technopolis |group|

Evaluating research-performing people and organisations

Bibliometrics in context

Erik Arnold
ESSS, Berlin
10 September 2012

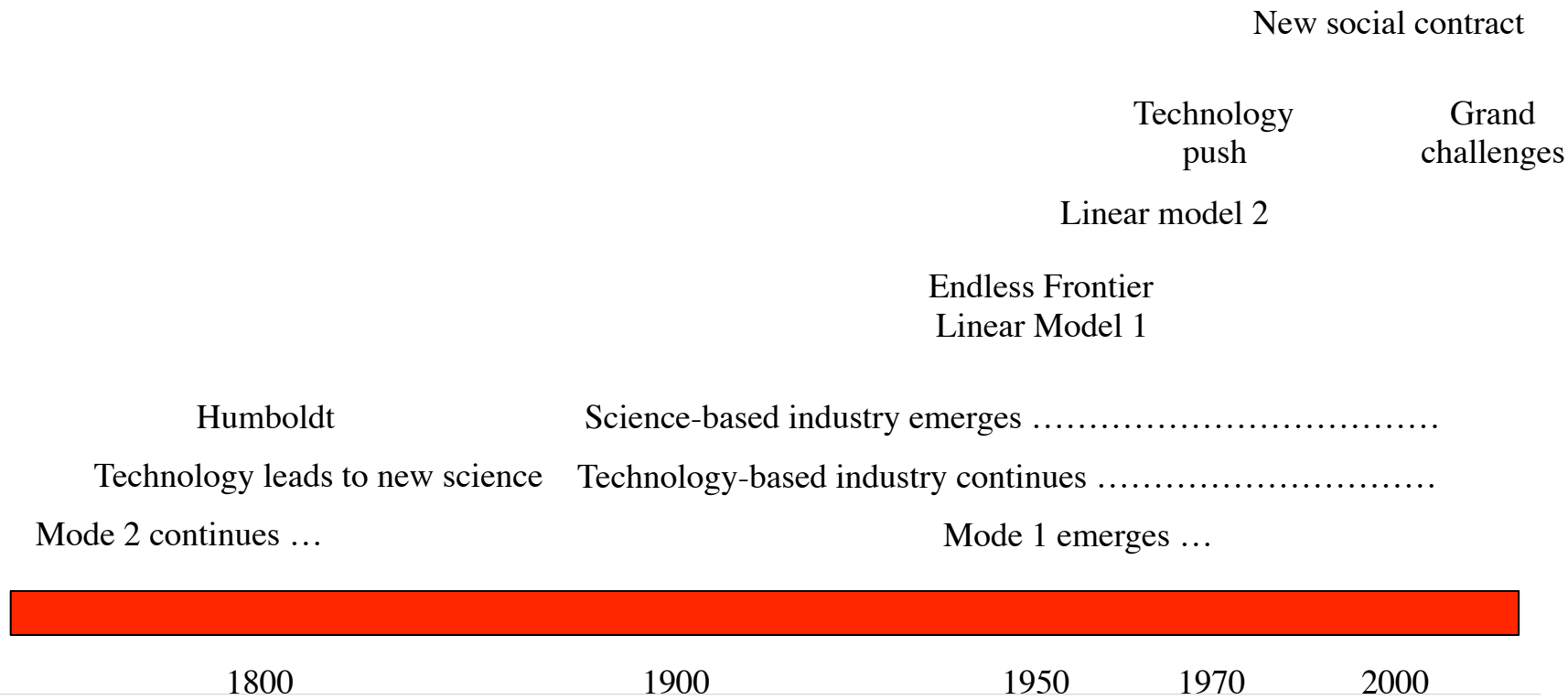
technopolis |group|

Road map

- Some historical background
- Evaluating RCN
- Performance-based research funding
- R&D evaluation challenges

technopolis |group|

The long-term research landscape



technopolis |group|

Road map

- Some historical background
- Evaluating RCN
- Performance-based research funding
- R&D evaluation challenges

technopolis |group|

Policy frameworks

- First generation
  - *Cultural lags*
  - *Linear model of innovation*
- Second generation
  - *Accounting*
  - *Economic growth*
  - *Industrial competitiveness*
- Third generation
  - *National Innovation System*
  - *Knowledge-Based Economy*
  - *Information Economy (or Society)*

Source: Benoît Godin, Making Science, Technology and Innovation Policy: Conceptual Frameworks as Narratives, Innovation/Innovacíon/ Innovaçao-RICEC, 1(1), 2009,

technopolis |group|

OECD is instrumental in the death of the linear model in policy

- Frascati Manual, 1963 – implicit systems perspective
- 1963 First OECD Ministerial Meeting on Science
- 1964 Science and Technology Reviews start, eg Sweden
- 1965 Committee for Science Policy and the DSTI
  - *Jean-Jacques Salomon*
  - *Introducing the idea of planning and science policy*
  - *The idea of a 'technology gap' with the USA becomes a dominant justification for S&T policy*
- 1966 – OECD training in science policy for people from member states

technopolis |group|

The 'OECD line' set by DSTI in 1966 underpins the change in the social contract

1. Research should help reach national, politically-determined goals
2. Research should be planned and organised to that end
3. Research should be more interdisciplinary, in order to solve real-world problems
4. The universities were rigid, organised by discipline and unable to change themselves. They should be 'reorganised' in order to contribute more to the solution of societal problems and to reach national goals

- Source: Committee of Science Policy (CSP), Draft Report on the Promotion and Organisation of Fundamental Research. With Particular Reference to the Situation in Western Europe, 20 December 1967. The Promotion and Organisation of Fundamental Research. Issues and Recommendations, 21 February 1968.

technopolis |group|

R&D evaluation really starts in the 1990s

- First distinctive evaluation activities clearly visible ca. 1980
  - *Key UK effort was to evaluate the Alvey Programme, 1983-7*
- Slow growth of programme evaluation through the 1980s
  - *eg of ESPRIT and later other parts of the Framework Programme*
  - *Autodidact R&D evaluation community developed, with poor separation from those being evaluated and little distinction between evaluation and planning*
- European Commission's SPEAR/Monitor network generated a community of European R&D evaluators and evaluation customers
- Growth of New Public Management in the 1990s
  - *Promoted programme evaluation*
  - *Enabled R&D evaluation specialists to enter*
  - *Led R&D evaluators to connect with the wider evaluation community*

technopolis |group|

R&D evaluation since 2000

- ASTPP: Evaluation and imperialism *in lingua latina*?
- Growing interest in organisations and portfolios
  - *TFK, KISR, RCN, FFF, FWF, Marsden Fund, NWO, KFAS, FNR, RCN again, Tekes, Academy of Finland, BOKU, Saltzburg University of Applied Science, Danube-University Krems*
  - *Finnish energy programmes, Swedish long-run energy programme, Ludwig Boltzmann Society, Czech Research Audit*
  - *Competence centres: Sweden, Austria, Netherlands, Estonia*
- Innovation system evaluations – key issue is governance
  - *OECD: Luxembourg, S Africa, Norway, Croatia*
  - *EC: France, Netherlands, Latvia, Slovenia, Belgium*
- Re-emergence of long-term impact studies
  - *Driven by VINNOVA*
  - *Now taken up at EU level and in the USA*

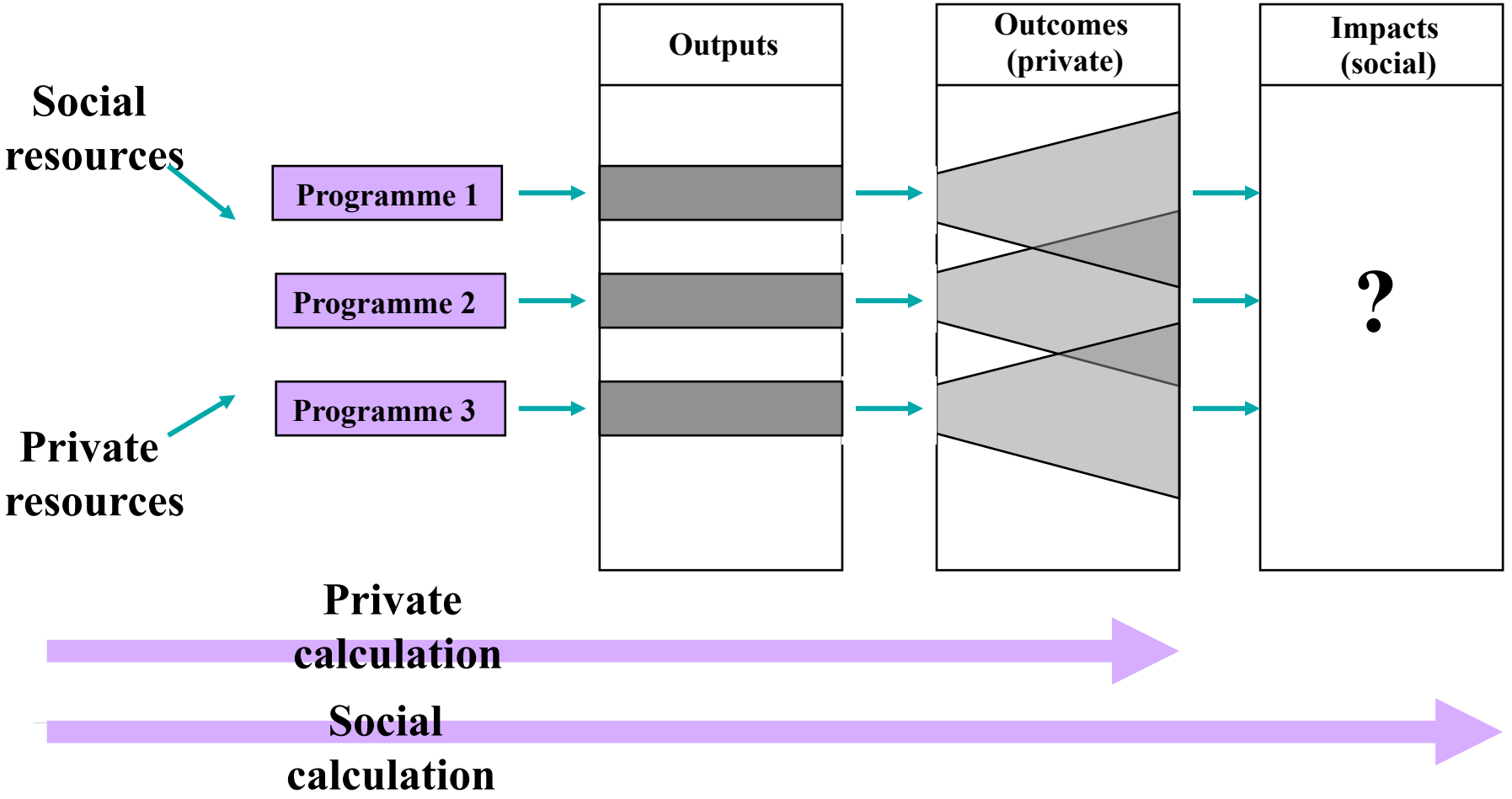
technopolis|group|

How methods have changed

- Peer review focused on assessment more than evaluation, but use of expert review at aggregate levels (eg Framework Programme)
- Growing dominance of social scientific methods and skills
- More explicit analysis and charting of intervention logics
- Changing survey economics: growth then survey overload
- Cost disadvantage reduces role of qualitative interviews, with growth of case study as a substitute
- Bibliometrics: declining cost and entry barriers; greater granularity; simplistic use in assessment
- SNA, co-word and content analysis: getting more interesting but largely not robust; macro-micro link generally weak
- Cost/benefit, economic impact analysis gaining ground, despite the uncertain status of the numbers

technopolis |group|

Effectiveness - the measurement problem



technopolis|group|

Effectiveness - practical issues

- **Data quality and collection**
- **Understanding and describing skew in innovation processes**
- **Attribution, multiple causality**
- **Additionality, dead weight, free riders**
- **Identification of complete set of effects**
- **Monetarisation of non-financial costs and benefits**
- **Understanding interplay between effectiveness and process**
- **Tuning methodology to policy objectives (eg technology diffusion vs technology creation)**

technopolis|group|

Relative effectiveness measurement (1)

- **Methodological obstacles ….**

- **The variation in the 'time gap' between interventions and effects**
- **The attribution of effects to the programme; multiple causality**
- **What works well at one time may not work in future due to change in context, eg learning, development**
- **If individual measurements are incomplete or use different ways to describe effects, they are not comparable**
- **Monetarisation of non-financial effects is value-laden and therefore political**

technopolis|group|

Relative effectiveness measurement (2)

- **Policy objections …**

- **If NIS system performance drives welfare, different interventions are not alternatives**
- **Instruments address different policy objectives so they are not easily interchangeable**
- **Not all desired policy effects can be quantified**
- **Results take too long for flexible policy decisions**
- **Different instruments address different target groups**

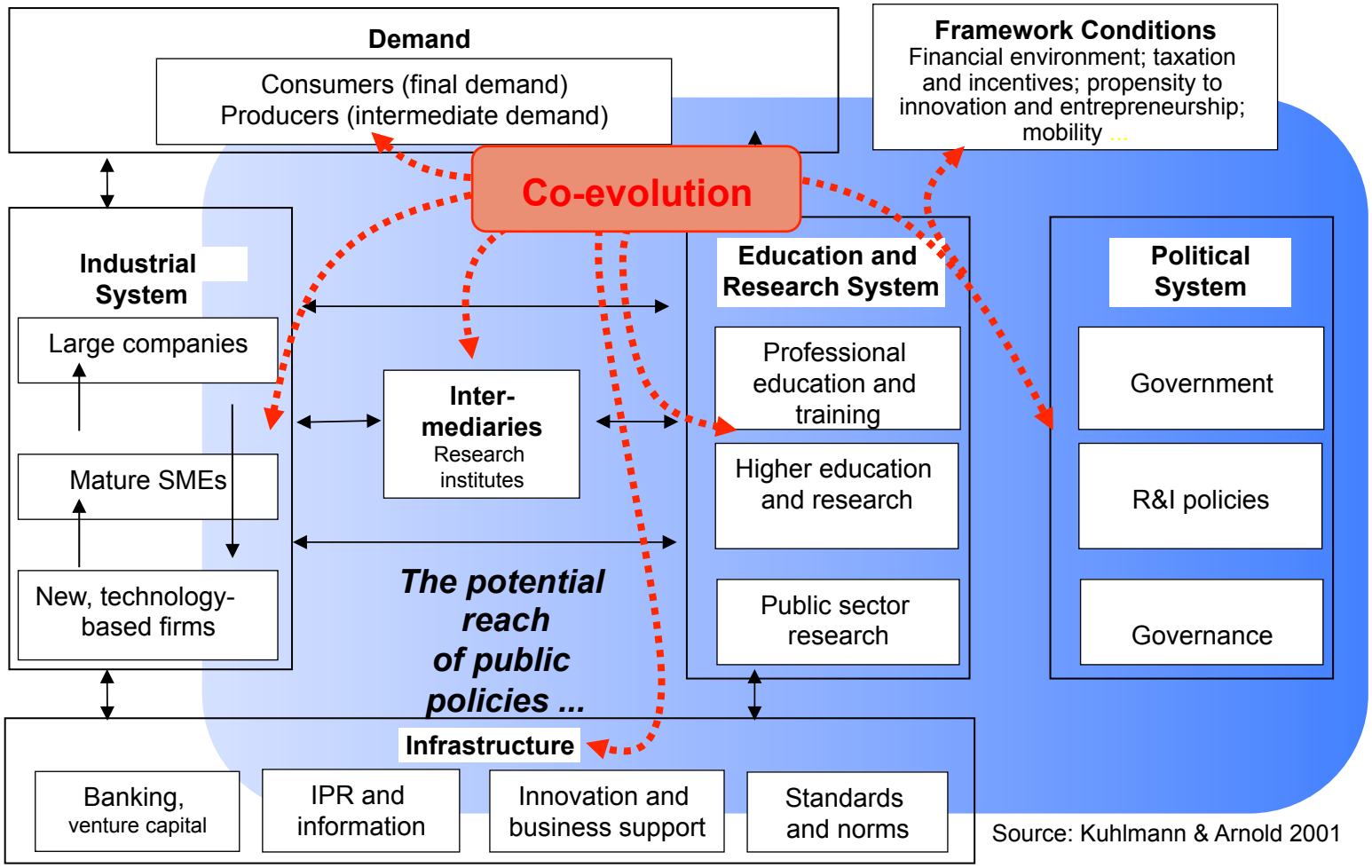
technopolis |group|

Road map

- Some historical background
- Evaluating RCN
- Performance-based research funding
- R&D evaluation challenges

technopolis |group|

Major organisational evaluations are systems evaluations



Source: Kuhlmann & Arnold 2001

technopolis|group|

The NIS perspective has important implications for how we understand performance

- The bounded rationality of actors has important consequences
  - *Knowledge, learning and institutions are key*
  - *Path dependence*
- Institutions and their environments are inter-dependent – they co-evolve, so institutions are always context dependent
- In many cases, the relevant unit of analysis is not the individual but networks, clusters and institutions
- Governance matters
- We cannot deal in static optima – we have to understand how to deal with system dynamics
- Demand, not just supply, drives innovation systems
- National systems are internationally open
- A NIS is time- and context-specific. Forget 'physics envy'

technopolis |group|

Some challenges for the evaluation

- Understanding the specificities of the Norwegian research and Innovation System and RCN's systemic impacts
  - *Characteristics*
  - *Performance*
  - *History and culture*
- Tackling governance and principal-agent relations – especially where the principal is the evaluation customer
- Assessing change agency within a unique NRIS and history
  - *No control*
  - *No universally applicable or general 'theory'*
  - *No way statistically to specify expected performance*

Formative questions. How well does RCN .....

- Create and provide strategic intelligence on research and innovation to stakeholders in the National Research and Innovation System (NRIS), including itself?
- Operate effective organisation and governance structures and processes within a national division of labour among government authorities?
- Implement and add value to national research and innovation priorities and policies?
- Play a developmental role in the NRIS, supporting the needs of the various component communities and institutions?
- Embed Norway in the changing international Research and Innovation System?

technopolis |group|

Summative questions ….. (MRS goals)

- Goal 1: Increase quality, capacity and relevance in Norwegian research nationwide
  - *Strengthened research in nationally prioritised areas*
  - *Strengthened breadth of long term, fundamental research and concentration of resources on the best research*
  - *Strengthened breadth of research-based innovation and concentration of resources on the best research and innovation milieux*
  - *More research in business, both in terms of breadth and excellence*
  - *Strengthened research to serve the knowledge needs of the sectors and the administration*
  - *Increased returns from international research cooperation*

technopolis |group|

Summative questions (continued) .....

- Goal 2: Good use of resources and appropriate division of labour, cooperation and structure in the research system
  - *A good relationship between the national research effort and international cooperation*
  - *Value added through efficient use of research funds*
  - *Dynamic, efficient and effective cooperation and division of labour in the research system*
  - *Assuming strategic responsibility for the research institutes*
- Goal 3: Research results are used by business, society and administration throughout the country
  - *Good and appropriate communication of research to society and good learning arenas among companies, institutes and the higher education sector, the health sector and the administration*
  - *Increased commercialisation of research results*
  - *A strengthened knowledge base in RCN's research policy and advisory work*

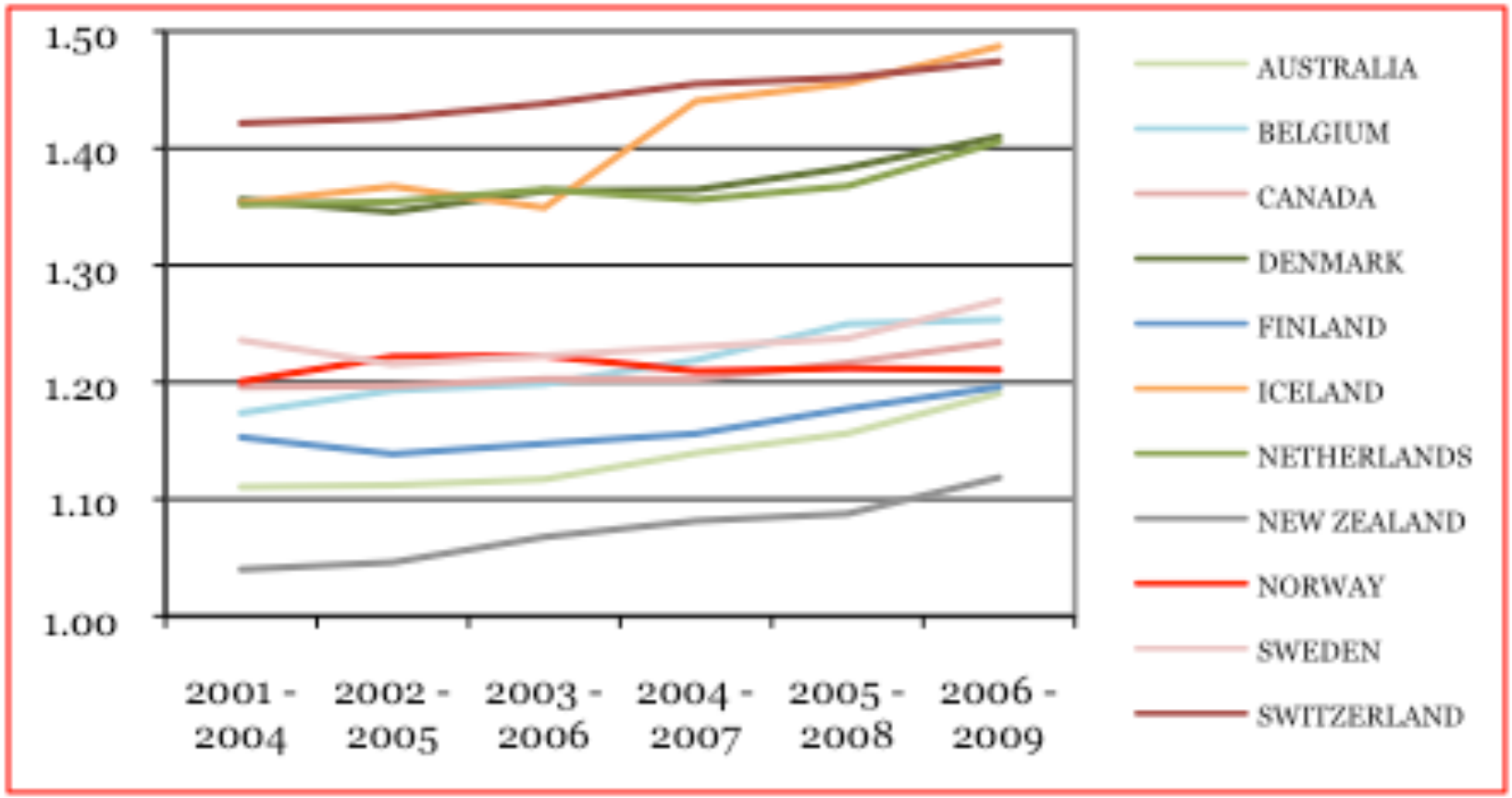
technopolis |group|

How methods map to evaluation questions

Evaluation Questions	Inter-views	Surveys	Comp-osition	Meta-evaluat-ion	Docu-ments	Biblio-metrics	Econo-metrics	Inter-national Comp-arisons
Strategic intelligence	X	X	X	X	X			X
Organisation, governance	X	X			X			X
Adding policy value	X		X	X	X			X
Developing the NRIS	X	X	X		X	X	X	X
Internation-alisation	X				X	X		X
Quality, capacity, relevance	X	X	X	X	X	X		X
Good use of resources	X	X		X	X			X
Use of results	X	X		X	X			X

technopolis |group|

Research quality: not at all bad, just not good enough



Ratio of field-normalised impact factors to world average

technopolis |group|

Low degree of internationalisation in RCN-funded papers

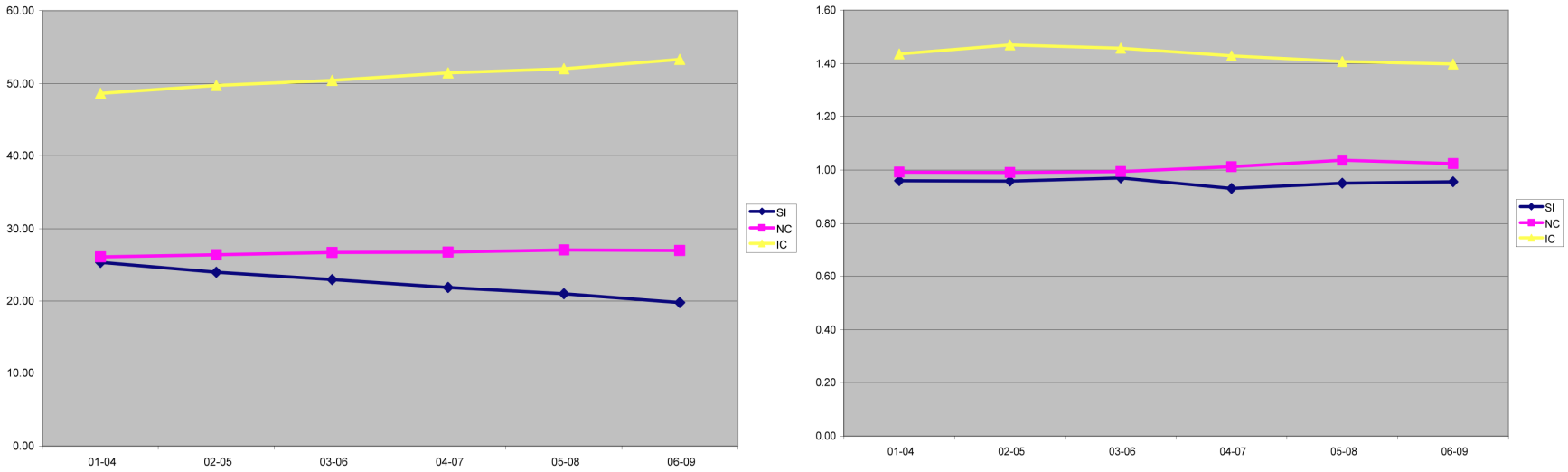
- Proportions of internationally co-authored papers, Norway vs selected other countries

	NO	BE	CH	DK	NL	SE
RCN Funded	40%					
RCN Funded Soc+A&H	43%					
National Public Bodies Funded	45%	58%	51%	48%	51%	51%
Other Nations' Public Bodies Funded	92%	94%	90%	91%	89%	89%
EU Funded	86%	80%	56%	78%	58%	58%
Privately Funded	64%	77%	79%	68%	75%	75%

technopolis |group|

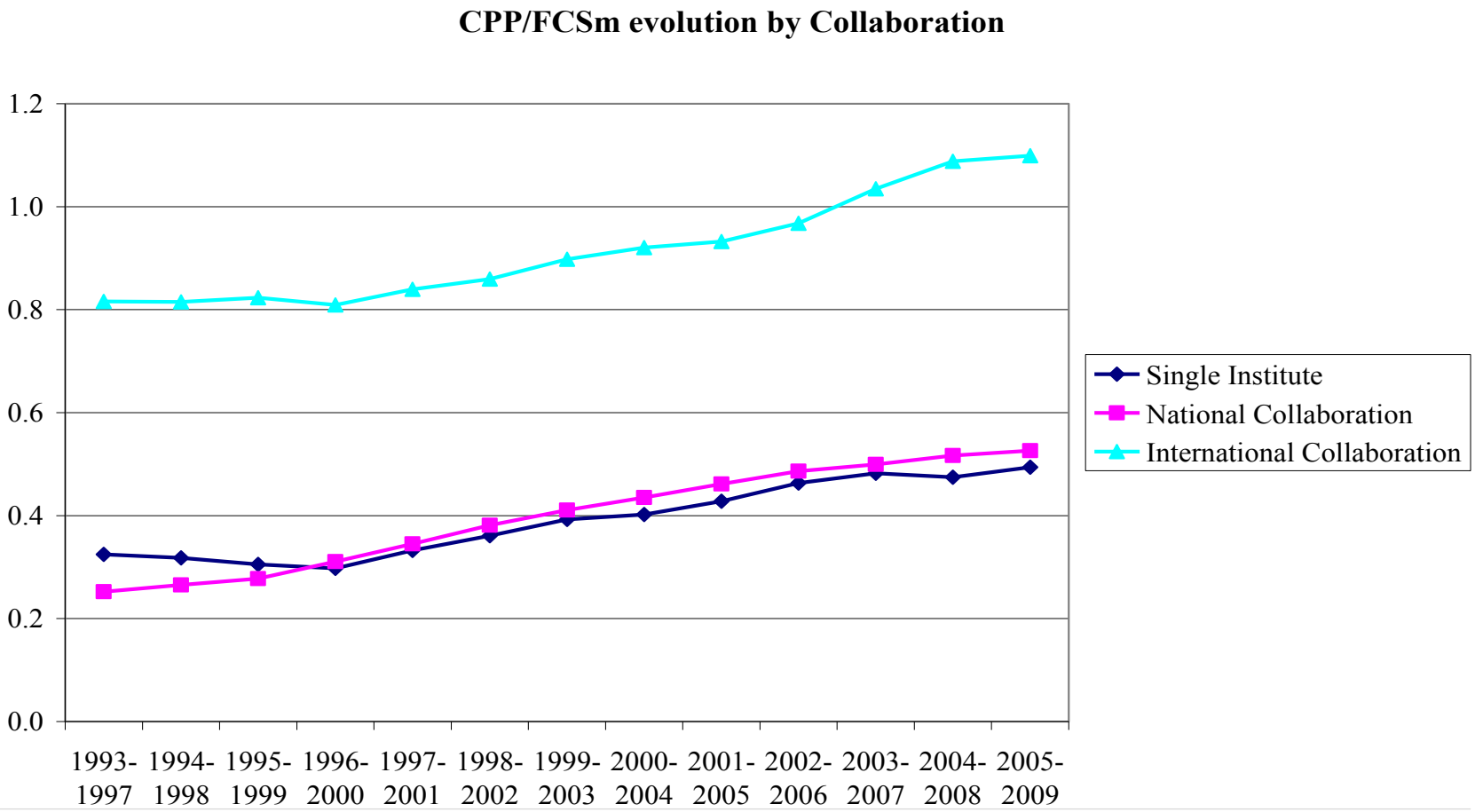
Reaching the limit for benefits of international collaboration?

- Frequency vs impact of international collaboration, national collaboration and single institute papers



technopolis |group|

Compare the Czech case (more typical of countries in development or transition)



technopolis |group|

Finland – an expensive but crucial result

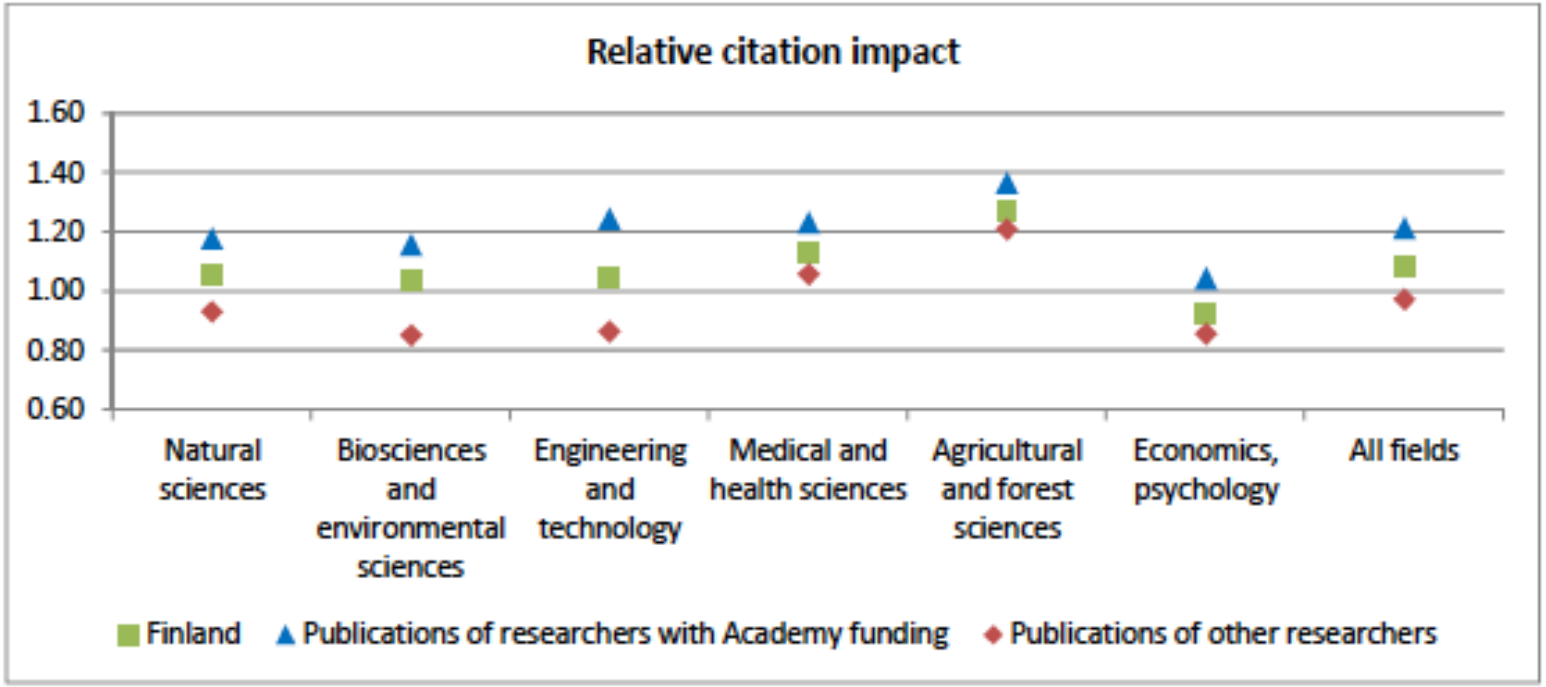


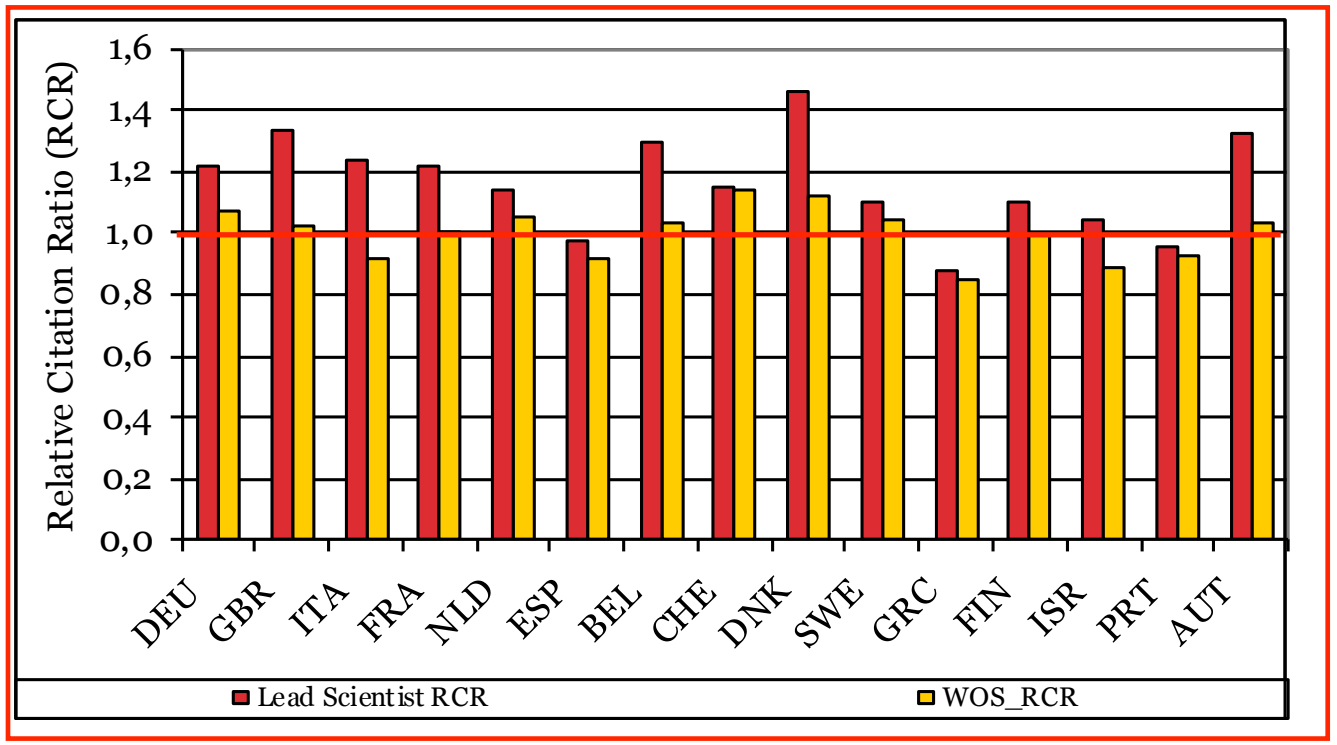
Figure 4. Relative citation impact by researcher's funding background and main field of science, based on citations accumulated during 2008–2011 to publications in 2008–2009. The world average level is one.

Source: Anu Nuutunen, Yrjö Leino and Hanna-Mari Puuska, 2013

technopolis |group|

Most of FP6's Lead scientists have a higher visibility than the average in their country

Question : Are the most highly-cited FP participants also the most highly cited authors in their country/field?



Lead scientists RCR

versus

*Entire scientific community RCR**

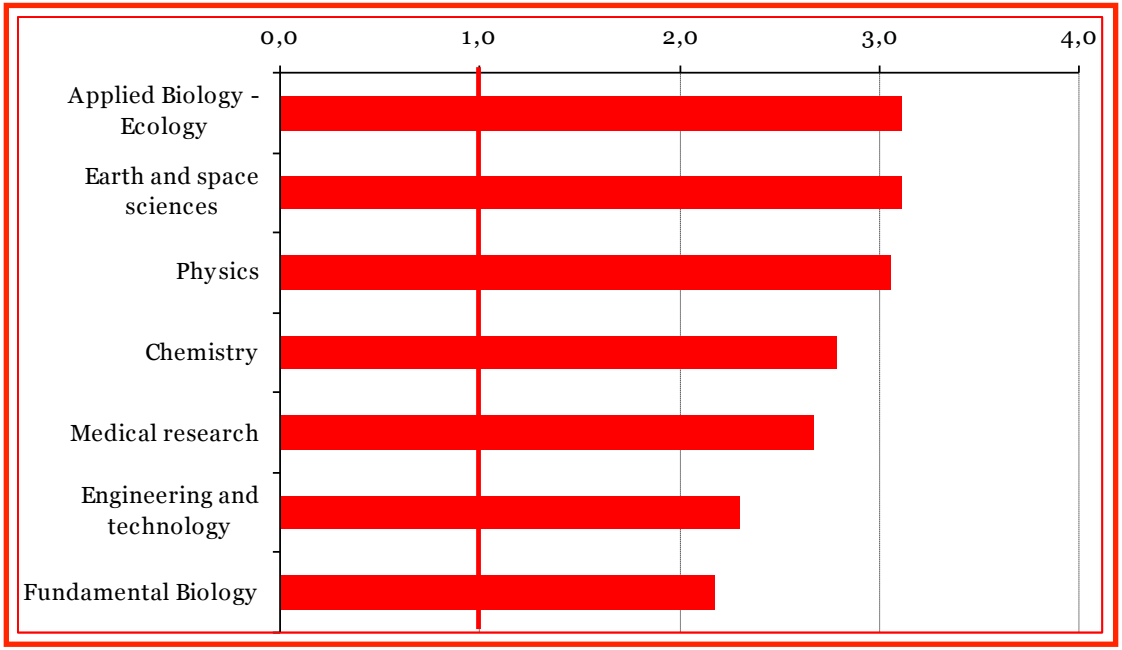
=> by country, 2006

•Compares the citations of Lead scientists to citations that are expected given the Journal in which they publish

technopolis |group|

The rate of publications in the 1% most cited publications is usually higher for FP6 Lead scientists than for the scientists of the same large field of science

Question : How do the publication profiles of lead scientists compare to the average performance of authors in the same field of science?



Activity index = comparison of:

- the distribution of the Lead scientists publications
- & the distribution of the entire scientific community publications

=> in Top1% class of citation, by fields of science (2006)

technopolis |group|

Road map

- Some historical background
- Evaluating RCN
- Performance-based research funding
- R&D evaluation challenges

technopolis |group|

Key characteristics of PRFS (1)

- Applied in many countries
  - *Austria, Australia, Denmark, Finland, Flanders, France, Hong Kong, Italy, New Zealand, Norway, Poland, Slovak Republic, Spain, Sweden, UK*
- Used in different ways
  - *A competitive source of descretionary income*
  - *A reward for quality and / or volume of output*
  - *An instrument of policy*
  - *Or a mix of these (UK: all three)*
- Used at different levels
  - *Individual performance (Spain)*
  - *"Units of assessment", e.g. department, research group (UK, Hong Kong)*
  - *Institutions (most systems)*

technopolis |group|

Key characteristics of PRFS (2)

- Typical output measures applied
  - *Volume*
  - *Quality*
  - *Impact*
  - *Utility*
- Trend: from peer-review to indicator-based systems
  - *To de-politicise and to de-personalise (esp. CR)*
  - *To cut costs and to simplify*
- Intervals between successive allocations & period assessed vary
  - *CR: annual decisions, based on number of results of past five years*
  - *AT: decisions every three years, based on the differences between the past two funding periods*
- Percentage of funding moved in one allocation round (→)

technopolis |group|

How much money is moved by a PRFS?

- International practice: modest amounts of money are moved, compared to total institutional funding
  - *NO: reallocation is limited to 2% of university income*
  - *AUS, NZ: 10% of national university funding governed by PRFS*
  - *AT: formula based funding rules 18'% of national "block" funding for universities and the max. annual reduction is limited by law to 2% p.a.*
- The situation in the Czech Republic
  - *Implementation of the new system in various phases and with differences between institutions*
  - *By 2011, for most research institutions, 2/3 of institutional funding will be determined by the PRFS – for others even 100%*
  - *Rough estimate: 30% of universities' and 80% of the Academy's total national public funding affected*

technopolis |group|

Effects of PRFS

- PRFS are intended to affect performance
- PRFS are relatively new, hence, there is a limited amount of evidence about their effects
- UK: measured quality of university research has clearly increased
- The effects on collective behaviour can be surprisingly large
  - *Potential funding change*
  - *Role of rankings & grades as indicators of esteem*
- … and not all effects might have been intended
  - *See the example of Australia*
- Care needs to be taken to avoid such unintended, perverse effects
- Critical factor: inter-field differences (→)

technopolis |group|

How to tackle inter-field differences in PRFS?

- Differences between scientific fields are large and need to be taken into account
- Approach 1: use methods for field normalisation
  - *Either in calculating indicators*
  - *Or in the translation of indicators into the allocation of money*
  - *NO considered field normalisation but decided it was too hard*
- Approach 2: avoid putting fields in competition with each other
  - *Requires separate decisions on the allocation of money among fields*
  - *UK uses approx. 60, Sweden 34 disciplinary categories*
  - *CZ EM 2010 is an attempt in this direction, but with only 10 categories → still inter-field competition*

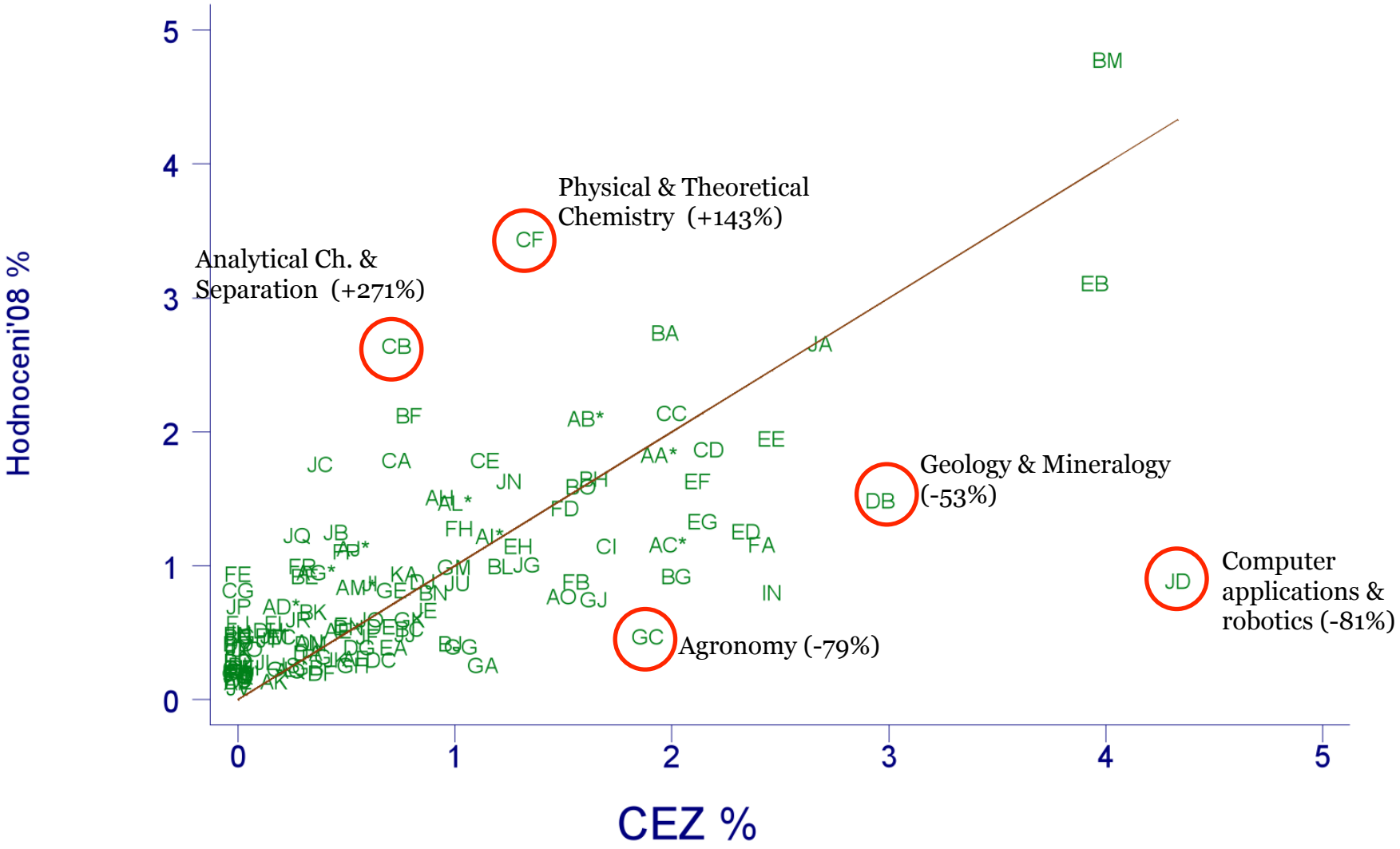
technopolis |group|

Inter-field differences in the Czech EM?

- How would the Czech EM affect different fields?
- Simulations by D. Münich: *Project for the preparation of the Methodology to evaluate the results of research institutions and of programmes finished in 2010*, Secretariat of the Board of the R&D&I Council, 2010
- Simulations based on the EM 2008
- Major conclusions valid for EM 2010:
  - *differences within the ten fields defined by the new EM remain large (e.g. history vs. economics within 'Humanities and Social Sciences')*
  - *J_{imp} are excluded from the 'damping', but they account for the lion's share of point values (65% in 2009!)*

technopolis group

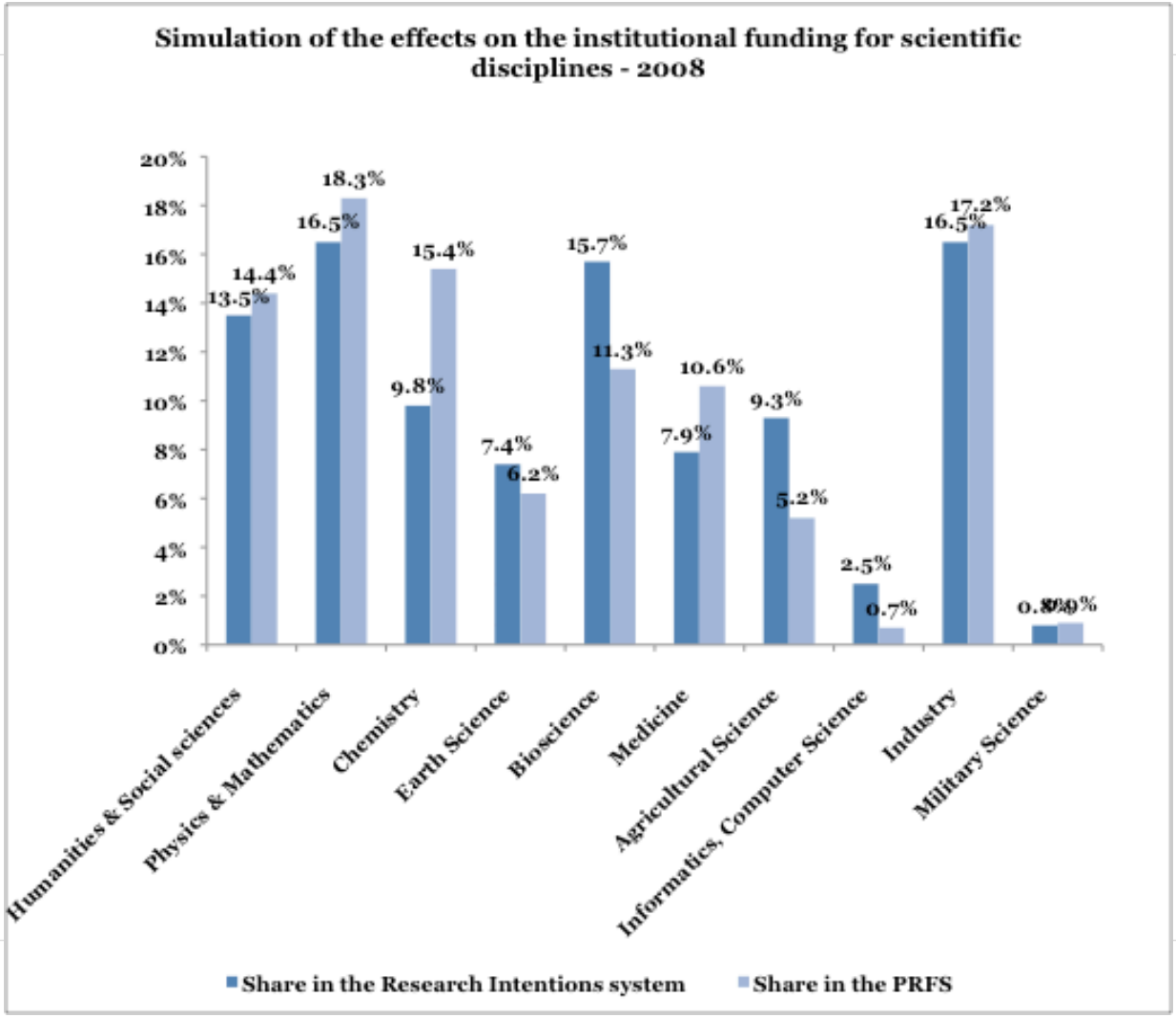
Field specific shares (in %) of institutional funds allocation by Research Intentions and by the PRFS



Source: Report of the "Project for the preparation of the Methodology to evaluate the results of research institutions and of programmes finished in 2010", Secretariat of the Board of the R&D&I Council, 2010

technopolis |group|

Shift in focus of institutional funding for research



Source: "Project for the preparation of the Methodology to evaluate the results of research institutions and of programmes finished in 2010“ Secretariat of the Board of the R&D&I Council, 2010

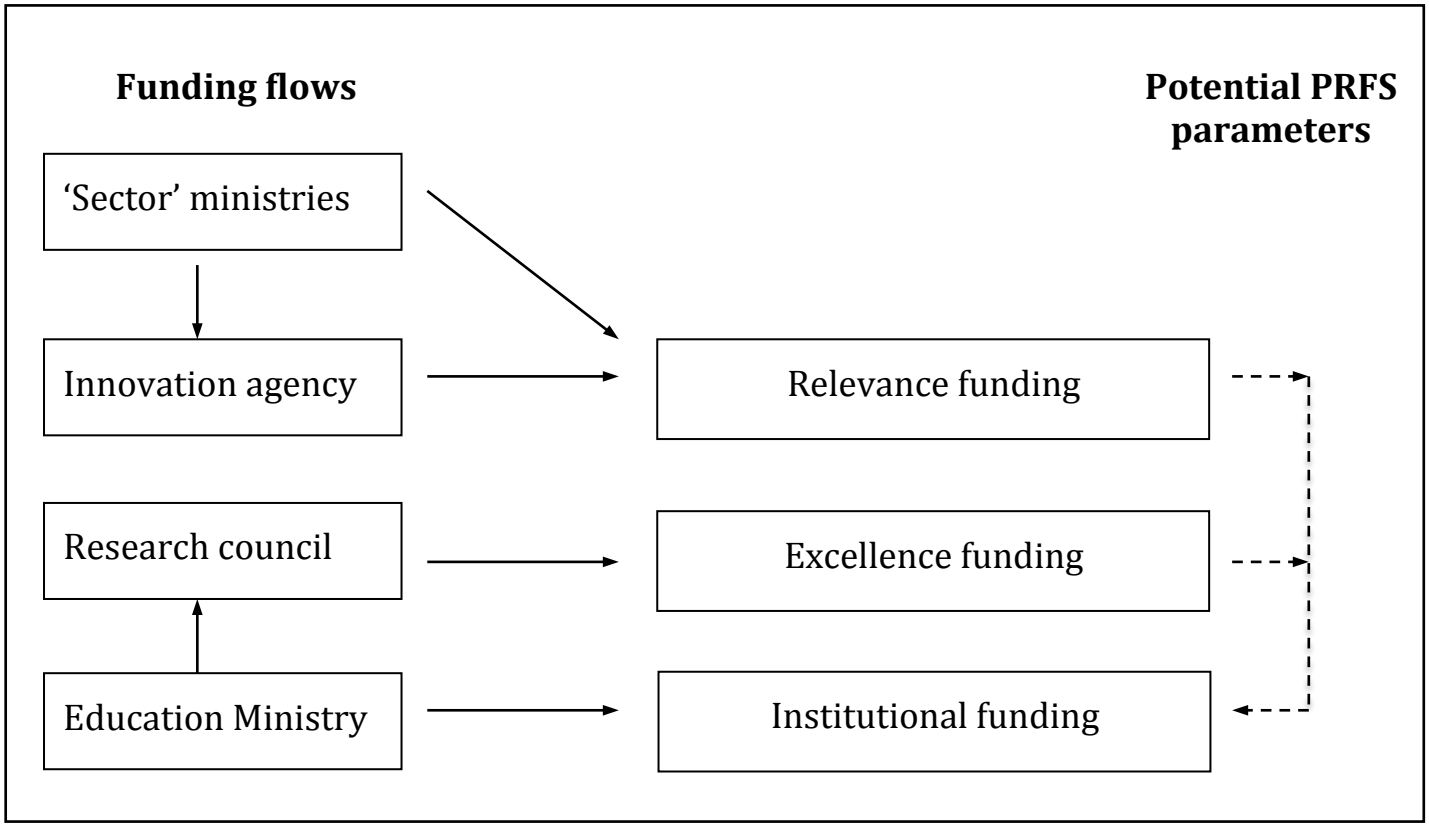
technopolis |group|

Main preliminary findings from the simulation of the CZ system

- Expected shifts in the focus of institutional funding
- "Winners":
  - *Chemistry: inorganic c., analytical c. & separation, physical & theoretical c.*
  - *Medicine: traumatology & orthopedics, psychiatry, pharmacology etc.*
  - *Physics & mathematics: Theoretical ph., elementary particles & high energy ph., general mathematics*
  - *Humanities & Soc. Sci.: Politology, Legal Sci., Letters*
- "Losers":
  - *Bioscience: (all but immunology), among them genetics & molecular biology*
  - *Agricultural sciences: (all but plant pathology)*
  - *Informatics & computer sciences*
  - *Industry: computer application and robotics*
- Some PRFS results would contradict the thematic priority set in the National Priorities for Applied R&D&I 2009 -2011 (highlighted cases)

technopolis |group|

So far, a lack of systems thinking about PRFS



technopolis |group|

Road map

- Some historical background
- Evaluating RCN
- Performance-based research funding
- R&D evaluation challenges

technopolis|group|

Challenges for R&D evaluation

- Internationalisation and globalisation of programmes
- Increased self-organisation and empowerment of stakeholders
  - *Dissipates responsibility for evaluation*
  - *Legitimacy*
  - *Means new actors have to learn*
  - *Leads to data loss – just when we're finally trained the agencies!*
- Reintegration of evaluation and strategic planning
  - *Better learning*
  - *More technical content*
- As ever – delivering simpler message to the political level through increasingly sophisticated evaluations

Methods challenges

- Connecting micro- to macro-phenomena in economic and social analyses
- More sophisticated qualitative analytic methods
- Bibliometrics
  - *Cheapening and reduced barriers to entry*
  - *Moving the unit of analysis towards micro-phenomena*
- Making sense of the Web
- Moving social network analysis from pretty pictures to having theories and evidence about what networks mean
- Paradox of using linear intervention logics while claiming to live in a complex systems world

technopolis |group|

Thank you

technopolis |group| has offices in Amsterdam, Ankara, Brighton, Brussels, Frankfurt/Main, Paris, Stockholm, Tallinn and Vienna