AN ASSOCIATIVE AND PROJECTIVE DIAGNOSIS FOCUSED ON THE ANALYSIS OF STATISTICAL DATABASES CHARACTERISTIC OF ACADEMIC EDUCATION

Assoc. prof. habil. Gheorghe Săvoiu
University of Piteşti

Abstract

The paper presents a distinct manner of statistical analysis, defined by an original associative and projective diagnosis of a set of 10 academic study programmes of the University of Pitesti, generating 244 variables drawn from a larger study that is part of the POSDRU (HRD) contract 156/1.2/G/141632, part of the project known, in its abbreviated form, as NOVA-CURRICULA, a project co-funded by the European Social Fund through the Operational Programme Human Resources Development 2007-2013: "Investing in people". The associative and projective diagnosis hereby proposed as an innovative method is at the same time regional and county-oriented, referring strictly to the South-Muntenia region and the Argeş county, as the extended pool and the representative pool, in keeping with the statistic 60/60 principle, for the University of Pitesti, and the 10 programmes analysed (in a diverse range, from academic programs in engineering to nursing programmes, etc.). The article tries to find, and – in the author's opinion – succeeds in identifying, among the multitude of variables aggregated in the database of a multidimensional analysis, associations and correlations that should meet the requirement of specifically modelling the educational supply, in the selection pool, an extensive set of ten academic programmes, using as assumptions and priorities the Europe 2020 strategy, that is an academic education able to link with the labour market, and especially lead to intelligent, sustainable and socially inclusive growth, by means of a software package (EViews); finally, there are a number of remarks and anticipations necessary for a general delimitation of new skills, which are in fact requisite throughout the national academic system, not only in the regional and county real market.

Keywords: associative and projective diagnosis, matrix and correlation ratio, multidimensional analysis, statistic association and correlation, educational modelling.

Introduction

The associative and projective diagnosis briefly presented below represents a deeply original application of econometric thinking to a set of 244 variables, extracted, according to multidimensional analysis, in the previous chapters of the study, variables that are at the same time regional and county-oriented, with strict reference to the South-Muntenia region and the Argeş county, as the extended pool and, respectively, the representative pool in accordance with the 60/60 statistic principle, for the

University of Piteşti, and of the 10 programmes analysed (in a diverse range, from academic programmes in engineering to nursing programs, etc.).

This approach has tried to find, and has succeeded in identifying, in the variables aggregated in the database of multidimensional or multivariate analysis, associations and correlations that should meet the requirement of modelling the educational supply in the selection pool of the ten programmes of the University of Piteşti, using as assumptions and priorities the Europe 2020 strategy, i.e. an academic education able to secure a connection with the labour market, and especially lead to "an intelligent, sustainable or environment-friendly and inclusive growth, that is focused on a high employment rate, which could promote both social cohesion and spatial structural consistency, i.e. regionally and county-wise" (European Commission, Europe 2020).

1. Methodology, main methods and statistical tools

Three previous studies, focused on three distinct questionnaires of the diagnosis, were correlated in the project *Nova Curricula*, so that one might define and draw, from questionnaire 1, the samples likely to provide freely expressed answers and opinions, using voluntary service, which provided a number of 400 students for questionnaire 2, and 100 companies and institutions for questionnaire 3 (due to reasons of brevity in drafting the paper, questionnaires 1, 2 and 3 are not attached, but they can be at any time requested of the author (address: gheorghe.savoiu@upit.ro).

The multidimensional character of the analysis was generated by the initial identification of a set of key dimensions, as studies 1, 2 and 3 focused, in their initial part, on selecting a large number of characteristic variables, and variables of interest with respect to the relationship between academic education and labour market in terms of those dimensions. The construction of the 244 variables included in the final database, which was capitalized on in this associative and projective diagnosis, was first made by processing the 10 type 1 questionnaires completed by the ten managers responsible for the ten undergraduate academic educational programmes (which generated 40 endogenous variables), and the remaining 204 exogenous variables resulted from the multidimensional analysis carried out in the South-Muntenia region and the Argeş county (by forming a set of characteristic regional and county-oriented dimensions).

Thus, the final set of dimensions distinctly addressed the characteristics pertaining to, or connected with, geographical location and regional network, demography, labour, education, school population and its specific institutions, justice and crime, health and the institutions that contribute to ensuring it, or those characteristics derived from such economic performance features and results as GDP per capita and net incomes, as well as those characteristics associated with motor vehicle the industry and specific to route transport.

The intrinsic logic of statistical thinking, which was behind this analysis, is supported, in point of methods and tools, by the determination coefficient initially quantified as the squared value of the correlation ratio (R^2) belonging to the known

interval [0; 1] and tending to a unit value, i.e. to one, all the more as the intensity of the correlation is higher between the two associated variables, one of which is by definition independent or exogenous (xi) and belongs to the labour market, and the other is endogenous or dependent (yi), i.e. the number of students enrolled or of the graduates of a BA academic education programme, in this article. The specialized software packages have of the EViews type allowed very fast calculations of the intensity of associations or correlations, which they also display promptly via a set of correlation matrices, bringing together all the possible interrelationships between factors by means of all possible values of the correlation ratio (R).

The correlation ratio (R) is a parameter of normal bi-dimensional distributions that characterize the relationship between the independent variable x_i and the dependent variable y_i, expressed mathematically in the general form yi = f(xi) belongs to interval [-1,0) È (0,1]. For R=0 there is no correlation or statistical association (a relation, in general terms, or in common language), with x_i and y_i independent variables. The interpretation of a correlation ratio is achieved briefly in mathematical mode (R), and the following statements are considered accurate empirically: i) if |R| Î (0;0,2] there is virtually no statistical relationship, or it is very weak; ii) if |R| Î (0,2;0,5] the link is weak and it should be tested statistically (F-test, etc.); iii) if |R| Î (0,5;0,75] the link or connection is of average intensity; iv) if |R| Î (0,75;0,95] the connection is strong; v) if |R| Î (0,95;1] the link or connection is strong, of a deterministic type (so, functional).

2. Useful associations and correlations for modelling educational supply in the selection pool

Associative and projective diagnosis of the number of students enrolled in the 10 undergraduate academic education programs in the South - Muntenia region and the Argeş county as an extended pool, and respectively, the representative pool of the University of Piteşti offer useful associations and correlations for the regional academic educational evolution.

The 244 resulting variables or data sets were analysed one by one, starting from the first ten variables, i.e. the number of students enrolled in the 10 undergraduate academic educational programmes, and multiple correlations were identified, which were significant by their intensity and major in point of essence modelling the associative and projective diagnosis. With respect to simplifying and accessibility of the presentation of the diagnosis, the results were summarized in Tables no. 1-10 (with the usual abbreviations, RSM for the South-Muntenia Region, and AG for the Argeş county):

Total number of students enrolled in the 1st programme (SER01)

Table no. 1

Associated / correlated exogenous variables of average intensity at least, or |R|∈(0,5;1]	Value of R
SER49 = Stable population as of 1 July, county of AG, male	0.512
SER79 = Internal migration – RSM – Rural – Migratory balance	0.521
SER142 = Number of students – county of AG – post-high-school units	0.512
SER157 = University-age population – RSM	0.575
SER158 = University-age population – county of AG	0.536

Software used: Eviews.

The first identified aspect is a correlation of an average intensity of the SER01 with the evolution of school population in the selection pool.

Total number of students enrolled in the 2nd programme (SER02)

Table no. 2

Associated / correlated exogenous variables of average intensity at least, or $\|R\| \in (0,5;1]$	Value of R
SER43 = Stable population as of 1 July, RSM rural	0.644
SER48 = Stable population as of 1 July, county of AG rural	- 0.661
SER51 = Stable population as of 1 July, RSM 0–14 years	- 0.685
SER123 = Number of the unemployed out of the total – RSM	- 0.796
SER174 = Network of health units – RSM – medical units – medical specialty	0.651
SER195 = total GNP – in million lei RON current prices-county of AG	0.665
SER205 = Average net nominal earnings per month – RSM – lei RON current prices Education	0.691
SER206 = Average net nominal earnings per month – RSM – lei RON current prices Health and social welfare	0.661
SER231 = Evolution of turnover in local units of manufacturing activities – RSM – million lei RON current prices 10 – 49 employees	0.657
SER234 = Evolution of turnover in local units of manufacturing activities – county of AG	0.684
SER235 = Evolution of the number of enterprises active in the activities of transport, storage and communications – RSM – total	0.667
SER236 = Evolution of the number of enterprises active in the activities of transport, storage and communications – RSM – 1– 9 employees	0.677
SER240 = Evolution of the number of enterprises active in the activities of transport, storage and communications – county of AG – total	0.634
SER241 = Evolution of the number of enterprises active in the activities of transport, storage and communications – county of AG – 1– 9 employees	0.640

Software used: Eviews.

SER02 is correlated with the evolution of school population in the selection pool, and with the labour market, but with alternative signs or directions/

Total number of students enrolled in the 3th programme (SER03)

Table no. 3

Associated / correlated exogenous variables of average intensity at least, or $\|R\| \in (0,5;1]$	Value of R
SER172 = Network of medical and health units – school medicine	0.650
SER76 = Internal migration – RSM – Urban – Migratory balance	0.660
SER180 = Network of medical and health units – county of AG – medical units – school medicine	-0.669
SER200 = Average net nominal earnings per month – RSM – lei RON current prices Agriculture	-0.608
SER204 = Average net nominal earnings per month – RSM – lei RON current prices Transport and storage	-0.603
SER239 = Evolution of number or companies active in activities of transport, storage and communications – RSM – $\geq$ 250 employees	-0.607

Software used: Eviews.

SER03 is not correlated with the evolution of school population in the selection pool, and only relatively with the labour market.

Total number of students enrolled in the 4th programme (SER04)

Table no. 4

Associated / correlated exogenous variables of average intensity at least, or $\|R\| \in (0,5;1]$	Value of R
SER125 = Number of de unemployed that are university graduates – RSM	0.638
SER126 = Number of de unemployed that are university graduates – county of AG	0.674
SER169 = Network of sanitary units – RSM – medical laboratories	0.568
SER174 = Network of sanitary units – RSM – medical units – medical specialities	0.512
SER180 = Network of sanitary units – county of AG – medical units – school medicine	0.675
SER182 = Network of sanitary units – county of AG – medical units – medical specialities	0.516
SER201 = Average net nominal earnings per month – RSM – lei RON current prices Industry	0.554
SER202 = Average net nominal earnings per month – RSM – lei RON current prices Manufacturing industry	0.554
SER203 = Average net nominal earnings per month – RSM – lei RON current prices Constructions	0.538
SER204 = Average net nominal earnings per month – RSM – lei RON current prices Transport and storage	0.553
SER205 = Average net nominal earnings per month – RSM – lei RON current prices Education	0.558
SER206 = Average net nominal earnings per month – RSM – lei RON current prices Health and social welfare	0.553
SER213 = Number of private entrepreneurs – RSM – family businesses	-0.523
SER214 = number of private entrepreneurs – county of AG – family businesses	-0.522
SER223 = Evolution according to the share of the number of employees by size class – Romania – no employees	0.565
SER239 = Evolution of the number of companies active in activities of transport, storage and communications – RSM – ≥ 250 employees	0.870
SER244 = Evolution of the number of companies active in activities of transport, storage and communications – county of AG – ≥ 250 employees	0.685

Software used: Eviews.

SER04 is reduced as number of terms (years) and is not correlated with the evolution of school population in the selection pool, and relatively or somewhat apparently with the labour market.

Total number of students enrolled in in the 5th programme (SER05)

Table no. 5

Associated / correlated exogenous variables of average intensity at least, or $\|R\| \in (0,5;1]$	Value of R
SER52 = Stable population as of 1 July, RSM 15–59 years	-0.799
SER53 = Stable population as of 1 July, RSM ≥ 60 years	0.781
SER55 = Stable population as of 1 July, county of AG 15–59 years	-0.790
SER56 = Stable population as of 1 July, county of AG ≥ 60 years	0.808
SER129 = Network of education units-RSM- secondary schools	-0.748

Software used: Eviews.

SER05 is low in point of number of years and is correlated neither to the school population in the selection pool, nor to the labour market (this is just a stable variable associated with stable population only).

Total number of students enrolled in in the 6th programme (SER06)

Table no. 6

Associated / correlated exogenous variables of average intensity at least, or |R|∈(0,5;1]	Value of R
SER145 = Graduates – RSM – university	0.683
SER184 = Staff of health units – RSM – auxiliary medical personnel	0.684
SER185 = Staff of health units – RSM – average medical personnel	0.682
SER186 = Staff of health units – RSM – doctors (exclusive of dentists)	0.698
SER187 = Staff of health units – RSM – family doctors	0.709
SER188 = Staff of health units – RSM – dentists	0.663
SER189 = Staff of health units – county of AG – auxiliary medical personnel	0.654
SER190 = Staff of health units – county of AG – average medical personnel	0.684
SER222 = Share of start-ups going on for one year – RSM – closed	0.652

Software used: Eviews.

SER06 is correlated not with the evolution of the school population in the selection pool and the labour market, but is associated with health units.

Total number of students enrolled in in the 7th programme (SER07)

Table no. 7

Associated / correlated exogenous variables of average intensity at least, or |R|∈(0,5;1]	Value of R
SER43 = Stable population as of 1 July, RSM rural	-0.702
SER48 = Stable population as of 1 July, county of AG rural	-0.853
SER51 = Stable population as of 1 July, RSM 0–14 years	-0.722
SER54 = Stable population as of 1 July, county of AG 0–14 years	-0.700
SER123 = Number of unemployed out of the total – RSM	-0.741
SER128 = Network of education units – RSM – schools – total	-0.749
SER133 = Network of education units – county of AG – schools – total	-0.729
SER158 = University-age population – county of AG	-0.713

Software used: Eviews.

SER07 is intensely correlated with the evolution of school population in the selection pool and relatively with the specific labour market.

Total number of students enrolled in in the 8th programme (SER08)

Table no. 8

Associated / correlated exogenous variables of average intensity at least, or $\|R\| \in (0,5;1]$	Value of R
SER145 = Graduates – RSM – university	0.717
SER148 = Graduates – county of AG – university	0.719
SER194 = Total GNP – in million lei RON current prices – RSM	0.700
SER195 = Total GNP – in million lei RON current prices – county of AG	0.726
SER196 = GNP per capita – lei RON current prices – RSM	0.708
SER199 = Average net nominal earnings per month – RSM – lei RON current prices	0.707
SER203 = Average net nominal earnings per month – RSM – lei RON current prices Constructions	0.706
SER204 = Average net nominal earnings per month – RSM – lei RON current prices Transport and storgae	0.716
SER205 = Average net nominal earnings per month – RSM – lei RON current prices Education	0.770
SER206 = Average net nominal earnings per month – RSM – lei RON current prices Health and social welfare	0.750
SER231 = Evolution of turnover in local units of manufacturing activities – RSM – milioane lei RON current prices 10 – 49 employees	0.710
SER234 = Evolution of turnover in local units of manufacturing activities – county of AG	0.717
SER235 = Evolution of number of companies active in activities of transport, storage and communications – RSM – total	0.718
SER236 = Evolution of number of companies active in activities of transport, storage and communications – RSM – 1– 9 employees	0.730

Software used: Eviews.

SER08 is correlated with the evolution of school population in the selection pool and rather relatively and apparently with the labour market

Total number of students enrolled in in the 9th programme (SER09)

Table no. 9

Associated / correlated exogenous variables of average intensity at least, or $\|R\| \in (0,5;1]$	Value of R
SER48 = Stable population as of 1 July, county of AG rural	-0.814
SER49 = Stable population as of 1 July, county of AG male	-0.821
SER50 = Stable population as of 1 July, county of AG female	-0.804
SER51 = Stable population as of 1 July, RSM 0–14 years	-0.850
SER123 = Number of the unemployed out of the total – RSM	-0.844
SER133 = Network of education units – county of AG – schools – total	-0.859
SER157 = University-age population – RSM	-0.813
SER158 = University-age population – county of AG	-0.856

Software used: Eviews.

SER09 is only correlated with the evolution of school population in the selection pool.

Total number of students enrolled in the 10th programme (SER10)

Table no. 10

Associated / correlated exogenous variables of average intensity at least, or $\|R\| \in (0,5;1]$	Value of R
SER41 = Stable population as of 1 July, RSM total	0.912
SER42 = Stable population as of 1 July, RSM urban	0.824
SER43 = Stable population as of 1 July, RSM rural	0.927
SER44 = Stable population as of 1 July, RSM male	0.915
SER45 = Stable population as of 1 July, RSM female	0.909
SER46 = Stable population as of 1 July, county of AG total	0.856
SER48 = Stable population as of 1 July, county of AG rural	0.827
SER49 = Stable population as of 1 July, county of AG male	0.847
SER50 = Stable population as of 1 July, county of AG female	0.866
SER51 = Stable population as of 1 July, RSM 0–14 years	0.910

Software used: Eviews.

SER10 is correlated exclusively with the evolution of the stable population in the selection pool.

The associative and projective diagnosis that was proposed in an original manner in this paper is effective, synthetic and quite prompt (although devising and making up the database represents, as could be noticed, a very laborious process), which indicates the main trends of association and finally focuses on hypotheses scientifically grounded and formulated. Most of the programmes are correlated with the declining dynamics of the population, and especially with school age groups, while more than half (about seven out of ten) are correlated with the labour market, though many of them only in an apparent or rather illusory manner (about half appear to be somewhat correlated with both the inputs in the system and the outputs of the labour market). Illusory correlations can be also generated by some hidden associations, but there is no doubt that such analyses based on correlation matrices are necessary in order to identify the extent to which the programmes have their own policies for labour market adjustment and suitability.

3. Some final remarks and expectations necessary to adapt the curricula analyzed

Some general remarks show that the major endogenous variables generated by the number of students enrolled, or the series representing the dependent variable, are nearly all correlated with the school resources in the main demographic pool (except for the series that describe programmes III, IV, V, since the last two have been active for only 5 and 3 years, respectively).

There are also co-existing objective arguments relative to the mismatch of the number of enrolled students and graduates with the activities and the variables derived from regional and county labour market (programme VIII has been asked for a long time in the extended EU market), but many of the programmes train graduates in a market characterized by contraction or relatively saturated within specific sectors, aspects that are reflected in the correlation matrices that have over 30% illusory associations.

About half of the programmes are heterogeneous and strongly polarized over time, according to the databases in the last 15 years, but only three of those show risks and high developmental vulnerabilities or abnormalities, more specifically programs III, IV and

X. The dynamics of the graduates, alongside the evolution of those who were employed in jobs requiring higher education, identifies, after the multidimensional analyses and the descriptive statistical analysis with the analysis of association through correlation matrices in three other programs I, VI and IX, the only completely risk-free programmes in the descriptive statistics of the data, despite the shrinking school population in the selection pool of the University of Piteşti. The current skills need being completed by two necessary meta-competencies related to inter-, trans- and multi-disciplinarity, as well as ethics and professional conduct. In all dimensions there are favourable as well as unfavourable aspects, and the analysis of the 244 variables clearly demonstrates that. Among the things that are requisite in the short term future, in keeping with the present analysis, one can mention: a) the need to generalize the various valuable active practices that can substitute other practices, which proved harmful; b) continuous documentation and ongoing investigation are required of the aspects revealed by the study presented; c) irrespective of the expected results, the effectiveness of the academic education action, regionally, and the effects expected as the result of the new skills, or the additional meta-competences that will be proposed, cannot come into being without modern management being provided, centring on stimulating educational, economic and social policies, emphasizing the role of the monitoring component in the academic management decision-making for the undergraduate (BA) programmes, a dual and concurrent kind of monitoring of the results of the programmes and the quality of the graduates, as well as the labour market, and ensuring adjustment and appropriateness to the dynamism of the latter.

Acknowledgements

The article is based on a set of issues drawn from a larger study that is part of the POSDRU (HRD) contract 156/1.2/G/141632 within the project "Adapting academic study programs to the National Framework of Qualification in Higher Education in order to improve the educational supply, creating and implementing specific information instruments intended to ensure the expansion of learning opportunities and interaction with the business milieu (NOVA-CURRICULA)", co-funded by the European Social Fund through the Sectoral Operational Programme Human Resources Development 2007-2013: "Investing in people", which was also conducted in the University of Piteşti. We hereby thank the project manager and the entire project team for the opportunity to apply elements of statistical thinking to analyses of academic teaching processes.

References

1. Radermacher, W., *Foreword*, in Eurostat Regional Yearbook, 2013, pag. 3 disponibil online la: http://epp.eurostat.ec. europa index. php/ Regional_yearbook_foreword accesat în 25 iulie 2014.
2. Săvoiu G., 2004. *Statistică aplicată în domeniul economic şi social*. Ed. Independenţa Economică, Piteşti.
3. Săvoiu G., 2006. *Populaţia lumii între explozie şi implozie demografică*,Ed. International University Press, Bucureşti.
4. Săvoiu, Gheorghe, Ion Iorga–Simăn. 2008. *Some Relevant Econophysics' Moments of History, Definitions, Methods, Models and New Trends*. *Romanian Economic and Business Review* 3 (3): 29-41.

5. Săvoiu, G.,2009. *Statistica. Mod de gândire şi metode* Editura Universitară, Bucuresti.
6. Săvoiu, G., Crăciuneanu, V., Ţaicu, M. 2010. A New Method of Statistical Analysis of Markets' Concentration or Diversification. *Romanian Statistical Review*, 58(2), pp.15- 27.
7. Săvoiu, G., 2011. *Statistică pentru afaceri*, Editura Universitară Bucureşti.
8. Săvoiu, G, 2012. *GDP Indicator for Statistical Comparisons at National/Regional and International*, Romanian Statistical Review, vol.60(12), pp. 54-62.
9. Săvoiu, G., 2013. *Modelarea Economico-Financiară: Gândirea econometrică aplicată în domeniul financiar,* Editura Universitară, Bucureşti.
10. Săvoiu G., Vladu, M., 2012. *Decizia şi modelarea în procese educaţionale innovative Educaţia din perspectiva valorilor, Idei concepte, modele, Cluj Napoca Eikon, pp. 189 -207.*
11. *** Comisia Europeană, Europa 2020, *O strategie europeană pentru o creştere inteligentă, ecologică şi favorabilă incluziunii* , Bruxelles, 3.3.2010 COM(2010) 2020 final disponibilă on line la: http://www.mae.ro/sites/default/files/file/Europa2021/Strategia _Europa_2020. pdf accesat 8 august 2014
12. *** *Statistica teritorială, repere economice şi sociale regionale*, colecţia 2000- 2014, Ed. INS, Bucureşti şi seriile de date disponibile on line la şi http://www.insse.ro/cms/files/IDDT%202012/index_ IDDT. htm accesate în iulie 2014.
13. *** *Educaţia în România*, 2014, Ed INS. Bucureşti.