Factors Influencing the Stability and Instability of International Rankings in Higher Education

Candidate Number: 698492

Dissertation submitted to the University of Oxford in part-fulfillment of the requirements for the degree of M.Sc. in Education Research Methodology

Trinity Term, 2009

Abstract

Whether rankings can reflect the real-life situation within universities or create an artificial illusion of universities is a question in our mind when discussing rankings of universities. This question is even more interesting when associated with a discussion of the world class universities. The two most influential rankings of world universities: the Academic Ranking of World Universities (ARWU) by Shanghai Jiao Tong University and the World University Rankings (WUR) by Times Higher Education-Quacquarelli Symonds commenced in 2003 and 2004 respectively. The short period of publication of these rankings means little research has been conducted to answer this question. This dissertation addresses this issue.

To grasp the features of the ARWU and the WUR, various statistical methods are employed first, followed by in-depth multiple case studies to verify the values in the two rankings and to test if they reflect the actual situation in universities. To ensure the accuracy of data, data for the multiple case studies are acquired from different sources for triangulation.

It is found that all indicators in the ARWU are positively correlated. The multiple regression analysis reveals that the actual weightings in the ARWU correspond with the alleged weightings. Two components are extracted by the principal components analysis. The biggest contributors to the first component are Alumni Prize Winners and

Staff Prize Winners. The biggest contributor to the second component is Publication. The two components are positively correlated. Thus, it is reasonable to combine the indicators into one ranking.

In contrast, not all indicators in the WUR are positively correlated. The correlation coefficients between International Student Number and International Staff Number with Citation are negative. The multiple regression analysis reveals that the actual weightings in the WUR do not correspond with the alleged weightings. The actual weightings change year by year. Three components are extracted by the principal components analysis. The biggest contributors to the first component are Peer Review and Employer Review. The biggest contributors to the second component are International Student Number and International Staff Number. The only contributor to the third component is Student to Staff Ratio. The three components are not correlated with each other. Actually, within the second components, the relationship between International Student Number and International Staff Number with Citation are negative. All the evidences show that the indicators of the WUR should not be combined into a single ranking.

The multiple case studies show that universities' rankings in the ARWU are relatively stable but shift dramatically in the WUR. The major factors help universities improve their rankings in the ARWU are Alumni Prize Winners and Staff Prize Winners. The dramatic changes of rankings in the WUR are not attributed to the actual changes in universities but the frequently changes in ranking methodology and the use of incorrect data.

Table of Contents

Abstract I
Table of Contents III
Abbreviations VIII
List of Figures and Tables IX
Figures IX
Tables X
Chapter One: Introduction 1
Research Rationale 1
The History and the Development of Rankings 1
Two International Rankings 4
Definition of Ranking 5
The Concept Underpinning Ranking 8
Criticism of Rankings 11
Research Questions 11
Research Design 12
Research Strategy 12
Data collection 13
Data Analysis 14
Overview of Study 14
Chapter Two: Literature Review 16
Theoretical Arguments 16
The World Class University 16
Principles of Selecting Indicators 17
Dilemmas 17
University-based versus Discipline-based 18
Rankings versus Clusters 18
Value-added or Not 19

Indicators 20
Classifications of Indicators 20
Denotation Approach 20
Input-Process-Output Approach 21
Objective-subjective Approach 23
A New Classification 23
Arguments about Indicators 25
Institution Related Indicators 25
Staff Related Indicators 26
Student Related Indicators 27
Impact of Ranking 29
Chapter Three: Methodology 31
Research Questions 31
Research Design 31
Data Collection 32
Data Analysis 34
Data Inclusion 34
The Method of Standardisation 34
Statistical Analysis 36
Correlation 36
Multiple Regression 36
Principle Component Analysis 37
Case Study 37
Chapter Four: First Stage Findings and Analysis 40
The Ranking Methodologies of the ARWU and the WUR 40
Indicators 40
The ARWU's Indicators 40
The WUR's Indicators 41
Ranking Method 42
Methodology Changes 42

The ARWU's Methodology Changes 43
The WUR's Methodology Changes 44
Results of Standardisation 44
Relationship between Total Scores and Rankings 49
Trends of the ARWU and the WUR 49
The Relationship between the Indicators 55
Results from Correlation 55
Results from Multiple Regression 57
Results from Principal Component Analysis 61
Chapter Five: Multiple Case Studies 64
Choosing Cases 64
Significant Changes over the Years 64
Universities for Case Study 68
University of California-Berkeley (UCB) 69
UCB in the ARWU 71
UCB in the WUR 73
University of California - Los Angeles (UCLA) and University of California- San Diego (UCSD) 76
UCLA in the ARWU 78
UCLA in the WUR 80
UCSD in the ARWU 83
UCSD in the WUR 84
University of California-San Francisco (UCSF) 86
UCSF in the ARWU 87
UCSF in the WUR 88
University of California - Santa Barbara (UCSB) 90
UCSB in the ARWU 91
UCSB in the WUR 93
University of California – Davis (UCD) and University of California – Irvine (UCI) 95

UCD in the ARWU 97
UCD in the WUR 99
UCI in the ARWU 101
UCI in the WUR 103
University of California – Santa Cruz (UCSC) and University of California Riverside (UCR) 104
UCSC in the ARWU 105
UCSC in the WUR 106
UCR in the ARWU 108
UCR in the WUR 109
Summary of the ARWU 110
Summary of the WUR 111
Chapter Six: Discussion 113
The ARWU and the WUR's Choice of Indicators 113
Size of Institution 113
Citation vs. International Student Number and International Staff Number 113
Peer Review and Employer Review 114
Validity 116
Validity of the ARWU's 116
Validity of the WUR's 116
Reliability 117
Questions for Future Research 119
What are the indicators of quality? 119
What is the appropriate ranking method? 120
Chapter Seven: Conclusions 121
Statistical Analysis 121
Relationship between the Indicators 121
Relationship between the Indicators and the Total Scores 123
Case Study 124
References 126

Appendices 133
Appendix A: Background of Rankings 133
Appendix B: Current Rankings in Different Countries 136
Appendix C: Webster's Categories of Indicators 140
Appendix D: Van Dyke's Categories of Indicators 142
Appendix E: Dill and Soo's Categories of Indicators 151
Appendix F: Usher and Savino's Categories of Indicators 156
Appendix G: New Classification 170
Appendix H:The Third Missions Indicators 180
Appendix I: VALUE Rubrics 183
Appendix J: The Indicators of the ARWU and the WUR. 184
Appendix K: Correlation 186
Appendix L: Multiple Regression 188
Appendix M: Principal Component Analysis 201

Abbreviations

ARWU	Academic Ranking of World Universities
FTE	Full Time Equivalent
HEIs	Higher Education Institutions
ISI	Institute for Scientific Information
LTRS	League Tables and Ranking Systems
NCES	National Center for Education Statistics
OECD	Organisation for Economic Co-operation and Development
PRSPWU	Performance Ranking of Scientific Papers for World Universities
UC	University of California
UCB	University of California—Berkeley
UCD	University of California—Davis
UCI	University of California—Irvine
UCLA	University of California—Los Angeles
UCR	University of California—Riverside
UCSB	University of California—Santa Barbara
UCSC	University of California—Santa Cruz
UCSD	University of California—San Diego
UCSF	University of California—San Francisco
WUR	World University Rankings

List of Figures and Tables

Figures

Figure 1 Instruments of quality assurance in higher education 9
Figure 2 2004 vs. 2008 WUR Total Score 45
Figure 3 Standardised Results of the WUR from 2004 to 2008 47
Figure 4 Standardised Results of the ARWU from 2003 to 2008 48
Figure 5 2003 ARWU Standardised Total Score vs. Ranking 50
Figure 6 2005 ARWU Standardised Total Score vs. Ranking 50
Figure 7 2007 ARWU Standardised Total Score vs. Ranking 50
Figure 8 2004 ARWU Standardised Total Score vs. Ranking 50
Figure 9 2006 ARWU Standardised Total Score vs. Ranking 50
Figure 10 2008 ARWU Standardised Total Score vs. Ranking 50
Figure 11 2004 WUR Standardised Total Score vs. Ranking 51
Figure 12 2006 WUR Standardised Total Score vs. Ranking 51
Figure 13 2008 WUR Standardised Total Score vs. Ranking 51
Figure 14 2005 WUR Standardised Total Score vs. Ranking 51
Figure 15 2007 WUR Standardised Total Score vs. Ranking 51
Figure 16 Six Years' Trends for the ARWU 54
Figure 17 Five Years' Trends for the WUR 54
Figure 18 UCB's Positions in Columns 70
Figure 19 UCLA's Positions in Columns 76
Figure 20 UCSD's Positions in Columns 77
Figure 21 UCSF's Positions in Columns 86
Figure 22 UCSB's Positions in Columns 90
Figure 23 UCD and UCI's Positions in Columns 96

Tables

Table 1 Chronology of Ranking Activities in the United States, 1870-1982 2
Table 2 Ranking Systems Worldwide, 2006 3
Table 3 Criteria and Weights for the 2004 ARWU 4
Table 4 Indicators and weighting 2008 THE-QS 5
Table 5 The Correlation between the ARWU and the WUR 55
Table 6 Alleged and Actual Weightings of the ARWU 59
Table 7 Alleged and Actual Weightings of the WUR 59
Table 8 The comparison of the results of the Principal Component Analysis for the ARWU and the WUR's 61
Table 9 Rankings of UCB 69
Table 10 UCB Yearly Changes in the ARWU 71
Table 11 UCB Yearly Changes in the WUR 73
Table 12 Triangulating Data for UCB 75
Table 13 Rankings of UCLA 76
Table 14 Rankings of UCSD 76
Table 15 UCLA Yearly Changes in the ARWU 78
Table 16 UCLA Yearly Changes in the WUR 80
Table 17 Triangulating Data for UCLA 82
Table 18 UCSD Yearly Changes in the ARWU 83
Table 19 UCSD Yearly Changes in the WUR 84
Table 20 Triangulating Data for UCSD 85
Table 21 Rankings of UCSF 86
Table 22 UCSF Yearly Changes in the ARWU 87
Table 23 UCSF Yearly Changes in the WUR 88
Table 24 Triangulating Data for UCSF 89
Table 25 Rankings of UCSB 90
Table 26 UCSB Yearly Changes in the ARWU 91

Table 27 UCSB's scores in the ARWU and the PRSPWU (HiCi: Highly Cited Researchers)..... 92
Table 28 UCSB Yearly Changes in the WUR..... 93
Table 29 Triangulating Data for UCSB..... 95
Table 30 Rankings of UCD..... 95
Table 31 Rankings of UCI..... 96
Table 32 UCD Yearly Changes in the ARWU..... 97
Table 33 UCD Yearly Changes in the WUR..... 99
Table 34 Triangulating Data for UCD..... 100
Table 35 UCI Yearly Changes in the ARWU..... 101
Table 36 UCI Yearly Changes in the WUR..... 103
Table 37 Triangulating Data for UCI..... 104
Table 38 Rankings of UCSC..... 104
Table 39 Rankings of UCR..... 105
Table 40 UCSC Yearly Changes in the ARWU..... 105
Table 41 UCSC Yearly Changes in the WUR..... 106
Table 42 Triangulating Data for UCSC..... 107
Table 43 UCR Yearly Changes in the ARWU..... 108
Table 44 UCR Yearly Changes in the WUR..... 109
Table 45 Triangulating Data for UCR..... 110

Chapter One: Introduction

Research Rationale

The History and the Development of Rankings

The ranking of higher education institutions (HEIs) has always been a controversial issue in academia. League tables, the synonym of rankings, imply that rankings tend to be associated with commercial games and mass media. Although the first well-known compilation of university ranking was the U.S. News, which compiled a league table of American undergraduate programmes in 1983, the early history of ranking actually can be traced back to the nineteenth century. According to Stuart (1995), the United States Bureau of Education began publishing an annual report of statistical data in 1870, which gradually developed into a form of ranking. In 1910, James Mckeen Cattell, a psychologist, published the first true ranking, American Men of Science, by counting the number of eminent scientists to rank institutions (Cattell, 1910). In 1925 Raymond M. Hughes, president of Miami University of Ohio and president of the Association of American Colleges, proposed the first reputational ranking by using peer review to rank universities (Hughes, 1925). Since then, ranking universities by reputation has become a major methodology, employed by Chesly Manly in 1946, Hayward Keniston 1959, Allan M. Cartter 1966, Kenneth D. Roose and Charles J. Anderson 1970, Blau and Margulies 1973, and National Academy of Science 1982 (Stuart, 1995).

Although Chesly Manly, a reporter from the Chicago Tribune, first drew public attention to university ranking in 1946, the mass media was not interested in ranking until the mid 1980s and 1990s, when publications such as the Fiske Guide to Colleges, the National Review College Guide, the Insider's Guide to Colleges, the Barron's Best Buys in College Education, and the Money Guide: Your Best College Buys Now appeared. The best-known one is the U.S. News, which publishes America's Best Colleges each year (Stuart, 1995). The history of rankings is listed in Table 1.

1870-1890	The Commission of the US Bureau of Education begins publishing an annual report of statistical data, classifying institutions.
1910	The Association of American Universities urges the US Bureau of Education to reinstate classifications.
1910-1933	James Cattell, one of America's first psychologists, professor at the University of Pennsylvania and then Columbia, publishes "American Men of Science" in which he ranks institutions on the basis of the number of eminent scientists associated with an institution either as a student or a faculty member, and factors in the ratio of scientists at a given institution to the total number of faculty.
1925	Raymond Hughes, president of Miami University and later chair of the American Council on Education and its Committee on Graduate Instruction publishes "A Study of the Graduate Schools of America" in which he uses reputational ranking of 26 disciplines in 36 institutions.
1957	Chesley Manly of the *Chicago Tribune* publishes six different rankings: ten best universities, co-educational colleges, men's colleges, women's colleges, law schools and engineering schools.
1959	Hayward Keniston of the University of Pennsylvania publishes reputational ranking of 25 universities in a range of disciplines.
1966	Allan Cartter of the American Council of Education publishes "An Assessment of Quality in Graduate Education" which ranks 106 institutions.
1973-1975	Blau and Margulies conduct reputation ranking of professional schools.
1982	The US National Academy of Science commissions an assessment of research and doctoral programmes in the United States.
1982	Rankings begin to be extended to undergraduate education (*e.g. Fiske Guide to Colleges*, 1982; *US News and World Report*, 1983; etc.).

Table 1 Chronology of Ranking Activities in the United States, 1870-1982
Adapted from (Salmi and Saroyan, 2007)

Since the 1990s, the ranking of universities has become a global phenomenon. Salmi and Saroyan (2007) summarised the distribution of league tables by country and

responsible entity (see Table 2). The majority of rankings are compiled by the media, for example, in Canada, France, the United Kingdom and the United States. In thirteen out of thirty-one countries, there are rankings produced by independent organisations and professional associations, including Australia, China, Japan, Kazakhstan, Romania, Russia, Slovakia, Ukraine, Canada, Germany, Spain, Switzerland, and United Kingdom. Some rankings are initiated by government agencies such as the Ministry of Higher Education or the University Grants Council in China, Korea, Malaysia, New Zealand, Thailand, Kazakhstan, Brazil, Tunisia, Pakistan, Nigeria, Netherlands, and United Kingdom. In Chile and India, rankings are conducted by accreditation agencies.

Region	National and international ranking system
East Asia and Pacific	Australia (B), China (B, C, IB), Hong Kong (C), Japan (B, C), Korea (A), Malaysia (A), New Zealand (A), Thailand (A)
Eastern Europe and Central Asia	Kazakhstan (A, B), Poland (C), Romania (B/C), Russia (B), Slovakia (B), Ukraine (B/C)
Latin America and the Caribbean	Argentina (D), Brazil (A), Chile (C,D)
Middle East and North Africa	Tunisia (A)
North America	Canada (B, C, B/C), United States (C, IC)
South Asia	India (C, D), Pakistan (A)
Sub-Saharan Africa	Nigeria (A)
Western Europe	Germany* (B/C, C), Italy (C), Netherlands (A), Portugal (C), Spain** (B, C, IC), Sweden (C), Switzerland (B/C), United Kingdom (A, B, IC)

A = Ranking prepared by a government agency (Ministry of Higher Education, Higher Education Commission, University Grants Council, etc.). B = Ranking prepared by an independent organisation, professional association, university or preparatory school. B/C = Ranking prepared and published through a partnership between an independent agency and a newspaper or magazine. C = Ranking prepared and published by a newspaper or magazine. D = Ranking prepared by an accreditation agency. I = International ranking (IA, IB, IC and ID linking the international dimension to the type of institution conducting the ranking).

* Austrian and Swiss universities are included in the German ranking prepared by the Centre for Higher Education Development (CHE).

** A consortium of Spanish, Portuguese and Latin American universities, Universia, computes a ranking of Iberian and Latin American universities based exclusively on publications in internationally recognised journals (http://investigacion.universia.net/).

Table 2 Ranking Systems Worldwide, 2006

Adapted from (Salmi and Saroyan, 2007)

Two International Rankings

The two influential international rankings are: the Academic Ranking of World Universities (ARWU) by the Graduate School of Education at Shanghai Jiao Tong University and the World University Rankings (WUR) by Times Higher Education-Quacquarelli Symonds. These two rankings were established for different purposes. The intention of ARWU is to assess the gap between Chinese universities and world-class universities in terms of research performance, based on internationally comparable data (Liu and Cheng, 2005). The indicators are chosen from prestige prize winners to the number of publications and citations, with the intention to rank universities by objective indicators rather than subjective judgement (see Table 3).

Criterion	Indicator	Code	Weight (percentage)
Quality of education	Alumni of an institution winning Nobel Prizes and Fields Medals	Alumni	10
Quality of faculty	Staff of an institution winning Nobel Prizes and Fields Medals	Award	20
	Highly cited researchers in 21 broad subject categories	*HiCi*	20
Research output	Articles published in *Nature* and *Science*	*N&S**	20
	Articles Indexed in SCIE and/or SSCI	*SCI*	20
Size of institution	Academic performance relative to institutional size	*Size*	10
Total			100

* For institutions specialized in the Humanities and Social Sciences, such as the London School of Economics, *N&S* is not considered, and the weighting redistributed to other indicators.

Table 3 Criteria and Weights for the 2004 ARWU
Adapted from (Liu and Cheng, 2005)

In contrast, the purpose of WUR is to 'provide a global comparison of their success against the notional mission of remaining or becoming world-class' (QS, 2009b, p.1).

Most weightings are allocated to subjective indicators, the reputational surveys (see Table 4).

Indicator	Explanation	Weighting
Academic Peer Review	Composite score drawn from peer review survey (which is divided into five subject areas). 6,354 responses in 2008.	40%
Employer Review	Score based on responses to employer survey. 2,339 responses in 2008.	10%
Faculty Student Ratio	Score based on student faculty ratio	20%
Citations per Faculty	Score based on research performance factored against the size of the research body	20%
International Faculty	Score based on proportion of international faculty	5%
International Students	Score based on proportion of international students	5%

Table 4 Indicators and weighting 2008 THE-QS
Adapted from (QS, 2009a)

Definition of Ranking

Webster (1986) defines ranking as:

> [M]ust be arranged according to some criterion or set of criteria which the compiler(s) of the list believed measured or reflected academic quality [; and] it must be a list of the best colleges, universities, or departments in a field of study, in numerical order according to their supposed quality, with each school or department having its own individual rank, not just lumped together with other schools into a handful of quality classes, groups, or levels.
>
> (Webster, 1986, p.5)

Stuart (1995, p.13) notes:

> … to issue a ranking, as opposed to purely descriptive information, publications must assign a position to each unit on the basis of the chosen criterion or criteria.

Morrison, Magennis and Carey (1995, p.128) share a similar view: 'League tables are weighted combinations of scores on performance indicators'. Agreeing with Morrison, Magennis and Carey's definition, Hendel and Stolz (2008, p.175) add:

> Additionally, however, ranking also includes the act of providing the outcomes of this process as entities of information to a wide variety of stakeholders; it is those stakeholders that then engage in acts of interpreting and evaluating the ranked data.

One common theme emerging from these definitions is that ranking is based on a set of criteria or indicators. Ideally, the choice of indicators should reflect the quality of education (Webster, 1986). However, it could be made to address the interests of stakeholders (Hendel and Stolz, 2008), which brings the value or bias of stakeholders into the ranking. Furthermore, the choice of indicators is also restricted by the availability of data. This problem becomes even more pronounced when compiling international rankings (Liu and Cheng, 2005).

In fact, when the indicators are chosen carefully to reflect the quality of education, the potential disagreement about indicators is intensified, since it is difficult to define "quality". According to Green (1994) there are three perspectives towards quality: the traditional concept of quality, conformance to specifications or standards, and quality as fitness for purpose. The traditional concept of quality is closely associated with the traditional concept of a university, an elite institution and a centre of professional training and pure research. The notion of quality as conformance to specification or standards originates from quality control in business. Products and service are measured to meet required specifications and standards, and rejected if they fail to conform. Specifications are the quantitative data that can be calculated or substantive evidence that can be seen, while standards are statements that describe

desirable practice and outcome. Quality as ‘fitness for purpose’ is justified by the argument that ‘quality has no meaning except in relation to the purpose of the product or service’ (Green, 1994, p.15). But what is traditional excellence, what are the required specifications and standards and what should be regarded as fitness for purpose are questions hard to reach consensus, depending on different perspectives and values.

Another common theme from these definitions is that ranking should assign a score or position to each individual or organisation. However, recent developments mean this is not always true. The strength of the university ranking compiled by the Centre for Higher Education Development in Germany is to group universities in clusters, rather than assigning a particular score or position to each university when they are not significantly different (van der Wende, 2008).

Van Dyke (2005) calls rankings as university report cards. Indeed, rankings should provide concise information on universities. The information could be presented in the form of scores, positions, levels or even descriptions. As Hendel and Stolz (2008) point out, ranking should help readers engage with the information and interpret and evaluate it. Furthermore, rankings should enable readers to make comparison between universities. Ranking is based on concise information, which enables readers to make comparisons between universities.

The Concept Underpinning Ranking

Wherever there are rankings, there are criticisms. Rankings do not stop with these criticisms and, in contrast, they are blooming. The demands for rankings come from those who finance higher education, such as the public, governments and funding authorities requiring transparency for decision-making. Moreover, universities themselves need rankings too. They try to legitimise their positions by rankings for reasons of prestige, student and staff recruitment, and funding (Altbach, 2006).

But rankings are more than these. Salmi and Saroyan (2007, p.2) believe ranking is a form of accreditation:

> Accreditation, cyclical reviews, external evaluation by peers, inspection, audits, performance contracts based on predetermined indicators, benchmarking and research assessments are among the most common forms of accountability. Some are initiated by the institutions themselves; some are imposed on tertiary education institutions externally by funding bodies, quality assurance agencies, committees of presidents and vice chancellors, as well as stakeholders at large. One example of the latter is institutional rankings by league tables.

Vaughn (2002) regards ranking as one aspect of higher education assessment. He considers that ranking is created as an assessment procedure when the public demands more information and transparency in higher education, which the dichotomous yes or no answer from accreditation fails to address.

Federkeil (2008) views ranking as an external assessment, offering transparency and contributing to quality assurance at a system level. He positions ranking in the bigger picture of quality assurance (see Figure 1):

> Analytically, the existing instruments of quality assurance can be ordered along two lines: One is the level of reference (institution vs. system); the other refers to the main aims of the instruments and could draw on the distinction between enhancement and accountability.
>
> (Federkeil, 2008, p.220)

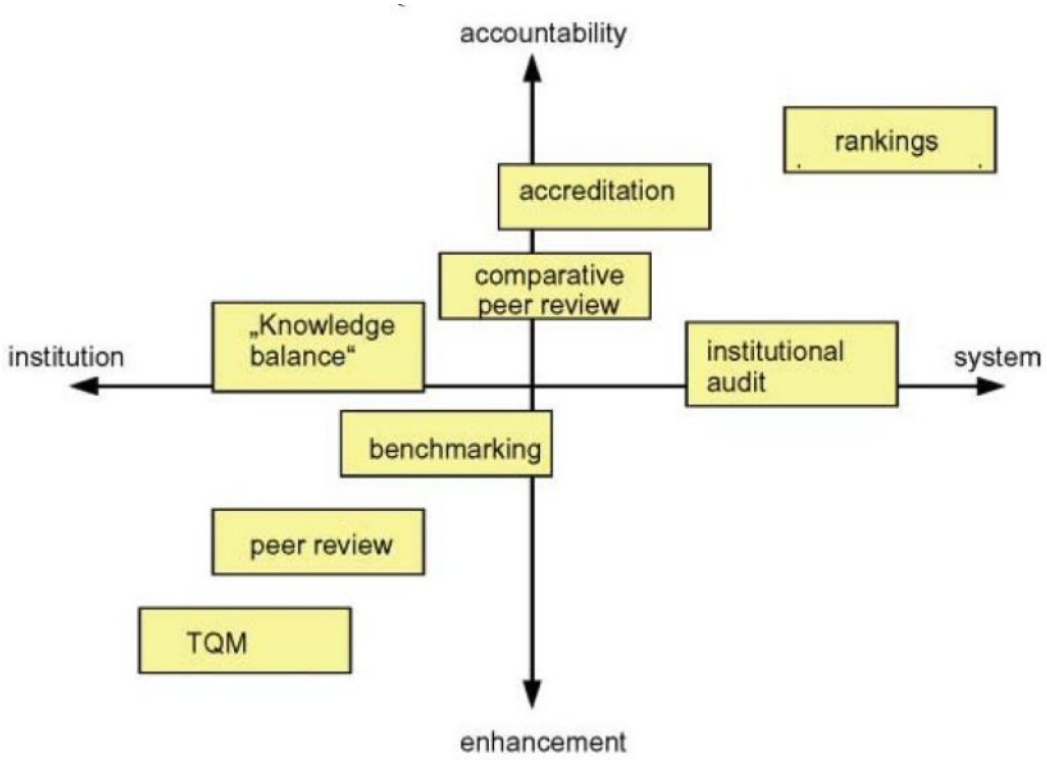


Figure 1 Instruments of quality assurance in higher education
Adapted from (Federkeil, 2008)

Federkeil (2008) also elaborates on the differences between evaluations and rankings.

> Whereas evaluations normally include only a small number of institutions, rankings strive to encompass a whole system – however, this is defined (e.g. a national system of higher education or 500 world-class universities). While rankings are aiming towards a broad inclusion of institutions, evaluations normally are focused more on in-depth analysis. The most important difference between rankings and evaluations is that rankings as a system of quantitative indicators may completely neglect causes and reasons for bad results of institutions, while the identification of causes of weaknesses is the very task of evaluation. Yet rankings offer a variety of data on individual institutions that can be used for internal analysis in terms of comparison or some form of benchmarking. While ranking data normally cannot deliver causal analysis, they can be a starting point for a more detailed analysis of the strengths and weaknesses of an individual institution.
>
> (Federkeil, 2008, p.222)

Harvey (2008) summarises five purposes of rankings based on the Berlin Principles on Ranking of Higher Education Institutions, a set of principles of quality and good practice in HEIs rankings established by the United Nations Educational, Scientific and Cultural Organization European Centre for Higher Education and the Institute for Higher Education Policy (CHE, CEPES and IHEP, 2006).

> First, as implied above, rankings respond to demands from consumers for easily interpretable information on the standing of higher education institutions, satisfying a need for consumer guidance among the student population 'they help students and their parents make their decisions concerning a fitting place for college education' (RFSU, 2008). This is seen as the primary purpose by many of the creators and publishers of rankings.
> Second, rankings stimulate competition among institutions.
> Third, rankings help differentiate among different types of institutions and different programmes and disciplines.
> Fourth, rankings can also serve as part of a framework for national assessment, accountability and quality assurance in the higher education system and arguably can provide some of the rationale for allocation of funds.
> Fifth, linked to a national framework for quality, rankings, it is claimed, serve to generate a debate that contributes to the definition of 'quality' of higher education, 'complementing the rigorous work conducted in the context of quality assessment and review performed by public and independent accrediting agencies'
>
> (Harvey, 2008, p.189)

Thus, the in-depth rationale for conducting ranking is that it serves as one of the instruments for quality assurance in higher education. Unlike Total Quality Management and Peer review, ranking is initiated by external bodies. In this sense, ranking, accreditation, and audit share more similarities. However, accreditation and audit are often conducted by government bodies and agencies, asking for accountability. Ranking provides transparency for the public. To achieve transparency, results and data should be accessible to enable readers to compare universities and

make their own judgement. Ranking can not only contribute to quality assurance at system level but also serve as a diagnostic signpost for internal quality assurance. In addition, unlike accreditation and audit, rankings can be initiated without governmental authorisation. In theory, anyone can compile a ranking, thus promoting academic freedom.

Criticism of Rankings

Criticism of rankings tends to focus on the validity and reliability of ranking, particularly on ranking techniques and the choice of indicators. Problems related to ranking techniques are arbitrary weighting (van der Wende, 2008;Williams, 2005;Locke et al., 2008)，lack of transparency (Locke et al., 2008;Locke, 2008), and inappropriate use of bibliometrics (Van Raan, 2005). Problems associated with the choice of indicators suggest that rankings do not always reflect the quality of education. The positions in rankings are biased towards reputation (van der Wende, 2008;Williams, 2005;Locke et al., 2008;Hazelkorn, 2008a), research (van der Wende, 2008;Locke, 2008;Lindblad, 2008), natural and medical science (van der Wende, 2008;Lindblad, 2008), publication (Locke et al., 2008;Lindblad, 2008), and gross performance (Williams, 2005).

Research Questions

The arguments listed above question the validity and reliability of rankings from a theoretical perspective. However, there are little empirical evidences collected to judge whether the rankings reflect the actual situation within universities and to

understand whether shifts in ranking position are due to actual changes within universities or to flaws in the ranking methodology.

To explore these questions, this research starts from the position of the two international rankings, the ARWU and the WUR. They demonstrate somewhat different characteristics. The indicators employed in the ARWU are all objective indicators. The ARWU's ranking methodology is claimed to have remained the unchanged after 2005. On the other hand, the WUR places a 50% weightings on subjective indicators and remaining 50% weightings on objective indicators. Referring to ranking methodology, the compiler of the WUR admits that the ranking methodology is adjusted each year. These observations lead to the research questions:

- How are the ARWU and the WUR compiled?
- How do universities' positions change over time in the ARWU and the WUR?
- What are the factors influencing the ranking?
- How do the factors influence the rankings?

Research Design

Research Strategy

The nature of the research questions demands a combination of statistic analysis and the in-depth case study approach. The statistic analysis can draw an overall picture of the ARWU and the WUR. Multiple regression can test whether the rankings are conducted based on the claimed weightings and how much each indicator contributes to the final results. Principal component analysis can uncover factors underpinning the indicators.

The multiple case studies can check the validity and the reliability of the two rankings by testing whether the values in rankings match the real situation within the universities. The case study is a well-established research strategy with the basic concept that the case is studied in detail to 'understand the case in depth, and in its natural setting, recognizing its complexity and its context' (Punch, 2005, p.144). Multiple case studies are helpful in testing whether the hypothesis fits the circumstances to improve theory building (Yin, 1984;Eisenhardt, 1989). In this dissertation, multiple cases, namely universities, are selected for comparison to find out whether the shift of position in the ARWU or the WUR reflects actual changes within universities.

Data collection

A major strength of case study is that it provides a good opportunity to gather data from multiple sources. The need for diverse sources is much greater than other with research approaches. Thus, the findings from a case study tend to be more convincing and accurate (Yin, 1994). Data for this research project are mainly obtained from the public domain, including the Graduate School of Education at Shanghai Jiao Tong University, Times Higher Education Supplement, Quacquarelli Symonds, U.S. News, Higher Education Evaluation & Accreditation Council of Taiwan, individual campuses of University of California, University of California central office, U.S. National Center for Education Statistics, Nobel Prize and Fields Prize committees.

Data Analysis

The general approach to analyse the data is secondary analysis.

> Secondary analysis is the analysis of data by researchers who will probably not have been involved in the collection of those data for purposes that in all likelihood were not envisaged by those responsible for the data collection. Secondary analysis may entail the analysis of either quantitative (Dale, Arber and Procter, 1988) data or qualitative data (Corti, 1993) …
>
> (Bryman, 2004, p.201)

In secondary analysis, it is important to note there are limitations inherent in the method:

- Lack of familiarity with data
- Complexity of the data
- No control over data quality
- Absence of key variables

(Bryman, 2004, p.205-206)

However, for this research the advantages of secondary analysis outweigh the limitations, namely:

- Cost and time.
- High quality data
- Opportunity for longitudinal analysis
- Subgroup analysis
- Opportunity for cross-cultural analysis
- More time for data analysis.
- Reanalysis may offer new interpretations.
- The wider obligations for the social researcher

(Bryman, 2004, p.202-205)

Importantly, secondary analysis can maximize the useful sources of data and help form a full picture of the situation.

Overview of Study

Chapter Two: Literature Review summarises the researches related to rankings.

Chapter Three: Methodology addresses the methodologies employed for data

collection and data analysis. Chapter Four: First Stage Findings and Analysis presents the statistical findings and analysis. Chapter Five: Multiple Case Studies takes an in-depth approach to explore factors influencing the ranking. Chapter Six: Discussion discusses the limitations of the ARWU and the WUR and possible future development. Chapter Seven: Conclusions summarise the whole research project.

Chapter Two: Literature Review

To achieve a wide coverage of the literature search, OxLIP+ CrossSearch is first used to cross search between the Education Resources Information Center (ERIC), the Scopus, and Web of Science with key words: university, league table and ranking. Then, an advanced search is conducted within ERIC with key words: ranking(s), league table(s), university(ies), higher education, quality, evaluate(s)(ing)(ed)(tion).

Researches related to rankings mainly focus on six aspects: background of rankings, theoretical arguments, current rankings in different countries, indicators, and the impact of rankings. The background of rankings is addressed in the Appendix.

Theoretical Arguments

The World Class University

The ARWU and the WUR try to identify world class universities by rank. But neither provides a definition of a world class university. Birnbaum (2007, p.1) suggest that world class universities should 'help define our civilization, serve as bastions of original thought, and respond to the quintessential human need to know'. Similarly, Salmi (2009, p.19-20) proposes a definition of the world class university:

a) a high concentration of talent (faculty and students);
b) abundant resources to offer a rich learning environment and conduct advanced research;
c) favorable governance features that encourage strategic vision, innovation and flexibility, and enable institutions to make decisions and manage resources without being encumbered by bureaucracy.

These abstract definitions leave ample room to debate appropriate indicators.

Principles of Selecting Indicators

Harvey (2008, p.190) summarises the points made by Lazarsfeld et al. (1972) and put forward an idealised positivistic process to measure complex concepts:

> First, a clear statement and definition of what the concept is that is being measured is developed.
> Second, the different dimensions of the theoretical concept are identified.
> Third, sets of indicators that represent aspects of each dimension are specified.
> Fourth, an appropriate indicator or suite of indicators from the range of possible indicators is selected.
> Fifth, each indicator within each dimension is weighted on the basis of a theoretically-sound understanding of the importance of the indicator as a measure of the dimension.
> Sixth, each dimension of the concept is weighted and the dimensions are combined into a single index.

When choosing indicators, Linn (1993) suggests checking the validity, the reliability, and the comparability by asking whether the indicator measures what it purports to measure; does it do so in a consistent/error-free fashion; and can it be interpreted in a similar way across different kinds of programmes or institutions.

Dilemmas

Besides arguments on indicators, there are certain dilemmas when compiling rankings. These include university-based versus discipline-based; rankings versus clusters; value-added or not, which are all hotly debated.

University-based versus Discipline-based

Van Dyke (2005, p.106) notes that ‘on the one hand, students may have to apply to a discipline rather than to a university, and quality within a university may vary by discipline so averaging across departments can produce a distorted view of the university’s quality. On the other hand, variability across disciplines is diminishing, particularly among the best universities; students are not always certain in which discipline they are most interested when they investigate universities; and there is arguably an important ‘brand effect’ for a university as a whole’.

Rankings versus Clusters

Clarke (2002) and Merisotis (2002) point out one problem of using ranking is that the differences between universities are often statistically insignificant. Thus, universities should be grouped into clusters—high, medium, and low—instead. Van Dyke (2005, p.106-107) argues that ‘this solution is not without its own methodological difficulties, however, as choice of cut-off points between clusters may be rather arbitrary, and the same issue of true differences remains: the apparent distinction between the lowest-rated university and discipline in the ‘top’ group, for example, and that of the highest-rated one in the next group may not be real. Moreover, clusters are simply not as satisfying as rankings’.

Value-added or Not

Marginson (2007) argues that a value-added measure should be added in rankings. Coates (2009) shares this view and puts forward two approaches to measure the value-added:

> The first approach computes value-added estimates by comparing predicted against actual performance using data from entrance tests and routine course assessments. In the second approach, comparisons are made between outcomes from objective assessments administered to cohorts in the first and later years of study.
>
> (Coates, 2009, p.1)

However, Morrison, Magennis, and Carey (1995) put forward potential problems in using the value-added approach. Test-retest scores (A level vs. degree scores) are not measured by the same test metric; most disciplines do not use objective measures but base assessment on professional judgment; criterion-referenced testing of the student's progress is compared with established assessment criteria; norm-referencing requires the student's achievement to be compared with those of other students.

Furthermore, there is a paradox related to the concept of value-added. Morrison, Magennis, and Carey (1995) argue that value-added is biased against high achieving students, because they have less room for improvement due to the small margin below the ceiling. 'Gain scores are negatively correlated with initial scores and, under certain conditions, can have zero reliability' (Morrison, Magennis and Carey, 1995, p.132). However, Matthew Effects (Walberg and Tsai, 1983) in education

illustrates that value-added can be biased towards the high achieving students as students with higher ability at their entrance improve academically at a higher rate.

There is no easy answer to these dilemmas due mainly to the lack of empirical research evidence on the ranking of universities. Furthermore, it also depends on the purpose of the ranking—different purposes demand different strategies.

Indicators

Different countries have different national rankings. The U.S. News compiles two parallel domestic league tables for undergraduate and postgraduate education. The Canadian counterpart is MacLeans. The longest running league table for universities in UK is the Times Good University Guide. A relatively new ranking in Australia is the Melbourne Institute Index. The Center for Higher Education Development in Germany uses an innovative approach to compile both domestic and European ranking. A detailed review of indicators used in these rankings is in the Appendix.

Classifications of Indicators

Denotation Approach

Each ranking has its own indicators. There are huge varieties of indicators and there are considerable overlaps between different rankings. Clarke (2002) suggests three dimensions of indicators. Building on Webster's (1986) categories, she organises indicators into three categories: student achievements, faculty accomplishments, and institutional academic resources (Clarke, 2002). (For detail on Webster's work, see

Appendix.). We can call this the denotation approach as the titles of categories summarise the meanings of indicators.

Van Dyke (2005) gave a comprehensive summary of 10 rankings, in which he grouped 66 different indicators into six categories, leaving 6 indicators without grouping. The six categories are:

- Quality of Academic Staff/Faculty
- Quality of Incoming Undergraduate Students
- Quality of Undergraduate Programme
- Quality of Graduate Programmes
- Resources
- Stakeholder Opinions

(For detail on Van Dyke's classification see Appendix.)

Input-Process-Output Approach

The second dimension of indicator classification is in terms of inputs, processes, and outputs. Echoing the indicators listed above, inputs can be the scores of incoming students on standardised tests and library resources, while outputs are faculty citations and awards or effectiveness of graduates in the workplace. A frequently mentioned problem with indicators is the lack of process measures such as teaching quality, and more generally because of their being difficult to identify and more difficult and costly to measure (Clarke, 2002).

Dill and Soo (2005, p.500-502) reviewed five rankings in four counties and categorised indicators in the following approach:

Input

- Faculty

- Students
- Financial resources and facilities

Process

- Teaching

Output

- Satisfaction
- Graduation
- "Value-added"
- Learning progress
- Employment

Reputation

(For detail on Dill and Soo's classification please see Appendix.)

Usher and Savino (2006, p.15) committed to a comprehensive review, covering 19 league tables in 10 counties, including three international ranking, and drew a different framework:

Research
Beginning characteristics — Learning inputs: staff / Learning inputs: resources — Learning outputs —Final outcomes
Reputation

(For detail on Usher and Savino's classification see Appendix.)

Hendel and Stolz (2008, p.180) followed the same route and classified 150 indicators from 23 rankings in 11 countries in Europe:

Input

a) Beginning characteristics students
b) Faculty and staff
c) Financial resources
d) Program design
e) Amenities

Thoughput

a) Program/course proceedings
b) Organization and management
c) Student support systems
d) Teaching quality

Output

a) Career prospects
b) Reputation
c) Outcome measurements
d) Value added

Research

Objective-subjective Approach

Indicators can also be grouped by their objective or subjective nature. 'Objective indicators are those not dependent on the person doing the counting…Subjective indicators are those that can vary depending on who is responding' (Clarke, 2002, p.446). Taylor and Braddock (2007) argue subjectivity is introduced by deciding weightings, peer review, employer survey, and course satisfaction surveys from students and graduates. Objective indicators include publication, citation, funding, staff student ratio, class size and annual university spending per student. The benefits of using objective indicators are:

> a) It is focused on an aspect of university performance—namely, research— which can be rationally justified as an indication of university excellence;
> b) It allows a university's 'score' to be automatically generated by a simple counting process that minimises human error by avoiding any reliance on inference or judgement; and
> c) It allows for a comprehensive survey of all universities, resulting in an exact performance ranking for each university using the identical data source in each case.
>
> (Taylor and Braddock, 2007, p.248)

A New Classification

Webster's categories are concise but they omit new indicators lately developed. The problem of categorising indicators by input-process-output is that some indicators can be regarded as both input and output, for example, research grants. This causes unavoidable ambiguity. The objective-subjective approach is too general and actually

just draws a line between quantitative and qualitative data. Van Dyke and Usher and Savino's categories cover almost all kinds of indicators used in rankings. However, their titles of the categories do not fit the indicators neatly, with some overlaps or some unclassified aspects. I attempt to make a new classification of indicators to include all indicators reviewed by the researchers mentioned and group them under the proper titles that articulate the denotation of indicators. The new classification covers aspects of students, staff, and programmes/university.

Students

- **Student Pre-achievements**: students' incoming grades
- **Student Learning Process**
- **Student Post-achievements**: academic and career achievements

Staff

- **Staff Accomplishment**: projects, grants, patents, chair in research centres, and conference presentations
- **Staff Productivity**: publication and citation
- **Staff Reputation**: awards and membership in prestige association
- **Staff Teaching**

Programme/University

- **Admission**: entry standard
- **Teaching and Learning**: number or ratio of faculty; number, size or time of class; classes taught by tenured staff; ratio of postgraduates
- **Student Support**: expenditure on library, IT facilities, space available, service, and findings
- **Staff Support**: salary, spending on laboratory, and assistants
- **Revenue**: public and private sources; student fees; giving
- **Environment**: culture and tradition created by the university and the surrounding of the city in which the university nests
- **Institution Reputation**

For detail of this new classification along with Webster, Dill and Soo, Hendel and Stolz, Usher and Savino, and Van Dyke's classifications, see Appendix.

Arguments about Indicators

Institution Related Indicators

Reputation is one of the most important proxies of quality. Brooks (2005, p.6-7) argues that reputation has a high level of credibility.

> When multivariate regression models are used to predict the 1995 reputational scores, measures of faculty size, faculty research productivity, and doctoral program success are all significantly related to reputation for the majority of fields (Ehrenberg and Hurst, 1996;Ehrenberg and Hurst, 1998). Thus, reputational scores demonstrate a high level of overall consistency with other program measures. In addition, the 1995 NRC and U.S. News & World Report rankings are highly correlated, lending further credibility to the data (Rogers and Rogers, 1997).

Yet, the critics against reputational survey argue that it 'grow(s) out of the lack of a common frame of reference for the raters, which varies because individuals provide quality judgments based on what they personally know of an institution or program and how it relates to all others they know' (Brooks, 2005, p.7). Furthermore, even when raters are comparably informed, Jones, Lindzey and Coggeshall (1982) found no substantial differences in comparisons of ratings from raters who received such information with those who did not.

Other critiques are related to the Halo Effect. Programmes or departments may receive higher rating in prestige institutions (Cartter, 1966;Diamond and Graham, 2000). The extreme example of the Halo Effect was that Princeton received a high rating for a law school, which, in fact, it did not have (Webster, 1981). Moreover,

‘reputational surveys generate circular reputational effects that tend to reproduce the pre-given hierarchy’ (Marginson, 2007, p.140).

Enrolment, degrees awarded, faculty resources, student selectivity, financial resources, and alumni giving are biased towards large institution (Brooks, 2005;Graham and Diamond, 1997). Analysing the MacLeans’ ranking, Cramer and Page (2007) revealed that the final rank standings of 2006 were correlated with operating budget; in the 2005 ranking, the operating budget was correlated with the reputational index and library holdings showed the highest correlation with final rank standings.

Looking beyond teaching and research, Montesinos et al.(2008) draw the attention to the interaction between university and society, in other words, the third missions: Social Third Mission, Enterprising Third Mission, and Innovative Third Mission. (For detail on the third missions indicators see Appendix)

Staff Related Indicators

Brooks (2005) believes that the indicator—research grants is biased towards sciences and only shows that the research proposal is distinctive, not necessarily pointing to the quality of the research product.

Johnston (2009) suggests that, instead of citation, impact score of journals should be used to measure the real impact of research. But Brooks (2005) argues that the

impact factor counts negative citations and self-citations, distracting from the quality of research. In addition, the mobility of staff makes it difficult to trace the contribution of each institution.

Noting that Humanities, Arts, and Social Sciences are disadvantaged in Thomson Institute for Scientific Information (ISI), Cunningham (2008) advocates the build of a citation database for these disciplines, such as the European Reference Index for the Humanities in Europe and the Research Quality Framework in Australia to enable these disciplines to compete fairly with sciences and medicine in rankings.

Indicators of awards and honours with a narrow frame of reference are limited to the distinguishing top-tier institutions, making it not applicable to other institutions (Brooks, 2005).

Marginson (2007) is concerned with the lack of indicators in teaching and learning.

Student Related Indicators

Generally speaking, past performance is one of the best predictor of future academic performance (Himmel, 1967). But Ting (2003) found that for students of colour, non-cognitive indicators were better predictors of academic success.

Elbeck (2009) draws attention to indicators on students' learning experiences and outcomes at business schools, by suggesting interviews, stimulation, case

competition, internship, reports, and presentations to measure students' verbal ability, problem solving skill, professional knowledge, research ability, and adaptability.

Pike (2004), Cramer and Page (2007) found that student satisfaction seldom correlates highly with overall rank standings in U.S.News or MacLeans'.

Brooks (Brooks, 2005, p.14) points problems associated with student experience surveys. 'Students may not have a broad perspective that includes other educational institutions or programs with which to compare their own. In addition, cross-institutional comparisons of these data must be made with attention to any unique or contextual factors that may shape student expectations'.

Large scale projects are underway to assess students' learning results, including the Valid Assessment of Learning in Undergraduate Education and the Assessment of Higher Education Learning Outcomes by Organisation for Economic Co-operation and Development (OECD). (For detail of the former project, see Appendix)

Finally, Ehrenberg (2005) warns that universities can manipulate improvement to the indicators for rankings without actually improving the quality of education.

Since 'Quality is both a subjective and nebulous term' (Hattendorf, 1996, p.1), it is almost impossible to observe quality directly. The debate on what the appropriate proxies of quality are will continue. To choose suitable quality proxies of world class

universities demands more international cooperative research projects, such as the Assessment of Higher Education Learning Outcomes, to discover the empirical evidence for theoretical arguments.

Impact of Ranking

Researches on UK rankings indicated there was no influence of rankings on students' choice (Wagner, 1998;Eccles, 2002;Gunn and Hill, 2008). In the USA, contradictory results were found. In the University of California, Los Angeles Freshman Survey 2007, rankings in national magazines were placed tenth in the twenty-one criteria on which students based their college choice. Gunn and Hill (2008, p.250) examined the data from 1995 to 2004 and found that 'the influence of the U.S. News rank on school choice is important and the effect is different for lower ranked schools'.

Jin and Whalley (2007) argue that rankings have positive influence which can attract public attention. They noted that being included in the U.S. News rankings made public colleges increase educational and general expenditure by 3.2 percent per student.

Supported by the OECD, Hazelkorn (2007) conducted the first global survey on the impact of rankings on institutions from 41 different HEIs in 41 different countries. She found rankings have considerable impact on HEIs. HEIs believed that rankings drew wide attention to them. 'Ninety percent of respondents said league tables and

ranking systems (LTRS) are published widely, and cited the media as the principal developer, followed by government departments, accreditation and higher education agencies, and independent research organisations' (Hazelkorn, 2007, p.9). Thus, they used rankings for publicity. 'Almost 50% use their institutional position for publicity purposes: press releases, official presentations and their website' (Hazelkorn, 2007, p.11). HEIs also believed that their ranking could influence cooperation with other institutions. '57% of respondents said they believed LTRS were influencing the willingness of other HEIs to partner with them' (Hazelkorn, 2007, p.14). Due to the importance of rank standings, HEIs wanted to improve their rankings.

> 92.8% and 82%, respectively, want to improve their national or international ranking….
> Comparing current with preferred rank, 70% of all respondents wish to be in top 10% nationally, and 71% want to be top 25% internationally.
>
> (Hazelkorn, 2007, p.8)

Chapter Three: Methodology

Research Questions

To research factors influencing rankings there are four basic questions and several sub-questions that need to be answered:

How the ARWU and the WUR are compiled?

- What are the indicators in the ARWU and the WUR?
- What are the ranking methods of the ARWU and the WUR?
- What are the changes in methodology over the years?

How do universities' positions change over time in the ARWU and the WUR?

- What are the relationships between indicators and ranking?
- What is the trend of the ARWU?
- What is the trend of the WUR?

What are the factors?

- What are the relationships between the indicators?

How do the factors influence the rankings?

- What are the changes of an individual university over the years?
- What are the relative changes of one university compared with others?

Research Design

A detailed review on the definition of indicators, the methodology of compiling ranking, and changes of methodology is conducted first. To investigate changes in

universities ranking standings, it is necessary to research the sensitivity (Chatterjee and Hadi, 1988). Sensitivity is divided into two dimensions: the relationship between the total score and the rank standing; the relationship between the indicators. By exploring the first dimension of sensitivity, we can understand the impact of fluctuations of total score on rank standing; by exploring the second dimension of sensitivity, we can understand the impact of changes of different indicators on rank standing. The relationship between the total score and the rank standing is revealed by scatterplotting, while the relationship between the indicators is analysed by correlation, multiple regression, and principal component analysis. Finally, multiple cases are selected to check whether the results of ranking correspond with the real situation within the institutions where the raw data are available. When the raw data are not available, the variables will be triangulated with similar variables from other rankings to check if there are discrepancies or not. Universities will be analysed in both an absolute and a relative manner. Absolute changes are the changes of an individual university compared with it's self over the years. Relative changes are changes of relative position compared with other universities over the years.

Data Collection

The results of rankings and value indicators are obtained from the compliers' official websites. Data are available from the Internet, but are not in an analysable format. A considerable amount of time and care is invested to correctly transfer data from

websites onto spreadsheets ready for analysis. Altogether 20,384 values are keyed in and organised by university.

For the multiple case studies, original data to verify value of indicators of ranking are available for Nobel Prize Laureates, Fields Prize winners, Student to Staff Ratio, International Staff and Students number. Data of Nobel Prize Laureates and Fields Prize winner are obtained from their official website. Students and staff related data are gathered from the affiliated individual universities, university central office and also the National Center for Education Statistics to triangulate the data.

Original data of Highly Cited Researchers, Articles Published in Nature and Science, Publication, Citation, Peer Review and Employer Review for individual university are not easily obtainable due to access restrictions and the limited scope of time and human resource for this dissertation. But some indicators have counterparts in other rankings. The Peer Review has its counterpart in America's Best Colleges by U.S. News. Highly Cited Researchers and Citation have their counterparts in Performance Ranking of Scientific Paper for World Universities by Education Evaluation & Accreditation Council of Taiwan. So we can compare same indicators from different compilers.

Data Analysis

Data Inclusion

The presentation of the ARWU and the WUR includes the rank standings of universities, the scores of each category of indicators, and the total scores of universities.

Both the ARWU and the WUR rank around 500 universities. But the ARWU only provides total scores for the top 100 and the WUR for the top 200, which confines the statistical analysis within this limit. Due to shifts of universities' standings each year, some universities shift out from top 100 in the ARWU or top 200 in the WUR in certain years, leaving these universities with no total scores in these years. To eliminate errors caused by missing data, the universities with incomplete data are excluded from the statistical analysis. Thus the data for this analysis comprise 86 universities in the ARWU from 2003 to 2008 and 128 universities in the WUR from 2004 to 2008.

The Method of Standardisation

To check the sensitivity on the total score and rank standing, in other words, the impact of fluctuations of total score on ranking, the total scores across years should be comparable. To ensure comparability, a standardisation is conducted. The method to for standardisation is to 'equate the forms—to convert the system of units of one

form to the system of units of the other—so that scores derived from the two forms after conversion will be directly equivalent' (Angoff, 1971, p.562). Before such equating, two conditions have to be fulfilled.

> The first is that the two instruments in question be measures of the same characteristic…
> The second restriction implied by the notion of equating is that, in order to be truly a transformation of only systems of units, the conversion must be unique…
>
> (Angoff, 1971, p.563)

The definition of equivalent scores is:

> Two scores, one on Form X and the other on Form Y (where X and Y measure the same function with the same degree of reliability), may be considered equivalent if their corresponding percentile ranks in any given group are equal.
>
> (Angoff, 1971, p.563)

When Form X and Form Y are different, equivalent scores yielded by equating will stretch and compress the scale of Form X so its distribution will coincide with the distribution of Form Y (Angoff, 1971).

The equation is (Angoff, 1971, p.564):

$$\frac{Y - M_y}{S_y} = \frac{X - M_x}{S_x}$$

Total scores in the ARWU and the WUR can be regarded on a percentile rank scale. 'Distributions of percentile ranks for the groups on which the ranks are based are necessarily rectangular. The percentile rank scale itself is clearly ordinal and, according to most points of view, its units are unequal since they are intended to

provide equal proportions of a group, not equal intervals on a scale…' (Angoff, 1971, p.515).

Statistical Analysis

To explore the second dimension of sensitivity, in other words, the relationship between indicators, several statistical analyses are undertaken.

Correlation

Correlation coefficients matrices are established to measure the extent to which indicators are linearly related. There are two frequently used correlation coefficients: Pearson's and Spearman's. Pearson's correlation coefficient is a parametric statistic, which requires that data are interval and normally distributed. Spearman's correlation coefficient is a non-parametric statistic and can be used when the data have violated parametric assumptions (Field, 2005). As the data here are not always normally distributed, the Spearman's correlation coefficient is employed.

Multiple Regression

To reveal the consequences of changes in different indicators to rank standing, it is necessary to check how much each indicator's weight actually, comparing the alleged weightings. Multiple regression can measure how much each indicator accounts for the variance in the total score. The standardised coefficient Beta (b) reveals the real weighing of each indicator. The R square change shows the amount of variance

explained by the variable. Hierarchical entry is used as the method of variable entry. The variables are added based on the substantive theoretical importance.

Principle Component Analysis

To unearth the factors underneath the indicators, two statistical methods are frequently used, namely factor analysis and principal component analysis. The major differences between the two are in the communality estimates that are used. 'Factor analysis derives a mathematical model from which factors are estimated, whereas principal component analysis merely decomposes the original data into a set of linear variates. As such, only factor analysis can estimate the underlying factors and it relies on various assumptions for these estimates to be accurate. Principal component analysis is concerned only with establishing which linear components exist within the data and how a particular variable might contribute to that component' (Field, 2005, p.630-631). In this case, principal component analysis can fulfil the task of revealing what the linear components underneath and the indicators contribute to the component.

Case Study

To investigate how the factors influence the rank standing, we need to take a close look at universities to compare shifts in ranking and changes on the ground. The multiple cases study can serve this purpose. As mentioned in the research strategy, the case can be a critical case, a unique case, a revelatory case (Yin, 1984), or an

exemplifying case (Bryman, 2004). In this study, multiple cases are chosen based on these criteria:

- **Critical**: case can be chosen from those who demonstrate significant change.
- **Exemplifying**: case can be chosen from the majority of universities, who demonstrate random changes. Compared with the critical case, this can help find whether there are any substantial differences within universities between those who demonstrate significant changes and those who do not; and the reasons.
- **Revelatory**: cases can be chosen from different columns to explore the different magnitudes in changes.
- **Unique**: the case should be worthwhile to research.

To select cases, it is important to 'distinguish here between annual changes in the rankings due to short-term fluctuations and significant changes over a longer period of time. While schools often do switch places on an annual basis, moving up or down a notch or two, most of these annual changes are, in fact, "blips" that tend to cancel themselves out over time. Errors in reporting data are a prime example of this, since they tend to be found and corrected the following year. Such change is not considered to be significant, since an institution's ranking usually reverts back to its original location. Given the stability of national universities, significant change, either upward or downward, usually takes multiple years to manifest itself' (Daly, Machung and Roque, 2007, p.8)

To check if changes are random or significant changes, a statistical method, the runs test, can be applied. The runs test is a non-parametric test that checks a randomness hypothesis (Sheskin, 2004). In this case, if the changes are random, the results of the runs test will be significant. One potential problem of using a runs test in this circumstance is that the runs test may not be sensitive to such a short time span. Consequently, any change patterns within five or six years may be identified as random changes by a runs test. If a runs test fails to identity any significant changes, we need to wait a further period until there are enough data being published to run any statistical tests. In the current situation, we can only resort to checking the direction of moves over the years by a simple definition: if one university moves continuously up or down, the changes are considered significant changes.

Besides the direction of changes, the magnitude of changes is of interest. Even when the directions of changes are the same, the magnitude of changes can be different. For some universities, the small changes in total score cause only small changes in ranking standings; for others, small changes in total score can cause the large shifts in ranking standings, and vice versa. By choosing universities from different columns of magnitude, we can learn what the actual gaps between these universities are in reality and understand whether the gaps between these universities are illusion created by rankings or differences in the universities capacity.

Chapter Four: First Stage Findings and Analysis

The Ranking Methodologies of the ARWU and the WUR

Indicators

The ARWU's Indicators

The indicators of the ARWU are Alumni Prize Winners (10% of weightings), Staff Prize Winners (20% of weightings), Highly Cited Researchers (20% of weightings), Articles Published in Nature and Science (20% of weightings), Publication (20% of weightings), and Size of Institution (10% of weightings). A detailed summary of each indicator can be found in the Appendix.

Based on the new classification of indicators outlined in Chapter Two, the ARWU includes Student Post-achievements (Alumni Prize Winners), Staff Reputation (Staff Prize Winners), and Staff Productivity (Highly Cited Researchers and Articles Published in Nature and Science). Alumni Prize Winners is a prestigious indicator of student academic achievements; thus, its frame of reference is very narrow. It can help distinguish the world's top universities but cannot differentiate the majority of universities that have no alumni prize winners. Staff Prize Winners, a prestigious staff reputation indicator, shares the same strength and weakness as Alumni Prize

Winners. The Articles Published in Nature and Science and Publication are strong indicators for Staff Productivity with a limited frame of reference.

Size of Institution consists of the weighted scores for the later five indicators divided by the number of Full Time Equivalent (FTE) academic staff. The numbers of FTE are not obtained for all countries; they are only available for institutions in the U.S., the UK, France, Japan, Italy, China, Australia, the Netherlands, Sweden, Switzerland, Belgium, South Korea, Czech, Slovenia, New Zealand and so forth. The different treatments for institutions in different countries may introduce some unpredicted bias. Further discussion of this indicator follows in Chapter Six.

The WUR's Indicators

The indicators of the WUR are Peer Review (40% of weightings), Employer Review (10% of weightings), Student to Staff Ratio (20% of weightings), Citation (20% of weightings), International Student Number (5% of weightings), and International Staff Number (5% of weightings). A detailed summary of each indicator is given in the Appendix.

The WUR covers Institution Reputation (Peer Review and Employer Review), Teaching and Learning (Student to Staff Ratio), Staff Productivity (Citation), and Environment (International Student Number and International Staff Number), according to the new classification. Both Peer Review and Employer Review are good indicators of an institution's reputation. However, the quality of the peer and

employer opinion surveys is subject to the design of the questionnaire, the sampling method, and the response rate. More discussion on this issue is provided in Chapter Six. Citation is a good indicator of Staff Productivity. With a wide frame of reference, this indicator allows all institutions to enter the competition. International Student Number and International Staff Number are related to the topic of Environment. Although their original purpose is to address the diversity of an institution, these indicators actually measure the size of international bodies rather than diversity. this topic is discussed further in Chapter Six.

Ranking Method

Both the ARWU and the WUR use the weight-and-sum ranking method, presenting the rankings of universities, the scores for each category of indicators and the total scores of universities in their final publications. Category scores are presented in a rescaled manner with the top institution assigned a value of 100. The scores of other universities are calculated as a proportion of that top score and the same applies to the total scores. After each category is weighted, universities are ranked against their peers in descending order, based on their composite weighted score. The weightings are arbitrary. No explanation of the theoretic basis is provided by the compilers for assigning the weightings.

Methodology Changes

To determine the factors that influence the rankings, it is first necessary to verify any changes in the ranking methodology. Since the first publication of their rankings, the

ARWU and the WUR have adjusted their methodology to improve their rankings, according to received feedback. The ARWU claims that they have resolved their method after 2005, while the WUR changes its methods each year.

The ARWU's Methodology Changes

The methodology changes for the ARWU are (SJTU, 2008):

2004

- Fields Medal is added.
- Indicator of Alumni Prize Winners is added.
- The weighting of size decreases from 20% to 10%, with the 10% weighting allocated to Alumni Prize Winners.

2005

- Article in Arts and Humanities Index is added.
- Special weight is allocated to articles in the Social Science Citation Index and Arts and Humanities Citation Index.

2006

No changes

2007

No changes

2008

The name of SCI is changed into PUB, and the name of Size is changed to PCP with no substantial changes.

The WUR's Methodology Changes

The methodology changes for the WUR are as follows (Sowter, 2009, p.1):

2005
Add employer review component
Collect full data on 200 additional universities
2006
Switch from 10 years to 5 years for citations measure
2007
Switch from ESI to Scopus for citations
Adopt new normalisation methodology
Insist upon FTE for all personnel metrics
Peer reviewers unable to select own institution
2008
Separate int'l and domestic responses to surveys

The effects of these changes will be revealed by the statistical analysis and multiple case studies.

Results of Standardisation

To standardise the data, seven points are chosen: 100, value at 90%, value at 70%, value at 50%, value at 30%, value at 10%, and zero. Data for early year are made equivalent to the data for 2008. For example, to covert the 2004 WUR total score to the 2008 WUR total score, value at 90% of the 2004 WUR total score is identified as 44.4 and its counterpart in the 2008 WUR is identified as 95.7. Other points are identified in a similar manner (see Figure 2).

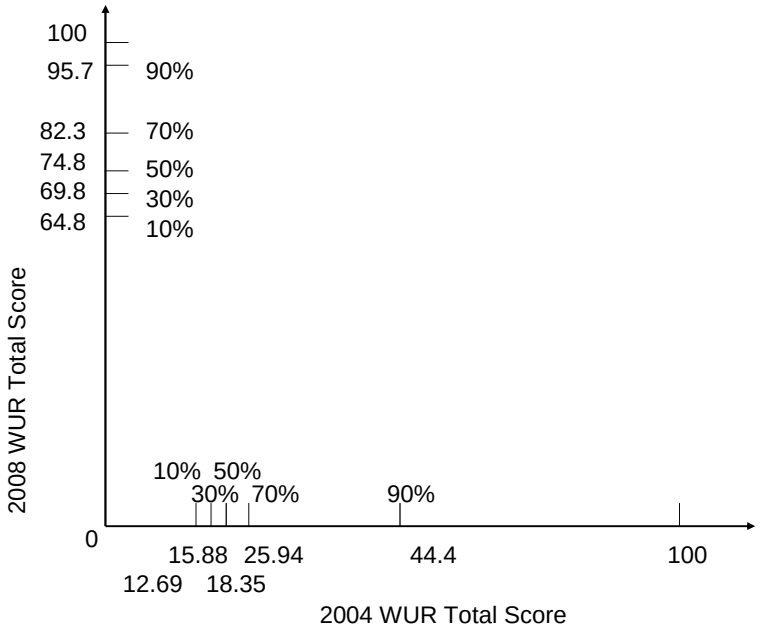


Figure 2 2004 vs. 2008 WUR Total Score

After all points are identified, the 2004 WUR original total score (Y_1) from 100 to 90% are standardised to Y_{1S} with the following equation:

$$Y_{1S} = 95.7 + \frac{(100-95.7)(Y_1-44.4)}{(100-44.4)}$$

The same equation applies to the values at 70%, 50%, 30%, 10%, and zero. The 2005 WUR scores (Y_{2S}), 2006 WUR scores (Y_{3S}) and 2007 WUR (Y_{4S}) scores are calculated by the same process. The same applies to data in the ARWU from 2003 to 2007. All calculations were performed by SPSS.

The standardised distributions of the WUR total scores from 2004 to 2008 are presented in Figure 3; the standardised distributions of the ARWU total scores from 2003 to 2008 are given in Figure 4.

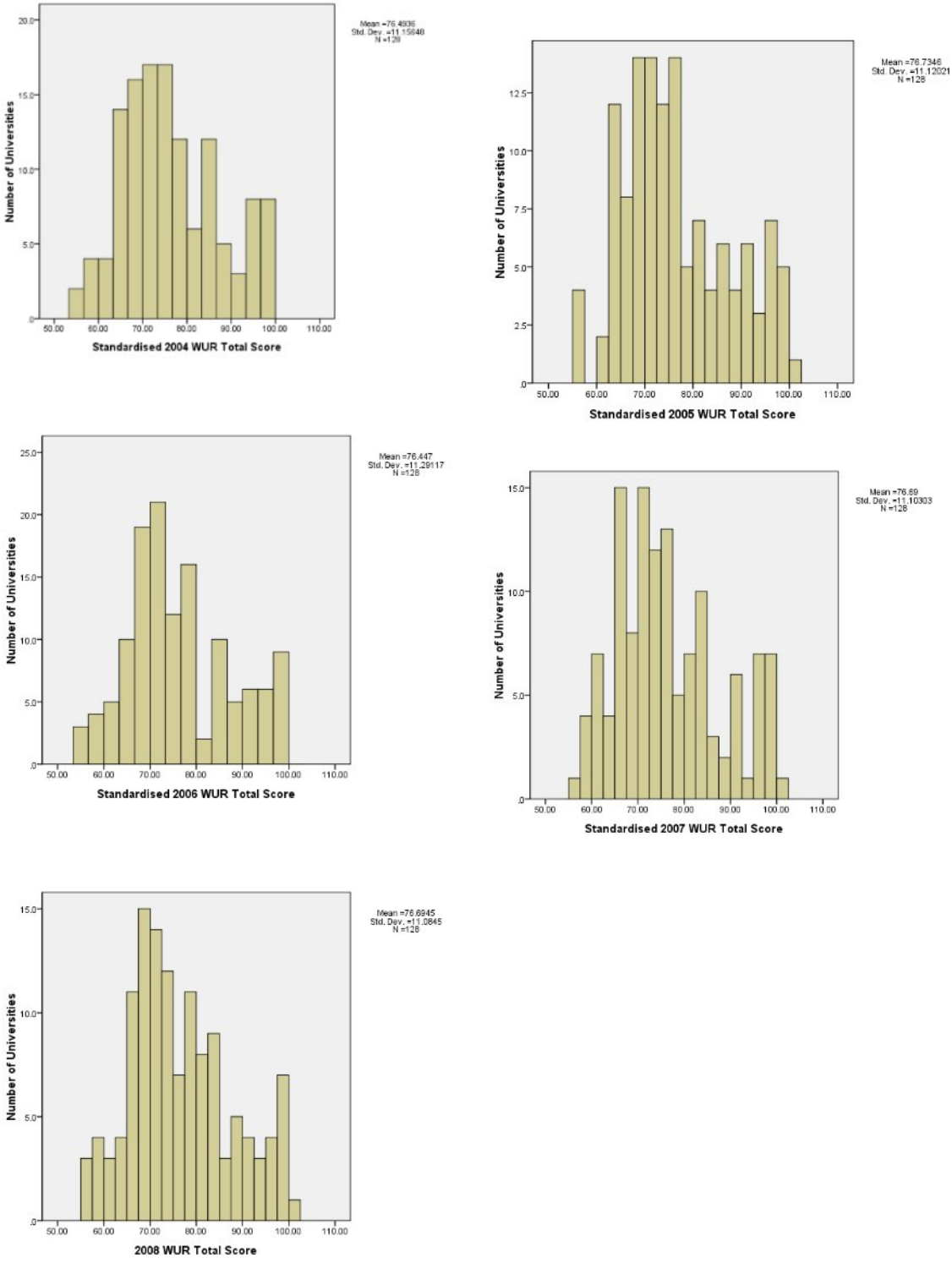


Figure 3 Standardised Results of the WUR

from 2004 to 2008

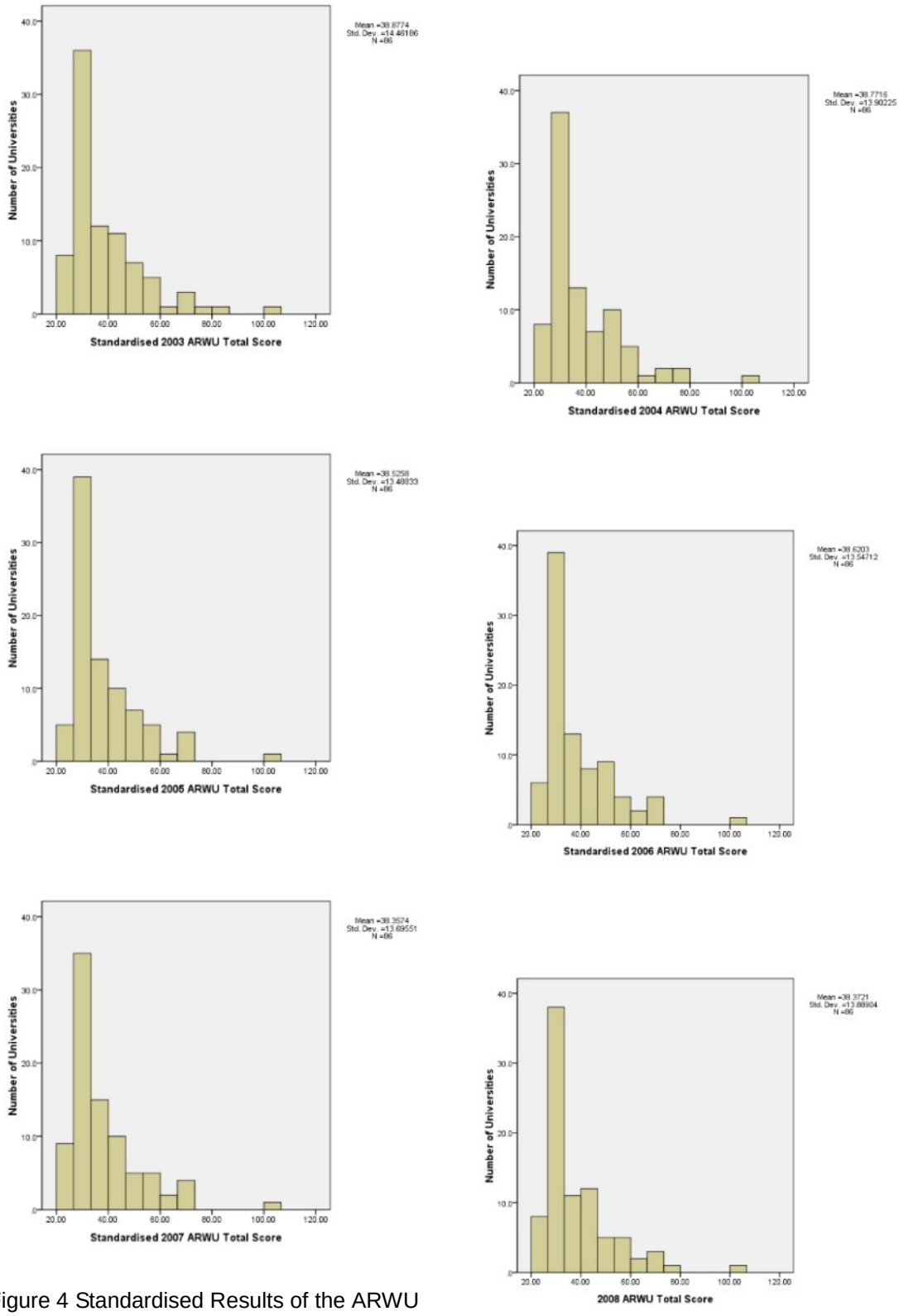


Figure 4 Standardised Results of the ARWU

from 2003 to 2008

Relationship between Total Scores and Rankings

The changes seen in the published total scores (the non-standardised scores) do not correspond to changes in the rankings. In other words, an increase in total score does not correlate with a raise in rank, and vice versa. This finding indicates that the published total scores and the scores of indicators are unable to show the absolute changes experienced by universities. It is not possible to judge whether a university has made overall progress or improved in the aspects related to particular indicators solely by the scores. To draw comparisons between universities, we need the assistance of the overall rankings and the rankings by particular indicators.

Trends of the ARWU and the WUR

By defining the standardised total scores on the x-axis and the ranking on the y-axis, the scatter plots are drawn to show different trends of the ARWU and the WUR, in which each point represents a university. For the ARWU, six graphs (Figure 5 to Figure 10) are drawn to show the trends from 2003 to 2008; for the WUR, five graphs (Figure 11 to Figure 15) are drawn to show the trends from 2004 to 2008.

The ARWU's Trends

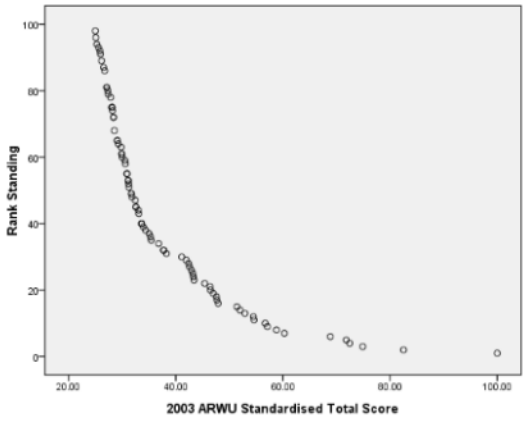


Figure 5 2003 ARWU Standardised Total Score vs. Ranking

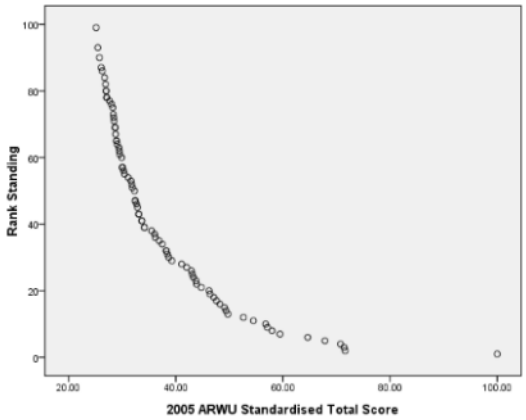


Figure 6 2005 ARWU Standardised Total Score vs. Ranking

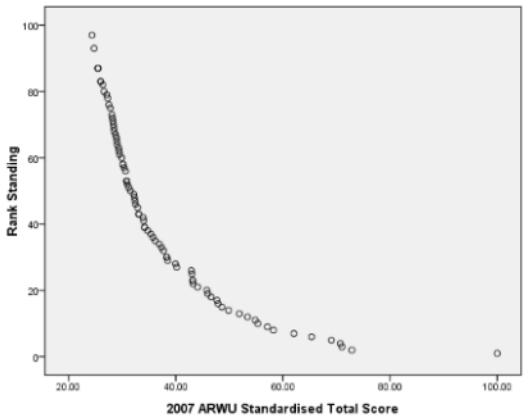


Figure 7 2007 ARWU Standardised Total Score vs. Ranking

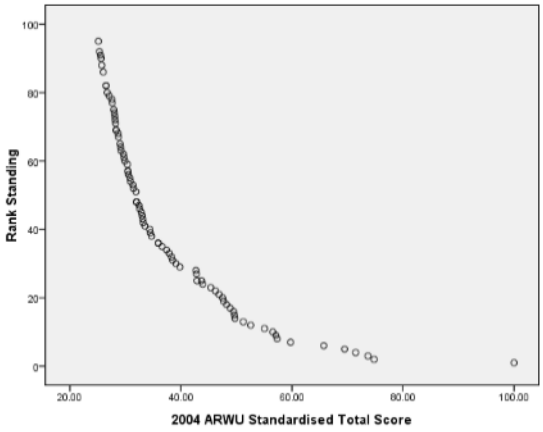


Figure 8 2004 ARWU Standardised Total Score vs. Ranking

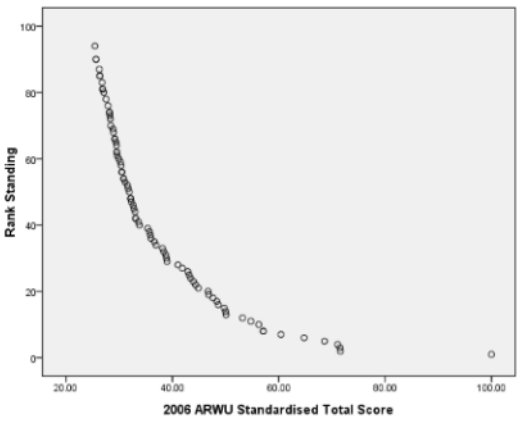


Figure 9 2006 ARWU Standardised Total Score vs. Ranking

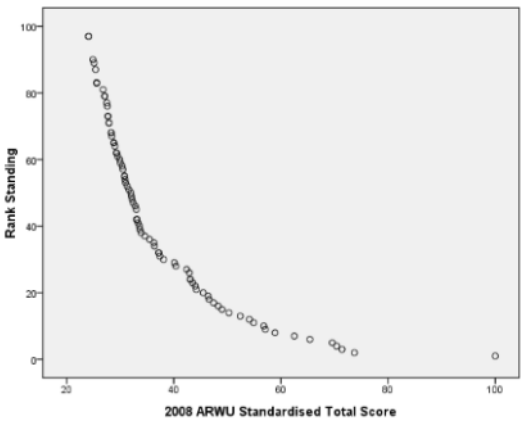


Figure 10 2008 ARWU Standardised Total Score vs. Ranking

The WUR's Trends

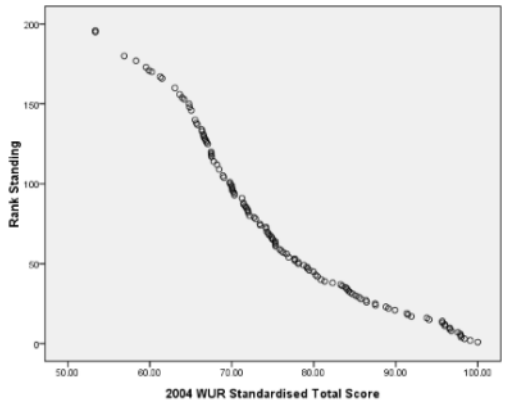


Figure 11 2004 WUR Standardised Total Score vs. Ranking

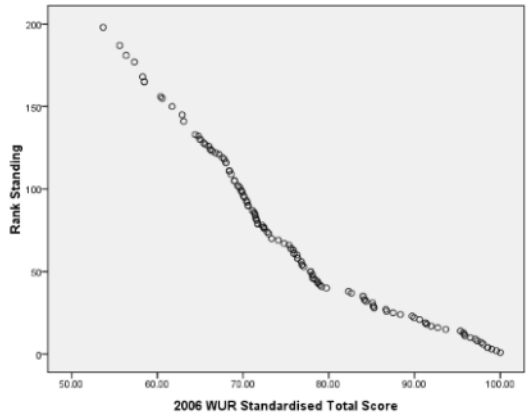


Figure 12 2006 WUR Standardised Total Score vs. Ranking

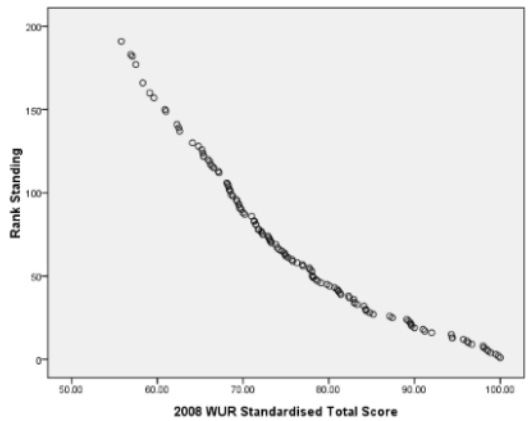


Figure 13 2008 WUR Standardised Total Score vs. Ranking

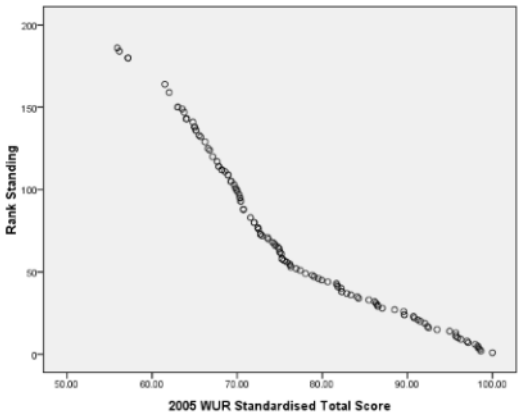


Figure 14 2005 WUR Standardised Total Score vs. Ranking

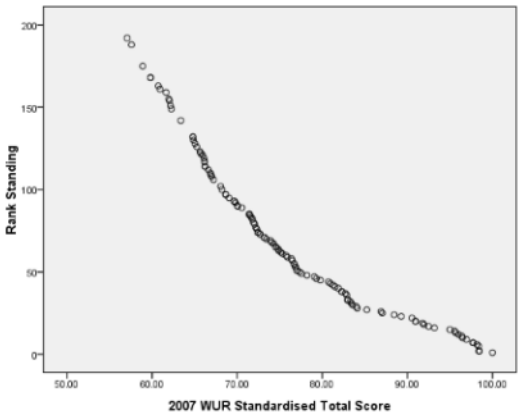


Figure 15 2007 WUR Standardised Total Score vs. Ranking

As illustrated above, the ARWU and the WUR demonstrate consistent trends over the years. Consequently, the ARWU's trends for the six years can be merged into a single graph (see Figure 16). This applies to the WUR as well (see Figure 17). However, the ARWU and the WUR demonstrate different patterns.

In the plot of the ARWU's trends, three columns emerge (see Figure 16). The first column contains huge gaps between the standardised total scores of neighbouring rankings. In the second column, the gaps decrease. In the third column, not much difference is shown between the standardised total scores of neighbouring rankings. The subtle differences in scores create huge differences in the rankings. Thus, the rankings of universities in the first column are most reliable, because considerable differences must exist in the total scores to achieve different ranks. The rankings of universities in the third column are the least reliable. Since any fluctuations due to methodology changes would shift their ranking when their total scores are similar.

When the six years' data for the WUR are merged together, the overlap seen is not as neat as for the ARWU. Three columns also emerge from the plot (see Figure 17). In the first column, the differences in rank require large differences in the standardised total scores. In the second column, the universities' rankings are differentiated by subtle differences in the standardised total scores. In the third column, the tension is relieved, but the five years' data are not perfectly matched. Thus, the rankings of the universities in the first column are more reliable, because they need considerable

differences in the standardised total scores to achieve different ranks. The rankings of the universities in the second column are less reliable since any fluctuations due to methodology changes would shift their rankings when their total scores are similar.

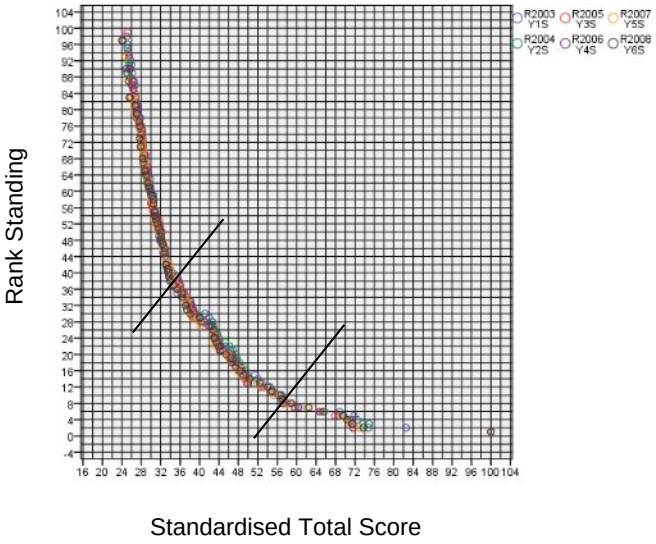


Figure 16 Six Years' Trends for the ARWU

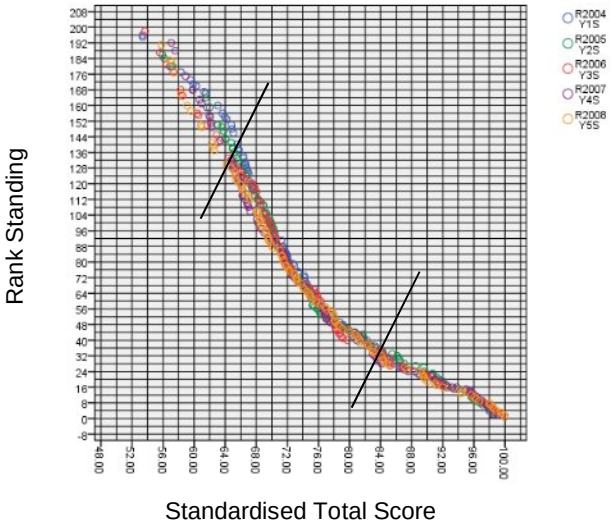


Figure 17 Five Years' Trends for the WUR

The Relationship between the Indicators

The relationships between the indicators are explored by correlation, multiple regression and principal component analysis.

Results from Correlation

The ARWU and the WUR differ in some respects. For example, the ARWU gives 60% weight to Citation and Publication, while the WUR assigns 50% weight to reputation survey. On the other hand, they share some similarities. The bibilometric indicators tend to be correlated with reputation survey. The correlation coefficients between the ARWU and the WUR are shown in Table 5.

	Correlation between the ARWU and the WUR
2004	.71
2005	.68
2006	.70
2007	.61
2008	.63

Table 5 The Correlation between the ARWU and the WUR

As illustrated in Table 5 the two rankings share more similarity, around 49%, from 2004 to 2006. After 2007, the commonality reduces to around 36%.

Moreover, the ranking based on Peer Review in the WUR has a higher correlation with the ARWU's ranking ($r = .715$, $p < .01$ in 2007; $r = .693$, $p < .01$ in 2008), which suggests the rankings by peer review are more similar to the ARWU overall rankings. In fact, the rankings by Peer Review in the WUR share even more similarity with the ARWU overall rankings one year earlier ($r = .716$, $p < .01$ in 2007; $r = .696$, $p < .01$ in

2008), which suggests that the respondents may consult the previous year's ARWU results when answering the WUR's peer review questionnaire.

The yearly correlation coefficients between the indicators of the ARWU and the WUR are presented in Appendix.

The correlation coefficients differ from year to year for both the ARWU and the WUR. In the ARWU, all correlation coefficients among the variables are positive. All of the variables are highly correlated with the standardised total scores, and the correlation coefficients among the variables are very high, except for that of Publication, which suggests that considerable overlaps in the measured elements. Yet, in the WUR, the correlation coefficients for International Staff Number, International Student Number, and Citation are sometimes negative, which suggests underlying conflicting factors. All of the variables are correlated with the standardised total scores, Peer Review the most highly and International Staff the least highly. The correlation coefficients among the variables are not very high, which suggests that there are limited overlaps in what is being measured.

Clarke raises a relevant question regarding the degree of correlation that should be present between indicators:

> Are high correlations a sign of validity (i.e., are they all measuring quality)—or of redundancy (are they all measuring the same thing, and are therefore not all needed?) Are low correlations a sign of invalidity (i.e., are some or all of the indicators not measuring quality?) or unique information (i.e., are they all

> measuring different aspects of academic quality?). In addition, if efficiency is of value, and if the indicators are highly correlated, should some of the indicators be dispensed with since doing so would probably not affect the final ordering of schools?
>
> Clarke (2002, p.450)

Results from Multiple Regression

The data for the ARWU are quite consistent during the six-year period. All of the variation in the standardised total scores is explained on the basis of their linear relationship with the six indicators ($R^2 = 1.00$). The correlation coefficient between the published ranking and the predicted ranking is 1.00. By contrast, there are many variations in the WUR's data for 2004 to 2008. In 2004 and 2007, not all of the variation in the Overall Score can be explained based on its linear relationship with the six indicators ($R^2 = 0.997$ in 2004;$R^2 = 0.994$ in 2007), although, in the other years $R^2 = 1.00$. The total weightings of the indicators should be within the limit of 100%, but they go beyond the limit in 2005 and 2006. The correlation coefficient between the published ranking and the predicted ranking is 1.00 in 2005, 2006, and 2008, but is 0.998 in 2004 and 0.997 in 2007.

By multiple regression, we can test whether the alleged weightings are the actual weightings contributing to the standardised total scores. As illustrated in Table 5, the ARWU's actual weightings generally match the alleged weightings. However, the WUR's actual weightings do not match the alleged weightings before the 2007, the first year of WUR's standardisation (see Table 7). After 2007, the actual weightings

move closer to the alleged weightings. Using the non-standardised data to conduct rankings violates the original purpose of assigning weightings to each indicator.

Weightings	Alleged 2003	Actual 2003	Alleged 2004-2008	Actual 2004	Actual 2005	Actual 2006	Actual 2007	Actual 2008
Alumni Prize Winners			0.10	0.10	0.10	0.10	0.10	0.10
Staff Prize Winners	0.20	0.21	0.20	0.21	0.21	0.21	0.21	0.21
Highly Cited Researchers	0.20	0.21	0.20	0.21	0.21	0.21	0.21	0.21
Articles in Nature or Science	0.20	0.21	0.20	0.21	0.21	0.21	0.21	0.20
Publications	0.20	0.21	0.20	0.21	0.21	0.20	0.21	0.21
Size of Institution	0.20	0.21	0.10	0.10	0.10	0.10	0.10	0.10

Table 6 Alleged and Actual Weightings of the ARWU

Weightings	Alleged 2004	Actual 2004	Alleged 2005-2008	Actual 2005	Actual 2006	Actual 2007	Actual 2008
Peer Review	0.50	0.11	0.40	0.59	0.56	0.41	0.41
Employer Review			0.10	0.15	0.14	0.10	0.10
Student to Staff Ratio	0.20	0.09	0.20	0.30	0.28	0.20	0.21
Citation	0.20	0.10	0.20	0.30	0.28	0.20	0.21
International Staff	0.05	0.1	0.05	0.08	0.07	0.05	0.05
International Students	0.05	0.11	0.05	0.07	0.07	0.05	0.05

Table 7 Alleged and Actual Weightings of the WUR

In the ARWU, Alumni Prize Winners and Staff Prize Winners account for most of the variance (69% in 2003, 77% in 2004, 77% in 2005, 76% in 2006, 77% in 2007, 76% in 2008) in the standardised total scores. Highly Cited Researchers and Articles in Nature and Science are the second largest factors that account for the variance in the rankings (28% in 2003, 21% in 2004, 22% in 2005, 23% in 2006, 22% in 2007, 23% in 2008). Publication and Size play little role in contributing to the rankings.

In the WUR, Peer Review accounts for most of the variance in the rankings, but its importance decreases throughout the years (89% in 2004, 83% in 2005, 76% in 2006, 62% in 2007, and 55% in 2008). This could explain the decreasing correlation between the ARWU and the WUR found by correlation analysis. Before 2007, the WUR's results are highly dependent on Peer Review, which is highly correlated with the bibilometric indicators, to which the ARWU allocates 60% weight.

The importance of Student to Staff Ratio increases from 7% in 2004 to 6% in 2005, 11% in 2006, 19% in 2007 and 21% in 2008. The importance of Citation also grows from 3% to 11% from 2004 to 2008. Employer Review accounts for around 6% to 8% of the rankings. The contributions from International Staff Number and International Student Number remain low. (Detailed results are included in Appendix.)

Results from Principal Component Analysis

Table 8 shows a comparison of the results of the Principal Component Analysis for the ARWU and the WUR, which includes the number of components extracted, the percentage explained by each component, and the correlation coefficients (r) between the components.

	ARWU			WUR					
	Components		r	Components			r		
Year	Ⅰ	Ⅱ	r- Ⅰ & Ⅱ	Ⅰ	Ⅱ	Ⅲ	r- Ⅰ & Ⅱ	r- Ⅱ & Ⅲ	r- Ⅰ & Ⅲ
2003	64.1%	20.8%	.35						
2004	63.6%	17.5%	.40	36.2%	28.3%		-.125		
2005	68.0%			34.1%	27.4%		-.059		
2006	67.0%	17.0%	.43	34.6%	27.3%		-.022		
2007	67.3%	17.0%	.43	33.6%	22.3%	17.6%	.090	.050	.220
2008	67.0%	16.8%	.42	33.8%	23.3%		.096		

Table 8 The comparison of the results of the Principal Component Analysis for the ARWU and the WUR's

As shown in Table 8, two components are extracted from the ARWU in most of the years, except 2005, and two components are extracted from the WUR in most of the years, except 2007. The 2005 for the ARWU and the 2007 for the WUR are special years, because they are the years when the two rankings experience dramatic changes. In 2005, the ARWU added the Article in Arts and Humanities Index and assigned special weight to articles in the Social Science Citation Index and Arts and Humanities Citation Index. In 2007, the WUR switched from ISI to Scopus for citations, adopted a new normalisation methodology, insisted upon FTE for all

personnel metrics and introduced a new rule for Peer Review such that the peer reviewers cannot select their own institution.

The importance of the first component in the ARWU increases from 2003 to 2008, while that of the second component decreases. On the other hand, the importance of the first component in the WUR first decreases but later increases, whereas the importance of the second component first increases and then drops.

In the ARWU, the indicators associated with the first component are Staff Prize Winners, with the highest loadings, followed by Alumni Prize Winners, Size, Articles Published in Nature and Science, and Highly Cited Researchers. The indicators related to the second component are Publication, with the highest loadings, followed by Highly Cited Researchers, and Articles Published in Nature and Science. The detailed results of the principal component analysis for the ARWU are reported in Appendix.

In the WUR, the indicators associated with the first component are International Staff number, with the highest loadings. International Staff Number and International Student Number and Citation are negatively loaded with each other. The indicators related to the second component are Peer Review, with highest loadings, followed by Employer Review and Citation. Student to Staff Ratio contributes to neither the first component nor the second component. Thus, it should be regarded as the third

component, as indicated in the 2007 analysis. The detailed results of the principal component analysis for the WUR are reported in Appendix.

To sum up, with a modest correlation between the two components, the ARWU provides a unidimensional scale, generally measuring staff reputation and publication. The WUR suffers from a lack correlation between the three components and even negative relationship between the indicators within one component. Moreover, it measures different dimensions within one single scale, which creates conflicts between the indicators and harms the validity of the rankings.

Chapter Five: Multiple Case Studies

Choosing Cases

As indicated in the methodology, cases should be chosen from critical, exemplifying, revelatory, and unique cases. The critical cases consist of universities that demonstrate significant changes, while the exemplifying cases comprise the majority of universities that demonstrate random changes.

Significant Changes over the Years

To identify whether the changes are significant changes or random changes, the runs test is conducted first. It turns out that the results of the runs test for all 86 universities in the ARWU and 128 universities in the WUR are statistically significant, which means the yearly changes are random. Consequently, no university's change is identified as significant by the runs test, possibly because runs test is not sensitive to the data for such a short period (six years for the ARWU and five years for the WUR).

Since no statistical method is available for such a short period, we resort to checking the movement of each university over the years by a simple definition: if a university moves continuously up or down in rank, the changes are considered significant. From 2003 to 2008, in the ARWU, seven out of 86 universities (8.1%) demonstrate continuous upward movement and 14 (16.3%) demonstrate continuous downward

movement. Two (2.3%) universities remain at the same rank. Sixty-three (73.3%) universities move both up and down.

In the WUR, from 2004 to 2008, six out of 128 (4.7%) universities show continuous upward movement and nine universities (7.0%) show continuous downward movement. One (0.8%) university maintains a stable ranking. Of the 128 universities, 112 (87.5%) demonstrate fluctuation.

In the ARWU, the following universities continuously move upwards from 2003 to 2008:

Columbia University

Massachusetts Institute of Technology

University of California—Berkeley

University of Colorado—Boulder

University of Manchester

University of Maryland—College Park

University of Pennsylvania

In the ARWU, the following universities continuously move downward from 2003 to 2008:

California Institute of Technology

Case Western Reserve University

Imperial College London

Princeton University

Rice University

Rockefeller University

Tohoku University

University of California—San Francisco

University of Edinburgh

University Munich

University of Texas Southwestern Medical Center

Vanderbilt University

Washington University—St. Louis

Yale University

In the ARWU, the following universities maintain the same rank from 2003 to 2008:

Cornell University

Harvard University

In the WUR, the following universities continuously move upwards from 2003 to 2008:

King's College London

Seoul National University

University College London

University of Utrecht

University of Alberta

Yale University

In the WUR, the following universities continuously move downwards from 2003 to 2008:

London School of Economics

Massachusetts University—Amherst

Monash University

Sussex University

University of California-Berkeley

University of California-San Diego

University of Paris VI

University of Texas—Austin

University of New South Wales

In the WUR, the following university maintains the same rank from 2003 to 2008:

Harvard University

Interestingly, the direction of significant changes for University of California —Berkeley is upward in the ARWU but downward in the WUR. By contrast, the

direction of significant changes for Yale University is downwards in the ARWU, but upward in the WUR.

Universities for Case Study

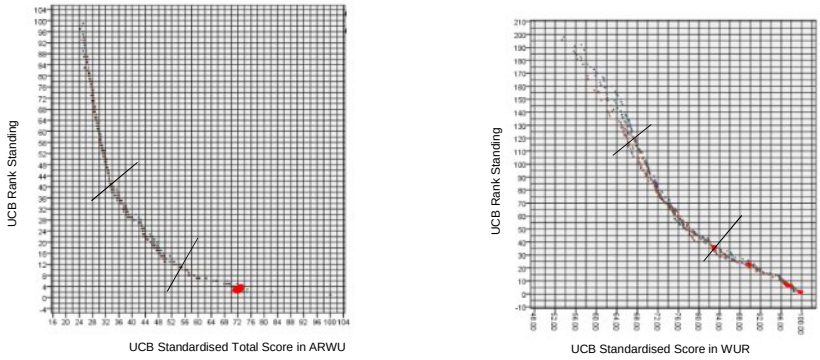
As mentioned in the methodology, the cases should be critical, exemplifying, revelatory, and unique. Given these criteria, the campuses of the University of California draw our attention. University of California—Berkeley (UCB) demonstrates significant changes in both the ARWU and the WUR. Interestingly, it is significant on moving upwards in the ARWU but downwards in the WUR. University of California—San Diego (UCSD) shows significant changes in the WUR but not in the ARWU. University of California—San Francisco (UCSF) demonstrates significant change downwards in the ARWU, dropping from 13th place in 2003 to 18th place in 2008, while in the WUR, it ranks 20th place in 2004, 17th in 2005, but disappears from the top 500 afterward. Neither University of California—Los Angeles (UCLA) nor University of California—Santa Barbara (UCSB) demonstrates consistent change. University of California—Davis (UCD), University of California—Irvine (UCI), University of California—Riverside (UCR), and University of California— Santa Cruz (UCSC) are excluded from the statistical analysis due to missing data. Consequently, UCB, UCSD and UCSF, those that demonstrate significant changes, serve as the critical cases. UCLA, UCSB, UCD, UCI, UCR and UCSC serve as exemplifying cases to represent the majority of the universities. All of these universities are also revelatory cases, because they represent universities within different columns. UCB belongs to the first column; UCLA,

UCSD and UCSF are in the second column; UCSB is in the second column in the ARWU, but in the WUR, it moves from the second to the third column. The graphs illustrating the universities' positions in the columns are given in Figure 18 through Figure 23 . Moreover, the nine universities enjoy a unique status as different campuses within the University of California system. The entire range of University of California campuses provides unique cases to study universities with various features within the same macro political, economic, and cultural environment. Thus, the nine universities chosen for multiple case studies fulfil the criteria of being critical, exemplifying, revelatory and unique.

University of California-Berkeley (UCB)

UCB	2003	2004	2005	2006	2007	2008
ARWU Rankings	4	4	4	4	3	3
WUR Rankings		2	6	8	22	36

Table 9 Rankings of UCB



ARWU WUR

Figure 18 UCB's Positions in Columns

Table 9 shows UCB's yearly changes in rank in the ARWU and the WUR. As shown in Table 9, UCB's rank is quite stable in the ARWU, but drops dramatically in the WUR.

Figure 18 illustrates UCB's positions in the columns for all universities, with the standardised total scores in the x-axis and standings in the y-axis. The black line is composed of all universities in the ARWU and the WUR, respectively. The line is also divided into three columns. The points in red represent UCB's yearly position the ARWU (on the left) and the WUR (on the right). This shows that UCB belongs to the first column in both ranking systems.

UCB in the ARWU

UCB	2003	2004	2005	2006	2007	2008
Rankings	4	4	4	4	3	3
Total Scores	74	74.2	72.8	72.1	71.9	71.4
Alumni Winners		70	71.8	70.6	72.5	69
Rankings by Alumni		6	6	6	5	5
Staff Winners	75	76	76	74.5	77.1	77.1
Rankings by Staff Winners	6	6	6	6	7	7
HiCi	70.3	74.1	69.4	70.5	69.7	68.8
Rankings by HiCi	3	3	3	3	3	3
N&S	76.1	75.6	73.9	72.2	72.9	70.6
Rankings by N&S	2	2	2	2	2	2
Pulication	72.8	72.7	72.2	71.9	69.2	70
Rankings by Publication	8	8	9	9	10	10
Size	51.8	45.1	52.7	53.1	52.6	53
Rankings by Size	11	14	8	7	7	8

Table 10 UCB Yearly Changes in the ARWU

HiCi: Highly Cited Researchers; N&S: the Articles Published in Nature and Science

The reliability of the ARWU is checked first. As shown in Table 10, the total scores in the ARWU are unable to show the movements in the rankings. For example, UCB ranks higher in 2007 and 2008, but its total scores are lower than the scores in previous years.

To verify the data for Alumni Prize Winners and Staff Prize Winners, it is necessary to check the real number of prizewinners:

2004 David J. Gross, a UCB alumnus, won a third of the Nobel Prize in Physics.

2005 Thomas C. Schelling, a UCB alumnus, won half the Nobel Prize in Economics.

2006 George F. Smoot, a UCB staff member, won half the Nobel Prize in Physics.

2006 Andrew Z. Fire, a UCB alumnus, won half the Nobel Prize in Physiology or Medicine.

Again this shows the scores of indicators do not reflect the real changes in the university. For example, with the increase in number of alumni and staff prizewinners, the two indicators should increase as well. However, such an increase is not reflected in the scores of the two indicators.

Because the scores of the indicators are not reliable, it is better to use the rankings by the indicators to analyse the changes in indicators. The rankings by indicators are also shown in Table 10.

By comparing the rankings by indicators, the changes in UCB's rank in the 2007 ARWU is attributed to the one-place rise in Alumni Prize Winners, despite the drops in Staff Prize Winners, Publication and Size. In fact, the one-place rise in Alumni Prize Winners does not solely depend on the half of the Alumni Nobel Prize won in 2006. Instead, it is the result of three years' accumulation of alumni Nobel Prizes wins. As listed above, one-third of an alumni Nobel Prize was won in 2004 and half an alumni Nobel Prize was won in 2005. As a university in the first column (see Figure 18), UCB needs a considerable increase in scores to move upward by one place. A single alumni or staff prizewinner is not enough to help it move up the list. This is also evidenced in the UCB's 2007 change in rank by the Staff Prize Winners. UCB

has half a staff Nobel win, but the ranking of this indicator falls from sixth to seventh place. UCB's rankings by Highly Cited Researchers and Articles Published in Nature and Science do not change from 2003 to 2008. Therefore, UCB's change in rank in the ARWU is caused by three years of accumulated increases in Alumni Prize Winners.

UCB in the WUR

UCB	2004	2005	2006	2007	2008
Rankings	2	6	8	22	36
Total Scores	880.2	80.6	80.4	89.7	82.9
Peer Review	665	95	92	100	100
Rankings by Peer Review	1	3	4	1	1
Employer Review		62	75	98	100
Rankings by Employer Review		11	9	11	1
Student to Staff Ratio	7	7	22	59	24
Rankings by Ratio	112	108	97	79	120
Citation	169	39	39	92	100
Rankings by Citation	7	7	7	15	1
International Staff	6	7	6	73	88
Rankings by Inter Staff	107	105	111	51	36
International Students	7	13	13	88	36
Rankings by Inter Students	106	95	93	36	104

Table 11 UCB Yearly Changes in the WUR

As shown in Table 11, the total scores and the scores of indicators in the WUR are unable to show the movement in the rankings, because they use different scales. The values in 2004 have a full score of 1,000, but in the rest of the years, the scale uses 100 as the full score. Even within the same scale, the values are not comparable. For example, UCB has a higher rank in 2005 and 2006, but the total scores in 2005 and

2006 are lower than 2007 and 2008. Therefore, to ensure comparability, rankings by indicators should be employed; these are also given in Table 11.

UCB's rank falls continuously from 2004 to 2008. However, no ranking by indicators show a continuous downward movement. By contrast, its rankings by Citation, International Staff Number and International Student Number, and Employer Review show an upward movement between 2004 and 2008. The rankings by Peer Review remain at number one in 2004 and 2008. Thus, the continuous drops in the UCB's rank standings are not supported by its indicators, a phenomenon that is difficult to explain. Still, one likely explanation could be changes in the real weightings, identified by multiple regression analysis.

To verify the scores of indicators, it is necessary to gather data from different sources to triangulate the values. To check if the scores for Student to Staff Ratio, International Student Number and International Staff Number reflect the actual situation within the university, data are gathered from the Common Data Set provided by university itself, University of California's (UC) central office, the National Center for Education Statistics (NCES) and U.S. News America's Best Colleges. For Peer Review, no raw data are available; thus, the values are compared with the same indicator in U.S. News America's Best Colleges. The data from these sources are listed in Table 12.

UCB	2003	2004	2005	2006	2007
Common Data Set Full-time Undergraduates	21941	21771	22295	23107	23863
Common Data Set Full-time Graduates	8632	8641	8678	8678	8879
Common Data Set Part-time Undergraduates	1264	1109	1187	756	773
Common Data Set Part-time Graduates	471	539	633	613	665
Common Data Set Full-time Instructors	1475	1496	1543	1545	1601
NCES Full-time Faculty	1494	1515	2309		2214
UC Office FTE	3089.23	3022.8	3215.2	3215.04	3372.52
Common Data Set Student to Staff Ratio	15.5 to 1	15.5 to 1	15.4 to 1	15.3 to 1	15.1 to 1
U.S. News Student to Staff Ratio	16 to1	16 to1	16 to1	15 to 1	15 to 1
Common Data Set International Students	2572	2537	2540	2521	2625
Common Data Set International Staff	57	61	58	51	56
NCES International Staff		62	186		80
U.S. News Peer Review	4.7	4.8	4.8	4.7	4.8

Table 12 Triangulating Data for UCB

FTE: Full Time Equivalent

There are discrepancies among the data from different sources. According to the data from the Common Data Set, UC's central office and the U.S. News, the values for Student to Staff Ratio, International Student Number and International Staff Number and Peer Review are relatively stable over the years. However, the data from the NCES are incomplete and demonstrate huge fluctuations. This causes concern about

the reliability of the data from the NCES, a third party, which could result in inaccurate rankings.

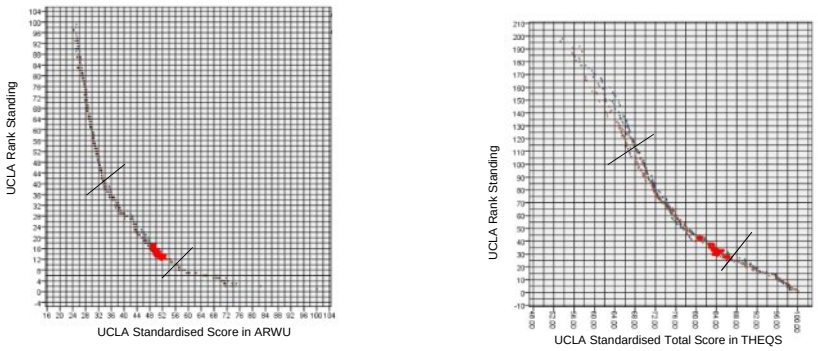
University of California - Los Angeles (UCLA) and University of California- San Diego (UCSD)

UCLA	2003	2004	2005	2006	2007	2008
ARWU Rankings	15	16	14	14	13	13
WUR Rankings		26	37	31	41	30

Table 13 Rankings of UCLA

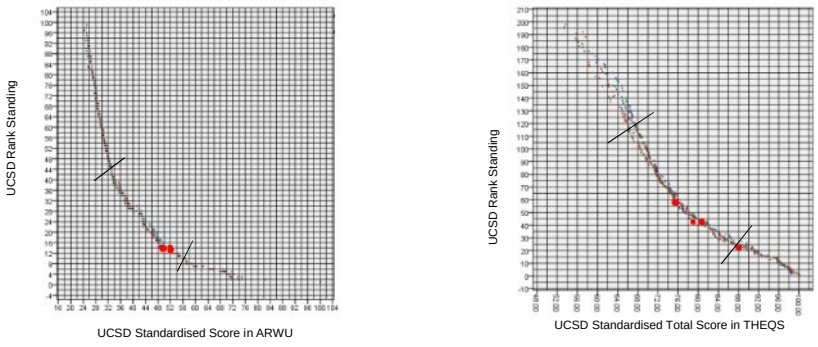
UCSD	2003	2004	2005	2006	2007	2008
ARWU Rankings	14	13	13	13	14	14
WUR Rankings		24	42	44	58	58

Table 14 Rankings of UCSD



ARWU WUR

Figure 19 UCLA's Positions in Columns



ARWU WUR

Figure 20 UCSD's Positions in Columns

As shown in Table 13 and Table 14, UCLA's rank in the ARWU is relatively stable, while its position in the WUR ranges from 26th to 41st. UCSD encounters a similar situation. UCSD maintains 13th or t 14th place in the ARWU, while in the WUR, its rank drops continuously from 24th to 58th place, showing significant decreases. When compared with the two universities in the two ranking systems, an interesting phenomenon occurs. In the WUR, UCLA ranks higher than UCSD, except in 2004. In the ARWU, UCLA ranks lower than UCSD from 2003 to 2006. However, after 2007, UCLA takes a higher position than UCSD.

Figure 19 and Figure 20 show that UCLA and UCSD are universities in the second column in both the ARWU and the WUR. In the following analysis, we will find that

as universities in the second column, their rankings will change one or two places with a win of even a half or a third of a prize.

UCLA in the ARWU

UCLA	2003	2004	2005	2006	2007	2008
Rankings	15	16	14	14	13	13
Total Scores	53.8	51.6	50.6	50.4	52.6	52.4
Alumni Winners		27.4	27.3	26.4	25.6	24.4
Rankings by Alumni		37	40	40	40	40
Staff Winners	37.3	32.8	32.8	32.1	42.8	42.8
Rankings by Staff Winners	20	27	29	29	14	15
HiCi	58	60.5	56.7	57.6	57.4	57.4
Rankings by HiCi	8	11	10	10	10	9
N&S	48	48.1	50.1	47.5	49.1	48.9
Rankings by N&S	15	16	14	16	15	15
Pulication	78	79.9	75.6	77.3	75.9	75.7
Rankings by Publication	3	3	5	4	5	5
Size	30.3	24.8	34.3	34.9	35.5	36
Rankings by Size	46	60	28	29	25	25

Table 15 UCLA Yearly Changes in the ARWU

HiCi: Highly Cited Researchers; N&S: the Articles Published in Nature and Science

As illustrated in Table 15, UCLA moves down one place in the 2004 ARWU. All the rankings by indicators fall except for Publication, which remains the same. In 2003, no Nobel Prize or Fields Medal was awarded to UCLA. UCLA's decrease in ranking by Staff Prize Winners in 2004 may result form being overtaken by other institutions awarded with a Nobel or Fields Medal. UCLA's ranking by Size suffers a great slip in 2004, dropping from 46th to 60th place. However, the actual number of FTE, the determiner of Size, remains quite stable from 2003 to 2007, as shown in Table 17.

Thus, the slip in Size may have been caused by a technical mistake. In 2005, UCLA's rank moves up to 14th place, because Size bounces back from 60th to 28th place. It seems the mistake related to Size was corrected. In 2006, UCLA rises to 13th place as its ranking by Staff Prize Winners increases from 29th to 14th place. The rise of in the indicator matches the actual situation: in 2006, Terence Tao, a UCLA staff member, won the Fields Medal. Thus, the factors contributing to the changes in UCLA's rank in the ARWU are Staff Prize Winners and likely some methodological mistake in calculating the score of the indicator.

UCLA in the WUR

UCLA	2004	2005	2006	2007	2008
Rankings	26	37	31	41	30
Total Scores	316.3	43.3	55.9	82.8	84.3
Peer Review	180	52	58	100	100
Rankings by Peer Review	22	20	19	1	1
Employer Review		6	42	92	98
Rankings by Employer Review		84	33	47	23
Student to Staff Ratio	12	12	34	56	48
Rankings by Ratio	77	81	70	83	96
Citation	106	24	25	91	100
Rankings by Citation	11	11	15	21	1
International Staff	2	2	2	20	23
Rankings by Inter Staff	124	122	122	118	122
International Students	8	11	12	36	36
Rankings by Inter Students	100	104	98	105	104

Table 16 UCLA Yearly Changes in the WUR

UCLA's rank falls from 26th to 37th place in the 2005 WUR, mainly due to the decrease in Student to Staff Ratio. The university's rank by this indicator falls from 77th to 81st place, but as shown in Table 17, the data for Student to Staff Ratio from the Common Data Set and the U.S. News remain stable. In the 2006 WUR, UCLA bounces back to 31st place because its rankings by Peer Review, Employer Review, Student to Staff Ratio, and International Student Number move up. The biggest jump occurs in Employer Review, rising from 84th to 33rd place, and Student to Staff Ratio from 81st to 70th place. Again, as shown in Table 17, the data for Student to Staff Ratio

remain stable. It is unclear why Student to Staff Ratio shifts dramatically while the data from other sources remain stable. In the 2007 WUR, UCLA drops 10 places, despite its ranking by Peer Review increasing from 19th to the first place. The most serious drop is in Student to Staff Ratio, falling 13 places. Again, the data for Student to Staff Ratio in Table 17 remain relatively stable, which does not support the drop of this indicator in the WUR. In the 2008 WUR, UCLA gains back 11 places. The major factors contributing to this recovery are Employer Review, moving up 24 places, and Citation, up 20 places. Citation is checked against its counterpart in another ranking: the Performance Ranking of Scientific Papers for World Universities (PRSPWU) by the Higher Education Evaluation and Accreditation Council of Taiwan. UCLA's score for Citation in the PRSPWU is 39.65 in 2007, which falls to 37.87 in 2008, a move in the opposite direction compared with the WUR. The same indicator demonstrates different features in different ranking systems. To sum up, UCLA's changes in rank in the WUR are caused by huge shifts in several indicators. However, these shifts cannot be verified by data from other sources. This causes serious concern about the reliability of the WUR.

UCLA	2003	2004	2005	2006	2007
Common Data Set Full-time Undergraduates	24560	23973	23815	24338	24783
Common Data Set Full-time Graduates	8796	8496	8414	8734	8933
Common Data Set Part-time Undergraduates	1117	941	959	1000	997
Common Data Set Part-time Graduates	155	160	110	82	162
Common Data Set Full-time Instructors		1859	1890	1948	
NCES Full-time Faculty	2676		3624		3476
UC Office FTE	4478	4489	4464.9	4607	4751.1
Common Data Set Student to Staff Ratio	17.6 to 1	18 to 1	17.62 to 1	16 to 1	16 to 1
U.S. News Student to Staff Ratio	18 to 1	18 to 1	18 to 1	18 to 1	16 to 1
Common Data Set International Students	2696	2633	2476	2609	2802
Common Data Set International Staff		17	19	35	
NCES International Staff			118		51
U.S. News Peer Review	4.3	4.3	4.3	4.3	4.2

Table 17 Triangulating Data for UCLA

FTE: Full Time Equivalent

UCSD in the ARWU

UCSD	2003	2004	2005	2006	2007	2008
Rankings	14	13	13	13	14	14
Total Scores	54.4	53.8	51	50.5	50.4	50.3
Alumni Winners		17.8	17.7	17.1	16.6	15.8
Rankings by Alumni		60	62	64	66	66
Staff Winners	14.2	34.7	34.7	34	34	34
Rankings by Staff Winners	66	20	21	21	23	23
HiCi	58	63.6	59.8	59.6	59.3	59.7
Rankings by HiCi	8	8	8	6	6	7
N&S	59.8	59.4	56.5	54.8	55.5	53
Rankings by N&S	7	7	9	8	8	8
Pulication	67.5	67.2	64.5	65.6	64.6	66.7
Rankings by Publication	18	18	22	19	20	15
Size	55.2	47.9	46.6	47.1	46.6	47.4
Rankings by Size	8	11	11	11	11	11

Table 18 UCSD Yearly Changes in the ARWU

HiCi: Highly Cited Researchers; N&S: the Articles Published in Nature and Science

As illustrated in Table 18, UCSD moves one place up in 2004 ARWU. Scanning all indicators, the only increasing indicator is Staff Prize Winners from 66th to 20th place. Size drops three places, and the other indicators remain the same. The increase in Staff Prize Winners corresponds to Clive W.J. Granger, a UCSD staff member, winning half the Nobel Prize in Economics in 2003. UCSD experiences a one-place

drop in 2007. The only indicators moving down in ranking are Alumni Prize Winners and Staff Prize Winners, drops that are likely caused by other institutions overtaking them in these indicators. In this case, if UCSD is compared with UCLA, UCLA overtakes UCSD because UCLA has a Field Prize winner. Thus, the changes in UCSD's rank are mainly caused by Staff Prize Winners.

UCSD in the WUR

UCSD	2004	2005	2006	2007	2008
Rankings	24	42	44	58	58
Total Scores	331.5	41.9	47.5	76.3	76.3
Peer Review	96	43	46	98	98
Rankings by Peer Review	55	36	34	27	27
Employer Review		0	16	39	58
Rankings by Employer Review		107	105	118	103
Student to Staff Ratio	7	10	26	51	35
Rankings by Ratio	112	94	90	91	111
Citation	208	43	42	95	100
Rankings by Citation	4	6	6	8	1
International Staff	3	3	3	23	20
Rankings by Inter Staff	116	119	117	115	127
International Students	6	9	9	30	29
Rankings by Inter Students	110	111	109	114	118

Table 19 UCSD Yearly Changes in the WUR

As shown in Table 19, USCD shows significant drops in the WUR from 2004 to 2008. Still, no indicator decreases continuously over the five years. By contrast, Peer Review increases continuously from 55th to 27th place. The consistent decreases in

USCD's ranking drop each year are difficult to understand when the ranking movement of the indicators do not experience the same pattern.

When verifying the values for the indicators by data from different sources (see Table 20), the data from other sources are found to be quite stable. However, some data are missing from one source. If the data for rankings were from only one source, these missing data would cause problems for compilers.

UCSD	2003	2004	2005	2006	2007
Common Data Set Full-time Undergraduates	19013	19384	19578	20341	21326
Common Data Set Full-time Graduates	3242	3306	3571	3757	3791
Common Data Set Part-time Undergraduates	859	955	1102	1028	722
Common Data Set Part-time Graduates	111	98	103	103	105
Common Data Set Full-time Instructors					943
NCES Full-time Faculty	1287		2155		2206
UC Office FTE	2701	2805	2960.5	3008.6	3199.1
Common Data Set Student to Staff Ratio					
U.S. News Student to Staff Ratio	19 to1	19 to1	19 to1	19 to1	19 to1
Common Data Set International Students	1424	1367	1406	1433	1530
Common Data Set International Staff					12
NCES International Staff			90		31
U.S. News Peer Review	3.9	3.8	3.9	3.8	3.8

Table 20 Triangulating Data for UCSD

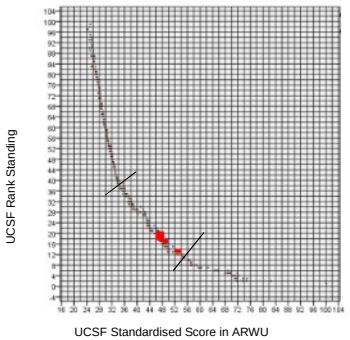
FTE: Full Time Equivalent

University of California-San Francisco (UCSF)

UCSF	2003	2004	2005	2006	2007	2008
ARWU Rankings	13	17	18	18	18	18
WUR Rankings	20	17				

Table 21 Rankings of UCSF

As Table 21 shows, UCSF occupies similar positions in the ARWU and the WUR before 2005. After 2005, UCSF disappears from the WUR.



ARWU

Figure 21 UCSF's Positions in Columns

UCSF in the ARWU

UCSF	2003	2004	2005	2006	2007	2008
Rankings	13	17	18	18	18	18
Total Scores	55.3	50.8	47.8	47.7	46.8	46.6
Alumni Winners		0	0		0	0
Rankings by Alumni		80	80	80	81	81
Staff Winners	41.6	37.6	37.6	36.8	36.8	36.8
Rankings by Staff Winners	15	15	15	15	17	17
HiCi	57.1	59.3	55.6	55.5	54	54.1
Rankings by HiCi	11	13	13	14	14	13
N&S	60.1	59.5	57.9	54.8	53.7	51.5
Rankings by N&S	6	6	6	8	10	13
Pulication	60.9	62.9	58.8	61.6	59.8	60.8
Rankings by Publication	29	27	33	30	32	32
Size	39.2	48.8	45.2	48.2	46.7	47.5
Rankings by Size	28	10	13	10	10	10

Table 22 UCSF Yearly Changes in the ARWU

HiCi: Highly Cited Researchers; N&S: the Articles Published in Nature and Science

As Table 22 shows, UCSF shows significant downward movements over the years in the ARWU. It suffers the biggest slip in 2004, dropping from 13th to 17th place, as a result of the methodology change in 2004, when Alumni Prize Winners was added. UCLA, UCSD and UCSF all belong in the second column (see Figure 19, Figure 20, and Figure 21). In 2003, UCSF actually ranks higher than UCLA and UCSD but UCSF is badly affected in 2004 because it has no Alumni Prize Winners. UCLA and UCSD, both with Alumni Prize Winners, are not affected by this methodology change. In 2005, UCSF falls another place because its ranking by Publication decreases from 27th to 33rd and Size drops three places. Afterward, UCSF maintains

the same rank from 2005 to 2008. Therefore, the changes in UCSF's rank are mainly due to ranking methodology changes.

UCSF in the WUR

UCSF	2004	2005	2006	2007	2008
Rankings	20	17			
Total Scores	376.5	54.9			
Peer Review	21	24			
Rankings by Peer Review					
Employer Review		0			
Rankings by Employer Review					
Student to Staff Ratio	39	91			
Rankings by Ratio					
Citation	300	44			
Rankings by Citation					
International Staff	5	4			
Rankings by Inter Staff					
International Students	0	6			
Rankings by Inter Students					

Table 23 UCSF Yearly Changes in the WUR

As Table 23 shows, UCSF ranks highly in the WUR in 2003 and 2004 but then suddenly disappears from the top 500 universities. One possible reason could be that the missing data exclude UCSF from the ranking. Another possible reason could be the compiler decided to exclude UCSF as a university specialised in medicine with only graduate programmes.

The problem caused by missing data is illustrated in Table 24, most of the information was not available from the Common Data Set provided by UCSF. The data from the

NCES are also incomplete. Another strange phenomenon is that UCSF's data from the NCES show dramatic shifts while the data from UC's central office are quite consistent. Again, this leads to doubt about the reliability of data from the NCES.

UCSF	2003	2004	2005	2006	2007
Common Data Set Full-time Undergraduates					
Common Data Set Full-time Graduates					
Common Data Set Part-time Undergraduates					
Common Data Set Part-time Graduates					
Common Data Set Full-time Instructors					
NCES Full-time Faculty	990		2308		2422
UC Office FTE	3019.92	3120.6	3187.03	3280.26	3499.56
Common Data Set Student to Staff Ratio					
U.S. News Student to Staff Ratio					
Common Data Set International Students	119	142	135	133	148
Common Data Set International Staff					
NCES International Staff			258		46
U.S. News Peer Review					

Table 24 Triangulating Data for UCSF

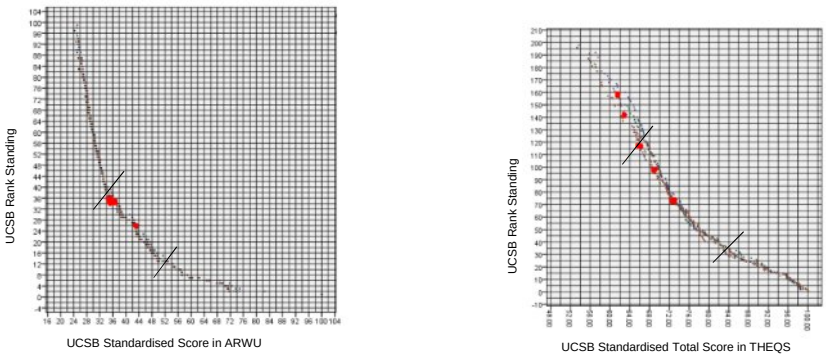
FTE: Full Time Equivalent

University of California - Santa Barbara (UCSB)

UCSB	2003	2004	2005	2006	2007	2008
ARWU Rankings	26	35	34	35	35	36
WUR Rankings		72	159	141	117	98

Table 25 Rankings of UCSB

As Table 25 shows, UCSB suffers a slip in 2004 in the ARWU but its rank remains stable afterward. In the WUR, UCSB drops 87 places in 2005 and shows a rapid recovery afterward.



ARWU WUR

Figure 22 UCSB's Positions in Columns

UCSB in the ARWU

UCSB	2003	2004	2005	2006	2007	2008
Rankings	26	35	34	35	35	36
Total Scores	45.3	37	36.9	36.1	35.8	35.5
Alumni Winners		0	0	0	0	0
Rankings by Alumni		80	80	80	81	81
Staff Winners	32.7	28.5	36	35.3	35.3	35.3
Rankings by Staff Winners	26	32	19	19	21	21
HiCi	47	45.4	42.3	42.1	42.6	41.9
Rankings by HiCi	22	28	28	28	27	29
N&S	40.3	41.4	39	37	36.2	34.9
Rankings by N&S	29	26	27	30	32	33
Pulication	42.5	44	44.1	43.7	42.7	43.2
Rankings by Publication	77	77	75	76	77	78
Size	49.4	36.2	35.8	35.7	35.1	35.4
Rankings by Size	14	26	25	25	27	27

Table 26 UCSB Yearly Changes in the ARWU

HiCi: Highly Cited Researchers; N&S: the Articles Published in Nature and Science

As a university in the second column in the ARWU (see Figure 22), UCSB is disadvantaged by a lack of Alumni Prize Winners in 2004, as is UCSF. Therefore, its rank falls from 26th to 35th place in 2004. In 2005, UCSB gains one place, largely because of a rise in Staff Prize Winners, from 32nd to 19th place, which matches the real-life situation. In 2004, David J. Gross, a UCSB staff member, won a third of the Nobel Prize in Physics, and Finn E. Kydland, also a UCSB staff member, won half of the Nobel Prize in Economics. In the following years, due to small falls in ranking by Articles Published in Nature and Science, Publication, Highly Cited Researchers, and

Alumni Prize Winners and Staff Prize Winners, UCSB drops to 35th place in 2006 and 2007, 36th place in 2008.

Since no Alumni and Staff were awarded any prizes from 2005 to 2007, the drops in rank of the two indicators are caused by being overtaken by other institutions. The scores for Highly Cited Researchers and Publication are compared with their counterparts in the PRSPWU in Table 27. Although the scores differ between the two ranking systems, the two indicators are quite stable in both the ARWU and the PRSPWU. However, it is impossible to determine which one reflects reality.

UCSB	ARWU		PRSPWU	
	HiCi	Publication	HiCi	Publication
2007	42.6	42.7	15.85	18.18
2008	41.9	43.2	15.74	18.09

Table 27 UCSB's scores in the ARWU and the PRSPWU (HiCi: Highly Cited Researchers)

Thus, the rise of UCSB's rank is attributed to an increase in Staff Prize Winners, whereas its the drop in ranking could be caused by different indicators in the ARWU.

UCSB in the WUR

UCSB	2004	2005	2006	2007	2008
Rankings	72	159	141	117	98
Total Scores	180.6	24.4	32.9	65.5	68.8
Peer Review	64	24	31	88	88
Rankings by Peer Review	81	102	90	51	61
Employer Review		0	11	31	50
Rankings by Employer Review		107	113	122	114
Student to Staff Ratio	6	6	22	25	21
Rankings by Ratio	118	119	97	120	123
Citation	93	23	24	89	99
Rankings by Citation	16	12	17	32	11
International Staff	9	10	7	43	39
Rankings by Inter Staff	96	96	106	79	88
International Students	3	6	8	21	22
Rankings by Inter Students	118	118	116	124	124

Table 28 UCSB Yearly Changes in the WUR

In the WUR, UCSB's rank drops from 72nd to 159th place in 2005. The major cause of this fall is the Peer Review. UCSB's ranking by Peer Review falls from 81st to 102nd place. However, in U.S. News, it remains the same (see Table 29). UCSB's big drop in the WUR's Peer Review may be caused by the immature survey methodology given that, in the following years, UCSB bounces back in the rankings by Peer Review.

From 2005, UCSB's rank rises continuously. However, none of the rankings by indicators demonstrate the same pattern. This phenomenon is hard to explain, but one likely explanation could be changes in the real weightings, identified by the multiple regression analysis.

As illustrated in Table 29, missing data could cause a problem when conducting the rankings. Again, large shifts in the data from the NCES cause concern.

UCSB	2003	2004	2005	2006	2007
Common Data Set Full-time Undergraduates	17210	17533	17436	17709	17972
Common Data Set Full-time Graduates	2956	2872	2921	2851	2963
Common Data Set Part-time Undergraduates	652	599	650	509	457
Common Data Set Part-time Graduates	29	22	9	13	18
Common Data Set Full-time Instructors					
NCES Full-time Faculty	907	923	1227		1142
UC Office FTE	1493.07	1492.33	1504.08	1507.79	1534.68
Common Data Set Student to Staff Ratio					
U.S. News Student to Staff Ratio	17 to 1	17 to 1	17 to 1	17 to 1	17 to 1
Common Data Set International Students	841	832	804	744	774
Common Data Set International Staff					
NCES International Staff		40	119		82
U.S. News Peer Review	3.5	3.5	3.6	3.5	3.5

Table 29 Triangulating Data for UCSB

FTE: Full Time Equivalent

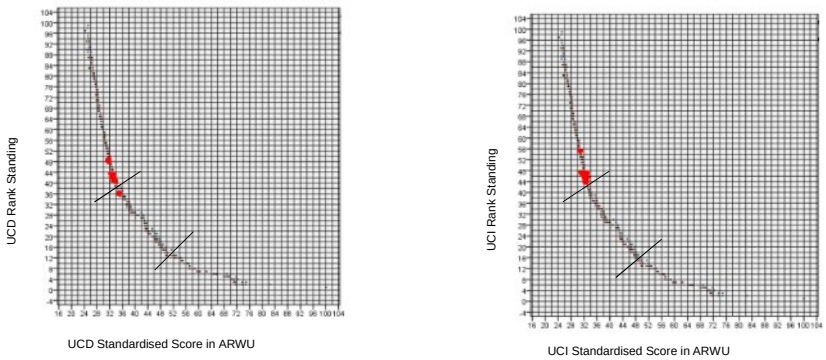
University of California – Davis (UCD) and University of California – Irvine (UCI)

UCD	2003	2004	2005	2006	2007	2008
ARWU Rankings	36	42	41	42	43	48
WUR Rankings		182	206	170	96	89

Table 30 Rankings of UCD

UCI	2003	2004	2005	2006	2007	2008
ARWU Rankings	44	55	47	44	45	46
WUR Rankings			253	198	140	132

Table 31 Rankings of UCI



UCD ARWU UCI ARWU

Figure 23 UCD and UCI's Positions in Columns

As Table 30 shows, UCD's rank ranges from 36th to 48th place in the ARWU and from 89th to 206th place in the WUR. As Table 31 shows, UCI's rank ranges from 44th to 55th place in the ARWU and from 253rd to 132nd place in the WUR. Figure 23 shows UCD and UCI are both universities in the third column in the ARWU. As such, the range of their ranking is around 10 places, while universities in the first column or the top part of the second column normally move within five places. However, UCD's rank shifts 117 places and UCI's rank standings shift 121 places in the WUR. Compared with the WUR, the rankings in the ARWU are still relatively stable. The dramatic shifts in the WUR cast doubt on the reliability of the ranking system.

UCD in the ARWU

UCD	2003	2004	2005	2006	2007	2008
Rankings	36	42	41	42	43	48
Total Scores	38.1	33.6	32.9	32.7	32.7	32.2
Alumni Winners		0	0	0	0	0
Rankings by Alumni		80	80	80	81	81
Staff Winners	0	0	0	0	0	0
Rankings by Staff Winners	70	70	69	71	71	74
HiCi	52.3	48.6	46.5	47.5	46.9	46.7
Rankings by HiCi	17	23	23	23	20	21
N&S	34.6	37.2	34.5	33.3	33.1	31.7
Rankings by N&S	39	34	37	39	39	39
Pulication	65.1	64.7	64	63.3	64.2	63.2
Rankings by Publication	24	24	25	24	21	25
Size	26.4	20.7	29.8	30.1	30	30
Rankings by Size	62	64	44	47	45	47

Table 32 UCD Yearly Changes in the ARWU

HiCi: Highly Cited Researchers; N&S: the Articles Published in Nature and Science

UCD suffers its first slip in the 2004 ARWU because of the same ranking methodology change affecting UCSF and UCSB, the addition of Alumni Prize Winners to the ranking. In 2005, UCD moves up one place, which could result from the rise of Size, from 64th to 44th place. This rise of Size could result from the fall of FTE in 2004. As shown in Table 34, FTE decreased from 3,723.07 in 2003 to 3,682.99 in 2004, according to UC's central office. However, data from the Common Data Set shows that the number of full-time instructors increased from 1545 in 2003 to 1,621 in 2004. Data from different sources could cause different ranking results. In 2006, UCD drops another place because of falls in the Staff Prize Winners, Articles

Published in Nature and Science and Size. In 2007, UCD falls one more place, while Highly Cited Researchers, Publication and Size rise two or three places. The only falling indicator is Alumni Prize Winners. In 2008, UCD drops five places because Staff Prize Winners drops three places, Highly Cited Researchers one place, Publication four places and Size two places. Thus, the changes in the ranking of UCD, a university without alumni or staff prizewinners, are also dominated by the two indicators. In other words, being overtaken by universities with alumni and staff prizewinners affects the universities with none.

UCD in the WUR

UCD	2004	2005	2006	2007	2008
Rankings	182	206	170	96	89
Total Scores	110.8		30.3	69.1	69.9
Peer Review	27		30	83	84
Rankings by Peer Review					
Employer Review			0	34	46
Rankings by Employer Review					
Student to Staff Ratio	10		30	60	42
Rankings by Ratio					
Citation	65		17	82	96
Rankings by Citation					
International Staff	1		2	31	28
Rankings by Inter Staff					
International Students	4		8	26	26
Rankings by Inter Students					

Table 33 UCD Yearly Changes in the WUR

Due to the missing data in the WUR, UCD's rankings by indicators cannot be produced. Consequently, the comparison of the yearly changes in indicators cannot be achieved. However, the ranking of UCD can be used for comparison, which shifts dramatically over the years (see Table 33). The data available for verification are the counterparts of Peer Review in U.S. News; the counterparts of Student to Staff Ratio in the Common Data Set and U.S. News; and the raw numbers of students, staff, international students and staff from the Common Data Set, UC's central office, and

the NCES. These data are listed in Table 34. These numbers are quite stable over the years and do not correspond to the dramatic shifts in the UCD's ranking in the WUR.

UCD	2003	2004	2005	2006	2007
Common Data Set Full-time Undergraduates	20999	20933	20674	21357	21373
Common Data Set Full-time Graduates	4176	4234	4345	4355	4400
Common Data Set Part-time Undergraduates	2510	2238	2061	2101	2126
Common Data Set Part-time Graduates	387	455	487	556	586
Common Data Set Full-time Instructors	1545	1621	1610	1595	1586
NCES Full-time Faculty	1968		2964		2772
UC Office FTE	3723.07	3682.99	3685.22	3668.22	3649.67
Common Data Set Student to Staff Ratio	18.9 to 1	19 to 1	19 to 1	19 to 1	19 to 1
U.S. News Student to Staff Ratio	19 to 1	19 to 1	19 to 1	19 to 1	19 to 1
Common Data Set International Students	1375	1442	1368	1418	1491
Common Data Set International Staff	18	24	31	36	30
NCES International Staff			193		105
U.S. News Peer Review	3.8	3.8	3.8	3.7	3.8

Table 34 Triangulating Data for UCD

FTE: Full Time Equivalent

UCI in the ARWU

UCI	2003	2004	2005	2006	2007	2008
Rankings	44	55	47	44	45	46
Total Scores	35.9	31.4	31.8	32.6	32.5	32.8
Alumni Winners		0	0	0	0	0
Rankings by Alumni		80	80	80	81	81
Staff Winners	27.6	25	30	29.4	29.4	29.4
Rankings by Staff Winners	34	39	32	33	33	33
HiCi	29	33.8	32.4	35.3	35.5	37.2
Rankings by HiCi	49	48	49	46	46	43
N&S	27.6	29.6	28.5	28.9	28	27.5
Rankings by N&S	56	51	52	48	50	50
Pulication	45	47.2	48.2	49	48.9	49.3
Rankings by Publication	74	68	58	58	57	61
Size	38.8	29.9	31.1	32.4	32.1	32.9
Rankings by Size	30	43	33	37	36	34

Table 35 UCI Yearly Changes in the ARWU

HiCi: Highly Cited Researchers; N&S: the Articles Published in Nature and Science

In the 2004 ARWU, UCI is disadvantaged by the same methodology change that affects UCD, UCSF and UCSB. Without Alumni Prize Winners, UCI's rank drops 11 places. However, in the 2005 ARWU, UCI moves up eight places from 55th to 47th place, as the rankings by Staff Prize Winners is up from 39th to 32nd place, Publication from 68th to 58th place and Size from 43rd to 33rd place. The rise of Staff Prize Winners matches the real-life situation. In 2004, Irwin Rose, a UCI staff member, won a third of the Nobel Prize in Chemistry. The higher ranking by Size

may result from changes in FTE. However, as Table 37 shows, the numbers of FTE in 2003 and 2004 are quite similar according to UC's central office, at 2,237.08 and 2,242.81, respectively. In the 2006 ARWU, UCI gains another three places. This is attributed to rises in Highly Cited Researchers and Articles Published in Nature and Science, despite the drops in Staff Prize Winners and Size. In the 2007 ARWU, UCI falls one place, owing to the drops in Alumni Prize Winners and Articles Published in Nature and Science, despite raises in Publication and Size. In the 2008 ARWU, UCI drops one more place because its ranking by Publication falls even when the rankings by Highly Cited Researchers and Size increase. Thus, when there is an increase in Staff Prize Winners, UCI's rank is lifted eight places. However, with no increases in Alumni Prize Winners or Staff Prize Winners, the movement of UCI's rank does not follow a particular pattern.

UCD ranks higher than UCI from 2003 to 2007 in the ARWU. UCD always ranks higher than UCI by Highly Cited Researchers, Articles Published in Nature and Science and Publication from 2003 to 2008. UCI always ranks higher than UCD by Staff Prize Winners from 2003 to 2008. Yet, in 2008, UCI suddenly ranks two places higher than UCD. This is hard to decode.

UCI in the WUR

UCI	2004	2005	2006	2007	2008
Rankings		253	198	140	132
Total Scores			28		63.8
Peer Review			24		80
Rankings by Peer Review					
Employer Review			16		30
Rankings by Employer Review					
Student to Staff Ratio			19		30
Rankings by Ratio					
Citation			21		94
Rankings by Citation					
International Staff			2		24
Rankings by Inter Staff					
International Students			10		29
Rankings by Inter Students					

Table 36 UCI Yearly Changes in the WUR

Due to the missing data in the WUR, UCI's rankings by indicators cannot be produced. Consequently, the comparison of the yearly changes of indicators cannot be achieved. However, when comparing UCI's dramatic shifts in rank (see Table 36) with the data from different sources (see Table 37), it is obvious the huge shifts in rank are not supported by the data from different sources.

UCI	2003	2004	2005	2006	2007
Common Data Set Full-time Undergraduates	19351	19276	19464	20192	21131
Common Data Set Full-time Graduates	3495	3693	3678	3713	3943
Common Data Set Part-time Undergraduates	782	718	597	630	602
Common Data Set Part-time Graduates	114	123	106	86	85
Common Data Set Full-time Instructors			1038	1419	1100
NCES Full-time Faculty	1340		1987		1943
UC Office FTE	2237.08	2242.81	2343.46	2447.51	2498.77
Common Data Set Student to Staff Ratio			17 to 1	19.2 to 1	19.2 to 1
U.S. News Student to Staff Ratio	18 to 1	18 to 1	19 to 1	17 to 1	19 to 1
Common Data Set International Students	1404	1533	1516	1439	1545
Common Data Set International Staff			26	4	15
NCES International Staff			108		46
U.S. News Peer Review	3.6	3.6	3.7	3.6	3.6

Table 37 Triangulating Data for UCI

FTE: Full Time Equivalent

University of California – Santa Cruz (UCSC) and University of California Riverside (UCR)

UCSC	2003	2004	2005	2006	2007	2008
ARWU Rankings	102-151	101-152	101-152	102-150	102-150	101-151
WUR Rankings		175	326	311	259	235

Table 38 Rankings of UCSC

UCR	2003	2004	2005	2006	2007	2008
ARWU Rankings	88	101-152	101-152	102-150	102-150	101-151
WUR Rankings			491	412	298	262

Table 39 Rankings of UCR

As Table 38 and Table 39 show, UCSC and UCR are in the same tier from 2003 to 2008 in the ARWU, while in the WUR the two universities show dramatic changes in rank: UCSC shifts 151 places and UCR shifts 229 places.

UCSC in the ARWU

UCSC	2003	2004	2005	2006	2007	2008
Rankings	102-151	101-152	101-152	102-150	102-150	101-151
Total Scores						
Alumni Winners		0	0	0	0	0
Rankings by Alumni						
Staff Winners	0	0	0	0	0	0
Rankings by Staff Winners						
HiCi	22.9	29	28.3	27.7	26.7	27.3
Rankings by HiCi						
N&S	26.9	28.6	28.5	29.2	29.3	30.7
Rankings by N&S						
Pulication	30	29.7	31.1	29.5	29.3	30.7
Rankings by Publication						
Size	39.9	33.4	29.2	29.2	28.7	30.1
Rankings by Size						

Table 40 UCSC Yearly Changes in the ARWU

HiCi: Highly Cited Researchers; N&S: the Articles Published in Nature and Science

Due to the missing data in the ARWU, UCSC is excluded from the statistical analysis. Thus, UCSC's rankings by indicators cannot be produced. Consequently, the

comparison of the yearly changes of indicators cannot be achieved. The scores of indicators published in the ARWU (see Table 40) are unable to provide any information for sensible judgement.

UCSC in the WUR

UCSC	2004	2005	2006	2007	2008
Rankings	175	326	311	259	235
Total Scores	115.6				
Peer Review	14				
Rankings by Peer Review					
Employer Review					
Rankings by Employer Review					
Student to Staff Ratio	4				
Rankings by Ratio					
Citation	87				
Rankings by Citation					
International Staff	5				
Rankings by Inter Staff					
International Students	2				
Rankings by Inter Students					

Table 41 UCSC Yearly Changes in the WUR

To verify UCSC's dramatic shifts in rank in the WUR (see Table 41), the data from different sources are listed in Table 42. The table shows that the shifts in the WUR cannot be verified by these data.

UCSC	2003	2004	2005	2006	2007
Common Data Set Full-time Undergraduates	13002	12983	13139	13487	13909
Common Data Set Full-time Graduates	1275	1342	1387	1323	1333
Common Data Set Part-time Undergraduates	658	711	486	474	494
Common Data Set Part-time Graduates	62	80	75	80	89
Common Data Set Full-time Instructors	533	537	548	554	566
NCES Full-time Faculty	518	541	773		754
UC Office FTE	949.37	936.43	969.77	1029.96	1058.26
Common Data Set Student to Staff Ratio	19.38 to 1	19.37 to 1	19.34 to 1	19.34 to 1	19.3 to 1
U.S. News Student to Staff Ratio				19 to 1	19 to 1
Common Data Set International Students	355	420	325	292	308
Common Data Set International Staff	15	17	13	16	13
NCES International Staff	518	541	773		754
U.S. News Peer Review		3.2		3.2	3.2

Table 42 Triangulating Data for UCSC

FTE: Full Time Equivalent

UCR in the ARWU

UCR	2003	2004	2005	2006	2007	2008
Rankings	88	101-152	101-152	102-150	102-150	101-151
Total Scores	27.3					
Alumni Winners		0	0	16	15.6	14.8
Rankings by Alumni						
Staff Winners	0	0	0	0	0	0
Rankings by Staff Winners						
HiCi	30.8	31.5	28.3	27.7	27.7	27.3
Rankings by HiCi						
N&S	26.2	25.1	25.9	25.3	27.3	25.7
Rankings by N&S						
Pulication	34.8	35.4	36.5	36.5	35.8	35.9
Rankings by Publication						
Size	36	27.7	27	27.8	27.8	27.6
Rankings by Size						

Table 43 UCR Yearly Changes in the ARWU

HiCi: Highly Cited Researchers; N&S: the Articles Published in Nature and Science

Due to the missing data in the ARWU, UCR is excluded from the statistical analysis. Thus, UCSR's rankings by indicators cannot be produced. Consequently, the comparison of the yearly changes of indicators cannot be achieved. The scores of indicators published in the ARWU (see Table 43) are unable to provide any information for sensible judgement.

Still, one point that can be drawn from this case is that one-third of an Alumni Prize Winner is not enough to help UCR move out of this tier. In 2005 Richard R. Schrock,

a UCR alumnus, won a third of the Nobel Prize in Chemistry. Although UCR gains 16 points for this, its rank remains the same.

UCR in the WUR

UCR	2004	2005	2006	2007	2008
Rankings		491	412	298	262
Total Scores					
Peer Review					
Rankings by Peer Review					
Employer Review					
Rankings by Employer Review					
Student to Staff Ratio					
Rankings by Ratio					
Citation					
Rankings by Citation					
International Staff					
Rankings by Inter Staff					
International Students					
Rankings by Inter Students					

Table 44 UCR Yearly Changes in the WUR

To verify the dramatic shifts in the ranking of UCR in the WUR (see Table 44), the data from different sources are listed in Table 45. Clearly, the shifts in the WUR cannot be verified by these data.

UCR	2003	2004	2005	2006	2007
Common Data Set Full-time Undergraduates	14018	14068	13822	14334	14603
Common Data Set Full-time Graduates	1826	1858	1898	1952	2068
Common Data Set Part-time Undergraduates	1381	1106	827	526	438
Common Data Set Part-time Graduates	22	21	26	14	29
Common Data Set Full-time Instructors					
NCES Full-time Faculty	685	667	984		931
UC Office FTE	1137.36	1141.06	1154.46	1154.37	1210.31
Common Data Set Student to Staff Ratio					
U.S. News Student to Staff Ratio			18 to 1	18 to 1	
Common Data Set International Students	922	910	879	860	909
Common Data Set International Staff					
NCES International Staff		22	73		35
U.S. News Peer Review		3.2	3.1	3.1	3.1

Table 45 Triangulating Data for UCR

FTE: Full Time Equivalent

Summary of the ARWU

Generally, the rankings in the ARWU are stable. From the nine cases analysed above, we can conclude that Alumni Prize Winners and Staff Prize Winners are the major contributors of yearly changes in the ARWU. This is illustrated the findings from the multiple regression and principal component analysis. Alumni Prize Winners and

Staff Prize Winners account for around 77% of variance, and the first component related to the two indicators explains around 67% of variance. For universities in the first column, the changes in rank are not sensitive to the win of only a half or a third of a prize. To initiate a change of rank, more prize inners are needed. For universities in the second column, the win of a half or a third of prize can result in one-place rise in rank. For universities in the third column, the win of a half or a third of prize can cause several shifts in rank. The other major factor influencing rank is methodology change, such as the addition of Alumni Prize Winners as an indicator. This change puts the universities without Alumni Prize Winners at a disadvantage.

Summary of the WUR

For the WUR, the unpredictable changes in rankings and incomprehensible relationship between indicators are due to frequent changes in methodology and multiple components combined in a single ranking.

Before 2007, the data are non-standardised. The alleged weightings of indicators are undermined by the different scales of each indicator. The most striking changes unveiled by multiple regression analysis are found in Peer Review and Student to Staff Ratio. The weighting of Peer Review is reduced from 89% in 2004 to 83% in 2005, 76% in 2006, 62% in 2007 and 55% in 2008, even though it does not fit the

claimed 40% weight. The weighting of Student to Staff Ratio is increased from 7% in 2004 to 6% in 2005, 11% in 2006, 19% in 2007 and 21% in 2008.

Another issue is multiple components in a single ranking. In 2007, the principal component analysis identifies three components. In the rest of the years, the eigenvalues of Student to Staff Ratio are very close to one, the cutting point. If the cutting point is lowered to .9, Student to Staff Ratio would be identified as the third component. Another point supporting the separation of Student to Staff Ratio as a third component is that Student to Staff Ratio contributes to neither the first nor the second component. Thus, the first component explains 33.6% of variance, the second 22.3%, and the third 17.6%, according to the results of 2007. Changes in any of the components will result in changes in the rankings. Therefore, it is difficult for the rankings to remain stable. Furthermore, in the second component, the relationship between International Student Number and International Staff Number and Citation are negative. When the loadings of International Student Number and International Staff Number are positive, the loading of Citation is negative, and vice versa. Thus, it is not sensible to combine two conflicting indicators into a single scale.

Another serious problem with the WUR arises from the scores of indicators not being verified by data from other sources. Verification of the scores is vital to the validity and reliability of the WUR.

Chapter Six: Discussion

The ARWU and the WUR's Choice of Indicators

Size of Institution

In the ARWU, Size is a redundant indicator. As mentioned previously, Size provides no new information. In addition, it cannot fulfil the purpose for it was designed. If designed to eliminate the disadvantages associated with being a small institution, all indicators should take into account the size of the institution. The 10% weighting is unable to counteract the impact of large institutions on the other five indicators. Furthermore, in countries for which the number of staff was obtained, universities are weighted by their size, whereas universities in other countries are exempted from this weighting. Although the impact of Size is not great, it may introduce new bias to the ranking.

Citation vs. International Student Number and International Staff Number

As shown by the principal component analysis of the current WUR, Citation and International Student Number and International Staff Number measure the second component in opposite ways. However, it is not clear whether the contradiction is caused by the use of incorrect data or if it represents a true contradictory relationship between the two indicators. As mentioned previously, the scores of these indicators in the WUR are not verified by data from different sources. If the two sets of indicators

are contradictory in nature, they should not be combined into the same measure to rank universities.

The indicators of International Students Number and International Staff Number are designed to measure the international diversity of a university. However, the validity of the two indicators is problematic, as they mistake the real diversity in the student and staff communities with the size of the international students and staff bodies. For example, a university may have 90% percent of students and staff coming from a single foreign county. Although such a university will score high on these indicators, it fails to demonstrate diversity in the student and staff. To measure diversity, an indicator should take into account the percentage of students and staff from each country.

Peer Review and Employer Review

Peer Review and Employer Review are the highlights of the WUR. The results of a well conducted and mature reputation survey should be stable. The peer review results from U.S. News with a 26-year history become consistent from 2003 to 2008. After the 2007 standardisation of the WUR, almost all the nine universities' peer review results remained the same from 2007 to 2008. This indicates a reliable peer survey should produce consistent results. However, Employer Review is still unstable, even after 2007. This may be attributed to the sample number of the respondents. The number of peer review respondents was 3,703 in 2006, 3,069 in 2007, and 6,354 in 2008. The 2008 data include respondents who did not update their responses in 2006

and 2007. But the Employer Review respondents' number is 2,339 in 2008, including those who have not updated their responses in 2006 and 2007. This relatively small number of world employers is unable to provide consistent results.

Another problem with Peer Review and Employer Review is the over representation of certain countries. Peer review respondents from the U.S., the UK, Australia, Italy, Canada and India dominate around half of the responses, even though the respondents actually come from about 100 different countries. For the employer review, respondents from the U.S., the UK and Australia again dominate the results. This may introduce bias to certain countries.

A third problem relates to how to make sure respondents are informative and able to provide creditable answers. For example, most of the respondents of the employer review come from three areas: financial service/banking, consulting/professional services and manufacturing/engineering. Further, the majority of the respondents may lack adequate knowledge in other fields, which will introduce bias toward the institutions specialised in these particular areas.

Validity

Validity of the ARWU's

As its name indicates, the ARWU was designed to rank universities by their academic capacity. In fact, it measures universities by Staff Reputation, Staff Productivity and Student Post-achievements. As mentioned in my new classification of indicators, Staff Reputation refers to world-famous awards or membership of prestigious associations; Staff Productivity denotes Publication, Citation and Student Post-achievement refers to student achievements after graduation. The ARWU measures the most prestigious awards, publication and citation, because they comprise the available international comparable data. However, the academic features of universities are not confined to staff reputation and staff productivity and student post-achievements. Instead, they include other indicators listed in the new classification, such as Staff Teaching, Student Learning Process, Programme/University's Teaching and Learning. However, considering the data availability and international comparability, the indicators in the ARWU are respected indicators for staff reputation and staff productivity, but not representative of student post-achievements.

Validity of the WUR's

The WUR aims to identify world-class universities with consideration of the multi-faceted nature of universities. The ranking system attempts to measure four aspects from the new classification: institutional reputation, staff productivity, staff

teaching, and environment of programmes/university. Institutional reputation is measured by peer and employer surveys. If the surveys are designed and conducted carefully, they act as good indicators of institutional reputation. Another concern associated with the validity of reputational surveys is how informative the respondents actually are and whether their responses reflect the reputation of the universities. The correlation between the results of the peer review in the WUR and the overall rankings of the ARWU for the previous year are slightly higher than the correlation between the overall rankings of the WUR and the ARWU. This finding may suggest respondents consulted the results of the ARWU before answering the questionnaire. Staff Teaching is measured by Student to Staff ratio, the easiest obtainable data. Although Staff Teaching refers to more than Student to Staff Ratio, the availability of data prevents the use of other staff teaching indicators. Diversity, part of Programmes/University Environment, was designed to be measured by International Student Number and International Staff Number. As mentioned above, instead of diversity, these two indicators actually measure the size of the international student and staff bodies, which is not necessarily related to an institution's diversity. Consequently, the indicators' design of the indicators casts doubt on the validity of the WUR.

Reliability

The reliability of both the ARWU and the WUR causes great concern. For the WUR, one of the most serious problems is caused by mistakes in the data. For example, the

scores of Student to Staff Ratio, International Student Number and International Staff Number fail to match the raw data from different sources.

The bibilometric indicators may seem objective, but these are also affected by human judgement. Florian (2007) tried to reproduce the same ranking of Romanian universities by imitating the method described by the ARWU, but concluded it was impossible to replicate the results. Van Raan (2005) warns of the problem of misidentifying authors and institutions, which will create different results even when using the same database and method. This is evidenced by the discrepancies between the results for the ARWU and the PRSPWU, which use the same database and method but show different values for the same indicator.

The reliability-related concern of Peer Review and Employer Review lies in the sample sizes of the respondents, which are too small to produce a representative sample.

Another serious concern of reliability is related to changing methodology. The ARWU and the WUR's compilers openly admit to making yearly changes to the ranking methodology. The changes in the ARUW are tolerable, whereas the changes in the WUR totally distort the results of the ranking. The largest impact comes from the introduction of standardisation in 2007 to the WUR, as the non-standardised data violate the weightings. All of these changes result in the scores of indicators being

incomparable across the years and making it hard to interpret what contributes to the changes in the rankings. For the ARWU, the changes officially stopped in 2006, but some minor methodology changes are sill in progress after 2006. This is evidenced by changes in the scores of Alumni Prize Winners and Staff Prize Winners when no alumni or staff receive awards during the period.

Questions for Future Research

Before solving the problems underlying the ARWU and the WUR, two fundamental questions deserve our consideration.

What are the indicators of quality?

The ARWU measures Staff Reputation and Staff Productivity and Student Post-achievements, while the WUR attempts to measure some aspects of Staff Productivity, Staff Teaching, Institutional Reputation, and Environment, although it does not fulfil its purpose. Still, the academic capacity or quality of a world-class university is greater than these characteristics. As my new classification suggests, they can include Student Pre-achievements, Student Learning Process, Student Post-achievements, Staff Reputation, Admission, Teaching and Learning, Student Support, Staff Support and Revenue. However, which should be chosen as indicators for ranking, how to obtain access to the data and how to ensure international comparability are crucial questions that need to be addressed first.

What is the appropriate ranking method?

As shown by the case study, the results created by the weight-and-sum method are not stable. The rankings of universities in the third column or those who ranked even lower are easily affected by small changes in the ranking methodology rather than real changes within the universities. Whether universities should be divided into tiers or ranked by other statistical methods, such as latent method (Guarino et al., 2005), clustering analysis, regression (Valadkhani and Worthington, 2006), data envelopment analysis (Bonaccorsi, Daraio and Simar, 2006;Simar and Wilson, 2007), or archetypal analysis (Porzio, Ragozini and Vistocco, 2008) remains a question to be addressed.

Chapter Seven: Conclusions

To explore the factors influencing the stability and instability of international rankings, a longitudinal study of two influential international ranking systems, the ARWU and the WUR, is conducted. The ARWU contains six years of ranking data and the WUR five years of data. To study the changes in the universities' rankings over the years, a statistical analysis is conducted first, followed by multiple case studies.

Statistical Analysis

Relationship between the Indicators

Correlation, multiple regression, and principal component analysis are conducted to determine the relationship between the indicators and their underpinning factors. For the ARWU, the results from all analyses are consistent from 2003 to 2008. The correlation coefficients between indicators are high, which serves as a sign of high reliability. Further, the regression coefficients of the indicators are similar to the published weightings. Alumni Prize Winners and Staff Prize Winners account for most of the variance in the total score. Highly Cited Researchers and Articles Published in Nature and Science are the second largest factors. For most of the years, two components are extracted. The first component is related to Staff Reputation and Student Post-achievements, according to the new classification of indicators. Alumni Prize Winners contributes to Student Post-achievements, and Staff Prize Winners is

the major contributor to Staff Reputation. The second component is associated with Staff Productivity, whose main contributor is Publication. Articles Published in Nature and Science and Highly Cited Researchers also contribute to both components too. However, Articles Published in Nature and Science is inclined toward Staff Reputation while Highly Cited Researchers is inclined toward Staff Productivity.

In terms of the WUR, the results from all analyses are inconsistent from 2004 to 2008. The correlation coefficients between indicators are low, which shows each indicator contributes new information in ranking. However, it also signals a lack of reliability. The regression coefficient indicators change each year and differ from the published weightings, but the situation improves after 2007. Peer Review accounts for most of the variance in the overall score, but its importance decreases throughout the years. At the same time, the importance of Sstudent to Staff Ratio increases. Three components underpin the indicators. The first component is Programme/ University Reputation, measured by Peer Review and Employer Review. The second component is Environment, measured by International Student Number and International Staff Number. The third component is Staff Teaching, measured by Student to Staff Ratio. Citation contributes to both the first and second components, but relationship between Citation and International Student Number and International Staff Number is negative. This means that indicators should not be combined in the same ranking.

Relationship between the Indicators and the Total Scores

To analyse the relationship between the indicators and the total scores, it is necessary to ensure the comparability between the data in different years by standardising the total scores on the basis of equivalent scores.

Plotting the standardised total scores and rankings shows the ARWU data for six years and the WUR data for five years merge into the same patterns, respectively. According to these patterns, universities can be grouped into three columns. In the ARWU, universities in the first column have high scores and are the top- ranked institutions. However, to enable changes to take place in the rankings change, a considerable amount of change in the scores must occur. For example, a half or one-third Nobel Prize winner would not change a first-column university's rank. A greater number of prize winners are needed. For universities in the second column, an increase in score can result in a modest change in ranking. For example, a half or one-third Nobel Prize winner would help a university in the second column move up one place. By contrast, the ranking of universities in the third column can be boosted by a minor change in score. For example, a half or one-third Nobel Prize winner can result in an increase of six places in the ranking. In the WUR, the difference between columns is not as pronounced as in the ARWU.

Case Study

Nine campuses of the University of California are chosen for the multiple case studies. They exist in the same macro-economic and social environments, but they demonstrate different characteristics. Some show significant changes over the years while others do not. The campuses also belong to different columns. They fulfil the criteria of being critical, exemplifying, revelatory and unique. The values for the nine universities in the rankings are verified by data obtained from three different sources or their counterparts in other ranking systems.

The values for Alumni Prize Winners and Staff Prize Winners in the ARWU can reflect the actual prize-winning situation of every university. The two indicators are also the major factors influencing the rankings of universities. The values for Highly Cited Researchers and Publication remain relatively stable over the years. This is verified by their counterparts in another ranking system. Generally, the ranking of the ARWU is quite stable. The ARWU provides a relative reliable ranking of universities' Staff Reputation and Staff Productivity.

For the WUR, the dramatic changes in the values for Students to Staff Ratio, International Student Number and International Staff Number do not match the data obtained from other statistical sources. Similarly, the dramatic changes in the values for Peer Review do not match similar values in another ranking system. This variation causes serious concern about the reliability of the WUR. The two indicators

that measure the size of the international student and staff bodies are unable to fulfil their original purpose of measuring the diversity in the environment of the programme/university. This raises concern about the validity of the WUR. In addition, the ranking methodology of the WUR changes each year. Ultimately, mistakes in the data, inadequate indicator design and the frequent changes in methodology jeopardise the reliability, validity and the credibility of the WUR, which fails to produce a well-rounded ranking of the world's top universities.

Finally, the ranking of universities as a quality assurance method calls for more empirical research on which indicators should be chosen as the appropriate proxies of quality, as well as further research on ranking methodologies to create a more stable ranking. In addition, to promote transparency, the development of a ranking also relies on a greater availability of data at the national and international levels.

References

Altbach, P. (2006) The dilemmas of ranking. *International Higher Education,* 42, pp. 2-3.

Angoff, W.H. (1971) Norms, scales, and equivalent scores. In *Educational Measurement.*, ed. R.L. Thorndike, Second ed. Washington DC: American Council on Education.

Birnbaum, R. (2007) No world-class university left behind. *International Higher Education,* 47, pp. 7-9.

Bonaccorsi, A., Daraio, C. and Simar, L. (2006) Advanced indicators of productivity of universities. An application of robust nonparametric methods to Italian data. *Scientometrics,* 66 (2), pp. 389-410.

Brooks, R.L. (2005) Measuring university quality. *Review of Higher Education,* 29 (1), pp. 1-21.

Bryman, A. (2004) *Social Research Methods.* Second ed. Oxford: Oxford University Press.

Cartter, A.M. (1966) *An Assessment of Quality in Graduate Education.* Washington DC: American Council on Education.

Cattell, J.M.K. (1910) A further statistical study of American men of science. *Science,* 32 (827), pp. 633-648.

Chatterjee, S. and Hadi, A.S. (1988) *Sensitivity Analysis in Linear Regression.* New York: Wiley.

CHE, CEPES and IHEP (2006) *Berlin Principles on Ranking of Higher Education Institutions.* CHE & CEPES & IHEP. [Online]. Available from: <http://www.che.de/downloads/Berlin_Principles_IREG_534.pdf> (Accessed Nov/5, 2008).

Clarke, M. (2002) Some guidelines for academic quality rankings. *Higher Education in Europe,* 27 (4), pp. 443-459.

Coates, H. (2009) What's the difference? A model for measuring the value added by higher education in Australia. *Higher Education Management and Policy,* 21 (1), pp. 1-19.

Corti, L. (1993) Using diaries in social research. *Social Research Update,* 2, pp. 1-4.

Cramer, K.M. and Page, S. (2007) Calibrating Canadian universities: rankings for sale once again. *Canadian Journal of School Psychology,* 22 (1), pp. 4-13.

Cunningham, S. (2008) University and discipline cluster ranking systems and the humanities, arts, and social sciences. *Higher Education in Europe,* 33 (2/3), pp. 245-258.

Dale, A., Arber, S. and Procter, M. (1988) *Doing Secondary Analysis.* London: Unwin Hyman.

Daly, R., Machung, A. and Roque, G. (2007) Running to Stay in Place: the stability of US News' ranking system. *The 46th Annual Forum of the Association for Institutional Research.*Chicago 2006. Washington D.C.: ERIC, pp. 1-25.

Diamond, N. and Graham, H.D. (2000) How should we rate research universities?. *Change,* 32 (4), pp. 20-33.

Dill, D.D. and Soo, M. (2005) Academic quality, league tables, and public policy: a cross-national analysis of university ranking systems. *Higher Education,* 49 (4), pp. 495-533.

Eccles, C. (2002) The use of university rankings in the United Kingdom. *Higher Education in Europe,* 27 (4), pp. 423-432.

Ehrenberg, R.G. and Hurst, P.J. (1996) The 1995 NRC ratings of doctoral programs: a hedonic model. *Change,* 28 (3), pp. 46-50.

Ehrenberg, R.G. and Hurst, P.J. (1998) The 1995 ratings of doctoral programs: a hedonic model. *Economics of Education Review,* 17 (2), pp. 137-148.

Ehrenberg, R.G. (2005) Method or madness? Inside the USNWR college rankings. *The Advancement of Postsecondary Education Forum on The Use and Abuse of College Rankings.* Madison 2003. New York: Cornell University, pp. 1-21.

Eisenhardt, K.M. (1989) Building theories from case study research. *Academy of Management Review,* 14, pp. 532-550.

Elbeck, M. (2009) Advancing the relationship between business school ranking and student learning. *Journal of Higher Education Policy and Management,* 31 (1), pp. 81-88.

Federkeil, G. (2008) Rankings and quality assurance in higher education. *Higher Education in Europe,* 33 (2), pp. 219-231.

Field, A. (2005) *Discovering statistics using SPSS.* Second ed. London: Sage.

Florian, R.V. (2007) Irreproducibility of the results of the Shanghai academic ranking of world universities. *Scientometrics,* 72 (1), pp. 25-32.

Graham, H.D. and Diamond, N. (1997) *The Rise of American Research Universities: elites and challengers in the postwar era.* Baltimore: Johns Hopkins University Press.

Green, D. (1994) What is quality in higher education? concepts, policy and practice. In *What Is Quality in Higher Education?*, ed. D. Green,Buckingham: SRHE and Open University Press.

Guarino, C., Ridgeway, G., Chun, M. and Buddin, R. (2005) Latent variable analysis: a new approach to university ranking. *Higher Education in Europe,* 30 (2), pp. 147-165.

Gunn, R. and Hill, S. (2008) The impact of league tables on university application rates. *Higher Education Quarterly,* 62 (3), pp. 273-296.

Harvey, L. (2008) Rankings of higher education institutions: a critical review. *Quality in Higher Education,* 14 (3), pp. 187-207.

Hattendorf, L.C. (1996) Educational Rankings of Higher Education: Fact or Fiction? *The International Conference on Assessing Quality in Higher Education.*Queensland 1996. Chicago: The University of Illinois at Chicago, pp. 1-11.

Hazelkorn, E. (2007) The impact of league tables and ranking systems on higher education decision making. *Higher Education Management and Policy,* 19 (2), pp. 1-24.

Hazelkorn, E. (2008) Globalization, internationalization and rankings. *International Higher Education,* 53 (Fall), pp. 8-10.

Hendel, D.D. and Stolz, I. (2008) A comparative analysis of higher education ranking systems in Europe. *Tertiary Education and Management,* 14 (3), pp. 173-189.

Himmel, K.L. (1967) *The Number of Years of High School Science as a Predictor of First Year College Success Compared with Other Predictors, rapport de recherches.* Bozeman: ERIC.

Hughes, R.M. (1925) *A Study of the Graduate Schools of America.* Oxford: Miami University Press.

Jin, G.Z. and Whalley, A. (2007) *The Power of Information: Do Rankings Affect the Public Finance of Higher Education?* Cambridge: NBER Working Papers.

Johnston, R. (2009) Where there are data...quantifying the unquantifiable. *Political Studies Review,* 7 (1), pp. 50-62.

Jones, L.V., Lindzey, G. and Coggeshall, P.E. (1982) *An Assessment of Research-doctorate Programs in the United States.* Washington DC: National Academy Press.

Lazarsfeld, P.F., Pasanella, A.K. and Rosenberg, M. (1972) *Continuities in the Language of Social Research.* New York: The Free Press.

Lindblad, S. (2008) Navigating in the field of university positioning: on international ranking lists, quality indicators and higher education governing. *European Education Research Journal,* 7 (4), pp. 438-450.

Linn, R.L. (1993) *Educational Measurement.* Washington DC: Oryx Press.

Liu, N. and Cheng, Y. (2005) The academic ranking of world universities. *Higher Education in Europe,* 30 (2), pp. 127-136.

Locke, W. (2008) *A 'League Table Culture'? The Policy and Governance Implication of University Rankings in a Hierarchical Higher Education System.* London: SRHE.

Locke, W., Verbik, L., Richardson, J.T.E. and King, R. (2008) *Counting What Is Measured or Measuring What Counts? League Tables and Their Impact On Higher Education Institutions in England.* Bristol: Higher Education Funding Council for England.

Marginson, S. (2007) Global university rankings: implications in general and for Australia. *Journal of Higher Education Policy and Management,* 29 (2), pp. 131-142.

Merisotis, J.P. (2002) Summary report of the invitational roundtable on statistical indicators for the quality assessment of higher/tertiary education institutions: ranking and league table methodologies. *Higher Education in Europe,* 27 (4), pp. 475-480.

Montesinos, P., Carot, J.M., Martinez, J.M. and Mora, F. (2008) Third mission ranking for world class universities: beyond teaching and research. *Higher Education in Europe,* 33 (2), pp. 259-271.

Morrison, H.G., Magennis, S.P. and Carey, L.J. (1995) Performance indicators and league tables: a call for standards. *Higher Education Quarterly,* 49 (2), pp. 128-145.

Pike, G.R. (2004) Measuring quality: a comparison of US News rankings and NSSE benchmarks. *Research in Higher Education,* 45 (2), pp. 193-208.

Porzio, G.C., Ragozini, G. and Vistocco, D. (2008) On the use of archetypes as benchmarks. *Applied Stochastic Models in Business and Industry,* 24 (5), pp. 419-437.

Punch, K.F. (2005) *Introduction to Social Research: quantitative and qualitative approaches.* Second ed. London: SAGE.

QS (2009a) *The Methodology: a simple overview.* QS. [Online]. Available from: <http://www.topuniversities.com/worlduniversityrankings/methodology/simple_overview/> (Accessed June/16, 2009).

QS (2009b) *Purpose and Approach.* QS. [Online]. Available from: <http://www.topuniversities.com/worlduniversityrankings/methodology/purpose_amp_approach/> (Accessed June/16, 2009).

Robson, C. (2002) *Real World Research.* Second ed. Oxford: Blackwell.

Rogers, E. and Rogers, S.J. (1997) High science vs. just selling magazines: how the NRC and U.S. News graduate rankings compete. *AAHE Bulletin,* 49 (9), pp. 7-10.

Salmi, J. (2009) *The Challenge of Establishing World-Class Universities.* Washington DC: The World Bank.

Salmi, J. and Saroyan, A. (2007) League tables as policy instruments: uses and misuses. *Higher Education Management and Policy,* 19 (2), pp. 24-62.

Sheskin, D. (2004) *Handbook of Parametric and Nonparametric Statistical Procedures.* Boca Raton: CRC.

Simar, L. and Wilson, P.W. (2007) Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics,* 136 (1), pp. 31-64.

SJTU (2008) *The Academic Ranking of World Universities.* Shanghai Jiao Tong University. [Online]. Available from: <http://www.arwu.org/rank2008/EN2008.htm> (Accessed August/12, 2009).

Sowter, B. (2009) *Domestic Rankings Slow to Reveal Their Past.* QS. [Online]. Available from:

<http://qsiu.wordpress.com/2009/04/23/domestic-rankings-slow-to-reveal-their-past/> (Accessed August/12, 2009).

Stuart, D.L. (1995) Reputational rankings: background and development. *New Directions for Institutional Research,* (88), pp. 13-20.

Taylor, P. and Braddock, R. (2007) International university ranking systems and the idea of university excellence. *Journal of Higher Education Policy and Management,* 29 (3), pp. 245-260.

Ting, S.M.R. (2003) A ongitudinal tudy of non-cognitive variables in predicting academic success of first-generation college students. *College and University,* 78 (4), pp. 27-31.

Usher, A. and Savino, M. (2006) *A World of Difference: A global survey of university league tables.* Toronto: Educational Policy Institute.

Valadkhani, A. and Worthington, A. (2006) Ranking and clustering Australian university research performance, 1998-2002. *Journal of Higher Education Policy and Management,* 28 (2), pp. 189-210.

van der Wende, M. (2008) Rankings and classifications in higher education: a European perspective. In *Higher Education: handbook of theory and research*, ed. J.C. Smart, New York: Springer.

Van Dyke, N. (2005) Twenty years of university report cards. *Higher Education in Europe,* 30 (2), pp. 103-125.

Van Raan, A.F.J. (2005) Fatal attraction: conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics,* 62 (1), pp. 133-143.

Vaughn, J. (2002) Accreditation, commercial rankings, and new approaches to assessing the quality of university research and education programmes in the United States. *Higher Education in Europe,* 27 (4), pp. 433-41.

Wagner, L. (1998) What is the point of university league tables?. *The Use and Abuse of University League Tables Conference* London 1998. London: CVCP News.

Walberg, H.J. and Tsai, S.L. (1983) Matthew effects in education. *American Educational Research Journal,* 20 (3), pp. 359-373.

Webster, D.S. (1981) Advantages and disadvantages of methods of assessing quality. *Change,* 13 (7), pp. 20-24.

Webster, D.S. (1986) *Academic Quality Rankings of American Colleges and Universities.* Springfield: Charles C. Thomas.

Williams, R. (2005) Broadening the Criteria: lessons from the Australian rankings. In: *First International Conference on World Class Universities*. Shanghai 2005. Shanghai: Shanghai Jiao Tong University, pp. 1-17.

Yin, R.K. (1984) *Case study research: design and methods.* Beverly Hills: Sage.

Yin, R.K. (1994) *Case study research: design and methods.* Second ed. Thousand Oaks: Sage.

Appendices

Appendix A: Background of Rankings

Rankings are fuelled by the idea of knowledge economy (k-economy) and discussed in conjunction with diversity and classification.

Knowledge Economy

In current k-economy era, economy is no longer solely powered by finance and industry. Without the confine of natural scarcity, knowledge generates sustainable power for economic development.

Knowledge is 'a mixture of: 1) tradeable knowledge-intensive products, from intellectual property and commercial know-how to industrial goods; and 2) free knowledge goods produced and exchanged on an open source basis. Together the production, exchange and circulation of research, knowledge and information constitute the k-economy' (Marginson, 2009, p.3).

In theory, knowledge should flow freely with equal value. 'But when it passes through institutional settings and publications systems it becomes structured and acquires new social meanings. It is channelled and restricted in often one-way flows. The means of knowledge production are concentrated in particular universities, cities, national systems, languages, corporations and brands with a superior capacity in production or dissemination that stamp their presence on the k-economy and pull the flows in their favour' (Marginson, 2009, p.7). In k-economy, rankings play an important role in assigning different values to different parcels of knowledge (Marginson, 2009).

Diversity

Rankings often are criticised by jeopardising diversity. Diversity has been viewed positively in higher education system for its crucial role in policy making context. Diverse higher education system meets student needs, provides for social mobility, meets the needs of the labour market, serves the political needs of interest groups, permits the crucial combination of elite and mass higher education, increases the level of effectiveness of higher education institutions, and offers opportunities for experimenting with innovation (van Vught, 2008).

But whether diversification or dedifferentiation is the current trend in education is arguable. Parsons and Platt (1973) believe that the diversification is occurring with the emergence of new functions of education. Clark (1978) shares similar view and argues that the increasing complexity of higher education systems is the result of increasing

variety of the student population, growth of the labour market for academic graduates and emergence and growth of new disciplines.

However, other researchers believe that the current higher education system is characterized by dedifferentiation, because of Academic Drift, the imitating behaviour by lower status institutions (Riesman, 1958); centralised and uniform governmental policies (Birnbaum, 1983) ;and academic conservatism who defending the traditional academic norms and values (Rhoades, 1990).

The empirical evidences collected by Birnbaum (1983) and Morphew (2009) in USA; Skolnik (1986) in Canada; Rhoades (1990) in UK, France, Sweden and the USA; Maassen and Potman (1990) in Netherlands; Morphew and Huisman (2002) in USA and Netherlands; Meek (1991) in Australia support the idea that dedifferentiation, rather than diversification, is the current trend in education.

Yet, Teichler (2008) reminds us that diversity means more than only in respect to the institutional patterns of the higher education system. Birnbaum (1983) and Huisman (1995) refer to the administrative or organisational diversity; Huisman (1998) emphasises the internal diversity; Teichler (2006) stresses the intra-institutional diversity; Ayalon and Yogev (2006) distinguish diversity by the composition of the student body; Clark (1996) describes diversity according to the knowledge system; Tomusk (2003) puts emphasis on the ownership and control within diversity; Dill and Teixeira (2000) draw attention to programme diversity.

Due to the multi-dimensional definition of diversity and the lack of empirical research, it is hard to draw a general conclusion whether the current education system is moving towards diversification or dedifferentiation and whether rankings hamper or facilitate educational development. But what is clear is rankings are reflecting and deepening the pressures that already exist; they are not the major cause of these pressures (Ehrenberg, 2005).

Classification

Classification is another issue often associated with ranking. McCormick (2008) argues that classification directly or indirectly links to ranking. Classification can be regarded as ranking. Classification, such as the Carnegie Classification, uses objective measures to group universities. The membership can be viewed as a hierarchy. On the other hand, classification is often employed in the first step when compiling rankings. The rationale behind this is that any benchmarking or ranking should compare what is comparable (Teeter and Christal, 1987;Stralser, 1995).

Classifications differ in different countries due to different context. In USA the latest Carnegie Classification groups universities into: Associate's Colleges, Doctorate-granting Universities, Master's Colleges and Universities, Baccalaureate Colleges, Special Focus Institutions, and Tribal Colleges (The Carnegie Foundation

for the Advancement of Teaching, 2009). In England, Scott (2001) argues that there are seven types of higher education institutions: 1) Oxford and Cambridge, 2) University of London, 3) the old civics established in the Victorian period, 4) the 'redbricks' founded in the late 19th and early 20th century, 5) the new universities built on greenfield sites during the 1960s, 6) the technological universities and former colleges of advanced technology and 7) the new universities (former polytechnics). But the picture is not clear when classification crosses different countries. Thus, Van Vught et al. (2005) call for building a typology of higher education institutions in Europe to advance the Bologna process.

Appendix B: Current Rankings in Different Countries

USA

The U.S. News, a pioneer ranking compiler, currently produces two parallel domestic league tables for undergraduate and postgraduate education respectively. In the America's Best Colleges, universities are categorised into: national universities, liberal arts colleges, master's universities, and baccalaureate universities. The last two groups are further subdivided into north, south, midwest, and west region. The indicators are:

- Peer assessment (weighting: 25 percent)
- Retention (20 percent in national universities and liberal arts colleges and 25 percent in master's and baccalaureate colleges)
- Faculty resources (20 percent)
- Student selectivity (15 percent)
- Financial resources (10 percent)
- Graduation rate performance (5 percent; only in national universities and liberal arts colleges)
- Alumni giving rate (5 percent)

(Morse and Flanigan, 2008, p.2-3)

In the Best Graduate Schools, universities are grouped according to subjects, including business, education, engineering, law, medical, the Sciences, library and information studies, social sciences and humanities, health, public affairs, and fine arts and ranked by different sets of indicators (U.S.News, 2009).

Canada

Sharing some similarities with US higher education system, MacLeans' categorises Canadian universities into: primarily undergraduate, comprehensive, and medical doctoral universities and ranks them by Students/Classes (20%), Faculty (18%), Resources (12%), Student Support (13%), Library (15%), and Reputation (22%) (Dwyer, 2008).

UK

UK is also a competitive market for league tables. The longest running league table for universities is the Times Good University Guide. The indicators in the latest 2010 version are (Times, 2009):

- Student Satisfaction—National Student Survey (NSS)
- Research Quality—2008 RAE and eligible staff number
- Entry Standards—A and AS Levels and Highers and Advanced Highers
- Student-staff ratio
- Services and Facilities Spend

- Completion—Percentage of students projected to leave with a degree including students who transfer to other institutions
- Good Honours
- Graduate Prospects—Destinations of full-time first degree UK domiciled leavers

Australia

The ranking of Australian universities by Williams and Van Dyke (2007;2008) is relatively new. It is a ranking for universities' international standing, whose indicators cover research, teaching and research training. The changes of ranking methods introduced in 2008 (Williams, 2008) are the controlling for discipline areas, staff number, and student number. The indicators are:

Measurements in Research

- Publications and citations in journals—ISI
- Total publications including books
- Success in national competitive grants—Discovery and Linkage projects in Australian Research Council (ARC) and all National Health and Medical Research Council (NHMRC) projects
- Membership of academies—The ARC awards Federation Fellowships
- Doctoral completions

Measurements in Teaching and Research Training

- Student satisfaction graduation—Graduate Destination Survey (GDS) and Postgraduate Research Experience Questionnaire (PREQ)
- Progression rates—Tertiary Entrances Scores (TES) and GDS
- Retention rates students—GDS
- Resources available for learning and teaching—Commonwealth Grants Scheme and other grants; Australian Government funding; upfront student contributions; fees and charges for teaching; investment income; donations and bequests.

Germany

The Center for Higher Education Development (CHE) moves the concept of ranking to a new scope. Abstaining the overall score for universities, the CHE chooses subjects as the comparison units and cluster universities' subjects into top, median, and bottom for single indicator. Due to the interactive internet technology, users can customise the ranking according to their needs. The CHE also highlights the diversity of its indicators, comprising different data sources, hard facts, and subjective judgments (Federkeil, 2002). The indicators extensively cover almost every aspect of university life, even, for example, information about bicycle and students' parents. The indicators cover 9 modules (for details please see Appendix):

- Job market and career-orientation
- Equipment
- Research
- Overall opinions
- International orientation

- Result of study
- Town and University
- Students
- Academic studies and teaching

(ZEIT, 2009, p.1)

The CHE is also constructing an international ranking based on subjects—Excellence Ranking of European Graduate Programmes. The objective indicators include: number of publications in the Web of Science; citations; highly cited authors, Nobel Prize winners or field medallists; number of projects in the Marie Curie Programme. The subjective indicators are: students' judgements on training, examinations, advisor, laboratories, and career centre; the percentage of international and female staff; the percentage of female and international students (Berghoff, Brandenburg and Muller-Boling, 2008).

References:

Berghoff, S., Brandenburg, U. and Muller-Boling, D. (2008) Identifying the best: the CHE excellence ranking of European graduate programmes in the natural sciences and mathematics. *Higher Education in Europe,* 33 (2-3), pp. 273-283.

Dwyer, M. (2008) *Our 18th Annual Rankings.* MacLeans'. [Online]. Available from: <http://oncampus.macleans.ca/education/2008/12/19/our-18th-annual-rankings/> (Accessed July/17, 2009).

Federkeil, G. (2002) Some aspects of ranking methodology--the CHE-ranking of German universities. *Higher Education in Europe,* 27 (4), pp. 389-97.

Morse, R. and Flanigan, S. (2008) *How We Calculate the Rankings.* U.S.News. [Online]. Available from: <http://www.usnews.com/articles/education/best-colleges/2008/08/21/how-we-calculate-the-rankings.htm> (Accessed July/17, 2009).

Times (2009) *Good University Guide 2010: how the tables work.* Times Online. [Online]. Available from: <http://www.timesonline.co.uk/tol/life_and_style/education/good_university_guide/article2235223.ece> (Accessed July/16, 2009).

U.S.News (2009) *Best Graduate Schools.* U.S.News. [Online]. Available from: <http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools> (Accessed July/17, 2009).

Williams, R. (2008) Ranking Australian universities: controlling for scope. *Higher Education in Europe,* 33 (2), pp. 331-344.

Williams, R. and Van Dyke, N. (2007) Measuring the international standing of universities with an application to Australian universities. *Higher Education,* 53 (6), pp. 819-841.

Williams, R. and Van Dyke, N. (2008) Reputation and reality: ranking major disciplines in Australian universities. *Higher Education: The International Journal of Higher Education and Educational Planning,* 56 (1), pp. 1-28.

ZEIT (2009) *Cataloque of Criteria.* ZEIT Online. [Online]. Available from: <http://ranking.zeit.de/che10/CHE_en?module=Baustein#top> (Accessed July/16, 2009).

Appendix C: Webster's Categories of Indicators

Webster (1986) gives a comprehensive summary of the methods to collect data for these indicators and the advantages and disadvantages of them.

Categories of academic quality

Category	Method/indicator	Advantage
Faculty accomplishments		They produce results with face validity, i.e., results that most nearly match what the educated
	Surveys of reputation (e.g., ratings of faculty or programme reputation)	public considers the hierarchy of colleges and universities to be *Disadvantage* The overall reputation of an institution may influence the assessments by raters of the particular department(s) they are being asked to rank
	Counts of faculty awards, honours, and prizes	They are useful for ranking the best or the better institutions *Disadvantage* They may be years behind or ahead of reality
	Counts of faculty citations in citation indexes	Useful in assessing the influence and importance of the publications of faculty members, and not just their sheer volume *Disadvantage* The citation indexes on which the rankings are based do not distinguish between "good", "neutral", or "bad" citations
Student achievements	Distinguished alumni and the achievements of graduates after graduation	While only a small percentage of colleges and universities have faculties that produce much research, almost all of them attempt to prepare their students for rewarding careers in later life *Disadvantage*

	The scores of incoming students on standardized tests	Usually years, if not decades, behind reality The data are easy to obtain and are a measure by which most institutions can be ranked *Disadvantage* Based on the academic abilities of students before they enter college and thus fail to consider anything that these institutions do to educate their students once they enroll
Institutional academic resources	Compilation of measures of institutional resources, including educational expenditure per student, faculty–student ratios, and library resources	The data are easy to obtain and are a measure on which all institutions can be compared *Disadvantage* Offers little or no information about how often and how beneficially students use these resources

References:

Webster, D.S. (1986) *Academic Quality Rankings of American Colleges and Universities.* Springfield: Charles C. Thomas.

Appendix D: Van Dyke’s Categories of Indicators

Appendix E: Dill and Soo's Categories of Indicators

Appendix F: Usher and Savino's Categories of Indicators

Appendix G: New Classification

	New Classification	**Webster's**	**Dill and Soo's**	**Hendel and Stolz's**	**Usher and Savino's**
Students	Student Pre-achievements	Student Achievements	Students	Beginning characteristics students	Incoming grades
					Student status
	Student Learning Process		Satisfaction	Program/course proceedings	Student efforts : Hours spent in class pe
					Student efforts: % student participatio exchange projects
	Student Post-achievements	Student Achievements		Career prospects	Academic performanc
			Learning process	Outcome measurements	Graduation rate : Undergraduates only
			Graduation		Graduation rate : Master's only
					Graduation rate : Doctoral students only
					Graduation rate : International students
			Value-added	Value added	Type of degree obtain
			Employment		Retention : 1st to 2nd

					Work status
					Further / professional education
Staff	Staff Accomplishments	Faculty Accomplishments	Faculty	Faculty and staff	Faculty with research
					Funding and financial Science grants
					Funding and financial Social sciences and grants
					Research budget inclu
					Research budget: Tot of grants and projects
					Patents
					Number of research b per institution
					Number of research b affiliated research i centres for studies, etc
					Other output
					Academic quality of r

	Staff Productivity		Faculty	Faculty and staff	Citations :Science-indices(ie., the Scien Index; refers to natura engineering and oth fields)
				Research	Citations : Social oriented indices (ie., Science Citation Inde humanities)
					Citations : 'Highly determined by Thomsc
					Citations : Other
					Publications: Nature a (not quite the same cited' above)
					Publications: Publishe science-oriented indic Science Citation Index
					Publications: Nor Published in social oriented indices (ie., Science Citation Index
					Publications: Published papers in otl
					Publications: Books (c

	Staff Reputation	Faculty Accomplishments	Faculty	Faculty and staff	Teaching quality : Qua for teaching positio Master's, etc.)
					Awards : International
					Awards : National
					Awards : Regional (ie, state/ provincial or wit national borders)
	Staff Teaching		Teaching	Teaching quality	Teaching quality : performance on st 3rd-party tests if given
					Teaching quality : Pe on 'own metrics'
Programme/ University Resources	Admission				Percentage with grade set limit
					Performance on nation standardised tests or b
					Admittance : selectivit
					Admittance : nu applications to places
	Teaching and Leaning			Program design	Faculty/student ratio
				Organization and management	Social science faculty ratio

					Science faculty / stude
					Staff /student ratio 3rd-party : HESA (reg division)
					Course per teacher
					Number of full-time / p faculty
					Class size differentiati
					Classes taught by tenu
					Exchange programmes
					Number of classes taught'
					Aging and staff repl churn issues
					Number of doctoral an programmes
					Research staff : nu percentage of research (ie, as opposed to teach
	Student Support	Institutional Academic Resources	Financial resources and facilities	Financial resources	Administrative staff ratio

				Organization and management	
				Student support systems	Physical infrastructure Library : Acquisitions
					Physical infrastructure Library : total volumes
					Physical infrastructure Library : volumes per
					Physical infrastructure Library : Yearly ex outside of acquisitions
					Physical infrastructure Internet bandwidth
					Physical infrastru Computerisation of resources
					Funding and financial Student services
					Funding and financial resources: Expenditure
					Funding and financial Bursaries and sc disbursed by publi bodies

	Staff Support	Institutional Academic Resources	Financial resources and facilities	Financial resources	Per-teacher university
					Faculty pay rates for te
					Physical infrastructure Number of lecture spa

					Funding and financial Awards (not researcl subsidised or unsubsid
					Research budget : Expenditure (undefine
	Revenue	Institutional Academic Resources	Financial resources and facilities	Financial resources	Funding and financial resources: Public fund institutional budget
					Funding and financial resources: Private func (including foundations and organisations)
			Satisfaction		Funding and financial resources: Alumni sup
	Environment			Amenities	Out-of-locality percentage
					International student p
					Undergraduate student among all students : pe
					Ethnic diversity in stu
					% of international facu (v faculty as a whole)

	Institution Reputation		Reputation	Reputation	Among students/gradu
					Among academics
					Among general society sector / others outs connection to universit

Table 46 The comparison of different classifications

Appendix H:The Third Missions Indicators

As the non-financial benefit social service providers, universities should undertake the Social Third Mission.

> …social services can take the form of non-academic dissemination, media communication, volunteer contributions to the community (of labour, expertise, educational outreach, etc.), social networking or contributions to public policy. Other activities, such as services for retired senior personnel, summer school for the staff's children, courses offered to academic staff on didactics and/or andragogy, cultural activities such as art expositions, cinema forum etc., would fit into the social third mission.
>
> (Montesinos et al., 2008, p.262)

Indicators for the Social Third Mission are related with staff and students exchange networking and cooperative projects with developing countries.

> Staff magnitudes:
> – number of exchanges per year (students/staff);
> – number of active exchange agreements;
> – number of subjects taught in exchange students' first language;
> – percent of incoming and outgoing students/staff on exchange programmes.
> Networking magnitudes:
> – number of interchange associations/networks which university is member of
> – percentage used for international activity;
> – number of university offices abroad;
> – number of countries and institutions with which the university has agreements;
> – percent of work placements within international companies.
> Academic magnitudes:
> – number of joint bilateral academic degrees;
> – number of academic programmes abroad;
> – percent of university staff participating in international activities;
> – number of foreign languages offered
>
> (Montesinos et al., 2008, p.268)

The second dimension is the Enterprising Third Mission. The key success factor of the entrepreneur university lies in differentiated sources of funding (Clark, 1998).

> …enterprising activities can be found from a technology transfer point of view: consultancy to industry, patent registration, commercialisation of intellectual property, advisory work and contracts, shared development of research, problem-solving agendas or contract and collaborative research. From a continuing education point of view, the activities could be lifelong

> learning/teaching activities, curriculum alignment to societal needs, open and distance learning (ODL), commercialisation of facilities and conference organizing. From an employment perspective, the activities could be staff mobility (flow and exchange), student placements, former student employee links, and training for company creation and self-employment. International activities (student and teacher interchanges, doctoral and continuing education training in developing countries) require a special mention.
>
> (Montesinos et al., 2008, p.262)

Indicators for the Enterprising Third Mission are associated with Offered and tailor-made lifelong learning (LLL), E-learning material production, E-learning delivery, conference organisation.

> Academic magnitudes:
> – number of courses/LLL activities, hours/days committed;
> – number of university teachers involved;
> – percentage of university's teachers;
> – number of course participants.
> Economic magnitudes:
> – global incomes;
> – global expenses;
> – non-academic staff expenses;
> – academic staff expenses.
>
> (Montesinos et al., 2008, p.269)

The last dimension is related with innovation. Services include:

> …searching for seed or venture capital, business networking, company creation for patent exploitation, consulting for governments, joint ventures with industrial sectors, conferences for research, development and innovation in specific industries, and companies or regional innovation and networking with entrepreneurs. Science and technology parks are the logical and physical space where these services take place.
>
> (Montesinos et al., 2008, p.262-263)

Indicators for Innovative Third Mission are associated with seed or venture capital networking, company creation, patents exploitation, conferences for research, development and innovation (R&D&i) in specific industries.

> Entrepreneurship magnitudes:
> – number new companies created per year (spin-offs/star-ups);
> – number of students trained in entrepreneurship per year;
> – number of companies located in USTP shared by academics
> – number of companies located in the USTP;
> – number of international patents registered per year
> Networking magnitudes:
> – number of events open for industry organised per year;

– number of new projects per year;
– number of non-academic agreements with mature companies.

Knowledge-based jobs magnitudes:
– number of non-academic staff
– number of academic staff;
– number of yearly grants;
– total USTP sq.m.;
– sq.m dedicated to R&D&i groups;
– sq.m. dedicated to companies.

(Montesinos et al., 2008, p.269)

References:

Clark, B.R. (1998) *Creating Entrepreneurial Universities: organizational pathways of transformation.* London: Pergamon.

Montesinos, P., Carot, J.M., Martinez, J.M. and Mora, F. (2008) Third mission ranking for world class universities: beyond teaching and research. *Higher Education in Europe,* 33 (2), pp. 259-271.

Appendix I: VALUE Rubrics

Intellectual and Practical Skills

- Inquiry and analysis
- Critical thinking
- Creative thinking
- Written communication
- Oral communication
- Reading
- Quantitative literacy
- Information literacy
- Teamwork
- Problem solving

Personal and Social Responsibility

- Civic knowledge and engagement—local and global
- Intercultural knowledge and competence
- Ethical reasoning
- Foundations and skills for lifelong learning

Integrative Learning

- Integrative learning

(AAC&U, 2009, p. 1)

References:

AAC&U (2009) *Metarubrics*. AAC&U. [Online]. Available from: <http://www.aacu.org/value/metarubrics.cfm> (Accessed July/9, 2009).

Appendix J: The Indicators of the ARWU and the WUR.

The current definitions of the six categories of indicators used in ARWU are:

Alumni

The total number of the alumni of an institution winning Nobel Prizes and Fields Medals…

Award

The total number of the staff of an institution winning Nobel Prizes in Physics, Chemistry, Medicine and Economics and Fields Medal in Mathematics...

HiCi

The number of highly cited researchers in 21 subject categories. These individuals are the most highly cited within each category…

N&S

The number of articles published in Nature and Science between 2003 and 2007…

PUB

Total number of articles indexed in Science Citation Index-Expanded and Social Science Citation Index in 2007. Only publications of article type are considered. When calculating the total number of articles of an institution, a special weight of two was introduced for articles indexed in Social Science Citation Index.

PCP

The weighted scores of the above five indicators divided by the number of full-time equivalent academic staff. If the number of academic staff for institutions of a country cannot be obtained, the weighted scores of the above five indicators is used. For ARWU 2008, the numbers of full-time equivalent academic staff are obtained for institutions in USA, UK, France, Japan, Italy, China, Australia, Netherlands, Sweden, Switzerland, Belgium, South Korea, Czech, Slovenia, New Zealand etc.

The current definitions of the six categories of indicators used in WUR are:

- **Peer review**: data are collected from QS's academic peer survey.
- **Employer review**: data are collected from QS's employer survey.
- **Citation**: Total citation count for the last five years, adjusted by number of the Full Time Equivalent (FTE) faculty.
- **Student Faculty Ratio**: Ratio of FTE students and FTE faculty.
- **International Faculty**: proportion of FTE faculty that hold an overseas nationality.
- **International Students**: proportion of FTE students that hold an overseas nationality.

Appendix K: Correlation

Variable	1	2	3	4	5	6
1. Staff Winners	1.00	.35	.45	.07	.62	.68
2. Highly Cited Researchers	.35	1.00	.77	.51	.28	.78
3. Articles in Nature or Science	.45	.77	1.00	.59	.49	.87
4. Publications	.07	.51	.59	1.00	.08	.53
5.Size of Institution	.62	.28	.49	.08	1.00	.68
6. Total Score	.68	.78	.87	.53	.68	1.00

Table 47 2003 ARWU Correlation Coefficients

Variable	1	2	3	4	5	6	7
1. Alumni Winners	1.00	.52	.33	.40	.33	.36	.58
2. Staff Winners	.52	1.00	.44	.52	.13	.57	.70
3. Highly Cited Researchers	.33	.44	1.00	.81	.59	.28	.83
4. Articles in Nature or Science	.40	.52	.81	1.00	.57	.45	.86
5. Publications	.33	.13	.59	.57	1.00	.06	.59
6.Size of Institution	.36	.57	.28	.45	.06	1.00	.56
7. Total Score	.58	.70	.83	.86	.59	.56	1.00

Table 48 2004 ARWU Correlation Coefficients

Variable	1	2	3	4	5	6	7
1. Alumni Winners	1.00	.48	.33	.39	.32	.43	.55
2. Staff Winners	.48	1.00	.40	.51	.12	.62	.68
3. Highly Cited Researchers	.33	.40	1.00	.81	.65	.48	.84
4. Articles in Nature or Science	.39	.51	.81	1.00	.61	.65	.88
5. Publications	.32	.12	.65	.61	1.00	.29	.65
6.Size of Institution	.43	.62	.48	.65	.29	1.00	.69
7. Total Score	.55	.68	.84	.88	.65	.69	1.00

Table 49 2005 ARWU Correlation Coefficients

Variable	1	2	3	4	5	6	7
1. Alumni Winners	1.00	.50	.33	.37	.29	.41	.55
2. Staff Winners	.50	1.00	.39	.53	.09	.60	.68
3. Highly Cited Researchers	.33	.39	1.00	.81	.64	.44	.84
4. Articles in Nature or Science	.37	.53	.81	1.00	.59	.62	.90
5. Publications	.29	.09	.64	.59	1.00	.21	.62
6.Size of Institution	.41	.60	.44	.62	.21	1.00	.65
7. Total Score	.55	.68	.84	.90	.62	.65	1.00

Table 50 2006 ARWU Correlation Coefficients

Variable	1	2	3	4	5	6	7
1. Alumni Winners	1.00	.52	.32	.37	.28	.41	.55
2. Staff Winners	.52	1.00	.39	.52	.09	.60	.68
3. Highly Cited Researchers	.32	.39	1.00	.81	.64	.44	.84
4. Articles in Nature or Science	.37	.52	.81	1.00	.59	.62	.90
5. Publications	.28	.09	.64	.59	1.00	.21	.62
6.Size of Institution	.41	.60	.44	.62	.21	1.00	.65
7. Total Score	.55	.68	.84	.90	.62	.65	1.00

Table 51 2007 ARWU Correlation Coefficients

Variable	1	2	3	4	5	6	7
1. Alumni Winners	1.00	.52	.32	.36	.29	.35	.53
2. Staff Winners	.52	1.00	.40	.52	.11	.56	.69
3. Highly Cited Researchers	.32	.40	1.00	.82	.62	.47	.85
4. Articles in Nature or Science	.36	.52	.82	1.00	.61	.62	.90
5. Publications	.29	.11	.62	.61	1.00	.24	.63
6.Size of Institution	.35	.56	.47	.62	.24	1.00	.63
7. Total Score	.53	.69	.85	.90	.63	.63	1.00

Table 52 2008 ARWU Correlation Coefficients

Appendix L: Multiple Regression

		B	SE B	β
Step 1	Constant	26.40	1.38	
	Staff Winners	0.53	0.04	0.82*
Step 2	Constant	11.04	0.76	
	Staff Winners	0.24	0.02	0.37*
	Highly Cited Researchers	0.20	0.03	0.25*
	Articles in Nature or Science	0.45	0.04	0.48*
Step 3	Constant	5.41	1.18	
	Staff Winners	0.27	0.02	0.43*
	Highly Cited Researchers	0.19	0.03	0.24*
	Articles in Nature or Science	0.34	0.04	0.36*
	Publications	0.16	0.03	0.14*
Step 4	Constant	0.01	0.20	
	Staff Winners	0.21	0.00	0.34*
	Highly Cited Researchers	0.21	0.00	0.26*
	Articles in Nature or Science	0.21	0.00	0.23*
	Publications	0.21	0.00	0.19*
	Size of Institution	0.21	0.00	0.23*

Table 53 2003 ARWU Multiple Regression

Note R^2=0.69 for Step 1; ΔR^2=0.28 for Step 2; ΔR^2=0.01 for Step 3; ΔR^2=0.02 for Step 4 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	22.43	1.34	
	Alumni Winners	0.24	0.06	0.33*
	Staff Winners	0.39	0.05	0.60*
Step 2	Constant	8.53	0.59	
	Alumni Winners	0.16	0.02	0.22*
	Staff Winners	0.17	0.02	0.26*
	Highly Cited Researchers	0.23	0.03	0.28*
	Articles in Nature or Science	0.36	0.03	0.38*
Step 3	Constant	2.26	0.60	
	Alumni Winners	0.11	0.01	0.15*
	Staff Winners	0.24	0.01	0.37*
	Highly Cited Researchers	0.19	0.02	0.23*
	Articles in Nature or Science	0.27	0.02	0.28*
	Publications	0.19	0.02	0.17*
Step 4	Constant	-0.02	0.02	
	Alumni Winners	0.10	0.00	0.14*
	Staff Winners	0.21	0.00	0.32*
	Highly Cited Researchers	0.21	0.00	0.25*
	Articles in Nature or Science	0.21	0.00	0.22*
	Publications	0.21	0.00	0.18*
	Size of Institution	0.10	0.00	0.11*

Table 54 2004 ARWU Multiple Regression

Note R^2=0.77 for Step 1; ΔR^2=0.21 for Step 2; ΔR^2=0.01 for Step 3; ΔR^2=0.00 for Step 4 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	21.33	1.34	
	Alumni Winners	0.26	0.05	0.37*
	Staff Winners	0.36	0.05	0.58*
Step 2	Constant	7.84	0.49	
	Alumni Winners	0.15	0.01	0.22*
	Staff Winners	0.16	0.01	0.26*
	Highly Cited Researchers	0.27	0.02	0.30*
	Articles in Nature or Science	0.33	0.03	0.36*
Step 3	Constant	1.91	0.40	
	Alumni Winners	0.11	0.01	0.16*
	Staff Winners	0.22	0.01	0.36*
	Highly Cited Researchers	0.21	0.01	0.24*
	Articles in Nature or Science	0.26	0.01	0.28*
	Publications	0.19	0.01	0.16*
Step 4	Constant	-0.01	0.02	
	Alumni Winners	0.10	0.00	0.15*
	Staff Winners	0.21	0.00	0.33*
	Highly Cited Researchers	0.21	0.00	0.23*
	Articles in Nature or Science	0.21	0.00	0.22*
	Publications	0.21	0.00	0.18*
	Size of Institution	0.10	0.00	0.09*

Table 55 2005 ARWU Multiple Regression

Note R^2=0.77 for Step 1; ΔR^2=0.22 for Step 2; ΔR^2=0.01 for Step 3; ΔR^2=0.00 for Step 4 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	21.48	1.33	
	Alumni Winners	0.25	0.06	0.36*
	Staff Winners	0.36	0.05	0.57*
Step 2	Constant	8.17	0.48	
	Alumni Winners	0.15	0.01	0.22*
	Staff Winners	0.15	0.01	0.24*
	Highly Cited Researchers	0.27	0.02	0.31*
	Articles in Nature or Science	0.33	0.03	0.36*
Step 3	Constant	2.19	0.41	
	Alumni Winners	0.11	0.01	0.16*
	Staff Winners	0.22	0.01	0.36*
	Highly Cited Researchers	0.21	0.01	0.24*
	Articles in Nature or Science	0.25	0.01	0.28*
	Publications	0.19	0.01	0.17*
Step 4	Constant	0.03	0.02	
	Alumni Winners	0.10	0.00	0.15*
	Staff Winners	0.21	0.00	0.33*
	Highly Cited Researchers	0.21	0.00	0.24*
	Articles in Nature or Science	0.21	0.00	0.23*
	Publications	0.20	0.00	0.18*
	Size of Institution	0.10	0.00	0.09*

Table 56 2006 ARWU Multiple Regression

Note R^2=0.76 for Step 1; ΔR^2=0.23 for Step 2; ΔR^2=0.01 for Step 3; ΔR^2=0.00 for Step 4 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	21.30	1.33	
	Alumni Winners	0.26	0.06	0.35*
	Staff Winners	0.35	0.05	0.58*
Step 2	Constant	8.06	0.48	
	Alumni Winners	0.15	0.01	0.21*
	Staff Winners	0.16	0.01	0.26*
	Highly Cited Researchers	0.28	0.02	0.32*
	Articles in Nature or Science	0.32	0.03	0.35*
Step 3	Constant	2.18	0.41	
	Alumni Winners	0.11	0.01	0.15*
	Staff Winners	0.22	0.01	0.36*
	Highly Cited Researchers	0.21	0.01	0.24*
	Articles in Nature or Science	0.25	0.01	0.28*
	Publications	0.19	0.01	0.16*
Step 4	Constant	0.01	0.02	
	Alumni Winners	0.10	0.00	0.14*
	Staff Winners	0.21	0.00	0.34*
	Highly Cited Researchers	0.21	0.00	0.23*
	Articles in Nature or Science	0.21	0.00	0.24*
	Publications	0.21	0.00	0.18*
	Size of Institution	0.10	0.00	0.09*

Table 57 2007 ARWU Multiple Regression

Note R^2=0.77 for Step 1; ΔR^2=0.22 for Step 2; ΔR^2=0.01 for Step 3; ΔR^2=0.00 for Step 4 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	21.37	1.33	
	Alumni Winners	0.25	0.06	0.35*
	Staff Winners	0.36	0.05	0.58*
Step 2	Constant	8.45	0.49	
	Alumni Winners	0.15	0.02	0.20*
	Staff Winners	0.15	0.01	0.25*
	Highly Cited Researchers	0.27	0.02	0.30*
	Articles in Nature or Science	0.34	0.03	0.38*
Step 3	Constant	2.20	0.45	
	Alumni Winners	0.11	0.01	0.15*
	Staff Winners	0.22	0.01	0.36*
	Highly Cited Researchers	0.21	0.01	0.24*
	Articles in Nature or Science	0.26	0.01	0.28*
	Publications	0.19	0.01	0.16*
Step 4	Constant	-0.02	0.31	
	Alumni Winners	0.10	0.00	0.14*
	Staff Winners	0.21	0.00	0.34*
	Highly Cited Researchers	0.21	0.00	0.23*
	Articles in Nature or Science	0.20	0.00	0.24*
	Publications	0.21	0.00	0.18*
	Size of Institution	0.10	0.00	0.10*

Table 58 2008 ARWU Multiple Regression

Note R^2=0.76 for Step 1; ΔR^2=0.23 for Step 2; ΔR^2=0.01 for Step 3; ΔR^2=0.00 for Step 4 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	9.14	0.71	
	Peer Review	0.13	0.00	0.94
Step 2	Constant	7.74	0.89	
	Peer Review	0.13	0.00	0.93
	Staff/student Ratio	0.08	0.03	0.08
Step 3	Constant	5.14	0.58	
	Peer Review	0.11	0.00	0.81
	Staff/student Ratio	0.07	0.02	0.06
	Citation	0.09	0.01	0.29
Step 4	Constant	0.85	0.34	
	Peer Review	0.11	0.00	0.78
	Staff/student Ratio	0.10	0.01	0.09
	Citation	0.10	0.00	0.33
	International Staff	0.15	0.01	0.18
Step 5	Constant	-0.07	0.22	
	Peer Review	0.11	0.00	0.77
	Staff/student Ratio	0.09	0.01	0.08
	Citation	0.10	0.00	0.34
	International Staff	0.1	0.01	0.12
	International Students	0.11	0.01	0.10

Table 59 2004 WUR Multiple Regression

Note R^2=0.89 for Step 1; ΔR^2=0.01 for Step 2; ΔR^2=0.07 for Step 3; ΔR^2=0.03 for Step 4; ΔR^2=0.01 for Step5 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	7.35	1.39	
	Peer Review	0.81	0.03	0.91
Step 2	Constant	9.20	1.15	
	Peer Review	0.66	0.03	0.74
	Employer Review	0.26	0.03	0.30
Step 3	Constant	4.15	0.93	
	Peer Review	0.68	0.02	0.76
	Employer Review	0.19	0.02	0.26
	Staff/student Ratio	0.28	0.03	0.24
Step 4	Constant	2.86	0.60	
	Peer Review	0.63	0.02	0.70
	Employer Review	0.16	0.01	0.23
	Staff/student Ratio	0.31	0.02	0.27
	Citation	0.23	0.02	0.20
Step 5	Constant	0.64	0.20	
	Peer Review	0.59	0.01	0.66
	Employer Review	0.16	0.00	0.23
	Staff/student Ratio	0.30	0.01	0.26
	Citation	0.30	0.01	0.26
	International Staff	0.10	0.00	0.16
Step 6	Constant	0.07	0.06	
	Peer Review	0.59	0.01	0.66
	Employer Review	0.15	0.01	0.21
	Staff/student Ratio	0.30	0.01	0.26
	Citation	0.30	0.02	0.25
	International Staff	0.08	0.01	0.12
	International Students	0.07	0.02	0.06

Table 60 2005 WUR Multiple Regression

Note R^2=0.83 for Step 1; ΔR^2=0.06 for Step 2; ΔR^2=0.06 for Step 3; ΔR^2=0.03 for Step 4; ΔR^2=0.02 for Step5; ΔR^2=0.00 for Step6 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	14.04	1.80	
	Peer Review	0.80	0.04	0.87
Step 2	Constant	12.70	1.47	
	Peer Review	0.61	0.04	0.67
	Employer Review	0.26	0.03	0.35
Step 3	Constant	3.94	1.00	
	Peer Review	0.64	0.02	0.70
	Employer Review	0.18	0.02	0.24
	Staff/student Ratio	0.27	0.02	0.34
Step 4	Constant	2.96	0.63	
	Peer Review	0.60	0.02	0.65
	Employer Review	0.15	0.01	0.21
	Staff/student Ratio	0.27	0.01	0.35
	Citation	0.23	0.02	0.19
Step 5	Constant	0.30	0.23	
	Peer Review	0.56	0.01	0.61
	Employer Review	0.15	0.00	0.21
	Staff/student Ratio	0.29	0.00	0.37
	Citation	0.29	0.01	0.24
	International Staff	0.10	0.00	0.15
Step 6	Constant	0.10	0.06	
	Peer Review	0.56	0.00	0.61
	Employer Review	0.14	0.00	0.19
	Staff/student Ratio	0.28	0.00	0.36
	Citation	0.28	0.00	0.24
	International Staff	0.07	0.00	0.10
	International Students	0.07	0.00	0.07

Table 61 2006 WUR Multiple Regression

Note R^2=0.76 for Step 1; ΔR^2=0.08 for Step 2; ΔR^2=0.11 for Step 3; ΔR^2=0.03 for Step 4; ΔR^2=0.02 for Step5; ΔR^2=0.00 for Step6 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	30.31	3.26	
	Peer Review	0.56	0.04	0.79
Step 2	Constant	25.70	3.01	
	Peer Review	0.48	0.04	0.67
	Employer Review	0.15	0.03	0.31
Step 3	Constant	15.22	1.98	
	Peer Review	0.49	0.02	0.69
	Employer Review	0.11	0.02	0.22
	Staff/student Ratio	0.19	0.01	0.44
Step 4	Constant	4.29	1.62	
	Peer Review	0.43	0.02	0.61
	Employer Review	0.13	0.01	0.26
	Staff/student Ratio	0.20	0.01	0.47
	Citation	0.18	0.02	0.26
Step 5	Constant	-0.26	0.91	
	Peer Review	0.41	0.01	0.58
	Employer Review	0.12	0.01	0.22
	Staff/student Ratio	0.11	0.00	0.49
	Citation	0.21	0.01	0.30
	International Staff	0.08	0.01	0.20
Step 6	Constant	-0.39	0.59	
	Peer Review	0.41	0.01	0.58
	Employer Review	0.10	0.00	0.20
	Staff/student Ratio	0.20	0.00	0.47
	Citation	0.20	0.01	0.30
	International Staff	0.05	0.00	0.13
	International Students	0.05	0.00	0.12

Table 62 2007 WUR Multiple Regression

Note R^2=0.62 for Step 1; ΔR^2=0.08 for Step 2; ΔR^2=0.19 for Step 3; ΔR^2=0.06 for Step 4; ΔR^2=0.04 for Step5; ΔR^2=0.01 for Step6 (ps<.001). *p<.001.

		B	SE B	β
Step 1	Constant	29.37	3.87	
	Peer Review	0.57	0.05	0.74
Step 2	Constant	24.60	3.57	
	Peer Review	0.46	0.05	0.60
	Employer Review	0.18	0.03	0.33
Step 3	Constant	13.27	2.43	
	Peer Review	0.47	0.03	0.61
	Employer Review	0.14	0.02	0.26
	Staff/student Ratio	0.21	0.02	0.47
Step 4	Constant	4.91	1.34	
	Peer Review	0.40	0.02	0.53
	Employer Review	0.15	0.01	0.28
	Staff/student Ratio	0.21	0.01	0.46
	Citation	0.18	0.01	0.34
Step 5	Constant	0.30	0.63	
	Peer Review	0.41	0.01	0.54
	Employer Review	0.12	0.01	0.22
	Staff/student Ratio	0.21	0.00	0.47
	Citation	0.20	0.01	0.38
	International Staff	0.08	0.00	0.20
Step 6	Constant	-0.02	0.10	
	Peer Review	0.41	0.00	0.54
	Employer Review	0.10	0.00	0.20
	Staff/student Ratio	0.21	0.00	0.45
	Citation	0.21	0.00	0.38
	International Staff	0.05	0.00	0.12
	International Students	0.05	0.00	0.12

Table 63 2008 WUR Multiple Regression

Note R^2=0.55 for Step 1; ΔR^2=0.09 for Step 2; ΔR^2=0.21 for Step 3; ΔR^2=0.11 for Step 4; ΔR^2=0.03 for Step5; ΔR^2=0.01 for Step6 (ps<.001). *p<.001.

Variable	1	2	3	4	5	6
1. Peer Review	1.00	.10	.22	.03	.15	.86
2.Staff/student Ratio	.10	1.00	.07	-.20	-.13	.13
3.Citation	.22	.07	1.00	-.33	-.40	.41
4. International Staff	.03	-.20	-.33	1.00	.72	.20
5.Internatioanl Students	.15	-.13	-.40	.72	1.00	.28
6. Overall Score	.86	.13	.41	.20	.28	1.00

Table 64 2004 WUR Correlation Coefficients

Variable	1	2	3	4	5	6	7
1. Peer Review	1.00	.35	.05	.13	.17	.26	.88
2. Employer Review	.35	1.00	.10	.15	.17	.31	.54
3.Staff/student Ratio	.05	.10	1.00	-.12	-.09	.08	.28
4.Citation	.13	.15	-.12	1.00	-.38	-.16	.15
5. International Staff	.17	.17	-.10	-.38	1.00	.58	.26
6.Internatioanl Students	.26	.31	.08	-.16	.58	1.00	.40
7. Overall Score	.88	.54	.28	.15	.26	.40	1.00

Table 65 2005 WUR Correlation Coefficients

Variable	1	2	3	4	5	6	7
1. Peer Review	1.00	.42	.08	.15	.23	.28	.77
2. Employer Review	.42	1.00	.21	.17	.03	.32	.62
3.Staff/student Ratio	.08	.21	1.00	.01	-.15	.10	.52
4.Citation	.15	.17	.01	1.00	-.31	-.11	.23
5. International Staff	.23	.03	-.15	-.31	1.00	.62	.18
6.Internatioanl Students	.28	.32	.10	-.11	.62	1.00	.42
7. Overall Score	.77	.62	.52	.23	.18	.42	1.00

Table 66 2006 WUR Correlation Coefficients

Variable	1	2	3	4	5	6	7
1. Peer Review	1.00	.48	.11	.34	.14	.21	.83
2. Employer Review	.48	1.00	.22	-.04	.27	.40	.62
3.Staff/student Ratio	.11	.22	1.00	-.07	-.04	.18	.48
4.Citation	.34	-.04	-.07	1.00	-.25	-.19	.30
5. International Staff	.14	.27	-.04	-.25	1.00	.62	.26
6.Internatioanl Students	.21	.40	.18	-.19	.62	1.00	.41
7. Overall Score	.83	.62	.48	.30	.26	.41	1.00

Table 67 2007 WUR Correlation Coefficients

Variable	1	2	3	4	5	6	7
1. Peer Review	1.00	.53	.12	.30	.04	.15	.81
2. Employer Review	.53	1.00	.21	.07	.30	.45	.64
3.Staff/student Ratio	.12	.21	1.00	.04	-.05	.17	.50
4.Citation	.30	.07	.04	1.00	-.23	-.16	.43
5. International Staff	.04	.30	-.05	-.23	1.00	.64	.16
6.Internatioanl Students	.15	.45	.17	-.16	.64	1.00	.33
7. Overall Score	.81	.64	.50	.43	.16	.33	1.00

Table 68 2008 WUR Correlation Coefficients

Appendix M: Principal Component Analysis

	Component Matrix		Pattern Matrix		Structure Matr	
	Component 1 Loading	Component 2 Loading	Component 1 Loading	Component 2 Loading	Component 1 Loading	Comp 2 Lo
Staff Winners	.789	-.448	.906		.907	.3
Highly Cited Researchers	.871	.207	.422	.655	.653	.8
Articles in Nature or Science	.950	.124	.544	.620	.762	.8
Publications	.601	.722	-.186	.989	.162	.9
Size of Institution	.746	-.508	.926		.900	.2

Table 69 2003 ARWU Principal Component Analysis

	Component Matrix		Pattern Matrix		Structure Matr	
	Component 1 Loading	Component 2 Loading	Component 1 Loading	Component 2 Loading	Component 1 Loading	Comp 2 Lo
Alumni Winners	.801	-.186	.725	.194	.803	.4
Staff Winners	.842	-.406	.936		.935	.3
Highly Cited Researchers	.853	.296	.363	.693	.642	.8
Articles in Nature or Science	.929	.160	.530	.594	.770	.8
Publications	.606	.713	-.159	.989	.240	.9
Size of Institution	.715	-.479	.907	-.136	.852	.2

Table 70 2004 ARWU Principal Component Analysis

	Component Matrix	Component Score Coefficient
	Component 1 Loading	Component 1
Alumni Winners	.791	.194
Staff Winners	.817	.200
Highly Cited Researchers	.871	.213
Articles in Nature or Science	.937	.230
Publications	.653	.160
Size of Institution	.851	.209

Table 71 2005 ARWU Principal Component Analysis

	Component Matrix		Pattern Matrix		Structure Matr	
	Component 1 Loading	Component 2 Loading	Component 1 Loading	Component 2 Loading	Component 1 Loading	Comp 2 Lo
Alumni Winners	.791	-.244	.796		.825	.4
Staff Winners	.824	-.471	1.004	-.152	.939	.2
Highly Cited Researchers	.870	.309	.411	.670	.697	.8
Articles in Nature or Science	.939	.125	.612	.508	.829	.7
Publications	.614	.730	-.122	1.000	.305	.9
Size of Institution	.836	-.304	.878		.889	.4

Table 72 2006 ARWU Principal Component Analysis

	Component Matrix		Pattern Matrix		Structure Matr	
	Component 1 Loading	Component 2 Loading	Component 1 Loading	Component 2 Loading	Component 1 Loading	Comp 2 Lo
Alumni Winners	.797	-.254	.810		.835	.4
Staff Winners	.828	-.459	.999		.938	.2
Highly Cited Researchers	.874	.310	.416	.670	.702	.8
Articles in Nature or Science	.939	.127	.612	.507	.829	.7
Publications	.612	.734	-.123	1.002	.305	.9
Size of Institution	.837	-.307	.883		.891	.3

Table 73 2007 ARWU Principal Component Analysis

	Component Matrix		Pattern Matrix		Structure Ma	
	Component 1 Loading	Component 2 Loading	Component 1 Loading	Component 2 Loading	Component 1 Loading	Cor 2 I
Alumni Winners	.791	-.266	.817		.834	
Staff Winners	.824	-.468	1.002	-.153	.938	
Highly Cited Researchers	.879	.283	.449	.640	.718	
Articles in Nature or Science	.940	.142	.607	.520	.826	
Publications	.608	.736	-.118	.998	.302	
Size of Institution	.831	-.280	.859		.876	

Table 74 2008 ARWU Principal Component Analysis

	Component Matrix		Pattern Matrix		Structure Matrix	
	Component 1 Loading	Component 2 Loading	Component 1 Loading	Component 2 Loading	Component 1 Loading	Compo 2 Loa
Peer Review	-.215	.833	.257	.854	.151	.82
Staff/student Ratio	-.286	.338		.430	-.119	.43
Citation	-.551	.587	-.162	.769	-.258	.78
International Staff	.817	.390	.907		.905	
International Students	.843	.333	.899		.905	-.15

Table 75 2004 WUR Principal Component Analysis

	Component Matrix		Pattern Matrix		Structure Matrix	
	Component 1 Loading	Component 2 Loading	Component 1 Loading	Component 2 Loading	Component 1 Loading	Compo 2 Loa
Peer Review	.746	.394	.822	-.149	.831	-.19
Employer Review	.756	.346	.798	-.193	.809	-.24
Staff/student Ratio	.203	-.237		-.312		-.31
Citation	.293	.756	.721	.417	.696	.37
International Staff	.537	-.701		-.884		-.88
International Students	.709	-.498	.203	-.830	.252	-.84

Table 76 2005 WUR Principal Component Analysis

	Component Matrix		Pattern Matrix		Structure Matrix	

	Component 1 Loading	Component 2 Loading	Component 1 Loading	Component 2 Loading	Component 1 Loading	Compo 2 Loa
Peer Review	.777	.213	.743	-.297	.749	-.31
Employer Review	.786	.309	.810	-.226	.814	-.2
Staff/student Ratio	.293	.348	.446	.101	.444	
Citation	.362	.618	.667	.276	.661	.26
International Staff	.429	-.816	-.167	-.910	-.148	-.9(
International Students	.672	.572	.175	-.861	.193	-.8(

Table 77 2006 WUR Principal Component Analysis

	Component Matrix			Pattern Matrix			
	Component 1 Loading	Component 2 Loading	Component 3 Loading	Component 1 Loading	Component 2 Loading	Component 3 Loading	Com 1 Lo
Peer Review	.494	.688		.170	.806		.2
Employer Review	.688	.259	.248	.367	.455	.387	.4
Staff/student Ratio	.276	-.190	.867	-.142		.950	
Citation	-.130	.806		-.323	.735	-.238	-.
International Staff	.729	-.273	-.460	.931		-.228	.8
International Students	.823	-.197	-.129	.818		.112	.8

	Component Score Coefficient		
	Component 1	Component 2	Component 3
Peer Review	.077	.569	.061
Employer Review	.196	.314	.331
Staff/student Ratio	-.066	-.057	.805
Citation	-.196	.530	-.208
International Staff	.510	-.065	-.183
International Students	.450	.014	.104

Table 78 2007 WUR Principal Component Analysis

	Component Matrix		Pattern Matrix		Structure Matrix	
	Component 1 Loading	Component 2 Loading	Component 1 Loading	Component 2 Loading	Component 1 Loading	Compo 2 Loa
Peer Review	.408	.696		.810		.80
Employer Review	.723	.372	.355	.699	.422	.73
Staff/student Ratio	.234	.209		.302		.30
Citation	-.148	.719	-.554	.538	-.502	.48
International Staff	.756	-.405	.851		.856	.13
International Students	.830	-.228	.804	.241	.827	.31

Table 79 2008 WUR Principal Component Analysis

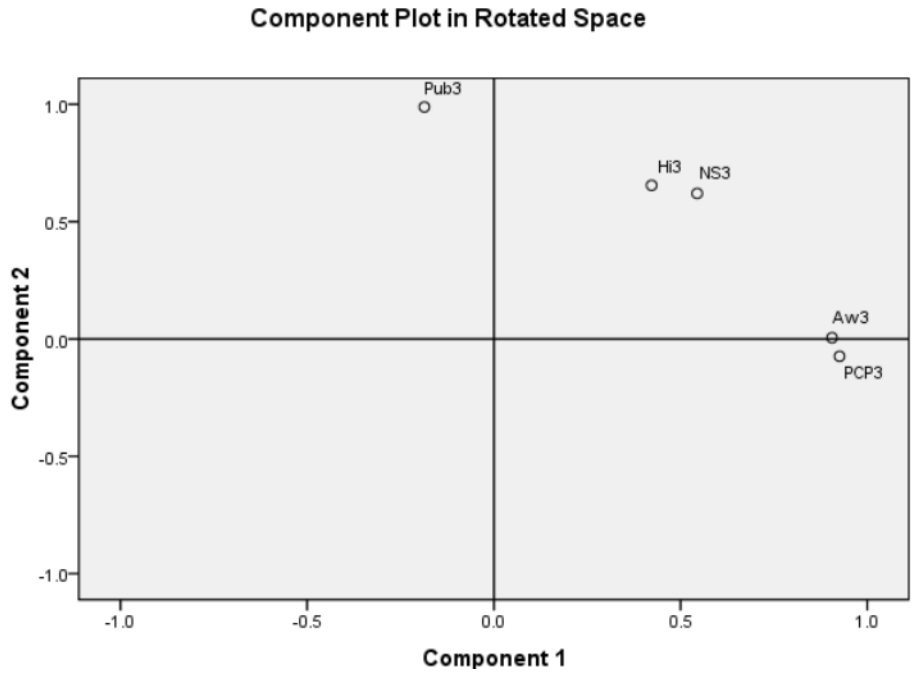


Figure 24 2003 SJTU Component Plot in Rotated Space

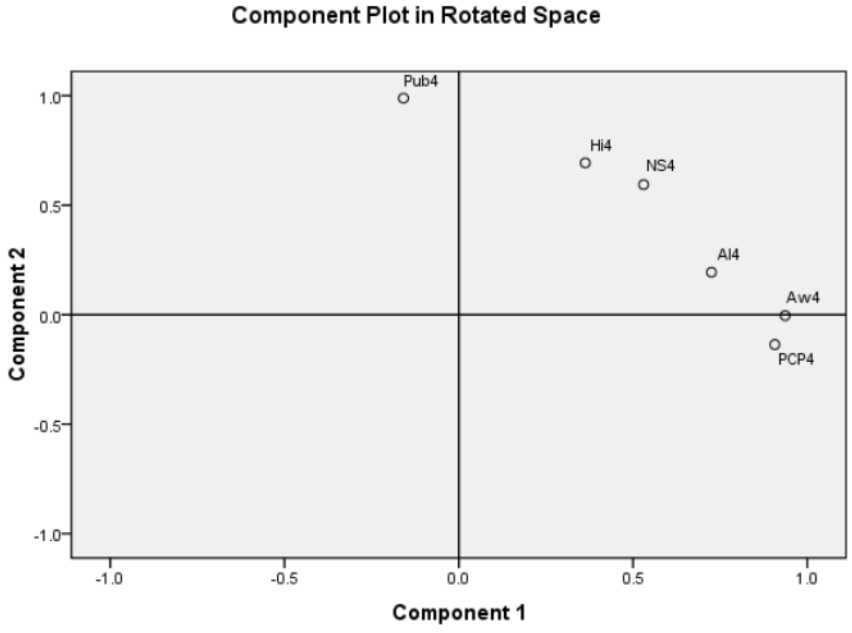


Figure 25 2004 SJTU Component Plot in Rotated Space

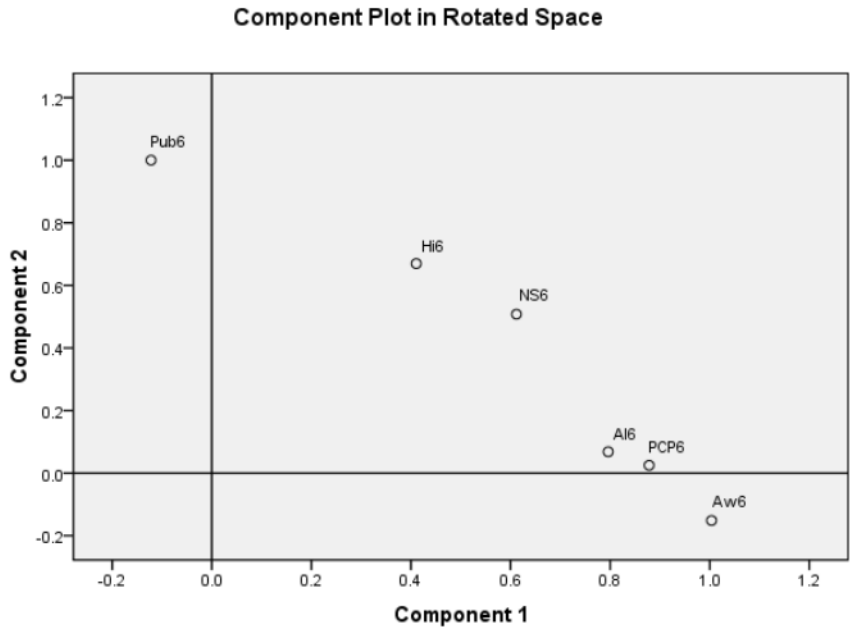


Figure 26 2006 SJTU Component Plot in Rotated Space

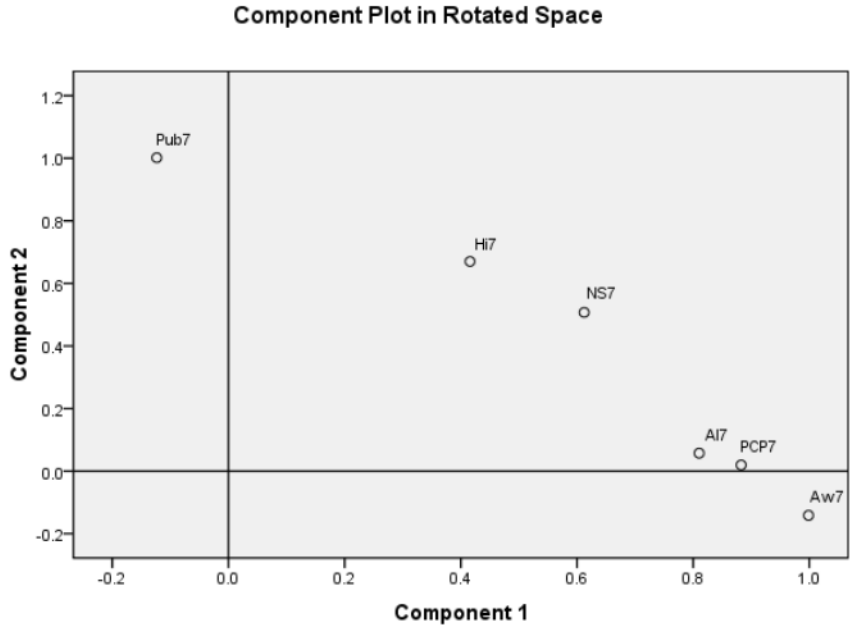


Figure 27 2007 SJTU Component Plot in Rotated Space

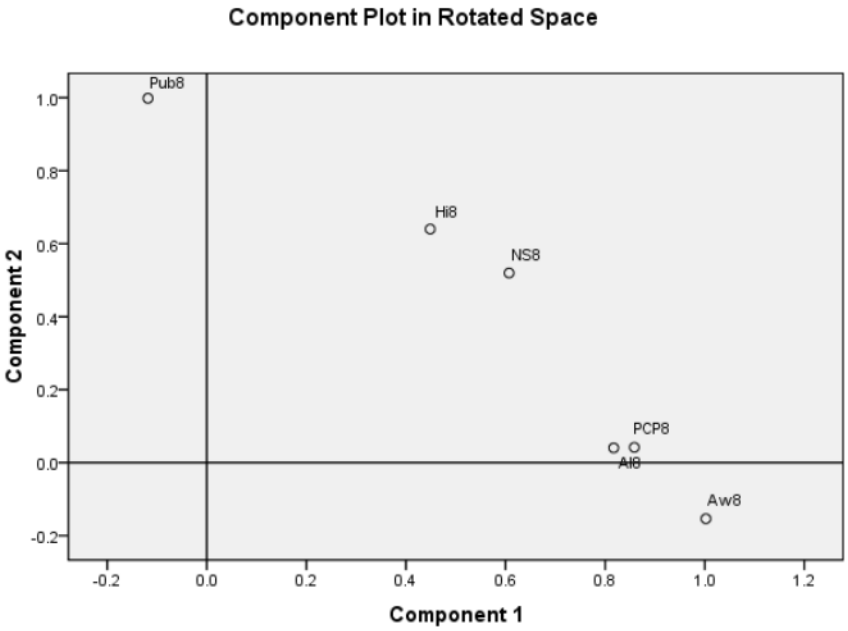


Figure 28 2008 SJTU Component Plot in Rotated Space

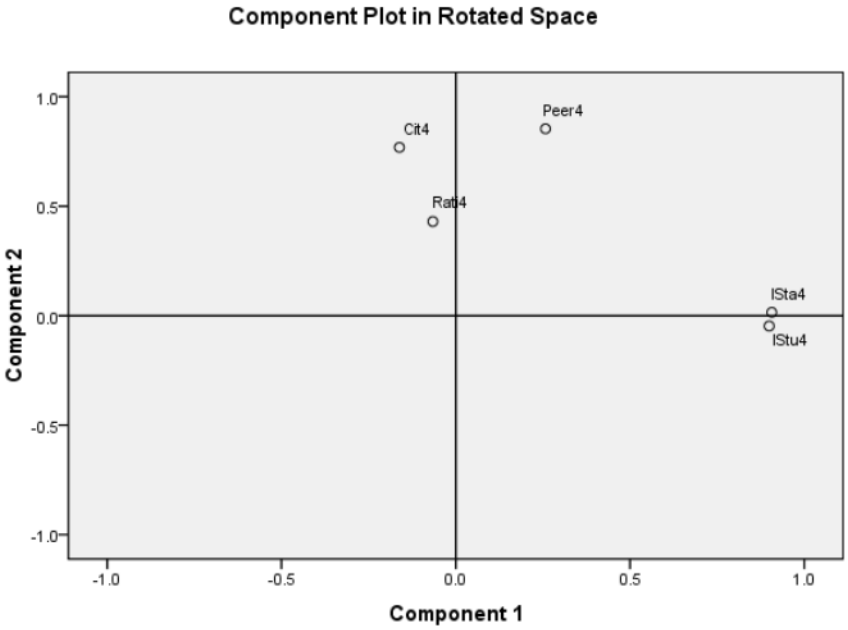


Figure 29 2004 WUR Component Plot in Rotated Space

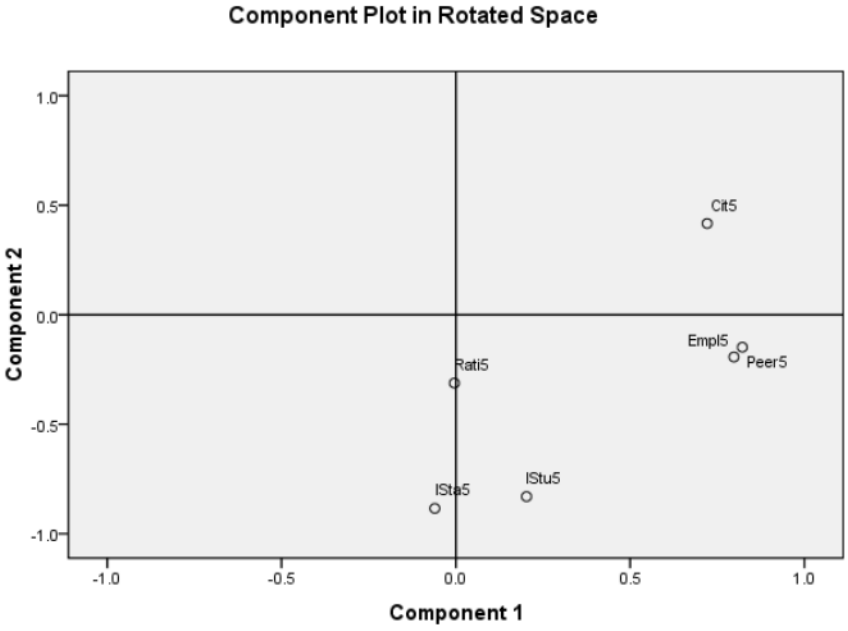


Figure 30 2005 WUR Component Plot in Rotated Space

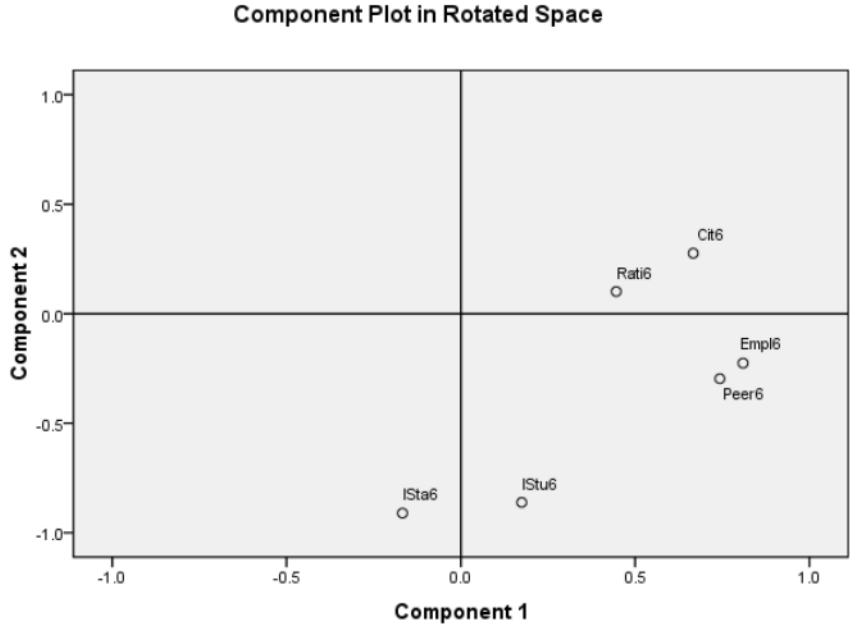


Figure 31 2006 WUR Component Plot in Rotated Space

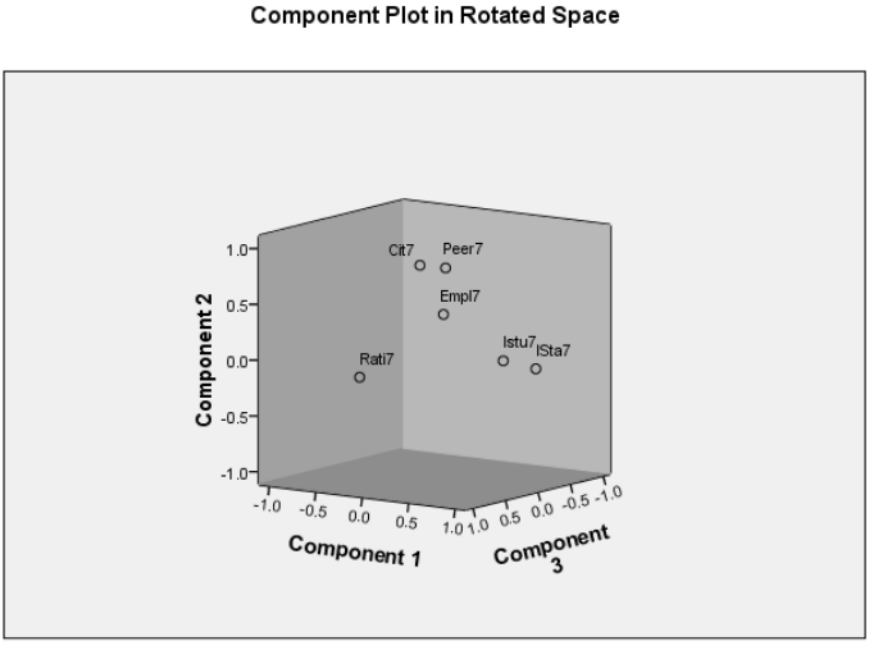


Figure 32 2007 WUR Component Plot in Rotated Space

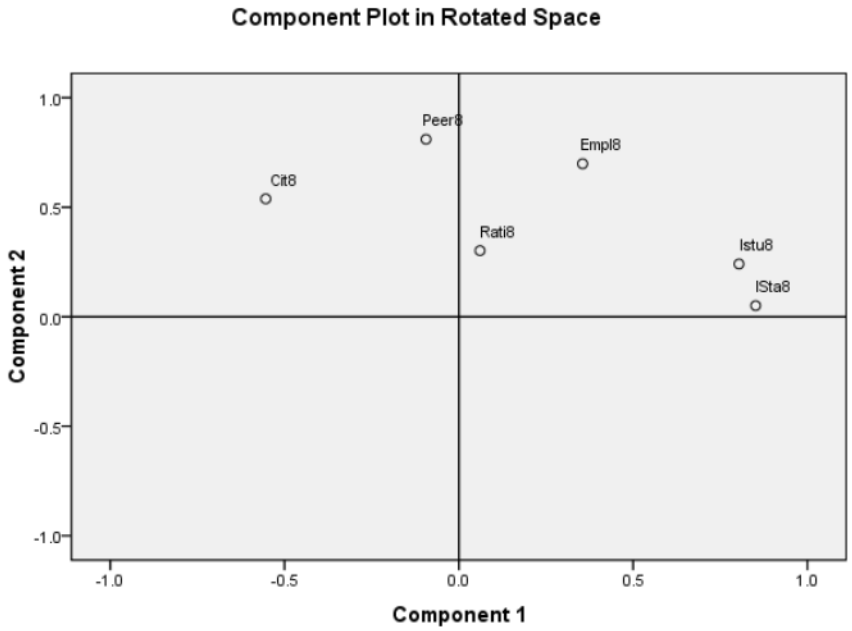


Figure 33 2008 WUR Component Plot in Rotated Space