

## THE IMPACT OF WEIGHTING PREFERENCES ON UNIVERSITY RANKINGS: THE EXAMPLE OF BULGARIA

ZAHARIEV, Boyan\*

BOYADJIEVA, Pepka

---

### **Abstract**

On the basis of data from the official Bulgarian University Ranking System (BURS launched in 2011 and updated in 2012), the selection of weights for different indicators in a university ranking system is discussed. Weights represent the importance assigned to the different dimensions of performance measured through ranking, and can therefore be regarded as quantified value statements. Weight preferences differ across groups of stakeholders or even from one agent to another. In this paper a number of stochastic simulations using 4 initial sets of weights are made. The actual weights applied to BURS—simple uniform weights—are derived from a survey encompassing 15,000 university students as well as weights determined by university rectors. The results show that rankings, whose value is relatively neutral—with a set of weights closer to uniformity—are more vulnerable to small random fluctuations. Value laden rankings with asymmetric weights as in the case of BURS are more resilient to slight disturbance(s) however are very unstable in the case of the rearrangement of priorities. A consistent ranking system demands asymmetric weighting, which means strong value statements and prioritisation. The problem however with such a ranking system is that if heavily weighted preferences do not capture important elements of performance, the whole system can become not just marginally imprecise but also very erroneous.

JEL classification: I21 Analysis of Education, I23 Higher Education and Research Institutions

*Key words:* university rankings, selection of indicators, stochastic weights

---

### **1. Introduction**

The Bulgarian system of higher education is changing. Changes include putting a bigger emphasis on the efficiency of the use of public funds. In 2011, the Ministry of Education commissioned a study on the elaboration of a university ranking system. This ranking system (<http://rsvu.mon.bg>) was developed by a consortium of private organisations, which included a policy institute, a public opinion agency, and an IT company. The system is deemed “official” in the strong sense as suggested by Codling and Lynn Meek (2006), since it is gradually becoming relevant to the way state funding for higher education institutions is allocated.

The main principles on which BURS is based are:

---

\* Boyan Zahariev is Programme Director at the Open Society Institute in Sofia, e-mail: [bzahariev@osi.bg](mailto:bzahariev@osi.bg) Pepka Boyadjieva is professor of sociology at the Bulgarian Academy of Sciences, e-mail: [pepka7@gmail.com](mailto:pepka7@gmail.com).

- *multidimensionality*: the ranking needs to take into account the main dimensions of quality in higher education – teaching, research and realisation;
- *differentiated approach* (“*compare the comparable*”): the ranking needs to take into account diversity in higher education via the ranking of professional areas instead of whole institutions;
- *empirical validity*: a combination of objective and subjective data habitually hail from different sources: administrative databases, student surveys, questionnaires completed by university rectors;
- *interactivity*: the online based system allows the users (prospective students, on-going students, employers, members of the academic community) to define their own preferences via the selection of indicators and weights;
- *flexibility*: the ranking system is open for improvement via the regular updating of data as well as via the modification of indicators and weights.

Initially the Bulgarian University Ranking System was designed as a compromise between the market-oriented paradigm of many first generation ranking systems - and the more diversity-based approach of the new generation ranking systems such as the European U-Multirank system.

The main feature of the Bulgarian university ranking is that it does not compare whole institutions but only professional areas. Hence if enough data are available for a particular university, it is included in as many different rankings as the number of professional areas represented in the university. In Bulgaria there are a total of 52 professional fields, 51 one of which are included in the university ranking. When a university has the objective of launching programmes in a new professional field, it must submit an application to the National Evaluation and Accreditation Agency for programme accreditation. At the beginning of 2012 there were 51 higher education institutions, 50 of which are included in the official ranking. In total there are 418 combinations of professional fields and higher education institutions. These are actually the units, which are ranked and for which data are collected. Metallurgy is only taught in one institution, so no ranking can be produced in this professional field yet all data are readily available on the website of the ranking system (BURS, 2012).

BURS tries to include all relevant dimensions of quality evaluation looking for a balance between realisation, teaching and research. The excessive emphasis on any particular dimension has been criticised, i.e. the focus on graduate employment and incomes (final outcomes) in some UK and South American rankings (Usher and Meadow, 2009) or the emphasis on research in Germany (Mayer and Ziegele, 2009) and in many of the global rankings.

A distinguishing feature of BURS is that along with the official ranking it also allows user defined rankings. Each user can choose their own set of preferred indicators and define their own weights. In this way it is practically possible to produce an unlimited number of rankings.

As in the case of other ranking systems, weights were one of the main targets for criticism in BURS. Various stakeholders were asked their opinion about the relative importance that should be given to the different aspects of quality of higher education.

The student survey and the questionnaire given to rectors and used in this analysis were part of the second edition of the ranking system. In the first edition a public opinion poll was also conducted. Despite a multi-stakeholder participation, the final set of weights was prepared by an expert-working group amid the consortium. This was due to the fact that stakeholder opinions also diverged widely. Discussions with university teachers and rectors have shown that even within this relatively homogeneous group, that consensus would be very difficult to reach. Discussions in the consultative bodies constituted by the Ministry of Education demonstrated that if the opinions of student organisations and representatives of the business were to be taken into account, that the task would become even more challenging.

Due to its design, BURS avoided the usual type of criticism for ranking entire institutions rather than faculties or programmes, but did not avoid criticism for the arbitrary selection of criteria and weights. It did however begin to shape the behaviour of institutions as has happened with other ranking systems around the world (OECD, 2006).

In this analysis we focus on the impact of the choice of weights on the university rankings. We create a stochastic disturbance in four different sets of weights: the official weights of the BURS, simple uniform weights, weights derived from a survey from among 15,000 university students, and weights arising from questionnaires carried out among 50 university rectors. The differences between these four groups of weights in terms of their resilience to small stochastic perturbations and to a general reshuffling of preferences and priorities are subsequently highlighted and explained.

## **2. Weights, preferences and value judgments in university rankings**

Weights are an important part of most university rankings, and are based on league tables. Despite the growing use of new approaches in for instance the case of the latent variable analysis, the most widely used approach in producing a single measure out of a multidimensional set of measures remains the weight-and-sum approach (Guarino *et al.*, 2005).

An alternative approach to league tables is the efficiency analysis, which has different version. In typical league tables only the “benefits” are measured so as to ensure that all indicators constitute the final score. When analysing efficiency, input indicators are expected to be minimised whilst output indicators are those that should be maximised for the sake of making an improvement. Data Envelopment Analysis is the most popular method employed to analyse university efficiency (for an example see Gonzalez-Perez *et al.* 2010). Data Envelopment Analysis was also used to analyse the performance of Bulgarian universities (Baikusheva *et al.*, 2011). Both input and output weights can be used in the efficiency analysis, so the concept of weighting remains applicable in this case as well.

A radical departure from the use of weights comes from the diversity approach, which insists that institutions of higher education have different missions. Therefore while it makes sense to divide them into separate categories, it makes less sense to make hierarchical comparisons (CHEPS 2008). This is the approach embraced by the pilot ranking of European universities U-Multirank. It can be argued however that people tend to frame their preferences hierarchically, in order for rankings to always be present in

consumer choice as noted by Codling (2006) whether they are chosen to be made explicit or to be ignored.

There are many questions related to the use of weights. The first is whether they should be constant or whether they must change with every edition of a weighed ranking. Bernardino and Marques (2009) point out that evaluation procedures including weights should be dynamic and change when society or the specific sector requires it. It seems that in most repeated rankings using weights, a degree of compromise is sought between stability and adaptability. BURS employs a set of weights which is determined through a series of expert discussions that were adjusted with feedback from stakeholders mainly between the first and the second edition of BURS. It is expected, however, that starting from the third year - the set of indicators and the weights will not change substantially. The predominant opinion among experts and decision makers in Bulgaria was that in order to set up incentives for development a ranking has to send clear signals to universities that do not change substantially from year to year.

The second question is how weights should be determined. The simplest and most straightforward approach is to use equal weights, which is often wrongly perceived as one that uses no weights at all. The second approach is to use expert assessment sometimes within the framework of a well-defined consultation procedure. This is how the official weights in BURS were determined. The third approach is to take-in the opinion of different stakeholders, such as students or members of the academic community. The fourth approach is to use certain statistical procedures that are independent of opinion – per se impartial. In this analysis a statistical procedure has been utilised so as to directly avoid the production of weights whilst ascertaining the sensitivity of other sets of weights to stochastic perturbations.

Expert opinion is probably the most widely used method of producing weights. This method can be very opaque and arguably unfair if the procedure is incoherent and if stakeholders are not consulted. Rector opinion is not used to determine weights for institutional rankings as it is assumed to be biased in favour of one's own institution. It is however very interesting to check whether this assumption is true and to conduct rankings based on rector-defined weights compared to other rankings. Student opinion occupies an important place in many rankings. It is of wide belief that this opinion is relevant for the assessment of the quality in higher education. Using student-determined weights can lead to the assumption that student choices are conditioned by attributes of institutions (Drewes, 2006) rather than by a number of factors that are unrelated to education.

### **3. Data and Methodology**

The data analysed originate in the Bulgarian University Ranking System (BURS). A table with the full description of all indicators used in the BURS is given in Annex I.

To rank the universities, BURS uses 47 indicators placed in 6 groups:

- 1) Teaching

The main indicator within the teaching group is the scores given by the National Agency for Assessment and Accreditation in the process of accrediting universities.

- 2) Research

Research production is assessed primarily on the basis of publication and citation indexes taken from Scopus by Elsevier. Other databases were not included during the first two years of the ranking system for technical reasons.

3) Educational environment

This section includes several indicators taken from university balance sheets showing the value of some categories of assets, the square metres of access per student to educational facilities, the number of library items per student, an assessment made by the students of the university library and the academic schedule.

4) Living conditions and administrative services

This group of indicators measure the availability of different administrative services, accommodation and scholarships.

5) Prestige

This category includes indicators measuring the reputation of the universities among candidate-students and students.

6) Realisation and linkages to the labour market

This group of indicators includes graduate incomes, chances for employment and the contribution of the university to the regional and national economy.

The data for the ranking system of Bulgarian universities originates from several sources:

- 1) The information system of the Ministry of Education Youth and Science. The system has a special database on higher education which goes by the name of AdminUni;
- 2) The National Social Security Institute, which shares with the National Revenue Agency a database on all people who work under different types of contracts and pay taxes and social security contributions;
- 3) The National Evaluation and Accreditation Agency;
- 4) Universities themselves through a special questionnaire sent to rectors;
- 5) A special survey among second and third year students at universities. In 2010, the first year of the ranking system, the survey was delivered to a sample of 10,000 students. In 2011, the sample was increased to 15,000 students.

BURS includes indicators in different measurement scales and with different units: monetary units, time measures, proportions, subjective evaluation scales among others. The aggregation of such diverse data usually undergoes the application of certain standardisation procedure(s). In BURS, raw values of indicators are standardised with the use of modified t-scores with an average of 50 and a standard deviation of 20. This procedure ensures that most of the score values fall in the range of 0 to 100. Scores above 100 and below 0 are truncated and take the value of 100 or 0 respectively.

To arrive at a final ranking, each group of indicators is given a weight determined by an expert group based on different sources of information. Each indicator within a group also has a specific weight. When the group weight is multiplied by the within-group weight of each indicator, a vector of individual weights is produced, the length of which

equals the number of indicators in the ranking system (47 indicators). The 47 weights add up to 1. The full list of indicators used in BURS with their definitions is given in Annex I. The official set of weights in the Bulgarian ranking system is a result of process of consultations combined with expert choices and decisions. The Bulgarian public debate is strongly driven by considerations for the growth of the economy, employment, increase in incomes and the quality of education and its adequacy for demand in the labour market. This atmosphere has influenced the process of prioritisation in the ranking system not as much through direct influence from business as by the set of values shared by experts who have participated in the design of the ranking system and of the policy priorities of the Ministry of Education as the commissioning body of the ranking system. At this stage, the question about what the real role of universities for the future salaries of graduates is not as relevant as the desire to push universities to do whatever may be possible to better respond to the needs of the labour market.

The set of weights produced by the rectors of Bulgarian universities is actually a vector of averaged weights given (more or less) independently by the rector of each university. The weights were taken from a questionnaire sent separately to each rector. Rectors were asked to complete a questionnaire whose main purpose was to provide additional data for the ranking system, which could not be collected from other sources. The questionnaire also included a section enquiring about the rector's opinion about the weights that should be assigned to each dimension (group of indicators) in the ranking system - based on the relative importance of this dimension for the evaluation of university performance. The questionnaire was completed by 50 rectors. Only one university was excluded because of its recent appearance on the scene (registered one year ago) and was therefore not yet part of the ranking.

Each rector may have discussed the choice of weights with certain of their colleagues yet there was no organised form of consultation or repeated feedback such as in the Delphi method. It therefore comes as no surprise that the weights produced by the rectors differed by a large margin. Typically it would be assumed that each rector had attempted to estimate which weights may benefit the overall ranking of their institution. As mentioned by Taylor (2007) since university presidents and vice-chancellors are humans, their opinions can be influenced 'by their own prejudices, loyalties and other positive and negative experiences'. This standard assumption about the behaviour of a rational agent can be further substantiated in the concrete case of rectors already being able to familiarise themselves with the 2011 rankings and obtain some information about the strengths and weaknesses of their university. Meanwhile, the Ministry of Education had started to experiment with using some of the indicators in the ranking system to differentiate funding.

As part of the making of BURS, a special questionnaire was delivered to 15,000 students sampled from each professional field in each university. Results from the questionnaire were used in the ranking for indicators with a total weight of 11%. In the first edition of the ranking system student opinion weighed more. There was however strong pressure from universities to reduce its weight and to use "objective" data from administrative databases. The questionnaire includes a question about the importance of each of the 6 dimensions (groups of indicators) included in the ranking system. Weights for the groups were then constructed by us based on the relative frequency of positive answers for each dimension. In other words, this equates to the relative number of

students who believed that this dimension is important for the evaluation of universities. Overall responses given by the students have shown that students tend to look benevolently on their own universities thereby attributing positive scores to them. Therefore we can suspect that students may have prioritised the dimensions of quality, thereby favouring their own university.

#### **4. Variables and Specifications of the Stochastic Simulation**

There are many ways to test if a system of weights used for a ranking is effective. One very important characteristic of a set of weights is their sensitivity to change. If a small change in the weights would produce a relatively large re-ranking, one could only argue on that ground that the ranking is unstable and is therefore not very reliable.

For the purpose of our analysis we generate random perturbations using two different methods around 4 initial vectors of weights. Both methods involve a Monte Carlo simulation, i.e. a large number of experiments based on random number generation, followed by the calculation of appropriate statistics across the experiments.

The first method uses uniform perturbations described in Benito and Romera (2011), who also give the precise algorithm. In principle their method is based on generating random uniform values within a given (small) radius around the initial weights. Not every simulation produces a usable vector of weights. If the vector is not deemed as “good” it is rejected and the procedure is iterated until the required number of usable vectors is produced.

The second method that we suggest involves the generation of random normally distributed perturbations around the initial vector of weights. The procedures include: a) generating around each weight in the initial vector a random normal disturbance with a standard deviation equal to 2/5 of the weight (negative values if any get 0 weight); b) normalising the new set of weights by dividing each of them to their sum.

The four vectors of weights are: a) the set of weights actually used in the ranking b) the set of uniform weights for all 47 indicators; c) the set of weights derived from a questionnaire completed by 15,000 university students; d) the set of weights derived from a questionnaire completed by 50 university rectors.

The small uniform perturbations proposed by the Benito and Romera (2011) are employed to represent the impact on a ranking system of small non-systemic fluctuations. Such fluctuations are actually equivalent to a kind of randomness in the underlying data when we keep weights fixed. If the chosen indicators with the chosen set of weights do not differentiate much, the whole system is in a delicate balance. Significant changes in rankings can occur even with a mere “puff of wind”. Such a system is to a degree unstable yet on the other hand it gives the chance for those lagging behind to make pace. Small uniform perturbations cannot move the system far away from the initial choice of priorities. If in the initial set there was no obvious priority, i.e. if we were presented with a relative large number of indicators with equal weights, the disturbed system will look broadly the same.

Nonetheless, what if we would like to model a random change in priorities. In this scenario, we require a significant disturbance, beyond the minimum weights. Since by definition in this case we are dealing with an asymmetric system with strong preferences reflected in a large difference across weights, a significant reshuffling would require

disturbances that are proportional to the initial weights. We model this situation by generating normally distributed perturbations proportional to each weight. Applying this procedure to a set of equal weights we can think of this procedure as the result of the action of an (slightly awkward) agent who would like to prioritise but does not care exactly which values or dimensions of assessment will be prioritised. In fact, this situation is not completely irrelevant. In the Bulgarian policy debate there have been statements of the type of “when you don’t have priorities, you are wasting resources, so please choose whatever is important for you and pool resources there”. This type of recommendation was given for example when funding for research was discussed.

For both types of perturbations we calculate a sequence of simple statistics given in Table I. The first column gives a list of the 8 cases of rankings that we analyse: the four types of weights with two different perturbations (uniform and normally distributed). In column 2, the figure 418 represents the number of ranked units. These are actually all the unique pairs between professional fields and universities. For example, general engineering is taught in 15 different institutions, so these are 15 ranked units. The professional field of Architecture, construction and geodesy is taught in 8 institutions therefore the total number of ranked units for the two professional fields amounts to 23. Adding up all professional fields in this way provides 418 ranked units. The range is the maximum observed change in the rank of any professional field in any institution observed across 100 random simulations (the minimum observed change is 0 across all units). The sum represents the sum of maximum changes (ranges) in rank for all 418 ranked units across the 100 simulations. The mean demonstrates the mean change in rank observed across the random simulations. This is simply the sum divided by the number of rankings ( $N=418$ ).

Table I. Descriptive statistics for four sets of weights with two different types of perturbations

|                                                                                                    | <b>N</b> | <b>Range</b> | <b>Sum</b> | <b>Mean</b> |
|----------------------------------------------------------------------------------------------------|----------|--------------|------------|-------------|
| <b>1. Uniform weights with normally distributed perturbations</b>                                  | 418      | 14           | 921        | 2.20        |
| <b>2. Uniform weights with uniform perturbations</b>                                               | 418      | 17           | 1245       | 2.98        |
| <b>3. Weights used in the Bulgarian university ranking with normally distributed perturbations</b> | 418      | 17           | 1083       | 2.59        |
| <b>4. Weights used in the Bulgarian university ranking with uniform perturbations</b>              | 418      | 4            | 108        | .26         |
| <b>5. Weights determined by students with normally distributed perturbations</b>                   | 418      | 11           | 955        | 2.28        |
| <b>6. Weights determined by students with uniform perturbations</b>                                | 418      | 8            | 632        | 1.51        |
| <b>7. Weights determined by rectors with normally distributed perturbations</b>                    | 418      | 11           | 921        | 2.20        |
| <b>8. Weights determined by rectors with uniform perturbations</b>                                 | 418      | 13           | 1021       | 2.44        |

Finally, we also produced rankings based on teach rector’s own set of weights. This procedure produced 50 non-random rankings, which were then compared to the rankings

based on stochastic perturbations. For 5 professional areas, we also verified to what extent each rector's own weights improved the ranking of their own university.

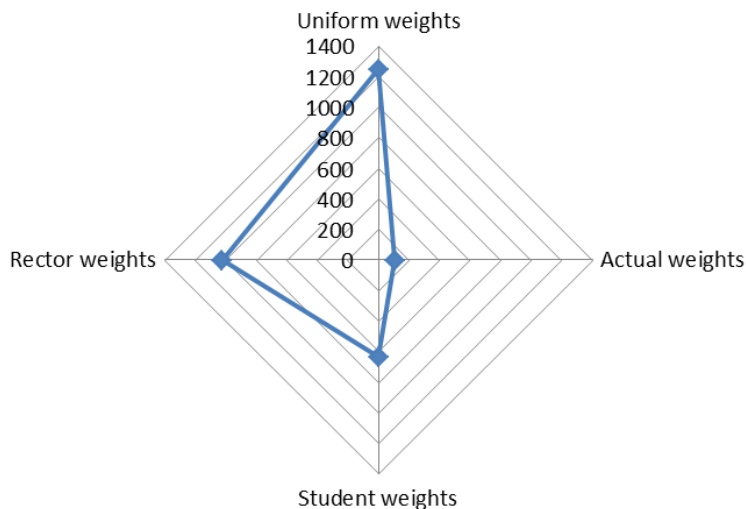
## 5. Analysis of Results

It can be imagined that for the purpose of our analysis that there is some finite number of independent measurable indicators, which truly characterise the performance of universities, or indeed of any units we would assess. Then a specific ranking should choose from this imaginary full set a subset of indicators based on the specific aspects of performance, which the authors or the users of the ranking would like to measure. What they subsequently have to do is to attribute weights to each indicator from the full set. All the indicators that are considered irrelevant and completely ignored are attributed a zero weight, and others which are considered relevant however unimportant receive small weights that can influence the ranking only *ceteris paribus*. The most important indicators are heavily weighted. In this framework picking up indicators for the ranking and assigning weights are parts of one and the same process, i.e. the process of measuring the valuable instead of valuing the measurable (HEFCE, 2008).

After generating perturbations we can look directly at the scores (from 0 to 100) produced in each of the 100 simulations. In this case, the standard deviation within each vector of scores is an appropriate measure of the impact of this particular disturbance. Standard deviations for the uniform and normally distributed perturbations are provided below. Note on Chart I, in the Annex, the difference in the left scale (for uniform perturbations) and right scale (for normal perturbations).

For each professional field in each university we can calculate the lowest and the highest rank with each specific type of disturbance. Subtracting the lowest from the highest rank we obtain the range for each type of disturbance. The sum of the 418 ranges obtained in this way is one of the possible measures of how much a disturbance affects rankings based on a specific vector of weights. On Chart II we have shown the application of the described procedure for aggregating ranges using small uniform perturbations.

Chart II. Sum of ranking range with small uniform perturbations

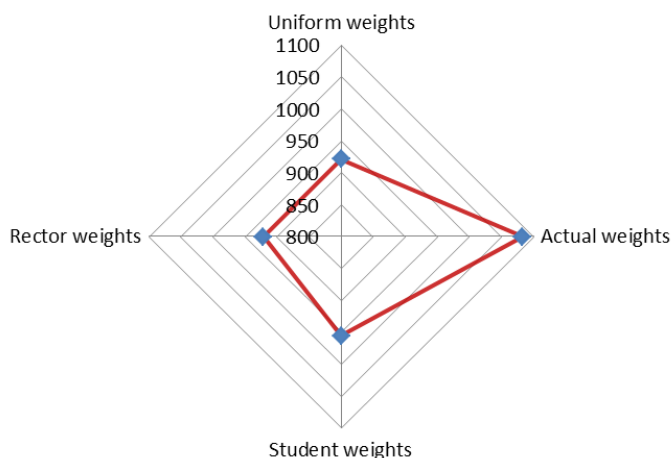


As seen from Chart II with uniform perturbations, the actual weights appear most stable. The meaning of this result could be that, asymmetrical as they are, the actual weights express certain very strong preferences. Such weights are unlikely to be disturbed very much by merely allowing them to fluctuate to a small degree around their initial values. We can say that a ranking system based on strong preferences is thus not very sensitive to non-systemic random fluctuations, i.e. fluctuations that do not challenge very much the ranking of priorities.

Just the opposite is true in the case of a ranking system with asymmetric weights or with very few indicators, which can be seen as equivalent from the point of view preferences. Such rankings are the result of a value-laden process whereby indicators are selected.

In the case of normally distributed perturbations we see a completely different picture as shown on Chart III.

Chart III. Sum of ranking range with normally distributed perturbations proportional to each weight



Obviously the actual weights used in the Bulgarian ranking system are most unstable to this kind of large systemic (we can call it *shattering*) disturbance. Unsurprisingly, uniform weights are resilient to this type of disturbance. Yet the same can be said of rector weights and student weights. We will try to explain this phenomenon.

The examples of concrete professional fields are even more revealing. In social work, as shown on Chart IV below in the Annex, the official ranking has 5 clusters of universities, which are also clearly distinguishable in rankings using student determined weights even after introducing small uniform perturbations.

The following Chart V, in the Annex, shows that the five clusters are still there even with a larger disturbance but normal perturbations have started to blur a bit the distinctions between the top 3 clusters. If we allow for a bit wider intervals of disturbance they will likely merge in a cloud and each successive trial will produce a different ranking.

The next chart in the Annex (Chart VI) demonstrates that the rankings based on preferences of university rector's have the effect of large random perturbations. It keeps more or less the three distinguishable groups but boundaries appear to be even fuzzier than in the case of the random simulation with normally distributed perturbations.

The same pattern is confirmed by example from other professional areas. Charts illustrating similar rankings for the most popular professional area of economics are given in Annex II.

We also produced rankings in 4 professional fields taking each rector's set of weights: a) Theory and management of education; b) Political science; c) Architecture, construction and geodesy; d) General engineering. The results for Architecture, construction and geodesy are shown on Chart VII below.

Chart VII. Differences in t-scores for 8 institutions teaching Architecture, construction and geodesy based on each rector's own weights and the average weights given by him and all of his colleagues.

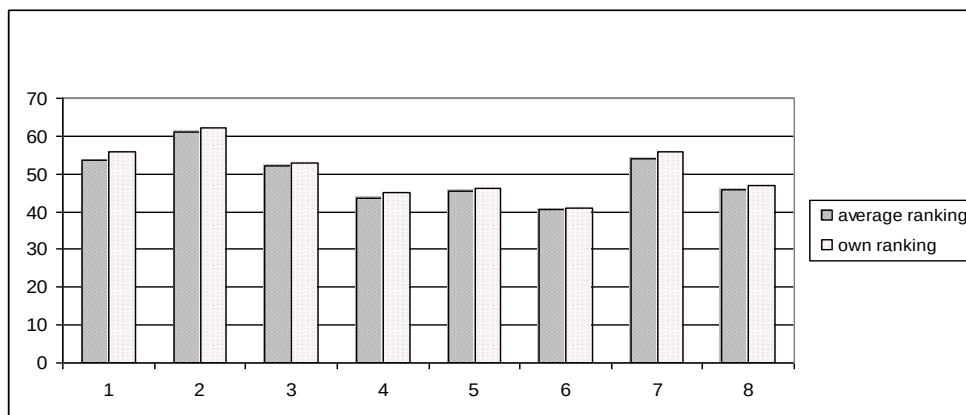
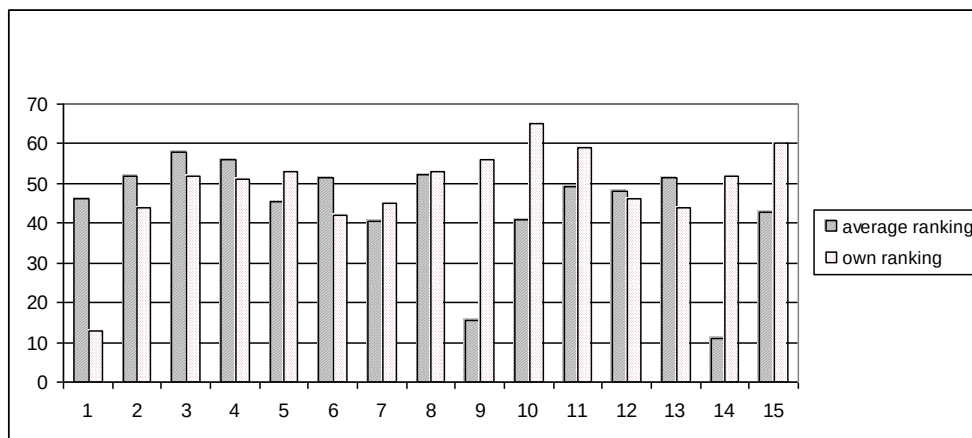


Chart VIII shows the same results for General engineering.

Chart VIII. Differences in t-scores for 15 institutions teaching General engineering based on each rector's own weights and the average weights given by him and all of his colleagues.



The data give some evidence that the set of weights chosen by each rector on average does benefit their own institutions. However the net effect is quite small and certain rectors are actually to an extent mistaken, meaning that their preferred set of weights actually produces a lower overall score for their institution than the averaged weights of all their colleagues or for the weights of the actual ranking. It seems that it is not that easy to manipulate a ranking system with a large set of indicators by picking up weights.

But we did not observe any systemic differences between the average ranking using the rector determined weights and the ranking of each university on the basis of the set of weights chosen by its management.

It is a very plausible assumption that university rectors would try to secure a better position for their universities in the ranking system. This can be done in two ways. Firstly by improving the performance of the university assuming that the ranking system itself is a true and fair measure of performance. Secondly they can try to somehow manipulate the figures capturing the performance of their university. Finally, they can try to directly influence the design of the ranking system, e.g. by changing the set of indicators used or by changing the weights assigned to different indicators. This last option is very limited or is not available at all when the ranking is produced by an independent agency, which does not have to consider the opinion expressed by the management or academic staff of the universities.

But in the case of BURS, the ranking system is commissioned by the Ministry of Education, which means that it is part of a complex policymaking process. It has to be discussed with a variety of stakeholders including university rectors both at the stage of designing the methodology and at that of presenting the results. This has many benefits in terms of improving the legitimacy of the ranking, but it also opens the way to attempts at manipulating the ranking.

## **6. Conclusions and final reflections**

In this paper, we use certain stochastic simulations using 4 initial set of weights: the actual weights applied in BURS, simple uniform weights, weights derived from a survey among 15,000 university students and weights determined by university rectors. The results show that rankings, which are relatively value neutral, with a set of weights closer to uniformity, are more vulnerable to small random fluctuations. Value laden rankings with asymmetric weights (such as in BURS) are more resilient to small disturbances but very unstable in the case of the rearrangement of priorities. A consistent ranking system demands asymmetric weighting, which means strong value statements and prioritisation. The problem with such a ranking system is however that if heavily weighted preferences do not capture important aspects of performance, the whole system not only becomes slightly imprecise but very wrong.

The rectors of big universities try to respond to the challenges of the new ranking system by thinking of how to maximise their scores. One option to maximise scores is to apply more value on the dimensions of the assessment they are better in. BURS provides some evidence that rectors might have tried to act in this way. Indeed the rectors of some large and influential universities had the chance to influence weights, albeit in a limited way and in a complicated game of public bargaining through their participation in committees, media activity and via pressure on the Ministry of Education. University rectors and department and programme management at universities became most active in

the days after the publishing of the official ranking by making statements to the media, participating in media events and expert discussions and sending letters with questions, comments and objections to the Ministry of Education and the private consortium producing the ranking. The choice of indicators and the set of weights have been among the most discussed topics. They can actually be viewed as two aspects of one and the same problem: choosing relevant aspects of quality to be evaluated from the finite number of possible measures.

However, weights picked-up by rectors failed to systematically improve the scores of their own institution. Although on average universities receive slightly better scores if the weights suggested by their management are used, the effect is not large enough to conclude that this strategy would actually work in practice. Some institutions actually receive worse scores with their management's own weighting of the ranking dimensions. Either university rectors have had a different motivation, i.e. to suggest what they believe are "fair" weights reflecting the true importance of the different dimensions of quality, or they have not been very successful in trying to guess which weights would improve their institution's score.

It should also be noted that the Bulgarian ranking system is based on normalised scores. It does therefore not provide clear benchmarks for the management of universities. If a rector has the chance to suggest a set of weights, which would maximise her institution's score, she is not required to only answer questions such as *what is the biggest, second biggest etc. strength of the institution*. The rector also needs to know "the competition" and answer questions with much more parameters, which also require information about the relative strengths of other institutions in each dimension.

It is also interesting to note that rankings based on rector weights bear a much closer similarity to stochastic rankings with normally distributed perturbations proportional to each weight, than to small uniform perturbations. This might be interpreted as an indication that among managers of universities, and probably in a broader sense within the academic community, at least at the starting point prior to any discussions and consultations, there is not much agreement on the importance that should be assigned to various dimensions of quality in higher education. In all probability there is no agreement even on what quality is about. Thus, for good or for bad, the making of a ranking system becomes an even more responsible task, since it has its chance to reshape higher education.

## 7. Bibliography

- Baikusheva, A., Zahariev, B. *et al.* (2011), 'Efektivnost na publichno finansiranite visshi uchilishta v Balgaria: prilozhenie na analiza na granitsata na efektivnostta' ('Efficiency of publicly funded higher education institutions in Bulgaria: an application of DEA') in *The quality of higher education in national and global context*, Kritika i Humanizam, pp. 69-101 (in Bulgarian).
- Benito, M. and Romera, R. (2011), 'Improving quality assessment of composite indicators in university rankings: a case study of French and German universities of excellence', *Universidad Carlos III de Madrid Statistics and Econometrics Series 15* **15 (11-20)**, 1-23.
- Bernardino, P. and Marques, R. (2009), 'Academic rankings: an approach to a Portuguese ranking', *MPRA Paper* **17297**, pp. 1-22.

CHEPS (2008), *Mapping Diversity: Developing a European Classification of Higher Education Institutions*, CHEPS, University of Twente, p. 11.

Codling, A. and Lynn Meek, V. (2006), 'Twelve Propositions on Diversity in Higher Education', *Higher Education Management and Policy* **18**, No. 3, p.17.

Drewes, T. and Michael, C. (2006), 'How d students choose a university? An analysis of applications to universities in Ontario, Canada', *Research in Higher Education* **47**, No. 7, p. 783.

Eurostat (2009), *Key Data on Education in Europe 2009*, European Commission.

González-Pérez, Begoña, López-González, E and Mendaña-Cuervo, C. (2010), 'La eficiencia de la rama de ciencias sociales de las universidades públicas españolas', *Regional and Sectoral Economic Studies* **10(3)**, pp. 117-128.

Guarino, C., Ridgeway G., Chun, M. and Buddin, R. (2005), 'Latent Variables Analysis', *Higher Education in Europe*, pp. 147-165.

HEFCE (2008), "Counting What Is Measured or Measuring What Counts? League Tables and Their Impact on Higher Education Institutions in England", HEFCE Issues Paper 14, April, [www.hefce.ac.uk/pubs/hefce/2008/08\\_14/](http://www.hefce.ac.uk/pubs/hefce/2008/08_14/).

Mayer, P. and Ziegele, F. (2009), 'Competition, Autonomy and New Thinking: Transformation of Higher Education in Federal Germany', *Higher Education Management and Policy*, pp. 1-20.

OECD (2006), *Education Policy Analysis: Focus on Higher Education*, OECD, p. 22.

Paxton, P. and Bollen, K. (2003), 'Perceived Quality and Methodology in Graduate Department Ratings: Sociology, Political Science, and Economics', *Sociology of Education* **76**, No. 1, pp. 71-88.

Taylor, P. and Braddock. R. (2007), 'International University Ranking Systems', *Journal of Higher Education Policy and Management* **27**, No. 3, pp. 245-260, p. 248.

Usher, A. and Medow, J. (2009), "chapter A Global Survey of University Rankings and League Tables". In: Kehm, B. and Stensaker, B. (ed.) (2009), *University Rankings, Diversity, and the New Landscape of Higher Education*, Sense Publishers, pp. 3-18.

Usher, A. and Savino, M. (2007), 'A Global Survey of University Ranking and League Tables', *Higher Education in Europe* **32**, No. 1.

West, P. (2009), 'A Faustian Bargain? Institutional Responses to National and International Rankings', *Higher Education Management and Policy* **21**, No. 1, 1-10.

## ANNEX I. List of indicators in the Bulgarian University Ranking System

| Indicator number | Indicator                          | Definition                                                                                                                                                                                                                                                                                                                                                                   |
|------------------|------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1                | Accreditation assessment grade     | National Evaluation and Accreditation Agency's assessment of the higher education institution's professional field on a scale from 0 to 10.                                                                                                                                                                                                                                  |
| 2                | Theoretical knowledge              | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on the diversity of courses included in the curriculum, the way they are taught, and the relevance of the content taught in the particular professional field.                                                                                                                       |
| 3                | Level of practical skills acquired | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on the lectures' practical relevance, the quality of practice-oriented activities, the opportunities for fieldwork and for consulting foreign language publications related to the studies in the particular professional field.                                                     |
| 4                | Participation in internships       | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on participation in internships, level of satisfaction from internships offered and the possibility to apply in practice what has been learnt in the particular professional field.                                                                                                  |
| 5                | Teaching                           | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on the approach to and quality of teaching in the particular professional field, the possibility for establishing a personal contact with and receiving adequate feedback from professors, and the motivation for further research and/or continuing education into a higher degree. |
| 6                | Student assessment                 | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on the availability of clear student assessment criteria and procedures, and their application by professors in the particular professional field.                                                                                                                                   |
| 7                | Student workload                   | Average number of hours spent by one student on lectures, seminars and independent learning per week. Data is obtained through a survey (Oct 2011 - Feb 2012).                                                                                                                                                                                                               |
| 8                | Exclusive full time academic staff | Share of full-time academic staff on a permanent employment contract only in the particular higher education institution in the total number of permanently employed academic staff at the higher education institution. Data from AdminUni as of January 2012.                                                                                                              |
| 9                | International mobility             | Number of students in the particular professional field who have participated in a student mobility program (Erasmus, Socrates, etc.) since the beginning of 2008 to December 2011. Data provided by the higher education institutions as of December 2011.                                                                                                                  |
| 10               | Bachelor level majors              | Number of professional bachelor and bachelor level majors offered in the particular professional field. Data provided by AdminUni as of January 2012.                                                                                                                                                                                                                        |
| 11               | Master level majors                | Number of Master level majors offered in the particular professional field. Data provided by AdminUni as of January 2012.                                                                                                                                                                                                                                                    |
| 12               | Regulated majors                   | Number of majors from the regulated professions list offered in the particular professional field. Data provided by AdminUni as of January 2012.                                                                                                                                                                                                                             |
| 13               | University                         | University citation index according to the international SCOPUS                                                                                                                                                                                                                                                                                                              |

| Indicator number | Indicator                                                         | Definition                                                                                                                                                                                                                                                                                             |
|------------------|-------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  | citation index                                                    | database for the period 2007–2011 as of December 2011 for the following publications: article, conference paper, review and book                                                                                                                                                                       |
| 14               | Citation index by scientific area                                 | Citation index according to the international SCOPUS database as of December 2011 for the period 2007–2011 in the scientific area to which the particular professional field belongs.                                                                                                                  |
| 15               | Citation index by scientific area, excluding self-citations       | Citation index excluding self-citations according to the international SCOPUS database as of December 2011 for the period 2007–2011 in the scientific area to which the particular professional field belongs.                                                                                         |
| 16               | Number of citations per paper                                     | Average number of citations per paper according to the international SCOPUS database as of December 2011 for the period 2007–2011 in the scientific area to which the particular professional field belongs                                                                                            |
| 17               | Papers with at least one citation                                 | Number of papers cited at least once according to the international SCOPUS database as of December 2011 for the period 2007–2011 in the scientific area to which the particular professional field belongs.                                                                                            |
| 18               | Articles in academic journals                                     | Number of articles in scientific journals according to the international SCOPUS database for the period 2007–2011 as of December 2011 in the professional fields included in the particular scientific area.                                                                                           |
| 19               | Student involvement in science and research                       | Percentage of surveyed students (Oct 2011 - Dec 2012) who have declared that they have participated in at least one science and research activity: research project, conference, writing and presenting a paper, writing and publishing articles or other papers in the particular professional field. |
| 20               | Total funds for science and research per student                  | Total funds for science and research, including government subsidy and additional funds attracted by university, per student (in leva). Data cover actual cash inflow for calendar year 2010. Data provided by the higher education institutions as of December 2011.                                  |
| 21               | Funds for science and research from non-state sources per student | Total number of additional funds (apart from government subsidy) attracted by the university for science and research per student, per year (in leva). Data cover actual cash inflow for calendar year 2010. Data provided by the higher education institutions as of December 2011.                   |
| 22               | PhD programmes in the professional field                          | Number of accredited Ph.D. programs in the particular professional field. Data provided by NEAA as of January 2012.                                                                                                                                                                                    |
| 23               | PhD programmes in the university                                  | Number of accredited Ph.D. programs in the particular university. Data provided by NEAA as of January 2012.                                                                                                                                                                                            |
| 24               | Doctoral-to-undergraduate and graduate students ratio             | Number of doctoral students per 100 undergraduate and graduate students in the particular professional field. Data provided by AdminUni as of January 2012.                                                                                                                                            |
| 25               | Assessment of material resources and infrastructure               | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on the condition of auditoria and technical equipment in the particular professional field, the access to computer labs and free internet at the university.                                                   |
| 26               | Fixtures and fittings                                             | Accounting value of the university's fixtures and fittings per student for 2010, in leva. Data provided by the higher education                                                                                                                                                                        |

| Indicator number | Indicator                                      | Definition                                                                                                                                                                                                                                                                                                                       |
|------------------|------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  |                                                | institutions as of December 2011.                                                                                                                                                                                                                                                                                                |
| 27               | Equipment                                      | Accounting value of the university's hardware, software and other equipment, machines, appliances and vehicles per student for 2010, in leva. Data provided by the higher education institutions as of December 2011.                                                                                                            |
| 28               | Class schedule                                 | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on the convenience and relevance of the class schedule in the particular professional field.                                                                                                                                             |
| 29               | Library stock per student                      | Number of library units per student available at the university. Data provided by the higher education institutions as of December 2011.                                                                                                                                                                                         |
| 30               | Library stock assessment                       | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on the convenience of the library's working hours, the availability of necessary literature, the access to and use of electronic databases in the particular professional field.                                                         |
| 31               | Library usage                                  | Number of library unit borrowings per student at the university for calendar year 2010. Data provided by the higher education institutions as of December 2011.                                                                                                                                                                  |
| 32               | Availability of information sources            | Number of international databases for which the university has a subscription. Data provided by the higher education institutions as of December 2011.                                                                                                                                                                           |
| 33               | Teaching premises per student                  | Teaching premises available per student at the university in square meters. Data provided by the higher education institutions as of December 2011.                                                                                                                                                                              |
| 34               | Scholarships                                   | Average university spending on scholarships per full-time student in the particular professional field for calendar year 2010 in leva. Data provided by the higher education institutions as of December 2011.                                                                                                                   |
| 35               | Student accommodation                          | Share of the full-time students at the particular university, using student housing. Data provided by the higher education institutions as of December 2011.                                                                                                                                                                     |
| 36               | Satisfaction with social and living conditions | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) of the social and living conditions, sports facilities and hygiene in the particular university, as well as the living and studying conditions at student housing facilities and the access to student canteens.                         |
| 37               | Assessment of administrative services          | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on working hours and efficiency of administrative service departments and the availability of and access to electronic administrative services in the particular professional field.                                                     |
| 38               | Career development support                     | Students' assessment on a 0-to-10 scale obtained through a survey (Oct 2011 - Feb 2012) on the availability and operation of a career development centre at the particular university, the opportunity to meet employers and alumni and to obtain information on existing career possibilities in the chosen professional field. |
| 39               | Secondary education diploma GPA                | Secondary education diploma GPA of the students in the particular professional field. Data provided by AdminUni as of January 2012.                                                                                                                                                                                              |
| 40               | Prestige among                                 | Percentage of students in the particular professional field who                                                                                                                                                                                                                                                                  |

| Indicator number | Indicator                                  | Definition                                                                                                                                                                                                                                                                                 |
|------------------|--------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                  | students                                   | have selected in a survey (Oct 2011 - Feb 2012) the university offering the best education in their professional field. Student's assessment of their own universities is not taken into consideration.                                                                                    |
| 41               | First preference                           | Percentage (0 to 100) of students in the particular professional field who have agreed in a survey (Oct 2011 - Feb 2012) with the statement: I was enrolled in the university of my FIRST preference.                                                                                      |
| 42               | Foreign students                           | Percentage (0 to 100) of foreign students in the total number of students in the particular professional field. Data provided by AdminUni as of January 2012.                                                                                                                              |
| 43               | Graduates' insurance income                | Average monthly insurance income of Bulgarian university graduates from the last 5 years for the particular professional field. Data is provided by National Social Security Institute as of November 2011.                                                                                |
| 44               | Unemployment among graduates               | Percentage of registered unemployed persons among those who have graduated from the particular professional field in the particular university in the last 5 academic years. Data is provided by National Social Security Institute as of November 2011.                                   |
| 45               | Applicability of degree acquired           | Percentage of employed persons among those who have graduated from the particular professional field in the particular university in the last 5 years, and are working in an area relevant to their education. Data is provided by National Social Security Institute as of November 2011. |
| 46               | Contribution to the social security system | Percentage of persons who make social insurance contributions in Bulgaria among those who have graduated from the particular professional field in the particular university in the last 5 years, according to NSSI data as of November 2011.                                              |
| 47               | Regional importance                        | Percentage of students in the particular professional field who plan to seek employment in the town or the region where their university is located. Data obtained through a survey (Oct 2011 - Feb 2012).                                                                                 |

## ANNEX II.

Chart I. Standard deviation of university scores obtained with small uniform (left scale) and large normally distributed (right scale) perturbations on the actual weights used in the Bulgarian University ranking system.

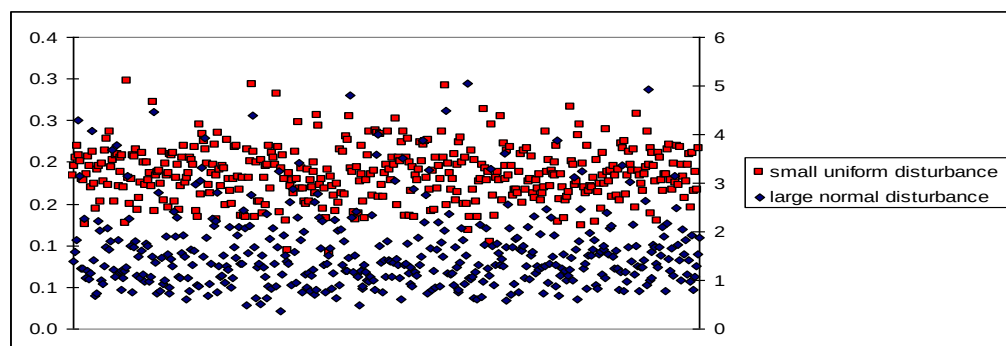


Chart IV. Ranking of 14 universities in *social work* with weights derived from a student survey (100 rankings with fixed width uniform perturbations)

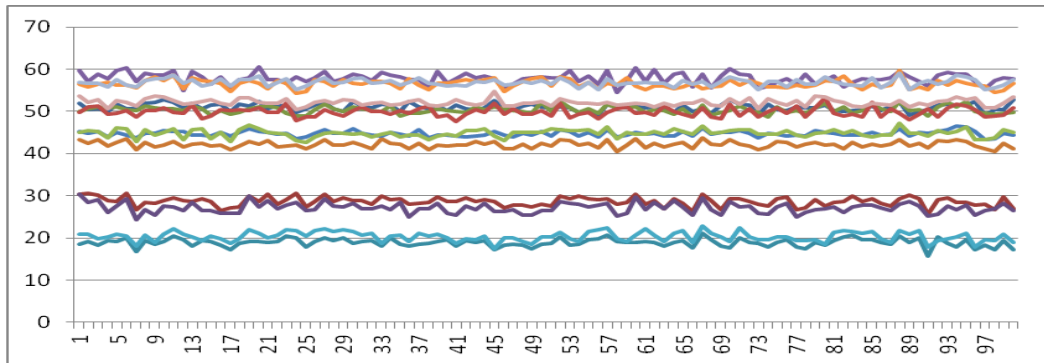


Chart V. Ranking of 14 universities in *social work* with weights derived from a student survey (100 rankings with normal perturbations, proportional to each weight)

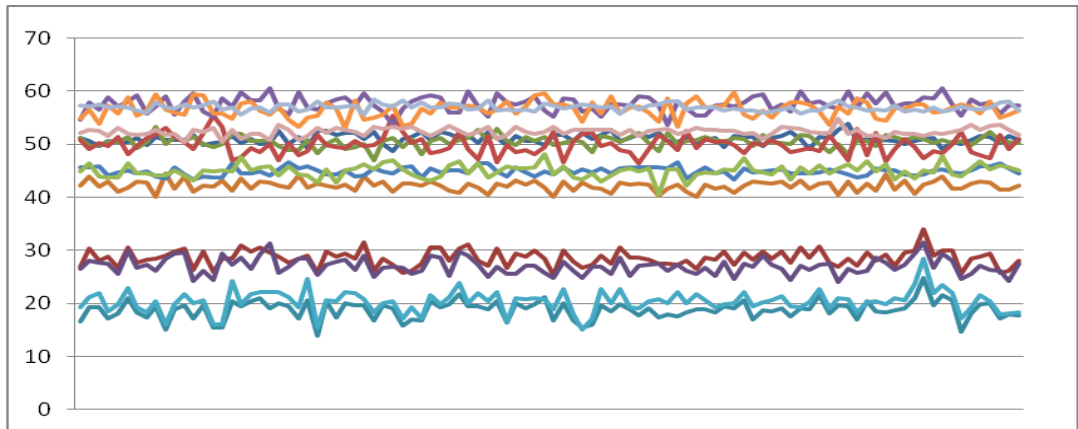
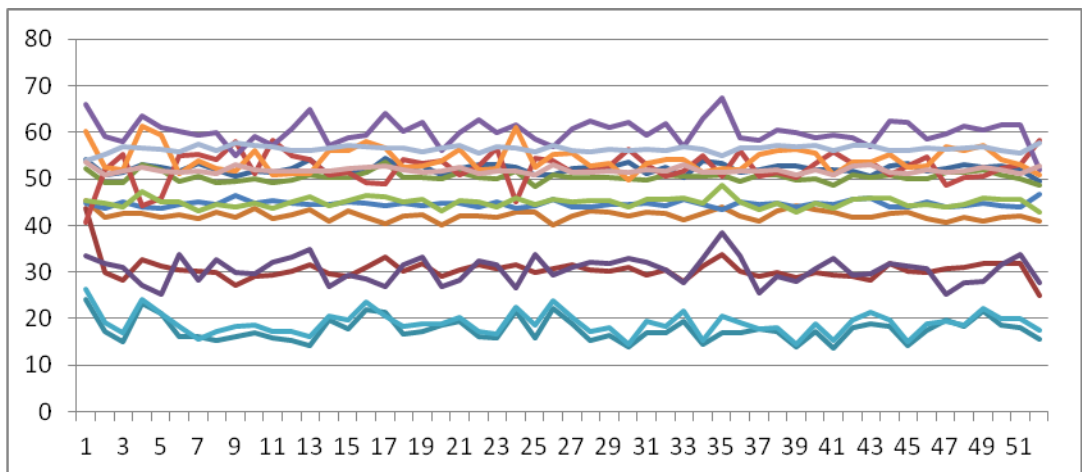
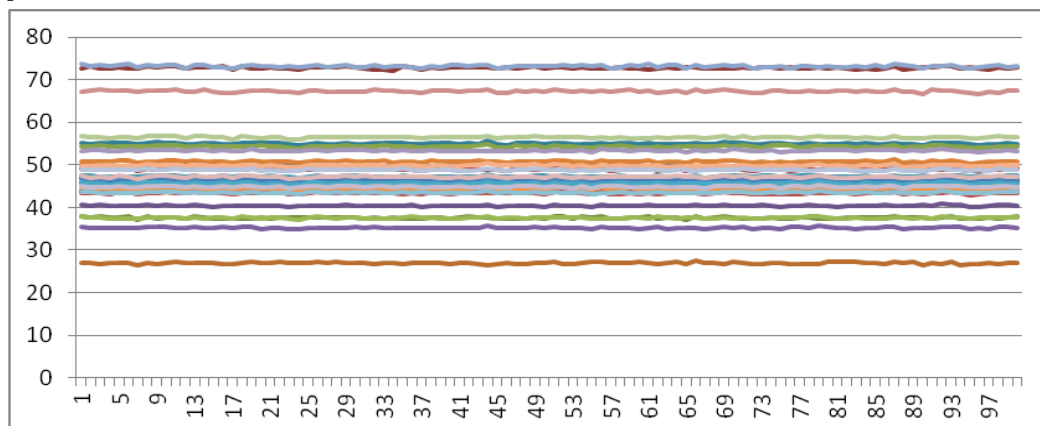


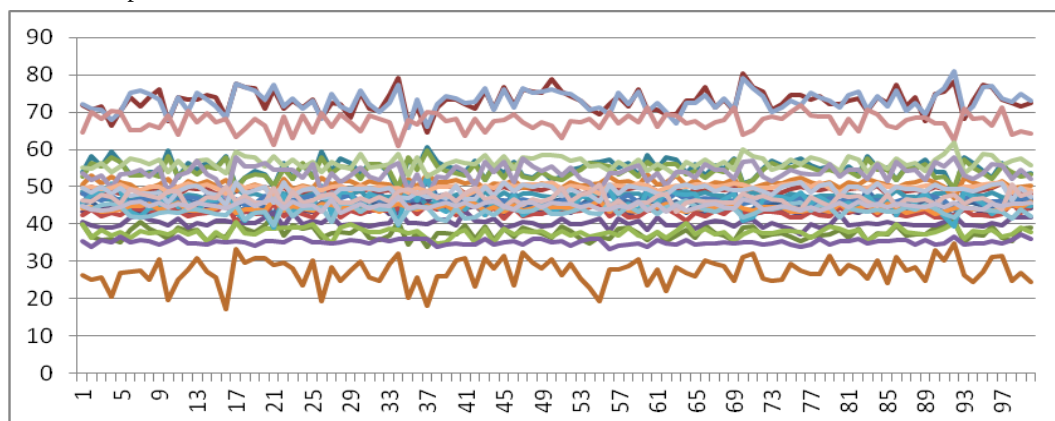
Chart VI. Ranking of 14 universities in *social work* with weights based on weights suggested by 50 university rectors.



Ranking of 14 universities in *economics* using the actual weights used in BURS with uniform perturbations



Ranking of 14 universities in *economics* using the actual weights used in BURS with normally distributed perturbations.



Ranking of 27 universities in *economics* based on weights suggested by 50 university rectors

