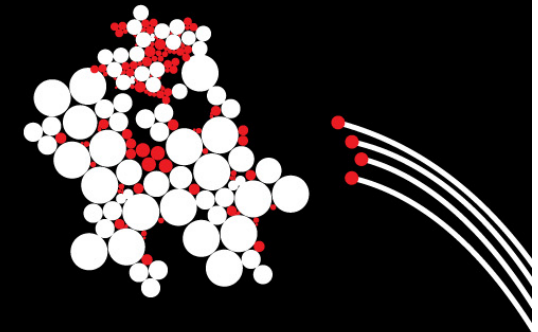


UNIVERSITY OF TWENTE.



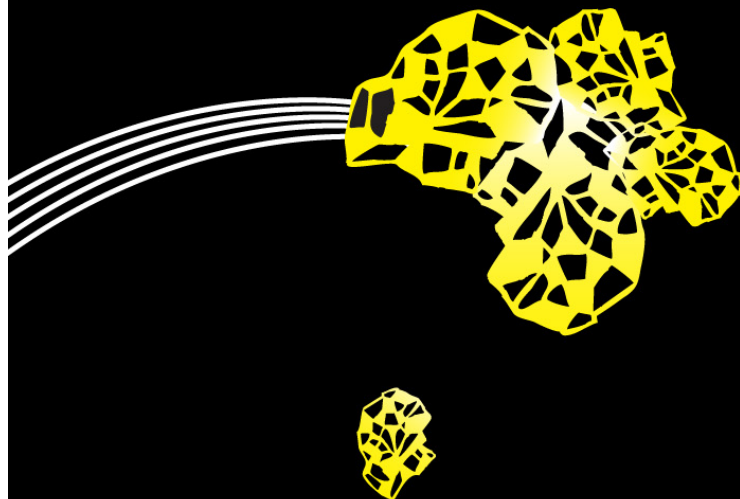
PER-Base: Data Base on Peer Evaluations of Research in the Netherlands

CHERPA-project

Marika Faber & Don F. Westerheijden

Presentation to CHERPA Symposium

Leiden, 2011-05-23





PER-Base: Data Base on Peer Evaluations of Research in the Netherlands – Contents of Presentation

- Project history and outlook
- Evaluation of research in the Netherlands
- Hypotheses





Project history and outlook

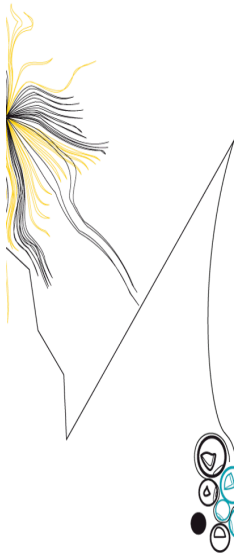
- Amazingly, a data base on the research evaluations did not exist
 - Although the evaluations are in the public domain
 - OCW, VSNU, QANU, KNAW
 - Thankful for cooperation by all, and by individual universities, academic hospitals
- CHERPA funding enabled CHEPS to set up this database
 - For analysis of the evaluations
 - This is what we present today
 - To link with other analyses of science, e.g. CWTS's analyses
- CHERPA funding ends, what happens to the database?
 - Rathenau Instituut shows interest in maintaining it





Evaluation of research in the Netherlands

- Research evaluations were the first formal, regular external evaluations in Dutch higher education
 - Conditional Funding policy (1982-1992)
 - Lasting impacts
 - Clustering of research in 'programmes'
 - New outside hard sciences
 - Broad umbrellas developed into meaningful programmes
 - Share of programmed research increased
 - Not successful: 'dynamisation' of funding



Evaluation of research in the Netherlands

- Self-regulation (since 1993)
 - VSNU acquired coordination of research evaluations
 - Meant to inform institutional research management, not for governmental policy
 - Introduced 4 dimensions, 5-point scales + verbal comments
 - Initial rounds organised by disciplinary area
 - Parallel to external quality assessment of education
 - Evaluation overload: VSNU, KNAW, NWO
 - Led to development of SEP
 - Since 2003: responsibility universities, no longer national approach per discipline, new rating.

Evaluation of research in the Netherlands: PER-Base

- PER-Base includes:
 - Covers period 1993-2010
 - Ca. 180 evaluation reports
 - Ca. 3640 evaluations of research programs / institutes
 - Information on: discipline, year publication, title publication, university /research institution, ratings on four criterion: Quality, Productivity, Relevance, Viability / Vitality & Feasibility.

VSNU 1992-2002		SEP 2003-2015	
5	Excellent	5	Excellent
		4	Very good
4	Good	3	Good
3	Satisfactory	2	Satisfactory
2	Unsatisfactory	1	Unsatisfactory
1	Poor		

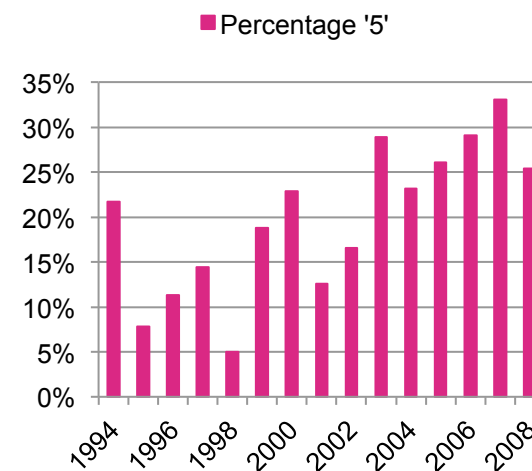
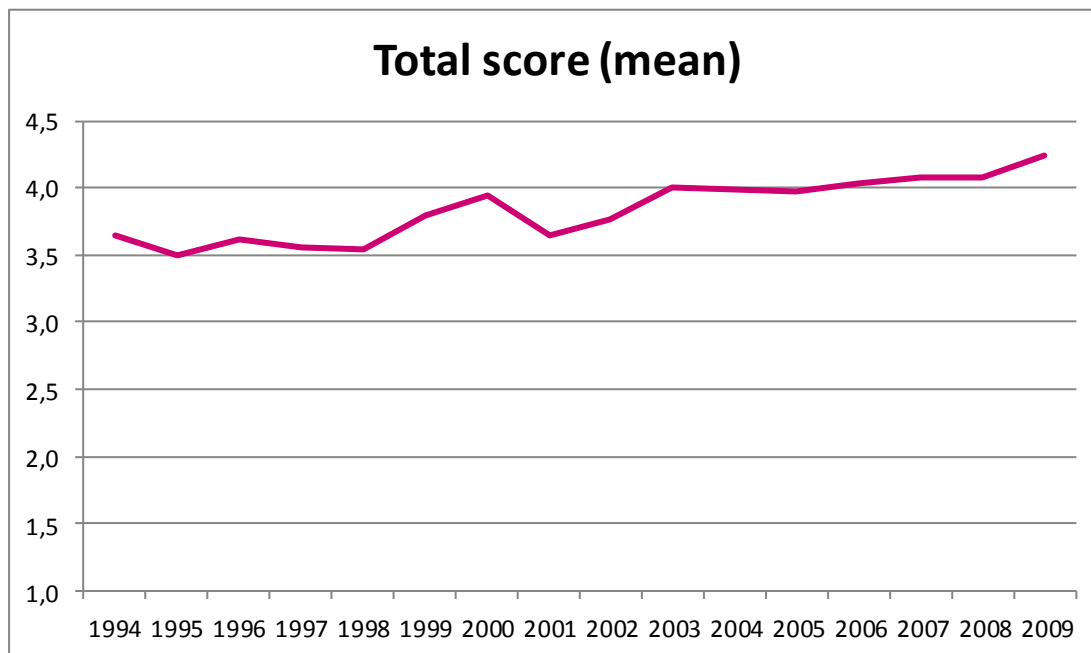


Hypotheses

- Grade inflation
 - Over time, ratings become higher
- Hard vs soft sciences: Disciplinary areas show different cultures of evaluation, in parallel to the task uncertainty in the discipline, ergo
 - Hard sciences, with lower task uncertainty, show less uncertainty about ratings, resulting in stricter/lower ratings than soft sciences



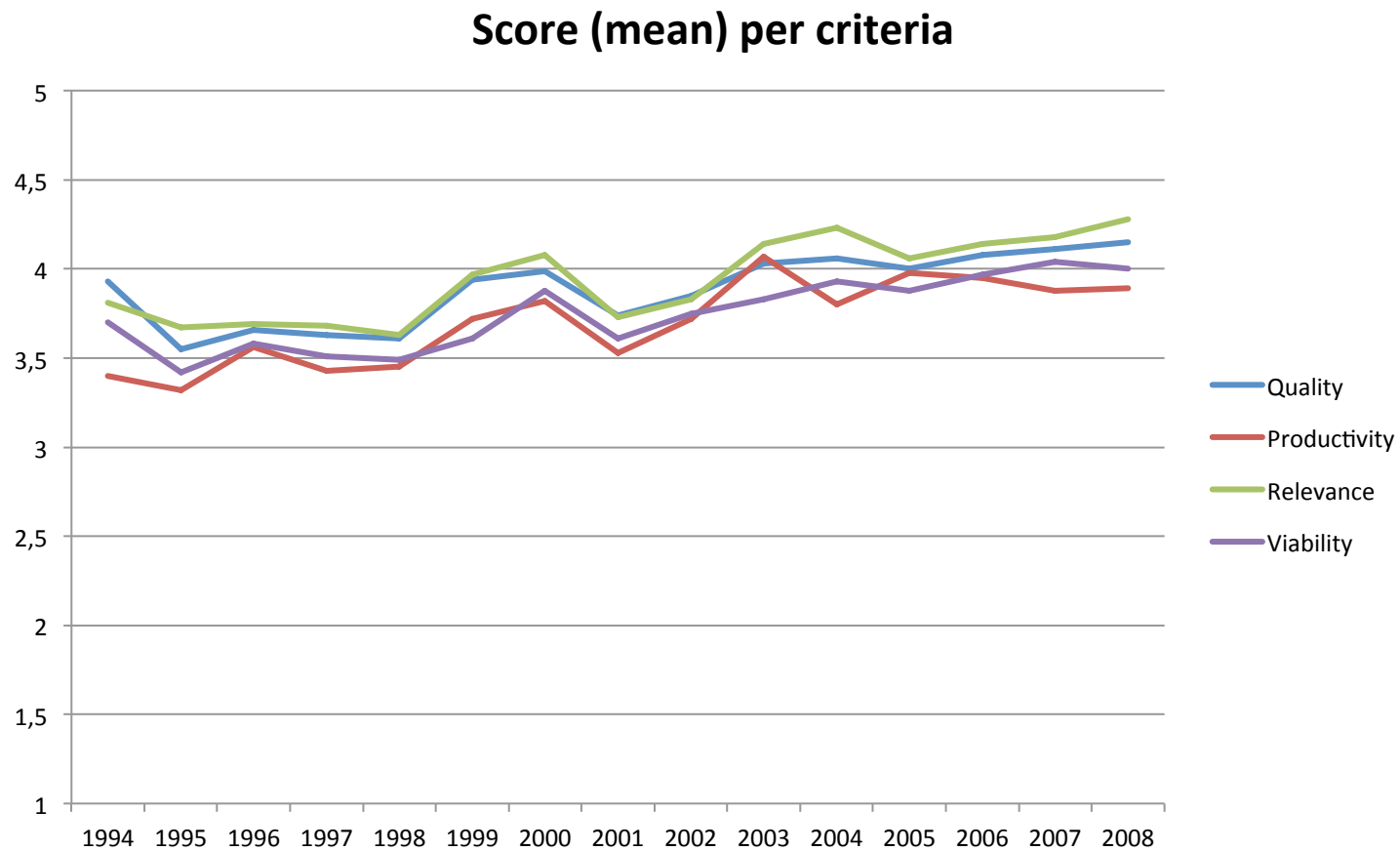
Grade inflation thesis



Number of evaluations (n)

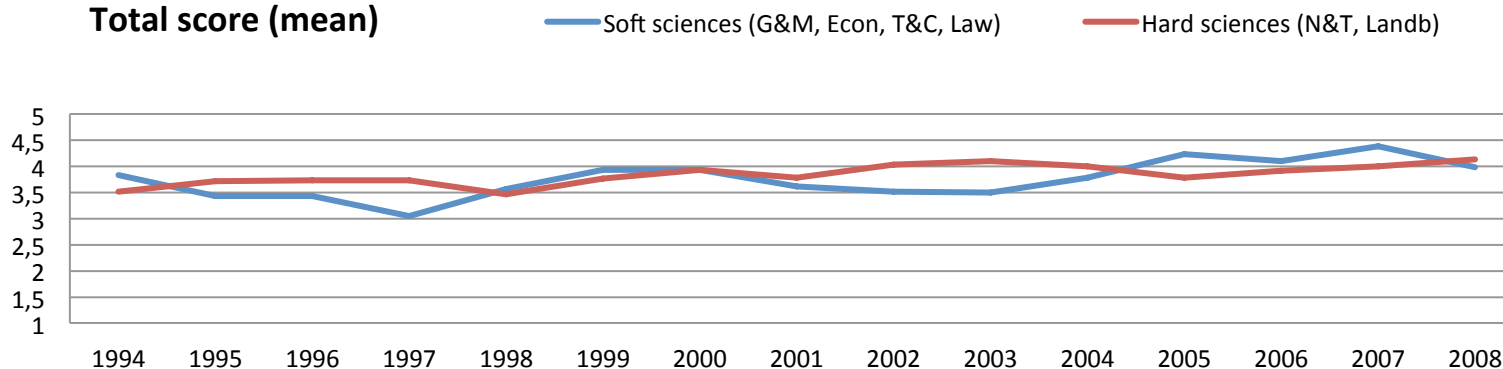
1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	Total
232	330	542	152	166	206	174	114	433	49	213	167	289	146	142	268	3623

Grade inflation thesis

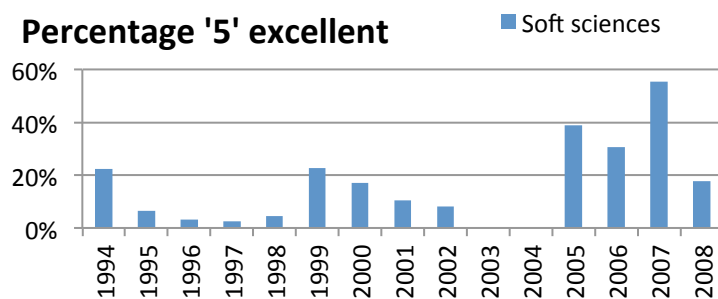


Hard science, hard judgement thesis

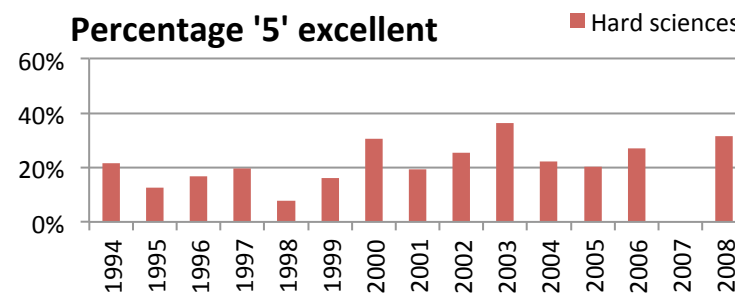
Total score (mean)



Percentage '5' excellent



Percentage '5' excellent

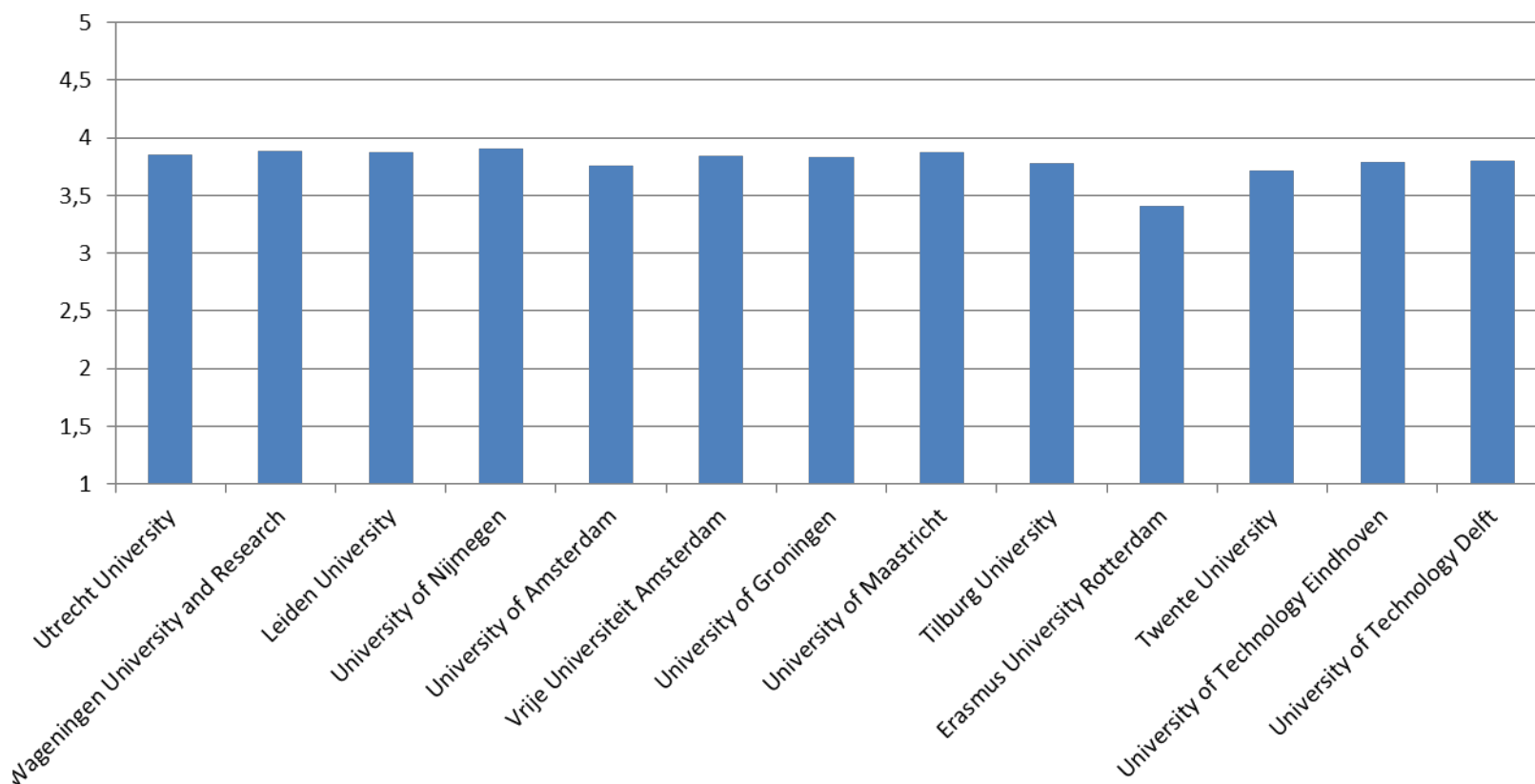


Evaluations (n)

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Soft sciences	122	265	235	41	140	58	84	87	228	11	8	41	143	38	75
Hard sciences	105	65	307	105	26	114	75	26	199	14	158	94	79	21	63

Evaluations per university

Total score (mean) per university



Concluding remarks

- Further analyses will be done in CHEPS...
- ...and also in cooperation with the Rathenau Instituut