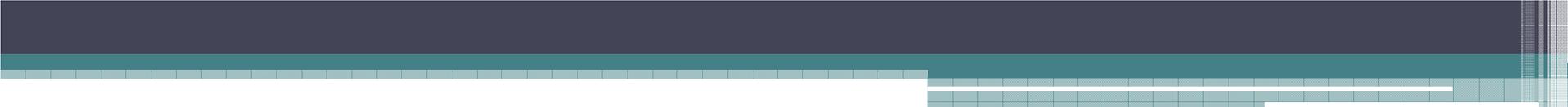# Soft and Hard Data and Definitions for University Rankings

Robert D. Reisz

Institut für Hochschulforschung
an der Martin Luther Universität
Halle Wittenberg, Germany

Introduction

- Complexity overwhelms. Just take the United States: over 4000 colleges and universities supply higher education.
- College and university rankings, as well as accreditation mechanisms, are intended to reduce complexity and criticized for reducing it too much.
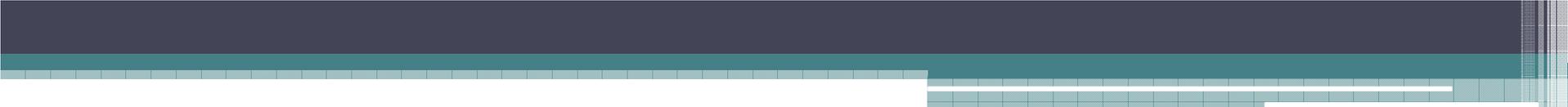
# Outline of the presentation

- Rankings and the data they use
- The nature of soft and hard data
- What do rankings measure?
  - One concept of „quality" or more?
  - Is „qualiy of higher education" a substantive concept?
- Conclusions and comments
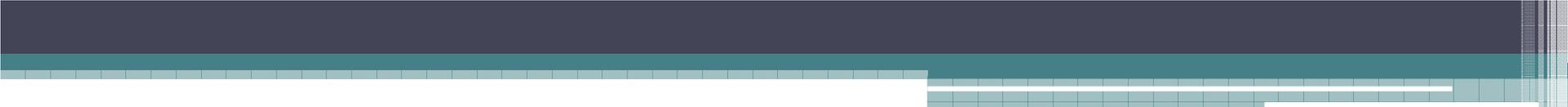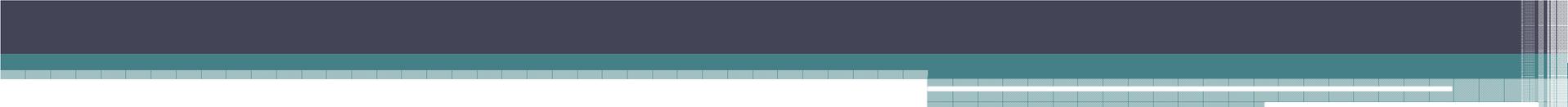
Rankings and the data they use

- While the number of universities might be overwhelming the number of different ranking systems also seems overwhelming at first site.
- Nevertheless, on a closer scrutiny these can be classified in:
  - universal rankings, intending to compare all universities in the world,
  - national rankings and
  - specialized rankings, comparing colleges from a particularistic point of view, be it scientific, professional or life-style related, such as Young America's Foundation's top ten conservative colleges.
- There also are rankings that rely only on one factor as, the internet presence of universities, the as G-Factor, or as Wuhan Universities ESI (Essential Science Indicators) based ranking.
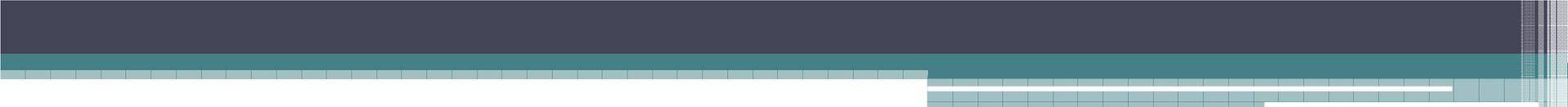
Rankings and the data they use

The Times Higher Education Supplement - QS World University Ranking (THES-QS)

- Peer Review Score (40%): "Composite score drawn from peer review survey (which is divided into five wide subject areas)"
- Faculty/Student Score (20%): "Score based on student faculty ratio"
- Citations/Faculty Score (20%): "Score based on research performance factored against the size of the research body"
- Recruiter Review (10%): "Score based on responses to employer survey. 2,339 responses in 2008."
- International Faculty Score (5%): "Score based on proportion of international faculty"
- International Students Score (5%): "Score based on proportion of international students"
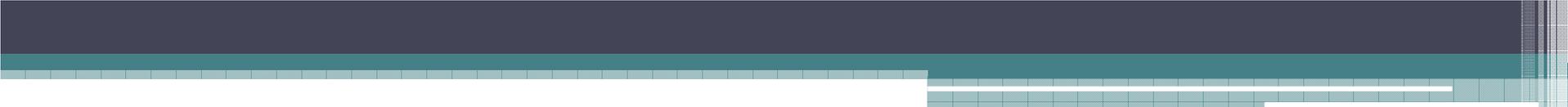
Rankings and the data they use

The Academic Ranking of World Universities of the Shanghai Jiao Tong University (ARWU)

- Alumni of an institution winning Nobel Prizes and Fields Medals (10%)
- Staff of an institution winning Nobel Prizes and Fields Medals (20%)
- Highly cited researchers in 21 broad subject categories (20%)
- Articles published in *Nature* and *Science* (20%)
- Articles Indexed in Science Citation Index-Expanded and Social Science Citation Index (20%)
- Academic performance with respect to the size of an institution (10%)

Rankings and the data they use

American's Best Colleges by the US News and World Report

- Peer assessment survey
- Student selectivity:
  - Acceptance rate (ratio of admitted students to applicants)
  - High school class standing—top 10%
  - High school class standing—top 25%
  - SAT/ACT score (average)
- Faculty resources:
  - Faculty compensation
  - Percent faculty with top terminal degree
  - Percent full-time faculty
  - Student/faculty ratio
  - Class size, 1-19 students (percentage of classes with less than 20 students)
  - Class size, 50+ students (percentage of classes with over 50 students)
- Graduation and retention rate
  - Average graduation rate
  - Average freshman retention rate
- Financial resources
- Alumni giving as percent of alumni who donate money.
- Graduation rate performance

Rankings and the data they use

Top American Research Universities by The Center for Measuring University Performance (The Center)

- Total research and development expenditures (in USD)
- Federally sponsored research and development expenditures (in USD)
- Number of members of the National Academies among an institution's faculty
- Number of significant faculty awards earned
- Number of doctorates awarded
- Number of postdoctoral appointments
- Median SAT scores of students
- Size of the institution's endowment
- Annual giving

Rankings and the data they use

CHE University Ranking by the Centrum für Hochschulentwicklung

- Indicators resulting from a questioning of the faculties and official statistical data:
  - Total number of students
  - Percentage of women
  - Number of first year students
  - Ratio of admitted students to applicants
  - Number of graduates
  - Percentage of graduates in the formal duration of studies
  - Research funding / professor
  - Patents / staff member
  - Scientific publications / professor and research staff member
  - Citations / publication
  - Ph.D.s awarded / professor
  - Student / staff ratio
  - Etc.
- Indicators resulting from a questioning of professors:
  - Research reputation, each professor will be asked to name the five German universities that are leading in its subject area.
  - Reputation of professors, according to the same method as above.
- Indicators resulting from a questioning of students:
  - Quality of teaching measured by a number of indicators
  - Quality of e-learning supply if available
  - Evaluation of study supply, different criteria
 - Evaluation of contacts between students
  - Etc.

Soft Data and Hard Data

- The first conclusion from the previous overview of variables used by five well-known ranking systems is the impressive variety of data used.
- In social science quantitative data are used to represent, to operationalize theoretical concepts.
- All these data are used to operationalize the same concept, that of a "good university".
- While all ranking agencies are very careful in not overstating the meaning of their results, the basic assumption remains: all these are ways to measure "quality in higher education".
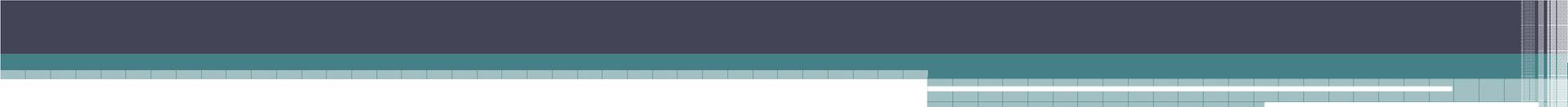
Soft Data and Hard Data

- Let us take the reverse of the usual empirical design, look at the operationalization and find properties of the concept it measures.
- What is the nature of quality in higher education if this is the way it is measured by some of the most influential agencies?
- This question is all the more interesting as research results show that it is common in many higher education systems to learn to "play the ranking game" in order to improve one's organizational position by speculating ranking criteria (or accreditation criteria or classification criteria, for that matter)
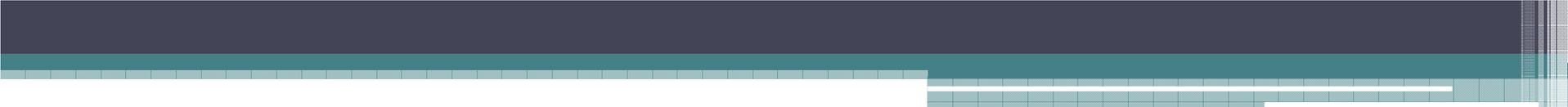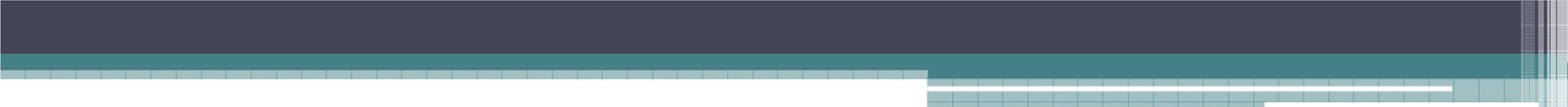
Soft Data and Hard Data

While data are very varied there are two major categories that appear in many methodologies. These are:

- Soft or survey data resulted from polls of academics, higher education administrators, students or potential employers.

- Hard or factual data resulted from official or organizational statistics.
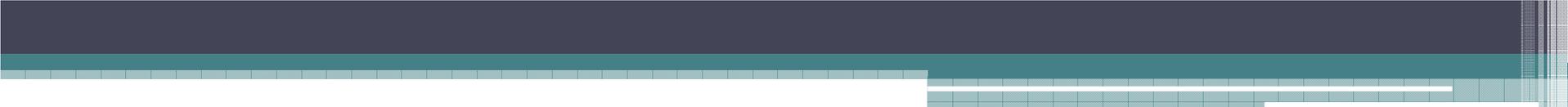
Soft Data and Hard Data

- In sociology soft data are information that are considered to be socially represented. Surveys are used to approximate the distribution of such variables in a population.
- As it is so, soft data are dependent on the way they are collected. The apparatus to measure such variables is a part of the result itself.
- An opinion collected with a questionnaire depends as much on the data collection method (questionnaire, sampling, selection, interviewing, etc.) as it does on the preexisting opinion in the population subject to the research.
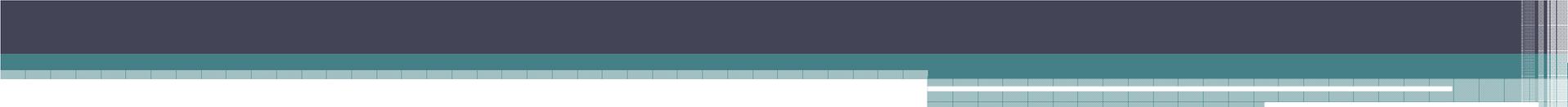
Soft Data and Hard Data

- Hard data, on the other hand, are supposed to be roughly independent from the way they were measured.

- To give examples from the ranking systems, the number of students and teaching staff should not depend on the way data was collected, be it by reading official statistics, phoning in or writing to the administration of the faculty.
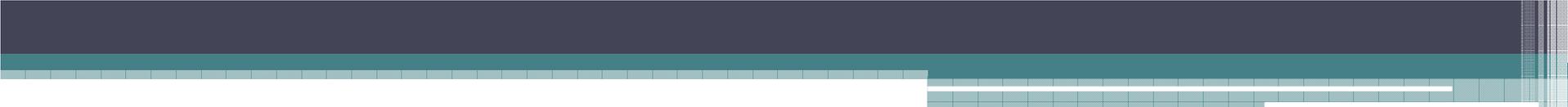
Soft Data and Hard Data

- The traditional view regarding academic quality is not that of a socially constructed or even socially represented quality, but that, that only peers are able to judge academic quality.
- Simply said a certain "amount" of quality is considered to exist in the substance of a research paper, or an exam paper or a book, the peer reviewer is the expert that is able to judge it.
- The soft data used in rankings follows the same logic. This is most obvious in the fact that none of the ranking systems that use soft data and that we have studied gives importance to sampling methods.
- The opinions resulted from the polls are not statistical sample survey results, they are peer reviews.
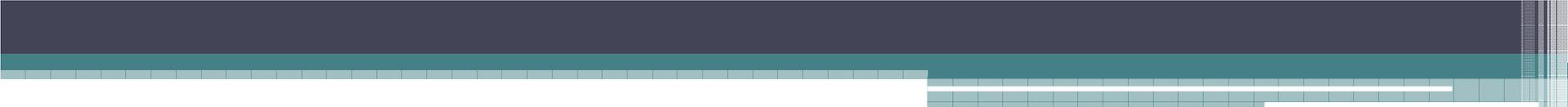
Soft Data and Hard Data

- Many of the hard indicators result from forms of collective action, e.g. indicators that are intended to measure research. The most common of these are citation numbers, publication numbers in peer-reviewed journals and different awards and memberships.
- One of the central points of this paper is that these data do not differ at all from the previously presented soft data.
- All these: citations, publications, awards, rely on the same logic of peer reviewing that is the central indicator of academic merit.
- Deciding to cite a paper, acceptance for publication or the receipt of an award are all decisions taken by relevant peers.
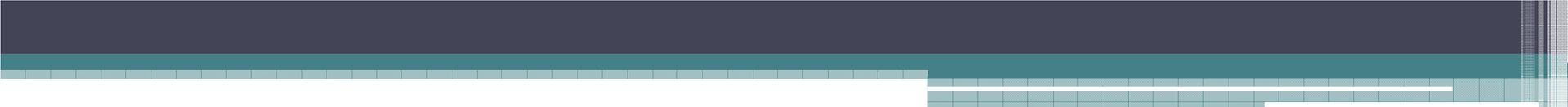
Soft Data and Hard Data

- Some research indicators, e.g. those used by the Center might seem different in nature at a first look.
- Nevertheless, receiving research funding from any agency is in the end also the result of a review process during which academic criteria are applied on a research proposal by peers.
- A similar logic can be applied to understand the relevance of the results of students at college exams or third part examinations. The CHE-ranking, for example, also uses results of graduates at the formal exams for a medical or legal career as indicators of quality.

Soft Data and Hard Data

- College selectivity, an important issue in the American rankings and one up and coming subject of debate in many European countries is also a materialization of choices of would be students and their families. As such, selectivity is the result of a socially existing desirability of studying in a specific college.
- And if research indicators parallel the surveying of academics, alumni giving is an apt parallel of an alumni survey.
- So most hard data are in fact also results of collective action and social representations of the concept of quality.

Soft Data and Hard Data

- Not many indicators remain.
- Most important of these are student / staff ratios, percentages of small classes or of large classes, percentage of full time faculty.
- A very simple logic lies behind all these; they state that the teacher should be able to focus on its students.
- Scanning all indicators of the studied rankings, maybe with the exception of the CHE method, this is all that remains as hard, factual operationalization of quality of higher education.
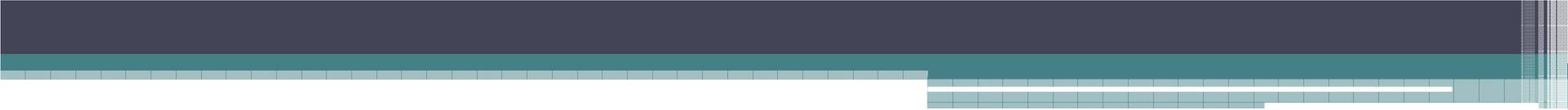
One "quality" or more?

- As mentioned, the ranking scales vary essentially in the indicators they use.
- Nevertheless, ranking scales do not consider that the quality they measure is a socially constructed concept but rather that it is a socially represented concept.
- The question that arises from the high variation of used indicators is if the ranking schemes refer to the same meaning of "quality of higher education" or not.

One "quality" or more?

- The simplest way to make an empirical check of the coherence of the different "quality" concept of ranking schemes is to consider that all succeed in reproducing in the rankings the quality they intend to measure and correlate the rankings between each others.
- The rank-order correlation of the two American rankings is 0.672** (2008).
- The rank-order correlation between the THES-QS and the ARWU is 0.583* (2007).
- If for the American case we control the influence of the indicator "total endowment" the correlation becomes statistically not relevant and almost disappears in size.

One "quality" or more?

- Let us here mention that the numerical values of the correlation coefficients might not look very low indeed, but they do not represent the relationship between different concepts but rather intend to test the coherence of two different proxy variables for the same concept.
- The simple empirical conclusion is: even if some relationship between the overall orders exists, this is not very high.
- It is safer to consider that rankings do not measure the same quality of higher education or if they do they do not measure it coherently.
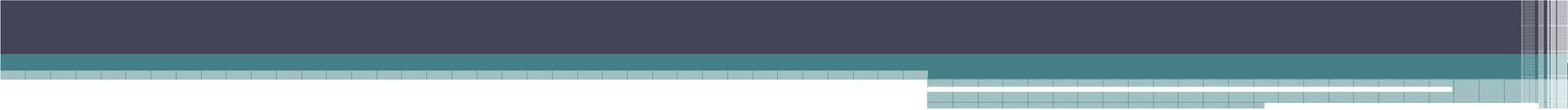
Is "quality of higher education" a substantive concept?

- While we can still accept that quality might have different facets and that rankings might still do a good job in each measuring some other facet of the overall concept of quality, there appears an even more important question.
- Is in fact quality an essential concept as all rankings seem to consider it? Does quality exist beyond the ranking systems themselves?
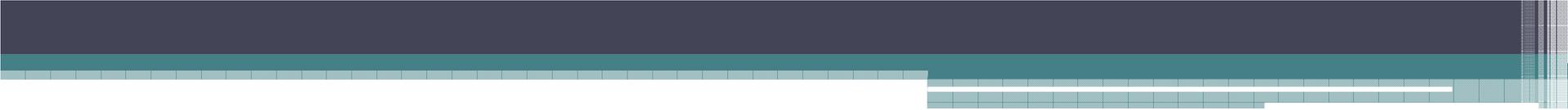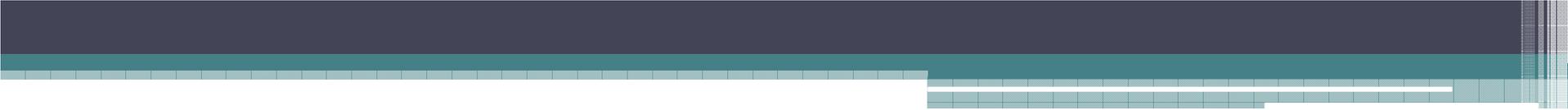
Is "quality of higher education" a substantive concept?

- If quality exists outside of the process of its evaluation and the ranking methods are sound and unbiased there should exist a stability of rankings over time.
- This mainly because we cannot imagine that in a system so stabile and even resistant to change as higher education, quality could change very much over a period of few years.
- In order to investigate this we have gathered time series of rankings and have done rank-order correlations of rankings of consecutive years.
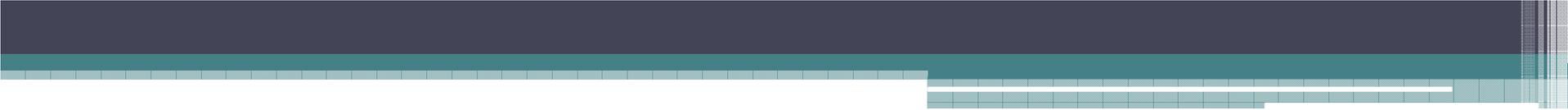
Is "quality of higher education" a substantive concept?

- The THES-QS year on year Spearman correlations are very high, e.g. the most recent 2008 on 2007 correlation is 0.954** if we consider the first four hundred universities.
- A year on year correlation of the ARWU system is 0.990** for 2008 on 2007, taking all 500 universities, while after the 99th position only the group of hundreds is mentioned. Other year on year correlations are as high.
- The US News and World Report ranking also shows very high year on year correlations, the Spearman coefficient for 2007 on 2008 being 0.992**. As in the case of the THES-QS ranking, older year on year values are as high or almost as high.
- Computing the year on year correlation of the Center's ranking of top research universities leads obviously to the same high Spearman coefficients. The value for the last years in our time series (2007 on 2006) was 0.984** for the number of top 25 rankings and 0.924** for top 26 to 50 rankings.
- The CHE rankings do not produce simple lists, so I could not compute similar correlations.

Is "quality of higher education" a substantive concept?

- The study of the time coherence of ranking scales shows very high year on year correlations and a natural reduction of these correlation with growing time delay.
- Do all these computations mean that "quality in higher education" is an essential concept having a real independence from its evaluation?
- Not necessarily, but it means that we cannot refute such a hypothesis.
- While rankings do not measure the same thing, or at least we have to accept that they measure different facets of the quality concepts, on the basis of our evidence we cannot refute the existence of a certain "quality in higher education" beyond its measurement.
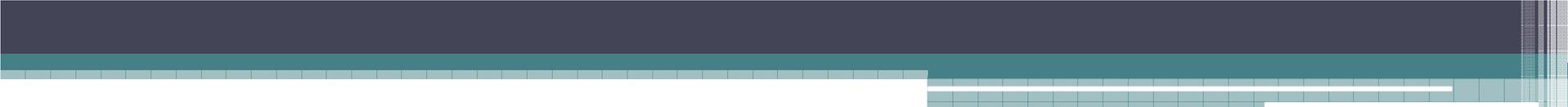
Conclusions and comments

- Ranking scales use very different indicators.
- Almost all these indicators are not substantive, but social representations of the concept of quality in higher education. They rely not on measuring what quality is but rather on measuring what peers or stake-holder consider as being quality.
- The collection of data is for most of these indicators not statistically sound, self-selection and low case numbers being widely present. The methodological documents of ranking do not consider sampling as essential for data collection.
- Different ranking scales do not correlate very highly between each other and cannot be considered as proxy-s for the same concept.
- Each of the ranking scales is coherent in itself, showing high year on year stability and a linear reduction of rank-order correlations with time delay.
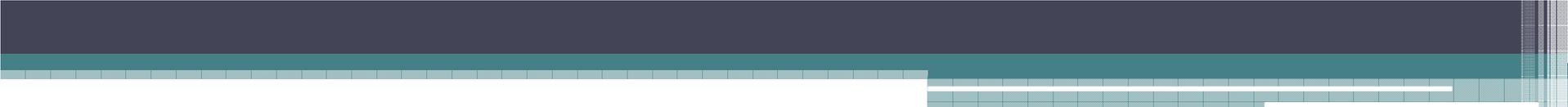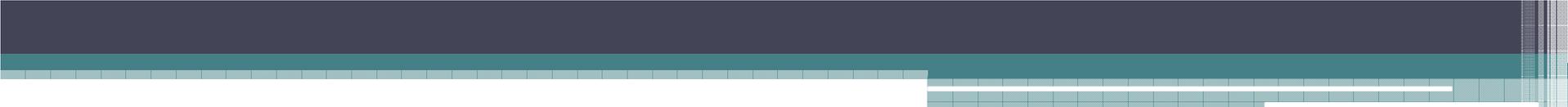
Conclusions and comments

- All these findings lead to a conceptualization of "quality in higher education" as a multivalent concept, socially represented but not necessarily socially constructed.
- This gives "quality in higher education", metaphorically speaking, a status similar to concepts of the natural sciences.
- It is typical for the natural sciences that a scientific definition, be it from the world of physics, chemistry or other hard science while being the formalization of a scientist or a group of such and being established by the social group of practitioners of the respective science, has nevertheless an independent empirically provable existence.

Conclusions and comments

- And yet, there is a big difference between the objects of the natural sciences and the "quality of higher education".
- While in the natural sciences, the definitions of scientific objects are accepted without any problems by the scientific community and have a univocal meaning "quality of higher education" has no such simple, accepted definition.
- As we have seen, the different rankings deal with different meanings or at least different facets of the concept. By using in the rankings largely soft data and hard data that result from the aggregated actions of many individuals, these methodologies seem to consider that "quality of higher education is what people who know about higher education consider quality of higher education to be".
- Such a consideration seems at first to contradict our previous findings.

Conclusions and comments

- Putting the objective existence of quality together with the fact that it cannot be defined in such a way as to be accepted by all parties seems rather inconsequential.
- But is not this the general case with quality in the arts, for instance? According to Bourdieu the formation of artistic value is a result of such an interplay between a field of artistic production and one of artistic validation.
- As with "quality of higher education", the "quality of an object of art" is supposed to have an objective existence, can still not be defined and is nevertheless recognized by experts.
- Hence, rankings treat higher education as an art form, something valuable in itself, having an objective quality but nevertheless the kind of quality that only the expert eye can judge and not even the expert can satisfyingly define.