A Structural Model of the Employment Pathways of the University of Foggia Graduates[1]

Corrado Crocetta*, Francesco d'Ovidio**

* *Department of Economic, Mathematic and Statistic Sciences, University of Foggia, Italy*
** *Department of Statistic Sciences, University of Bari, Italy*

Summary. The purpose of this study is to investigate the strategies used by graduates of the University of Foggia to enter the labour market. Using both quantitative and qualitative variables, quantified by means of optimal scaling, a structural equation model has been created to analyse the relations between latent variables tied to university education, and graduates' expectations and methods of job searching. Furthermore, we study if the correlation structure between these latent variables is constant observing separately female and male graduates.
Keywords: Graduates; Labour market; University of Foggia; Factor analysis; Structural equation models; Optimal Scaling; CATPCA; LISREL.

1. Introduction

In the following, we intend to analyse the strategies used by graduates to enter the labour market. Because of difficulties in the availability of the data, our analysis considers only the labour offer side.

Our objective is to construct an econometric model capable of detecting the variables that influence the placement of graduates and to study the relations between these variables and the *labour situation* (employed/unemployed) of graduates.

A great deal of the information collected with surveys can be considered *proxy* of latent variables which are particularly useful for describing a fact not

---

[1] In this joint work, C. Crocetta was responsible for the final editing of Sections 1, 2, 4, 6 and 7, whereas F. d'Ovidio was responsible for Sections 3 and 5. The authors wish to thank anonymous referees for their precious suggestions.

directly measurable, like the behaviour of graduates looking for employment (Crocetta & d'Ovidio, 2004).

If we have to analyse both quantitative and qualitative variables together, we can apply the Optimal Scaling procedure known as CATPCA (CATegorical Principal Component Analysis). Such a procedure may be used to quantify, in the factorial space, the observed variables (De Leeuw, 1984; Meulman & Heiser, 1999).

For this, we performed a series of analysis to verify:

1. which latent variables are to be considered,
2. the relations that exist between them,
3. whether it is possible to construct a model for explaining the variability of employment rates of graduates,
4. whether there are other variables, not considered in the model, which can explain the employment,
5. whether, by considering separately female and male graduates, the structure of the correlation between the latent variables changes significantly.

2. A survey on graduates of the University of Foggia

The data for this analysis came from the archives of the Student office of the University of Foggia and was integrated with telephone interviews of a sample of graduates[2]. The questionnaire is divided into four parts, which are dedicated to graduates who are already working, those who are looking for employment, and those with a work experience in progress, and, finally, to the satisfaction regarding the services and the preparation achieved.

We have found that, as the time upon graduation passes, the number of fixed term contract grows (Table 1). This situation regards all those who graduated in Agriculture, 92.5% of graduates in Economics and 83.3% of graduates in Law. However, by considering the graduates in the period 1997-1999, doctors in Medicine have the highest rate of stable position (83.3%), followed by their colleagues in Economics (75.6%) and Law (74.7%), while just 67.7% of graduates in Agriculture during the same period have a stable job.

The attainment of a stable job is not so difficult even for those who graduated less than three years before (59.3% of them are in this condition). In this case, too, graduates in Medicine (70.0%) have a small advantage over those in Economics (61.2%), in Agriculture (58.8%) and in Law (56.1%).

[2] The first objective was to contact all the 2,924 graduates of the Athenaeum of Foggia starting from when it was set up (1994). Up to seven attempts were made to contact them at different times, before they were considered unreachable. A number of 2,133 interviews was performed, which is the 72.3% of the eligible population.

Table 1. Percentage distribution of the employed graduates from the University of Foggia, according to the year of graduation, faculty, and work contract

Contract	Faculty				Total
	Economics	Agriculture	Law	Medicine	
	Years 1994-96				
Stable employment (full or part-time)	92.5	100.0	83.3	-	88.6
Temporary job or job training scheme	5.0	-	9.5	-	6.8
Occasional employment and the like	2.5	-	7.1	-	4.5
	Years 1997-99				
Stable employment (full or part-time)	75.6	67.7	74.7	83.3	74.7
Temporary job or job training scheme	11.9	29.0	12.4	16.7	13.5
Occasional employment and the like	12.5	3.2	12.9	-	11.9
	Years 2000-02s				
Stable employment (full or part-time)	61.2	58.8	56.1	70.0	59.3
Temporary job or job training scheme	27.2	41.2	30.2	30.0	28.9
Occasional employment and the like	11.6	-	13.8	-	11.8
	All employed graduates				
Stable employment (full or part-time)	69.2	68.5	68.0	75.0	68.7
Temporary job or job training scheme	19.7	29.6	19.3	25.0	20.1
Occasional employment and the like	11.1	1.9	12.7	-	11.2

The proportion of employed graduates with a fixed term contract or on a job training scheme tends to decrease as time passes. A period of precariousness seems, however, inevitable, especially for graduates in Agriculture (29.6%) and Medicine (25.0%). Graduates in these two faculties, unlike their colleagues of Economics and Law, tend not to accept seasonal and temporary work. Presumably, the number of graduates of the two former faculties is not so large, and this may prevent the competition that often makes graduates accept temporary or unsuitable jobs.

3. Analysis of the categorical components of the model

We carried out a descriptive analysis for screening the variables for the model. After that, we estimated a logit model, whose criterion variable is the binary position, employed *vs.* unemployed, of graduates. The variables whose coefficients were significant at 5% level are listed in Table 2.

Some variables are quantitative (such as the 0-100 score of the suitability of the university training received) or on an ordinal scale, but many others are nominal.

Table 2. Variables selected for the logit model

Quantitative variables	Nominal variables
▪ University graduation final grade.	▪ Faculty.
▪ Secondary school graduation final degree.	▪ Type of Secondary School degree.
▪ Age at university graduation.	▪ Pre-graduation work experience.
▪ Number of years between second. school graduation and university enrolment.	▪ Field of economic activity currently employed in or searching for.
▪ Number of years after the end of course before graduating. (*off-programme*)	▪ Current or desired professional position
▪ Number of months between university graduation and first employment.	▪ Professional or teaching qualification
▪ Overall score assigned to the adequacy of university education with respect to employment obtained or desired.	▪ Employment search methods. ▪ Knowledge of post-graduation prospects (orderable nominal variable).

We wanted to apply a LISREL model. Because the structural equation model, given the hypothesised normality of the latent variables, does not allow using categorical variables[3], it was necessary to quantify such variables with an Optimal Scaling (OS) method.

Given a population of n individuals described by a set of m categorical variables $\mathbf{x}_1 ... \mathbf{x}_j .... \mathbf{x}_m$, the OS procedure transforms the categories into real values $\boldsymbol{\omega}_j$. OS methods minimize a loss function regarding the categories of interest.

First, a scalar g_{ijh} has been defined with value 1 or 0 according to whether the i^{th} individual possesses the h^{th} category of the $\mathbf{x}_j$ variable. The vector $\mathbf{g}_{jh}$ is given by this scalar attached to the units in category h of x_j.

With all categories of $\mathbf{x}_j$, the column vectors $\mathbf{g}_{jh}$ originate the indicator-matrix $\mathbf{G}_j$ (of dimensions $n \times k_j$). Extending this procedure to all the m categorical variables, we obtain the disjunctive complete indicator-matrix, $\mathbf{G} = [\mathbf{G}_1 ... \mathbf{G}_j ... \mathbf{G}_m]$, of the order $n \times K$, where $K = \sum_j k_j$.

In this way, each categorical variable is a product of an indicator-matrix by a vector $\boldsymbol{\omega}_j = [\omega_{j1} \ldots \omega_{jh} ... \omega_{jk_j}]'$ of scaling parameters that, once estimated ($\hat{\omega}_{jh}$), originate quantitative variables:

$$^{os}\mathbf{x}_j = \mathbf{G}_j \hat{\boldsymbol{\omega}}_j \qquad \text{or, equivalently,} \qquad {}^{os}\mathbf{x}_j = \sum_{h=1}^{k_j} \mathbf{g}_{jh} \hat{\omega}_{jh} \qquad (j=1, 2, ..., p),$$

where the superscript "OS" indicates the optimally scaled variable. Extending this procedure to all the units of the population and all variables, we obtain the matrix of optimally scaled individual scores, ${}^{os}\mathbf{X} = ({}^{os}\mathbf{X}_1, {}^{os}\mathbf{X}_2, \ldots, {}^{os}\mathbf{X}_m)$.

---

[3] The estimates of the LISREL model parameters with the method of the maximum likelihood are asymptotically biased, because of the violation of the normality hypothesis of the latent variables and, implicitly, of the observed variables (Browne, 1984). This problem may be overcome by using non-parametric loss functions, such as WLS, GLS and ULS (see Lovaglio, 2000).

The vectors $\boldsymbol{\omega}_j$ may be estimated by optimizing a target function with identification constraints. It is worthwhile estimating simultaneously the quantities of the categorical variables, and the parameters of the model[4], by directly optimizing the target function with ALSOS (Alternative Least Squares Optimal Scaling) methods (De Leeuw *et al.*, 1976; Young *et al.*, 1978; Vittadini, 1999).

Among the available ALSOS procedures, we chose CATPCA (CATegorical Principal Component Analyses), a non-parametric algorithm using the main components of the transformed variables in a factorial *p*-dimensional space ($p \leq m$) (De Leeuw & Meulman, 1986; Meulman & Heiser, 1999).

In neither the simple case of no weighting for cases or variables and no supplementary nor multiple variables, the optimisation procedure estimates the ω_j scaling parameters (iteratively) by minimising the function

$$\sigma(\mathbf{Y};\boldsymbol{\Omega}) = n^{-1}\sum_{j}^{m} \mathrm{tr}\left[(\mathbf{Y}-\mathbf{G}_j\boldsymbol{\Omega}_j)'\,\mathbf{M}_j\,(\mathbf{Y}-\mathbf{G}_j\boldsymbol{\Omega}_j)\right],$$

where matrix $\mathbf{M}_j$ is diagonal (of $n\times n$ order) with elements 0 if the i^{th} observation is missing and 1 in the other cases; whereas $\mathbf{Y}$ (of $n\times p$ order) represents the *p*-dimensional standardised factor scores, with the following constraints of standardisation or centring (given $\mathbf{u}$, unit-vector of order n, and $\mathbf{M}=\sum_j \mathbf{M}_j$):

$$\mathbf{Y}'\mathbf{M}\,\mathbf{Y} = n\,\mathbf{I}_p, \qquad \mathbf{u}'\mathbf{M}\,\mathbf{Y} = \mathbf{0}. \qquad [1]$$

The algorithm begins with an estimate of $\mathbf{Y}$ which satisfies those constraints (unless otherwise specified, standardised and centred random numbers); the initial factor loadings $\mathbf{a}_j$ are calculated as cross-product between $\hat{\mathbf{Y}}$ and the categorical codes centred and re-scaled: ${}^{c}\mathbf{x}_j = [\mathbf{I}_n\text{-}\mathbf{M}_j\mathbf{uu}'/(\mathbf{u}'\mathbf{M}_j\mathbf{u})]\,\mathbf{x}_j$, with j= 1, 2, ..., *m* (De Leeuw *et al.*, 1976; Meulman & Heiser, 1999).

The first step of the iteration consists of calculating, given $\mathbf{D}_j = \mathrm{diag}(\mathbf{G}'_j\mathbf{G}_j)$,

$$\hat{\boldsymbol{\Omega}}_j = \mathbf{D}_j^{-1}\mathbf{G}_j'\,\hat{\mathbf{Y}}\,. \qquad [2]$$

After a first quantifications of categories[5] we standardised the data with $\hat{\boldsymbol{\omega}}_j^{\perp} = \hat{\boldsymbol{\omega}}_j\sqrt{n\;/(\hat{\boldsymbol{\omega}}_j'\,\mathbf{D}_j\;\;\hat{\boldsymbol{\omega}}_j)}$ to compute factor weights $\mathbf{a}_j = (\boldsymbol{\Omega}_j'\,\mathbf{D}_j\,\hat{\boldsymbol{\omega}}_j^{\perp})/n$. Through the standardised matrix of the scaling estimates $\hat{\boldsymbol{\Omega}}_j^{\perp} = \hat{\boldsymbol{\omega}}_j^{\perp}\mathbf{a}_j'$, the matrix $\hat{\mathbf{Y}} = [\mathbf{I}_n\text{-}\mathbf{Muu}'/(\mathbf{u}'\mathbf{Mu})]\,(\sum \mathbf{M}_j\mathbf{G}_j\hat{\boldsymbol{\Omega}}_j^{\perp})$ could be now calculated. The process keeps resuming the algorithm with a *singular value decomposition* of $\hat{\mathbf{Y}}$ starting again from [2]. After a certain number of iterations, the final estimates $\hat{\boldsymbol{\omega}}_j$ of the categories are obtained.

---

4 The scaling approach is not separable from the aims of the research, and quantification must be obtained under specific statistical models (Bradley *et al.*, 1962; Kruskal, 1965; de Leeuw *et al.*, 1976).

5 If variables are categorical, the factorial weights $\hat{\boldsymbol{\omega}}_j = \hat{\boldsymbol{\Omega}}_j\,\mathbf{a}_j$ are directly used; if variables are ordinal, the weights $\hat{\boldsymbol{\omega}}_j$ are obtained through a monotone regression of the weighted $\hat{\boldsymbol{\Omega}}_j\,\mathbf{a}_j$ with the diagonal elements of $\mathbf{D}_j$, whereas, if they are numeric, a weighted linear regression is used.

Table 3. Percent distribution of the University of Foggia graduates according to work condition and some other characteristic

	Labour condition			Labour condition	
	Unem-ployed	Employed		Unem-ployed	Employed
University Faculty			*Field of activity employed in/searching for*		
Medicine and Surgery	80.4	19.6	n.a.	85.0	15.0
Law	59.4	40.6	Public Administration	64.9	35.1
Economics	42.2	57.8	Industry	64.1	35.9
Agriculture	39.6	60.4	Commerce	46.8	53.2
Type of secondary school diploma			Agriculture	29.3	70.7
Languages	90.9	9.1	Services	26.0	74.0
Others	71.4	28.6	Other field	27.4	72.6
Classical	60.4	39.6	*Current or desired professional position*		
Scientific	55.2	44.8	n.a.	98.1	1.9
Teacher training	51.6	48.4	Consultant.	61.5	38.5
Technical commercial	50.2	49.8	Entrepreneur	50.0	50.0
Professional	48.7	51.3	Employee/Manager	40.9	59.1
Technical Geometer	33.3	66.7	Self-employed	35.0	65.0
Knowledge of post-graduate prospects			Teacher/Professor	11.3	88.7
Yes, quite well	49.3	50.7	Other Position	14.5	85.5
Yes, in a generic way	59.9	40.1	*Employment search method*		
No	48.4	51.6	n.a.	92.0	8.0
Pre-graduation employment			Newspaper/Internet	86.3	13.7
Never worked before	56.5	43.5	Specialized Agencies	80.9	19.1
Worked before	48.0	52.0	Local employm.agency	79.7	20.3
Professional/teaching qualification			Curriculum sent	39.4	60.6
Not licensed	60.0	40.0	Interviews/exams	36.8	63.2
Licensed	31.3	68.7	Personal Contacts	29.0	71.0
			Direct calls	-	100.0
			Suggestions	9.8	90.2
Total	***53.0***	***47.0***	Other search-methods	13.6	86.4

In Table 3, the categories are shown in decreasing order of employment rate.

CATPCA defines just four components with eigenvalues higher than one (Table 4). Altogether, they explain 52% of the overall variability. For a better definition and identification of the factors, we performed a Varimax rotation of the factorial axes. The first factor accounted for 16.1% of the variability, whereas the forth accounts for 9.6%.

Let us now identify the four latent variables (Table 5). The *first factor* is directly connected to the age at graduation and to time between high school graduation and university enrolment, whereas it presents a negative correlation with the University and high school final grade. We named this factor "*regularity and proficiency of educational career*".

Table 4. Variability explained by the main components and rotated factors*

Components	Weights of unrotated components			Weights of rotated factors		
	Eigenvalues	% of variance	% cumulated variance	Eigenvalues	% of variance	% cumulated variance
1	2.54	18.1	18.1	2.26	16.1	16.1
2	1.93	13.8	31.9	1.86	13.3	29.4
3	1.68	12.0	43.9	1.82	13.0	42.4
4	1.13	8.1	52.0	1.34	9.6	52.0

* 14 components have a non-zero eigenvalue. Extraction Method: Principal Components Analysis. Rotation with Varimax Method.

Table 5. Coefficients obtained from factor analysis of variables optimally scaled by means of CATPCA for the graduates of the University of Foggia

Variables	Commu-nality	Components			
		1	2	3	4
Time between graduation and first employment	0.66	*-0.17*	*0.27*	0.75	*-0.05*
Time between last exam and graduation	0.63	0.63	*0.14*	-0.46	*0.08*
Professional or teaching qualification	0.63	*-0.05*	*-0.16*	0.78	*-0.03*
Age at graduation	0.62	0.74	*0.03*	*-0.03*	*0.26*
Graduation final grade	0.59	-0.76	*0.08*	*0.12*	*0.02*
Current or sought professional position	0.54	*0.00*	0.74	*-0.05*	*-0.03*
Employment search strategies	0.52	*0.15*	0.50	0.48	*-0.15*
High school final grade	0.51	-0.70	*0.08*	*-0.10*	*0.05*
Faculty	0.50	*-0.27*	0.59	*-0.17*	*0.22*
Knowledge of post-graduation prospects	0.46	*0.07*	*0.11*	*0.06*	-0.67
Current or sought field of activity	0.44	*0.03*	0.61	*0.25*	*-0.03*
Pre-graduation work experience	0.42	*0.15*	*0.13*	*-0.07*	0.62
Type of high school diploma	0.38	*0.00*	0.41	*-0.12*	0.44
Time between high school and university enrolment	0.35	*0.30*	*-0.10*	*0.28*	0.41

The *second factor* is directly connected to the field of economic activity, to the professional position achieved, to faculty and type of high school diploma and to job search strategy. Because it is strongly influenced by type of studies and working experience of the graduates, we named it "*training-professional project*".

The *third factor* is correlated to the time spent between graduation and first job, to possession of professional or teaching qualification and to job search strategies, and opposite to time between high school graduation and university enrolment. All these variables have in common the *job finding experiences* of the graduates.

The *fourth factor* is mainly related to the work before graduation, to the type of high school diploma and to the time between high school graduation and university enrolment. This factor has also a negative correlation with the knowledge of the job prospects after graduation[6]. For this reason, we name it "*work plans and experiences*".

4. The structural equation model

LISREL is a structural equation model. It is popular in social sciences to study the cause-effect relations within a system. Generally, structural equation models are constructed with very simple relations.

By applying Path Analysis graphs (Wright, 1934), it is possible to represent the model with flow diagrams in which the surveyed variables are represented with right angles, whereas the latent variables and the erratic components are contained in elliptic shapes. These geometric figures can be connected with arrows that indicate the existence of a relation[7].

There is a distinction between measuring models, which are useful for identifying and measuring the latent variables through the observed variables, and structural models, which explain causal relations between the latent variables. The latter may be *exogenous*, if variables are explicative, or *endogenous* if they can be interpreted also as response variables. The LISREL model (Jöreskog, 1973, 1977; Wiley, 1973; Bollen, 1989) is defined as

$$\boldsymbol{\eta} = \mathbf{B}\,\boldsymbol{\eta} + \boldsymbol{\Gamma}\,\boldsymbol{\xi} + \boldsymbol{\zeta}\,,$$

with measuring equations given by

$$\mathbf{x} = \boldsymbol{\Lambda}_x\,\boldsymbol{\xi} + \boldsymbol{\delta} \quad \text{e} \quad \mathbf{y} = \boldsymbol{\Lambda}_y\,\boldsymbol{\eta} + \boldsymbol{\varepsilon}\,.$$

In this model, the quantities $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are, respectively, the cause and effect latent variables; the observed quantities $\mathbf{x}$ and $\mathbf{y}$ are variables linearly connected to $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ through the matrices of factorial weights $\boldsymbol{\Lambda}_x$ and $\boldsymbol{\Lambda}_y$; $\boldsymbol{\Gamma}$ is the matrix of the coefficients of the cause variable in the structural relation, $\boldsymbol{\zeta}$ is the vector of random errors in the structural relation between $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$, whereas $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$ are the vectors of measurement errors of $\mathbf{x}$ and $\mathbf{y}$ respectively; $\boldsymbol{\zeta}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\delta}$ are not correlated to one another, nor with $\boldsymbol{\xi}$, $\boldsymbol{\delta}$ and $\boldsymbol{\eta}$.

---

6 Crocetta & d'Ovidio (2003) stated that working during the university studies helps job finding prospects. However, if it concerns a job started before university enrolment, it is likely that it is maintained.

7 For the relations of dependence, the previous character is the one represented in the graphic element from which the arrow goes, whereas the following one is the one indicated in the box the arrow points. The relations of interdependence are represented with arcs of circumference that have arrows at both ends.

Then we have:

$$E(\boldsymbol{\zeta}) = E(\boldsymbol{\varepsilon}) = E(\boldsymbol{\delta}) = E(\boldsymbol{\xi}) = E(\boldsymbol{\eta}) = \mathbf{0},$$
$$Cov(\boldsymbol{\zeta}) = \boldsymbol{\Psi}, \quad Cov(\boldsymbol{\varepsilon}) = \boldsymbol{\Theta}_{\varepsilon}, \quad Cov(\boldsymbol{\delta}) = \boldsymbol{\Theta}_{\delta}, \quad Cov(\boldsymbol{\xi}) = \boldsymbol{\Phi},$$

where $\boldsymbol{\Phi}$ is the matrix $k \times k$ of co-variance of latent factors and $\boldsymbol{\Theta}$ are diagonal matrices of only variances.

We can estimate the coefficients and the matrices of variances and co-variances with various techniques (Jöreskog, 1973; Jöreskog & Goldberger, 1975; Browne, 1974). We chose GLS estimators because they are robust to non-normality of the distribution of the latent variables (Browne, 1984)[8]

The LISREL models can be used to analyse the data coming from several groups thus giving the opportunity of making comparisons with control groups or between groups undergoing different treatments.

It is possible to impose constraints on some or all the parameters considered. If we want to compare two groups, it is necessary to estimate each group separately for there to be no bonds; whereas if the data has to be analysed simultaneously to have efficient estimates, crossed constraints must be imposed between groups (Bollen, 1989; Kline, 1998; Civardi & Zavarrone, 2000, 2002).

We have to check whether the matrices of co-variances and correlation of the observed variables are equal for each group. To verify the equality of the matrices of correlation of **x**, it is necessary to set $\boldsymbol{\Theta}_{i;\delta} = \mathbf{0}$ and $\boldsymbol{\Lambda}_{i;x}$, as diagonal matrices of the standard deviations of **x** (where i=1, …, m denotes the group) and $\mathbf{0}$ is a null matrix. Testing the hypothesis of equality between correlation matrices is like checking that $\boldsymbol{\Psi}_1 = \boldsymbol{\Psi}_2$, where $\boldsymbol{\Psi}_i$ is the correlation matrix between the latent factors of group i.

If the hypothesis of invariance of the model is refused without any restriction, more constraints may be imposed to verify the causes of the lack of equality. First, the hypothesis of invariance of the initial factor weights can be tested for the measurement model in each of the groups. If this hypothesis is not acceptable, the invariance of the covariance of the unique factors and factor weights can be tested. The third hypothesis foresees the test of invariance of the co-variances of the unique factors and of the variances of the common factors and factor loadings.

If the hypotheses are less rigid, we can use the first additional hypothesis of the structural model. This foresees covariance matrices of the unique invariant factors symmetrical with some similar elements. Then the case with other constraints, in which the covariance matrices of the unique factors are invariant and symmetric to some elements set to zero, can be tested.

---

8 The quantification of each categorical variable through optimal scaling is referred to a limited number of modalities. So, it is not recommended to assume the hypothesis of normality of the latent variables.

5. A structural model for evaluating how graduates work

LISREL results represent just the starting point of the analysis because to obtain a model with convergent estimates it has been necessary to do small changes. The latent variables used in our analysis are the factors determined by means of factor analysis, with the exception of the variable *irregular studies*, whose importance has been reduced by the relations between latent variables. The starting variables are those obtained with the optimal scaling quantification procedure. The resulting model (see Figure 1) is laid out as a network of relations and it is complex; so we will consider the main correlations between the observed variables surveyed and the latent variables.

The most correlated (observed) variables are age at graduation and time passed between graduation and employment (r=0.95); this variable is also connected to the business field of employment (r=0.49) and study irregularity (r=0.41). It is noticeable the relation between the graduation grade and school leaving grade (r=0.38), and the correlation between type of course chosen and time between graduation and first job (r=0.35). The other correlations considered are lower than 0.33 in absolute terms; they are significant and have to be kept in the model to help the convergence or to improve adaptability.

The standardised regression weights show the direction and intensity of relations between latent and observed variables. Those relations slightly differ from those that come from the explorative analysis, because of the causal relations hypothesised between latent variables and the structure of the factors themselves.

The first factor was called education curriculum because it was directly correlated to graduation and secondary school final grades, and inversely to age at graduation and time between graduating and the last exam. The structural model keeps the same relations and adds a slight connection to the evaluation based on the adequacy of university education and study irregularity. This latent variable is influenced by the graduates' work plans and influences, in its turn, the post-graduate activities.

The latent variable *training and professional path* is mainly measured by the faculty chosen (regression coefficient=0.73), whereas the relations between the types of diploma achieved, the job achieved or sought, and the business field are less strong. The *work plan* factor has a positive relation with the knowledge of post graduation prospects, and negative (-0.69) with work before graduation. It influences (0.32) the non-observed variable, education curriculum. The *postgraduate activity* factor[9] influences three observed variables: years between graduation and employment (0.95), professional or teaching qualification (0.42) and job finding strategies (0.28).

---

[9] Differently from exploratory factor analysis, there is no business field related with the educational and professional pathway, and time between last exam and graduation.

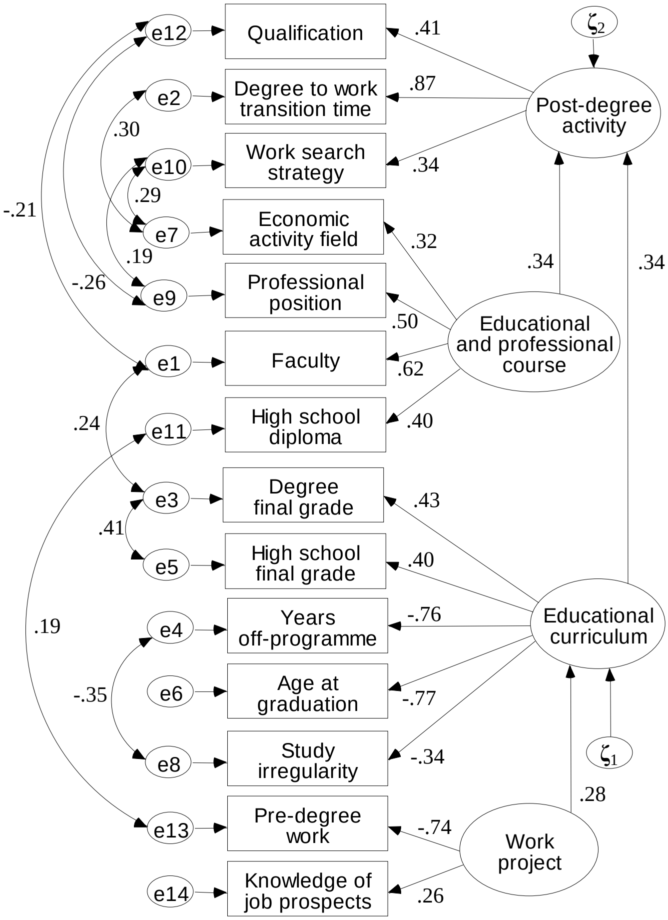

Figure 1. LISREL Model to describe the working way of the graduates of the University of Foggia (Italy)

Table 6. Fit Indexes of the LISREL model used for the analysis of working strategies of graduates from the University of Foggia

Fit Indexes	Models		
	Estimated	Saturated	Independence
ECVI *(Expected Cross-Validation Index)*	0.33	0.10	1.19
GFI *(Goodness of Fit Index)*	0.96	1.00	0.83
AGFI *(Adjusted Goodness of Fit Index)*	0.93	-	0.81
RMR (*Root Mean Square Residual*)	0.07	0.00	0.22
RMSEA *(Root Mean Square Error of Approximation)*	0.06	-	0.11
HOELTER critical N (α=0,05)	*296*	-	*98*

This model represents a first step of research in this field. Indeed, to obtain reliable estimates on employment rates or on the probability of entering the labour market, methods that are more sophisticated have to be used with information that is more detailed.

The obtained model fitted well the data, as is shown by the statistics reported in Table 6. The ECVI index is much closer the minimum value regarding the saturated model (0.10) than the independence one (1.19), indicating that the discrepancy index is rather low. In addition, the GFI index suggests that this model is very close to maximum adaptability[10].

The AGFI (adjusted goodness of fit index) verifies the adaptability of the model looking at the degrees of freedom available to test the model. In our case the value 0.93 is very close to 1, which indicates perfect adaptability.

The RMR (root mean square residual) index is given by the square root of the mean square deviations between the sample variance and its estimate obtained under the assumption that the model is correct. Obviously the lower this index is, the better is its adaptability. In our case, the value 0.07 is much closer to the value of the saturated model (0) than to that of the independence model (0.22).

The RMSEA index is 0.06. In general, a RMSEA lower than 0.08 indicates a good matching of data to the model.

The *N* statistic, for a significance level of 5%, is over the critical threshold (N=200) recommended by Hoelter (1983), whereas for the independence model this statistic is much lower than the suggested level.

Overall, the processed model seems to represent adequately the relations that exist within the data.

6. A structural model of invariance between genders

A variable that may influence the occupational possibility of graduates in Southern Italy is *gender*. Within the groups of graduates analysed here, 54.3% of males and 41.8% of female graduates actually work[11].

Gender in itself is not a determinant of employment, but it is connected to a series of socio-economical obstacles that make it more difficult for women to

[10] The *goodness of fit* (GFI) index is given by the complement to one of the relation between the discrepancy function minimum between model and the sample, in the hypothesis that the variability of the groups is null. This index varies between 0 and 1 where the value 1 indicates the perfect adaptability.

[11] In a segmentation analysis of graduates' placement (Crocetta & d'Ovidio, 2003), this variable appeared in the third branch of the classification tree, thus describing a situation of prevalence of male employment. However, this may depend on the interaction of gender with other structural variables.

enter the labour market than men. Gender is, therefore, a proxy of the lower possibility a woman be assigned roles of responsibility which require either total commitment, or frequent movements, or changes of residence, and implies work discontinuity caused by absences for personal reasons (pregnancy, children's illnesses, etc.)

Gender also influences the choice of educational courses: there are some faculties, such as Literature and Philosophy, with a very low male rate. Another aspect of university choice connected to gender is the distance from home: Antonucci *et al.* (2002) showed a particular inclination for women to choose the university close to their residence.

For these reasons, we decided to check whether the LISREL model described above remains the same with gender. Application of the same structure with males and females immediately produced convergence of the model.

The first hypothesis tested shows invariance of the correlation structure between models estimated separately for male (41.7% of the sample) and female graduates (58.3%): there is no constraint apart from the basic ones ($\boldsymbol{\Theta}_{i;\delta} = \mathbf{0}$; $\boldsymbol{\Lambda}_{i;x} = \mathbf{I}_p$; $\boldsymbol{\Phi}$ matrix with the elements of the main diagonal set to 1). If we test this hypothesis with the minimum value of the discrepancy function, which is distributed as a χ^2, we obtain $C_{MIN} = 40.8$ with 10 degrees of freedom, which corresponds to $p<0.0001$: the hypothesis of *invariance of the structures of correlation* (i.e. of the measuring model) has to be rejected.

The hypothesis of invariance of the *initial factorial weights* is then verified: $C_{MIN}=3.3$ with 3 degrees of freedom ($p=0{,}134$), therefore the hypothesis of invariance of the structural model can be accepted.

Looking at the fit indices (Table 7), it is evident that the adaptation compared to the saturated model does not perform significantly worse, even if the number of groups has more or less halved (the number of female graduates is 1,215, whereas the number the male ones is just 869).

Table 7. Fit Indexes of the LISREL model of invariance of the structural weights (compared to saturated model and model of independence) used for the analysis of working strategies of male and female graduates from the University of Foggia

Fit Indexes	Models		
	Invariance	Saturated	Independence
ECVI *(Expected Cross-Validation Index)*	0.39	0.20	1.23
GFI *(Goodness of Fit Index)*	0.95	1.00	0.83
AGFI *(Adjusted Goodness of Fit Index)*	0.93	1.00	0.80
RMR *(Root Mean Square Residual)*	0.09	0.00	0.22
RMSEA *(Root Mean Square Error of Approximation)*	0.04	-	0.08
HOELTER critical N (α=0,05)	*515*	-	*179*

Table 8. Standardised regression weights of the variables of the LISREL model for the employment analysis of male and female graduates of the University of Foggia

Observed variables	Stand.weights		Observed variables	Stand.weights	
	M	F		M	F
Work project			*Educat./professional course*		
Knowledge of post-degree prospects	0.28	0.27	Faculty	0.63	0.61
			Prof. position current/sought	0.54	0.51
Pre-degree employment	-0.70	-0.71	High school diploma	0.36	0.40
Educational Curriculum			Economic field curr./sought	0.30	0.32
University degree final grade	0.39	0.41	*Post-degree activity*		
High School final grade	0.34	0.35	Number of years between degree and first employment	0.89	0.85
Study irregularity	-0.20	-0.22	Professional/teaching qualification	0.45	0.47
Age at graduation	-0.62	-0.78			
Years off-programme	-0.82	-0.87	Work search strategies	0.35	0.33
Latent components of the variable ***Educational curriculum***			Latent components of the variable ***Post-graduation activity***		
Work project	0.34	0.21	*Educational curriculum*	0.43	0.42
			Educat./professional course	0.34	0.29

* The latent variables in bold type are endogenous.

Although the relations between latent and observed variables changed somewhat (Table 8), the factorial structure is invariant.

The two groups of graduates seem to differ in few aspects, such as the *age at graduation* and, to a lesser extent, the *number of off-programme years*, the *type of high school diploma* and the *number of years after completing the course before graduation*.

The relations between latent variables differ very little between genders, as the structural invariance study indicated.

The analysis confirms that male graduates' approach to the labour market is similar to the one of female graduates, but there are differences in female behaviour towards university training, because women study with more regularity.

7. Conclusions

We selected some variables that represent the pathways through which graduates access the labour market.

The obtained estimates allowed us to measure the external effectiveness of University of Foggia education and to study in non-monetary terms the effects of graduates' choices to enter the labour market[12].

Since the relation between education and employment is complex[13], we considered it worthwhile to underline the result based on the individual student (learning level, ability to finding employment, amount of human capital), that is, the final target of the university teaching in the current social and economical system (Gori, Crema & Vidoni, 2003).

It is thus particularly useful to have a model capable of examining quantitative, ordinal and nominal variables altogether through an appropriate scaling procedure.

The analysed models, both the one applied to the whole sample and those predicted to check the relation-to-gender invariance, showed a good level of matching with the observed data. This allows us to trust the reliability of our estimates and of the hypotheses put forward.

The analysis highlighted that the most influential latent variable on *placement* and success in the labour market is *postgraduate activity*, which depends on the *educational and professional pathway* chosen, on *curriculum* and indirectly on the work planning.

However, we showed that there is no difference between male and female graduates. This does not mean that choices and constraints are the same for the two genders, but that the relations between choices, constraints and latent variables are of the same order, and that the mental structure regarding the decisions is similar for both males and females.

---

12 The comparison between the professional skills of graduates and market demand means "the size and durability of the competences of the trainees, adaptability to situations they have to face, the propensity to learn from experience, the propensity to evolve from technical to managerial work" (Fabbris, 2003).

13 The knowledge determined by the training process and the consequent increase of capacity of finding work are "experience goods", whose effect can only be evaluated expost at different time intervals (Gori & Vittadini, 1999). The results of the training process, as well as the resources, can be measured in monetary or physical quantities (e.g, hours of lessons, number of graduates, etc.), in order to construct productivity indices for processes, structures, results, extending corporate techniques typical of industrial processes to the university case (Bini, 1999). However, a greater quantity of lessons does not imply a better learning and qualification for the labour world (Vittadini, 2001).

References

ANTONUCCI L., CROCETTA C., MECCARIELLO I. (2002) *La popolazione studentesca e l'Università di Foggia, Indagine sulle aspettative degli studenti dopo la riforma*, Uniongrafica Corcelli, Bari.

BINI M. (1999) *Valutazione dell'efficacia dell'istruzione universitaria rispetto al mercato del lavoro,* Rapporto di Ricerca 3/99, Comitato Nazionale per la Valutazione del Sistema Universitario. MIUR, Roma.

BOLLEN K.A. (1989) *Structural Equations with Latent Variables,* Wiley & Sons, New York-Toronto.

BRADLEY R.A., KATTY S.K., COONS I.J. (1962) Optimal scaling for ordered categories, *Psychometrika*, **27**: 355-374.

BROWNE M.W. (1974) Generalized least-squares estimators in the analysis of covariate structures, *South African Statistical Journal*, **8**: 1-24.

BROWNE M.W. (1984) Asymptotically Distribution-Free Methods for the Analysis of Covariance Structures, *British Journal of Mathematical and Statistical Psychology*, **37**: 62-83.

CIVARDI M., ZAVARRONE E. (2000) Il ricorso a modelli invarianti per valutare la qualità della didattica. In: CIVARDI M., FABBRIS L. (eds) *Valutazione della didattica con sistemi computer assisted*, CLEUP, Padova: 179-189.

CIVARDI M., ZAVARRONE E. (2002) Modelli strutturali multigruppo per l'analisi dei tempi di permanenza nel sistema universitario. In: PUGGIONI G. (ed) *Modelli e metodi per l'analisi di rischi sociali e sanitari*, CLEUP, Padova: 1-20.

CROCETTA C., D'OVIDIO F. (2003) La valutazione dell'inserimento lavorativo dei laureati all'Università di Foggia attraverso un'analisi di segmentazione. In: CIVARDI M. (ed) *Transizione Università-Lavoro: la definizione delle competenze*, CLEUP, Padova: 111-132.

CROCETTA C., D'OVIDIO F. (2004) La soddisfazione nella transizione Università-lavoro - Un modello ad equazioni strutturali. In: AURELI CUTILLO E. (ed) *Strategie metodologiche per lo studio della transizione Università-Lavoro*, CLEUP, Padova: 159-187.

DE LEEUW J. (1984) *Canonical Analysis of Categorical Data*, *2nd ed.* DSWO Press, Leiden (NL).

DE LEEUW J., MEULMAN J.J. (1986) Principal component analysis and restricted multidimensional scaling. In: GAUL W., SCHADER M. (eds.) *Classification as a Tool of Research*, Amsterdam (NL): 83-96.

DE LEEUW J., YOUNG F.W., TAKANE Y. (1976) Additive structure in qualitative data: an alternative least squares method with optimal scaling features, *Psychometrika*, 41: 471-504.

FABBRIS L. (2003) *Il monitoraggio dell'inserimento occupazionale e professionale dei laureati e dei diplomati dell'Università di Padova*, Quaderno Pharos n.6, Osservatorio sul Mercato Locale del lavoro, CLEUP, Padova.

GORI E., CREMA F., VIDONI D. (2003) Quali prospettive dalla ricerca sulla qualità e l'efficacia della scuola per la costruzione di sistemi di accountability dell'istruzione, *Non Profit*, **1**, Maggioli, Rimini.

GORI E., VITTADINI G., eds (1999) *Qualità e valutazione nei servizi di pubblica utilità*, ETAS, Milano.

HOELTER J.W. (1983) The analysis of covariance structures, *Sociological Methods and Research*, 11: 325-344.

KLINE R. B. (1998) *Principles and Practice of Structural Equation Modeling*, The Guilford Press, New York.

KRUSKAL J.B. (1965) Analysis of factorial experiments by estimating monotone transformations of the data, *Journal of Royal Statistical Society* (B), 27: 251-263.

JÖRESKOG K.G. (1973) A general method for estimating a linear structural equation system. In: GOLDBERGER A.S., DUNCAN O.D. (eds) *Structural Equation Models in the Social Sciences*, Academic Press, New York: 85-112.

JÖRESKOG K.G. (1977) Structural equation models in the social sciences. In: KRISHNAIAH P.R. (ed) *Application of Statistics*, North Holland, Amsterdam: 265-287.

JÖRESKOG K.G., GOLDBERGER A.S. (1975) Estimation of a model with multiple indicators and multiple causes of a single latent variable, *Journal of the American Statistical Association*, 70: 631-639.

LOVAGLIO P.G. (2000) *Modelli con variabili latenti e indicatori di tipo misto*, Doctoral Thesis, Università degli Studi di Trento.

MEULMAN J.J., HEISER W.J. (1999) *Categories 10.0*, SPSS Inc., Chicago.

VITTADINI G. (1999) Analysis of qualitative variables in structural models with unique solutions. In: VICHI M., OPITZ O. (eds) *Classification and Data Analysis: Theory and Application*, Springer Verlag, Heidelberg: 203-210.

WILEY D.E. (1973) The identification problem for structural equation models with unmeasured variables. In: GOLDBERGER A.S., DUNCAN O.D. (eds) *Structural Equation Models in the Social Sciences*, Academic Press, New York: 69-83.

WRIGHT S. (1934) The methods of path coefficients, *Annals of Mathematical Statistics*, **5**: 161-215.

YOUNG F.W., TAKANE Y., DE LEEUW J. (1978) The principal component of mixed measurement level multivariate data: an alternating least squares method with optimal scaling features, *Psychometrika*, **43**: 279-281.