

Scores and clusters of Hungarian universities

ZSOLT TIBOR KOSZTYÁN^{1,2,3*} , VIVIEN VALÉRIA CSÁNYI¹  and
ANDRÁS TELCS^{1,3,4,5} 

¹ MTA-PE Budapest Ranking Research Group, Veszprém, Hungary

² Institute of Advanced Studies, University of Pannonia – Kőszeg Campus, Kőszeg, Hungary

³ Department of Quantitative Methods, University of Pannonia, Veszprém, Hungary

⁴ Wigner Research Center for Physics, Budapest, Hungary

⁵ Department of Computer Science and Information Theory, Budapest University of Technology and Economics Budapest, Hungary

Received: April 20, 2017 • Revised manuscript received: April 21, 2018 • Accepted: November 4, 2019



© 2020 The Author(s)

ABSTRACT

We present an application preference, list-based framework to Hungarian universities, which allows different type of flexible aggregation, and hence, analysis and clustering of application data. A novel mathematical method is developed by which preference lists can be converted into scored rankings. The proposed approach is demonstrated in the case of Hungary covering the period of 2006–2015. Our method reveals that the efforts to leverage the geographical center–periphery differences did not fulfil the expectations of policy makers. Also, it turns out that a student’s top preference is very difficult to influence, while recruiters may build their strategy on the information of the first but one choice.

KEYWORDS

application preferences, preference clustering, Hungarian higher education

JEL CLASSIFICATION INDICES

C02, I23

* Corresponding author. Department of Quantitative Methods, University of Pannonia, Egyetem utca 10, Veszprém, H-8200, Hungary. E-mail: kzst@gtk.uni-pannon.hu

1. INTRODUCTION

University rankings attract the attention of students, parents, educators, Higher Education (HE) managers as well as politicians. Since rankings provide crucial information, ranking methodologies are of the utmost importance. Given that the vast literature on university ranking also contains excellent reviews and critical analytical works (see e.g. [Pittaway – Cope 2007](#)), we do not elaborate on the evaluation of them. We do not wish to promote ours as a ranking method either, or claim that it is better than other methods. Our aim here is to present a framework, which allows different types of flexible aggregations, and hence, the analysis and clustering of the application data. A novel mathematical method is developed by which preference lists can be converted into scored rankings. Our method is based on the aggregation of student's preference lists. It is free of *ad hoc* subjective weights and biased data provided by the institutions sometimes. We shall see that the choice of the input data and the aggregation method applied to the preference lists provide a very flexible foundation for further analysis in different directions.

When 18-year old students submit their applications to *one or more higher education institutions*, it is the result of a very complex evaluation process in which a lot of aspects are taken into consideration, including family traditions, financial background, and the students' position in the social network and in their sub-culture. Other important factors are, students' academic capabilities, their motivations, as well as exogenous factors like distance, prestige and quality of the institutions, tuition fees, living costs, expected salaries after graduation and so on ([Avery et al. 2013](#); [Hilmer – Hilmer 2012](#)). All these factors are projected down to a single virtual cost-benefit (utility) scale, where the universities receive their relative position. Of course, an individual choice does not reflect the excellence of a university, but the aggregation of the preference lists of *all* applicants reflects the students' perceptions very well about the benefits of choosing one or the other. The usual ranking methods and our earlier work on preference lists ([Telcs et al. 2016](#)) cannot provide a way to capture the similarity (or dissimilarity) measure of the universities. Thus, we present a method that turns our preference lists to preference scores and quantifies the distance between the elements on the list.

As a case study, we consider application system and data from the Hungarian HE application system. In Hungary, students can apply for admission to more than one majors of (the same or) several (different) universities indicating a strict preference order of them. A student is admitted to the first university on his/her list for which he/she meets the requirements. The ranking of the universities on an application form reflects a student's preferences ([Telcs et al. 2016](#)), the authors consider the list of applications as the results of paired comparisons of the institutions. If institution "A" precedes institution "B" in the list of a certain student, then "A" is more adequate for him/her than "B". If "A" and "B" are (not) included on the list of a certain student, they are considered equally (in)adequate. As the number of applicants is high every year, we have a lot of people's opinions which can be aggregated into a ranking of institutions. [Telcs et al. \(2016\)](#) have proposed a method to aggregate application preferences in order to obtain faculty or institutional preferences. They have compared this method to other ranking possibilities. The advantage of the method is that it does not use self-imposed weights, consequently it reflects the unobscured preferences of all applicants. The disadvantage of the method is that it provides ranks and not scores. Nevertheless, this method permits clustering, similarly to the rank order clustering methods (e.g. [Li et al. 2007](#)); the preference orders can also be clustered.



The lack of scores has prompted us to elaborate a new method for ordering. The method assigns random variables to the institutions and estimates their expectations by maximum likelihood method. The method improves the Thurstone's method (Mosteller 1951; Thurstone 1994) allowing for more than two options in the results of comparisons. The proposed method makes it possible to calculate institutional score values. Scores are more sensitive indicators than ranks, and allow to respond to early warning of score shifts. In addition, these values provide a refined basis for clustering and evaluation and they permit us to evaluate the significance of the difference of the institutions.

The application-based preference orders and the score values have a potential for characterizing and clustering institutions, regions, professional areas and so on (Kosztván et al. 2020). We shall demonstrate how the proposed methods can show the potentials, attraction of institutions, and how these methods can be used for grouping universities by the application preferences and the applicants' performances. The proposed preference scoring and its clustering methods use only objective data like application orders and admission points. In addition, the changes of the (institutional/professional) preferences can be analysed with our methods.

This paper is structured as follows. After the introduction, Section 2 summarises previous analyses on students' motivation and factors influencing their application to higher education. In Section 3, a possible solution is described for handling preference orders. Section 4 discusses the novel mathematical method in detail, which can convert preference lists into scored rankings, and this method is demonstrated in the case of Hungary (Section 5). Finally, in Sections 5 and 6, conclusions and possible future researches are presented.

2. RELATED WORKS

We collect here some works to perform a country level deep analysis of students' motivation, factors influencing their decisions and in some cases the measurable implication of policy changes to them (Perna – Titus 2004). Our study points out that while the institutional preferences are changing due to the changes in the system of financial support, professional area preferences are very robust.

As we mentioned before, there are many factors which can influence a student's application to higher education. Participating in the higher education is very costly for students, thus one of the main factors that influences a student's decision whether to participate in higher education at all is permissible within his/her *socioeconomic status*. There are several studies that focus on students' financial circumstances (e.g. Cabrera – La Nasa 2001; Dearden et al. 2014; English – Umbach 2016; Migin et al. 2015; Paulsen – John 2002), some of them discuss the ethnic differences as well (e.g. Hurtado et al. 1997; Hilmer – Hilmer 2012; Niu 2015; Perna 2000).

Student aid and loans are widely used tools to make higher education more accessible for students from low-income families. Several papers examine how these tools affect students' choices (e.g. Dearden et al. 2014; Paulsen – John 2002; Perna – Titus 2004; Migin et al. 2015). Furthermore, tuition fees also have a great impact on the decisions concerning the enrolments in higher education (e.g. Dearden et al. 2014; Paulsen – John 2002). There is evidence that during the process of choosing a college, the American students take into account the job opportunities, the cost of living, the expected earnings and the living standard as well (e.g. Hilmer – Hilmer 2012; Montmarquette et al. 2002)



Family has an even greater influence on the university choice. Several studies pointed out that the *parental education* has a positive effect on the prospects of graduating from higher education (e.g. Niu 2015; Paulsen – John 2002). The proximity of the education site is known as a key decision factor and has been the topic of several studies (e.g. Hilmer – Hilmer 2012; Drewes – Michael 2006; Shamsuddin 2016). In the case of Hungary, Jancsó – Szalkai (2017) showed that the accessibility of universities by car and the travel time are important factors for students.

Another key factor which has a considerable effect on students' choices is their performance. Students with good abilities (with high admission points, good grades) tend to apply to “elite” universities (e.g. English – Umbach 2016; Paulsen – John 2002). Thus, the *quality and the reputation of the university*, that is, the excellence of the HE institutions, are also important factors for students (e.g. Bruno – Improta 2008; Drewes – Michael 2006; Horstschräer 2012; Telcs et al. 2015). Avery et al. (2013) surveyed the high-achieving students and suggested a binary logit model for explaining the revealed preferences of students. However, their database contains only the first revealed preferences of the students, while our study considers all applications not only the first, but the second, third etc. applications.

We put emphasis on the students' sensitivity to the distance and econo-geographic factors, which, of course, depend on the population structure, the size and the cultural history of a given country. The method proposed in Section 4, regardless of the large variety of influencing factors, is applicable.

We have thoroughly searched the literature for alternatives to our scaling method, but we have found that compared to previous attempts our method is novel in assigning a scale to preference lists of HE institutions and in allowing us to test if the difference between the ranked institutions is significant.

While one can find numerous international ranking of the HE institutions, preference studies on the international level are missing, for good reason. There is no central database, where such aggregated information could be accessible. This impedes a study of students' preferences and motivation on the European or even a global level, even though it would be of utmost interest, given the increasing student mobility. It can be expected that the development of the pan-European student mobility will support the foundation of a unified database, which would eventually enable such an analysis.

Most of the studies we found are based on surveys (e.g. Dearden et al. 2014; Hilmer – Hilmer 2012; Hurtado et al. 1997; Migin et al. 2015; Perna 2000), interviews (e.g. Pérez – McDonough 2008) or a (sub)sample of a database (e.g. Montmarquette et al. 2002). Their sample sizes of the research database were considerable, however, our research is based on the application data of the entire student cohort. Our Hungarian database contains more than 400,000 applications per year and more than 6 million records (applications) within the research period (2006–2015). The proposed preference scoring and clustering methods use only the objective data like application orders and admission points. However, with these methods and the longitudinal data, the changes of the (institutional/professional) preferences can be analysed.

3. PRELIMINARIES

First of all, we have to handle two major problems with respect to the preference orders. The application preference orders are incomplete, because we do not know the positions of the



not-preferred institutions or majors. A possible solution to that is proposed in [Telcs et al. \(2016\)](#) and we adopt that in our study as well.

Assumption 1: *It is assumed that if an institution is not present on the student's application list then this institution is less preferred by her/him to those that are on the list.*

We also need the comparisons of the unlisted elements.

Assumption 2: *The institutions which are not present on the student's application list are not distinguished by her/him.*

Different kinds of distance functions are sensitive to different kinds of dissimilarities. For the sake of simplicity in most of the cases we use the Euclidean distance.

4. FROM RANKS TO SCORES

4.1. The evaluation method

In order to introduce scores and ranks, we assume that the evaluations are based on a kind of utility model. The institutions have a hidden utility value m_i , which is perceived by the applicants with Gaussian error $\xi_i \sim N(m_i, 1/2)$, $i = 1, 2, \dots, n$. The rank of the expectations provides the rank of the objects in our model, but that is hidden for us. We know the applicant's partial preference list which carries the pairwise rank differences. We assume that those pairwise rank differences are the proxies for the applicants' *hidden perceived differences* $\eta_{i,j} = \xi_i - \xi_j$. The sample concerns these differences. We suppose that the samples are independent and Gaussian distributed with equal unit variances, i.e. $\eta_{i,j} \sim N(m_i - m_j, 1)$, $i = 1, \dots, n-1$, $j = i+1, \dots, n$. We allow six categories for the samples and these are quantized by a parameter $0 < d$. In our model, we have the parameters m_i and d . Using standard maximum likelihood estimation procedure, we can obtain estimate of the centers and d as well (up to an irrelevant shift given that we only observe the proxies of the utility differences). Following [Orbán-Mihálykó et al. \(2019\)](#), we assume that the pairwise comparisons are independent though the revealed partial ranks impose a dependence between them. As it is pointed out, without this restriction the model is practically not computable, and given the large number of samples, we hope the bias is not influencing our results very much. Denoting the sample by V , the maximum likelihood estimation of the parameters is

$$\left(\hat{m}_1, \dots, \hat{m}_n, \hat{d} \right) = \underset{(m_1, \dots, m_n, d)}{\operatorname{arg\,max}} L(V | m_1, \dots, m_n, d) \quad (1)$$

We can use the theorem ([Orbán-Mihálykó et al. 2019](#)) which ensures (under mild conditions) that the maximum likelihood method can be applied and the maximum in (1) exists and is unique. We have checked and found that the assumptions of the theorem are satisfied in our case.

4.2. Testing equivalence

Our model on preference ordering makes it possible to judge on tight in the ranking. It can be investigated by the likelihood ratio test. In [Kosztján et al. \(2020\)](#), the authors used a different method to create scores for the institutes. In our analysis, we model the respondents as



non-consequent, the pairwise comparisons are not necessarily consistent, while in Kosztyán et al. (2020) the responses are consistent, that is $\eta_{i,j} + \eta_{j,k} = \eta_{i,k}$. Interestingly enough, the two models produce very similar results. Let us fix two institutions i, j and let $m_i = 0$.

Now, let

$$H_0 : m_j = 0 \tag{2}$$

be our null hypothesis and its alternative be

$$H_1 : m_j \neq 0. \tag{3}$$

Let, SS_i be the subset of admissible parameters if $m_i = 0$ and SS_{ij} are under the hypothesis H_0 . The test statistic (see Wilks 1938) is

$$D_{i,j} = -2 \log \left(\frac{\sup_{(m_1, \dots, m_n, d) \in SS_{ij}} L(V|m_1, \dots, m_n, d)}{\sup_{(m_1, \dots, m_n, d) \in SS_i} L(V|m_1, \dots, m_n, d)} \right). \tag{4}$$

If our model assumptions and H_0 hold, $D_{i,j}$ is asymptotically Chi-Square distributed with 1 degree of freedom. Consequently, if the test value in (4) is larger than the critical value of the Chi-Square distribution, we reject the equality of the expectations belonging to the objects i and j .

For the further analysis, we introduce the exponential transformation of the expected values:

$$w_i = \frac{\exp(\hat{m}_i)}{\sum_{j=1}^n \exp(\hat{m}_j)} \tag{5}$$

similar to the ones used in the discrete choice models or the utility theory.

4.3. Similarities and dissimilarities: distance of the score vectors and clustering

Let us recall and make it clear that the score vectors \underline{w} are in one-to-one correspondence to the vectors $\hat{\underline{m}}$, which characterize the estimated positions of the objects. The estimates themselves are based on the submitted applications, or a subset of them. In that way for an aggregation of the preference list, a score vector can be assigned to the given group of the list, would it be a pool of lists coming from a geographical region, from a decile of the students according to their performances or other partitioning of the submitted applications. The score vectors, representing groups (partitions of the whole submission) associated with the institutions, provide a spatial representation and possibility for comparison and clustering of them. We will use the usual Euclidean norm between them, that is

$$D^E(\underline{w}^{(1)}, \underline{w}^{(2)}) = \sqrt{\sum_{i=1}^n (w_i^{(1)} - w_i^{(2)})^2}. \tag{6}$$

The clustering is performed applying the ‘*k*-means method’. We arrange the subsets into the groups in such a way that the sum of squares from points to the assigned cluster centers is minimized. At the minimum, all cluster centers are equal to the mean of their Voronoi sets. The algorithm is generally accepted and it was published in Hartigan – Wong (1979). It uses the



previously determined scores and they are not recomputed during the clustering process. Consequently, the sizes of the subregions do not affect the clustering process. Further, we use KM-PSC as the acronym of this preference score-based clustering algorithm.

5. THE CASE OF HUNGARY

5.1. Data sources

The original data source is the totality of applications submitted to the Hungarian National Center of Higher Education – Education Nonprofit Ltd. This office collects and handles all HE applications – for the entire country and for all universities.¹ Their database contains the annual applications. Our subset contains the application records for the time period of 2006–2015. Each record has 12 fields as it is shown in List 1. The admission points cover the periods of (2006, 2010–2015), and calculated admission points (2007–2009) are computed from the raw student scores (results of final exams, marks, etc.).

List 1.

1. Year
2. Semester
3. Student ID
4. Number of personal preference order
5. HE institution
6. Faculty/School
7. Course
8. Level of study
9. Form of study
10. State or private financing of tuition
11. Location (subregion of the student's residence)
12. Admission points

Each record refers to a single application, one student may submit more than one. It is typical that a student applies for three places but there are cases of more than ten applications for different HE institutions. Our database contains more than 400,000 records per annum from more than 100,000 applicants. Each application provides a separate record, which contains information about the order, as well. Thus, several records may belong to the same applicant. The candidate is admitted to the first institution, the requirement of which he/she satisfies.

The calculation of the admission points has been very similar year by year. These points are mainly based on the grades of the last two school years in the secondary education and the results of the regular final exam (maturity exam). If the results of two chosen subject's advanced final exam are higher, they are taken into consideration instead of the last two years' grades. Additional points can be obtained if the applicant has a language proficiency exam, professional qualifications or he/she is medically disadvantaged.

¹In other words, direct application to universities is not permitted in Hungary.



While the calculation process of the admission points was similar year by year, the maximum admission points were different in 2006–2007, 2008–2011 and 2012–2015. In order to compare student performances in different years, we use the quantiles of admission points.

Individual preference orders can be aggregated for any subset from which different characteristics of the institutions can be calculated. As a result, a very interesting cross-sectional view of the student-institution relationship can be analysed. In what follows, we demonstrate some among the numerous possibilities.

5.2. The impact of the tuition fee on the preferences of fields of studies

In Hungary, most of the BA, BSc, MA, MSc courses were entirely state financed before 2012. However, since 2012, students at BA and BSc programmes on economic and business studies and also most programmes in social sciences have stopped receiving state support. Some authors found that an increase in students’ grants causes an increase in HE participation (e.g. Dearden et al. 2014). Thus, it was expected that the introduction of the tuition fees would lower the interest for participating in these fields of studies.

In order to examine this hypothesis, we partition students according to their performances into deciles, and then, aggregate the chosen programs into fields of studies. Table 1 contains the fields of studies coded by numbers 1–13, student performances’ deciles (header) and annual results in rows taking into consideration the first four listed preferences (places) of the BA and BSc applicants.

The first cursory investigation of Table 1 reveals the interesting fact that economic and business studies keep their top popularity. Engineering (code 8) is less preferred than economic sciences (code 7) in all deciles, while definitely more preferred than social sciences (code 3) since

Table 1. The changes of the preference order of professions (fields of studies)

Year/D	1 st place										Overall	Year/D	2 nd place										Overall																	
	D01	D02	D03	D04	D05	D06	D07	D08	D09	D10			D01	D02	D03	D04	D05	D06	D07	D08	D09	D10																		
2006						7					7											7	2006	8	3	8					2				3	9	2			
2007*						7					7												7	2007*									2							2
2008*						7					7												7	2008*																2
2009*						7					7												7	2009*									2							2
2010						7					7												7	2010																8
2011						7					7												7	2011																8
2012						7					7												7	2012																8
2013						7					7												7	2013																9
2014						7					7												7	2014																9
2015						7					7												7	2015																9

Year/D	3 rd place										Overall	Year/D	4 th place										Overall														
	D01	D02	D03	D04	D05	D06	D07	D08	D09	D10			D01	D02	D03	D04	D05	D06	D07	D08	D09	D10															
2006		10	8								3											3	2006	3												8	
2007*		3									8												8	2007*													3
2008*		3									8												8	2008*													3
2009*		3									8												8	2009*													3
2010											3												3	2010													2
2011											3												3	2011													2
2012											9												9	2012													2
2013											9												8	2013													5
2014											9												8	2014													4
2015											9												8	2015													4

Notes: The applied codes: 1. Agriculture, 2. Human Studies, 3. Social Science, 4. Informatics, 5. Legal and Public Administration, 6. National Defense and Military, 7. Economics, 8. Engineering, 9. Medical and Health Sciences, 10. Pedagogy, 11. Sports Science, 12. Natural Science, 13. Arts; D01-D10: deciles of admission points 2006, 2010-2015 and (*) calculated student performances between 2007-2009.



2012. It is interesting that the introduction of a tuition fee for economics/business programs did not effect their popularity. This observation calls for further investigations.

Besides calculating preference orders, the proposed preference scoring method allows us to follow the change of the difference of the preference scores between two professional areas. Figure 1 shows that the score differences between the economic studies (code 7) and engineering (code 8) have decreased with time (despite the increase of the distance of preference positions, see Table 1). This example demonstrates that if investigating the scores instead of the orders, the changes may be perceived at an earlier stage.

5.3. How are the students' performances reflected in the institutional preferences?

While the changes of preference orders of professions can be important for the educational policy decision makers, the admission-based institutional preference changes are important for the HE institutions. A more detailed picture can be obtained if the students' preferences for institutions are aggregated over the deciles. We determine the preference order in every decile and we cluster the institutions. We face four distinct clusters. By computing the average ranks of the institutions in the clusters during the available time, the change of preferences can be followed. It is interesting to observe that the preference of institutions as a function of students' performances follows four clearly separable patterns as depicted in Figure 2.

The first group (CL_1) is formed by the institutions where the preference orders are independent of the admission points. Most institutions within this group are countryside colleges and the applications to these institutions come mainly from their own subregions. The fact that most institutions are in this group is in line with other researchers' finding (e.g. Alm – Winters 2009; Bruno – Improta 2008; Drewes – Michael 2006; Telcs et al. 2015). They claimed that the most important factor in university applications is the distance between residence and university.

Lots of scholars have demonstrated that another main factor in applications is the quality (Bruno – Improta 2008), reputation (Drewes – Michael 2006) and institutional excellence

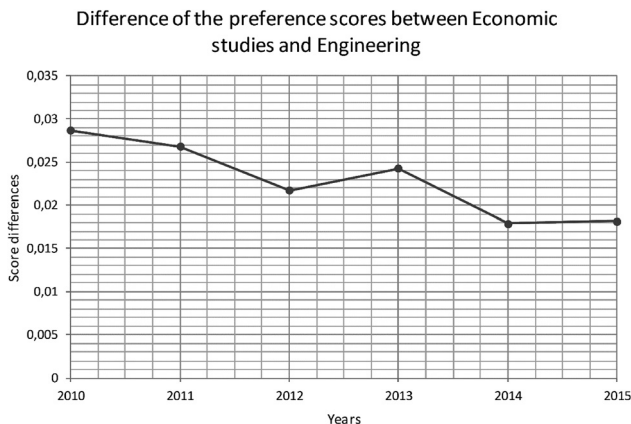


Figure 1. Difference between preference scores



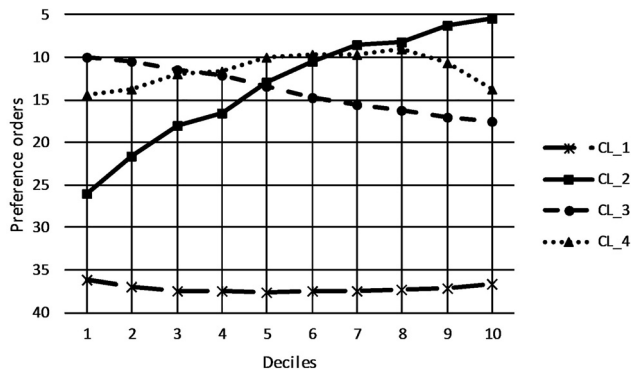


Figure 2. Cluster centers of institutional preference orders by deciles of students' admission points (2011)

(furthermore IE^2) (Horstschräer 2012; Telcs et al. 2015). This fact is again in line with the performance-preferences patterns in the other three groups. The second cluster (CL_2) of institutions contains the “elite” universities, which are preferred by the students who have higher admission points. It should be noted that all of them are located in the capital, Budapest. The third cluster (CL_3) includes several colleges and universities of mid-sized towns. In this group the relative position of the institutions decreases when the student performance increases. The fourth group (CL_4) is composed of different kind of universities of arts, which are preferred by the students who have only medium admission points. These universities and academies have different admission systems, which are based on individual talents rather than the students' performances in the secondary school. A similar phenomenon can be detected by plotting the changes of score values of the institutions in the deciles. Summarising the results, we can conclude that the institutions can be classified by the performances of the students and this classification coincides with another classification concerning the qualities of the institutions.

The changes of the clusters can be important signs for the institutions. The fine details of preference changes might be particularly interesting for decision makers at the institutions (see the Hungarian institutional preference clusters in Table 2 in Appendix).

5.4. Leaving or staying?

A student's application list contains information about the subregion³ of the applicant's residence. All proposed methods allow us to aggregate preferences by subregions. This way we can specify the subregional top preferences (the institution with the highest preference score). The proposed subregional preference map can be produced to all explored years and to all

²IE is a multidimensional rank, which is based on publication activities, the number of qualified professors etc. (See in detail: <http://eduline.hu/rangsor>)

³Subregions of Hungary were the subdivisions dividing the 20 counties of Hungary (including Budapest) into 175 administrative subregions (in Hungarian: *kistérség*). The subregions were abolished and replaced by 175 districts in 2013. However, in order to compare temporal results, all districts were assigned to the adequate subregions.



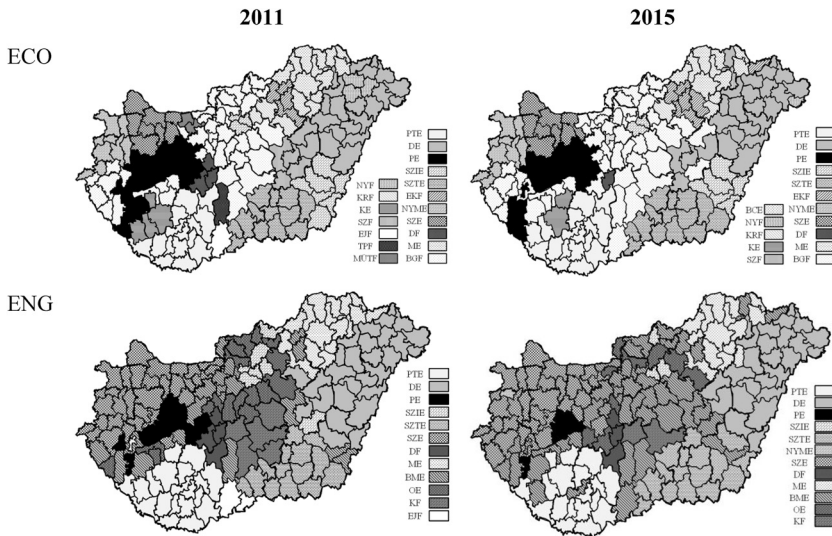


Figure 3. Comparing the subregional preference orders

Notes: ECO: Economic and Business Studies; ENG: Engineering

professional areas (for economics and business studies and engineering in 2011, 2015, see Figure 3). We can identify the geographically connected areas in the case of all fields of studies and also in the case of all applications. Each connected area has its own regional center institution. This phenomenon indicates the role of distance in the applications. We can identify similar shapes clustering the score vectors. The changes of the areas in time can indicate the realignments. The significance of the local institutions is decreasing, and *the increasing role of the capital is evident* and visualized.

Early warnings of changes are the result that we get by testing equivalence using (4). In Table 3 in Appendix, we find those subregions in which the top two preferences are not significantly different in 2011. Most of these subregions changed its top preference by 2015 and lots of further subregions have appeared with non-significant top two preferences (Table 4). Investigating HE institutions (HEIs), Figure 3 shows that the top preferences of the subregions turn more and more to the institutions in the capital, therefore the Budapest-centric attitudes of the students are increasing in time.

Turning to the role of the institutions locally, the 3D map of the preference scores (Figure 4(a)) and the preference contours (Figure 4(b)) show that the lowest relative scores are in Budapest and close to Budapest, where several universities and colleges compete for the students' applications, while these scores are relatively high in the peripheral subregions.

This might be explained by supposing that the institutions in the towns have more important roles locally in their surroundings than any institutions in the capital for the applicants living in Budapest. In the countryside the most outstanding universities are PTE (University of Pécs), SZTE (University of Szeged), DE (University of Debrecen), while the Northern part of



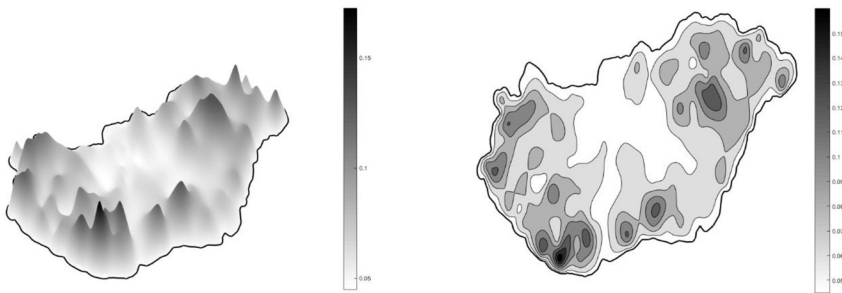


Figure 4. Preference scores of the top institution in the subregions and the contours of their territorials (based on 2015 applications)

Transdanubia has lower peaks. We can conclude that any changes which affect local institutions may have a profound effect on their applicants.

5.5. Increasing dominance of the capital

If we consider the number of applications for the most preferred institutions, we can see that lots of applications come from the subregion of the institution. Table 5 in Appendix shows that the single clusters' top preferences are always the countryside institutions. On the other hand, after the second or the third position the (sub)regional preference list follows the general (global) preferences, although most of these institutions are farther than 60 kms from the center of the subregion. (For clustering, the k-nearest neighbour (k-NN) method is used on rank order distances instead of the Euclidean one.)

If we consider the sums of the scores belonging to the top preferences, the increasing role of the institutions in Budapest is clearly demonstrated. Figure 5(a) represents the subregions where the total weights belonging to the institutions in Budapest in the top 6 have increased from 2011 to 2015. Figure 5(b) highlights those subregions where the differences between the total scores of the capital and the regional institutions in top 6 have raised. In the majority of the subregions, the increasing role of the universities in Budapest can be observed.

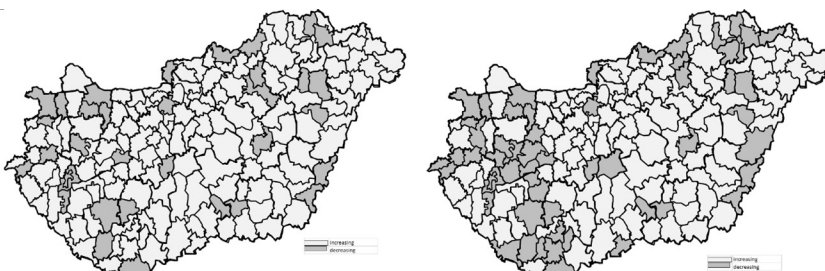


Figure 5. The subregions with increasing role of Budapest institutions



In 2011, almost all capital-dominant subregions were in the neighbourhood of Budapest, by 2015 the set has been enlarged remarkably (Figure 6).

Despite the government's intention to compensate Hungary's overcentralisation and strengthen the countryside universities, the students with good performance tend to apply to an institution in Budapest even if its IE position is worse than that of a countryside university. The most likely reason for the higher preference of the institutions in Budapest is the difference in job opportunities and living standards. This is in line with the findings of [Montmarquette et al. \(2002\)](#) and [Hilmer – Hilmer \(2012\)](#). The better the job opportunities and the higher the expected earnings, the more attractive a university or a major (city) is to a student. A more detailed analysis of these factors can be found in [Telcs et al. \(2015\)](#). As it has been demonstrated, in spite of the EU's and Hungarian government's efforts, the superiority of Budapest has not been eased but keeps increasing.

Although the competition among the institutions in Budapest is much stronger than among the regional universities, the number of applications is significantly larger in Budapest and near Budapest (Figure 7(a)) than in the far subregions. Most subregions have an average number of

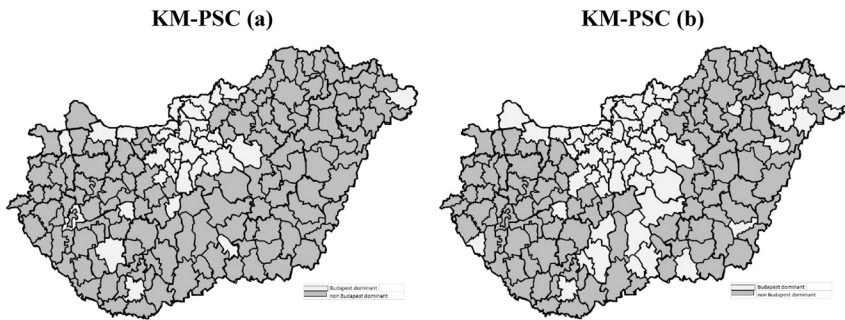


Figure 6. The subregions with Budapest-dominant HEI scores in top 6 during the time interval 2011–2015

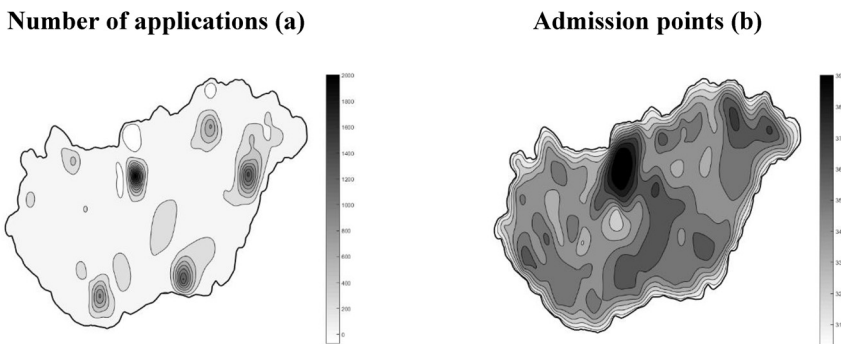


Figure 7. Number of applications and the mean of the admission points of the applicants in the subregions (2015)



applicants, although there are several underprivileged subregions with a very low number of applicants, such as the counties of Northern Hungary and Southern Transdanubia. It can be due to the fact that these regions have worse living conditions than other regions.

Figure 7(b) shows that most students with high admission points live in Budapest and in its agglomeration, while there are several subregions, where the admission points of students are 10% lower than the overall average.⁴ This result also highlights the subregional differences and the Budapest-centric comprehensive educational system.

6. CONCLUSION

6.1. Implications for decision-makers

6.1.1. Implications for education policy-makers. The results show that despite the European Union's Convergence Program and the Hungarian Government's intentions, decreasing the inequalities between the capital and the peripheral subregions *inter alia* by strengthening the regional institutions has not been successful so far. More students applied for the institutions in Budapest in 2015, than in 2011. Budapest's institutions can get students who have high admission points and the total weights of the institutions in Budapest are – in the majority of the cases – increasing. We believe that this effort will not lead to a success until the inequalities in the employment opportunities will not decrease.

The other main intention of the Hungarian Government was to increase the interest in engineering professions. However, according to the relative stability of the application preference orders, the result of this effort was only partially successful. While the majority of places in economic and business studies have not been financed by the state any more, economic studies have remained the most popular of all, and probably will remain in the future, too. Although the position of engineering has also been stable in the preference order of professions, the weights of engineering among the professions have been increasing and the distances between preference scores of the economic and business studies and the engineering have been decreasing in the time period of 2011–2015.

6.1.2. Implications for the decision-makers in the institutions. The management of a HEI has to balance between the quantity and quality of the applicants. The characteristics of the proposed institutional preference order show them the changes of the order in the function of the admission points. Both the changes in top preferences and the proposed clustering algorithms detect the impact and the increment of the impact of Budapest. The results (Table 2) present that one of the main challenges for regional institutions is how to acquire and retain the best applicants, and how to offer an attractive alternative against the increasing potential of the Budapest HEIs.

The other main challenges especially for the institutions in towns are how to stop the loss of applications and how to recover and later how to increase the potential application district. It is a long-term process to get into the elite cluster. The first step is to reach the third cluster, then the institution can progress.

⁴The maximal admission point is 500.



In order to organise an efficient promotion for a certain university, decision-makers should know where to go to promote the institution in order to keep or change the top preferences.

6.2. Implications for scholars

With the help of the proposed institutional preference scores we can set up a unique preference orders for professional areas and institutions. Moreover, we can specify the significant differences. The proposed institutional preference characteristics can be classified and this way the typical characteristics can be specified and distinguished.

The proposed preference order-based and score-based clustering algorithms specify the territorially connected preference clusters. Despite none of the proposed classifying methods uses geographical distance between the subregions, all of the proposed methods result in connected application districts. This result can be explained that in Hungary (similarly to several other countries) one of the main important factors in the choice of university is the distance between the residence of the applicant and the location of the university.

The proposed method can be applied on another large databases, like: National Center for Educational Statistics (NCES) (see: nces.ed.gov) and the change of the preferences can be analysed through the earlier significant parameters identified by other scholars (e.g. [English – Umbach 2016](#)).

7. FUTURE RESEARCH

The proposed preference orders, the preference score method and the preference score/order-based classification methods are generally applicable and provide useful results. The above presented results are based on the Hungarian HEIs' data and Hungarian students' applications. We would like to encourage decision-makers in other countries to produce similar preference score/order-based analysis in order to check the results of the governmental decisions and to improve the effectiveness and efficiency of the enrolment of the HEIs.

In this paper, we have investigated students' applications from the point of view of the universities and the subregions. In our next paper, we are going to investigate them from the applicants' points of view. According to the applicants' preference list, we believe that we can specify students who prefer institutions to programs (the first, second, third etc. places of the application list contain different programs, but these programs are announced in the same university), and there are several students who prefer programs to institutions. The important question is which factors influence institutional loyalty and insistence on the program.

Our work proposes methods which are typically applicable to those countries and higher educational systems where applications are systematically collected and stored in a centralised database. In that respect they are not applicable directly to develop international university ranking, but in a future work we will present such extension, as well. This information can also be important for the HEIs' decision-makers in the specification of an adequate program portfolio.

Since the application database can be merged into an employment database, a whole immigration path of a student can be tracked and analysed from his/her birthplace through the residence of secondary school and the university applied for to the residence of the employment. This immigration path shows underdeveloped subregions, from where it is very difficult to break



out. In the next paper, we are planning to analyse the students' immigration paths and the chances of rising of subregions.

ACKNOWLEDGEMENTS

The authors would like to express their special thanks to Éva Orbán-Mihálykó and Csaba Mihálykó for suggesting the stochastic pairwise comparison model and for some figures and tables. This publication was supported by the European Union and Hungary and co-financed by the European Social Fund through the project No. EFOP-3.6.2-16-2017-00017, titled "Sustainable, intelligent and inclusive regional and city models" and the "Analysis of Scientific Performance of Universities" – bilateral mobility project of HAS – FWO 2018–19. The research of the corresponding author has been supported by the Institute of Advanced Studies, Kőszeg (Hungary). The research reported in this paper has been supported by the National Research, Development and Innovation Fund (TUDFO/51757/2019-ITM, Thematic Excellence Program) and was supported by the Higher Education Excellence Program of the Ministry of Human Capacities in the frame of Artificial Intelligence research area of Budapest University of Technology and Economics (BME FIKP-MI/SC).

REFERENCES

- Alm, J. – Winters, J. V. (2009): Distance and Intrastate College Student Migration. *Economics of Education Review*, 28(6): 728–738.
- Avery, C. N. – Glickman, M. E. – Hoxby, C. M. – Metrick, A. (2013): A Revealed Preference Ranking of U.S. Colleges and Universities. *The Quarterly Journal of Economics*, 128(1): 425–467.
- Bruno, G. – Improta, G. (2008): Using Gravity Models for the Evaluation of New University Site Locations: A Case Study. *Computers & Operations Research*, 35(2): 436–444.
- Cabrera, A. F. – La Nasa, S. M. (2001): On the Path to College: Three Critical Tasks Facing America's Disadvantaged. *Research in Higher Education*, 42(2): 119–149.
- Dearden, L. – Fitzsimons, E. – Wyness, G. (2014): Money for Nothing: Estimating the Impact of Student Aid on Participation in Higher Education. *Economics of Education Review*, 43: 66–78.
- Drewes, T. – Michael, C. (2006): How Do Students Choose a University?: An Analysis of Applications to Universities in Ontario, Canada. *Research in Higher Education*, 47(7): 781–800.
- English, D. – Umbach, P. D. (2016): Graduate School Choice: An Examination of Individual and Institutional Effects. *The Review of Higher Education*, 39(2): 173–211.
- Hartigan, J. A. – Wong, M. A. (1979): Algorithm AS 136: A k-means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1): 100–108.
- Hilmer, M. J. – Hilmer, C. E. (2012): On the Relationship Between Student Tastes and Motivations, Higher Education Decisions, and Annual Earnings. *Economics of Education Review*, 31(1): 66–75.
- Horstschräer, J. (2012): University Rankings in Action? The Importance of Rankings and an Excellence Competition for University Choice of High-ability Students. *Economics of Education Review*, 31(6): 1162–1176.



- Hurtado, S. – Inkelas, K. K. – Briggs, C. – Rhee, B. S. (1997): Differences in College Access and Choice Among Racial/Ethnic Groups: Identifying Continuing Barriers. *Research in Higher Education*, 38(1): 43–75.
- Jancsó, T. – Szalkai, G. (2017): A Magyarországi Felsőoktatási Vonzáskörzetek Jellemzői, az Intézmények Elérhetősége és a Középszkolák Szerepe a Jelentkezésekben (Catchment Areas, Geographical Accessibility, and the Role of High Schools in the Hungarian Higher Education System). *Földrajzi Közlemények*, 141(4): 370–385.
- Koszttyán, Zs. T. – Orbán-Mihálykó, É. – Mihálykó, Cs. – Csányi, V. V. – Telcs, A. (2020): Analyzing and Clustering Student's Application Preferences in Higher Education. *Journal of Applied Statistics*, <https://doi.org/10.1080/02664763.2019.1709052>.
- Li, P. – Wu, Q. – Burges, C. J. (2007): Mcrank: Learning to Rank Using Multiple Classification and Gradient Boosting. In: *Advances Neural Information Processing Systems Conference Proceeding*, pp. 897–904.
- Migin, M. W. – Falahat, M. – Ab Yajid, M. S. – Khatibi, A. (2015): Impacts of Institutional Characteristics on International Students' Choice of Private Higher Education Institutions in Malaysia. *Higher Education Studies*, 5(1): 31–42.
- Montmarquette, C. – Cannings, K. – Mahseredjian, S. (2002): How Do Young People Choose College Majors? *Economics of Education Review*, 21(6): 543–556.
- Mosteller, F. (1951): Remarks on the Method of Paired Comparisons: I. The Least Squares Solution Assuming Equal Standard Deviations and Equal Correlations. *Psychometrika*, 16(1): 3–9.
- Niu, S. X. (2015): Leaving Home State for College: Differences by Race/Ethnicity and Parental Education. *Research in Higher Education*, 56(4): 325–359.
- Orbán-Mihálykó, É. – Mihálykó, C. – Koltay, L. (2019): A Generalization of the Thurstone Method for Multiple Choice and Incomplete Paired Comparisons. *Central European Journal of Operations Research*, 27(1): 133–159.
- Paulsen, M. B. – John, E. P. S. (2002): Social Class and College Costs. *The Journal of Higher Education*, 73(2): 189–236.
- Pérez, P. A. – McDonough, P. M. (2008): Understanding Latina and Latino College Choice. *Journal of Hispanic Higher Education*, 7(3): 249–265.
- Perna, L. W. (2000): Differences in the Decision to Attend College among African Americans, Hispanics, and Whites. *The Journal of Higher Education*, 71(2): 117–141.
- Perna, L. W. – Titus, M. A. (2004): Understanding Differences in the Choice of College Attended: The Role of State Public Policies. *The Review of Higher Education*, 27(4): 501–525.
- Pittaway, L. – Cope, J. (2007): Entrepreneurship Education: A Systematic Review of the Evidence. *International Small Business Journal*, 25(5): 479–510.
- Shamsuddin, S. (2016): Berkeley or Bust? Estimating the Causal Effect of College Selectivity on Bachelor's Degree Completion. *Research in Higher Education*, 57(7): 795–822.
- Telcs, A. – Kosztyán, Zs. T. – Neumann-Virag, I. – Katona, A. I. – Török, A. (2015): The Proceedings of 6th World Conference on Educational Sciences Analysis of Hungarian Students College Choices. *Procedia – Social and Behavioral Sciences*, 191: 255–263.
- Telcs, A. – Kosztyán, Zs. T. – Török, Á. (2016): Unbiased One-Dimensional University Ranking – Application-Based Preference Ordering. *Journal of Applied Statistics*, 43(1): 212–228.
- Thurstone, L. L. (1994): A Law of Comparative Judgment. *Psychological Review*, 101(2): 266–270.
- Wilks, S. S. (1938): The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9(1): 60–62.



APPENDIX

Table 2. Classification of the institutes (2011–2015) (Cluster); maximum (MAX) and minimum values (MIN) of the preference orders; I.E. Institutional Excellence see <http://eduline.hu/rangsor>;
 * = ecclesiastical universities/colleges; ** = universities/colleges/academies of arts/music
 *** = IE calculated)

Institutes	2011				2012				2013				2014				2015				
	Cluster	MAX	MIN	IE	Cluster	MAX	MIN	IE***	Cluster	MAX	MIN	IE	Cluster	MAX	MIN	IE	Cluster	MAX	MIN	IE	
ANNYE	1	51	48		1	48	44		1	49	41		1	53	47		1	54	45		
ÁNVF	1	28	22	32	1	32	27	33	1	32	26	30	1	35	30	24				40	
ÁVKF*	1	35	28	35	1	34	28	36	1	35	28	37	1	35	27	38	1	35	28	37	
BCE	2	25	1	8	2	24	2	14	2	19	1	27	2	19	1	8	2	20	1	19	
BGF	4	9	2	41	3	6	2	42	3	6	2	43	3	6	1	43	3	6	1	37	
BKF	3	16	11	37	1	20	18	32	1	21	17	30	1	22	17	32	1	20	13	34	
BKTF**	1	51	46	27	1	49	42	27	1	49	43	27	1	52	44	26	1	53	42		
BMÉ	2	33	3	11	2	26	2	14	2	24	3	20	2	25	3	20	2	21	3	17	
DE	3	6	1	3	1	5	1	9	3	5	1	12	3	5	1	13	3	7	1	12	
DF	3	30	16	29	3	33	15	29	1	29	23	32	3	34	21	35	3	37	24	26	
DRHE*					4	40	32		3	42	33	36		1	42	35	30	1	43	35	25
EGHF*														1	53	43		1	53	43	
EHE*														1	51	41		1	52	40	
EJF	1	37	29	28	3	40	27	22	1	38	28	19	2	37	22	22	1	35	28	22	
EKF	3	20	13	13	3	19	13	12	1	18	14	10	3	18	13	17	3	19	13	17	
ELTE	2	13	1	1	2	9	1	1	2	10	1	1	2	9	1	4	2	10	1	4	
GDF	1	34	23	42	1	35	31	42	1	33	28	41	2	32	20	44	1	38	24	43	
GFF*														1	47	37		1	37	30	
GFFH*	1	51	46		1	49	45		1	48	41		2	53	40		1	54	44		
HJF,MÜTF=>EDUTUS	1	38	30	32	1	29	26	36	3	36	27	41	3	34	26	41	3	38	30	42	
IBS	1	44	41	46	2	42	31	41	2	42	32	44	2	42	34	27	2	40	27	40	
KE	1	27	19	13	3	25	17	11	3	26	16	8	3	23	15	13	3	25	17	7	
KEE	1	51	46		1	49	44		1	49	44		1	53	42		1	53	44		
KF	3	36	17	40	3	30	14	39	3	23	13	38	3	26	14	36	3	24	15	35	
KJF	3	18	14	37	1	25	20	36	1	26	20	32	1	29	23	34	1	27	23	26	
KRE*	2	32	10	7	2	25	10	6	2	21	8	3	1	15	8	5	1	12	8	10	
KRF	1	25	15	26	1	26	16	23	3	34	20	15	1	27	21	18	3	29	18	7	
LFZE**	1	46	38	25	4	47	33	24	4	49	32	22	4	49	30	5	4	45	34	10	
ME	3	14	6	9	3	16	10	17	3	16	11	18	3	16	11	20	1	17	10	15	
MKE**	1	45	36	24	4	44	35	19	4	44	33	17	4	50	29	8	4	50	29	24	
MOME**	1	44	34	19	4	40	28	17	4	49	25	14	4	53	23	12	4	53	24	22	
MPANNI=>PAF	1	42	33	36	4	40	30	35	2	39	29	35	4	44	30	40	1	43	25	33	
MTF**	1	45	38	30	4	45	36	29	4	45	28	23	4	46	28	29	4	51	31	44	
RTF,ZNME=>NKE	1	37	28	23	2	22	15	24	4	15	9	25	4	22	11	24	2	26	11	26	
NVF	3	22	10	12	3	21	12	10	3	22	15	7	3	27	16	1	3	29	16	7	
NVME	3	17	5	3	3	16	5	6	3	17	8	3	3	17	6	8	3	17	9	12	
OE	3	17	7	38	3	12	6	38	3	14	4	39	3	13	6	39	3	13	5	31	
PE	3	15	11	10	1	16	11	11	3	17	9	12	3	17	10	11	3	17	8	3	
PPHF*																		1	54	39	
PPKE*	2	26	8	16	2	23	7	13	2	23	7	3	2	27	7	5	2	23	5	6	
PTE	3	8	2	6	3	9	1	6	3	10	4	6	3	10	4	2	3	10	4	2	
SE	2	27	7	5	2	21	8	7	2	18	8	10	2	18	6	13	2	19	9	1	
SSZHF*	1	51	46		1	49	45		1	49	40		1	53	39		1	54	39		
SZAGKF*														1	53	45		1	54	46	
SZE	3	16	8	20	3	17	7	27	3	15	7	29	3	14	7	31	3	15	5	29	
SZF	3	26	20	43	1	26	18	42	1	27	21	39	1	31	23	41	3	33	22	35	
SZFE**	1	51	33	18	2	49	38	21	4	45	24	23	4	53	36	23	4	54	29	16	
SZIE	3	11	3	13	3	13	6	22	3	14	5	25	3	14	5	19	3	14	6	21	
SZTE	4	5	1	2	1	4	2	2	1	4	1	2	4	4	2	3	1	4	2	4	
TE																		2	31	18	
TPF	1	33	27	17	1	36	29	15	1	37	29	9	1	42	32	16	1	41	31	14	
VHF*	1	47	39	38	1	44	39	37	1	44	37	34	1	43	38	37	1	54	39	31	
WJLF*	1	44	35	21	1	41	35	21	1	43	36	20	1	46	38	32	1	49	38	37	
WSUF	1	45	38		1	39	35		1	39	33		1	38	32		1	42	34		
ZSKF	1	21	18	21	1	29	24	27	1	28	24	28	1	29	21	28	1	30	21	20	



Table 3. Indistinguishable subregional top preferences (2011), where the significance level = 0.05. Inst 1, 2 = top 2 institutes with the largest preference scores; Score 1, 2 = preference scores of Inst 1 and Inst 2; Sign = significance value; n = number of applications from the subregion

Subregions	Inst 1	Inst 2	Score 1	Score 2	Sign	n
Baja	EJF	PTE	0.06088	0.06031	0.23	3,692
Bátonyterenye	EKF	SZIE	0.05314	0.05209	0.39	710
Devecser	PE	SZE	0.08438	0.08324	0.67	362
Dunakeszi	SZIE	SE	0.04269	0.04220	0.27	4,717
Hatvan	SZIE	KRF	0.05477	0.05443	0.56	2,398
Keszthely	PTE	PE	0.05825	0.05765	0.38	1,936
Mezőtúr	SZIE	SZTE	0.05562	0.05496	0.43	1,084
Nagykálló	DE	NYF	0.08673	0.08633	0.64	1,804
Nagykáta	SZIE	ELTE	0.05026	0.04953	0.22	2,603
Oroszlány	SZE	ELTE	0.05305	0.05139	0.07	1,270
Sátoraljaújhely	ME	NYF	0.06068	0.06045	0.82	1,169
Sümegeg	PE	PTE	0.06772	0.06532	0.12	625
Tokaj	DE	NYF	0.06764	0.06621	0.33	582
Vasvár	NYME	PTE	0.07380	0.07122	0.17	457
Zalakaros	PTE	PE	0.07829	0.07645	0.35	441



Table 4. Indistinguishable subregional top preferences (2015), where the significance level = 0.05. Inst 1, 2 = top 2 institutes with the largest preference scores; Score 1, 2 = preference scores of Inst 1 and Inst 2; Sign = significance value; n = number of applications from the subregion

Subregions	Inst 1	Inst 2	Score 1	Score 2	Sign	N
Abaúj-Hegyköz	DE	ME	0.07859	0.07528	0.17	391
Adony	DF	ELTE	0.06115	0.05970	0.37	627
Budaörs	BME	BCE	0.04711	0.04711	0.90	5,427
Cellódömök	SZE	PE	0.07251	0.07143	0.47	767
Devecser	PE	SZE	0.08115	0.07615	0.10	261
Dombóvár	KE	PTE	0.06853	0.06833	0.87	906
Enying	PTE	ELTE	0.06106	0.06084	0.93	347
Esztergom	PPKE	ELTE	0.05186	0.05172	0.86	1,831
Hévíz	PTE	PE	0.07452	0.07313	0.50	500
Keszthely	PE	ELTE	0.05278	0.05183	0.27	1,353
Kisbér	SZE	NYME	0.06810	0.06610	0.14	514
Komárom	SZE	NYME	0.05982	0.05912	0.31	1,164
Lengyeltót	PTE	KE	0.10690	0.09901	0.10	205
Mezőcsát	DE	EKF	0.11606	0.11084	0.26	298
Mezőtúr	ELTE	SZTE	0.05680	0.05561	0.33	902
Mór	SZE	PE	0.05557	0.05508	0.57	1,150
Nagykanizsa	PE	PTE	0.05575	0.05527	0.46	2,203
Pacsa	BGF	PE	0.08912	0.08242	0.09	225
Paks	PTE	DF	0.05299	0.05289	0.89	1,790
Sárvár	NMYE	SZE	0.08336	0.08324	0.92	1,077
Sátoraljaújhely	ME	ELTE	0.07025	0.06859	0.28	822
Tab	PTE	KE	0.08861	0.08488	0.29	227
Tatabánya	ELTE	SZE	0.05615	0.05527	0.13	3,078
Tokaj	ME	EKF	0.08519	0.08337	0.54	356
Vasvár	PE	NYME	0.08718	0.08460	0.32	351
Zalaszentgrót	PTE	NYME	0.06928	0.06908	0.94	370



Table 5. Institutes, which are farther than 60 kms from the center of the subregions (denoted by *)

Cluster no	Subregion	Regional Inst.	Top 4 Preferences
1	<i>Remains</i>	-	<i>ELTE, DE, SZTE, PTE</i>
2	Budapest	ELTE	ELTE, BCE, BGF, BME
3	Debrecen	DE	DE, NYF, BCE*, ELTE*
4	Miskolc	ME	ME, DE, EKF, ELTE*
5	Pécs	PTE	PTE, ELTE*, BCE*, BME*
6	Sopron-Fertőd	NYME	NYME, ELTE*, BCE*, BGF*
7	Szeged	SZTE	SZTE, PTE*, ELTE*, BCE*
8	Veszprém	PE	PE, ELTE*, BCE*, PTE*
9	Gödöllő	SZIE	SZIE, ELTE, BCE, BME
10	Győr	SZE	SZE, NYME, ELTE*, BCE*
11	Kecskemét	KF	KF, SZTE, ELTE*, BGF*

Open Access. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited, a link to the CC License is provided, and changes - if any - are indicated. (SID_1)

