Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia - Social and Behavioral Sciences 00 (2013) 000–000

Procedia Social and Behavioral Sciences

www.elsevier.com/locate/procedia

5th World Conference on Educational Sciences

Analysis of Hungarian students' college choices

András Telcs [a], Zsolt T. Kosztyán [a] *, Ildikó Neumann-Virág [b], Attila Katona [a], Ádám Török [c]

[a] *University of Pannonia, Department of Quantitative Methods, Egyetem utca 10., Veszprém, H-8200, Hungary*
[b] *University of Pannonia, Department of International Economics, Egyetem utca 10., Veszprém, H-8200, Hungary*
[c] *University of Pannonia, Department of Economics, Joint Research Unit with the Hungarian Academy of Sciences on Regional Innovation and Development Studies, Egyetem utca 10., Veszprém, H-8200, Hungary*

Abstract

The present work is the second in a series of studies in which we are going to present an unbiased picture on the attractivity of universities. In doing so we confined our study to Hungary, where we have access to all annual application data of students to universities and colleges. Our first study presented an unbiased one-dimensional preference list of higher educational institutions, schools and study programs alongside a bunch of methods to produce such preference lists.
In the present work we report on the first results of the second stage of our project in which we investigate students' choice of further studies. Our database contains more than a million application entries, covers student scores, place of residence, and GDP per capita and employment data of their regions of residence. Similar economic data have been collected about the institutions as well as their indicators of academic excellence. We incorporated into the database the distance between students' places of residence and colleges as well. Classical and novel econometric methods are used from logistic regression and gravity models to neural networks. The study reveals some common patterns of students' choices and striking differences between different fields of studies. Among other results it has beenfound that the most preferred place of study is selected with much care while descending on the preference list the choice is less and less sophisticated. To the best of our knowledge this article is one of the few attempts to analyse the behaviour of student mobility: an estimation of the quantitative direct impact of several determinants for student flows.

Selection and peer-review under responsibility of Prof. Dr. Huseyin Uzunboylu, Near East University, Faculty of Education, Cyprus.

Keywords: collage choice, discrete choice, higher education, gravitation model

* Corresponding author: Zsolt Tibor Kosztyán. Tel.: +36-20-208-5840
E-mail address: kzst@gtk.uni-pannon.hu

1. Introduction

Higher education (HE) has many facets and social aspects; it reflects the state of a society as well as influences its future. Parents and students face a complex decision in choosing the next level of study after maturity, completing secondary education. Their choice determines not only their coming years but the intake of the HE institutes as well. Last but not least entry into the job market and the sectorial labour needs, starting wages, future perspectives are important factors in their decision Given that very complex interrelation between education and society, the study of the students' motivation, choice of content, form and financial background all may contribute to the better understanding how HE complies or conflicts with the needs, reality and financial possibility of society. The present work is based on students' application records in Hungary over 2000-2010. This sizeable dataset is complemented with geographical, micro level and labour market indicators. The other side of the dataset describes the economic environment of Hungarian HE institutions and the institutions themselves as well.

Based on that relatively complete picture we are able to analyse how students' decisions are influenced by the standing of the regions they are coming from as well as the parameters of the region of the university and last but not least the indicators of the university. We have access to a limited number of university indicators.

Our investigations confirm the great attractivity of Hungary's capital city for students. This is why we made an attempt to re-evaluate several of our findings for a sample excluding the HE institutes of Budapest. As a result we obtained a much more elaborated and rich picture about the rest of the country. The top preference reveals the mix of students' dreams, preconceptions and practical considerations. Main motivations include the prestige of the institution, job opportunities, and the distance of the institution from home. If we remove the number one choice, there are some changes in the importance of decision factors which help to reveal motivations. Distance becomes more important. The economic condition of the region of the institution is more important to students choosing economics or management studies than those choosing liberal arts. Finally on the bottom of the preference list student choice more or less predetermined and as a consequence, little or no particular motivation can be identified.

In our work we do not present an encyclopaedic overview of all studies, faculties, universities but focus on illustrative examples. For instance we show how students' selection criteria change from the top to the second most preferred among students preparing for studying liberal arts or economics. We provide a detailed picture on preferences including the capital city in our sample and also excluding it.

2. Methods applied

We are going to analyse Hungarian students' college choices based on the ordered preference lists they submit in the form of applications. In the previous phase of this research (Telcs et. al. 2013) several methods are proposed to create unified preference list of institutes, faculties and programs. Here we apply the gravity, potential, logit models and neural networks to analyse student choices.

Gravity models have become the standard technique for the empirical analysis of flows of capital and goods (Frankel-Rose, 2002).

However, the gravity model helps to study motivations of migration. The gravity model of migration is based on the idea that as the importance of one or both of the location increases, there will also be an increase in movement between them. The model is used to predict the degree of interaction between two places (Rodrigue et al. 2009).

We use a gravity model to analyse distance elasticity of students.

Potential models included in the category of spatial models are based on physical analogies. The potential model is a quite good method to analyse and visualize the patterns of an economy's spatial layout. If there are several gravitating bodies, the forces among them build up a force-field, the potential space, in which every single body has its effect on the others.

Places with relatively high field-potentials are those with many opportunities for interaction with other; on the other hand, places with low scores have relatively poor opportunities for interaction.
These models are used for analysing the college choices of students. Figure 1 summarises the independent variables based on related studies[†].

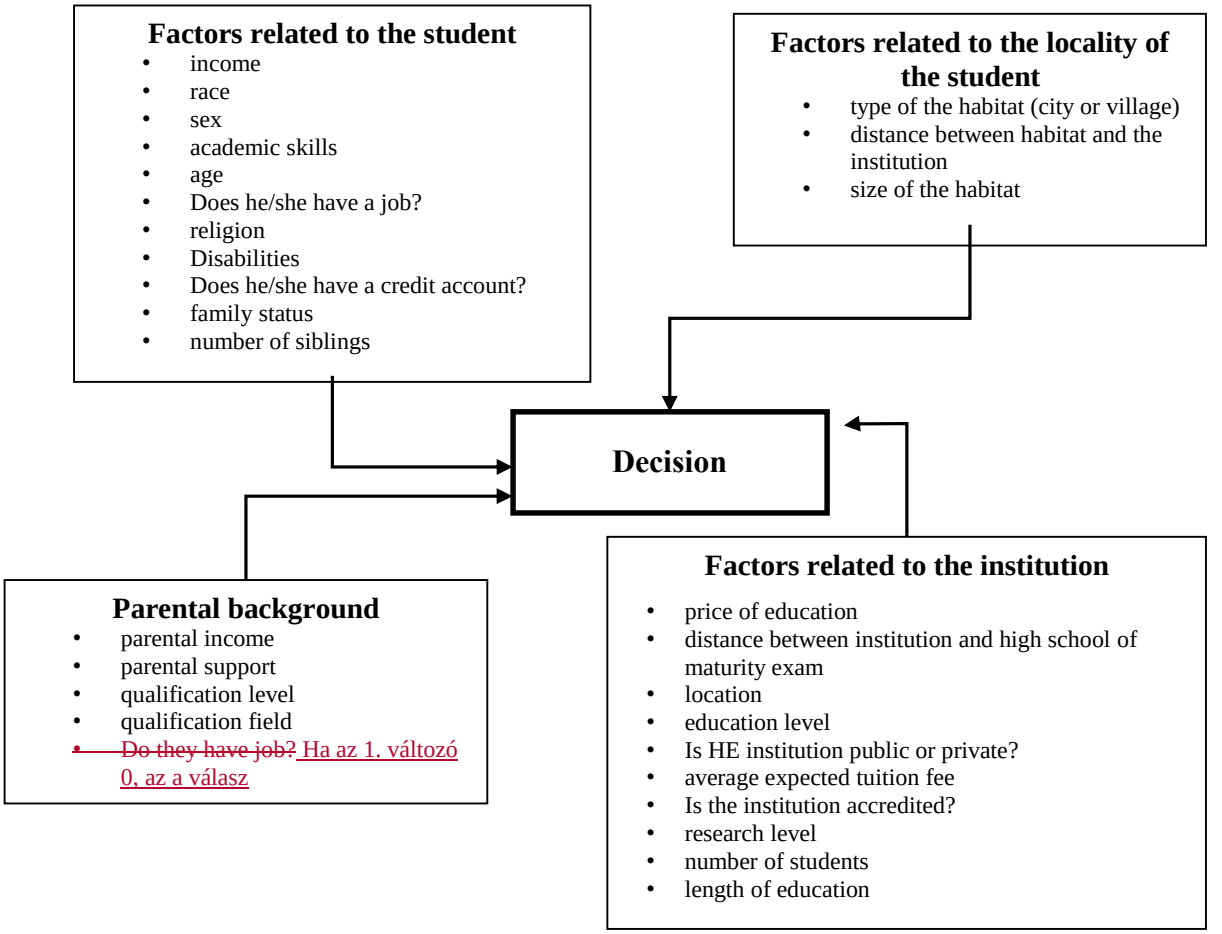


Figure 1: independent variables based on related researches

The sub-regional economic parameters and the excellence of institutions were not considered in these models. It is a generally accepted assumption that those factors contribute to students' decision a lot, so we should incorporate them in our investigations.
The ***consumer choice model*** introduced by McFadden (1974) is already a classical tool to investigate consumers' motivations in their decisions. To our best knowledge, it was first applied to analyse choices of college in Drewes (2006). The model is based on the extension of the ***statistical logit method***. Here we adopt the conditional rank

[†] (see i.e. Ahmet G., Büsra K., Ferda M., Çetin D., Kübra G., Hasan Y. H, 2011; Saisana, M., d'Hombres, B., Saltelli, A, 2011; Larose, S. Cyrenne, D., Garceau, O., Harvey, M., Guay, F., Deschęnes, C, 2009; Asari, F. F. A. H, Idris, A R, Daud, N. M., 2011; Rochat, D., Demeulemeester, J., 2011; Germeijs, V., Verschueren, K., 2007; Vrontis, D., Alkis Thrassou, A., Melanthiou, Y., 2007; Gibbons, S., Vignoles, A., 2012; Toutkoushian, R. K., 2001; Bruno, G., Improta, G., 2008; Alm, J., John V. Winters, J. V., 2009; Schwartz, B. 1985; Weiler, W. C., 1986,1989 Niu, S. X., Tienda, M., 2008; Montgomery, M., 2000; Montmarquette, C., Cannings, K., Mahseredjian, S., 2002; Reynolds, L. C.: 2012; DesJardins, S. L., Dundar, H., Hendel, D. D., 1999; Long, B. T., 2004; Coelli, M. B., 2011).

ordered logit model, which can handle partial preference lists to obtain the elasticity of the independent variables investigated, among others GDP/capita in the region of the institution.

3. Details of the methodology

In this section research database, dependent and independent model parameters are introduced.

3.1. Data and data preparation

For statistical analysis 3 kinds of master tables were used. The first one is *applications* collected by the Hungarian national centre of higher education (HE) - Educatio Nonprofit Ltd. This database contains all significant data of applications. The data table of *institute* contains the Institute Excellence parameter, which is a composite coefficient based on the qualified academic teachers & researcher per students, amount of academic degrees (PhD, CSc, DSc) etc.. The table of *sub-region* contains GDP per capita and Employment rate of the sub-region, and GPS coordinates for calculating distances between the centre of the sub-region of applicant's and the faculty or program of the institute. The third table contains the economic data of the sub-regions (see Fig.2).

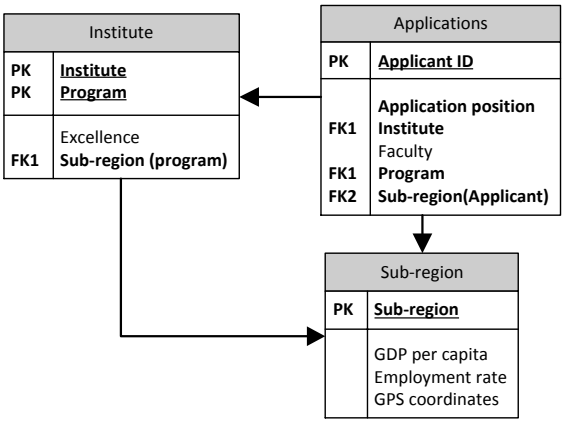


Figure 2: Research database

Each record of the table *applications* refers to a single application. Our database contains more than 400,000 records based on applications in the year 2011. Transactional tables for logit and gravity models generated by master data can contain more than one million entries.

3.2. Model variables

The independent variables are: (1) Institute Excellence; (2) Distance between the student's and institute (or program); (3) GDP per capita (sub-region of students); (4) GDP per capita (sub-region of the institute); (5) Employment rate (sub-region of students); (6) Employment rate (sub-region of the institute).
The dependent values are varied: in the case of gravity and potential model the number of applications are modelled. In case we use a binary and rank ordered conditional logit model the corresponding position in the preference list of student applications is investigated.

3.3. Investigated academic programs and studies

In this study we demonstrate our results on two specific fields; Economic & Business Studies and Human Studies. Given that the Budapest based HE institutions dominate the rest of the whole country we present the results with and without that Budapest's effect.

3.4. Neural networks for forecasting

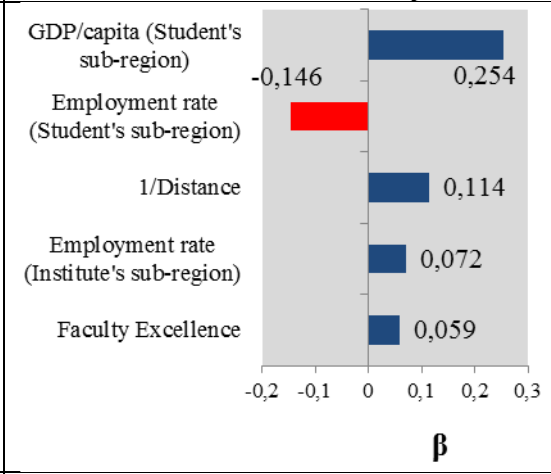
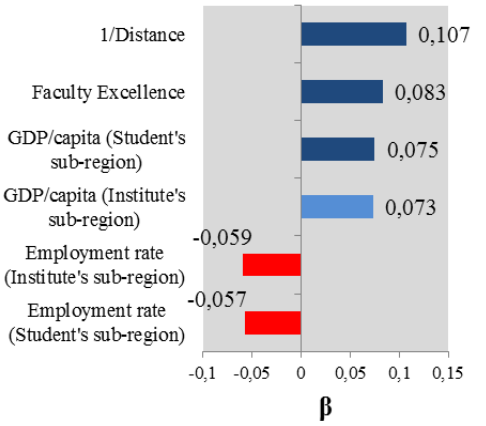
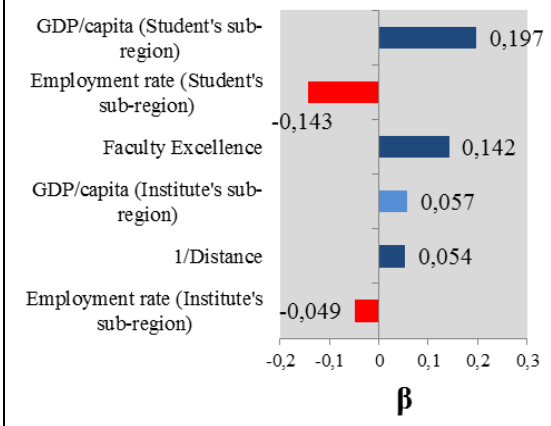
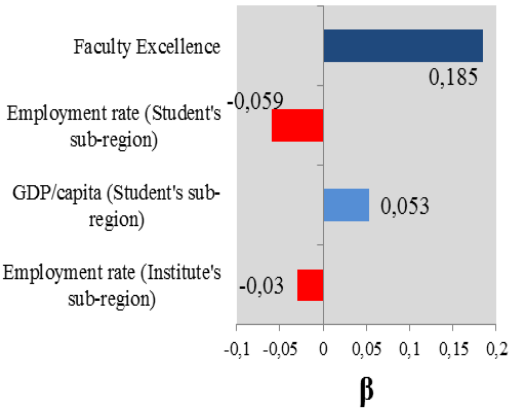
Besides standard statistical methods like logit models, analysing neural networks can also be used for forecasting. In this investigation a Radial Basis Function (RBF) of the Neural Network was applied to forecast student choices. The weights of input parameters can be described as importance values. Our database was separated into training (=70%) and test (=30%) data sets. Weights are calculated at the training phase and tested in the test database (see Table 4).

4. Results

4.1. Indicators of student applications

Number of students' applications are investigated by the gravity model. The gravity equation provides a benchmark analysis of the determinants of student migration within regions. The results indicate that wealthy regions attract more students.
The Table 1 shows the significant variables sorted by their impacts/importance.

Table 1: Results of gravity model

	With Institutions in Budapest	Without Institutions in Budapest
Economic and Business Studies		
Arts and Human Studies		

The economic coefficients are less, but faculty excellence are more important for students who wants to apply for arts and humanities. In the case of excluding institutions in Budapest faculty excellence is evaluated. Potential model can show the potential of institutions of higher education (see Figure 3).

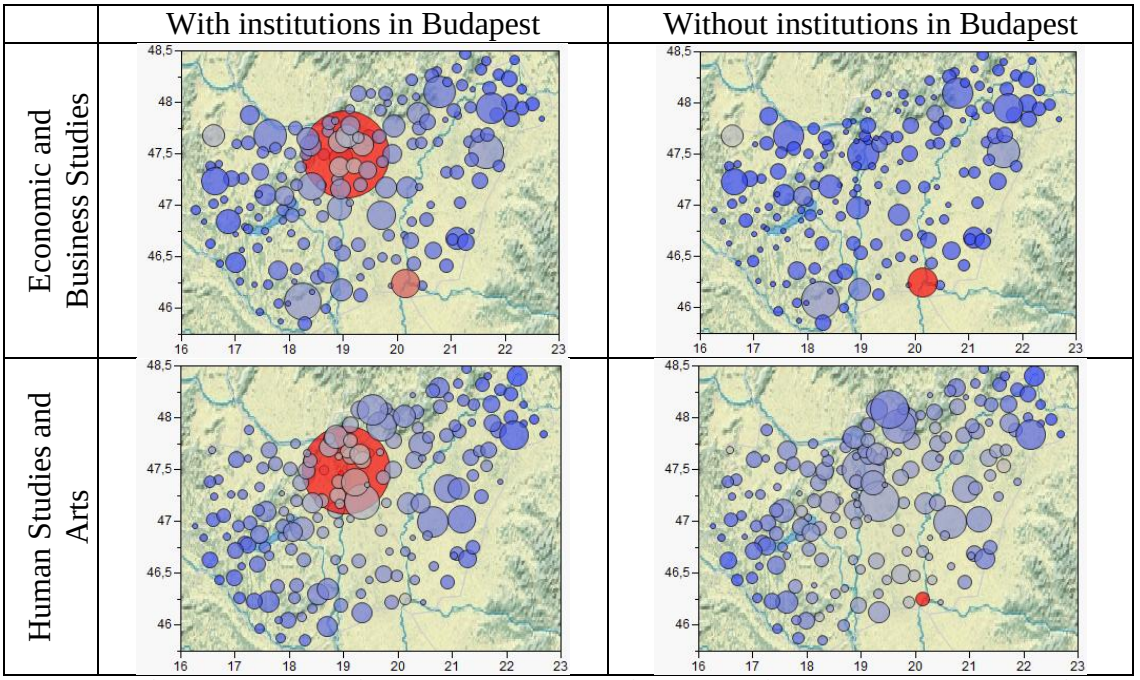


Figure 3: Results of potential model (size of circle is a proportional value of the potential of subregions. it shows the quality of access of subregions to universities; reddish circles represent the faculty excellence values)

Figure 3 shows, that capital city of Hungary, Budapest has the highest potential. In case of ignoring institutes in Budapest the potentials are more balanced. A former research (Langerné-Rédei M. (2007)) based on 2003-2005 student applications showed that at the time of EU accession the Western Hungarian Institutions had more student attracting potential

The colour of the circles in Figure 3 describe the importance of faculty excellence for students; reddish circles represent the faculty excellence values. Figure 3 shows that if we ignore institutions based in Budapest s the importance of faculty excellence becomes more important (see e.g. the University of Szeged N 46°16,247' E 20°05,333', Fig. 3.).

4.2. Indicators of preferences

In order to compare results independent variables are the same, but in this case the dependent variables are different. Applying binary logit, the significant parameters can be specified. This calculation also shows: most important factors are the distance between the institutions and the student's place of residence and Faculty excellence. The results of the logit model show that: economic parameters are less important for students applying to Humanities.

If we consider only the top priorities of applicants, the most two important factors include faculty excellence and distance. However the significance of these values can change if second, third and fourth order applications are also considered.

Table 2: Results of rank order logit

Order of application	Independent variables	With Institutes in Budapest				Without Institutes in Budapest			
		Economic and Business Studies		Arts and Human Studies		Economic and Business Studies		Arts and Human Studies	
		Sig.	Exp β	Sig.	Exp β	Sig.	Exp β	Sig.	Exp β
1st	Faculty Excellence	0,000	1,372	0,000	2,141	0,000	1,34	0,000	3,327
	1/Distance	0,000	2,482	0,000	2,117	0,000	2,038	0,000	1,7
	Emp. rate (Inst.)	0,000	0,966	0,000	0,898	0,000	0,904	0,000	0,879
	GDP per capita (Inst.)	0,000	1	0,000	1	0,000	1,001	0,009	1
	Emp. rate (Stud.)	-	-	0,006	1,048	0,029	0,983	0,048	0,95
	GDP per capita (Stud.)	0,090	1	0,000	1	0,000	1	0,003	1
2nd	Faculty Excellence	0,000	1,224	0,000	1,628	0,000	1,177	0,000	2,415
	1/Distance	0,000	2,224	0,000	2,113	0,000	1,836	0,000	1,702
	Emp. rate (Inst.)	0,415	0,986	0,007	0,924	0,105	0,961	0,008	0,913
	GDP per capita (Inst.)	0,000	1	0,000	1	0,009	1	0,285	1
	Emp. rate (Stud.)	-	-	0,975	1,001	0,676	0,993	0,126	0,96
	GDP per capita (Stud.)	0,000	1	0,000	1	0,000	1	0,002	1
3rd	Faculty Excellence	0,794	1,012	0,000	1,554	0,299	1,052	0,000	2,557
	1/Distance	0,000	2,095	0,000	2,024	0,000	1,731	0,000	1,627
	Emp. rate (Inst.)	0,757	1,005	0,576	0,985	0,221	0,968	0,190	0,961
	GDP per capita (Inst.)	0,003	1	0,001	1	0,110	1	0,037	1
	Emp. rate (Stud.)	-	-	0,926	1,002	0,379	0,984	0,017	0,941
	GDP per capita (Stud.)	0,001	1	0,037	1	0,000	1	0,153	1
4th	Faculty Excellence	0,943	0,995	0,000	1,38	0,000	1,7	0,049	1,784
	1/Distance	0,000	2,045	0,000	2,18	0,711	1,028	0,000	1,77
	Emp. rate (Inst.)	0,971	1,001	0,507	0,966	0,054	0,928	0,847	0,989
	GDP per capita (Inst.)	0,000	1	0,001	1	0,001	1,001	0,608	1
	Emp. rate (Stud.)	-	-	0,514	1,021	0,065	0,953	0,444	0,962
	GDP per capita (Stud.)	0,041	1	0,010	1	0,109	1	0,337	1

Table 3 shows that 2nd, 3rd and 4th applications have less significant variables. That indicates that the lower the priority the lower the freedom of choice and or? the more ad hock the choice is.

4.3. Results of importance estimation based on using neural networks

In this investigation a Radial Basis Function (RBF) of the Neural Network was applied to forecast that a given student would choose a HE institution in Budapest or not. In this case the input neurons were the: (1) Distance between the student's place of residence and the institution (or program) selected; (2) GDP per capita (sub-region of students) (3) Employment rate (sub-region of students residence). The most important factor was GDP per capita (by sub-region of student residence) and the second one was distance. The percentage of correct of prediction is 96%. The most important indicator is the GDP per capita at the student's sub-region of residence.

5. Conclusions

Higher education in Hungary is strongly concentrated to Budapest. Economic factors are important in student choices especially for students who want to learn Business Studies. If applying to non-Budapest institutes the importance of faculty excellence is more relevant.

For students applying to Business studies, the economic factors: GDP per capita, employment rate are more important than for students applying to Humanities and Arts. Students' first order application represents their

most conscious choices. Their second, third and fourth order choices are based on mainly the distance of the institution from their place of residence other factors gradually lose their relevance.

References

Ahmet G., Büsra K., Ferda M., Çetin D., Kübra G., & Hasan Y. H. (2011). Determining the relationship between students' choice of profession and mission and vision of their high school. *Procedia Social and Behavioral Sciences*, 15, 2595–2598.

Alm, J., John V., & Winters, J. V. (2009). Distance and intrastate college student migration. *Economics of Education Review*, 28, 728–738.

Asari, F. F. A. H, Idris, A R, & Daud, N. M. (2011). Modelling Education Tourism Using Gravity Model in Malaysian Public Higher Education Institutions. *Australian Journal of Basic and Applied Sciences*, 5, 1257-1274.

Bruno, G., & Improta, G. (2008). Using gravity models for the evaluation of new university site locations: A case study. *Computers & Operations Research*, 35, 436 – 444.

Coelli, M. B.(2011). Parental job loss and the education enrollment of youth. *Labour Economics*, 18, 25–35.

DesJardins, S. L., Dundar, H., & Hendel, D. D.(1999). Modeling the College Application Decision Process in a Land-Grant University. *Economics of Education Review*, 18, 117-132.

Dijk, B., Fok, D., & Paap, R.(2012). A Rank-Ordered Logit Model with Unobserved Heterogeneity in Ranking Capabilities. *Journal of Applied Econometrics*, 27(5), 831–846.

Duchessi, P., & Lauría, E.J.M.(2013). Decision tree models for profiling ski resorts' promotional and advertising strategies and the impact on sales. *Expert Systems with Applications*, 40, 5822–5829.

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.

Fayyad, U M., & Shapiro, P. G., Smyth, P. (1996). From data mining to knowledge discovery: An overview. In Advances in Knowledge Discovery and Data Mining., AAAI Press/The MIT Pres, 1–34.

Flannery, D., Kennelly, B., Doherty, E., Hynes, S., & Considine, J. (2013). Of mice and pens: A discrete choice experiment on student preferences for assignment systems in economics. *International Review of Economics Education*, 14, 57-70.

Frankel J., & Rose A. (2002). An Estimate of the Effect of Common Currencies on Trade and Income. *Quarterly Journal of Economics*, 117(2), 437-466. doi:10.1162/003355302753650292

Germeijs, V., & Verschueren, K. (2007). High school students' career decision-making process: Consequences for choice implementation in higher education. *Journal of Vocational Behavior*, 70, 223–241.

Gibbons, S., & Vignoles, A. (2012). Geography, choice and participation in higher education in England. *Regional Science and Urban Economics*, 42, 98–113.

Han, J., & Kamber, M. (2006). Data mining: concepts and techniques (Second Edition), Morgan Kaufmann Publisher

Jung, J., Diane M. Hall, H., & Rhoads, T.(2013). Does the availability of parental health insurance affect the college enrollment decision of young Americans? *Economics of Education Review*, 32, 49-65.

Langerné-Rédei M. (2007). The Hungarian migration regime - from talent loss to talent attraction. *Geographical Forum / Revista Forum Geographic*, 5(6), 134-145.

Larose, S. Cyrenne, D., Garceau, O., Harvey, M., Guay, F., & Deschęnes, C. (2009). Personal and social support factors involved in students' decision to participate in formal academic mentoring. *Journal of Vocational Behavior*, 74, 108-116.

Long, B. T.(2004). How have college decisions changed over time? An application of the conditional logistic choice model. *Journal of Econometrics*, 121, 271 – 296.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. Frontiers in Econometrics, ed. by P. Zarembka, Academic Press, New York.

McFadden, D., & Train, K. (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, 15(5), 447-470.

Montgomery, M.(2000). A nested logit model of the choice of a graduate business school. *Economics of Education Review*, 21, 471-480.

Montmarquette, C., Cannings, K., & Mahseredjian, S.(2002). How do young people choose college majors? *Economics of Education Review*, 21, 543-556.

Nagy, G. (2005). Changes in the Position of Hungarian Regions in the Country's Economic Field of Gravity. In: Barta, Gy., Fekete, É.G., Szorenyine Kukorelli, I., Tímár, J. (eds.): Hungarian Spaces and Places: Patterns of Transition, 124-142. Centre for Regional Studies, Hungarian Academy of Sciences, Pécs, ISBN 963 9052 46 9

Niu, S. X., & Tienda, M. (2008). Choosing colleges: Identifying and modeling choice sets. *Social Science Research*, 37, 416–433.

Reynolds, L. C. (2012). Where to attend? Estimating the effects of beginning college at a two-year institution. *Economics of Education Review*, 31, 345-362.

Rochat, D., & Demeulemeester, J. (2011). Rational choice under unequal constraints: the example of Belgian higher education. *Economics of Education Review*, 20, 15–26.

Rodrigue, J.-P., Comtois, C., & Slack, B. (2009). The Geography of Transport Systems. London, New York: Routledge. ISBN 978-0-415-48324-7.

Saisana, M., d'Hombres, B., & Saltelli, A. (2011). Rickety numbers: Volatility of university rankings and policy implications. *Research Policy*, 40, 165–177.

Schwartz, B. (1985). Student Financial Aid and the College Enrollment Decision: the Effects of Public and Private Grants and Interest Subsidies. *Economics of Education Review*, 4(2), 129-144.

Shtatland, E. S., Kleinman, K., & Cain, E. M. (2008). One more time about R2 measures of fit in logistic regression,NESUG 15, Statistics, Data Analysis & Econometrics

Telcs A., Kosztyán Zs. T., & Török Á. (2013). Unbiased One dimensional University Ranking – Application Based Preference Ordering. *Journal of Applied Statistics*. (under publication)

Toutkoushian, R. K. (2001). Do parental income and educational attainment affect the initial choices of New Hampshire's college-bound students? *Economics of Education Review*, 20, 245–262.

Vrontis, D., Alkis Thrassou, A., & Melanthiou, Y. (2007). A contemporary higher education student-choice model for developed countries. *Journal of Business Research*, 60, 979–989.

Weiler, W. C. (1989). A Flexible Approach to Modelling Enrollment Choice Behavior. *Economics of Education Review*, 8(3), 227-283.

Weiler, W. C. (1986). A Sequential Logit Modell of the Access Effects of Higher Education Institutions. *Economics of Education Review*, 5(1), 49-55.