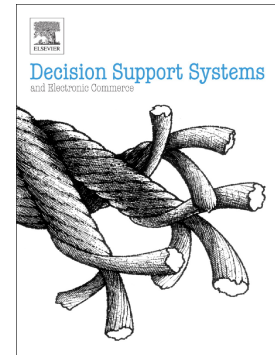


Applying data driven decision making to rank vocational and educational training programs with TOPSIS

J.M. Conejero, J.C. Preciado, A.E. Prieto, M.C. Bas, V.J. Bolós



PII: S0167-9236(20)30225-6

DOI: <https://doi.org/10.1016/j.dss.2020.113470>

Reference: DECSUP 113470

To appear in: *Decision Support Systems*

Received date: 8 March 2020

Revised date: 7 August 2020

Accepted date: 2 December 2020

Please cite this article as: J.M. Conejero, J.C. Preciado, A.E. Prieto, et al., Applying data driven decision making to rank vocational and educational training programs with TOPSIS, *Decision Support Systems* (2020), <https://doi.org/10.1016/j.dss.2020.113470>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Applying Data Driven Decision Making to rank Vocational and Educational Training Programs with TOPSIS

J. M. Conejero^a, J. C. Preciado^a, A.E. Prieto^a, M. C. Bas^b, V. J. Bolós^b

^a*Dept. Ingeniera Sistemas Informáticos y Telemáticos, Universidad de Extremadura, Cáceres, Spain*

^b*Dept. Matemáticas para la Economía y la Empresa, Universidad de Valencia, Valencia, Spain*

Abstract

In this paper we present a multi-criteria classification of Vocational and Educational Programs in Extremadura (Spain) during the period 2009-2016. This ranking has been carried out through the integration into a complete database of the detailed information of individuals finishing such studies together with their labor data. The multicriteria method used is TOPSIS together with a new decision support method for assessing the influence of each criterion and its dependence on the weights assigned to them. This new method is based on a worst-best case scenario analysis and it is compared to a well known global sensitivity analysis technique based on the Pearson's correlation ratio.

Keywords:

Vocational Education and Training ranking, Employability, TOPSIS, Data-driven decision making, Multicriteria decision analysis, Criteria influence analysis

1. Introduction

The 2008 financial crisis that hit the world's economies has had a particularly acute impact in Spain [1]. It is only since 2014 that Spain seemed to begin its recovery [2]. However, this recuperation is still far to be acceptable with regard to the labor landscape [3].

One of the main Spanish weaknesses that the crisis exposed was the so-called duality of the labor market. Thus, Spain is characterized by the existence of two very different types of workers.

Email addresses: chemacm@unex.es (J. M. Conejero), jcpreciado@unex.es (J. C. Preciado), aeprieto@unex.es (A.E. Prieto), maria.c.bas@uv.es (M. C. Bas), Vicente.Bolos@uv.es (V. J. Bolós)

On one hand, long term workers on indefinite contracts, having both a very high job security and a very high cost for companies (especially in terms of dismissals) and usually with university studies even for jobs that do not require them. On the other hand, short term workers on temporary contracts or seasonal contracts with low wages and, in most cases, with very little training.

Another structural weakness of the Spanish economy unveiled during the years of the crisis was the fact that it had been relied heavily on two pillars: construction and tourism (and their associated services). This productive model had its main Achilles heel in the low level studies required in many of the jobs created in both sectors. Moreover, the relatively high wages that a worker could earn before the crisis, mainly in construction, led many young people to abandon their studies to work in these industries, without prior quality training. When the crisis arose and the destruction of employment reached unprecedented level, Spain discovered that had to deal with a mass of unemployed, mostly young, people who, having no adequate training, had very difficult or even impossible reinstatement into the labor market.

This problem has been most pronounced in some regions of Spain as Extremadura. Extremadura is a European Union Objective 1 region located in western Spain that according to the Eurostat Regional Yearbook 2018¹, its GDP per inhabitant in relation to the EU-28 average is 61.47%, it has 23.7% of unemployment rate and, even worse, its early leavers from education and training of young people rate is 20.9% and its young people neither in employment nor in education or training rate is 20%.

In other countries where the economic model was more diversified, with large sectors of skilled employment and better trained workers, the crisis was less intense, the employment destruction less acute, and the recovery was faster. One of the differences between Spain and, particularly, Extremadura with respect to those countries is the importance they give to Vocational Education and Training (hereinafter VET). For the European Union, VET should “prepare young people for entering and successfully and sustainably participating in the labor markets as well as to enable high potentials (e.g. migrants, refugees, low-skilled and unemployed, inactive groups, including women) to stay and/or (re-)enter the labor market” [4]. In Germany, for example, VET studies are closely linked to the labor market, so that the majority of VET students are also trained in companies where, in many cases, they end up working. This means that there is a quarry of workers with a specific qualification for the needs of the labor market.

¹ <https://cutt.ly/as19ww2>

However, in the collective imagination of Spanish families, VET has been considered for decades a second-rate training and has not been much appreciated. This vision, together with the high drop-out rates, has caused a very marked duality in training: on the one hand people who either did not finish more than compulsory education or, at best, have VET studies (which in this last case are seen wrongly as a low level education), or people with university education who, due to the high unemployment rates suffered in Spain during the crisis (still persisting), are hired in positions for which such studies are not really required, causing another of the many big problems of the Spanish labor market, which is the overqualification of workers [5].

In order to try to alleviate the mentioned problems, the Government of Extremadura, providing its historical data, asked for a scientific analysis about the real impact of VET studies on accessing to the labor market with a two-fold goal: increase the resources of those VET studies with higher employment rate and promote such studies among their unemployed citizens enhancing the image of the VET studies that really help to get a job.

To this end, the aim of this work is to determine the efficiency of VET studies and the evaluation of the performance of VET graduates from Extremadura in the different degrees of VET programs in the labor market. Thus, we illustrate a data-driven multi-criteria decision-making methodology with the aim of classifying the different degrees of VET programs according to some criteria related to labor insertion. Concretely, we have applied TOPSIS (Technique for Order Performance by Similarity to Ideal Solution) [6] to this problem. TOPSIS is a well-known classical MCDM method, widely used by researchers and practitioners, that supports decision makers in performing analysis, comparisons and rankings to select the best alternative using a finite number of criteria. Moreover, since all criteria are weighted, their importance may be modified providing, thus, the flexibility for creating different rankings prioritizing different aspects. In our case, we have considered 8 different criteria whose values have been computed from a real dataset with more than 28000 VET student records containing both educational and labor information. This is a key contribution for this work since the information obtained is enriched by real data instead of being based on questionnaires as similar TOPSIS approaches [7] or the REFLEX project in the field of higher education [8]. Furthermore, to the best of our knowledge, the relationship between VET programs and labor market has not been explored by previous works with such level of detail.

The rest of the article is organized as follows. Section 2 reviews some works related with

predicting different outcomes using academic data, and some other works using TOPSIS as Multi-Criteria Decision Analysis method together with a weight assignment analysis, on different scientific domains. Section 3 describes the datasets used and the process applied to them for computing the different data used in our study. Section 4 describes the methodology followed to apply TOPSIS. Section 5 explains the influence of the criteria applied during the process. The results obtained and further considerations are detailed in Section 6. Finally, Section 7 concludes the paper.

2. Literature review

The analysis of educational data is considered today as one of the foundations to implement new educational policies and may be even more so in the future [9].

In this sense, there are several works focused on analysing success or failure in the pre-university stages. Thus, one example of this topic is the work of Sen et al. [10] that tries to predict the outcome of Turkish high school students in the examinations of national selection that these students must perform. Another example is the work of Sara et al. [11] where they analyze the data of more than 36.000 Danish students who, at least, have completed the first six months in secondary education institutes, to predict if they were going to abandon their studies in the next 3 months. Aguiar et al. [12] analyze 11.000 American students from different high school courses to try to detect who were at risk of abandonment and try to take appropriate measures so that it does not occur.

There are also a number of works within this field focused on university students. One example is the work of Campagni et al. [13], in which the authors analyze the academic data to define an ideal trajectory for university students and compare real students based on that ideal trajectory. Di Pietro et al. [14] analyze the academic performance in terms of student performance, teaching activities and student satisfaction among others. Continuing toward the goal of predicting the student performance in the university, Asif et al. [15] make use of data mining and clustering methods to predict the graduation performance of a student given only his pre-university marks. Fernández-García et al. [16] perform a comparison of algorithms that can predict which first-year university students are more likely to finish their studies and apply their conclusions in the students of the University of Almería (Spain).

More related with our work, that is, analyzing the possible relationship between the education of the students and their employability, we could start by mentioning the one of Jackson [17] where she analyzes data of 56255 Australian Bachelor degree graduates in 2011 and 2012 to identify which factors influence graduate employment. Also, in Australia, Mewburn et al. [18] perform an analysis of the extent of demand for Ph.D. student skills and capabilities in the Australian employment market. Bharambe et al. [19] analyse 5 skills of 91177 Indian students to predict the likelihood of each one of getting a job. Thakar et al. [20] use 151 attributes from 7143 Indian university students to conclude that only 8 attributes play a significant role in predicting students' employability in the first year of their enrollment. Garca-Peñalvo et al. [21] use information from 3000 Spanish students to build predictive models that define how these students get a job after finalizing their university degrees.

From the analysis of the state of the art in the domain of this project, we can observe that, although there are numerous works that use data from the educational field to try to predict possible dropouts, performance and even employability, we do not have evidence of anyone trying to analyze the possible relationship between VET programs and employment.

Concerning the methodological aspects of the present paper, several studies in different domains have used multi-criteria decision-making methods and, in particular TOPSIS methodology, to select the best alternative using a finite number of criteria. Indeed, in many studies new extensions or improvements for the TOPSIS methodology have been proposed. For instance, Yue [22, 23] proposes an extension of TOPSIS to determine the weights of the experts, considering a level of uncertainty assigned to the criteria. Olson [24] uses different distance metrics and shows how the results of the TOPSIS methodology depend on both the weighting scheme and the distance metric used. Other study [25] proposes an extension of the TOPSIS methodology (Fuzzy TOPSIS) combined with a weighting criteria methodology (Fuzzy AHP). In the last step of this method the authors apply sensitivity analysis based on the weights assigned by the experts to observe how the preferences of the decision makers would affect the final classification. To the best of our knowledge, the works that use sensitivity analysis as a final step of the TOPSIS or extended TOPSIS methodology are based on the weighting scheme assigned by the experts, regardless the weighting methodology used. The proposed decision support method (scenario comparison methodology) provides a useful tool for the experts that shows the sensitivity of each criterion to possible changes in the weighting scheme. It has the advantage of

not needing a previous definition of the criteria weights by the experts, and therefore the information delivered by the method is readily available to the expert panel before they make the weight assignment.

3. Data management

This section presents the datasets used to establish the different rankings presented in this paper together with the whole process followed to obtain them. To this purpose, a collaboration agreement was established with the data owners, the Education and Employment Board of the Government of Extremadura, to acquire the data. Based on this agreement, we established several meetings with them in order to identify the information available and needed to perform the analysis. Then, the process for obtaining the data was divided into three different steps: i) data sources identification; ii) data gathering and iii) data warehouse design. These steps will be explained in the next subsections.

3.1. Data sources identification

Since the main purpose of this study was to analyze the utility of the different VET programs in the region in terms of employability, the meetings with the Educational and Employment Board were mainly focused on identifying the data sources that provide information regarding those two areas. Fortunately, this information was acquired in the region by means of two different sources that directly depend on the aforementioned board:

- Rayuela² is the digital platform used by all the education centres to support the teaching staff in all the students management tasks such as exam realization, evaluation and so on. The platform includes data belonging to different education centres, concretely: Middle Schools, High Schools, VET Schools and Official Languages Schools. In this study we were mainly interested in the data from VET Schools since we wanted to analyze this kind of studies.
- Public Regional Employment Service (SEXPE)³ is the organization provided by the public administration to manage all the tasks related to labor in the region. This service manages

² <https://cutt.ly/Ws0d1Mq>

³ <https://cutt.ly/Gs0fobj>

any job contract signed in the region as well as the employment demands and subsidies for unemployed people in the region. Thus, all the information regarding the citizens' working life is provided by this public service.

Two additional sources of information were also used in this study:

- National Social Security Agency⁴ is a national organization that provides the public insurance to all the workers in the country. To this purpose, the Agency also recovers the labor information for the citizens independently of the region where they are working. The information provided was useful for complementing the data provided by the SEXPE service.
- Spanish National Statistics Institute (INE)⁵. This agency is responsible for publishing national statistics about different domains in the country. In this study we obtained the list of contract economical sectors from this source.

3.2. Data gathering

Based on the four data sources commented in previous section, the data owners provided us with the VET student's records stored in Rayuela for citizens since 1996. These data were also completed, on the one hand, with the job contracts stored by the SEXPE service for the citizens included in the previous dataset (the provided by Rayuela) and, on the other hand, the last information provided by the National Social Security Agency for these citizens (note that in both cases the information included the whole life for each citizen). The data provided by INE were also useful for completing the information available for each citizen.

Based on the different data sources, the files used in the study are summarized in Table 1 together with the number of instances, features and a description for each one.

File	Source	Features	Instances	Description
Graduates	Rayuela	14	28.272	Graduates data from VET Schools in Extremadura from 1996 to 2016.
Contracts	SEXPE	16	317.152	New employees contracts including the sector,

⁴ <https://cutt.ly/Xs0fELz>

⁵ <https://cutt.ly/gs0gzeP>

				economic area and location from 1990 to 2016.
SocSec	Soc. Sec. Agency	14	216.726	Complementary information about citizen labor periods. The range of dates is from 1984 to 2016
CNAE	INE	5	993	Spanish National Classification of Economic Activities that encode companies economic areas.

Table 1: Datasets Description

3.3. Data warehouse design

In order to be able to query and process all the datasets available, we needed to design a data warehouse so that all this information could be integrated into the same storage. To properly organize the data we used the typical *Snowflake Schema*[26] logical design, since the structure of our data perfectly fits with this type of design. Note that most of the datasets contained the personal data of the citizen as key identifier. Thus, the design contains a central table (person) with this information that makes as pivot table to connect to the rest of them. The data warehouse designed is illustrated in Figure 1.

Figure 1: Data Warehouse Schema.

Observe that *person* table plays the role of *fact table* in our design, whilst *graduate* and *job* act as *dimension tables*. Additionally, based on the normalization performed, *locality*, *trainingStudy* and *trainingFamily* tables are added as dimension tables of *graduate* and *job* ones. *Locality* stores the location of the family and education centre, *trainingStudy* stores the list of different VET programs included in our analysis and *trainingFamily* stores the list of VET professional families. These additional dimension tables are the reason why we used a snowflake design instead of a star one.

These tables were built based on the datasets explained in the previous section. Obviously, as it may be observed in Figure 1, there are some datasets that are not reflected in the design. Concretely, *SocSec* dataset was used to complete the information included in *job* table whilst *CNAE* dataset, which contains the relationships among VET professional families and job economical sectors, was used to derive the relation between *job* and *trainingFamily* tables.

4. Methodology

4.1. TOPSIS

The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) is a non-parametric multi-criteria ranking method which is model-free and data-driven. Following the notation of [27], let us assume we have m alternatives A_1, \dots, A_m and n criteria C_1, \dots, C_n , with weights $\omega_1, \dots, \omega_n$. Each criterion may be either a benefit (i.e. “the more the better”) or a cost (i.e. “the less the better”). Let J^+ and J^- denote the sets of indices j corresponding to the benefit and cost criteria, respectively, and let X denote the $m \times n$ matrix of performance ratings of each alternative at each criterion.

The TOPSIS algorithm encompasses the following steps:

1. Normalize and weight the columns of X , obtaining the normalized matrix N whose element n_{ij} is defined by

$$n_{ij} = \omega_j \frac{x_{ij}}{\sqrt{\sum_{k=1}^m x_{kj}^2}}, \quad i = 1, \dots, m; j = 1, \dots, n.$$

2. Find the ideal, A^b , and antiideal, A^w solutions:

$$A_j^b = \begin{cases} \max_{i=1, \dots, m} n_{ij} & \text{if } j \in J^+ \\ \min_{i=1, \dots, m} n_{ij} & \text{if } j \in J^- \end{cases} \quad A_j^w = \begin{cases} \min_{i=1, \dots, m} n_{ij} & \text{if } j \in J^+ \\ \max_{i=1, \dots, m} n_{ij} & \text{if } j \in J^- \end{cases}$$

3. Find the distances between each alternative and the ideal and antiideal solutions:

$$d_i^b = \|A_i - A^b\| = \sqrt{\sum_{j=1}^n (n_{ij} - A_j^b)^2}; \quad d_i^w = \|A_i - A^w\| = \sqrt{\sum_{j=1}^n (n_{ij} - A_j^w)^2}$$

4. Compute the score ranking for each alternative $r_i = \frac{d_i^w}{d_i^w + d_i^b}$.
5. Sort the alternatives according to the scores r_i .

This method is very easily implemented and applied, however, as usual, devil is in the details and in this case, the weight determination is a major concern since a slight change in them can lead to different final rankings.

4.2. Criteria definition and weight determination

In order to obtain the different rankings presented in this work, first the criteria that would allow us to compare the degree of employability of the different VET programs must be determined. A Delphi methodology was used for defining such criteria and their corresponding weights [28]. In particular, 10 experts were asked in rounds to define the main criteria that would take into account both, the employability when the VET program is related with the job economic sector, and the employability in general, i.e., regardless the job economic sector and the family of the VET program studied. After reaching a consensus, the criteria defined were:

- C₁** Time elapsed from the graduation to the signing of the first employment contract within the same professional field as the VET program.
- C₂** Average time elapsed between the end of a contract and the signing of the next employment contract within the same professional field as the VET program.
- C₃** Fraction of the labor life worked within the same professional field as the VET program.
- C₄** Fraction of the time worked within the same professional field as the VET program under temporary contracts.
- C₅** Time elapsed from the graduation to the signing of the first employment contract regardless the professional field.
- C₆** Average time elapsed between the end of a contract and the signing of the next employment contract regardless the professional field.
- C₇** Fraction of the labor life under temporary contracts.
- C₈** Time spent without contracts.

Criteria **C₁ – C₄** analyse the employability of a VET program in its corresponding professional field. On the other hand, criteria **C₅ – C₇** analyse the employability of a VET program in general, regardless of its professional family. The reason for including the latter is because it was found that for many degrees, most of the graduates did not get to work within the scope of their professional family, but still they managed reasonably well, to found a job in other fields and somehow this indicated that for some professions, the fact of just finishing a VET Program was enough. Finally, criterion **C₈** is meant to penalize long term unemployment.

Regarding the criteria weights, the experts were asked to set the minimum weight to 1 and then to quantify the rest of the weights in relative importance to that of the minimum. Afterwards some discussion was conducted among the experts until a consensus was reached. The values

obtained are given in Table 2. Those weights clearly reflect the importance given to the employability of the different VET programs within their professional family.

	C ₁	C ₂	C ₃	C ₄	C ₅	C ₆	C ₇	C ₈
W_i	4	2.5	1	1	3	2	1	1
ω_i	0.258	0.161	0.065	0.065	0.193	0.128	.065	0.065

Table 2: Relative and absolute weights of the criteria determined by the Delphi methodology.

6. Criteria influence determination

As was mentioned earlier, one of the most delicate parts in the design of a ranking score is the weighting of each of the components that define it. The use of expert opinion is common when determining those weights, however, the heterocedasticity and the correlation between the different criteria make these assigned weights rarely coincide with their real influence in the final score ranking [29].

In order to assess this influence, we are going to use two different approaches. The first one is the *sensitivity analysis* [29] which is based in the Pearson's correlation ratio, while the second one, which we introduce in this work, is based on the assessment of the performance under different scenarios.

5.1. Global sensitivity analysis based on the use of variance

The sensitivity analysis technique used here, is applied to assess the quality of the ranking score by finding the “*effective weights*” of the criteria (main effects), providing thus a measure of the real importance of each criterion.

The sensitivity analysis method determines the actual contribution of each criterion to the overall ranking by means of the Karl Pearson's correlation ratio, η^2 which, for each criterion C_j , is defined as [30]

$$\eta_j^2 = \frac{\text{Var}_{X_j}(E_{X_{-j}}(r | X_j))}{\text{Var}(r)}, \quad (1)$$

where X_j denotes the j -th column of the performance matrix X (i.e. the vector of the

performance of all m alternatives in the j -th criterion), X_{-j} denotes the vector containing all except the j -th input of vector X_j , and r denotes the vector of the final scores of all m alternatives. Parameter η_j^2 is known as the *main effect* of criterion C_j on the output vector r , in terms of the expected output variance explained by X_j [31]. The greatest difficulty in calculating these η_j^2 lies in obtaining the numerator in (1) which, in turn, depends on the conditional expectation of r , given X_j . To overcome this problem we use the “*dependent regression*” method [32], which mainly consists on a non-parametric smoothing algorithm based on a Kalman filter.

5.2. Comparing scenarios

The different scenario comparison we introduce here is inspired in the Kao-Liu fuzzy DEA method of best-worst case scenarios [33]. For a given criterion C_j we will consider two possible scenarios:

- Most weighted case: in this case we will assume that C_j has the highest relative weight among all criteria. As a reference measure we assign the relative weights as $W_j^+ = 2$, $W_i^+ = 1$ for $i \neq j$, i.e. in the most weighted scenario, criterion C_j is assigned twice as much importance as the rest of weights. Therefore, in our case, the absolute weights in this case will be $\omega_j^+ = 0.22$, $\omega_i^+ = 0.11$ for $i \neq j$.
- Least weighted case: in this case, criterion C_j has half the weight of the rest of criteria, i.e., $W_j^- = 1$, $W_i^- = 2$ for $i \neq j$. Hence, the weights used for this case in our study will be $\omega_j^- = 0.066$, $\omega_i^- = 0.133$ for $i \neq j$.

Thus, for each criterion we obtain two rankings R_j^+ and R_j^- which are represented by two permutations of the numbers $1, \dots, m$. The identity permutation is assigned to the first ranking, and the permutation for the second ranking is obtained using the first as a reference.

Finally, the relative Kendall - tau distance between both permutations, D_j is computed. This distance measures the fraction of all possible pairings of alternatives which change the order

from one ranking to the other.

If this value is small, it can be interpreted as a lack of influence of criterion j in the preference ordering of the alternatives, since no matter how big or small the weight of the criterion is, the final ranking is not varying that much.

6. Results and discussion

6.1. Data preparation

The original data set consisted on a total of 28.272 people graduating in 121 different VET programs from 1996 to 2016. Those VET programs were classified according to the professional family they are included in. For each VET program and for each graduating year the scores of each person verifying the conditions of the different criteria were obtained. Then the median of such scores was calculated and considered as the score of that program for that year in the corresponding criterion.

Figure 2: Number of total VET programs available each year (dashed blue line) and number of VET programs with all criteria computed as medians of more than 5 data (solid red line). The horizontal dashed line shows the minimum number of programs considered for the ranking.

Nevertheless, some filtering had to be carried out previously. First, not all programs had scores in all criteria all years. Indeed, for the first years in the series, there is a very small number of VET programs scoring in all criteria. Thus we only considered the years for which we had a reasonable amount of programs with data in all criteria. We set as threshold of 30 VET programs as the minimum for making the ranking.

Secondly, as previously stated, the scores in the different criteria are obtained as a median of the scores of individual persons. However, in some cases, such median was computed over very few people which could lead to biased results. Therefore we only considered VET programs for which the scores were obtained from more than 5 data.

Figure 2 depicts both, the number of VET programs available for each year (from 1996 to 2016) and the number of VET programs for which all the criteria were obtained from data from more than 5 people. According to the number of programs in each year, we considered the time

window from 2009 to 2016.

After the filtering, from the original 121 VET programs, a total amount of 106 were used in this study⁶.

As a brief descriptive summary of the data, Figure 3 shows the evolution of the performance scores of each criterion throughout the period 2009 - 2016.

Figure 3: Evolution of the general performance at each criterion in the period 2009-2016.

The general behaviour of the criteria reflects the evolution of the economy in Spain, particularly in Extremadura. The time required to find the first job (criteria C1 and C5) starts a clearly decreasing trend in 2012, together with the time between contracts (criteria C2 and C6). For the rest of the criteria there is no general improvement (in median). This turning point in 2012 coincides with the moment when the region's GDP growth bottomed out and began to recover, although it was not until 2014 when it began to have positive values, like the rest of the country.

6.2. VET rankings

Figure 4 shows the results for the rankings of the VET programs from 2009 to 2016. The figure shows, with a colour, the percentile of each VET program at each year. The VET programs have been sorted according to their average percentile during the whole period 2009-2016. This plot allows to easily find both, the overall performance of a program and the evolution over time of such performance.

Figure 4: VET programs ranking throughout the period 2009-2016. The colour reflects the percentile of each program at each year (the greener the better) and the overall sorting has been done according to the average percentile of each program during the whole time span.

A closer look at the top and bottom of the ranking, as shown in Table 3, gives us some insight about which VET programs are the best behaved regarding the employability. Seven out of

⁶ The list of all the VET names and Professional families is available in <https://github.com/chemacmunex/vetrankings.git>

the 10 best programs are from the professional families “*Agricultural and Livestock Activities*” (ALA) and “*Hostel and Tourism*” (HOT), which correspond with the two major economic activities (or sectors) in the region: Primary and Tertiary sectors, namely agriculture and tourism.

Pos.	ID	VET Program	Mean P.	Pr. F.
1	D102	Prod. Programming in Mech. Manufa.	0.968	MEM
2	D113	Catering Services	0.944	HOT
3	D9	Physical and Sport Activities Org.	0.898	PSA
4	D59	Oral hygiene	0.894	HEA
5	D99	Agroecological Production	0.888	ALA
6	D30	Cookery Management	0.885	HOT
7	D19	Cooking and gastronomy	0.855	HOT
8	D100	Agricultural and Livestock production	0.853	ALA
9	D56	Forest and Natural Envmnt. Mgmt.	0.852	ALA
10	D57	Mgmt. & Org. of Agr. & Livestock Co.	0.826	ALA
⋮	⋮	⋮	⋮	⋮
97	D96	Prepress in Graphic Arts	0.096	GRA
98	D112	Commercial services	0.089	COM
99	D116	Microcomputer systems and networks	0.081	ITC
100	D81	Vehicle maintenance	0.078	VTM
101	D95	Hairdressing and aesthetics	0.067	PIM
102	D40	Electricity and electronics	0.056	ELE
103	D65	IT and communications	0.044	ITC
104	D111	Administrative services	0.033	ADM
105	D88	Tesellations	0.012	ART
106	D64	Food industries	0.011	FOI

Table 3: Head and tail of the ranking. Position, mean percentile throughout the period 2009-2016 and the corresponding professional family is tabulated for the first and last 10 VET programs.

It is also interesting to analyse how the programs within each professional family are doing regarding the employability of their students. Figure 5 depicts, for each professional family, the evolution over time of their programs in the ranking. We show the mean of the percentiles

obtained by the VET programs of the corresponding professional family (solid lines) together with the minimum and maximum values (dashed lines) and the amount of programs that each year fulfilled all the conditions mentioned above, for being considered in the ranking (gray bars - right side y-axis).

It is clearly seen how Agriculture and Livestock Activities (ALA) and Hostel and Tourism (HOT) perform very well through most of the period with very few differences between the maximum and the minimum. Other professional families, on the contrary, perform poorly for most of the time, like Graphic Arts (GRA), Chemistry (CHE) and Arts and Artcraft (ART).

Finally, some other professional families, like, for example, Administration and Management (ADA), Health (HEA), IT and communications (ITC) or Personal Image (PIM), present huge differences between the maximum and the minimum. In some cases those differences can be explained because the programs within the same professional families are very different in nature: for example, in the Health professional family, Dentures (D103) and Oral Hygiene (D59) programs are ranked in the top of the list while Anatomical Pathology and Cytodiagnosis (D8), Radiotherapy (D107) or Clinical diagnostic Laboratory (D78) perform quite poorly in comparison. These professions differ both in the workplaces where they take place and in the activity itself.

However, in other cases, such as the Personal Image professional family, all programs, with the exception of Personal and Corporate Image Consulting VET program, are related to hairdressing and aesthetics, but their positions in the ranking range from as low as a 6.6% for Hairdressing and Aesthetics (E95), to an average percentile of 77% for Aesthetics (D45) programs. This high variability may be due to an over fragmentation of the professional family in too many programs with very similar characteristics.

Figure 5: Mean percentile of the VET programs in each professional families (solid line), minimum and maximum percentiles (dashed lines) and number of VET programs considered each year (gray bars).

6.3. *Criteria influence analysis*

Figure 6 depicts a boxplot of the Kendall - tau distance between the most weighted and the least weighted scenarios for each criterion for the whole 2009-2016 period, showing firstly that criteria C7 (Fraction of the labor life under temporary contracts) and C8 (Time spent without contracts) are

the least sensitive to weight changes with a median Kendall - tau distance of around a 5% and secondly, that criterion C4 (Fraction of the time worked within the same professional field as the VET program under temporary contracts) is the most sensitive to those weights changes (with a Kendall - tau distance reaching values larger than 35%).

Figure 6: Kendall - tau distance between the least and most weighted scenarios for each criterion.

On the other hand, we can compare these results with those obtained from the Global Sensitivity Analysis. Figure 7 depicts the Karl Pearson's correlation ratio for three different rankings: the least weighted scenario (left plot), the most weighted scenario (middle plot) and the weights proposed by the expert panel (right plot). Taking a closer look to the three plots we can see that the importance of criteria C7 and C8 in the construction of the final score is almost the same and it is very low. The importance of criteria C1 and C2 show more differences across the three rankings, and finally, criteria C3, C4, C5, and C6, show the largest differences in importance between the three rankings, showing a larger importance when the weights assigned are larger and lower values of importance when their assigned weights are smaller. This agrees with the results of the scenario comparison method that showed that these criteria were very sensitive to changes in the assigned weights.

However, the scenario comparison method introduced in this work, focuses not on the scores obtained by the MCDM method (TOPSIS score in this case), like the Global Sensitivity index does, but on the final ranking obtained. This is important since a change in the score of an alternative does not necessarily lead to a change in its position in the ranking. Also it is worth mentioning that the scenario comparison method introduced here has the advantage of not needing a previous definition of the criteria weights, and therefore the information about the sensitivity of the final ranking to the weight assigned to each criterion is available to the expert panel before they make the weight assignment.

Moreover, if we recall the weights assigned by the expert panel, it may look that there is a divergence between the weights assigned to each criterion and their sensitivity, as was found in the previous results. This divergence does not mean that the weights were wrongly assigned by the experts. In this case the experts (regional administration) gave more importance to criteria C1 and C2, mainly because those criteria were directly related to finding a first job rapidly and keep working on a labour sector directly related to the study field of the VET, and this could encourage

young people to enrol on such VET programs. On the other hand criterion C4 was regarded as a long term criterion which was considered to be less important for young people when deciding whether to take a VET course or not.

Notwithstanding the above, this tool does give information about the changes that the final ranking may have if, for political reasons, it is decided to focus the importance of the criteria in a slightly different way. In particular, the experts should know that a slight change is made in the weight of criterion C4, the final ranking will be far more affected than if this change is made to criteria C1 or C2 (for example). With this information, experts could establish the allocation of weights in a safer way and with information consistent with the allocation.

Therefore, to some extent, the proposed criteria analysis is independent from the weight definition, but decision-makers should be aware of the sensitivity of the different criteria to changes in the weights, since those changes can be determined by changes in some policies.

Figure 7: Global sensitivity analysis for the three cases considered: the least weighted scenario (left), the most weighted scenario (middle) and the ranking obtained with the original weights given by the expert panel (right).

7. Conclusions

In this work we have proven the effectiveness of the TOPSIS method as a multi-criteria analysis tool for the classification of VET programs in Extremadura during the period 2009-2016. It is very important to emphasize the fact that having quantitative data available relating different databases from the Education and Employment Board of Extremadura, allows us to have more criteria on which to base the performance of the different VET programs. It should be borne in mind that in this type of rankings, the criteria employed are usually either qualitative and based on surveys about the satisfaction on their employment status several years after completing their studies or, in the case of quantitative criteria, they are, in most cases, based on a single criterion (e.g. percentage of graduates who are working a year after completing their studies).

In the analysis, we have provided a classification considering 8 different criteria, that were calculated based on the real data provided by the Education and Employment Board of Extremadura. This is a significant contribution of this work with respect to similar ones that apply

multi-criteria analysis relying on the use of questionnaires.

In addition, in this work we have proposed a new decision support method for assessing the influence of the weights assigned to the different criteria in the final ranking, comparing it with other known technique for criteria influence analysis. This new method can be applied before the weight definition by the experts, so it can serve as support information for the determination of the final weight scheme. This analysis showed that it is of paramount importance a thorough analysis of the criteria, since the ranking of the different VET programs, and all the possible policy decisions based on such ranking, greatly depend on the choice of the weight assigned to each criterion.

Finally, we strongly believe that the approach presented in this work, in which the academic data of each VET student is linked to its full labor history, is a step forward in the analysis of the efficiency and usefulness of such VET programs. Moreover, since VET programs are widely consolidated all over Europe, this approach could be easily replicated in other countries and regions, using their own citizens' data, to try to reach the same goal.

Acknowledgements

This work has been developed with the support of (i) Ministerio de Ciencia e Innovación (MCI), Agencia Estatal de Investigación (AEI) and European Regional Development Fund (ERDF): RTI2018-098652-B-I00 and RTI2018-093608-B-C33 projects, and (ii) European Regional Development Fund (ERDF) and Junta de Extremadura: IB16055, IB18034 and GR18112 projects.

References

References

- [1] J. Guardiola, M. Guillen-Royo, Income, unemployment, higher education and wellbeing in times of economic crisis: Evidence from granada (spain), *Social Indicators Research* 120 (2) (2015) 395–409.
- [2] F. Martí, J. J. Pérez, Spanish public finances through the financial crisis, *Fiscal Studies* 36 (4) (2015) 527–554. doi:10.1111/j.1475-5890.2015.12079.
- [3] M. Casares, J. Vázquez, Why are labor markets in spain and germany so different?,

- Economic Modelling 75 (2018) 320 – 335.
doi:10.1016/j.econmod.2018.07.008.
- [4] E. U. Advisory Committee on Vocational Training, Opinion on the Future of Vocational Education and Training Post 2020, Tech. rep. (2018).
URL <https://cutt.ly/5s0lLNT>
- [5] S. Flisi, V. Goglio, E. C. Meroni, M. Rodrigues, E. Vera-Toscano, Measuring occupational mismatch: Overeducation and overskill in europe—evidence from piaac, Social Indicators Research 131 (3) (2017) 1211–1249.
- [6] C.-L. Hwang, K. Yoon, Methods for Multiple Attribute Decision Making, in: Multiple Attribute Decision Making. Lecture Notes in Economics and Mathematical Systems, vol 186, Springer Berlin Heidelberg, 1981, pp. 58–191.
doi:10.1007/978-3-642-48318-9_3.
- [7] A. Rad, B. Naderi, M. Soltani, Clustering and ranking university majors using data mining and AHP algorithms: A case study in Iran, Expert Systems with Applications 38 (1) (2011) 755–763. doi:10.1016/J.ESWA.2010.07.029.
- [8] J. Allen, R. Van Der Velden, The Flexible Professional in the Knowledge Society: First Results of the REFLEX projects (Jan 2006).
- [9] B. Williamson, Digital education governance: data visualization, predictive analytics, and ‘real-time’ policy instruments, Journal of Education Policy 31 (2) (2016) 123–141.
doi:10.1080/02680939.2015.1035758.
- [10] B. Sen, E. Uçar, D. Delen, Predicting and analyzing secondary education placement-test scores: A data mining approach, Expert Systems with Applications 39 (10) (2012) 9468 – 9476. doi:10.1016/j.eswa.2012.02.112.
- [11] N. B. Sara, R. Halland, C. Igel, S. Alstrup, High-school dropout prediction using machine learning: A Danish large-scale study, in: 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2015 - Proceedings, 2015.
- [12] E. Aguiar, H. Lakkaraju, N. Bhanpuri, D. Miller, B. Yuhas, K. L. Addison, Who, when and why: A machine learning approach to prioritizing students at risk of not graduating high school on time, in: ACM International Conference Proceeding Series, 2015.
doi:10.1145/2723576.2723619.

- [13] R. Campagni, D. Merlini, R. Sprugnoli, M. C. Verri, Data mining models for student careers, *Expert Systems with Applications* 42 (13) (2015) 5508–5521.
doi:10.1016/j.eswa.2015.02.052.
- [14] L. D. Pietro, R. G. Mugion, F. Musella, M. F. Renzi, P. Vicard, Reconciling internal and external performance in a holistic approach: A bayesian network model in higher education, *Expert Systems with Applications* 42 (5) (2015) 2691 – 2702.
doi:10.1016/j.eswa.2014.11.019.
- [15] R. Asif, A. Merceron, S. A. Ali, N. G. Haider, Analyzing undergraduate students' performance using educational data mining, *Computers & Education* 113 (2017) 177 – 194. doi:https://doi.org/10.1016/j.compedu.2017.05.007.
- [16] A. J. Fernández-García, L. Iribarne, A. Corral, J. Criado, A comparison of feature selection methods to optimize predictive models based on decision forest algorithms for academic data analysis, in: *Advances in Intelligent Systems and Computing*, Vol. 745, Springer Verlag, 2018, pp. 338–347. doi:10.1007/978-3-319-77703-0_35.
- [17] D. Jackson, Factors influencing job attainment in recent Bachelor graduates: evidence from Australia, *Higher Education* 68 (1) (2014) 135–153.
doi:10.1007/s10734-013-9696-7.
- [18] I. Mewburn, W. J. Grant, H. Suurinen, S. Kizimchuk, A Machine Learning Analysis of the Non-academic Employment Opportunities for Ph.D. Graduates in Australia, *Higher Education Policy*. (2108), doi:10.1057/s41307-018-0098-4.
- [19] Y. Bharambe, N. Mond, M. Mulchandani, R. Shankarmani, S. G. Shinde, Assessing employability of students using data mining techniques, in: *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, 2017, pp. 2110–2114. doi:10.1109/ICACCI.2017.8126157.
- [20] P. Thakar, A. Mehta, Manisha, A unified model of clustering and classification to improve students' employability prediction, *International Journal of Intelligent Systems and Applications* (9) (2017) 10–18. doi:10.5815/ijisa.2017.09.02.
- [21] F. García-Peñalvo, J. Cruz-Benito, M. Martín-González, A. Vázquez-Ingelmo, J. C. Sánchez-Prieto, R. Therón, Proposing a Machine Learning Approach to Analyze and Predict Employment and its Factors, *International Journal of Interactive Multimedia and Artificial Intelligence* (2) (2018) 39–45. doi:10.9781/ijimai.2018.02.002.

- [22] Z. Yue, An extended topsis for determining weights of decision makers with interval numbers, *Knowledge-Based Systems* 24 (1) (2011) 146 – 153.
doi:10.1016/j.knosys.2010.07.014.
- [23] Z. Yue, Extension of topsis to determine weight of decision maker for group decision making problems with uncertain information, *Expert Systems with Applications* 39 (7) (2012) 6343 – 6350. doi:10.1016/j.eswa.2011.12.016.
- [24] D. Olson, Comparison of weights in topsis models, *Mathematical and Computer Modelling* 40 (7) (2004) 721 – 727. doi:10.1016/j.mcm.2004.10.003.
- [25] H. Dincer, U. Hacıoglu, E. Tatoglu, D. Delen, A fuzzy-hybrid analytic model to assess investors' perceptions for industry selection, *Decision Support Systems* 86 (2016) 24 – 34.
doi:10.1016/j.dss.2016.03.005.
- [26] P. Potineni, *Database Data Warehousing Guide*, Oracle, 2018.
- [27] G.-H. Tzeng, J.-J. Huang, *Multiple attribute decision making: methods and applications*, CRC press, 2011.
- [28] N. Dalkey, O. Helmer, An experimental application of the delphi method to the use of experts, *Management Science* 9 (3) (1963) 458–467. doi:10.1287/mnsc.9.3.458.
- [29] P. Paruolo, M. Saisana, A. Saltelli, Ratings and rankings: voodoo or science?, *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 176 (3) (2013) 609–634.
doi:10.1111/j.1467-985X.2012.01059.x.
- [30] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index, *Computer Physics Communications* 181 (2) (2010) 259 – 270.
doi:10.1016/j.cpc.2009.09.018.
- [31] M. C. Bas, S. Tarantola, J. M. Carot, A. Conchado, Sensitivity analysis: A necessary ingredient for measuring the quality of a teaching activity index, *Social Indicators Research* 131 (3) (2017) 931–946. doi:10.1007/s11205-016-1297-2.
- [32] M. Ratto, A. Pagano, P. Young, State dependent parameter metamodelling and sensitivity analysis, *Computer Physics Communications* 177 (11) (2007) 863 – 876.
doi:10.1016/j.cpc.2007.07.011.
- [33] C. Kao, S. T. Liu, A mathematical programming approach to fuzzy efficiency ranking, *International Journal of Production Economics* 86 (2) (2003) 145 – 154.

doi:10.1016/S0925-5273(03)00026-4.

Journal Pre-proof