Bibliometric Research, Indicators and Science Policy-Making

Professor Ben R. Martin

SPRU – Science and Technology Policy Research,
University of Sussex

(B.Martin@sussex.ac.uk)

Keynote Speech at the 13th Nordic Workshop on Bibliometrics and Research Policy, University of Tampere, Finland, 11-12 September 2008

Introduction

Focus on **interaction** between

- bibliometric research
- indicators
- science policy making

What have we learnt from last 50 years?

Look briefly at history of each

Illustrate interactions by taking you on an intellectual journey

- Been working at interface of BR, I & SP for 30 years

Identify some important observations/lessons

Brief history

Bibliometric research

~50 years old, although

- earlier progenitors (e.g. Cole, Lotka, Bradford, Zipf)
- term ‘bibliometric’ not used till 1969 (before that, ‘science of science’)

Garfield

- 1st proposed *SCI* in 1955
- SCI began in early 1960s
- Library information tool

De Solla Price

- Analysis of exponential growth of science (1956)
- *Little Science, Big Science* (1963)

Development of basic tools of bibliometric analysis

- Publications
- Citations

First efforts to use bibliometrics for evaluation & policy

- e.g. Narin & CHI

Brief history

Science indicators

- Also ~50 years old
- Freeman, OECD, NSF/NSB

Initially focussed on input indicators e.g.

- R&D funding (Frascati Manual)
- QSEs

Bibliometric indicators for evaluation/policy purposes

- only began on significant scale in 1970s
- ISI e.g. Small
- Narin & CHI
  - work for NIH & NSF (SLID in *Science Indicators*)
  - *Evaluative Bibliometrics* (1976)
- Hungarian group (Braun, Glänzel etc.)
- SPRU
- CWTS
- …

Brief history

Science policy research

Also ~50 years old

Early key contributions e.g.

- Griliches (1957) – economics of technical change
- Nelson (1959) – simple economics of basic research
- Arrow (1962) – allocation of resources for invention

Mid-60s – establishment of 1st dedicated centres

- RPI Lund (1966)
- SPRU (1966)

Creation of dedicated journals

- e.g. *Research Policy* (1971)

Now established field – hundreds of researchers + own conferences, journals, PhD progs etc.

Interactions between bibliometric research, indicators & policy-making

At first sight, simple

Bibliometric research è Indicators è Policy

But linear models always misleading/wrong!

Bibliometric research

- Carried out for variety of reasons
- Construction of indicators just one use
- Original purpose = library/information search tool
- Another = historical & sociological analysis

Interactions between bibliometric research, indicators & policy-making

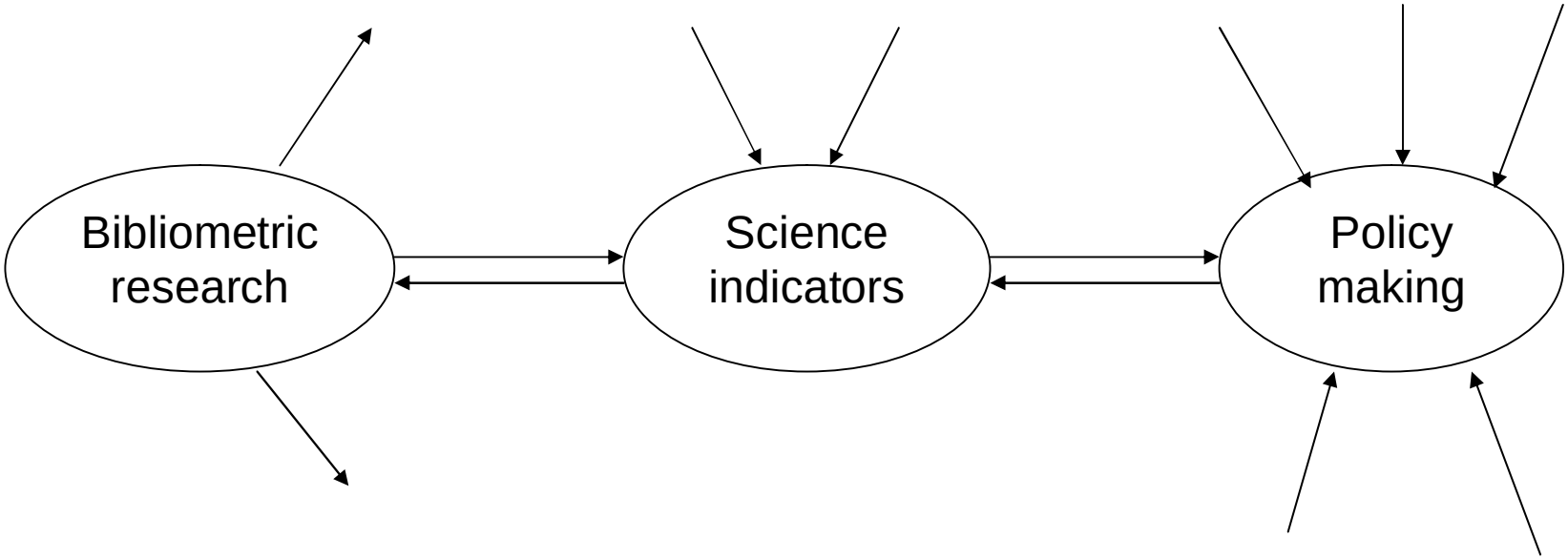
Indicators

- Variety of types e.g.
  - input VS output
  - science VS technology (e.g. patents) VS innovation (e.g. new products, sales, spin-offs)
- Bibliometric analysis just one of many sources
- Indicators used not just for policy purposes

Science policy

- Variety of inputs/influences e.g.
  - Scientific opportunities
  - Potential economic & social benefits
  - Broader government policy/ideology
  - Past performance of researchers involved
- Only for last of these are bibliometric indicators of use

So interactions complex


Bibliometric research
Science indicators
Policy making

An intellectual journey

1970-73

- What is progress?
- What is economic growth for?
- Is science being used for full benefit of humanity?
- Science and society

1975

- Manchester – LSS
- Discovered *Science Citation Index*!
- Used in sociological study to identify ‘cognitive elite’ in astronomy

Assessing basic research

1978 – SPRU

‘Big science’ project

- ~65% of SRC budget devoted to big science labs
- What benefits to UK?

Developed methodology for assessing

- scientific contributions
- technological spin-off etc.
- skills-related benefits

Assessment of scientific contributions based on

- bibliometric indicators
- extensive peer-review

Assessing basic research

Bibliometric analysis – done manually

- Collect publication lists from labs
- Write each reference on index cards
- Look up number of citations in each year in *SCI* volumes
- Analyse and calculate indicators (on a slide-rule or calculator!)

è Learnt about technical problems/limitations of SCI e.g.

- Mistakes in volume, page or date
- Variations in abbreviation of journal
- Only some journals included – biased coverage?
- Mis-spelt author (especially ‘foreign’ names), missing initials
- Variations in address etc.

Assessing basic research

Already by 1980, bibliometric indicators attracted much criticism from

- Scientists
  - Not a ‘true’ reflection of output/performance
  - Ex’s of top scientists who published little
  - Ex’s of important publications that were little cited
  - Ex’s of ‘mistaken’ publications that were heavily cited
- Sociologists
  - Publication and referencing behaviour influenced by vast range of social, institutional and other factors
  - Can’t quantify ‘quality’
  - Tribal warfare between bibliometricians and STS (e.g. at 4S confs and in *SSS*) eventually è schism

Conceptual clarification of what publications and citations do and do not measure

- e.g. citations relate to ‘impact’, not ‘quality’

Assessing basic research

SPRU study one of first

- to use bibliometric indicators for institutional evaluation and a specific policy purpose
- to stress that ‘partial indicators’ only
- to combine bibliometric indicators with extensive peer review
- to use HCPs as indicators of more radical scientific advances (cf. more ‘normal’ science)
- to demonstrate that peer review breaks down in ‘big science’ (no neutral peers)
  - Jodrell Bank VS Cambridge
  - Rutherford and Daresbury Labs
  - RGO and INT

Assessing basic research

Hostile reaction from

- scientists – threat to autonomy and monopoly of peer review
- STS – special issue of *SSS* devoted to criticisms; growing split between STS and science policy communities

Sceptical reaction from policy-makers

- "You've told us something about past performance but we're interested in the future."

CERN project

- More thorough approach to assessment of scientific performance (10,000 papers, citations for 10 years looked up manually, >1 person year)
- Factors affecting past performance
- Factors affecting future prospects of e.g. LEP, SSC
- Very controversial!

Indicators and women in science

Spin-off from big science project

Why do so few women reach top positions in science?

Data on women astronomers and male cohorts

- Women as good as men at PhD level
- But filtered out as advanced up career hierarchy
- Linked to reasons why moved job
  - Men – “promotion”, “better job”
  - Women – “partner moved”, “family commitments”

Conclusion – “Equality begins at home!”

International comparisons

Prior to 1980s, little more than anecdotal evidence

- e.g. UK scientists claiming that in decline

1980s – first systematic empirical comparisons e.g. by Braun et al.

SPRU study 'The decline of British science'

- UK's % share of publications & citations in decline
- Welcomed by scientists!
- Criticisms e.g. from Leydesdorff è more awareness of
  - technical problems (e.g. different versions of ISI database)
  - conceptual problems – all pubs VS ANR only; fixed VS varying journal set; fractional VS whole counting; % share of what?

This + another SPRU study on international comparisons of gov't funding of academic research

- è substantial increase in UK funding
- But also led to moral dilemma!

Bibliometric indicators become established

By 1990, bibliometric indicators become established for evaluation/policy purposes

- ISI, CHI – various evaluations for NIH, NSF etc.
- CWTS Leiden – showed possibilities for assessing university departments
- UK ABRC study – explored simple bibliometrics VS cocitation VS co-word analysis
- NSB *Science Indicators* – widely used
- Evaluation of research institutes

But indicators generally used in combination with peer review

UK Research Assessment Exercise

Established in 1986 by Thatcher Government

- Drastic cuts to UGC funding in early ’80s
- Demand for public accountability

Repeated in 1989 when methodology became more established

- Submitted 4 publications/researcher
- Assessed by ~50-60 discipline-based panels
- Ranked each unit/dept on 5-point scale

Stirred up storm of criticism

è Could peer review be complemented with bibliometric indicators?

SPRU study

Views of academics

- As critical of peer review as publication & citation analysis
- Best to use in combination

Purchased 10 years of UK data

- Expensive (& updates even more so)
- Huge effort to clean up and unify addresses
- Some correlation with RAE rankings but other significant influences (e.g. size of dept)

Major conceptual/methodological problem – wrong unit of analysis in RAE

- Academics say most important unit for research = (subfield-based) group, not the department
- But impossible/too expensive to carry out assessment at this level

BESST database

Previous bibliometric analysis tended to focus on ‘peaks’ (universities, large labs etc.)

BESST – aim = to produce ‘high resolution map’ of whole of British science

Identified ‘missing mass’ of British science

- ~40% of research not from universities or research institutes, but industry, hospitals and numerous small ‘players’
- Previously ignored in science policy

But very expensive – ISI

- monopoly supplier
- escalating charges
- tighter restrictions – only ‘leased’ data for specific project

The Chief Scientist gets involved!

1995 – Robert May appointed UK Chief Scientist

Invited to SPRU – showed him bibliometric data

1997 – 'The scientific wealth of nations'

- UK science strong cf. e.g. Fr, G, J
- UK science improving – effective gov't policy
- UK scientific system more efficient/competitive
- Ignored normal caveats e.g. re Eng-lang bias

2004 – David King, 'The scientific impact of nations'

Clear evidence that bibliometric indicators being used for top-level policy-making

But dilemma – concerns about possible misuse

Some observations – 1

Indicators – hidden problems and assumptions

- Technical issues/problems re how data coded e.g.
  - variations in journal abbreviations (cf. ‘General search’ VS ‘Cited reference search’)
  - books etc. not scanned (Google Scholar better?)
  - variations in addresses
  - tape-year VS publication year
  - …
- Also implicit assumptions e.g.
  - which journals scanned by ISI/Thompson (or Scopus or Google Scholar); also constant VS varying journal set
  - classification of publications (e.g. ‘note’, ‘letter’)
  - …
- Often not apparent to those doing computerised analysis
- Researchers of 30 years ago served apprenticeship in days of printed *SCI* and writing own computer programs!
- Indicators as much an ‘art’ as a ‘science’

Some observations – 2

Indicators only ‘indicate’, they do not measure

Research generates contributions to

- knowledge
- technology
- innovation
- economy
- society, health, environment etc.

Policy-makers would like info on all these

Bibliometric indicators only relate to 1st of these (and even then only provide ‘partial indicators’)

Some observations – 2

Publications counts

- Some indication of volume of output
- But not all publications are equal

Citations

- Some indication of impact on other researchers
- But not all citations are equal – variety of factors

HCPs

- Some indication of most influential scientific contributions
- But some HCPs short-lived ‘shooting stars’, then dismissed

JIF

- Some indication of most influential journals
- But only last 2 years è very ‘noisy’ data
- Also ‘game-playing’ (cf. van Raan)

Some observations – 3

‘The drunk and the street light’

“Not everything that counts can be counted, and not everything that can be counted counts.” (Einstein)

Science policy – important to measure

But just because you can measure it, doesn’t mean it’s important

There are also lots of important things that can’t be measured e.g.

- contributions to scientific knowledge
- contributions to ‘society’

Some observations – 4

Bibliometric indicators – designed for basic science

Work reasonably well for

- physical and life sciences
- ‘internationally homogeneous’ research fields

Work far less well for

- engineering/applied research
- social science (and virtually not at all for humanities)
- nationally/regionally specific research fields (e.g. law, literature etc.)

No info at all re contributions to technology, economy, society etc., even though these central to science policy decisions

Some observations – 5

Bibliometric indicators tend to work better at higher levels of aggregation e.g.

- International comparisons of national scientific efforts
- Comparisons at the field level
- Studies of institutions or departments
- Studies of outputs from major programmes etc.

Assumption that some of influence of ‘extraneous’ factors on indicators tend to cancel out (?)

Use of indicators at level of individuals requires great caution

- Sole reliance on indicators an admission of failure
- Read the publications!

Some observations – 6

Do not use indicators on their own for policy purposes

Triangulate with other sources (e.g. surveys, interviews, documentary sources etc.)

Tempting to cut corners

- Computerised bibliometric analysis now cheap and easy
- But can be misleading

Some observations – 7

Policy-makers need simple, accessible inputs

Have limited time and ‘absorptive capacity’

Some bibliometric analysis too complicated

e.g. ABRC study

- Simple bibliometrics
- Co-citation maps
- Co-word maps

Need appropriate balance between accuracy/ comprehensiveness and relevance/ comprehensibility

But need to beware misuse by policy-makers

Some observations – 8

Much bibliometric analysis has little connection with policy-making

- Not addressing specific policy issue
- Not timely
- No effort made to engage with policy-makers – may need to compromise

Growing divide between bibliometrics and science policy research communities from 1980s on

(Also split from STS community)

Why is academic research so ‘tribal’? (Becher)

Some observations – 9

Indicators relate only to past performance, but policy is concerned with the future

Some correlation between past and future performance (especially at dept/institutional level)

But bibliometric indicators only a partial indicator of future performance

If over-emphasise, è increased rigidity

- Little scope for individuals/groups to improve
- Young researchers at disadvantage

Some observations – 10

Cost of indicators must be < benefits

Implies cost of indicators for evaluation/policy purposes must be << cost of research being assessed

- SPRU – proposed 1% figure
- Reasonable target (that indicators will è policy being >1% more effective)
- NIH – also used 1% figure

But often lost sight off

e.g. UK RAE

- Total costs (including indirect) became ever larger
- Probably eventually exceeded benefits (Geuna & Martin)

Some observations – 11

If you measure a system, you change it

- Cf. Heisenberg uncertainty principle
- 'Hawthorn effect', 'Goodhart's law'

Use of bibliometric indicators è changes in behaviour e.g.

- More publications, 'salami publishing'
- More references/citations
- More 'instrumental' choice of references

Journal impact factors

- Switch in choice of journals
- More elitist hierarchy of journals? 'Winner takes all'?

Spin-offs – UK ex

RAE è more elitist structure + 'game-playing'

- Switch to REF will è new types of game-playing (worse?)

Some observations – 12

Performance indicators – 'the audit society'?

Performance indicators introduced for worthy reasons

- Greater benefits from science for society
- Opening up/democratisation of science
- Public accountability
- More effective policy and management

Applied not just to science but also education, health etc. – 'league tables' of

- schools
- hospitals
- universities

Some observations – 12

Performance indicators – 'the audit society'?

But è 'the audit society'

- Professionals try to maximise score on performance indicator
- May not improve actual performance (may even worsen)
- Tends to encourage risk-averse behaviour

Latter a major threat to science

- Need some bold/risky research where most fail but a few succeed è major advances
- Being squeezed out as indicators become more prominent?

Concluding questions

Following the introduction of bibliometric analysis and the development of indicators,

- Is science policy now more effective?
- Do we have evidence that evidence-based policy is better?
- Are the benefits from using indicators greater than the costs and unintended adverse consequences?
- Is science now better?
- Is the world a better place?