule 6: Research practices across difference scientific disciplines vary hugely and hence – not surprisingly – have major consequences for research performance measurement. Comparative research performance measurement – at whatever level – (institutional/ field/ individual) must adhere to the rule of comparing homogeneous practices/entities. *The rule of apples and pears*

Applications

Application: University rankings

Claims

- UCT is the top SA university as it is consistently ranked highest of all SA universities in the Shanghai ranking
  - Corollary: How can other SA universities - including SU – get onto the Shanghai ranking or improve its position?
- SU is the top SA university as it is now consistently ranked highest in the (unofficial) DHET-ranking

Rank	University	Pubs (20%)	N&S (20%)	HiCites (20%)	PCP (10%)
201 - 300	UCT	35.6	10	12.2	19.1
201 – 300	WITS	31.9	16.3	7.1	18.3
401 - 500	SU	30.2	5.8	5	14.6
401 - 500	UKZN	31.0	9.0	3.6	15.2

Ranking of SA universities i.t.o. weighted per capita research output (2008 - 2013)

Rank	University	2008	2009	2010	2011	2012	2013	Average
1	Stellenbosch University	2.13	2.3	2.37	2.38	3.06	2.97	2.54
2	University of Cape Town	2.01	2.3	2.21	2.24	2.38	2.57	2.29
3	Rhodes University	1.75	1.75	1.92	2.17	2.31	2.49	2.07
4	University of the Witwatersrand	1.53	1.75	1.68	1.99	1.94	2.32	1.87
5	University of Pretoria	1.35	1.4	1.43	2.03	2.14	2.4	1.79
6	University of KwaZulu-Natal	1.2	1.4	1.48	1.49	1.78	2.08	1.57
7	University of the Western Cape	0.93	1.11	1.31	1.48	1.51	1.75	1.35
8	North West University	1.17	1.19	1.22	1.21	1.41	1.69	1.32
9	University of the Free State	1.06	1.28	1.31	1.39	1.26	1.27	1.26
10	University of Johannesburg	1.04	1.05	1.13	1.42	1.47	1.43	1.26
11	Nelson Mandela Metropolitan University	0.95	0.89	1.14	1.37	1.42	1.39	1.19
12	University of Fort Hare	0.49	0.84	1.26	1.49	1.53	1.38	1.17
13	University of South Africa	0.72	0.67	0.7	0.84	1.05	1.19	0.86
14	University of Zululand	0.53	0.69	0.69	0.64	0.67	0.53	0.63
15	Tshwane University of Technology	0.34	0.41	0.44	0.55	0.58	0.57	0.48
16	University of Venda	0.32	0.29	0.41	0.61	0.55	0.57	0.46
17	University of Limpopo	0.28	0.34	0.34	0.38	0.54	0.54	0.40
18	Cape Peninsula University of Technology	0.25	0.35	0.36	0.37	0.46	0.43	0.37
19	Central University of Technology	0.3	0.31	0.24	0.34	0.34	0.49	0.34
20	Vaal University of Technology	0.14	0.19	0.21	0.32	0.36	0.35	0.26
21	Durban University of Technology	0.12	0.18	0.2	0.3	0.21	0.4	0.24
22	Walter Sisulu University	0.04	0.05	0.09	0.11	0.13	0.12	0.09
23	Mangosutho University of Technology	0.01	0.03	0.05	0.13	0.09	0.09	0.07

Comparing universities – Some responses

- Universities – despite common missions – are still very different in many respects. This is undoubtedly true globally, but even for universities within the same science or higher education system. Unless one corrects for basic factors such as size, staff qualification, staff-student ratio's, research income, faculty structure, and so on, it remains that one is comparing apples and pears! (Rule 6 violation)

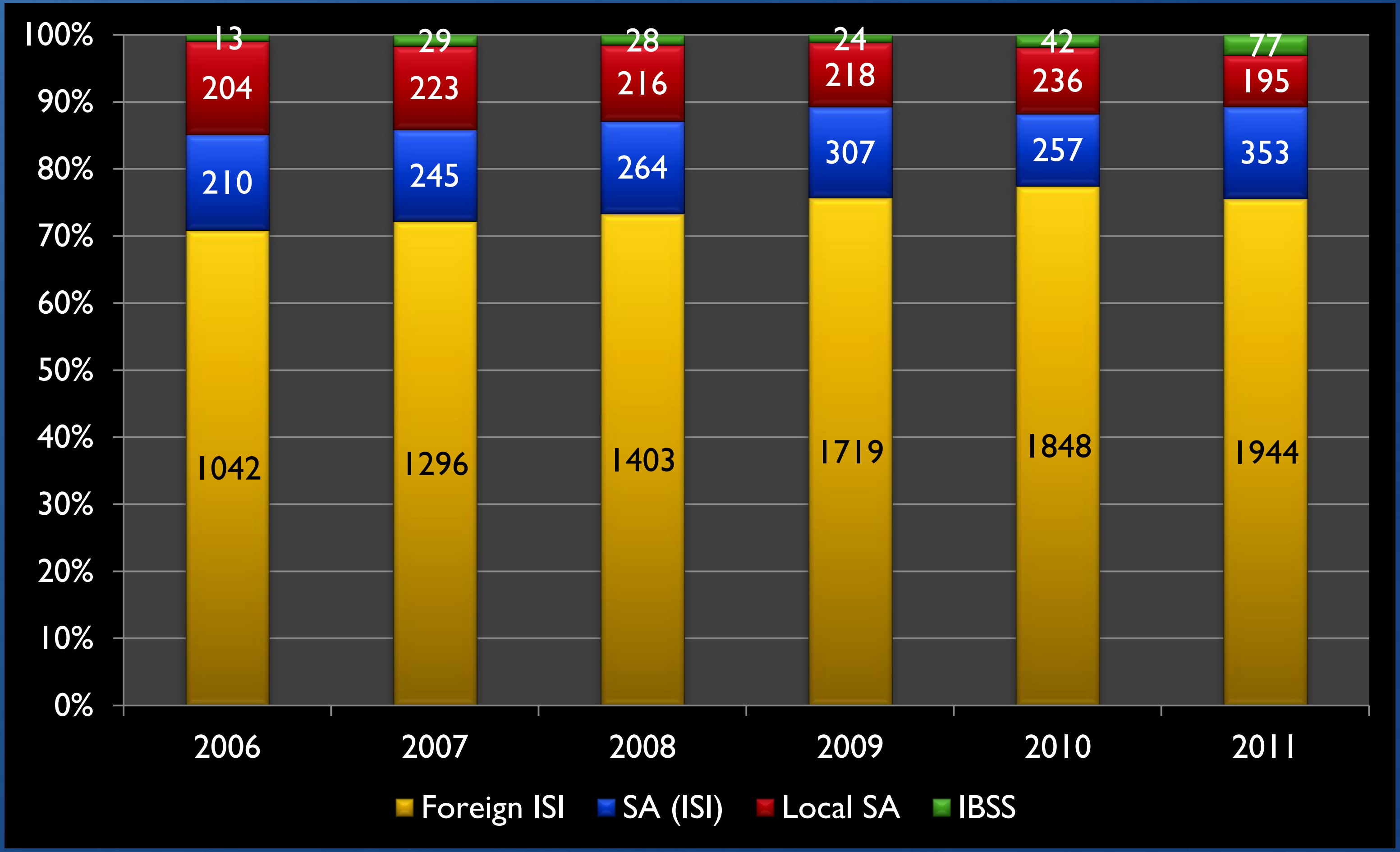
- Even within a very homogeneous set of institutions, it is unlikely that one or a few performance measures can adequately capture something as complex as research performance. (Rule 3 violation)

SciSTIP

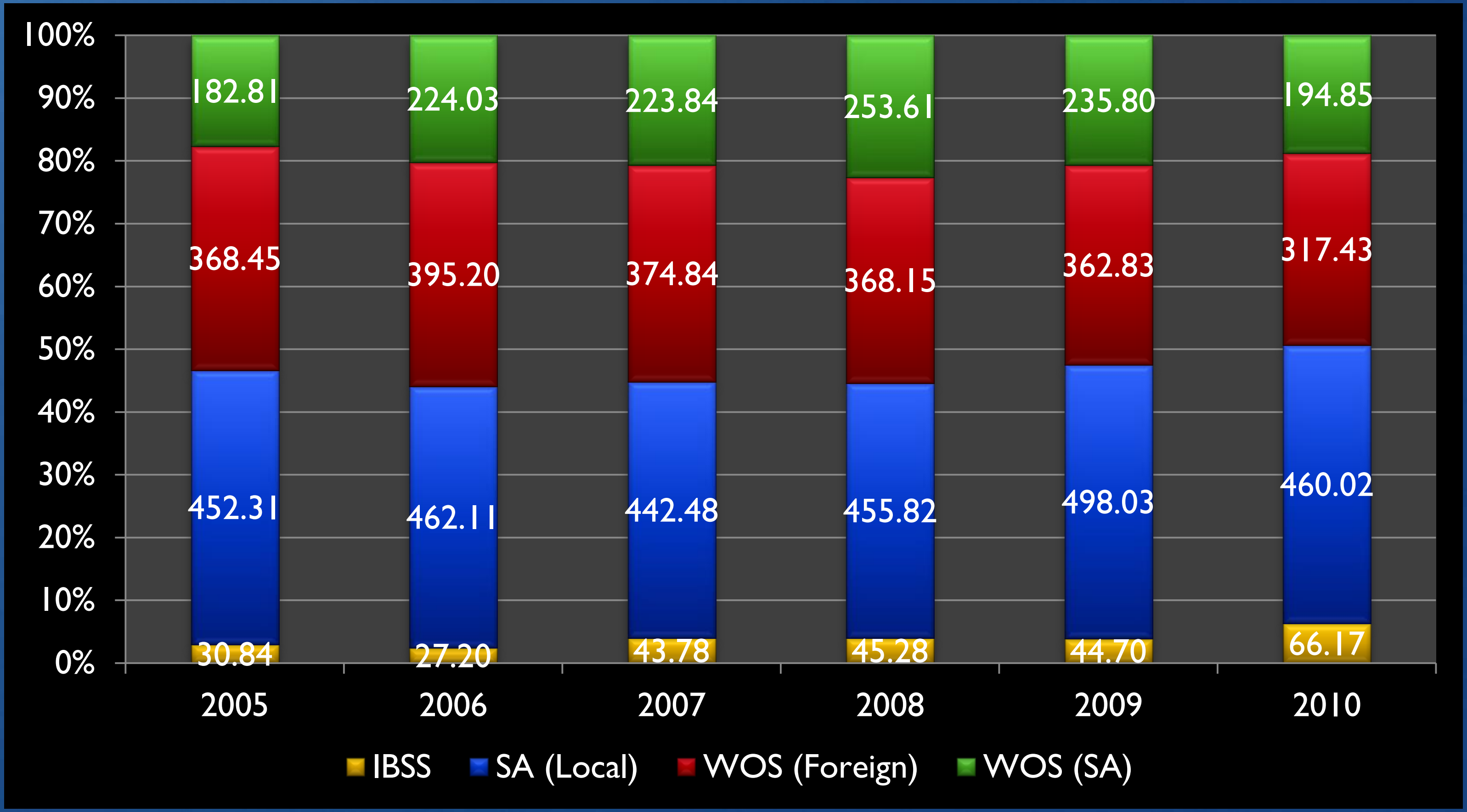
DST-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy

Understanding differences in rankings: The shape of knowledge production

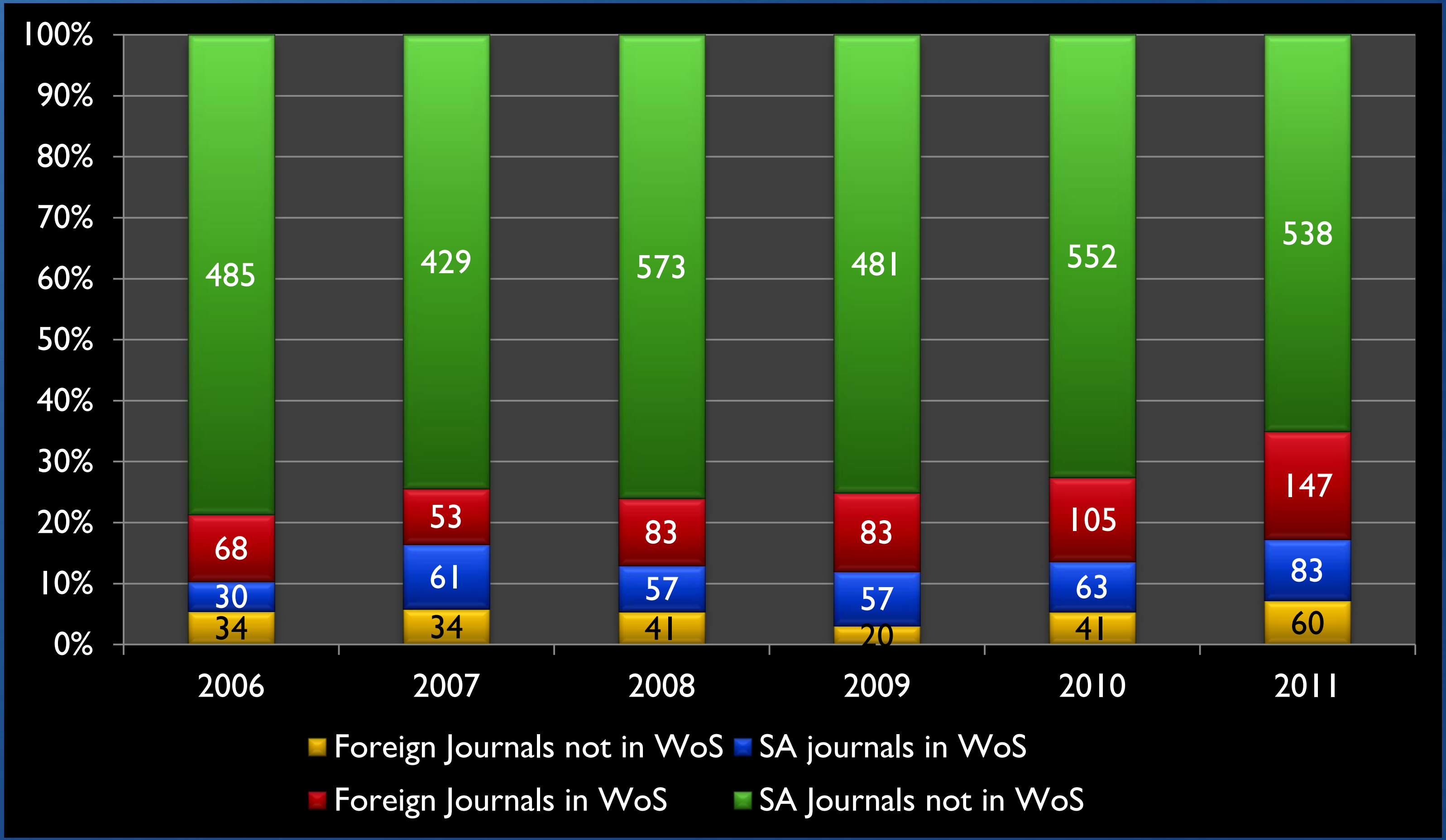
UCT Output (Full Papers) by Journal Index and year (2006 – 2011)



UP Output (Fractional Counts) by Journal Index and year (2005 – 2010)



UNISA Output (Full Papers) by Journal Index and year (2006 – 2011)



Application – Comparing scientific fields (and the JIF)

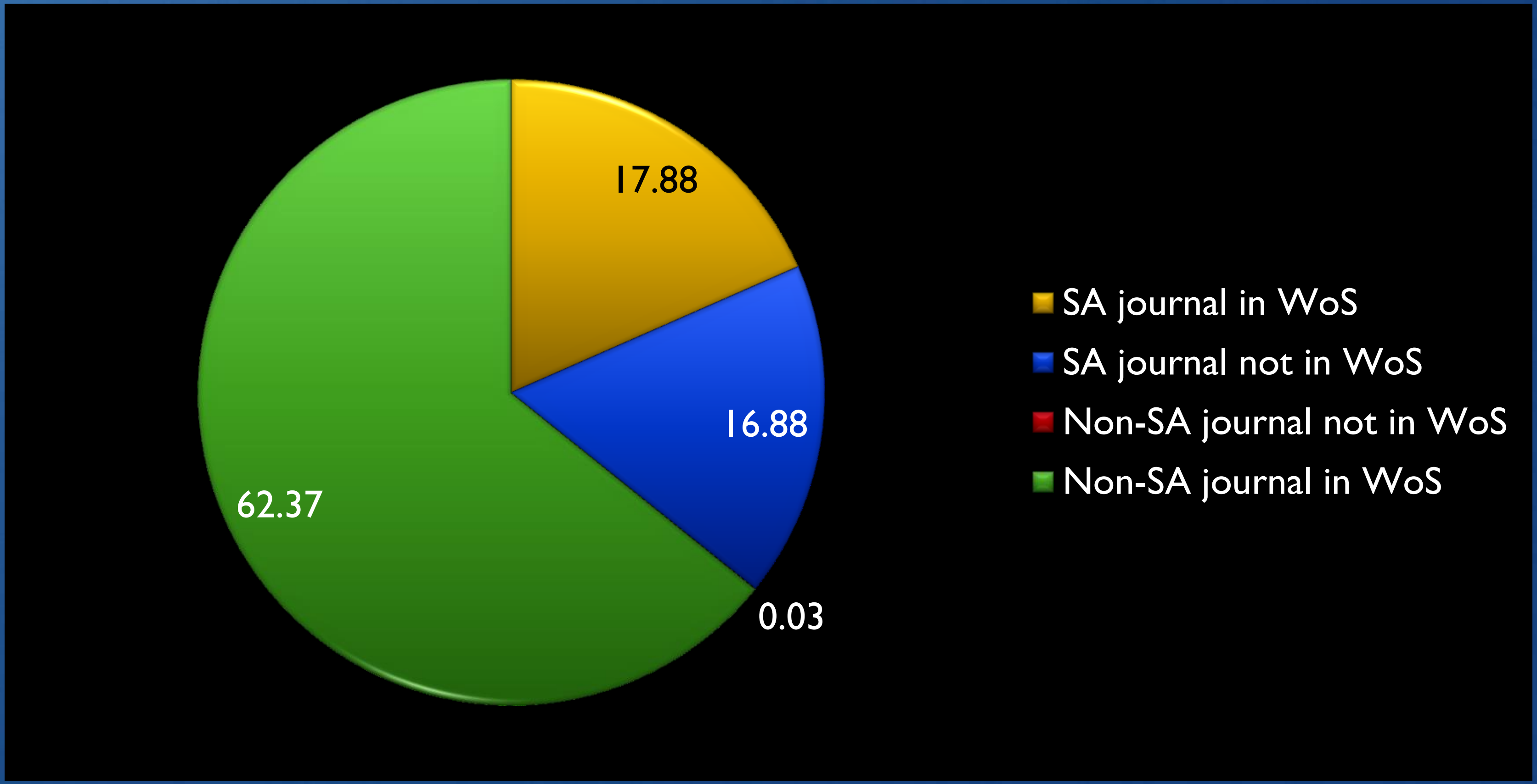
Research performance measurement across disciplines

- “The University Research Strategy requires/encourages that staff in all Faculties should aim to publish their papers in Web of Science (ISI) journals.
- The Dean of a Faculty of Arts announces higher research rewards (or incentives) if the staff in the Faculty publish in high-impact journals
- MRC Unit Director: Performance criterion = 0.5 points for every journal the unit publishes in with IF > 3 (but <5) and 1point for every article IF > 5

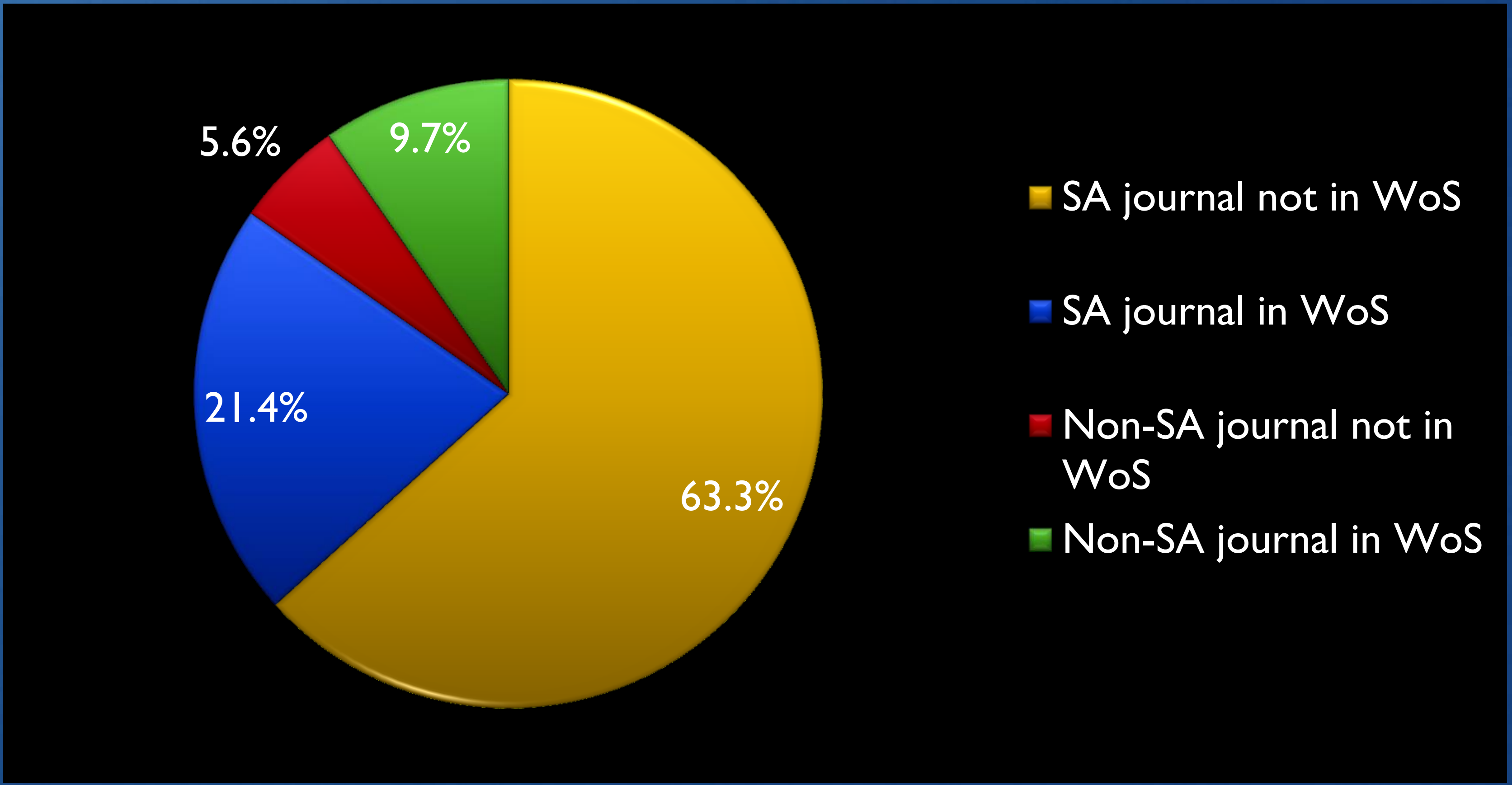
Research performance measurement across disciplines

- "The University Research Strategy requires/encourages that staff in all Faculties should aim to publish their papers in Web of Science (ISI) journals. **(Rule 1 violation)**
- The Dean of a Faculty of Arts announces higher research rewards (or incentives) if the staff in the Faculty publish in high-impact journals **(Rule 6 violation)**
- MRC Unit Director: Performance criterion = 0.5 points for every journal the unit publishes in with IF > 3 (but <5) and 1point for every article IF > 5 **(Rule 6 violation)**

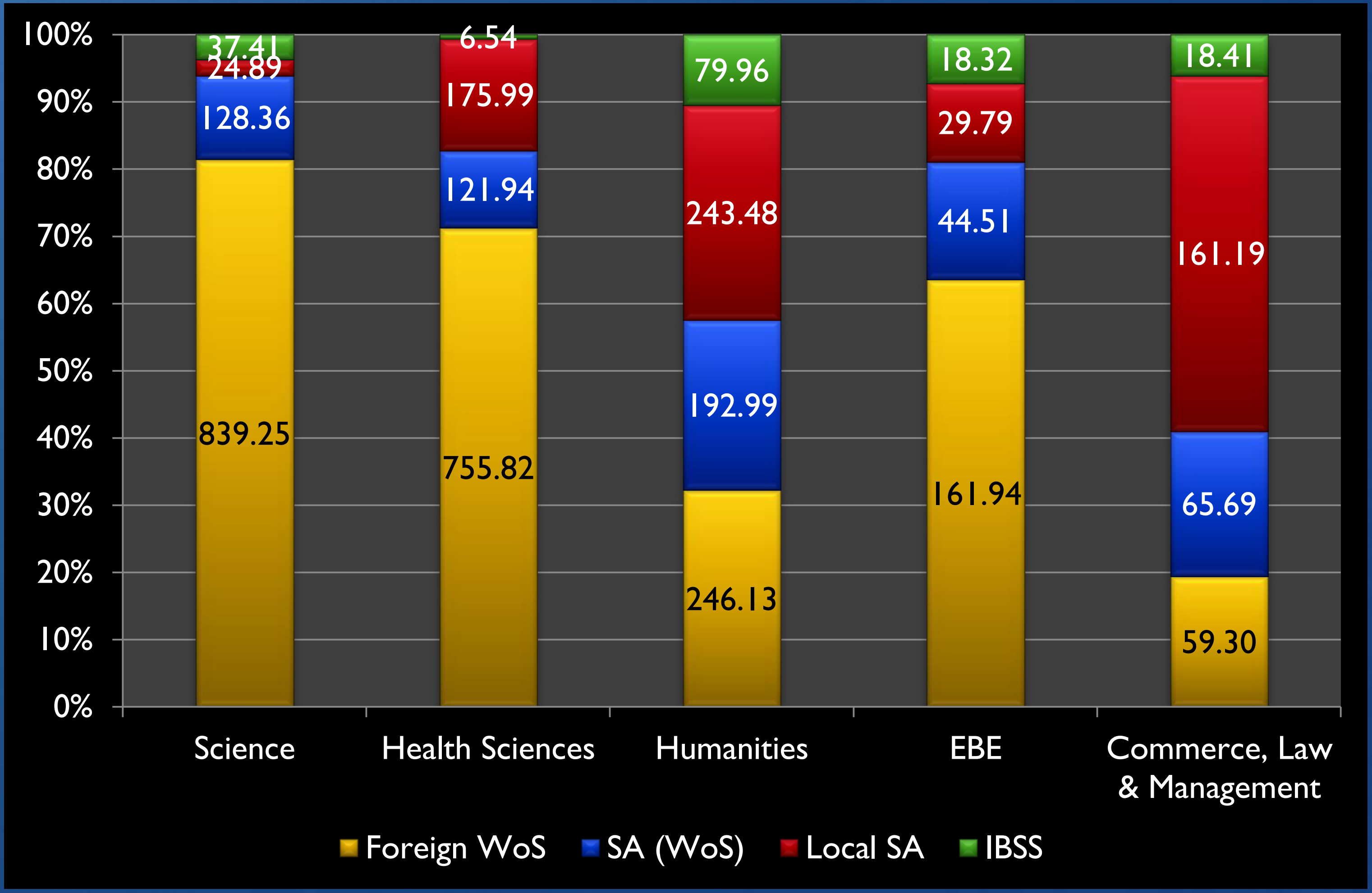
The UP "Sciences" Faculties (80%) in Web of Science: 2005 – 2010



The UP "Humanities" Faculties (31%) in Web of Science (2005 – 2010)



Comparison of article output (fractional counts) by Journal Index by Faculty (2007 -2011) WITS



Elaboration on Rule 6

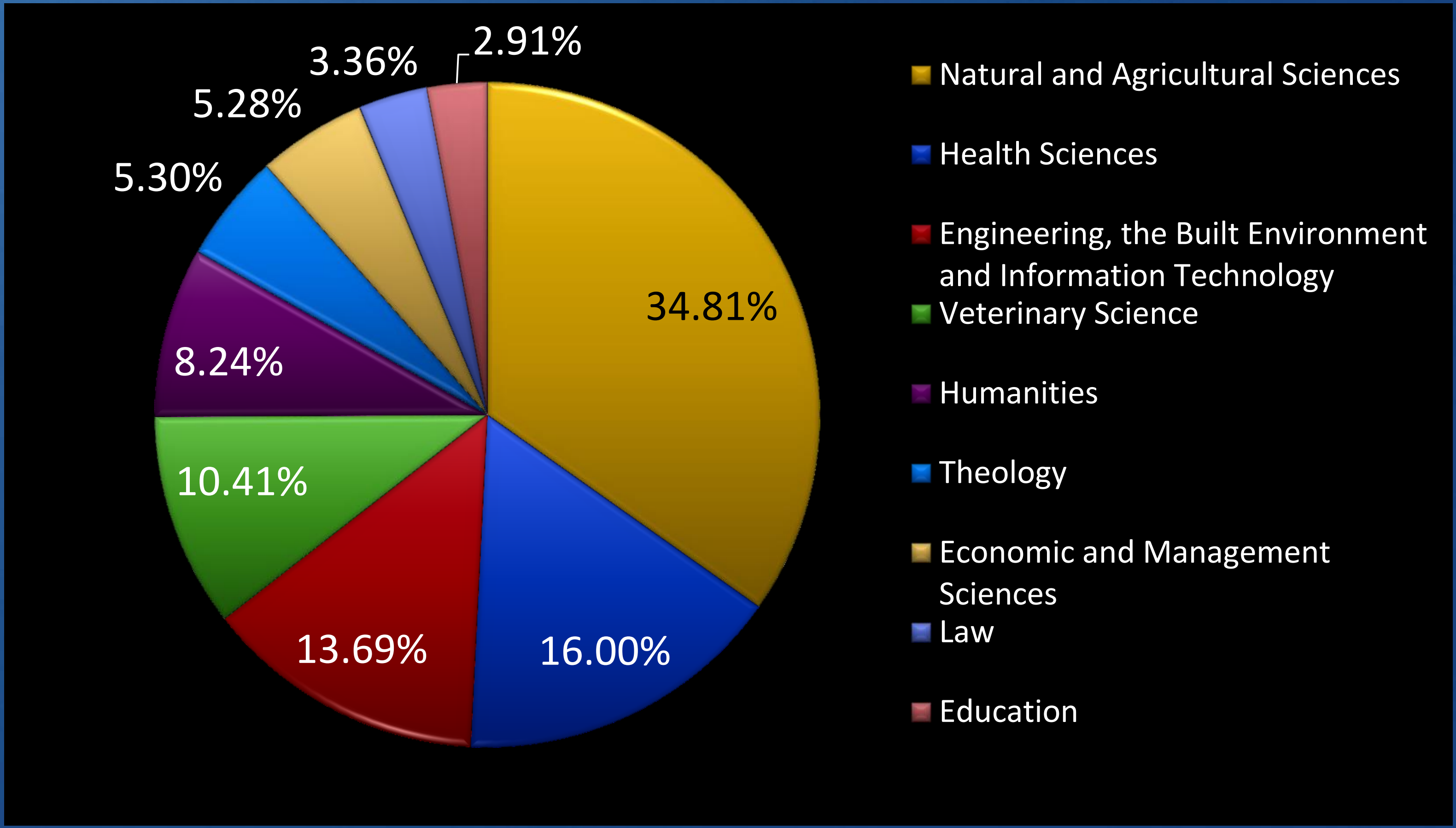
- Rule 6 requires that we recognize the very fundamental differences between scientific disciplines.
- These differences pertain both to differences in terms of publication behaviour and preferences as well as differences in citation behaviour.
- Citations and citation rates are the basis for calculating Journal Impact Factor values – This means that this discussion is especially relevant to any reference to Journal Impact Factors (IF) in research performance measurement

Application: Comparing research productivity across fields

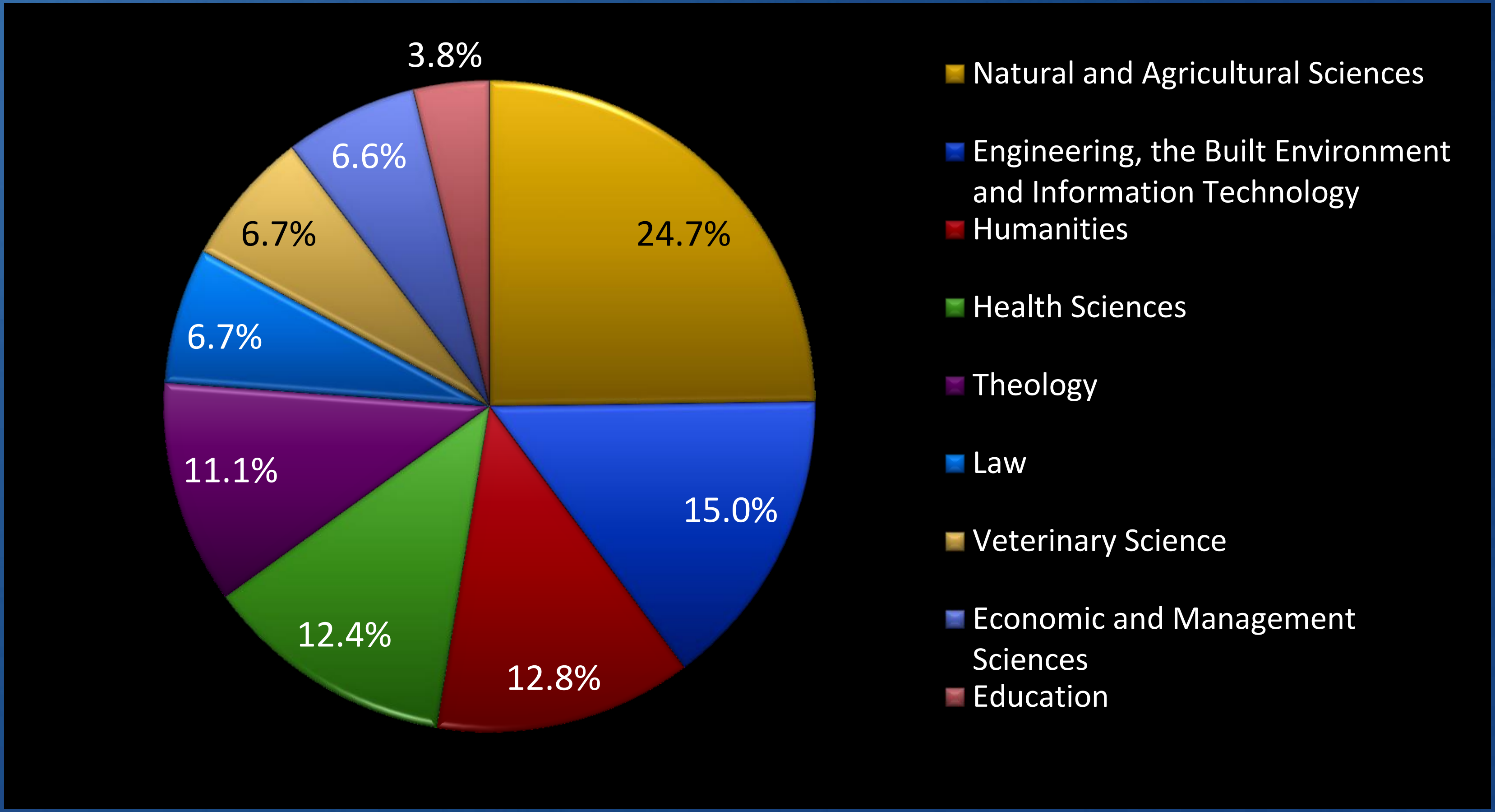
Research output measures

Fractional counts and Full Papers counts

Article (Full Paper) output by UP Faculty: 2005 - 2010



Article (Fractional count) output by UP Faculty: 2005 - 2010



UFS Research productivity by Faculty (2005 – 2011)

Faculty	Nr of authors (A)	Nr of papers (B)	Total fractional counts (C)	Paper Productivity B/A	Fractional Unit Productivity C/A
Economics/ Management	95	259	206.14	2.73	2.17
Education	51	175	127.86	3.43	2.51
Health	442	438	279.95	0.99	0.63
Humanities	267	956	788.16	3.58	2.95
Law	44	179	159.57	4.07	**3.63**
Natural/ Agriculture	636	1589	1089.89	2.50	1.71
Theology	90	339	305.96	3.77	**3.40**
TOTAL	1625	6505	2957.53	2.42	1.82

UP Research productivity by Faculty (2005 – 2010)

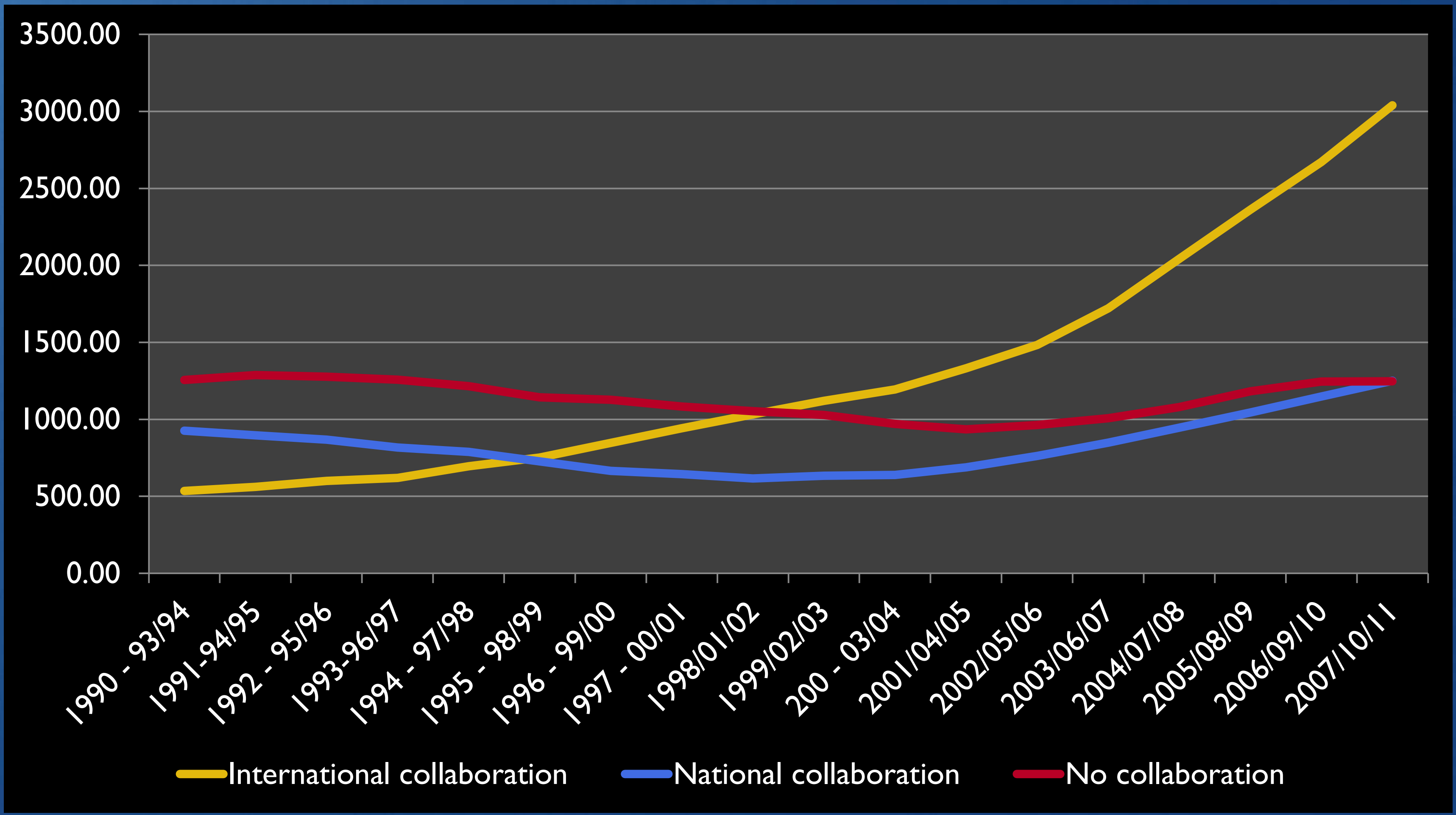
Faculty	Nr of authors (A)	Nr of papers (B)	Total fractional counts (C)	Paper Productivity B/A	Fractional Unit Productivity C/A
Veterinary	320	1656	445.94	5.18	1.39
Natural/ Agriculture	1169	5500	1626.19	4.70	1.39
Law	119	526	439.9	4.42	**3.69**
Health	774	2524	815.01	3.26	1.05
EBEIT	663	2150	982.27	3.24	1.48
Economics/ Management	273	818	426.69	2.99	1.56
Theology	301	833	724.42	2.77	**2.41**
Humanities	478	1273	828.59	2.66	1.73
Education	217	479	251.63	2.21	1.16
TOTAL	4314	15759	6540.64	3.65	1.52

Comments on measuring research productivity

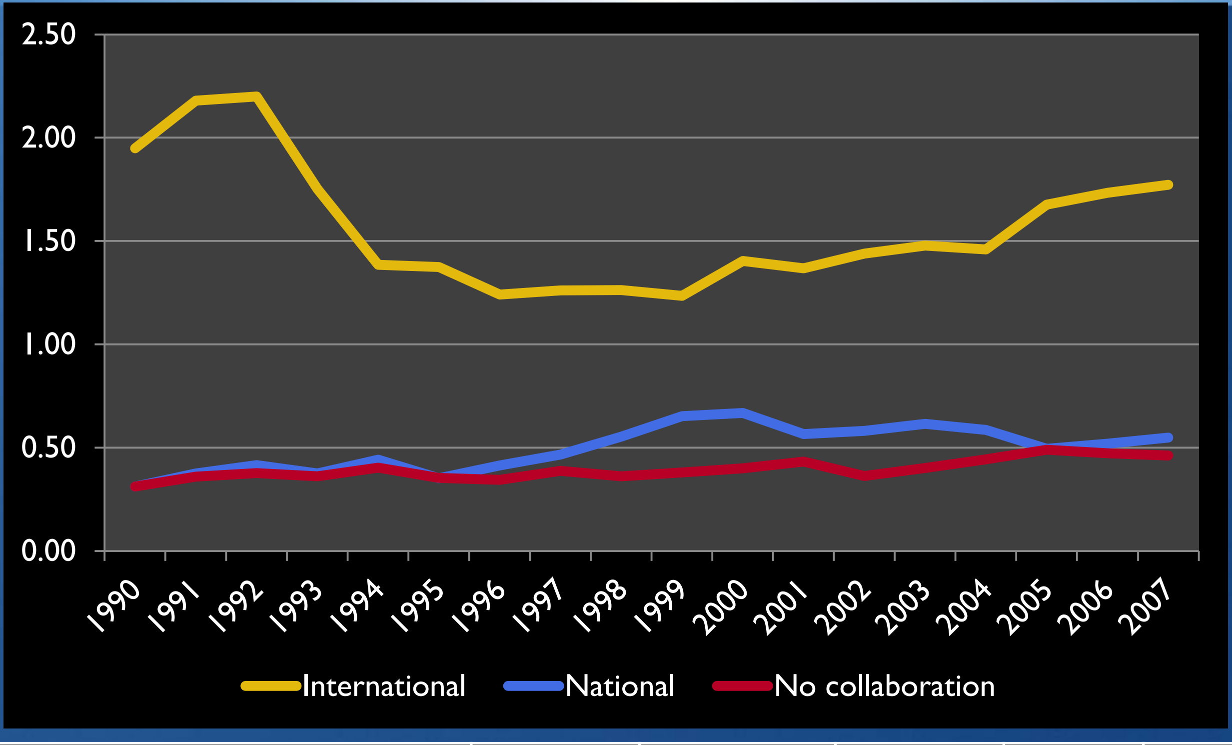
- Individual research productivity varies by scientific field. This means that any setting of performance targets by university managers must take these field differences into consideration rather than setting identical targets across the university.
- But we also need to understand what drives the differences in productivity? Besides the obvious differences that result from counting methods (full paper versus fractional counting) why, for example, do we find that academics in Law and Theology consistently register the highest productivity rates? In order to explain this properly we would have to have information on the age distributions of individual staff, student-staff ratio's as well as factor in the fact that there are greater (and arguably easier) opportunities for publishing in South Africa in these fields.

Application: Measuring collaboration and citation impact

Trends in collaboration patterns in UCT ISI-papers by 3-year rolling windows (1990 – 2011)



Collaboration type and impact (MNCS)of NWU papers in ISI-journals (1990 – 2011)



Collaboration Type	P	Mcs	MNCS	MNJS	PP top 10%	Proportion uncited	% self-citations
International collaboration	2377.00	19.35	1.60	1.22	16%	16%	24%
National collaboration	1228.75	5.70	0.57	0.77	4%	25%	23%
No collaboration	1824.25	3.92	0.43	0.66	2%	38%	26%

Some general lessons and caveats

Lessons (1)

1. Research performance management strategies are context-bound. This means that they are invariably embedded in and influenced by the policy and goal context (national or institutional) in which they operate. Policies are normative statements of what a system or institution wishes to achieve; strategies are high-level "means" or "ways" of achieving these goals. In this case: how to manage (steer/ shape/ incentivize/ reward/ monitor/ evaluate) research in order to achieve specific short- and medium-term outcomes.

2. In the SA context evaluative bibliometrics at the national level is intrinsically embedded in and influenced by the DHET Funding Framework. The recently revised framework will yet again impact on the way that we do bibliomtrics and use bibliometric indicators.

Lessons (2)

2. Research (performance) management strategies must be differentiated and customized in order to take into account:
   1. Different purposes of research assessment (and hence different expected outcomes for different target groups)
   2. Differences in research practices across different fields and scientific disciplines
   3. Differences that arise from managing research at different levels of the research system/ institution

Lessons (3)

3. As far as different purposes (and associated outcomes are concerned), research management strategies have to incorporate different strategies related to:
   1. Resource-allocation
   2. Incentivizing and rewarding performance
   3. Decision-making to inform strategic priorities
   4. Benchmarking performance

Lessons (4)

4. Research performance strategies (and their regular review) must be informed by robust and credible information systems. This implies that the system (nationally) as well as the individual institution must have appropriate information systems (databases) about its research. In addition there must be a capability at the system and institutional level to analyze and interpret both in-house information systems as well appropriate external data sources to inform intelligent decision-making.

Important caveat: There are very obvious dangers associated with using off-the-shelf bibliometric analytics such as Incites and SciValue. Although these products have their value, they are (1) expensive in the long run; (2) constrained by pre-defined queries and (3) do not take into account the peculiarities of the institutional research landscape.

Lessons (5)

5. Research management strategies must remain flexible and open to revision (as the environment changes rapidly). Such strategies must also recognize the inherent complexity of the research enterprise that it aims to manage and hence, must not fall into the traps of reductionism (the tendency to reduce complex measures to oversimplistic single measures) or absolutism (to make claims that are not based on robust and reliable data sources and seemingly not open to revision).

Thank you