A Multilevel Analysis of Graduates' Job Satisfaction

Leonardo Grilli, Carla Rampichini

Statistics Department "G. Parenti", University of Florence, Italy

Summary. In this paper, we analyse some aspects of job satisfaction by means of a multilevel factor model, decomposing the factor structure into the graduate and degree programme components, using data from a survey on the 1998 graduates of the University of Florence. Due to the ordinal scale of the response variables, we adopt a multilevel factor model for ordinal variables. The results show that the factor structures at the graduate and study programme levels are not the same, although they are similar; the study programmes with extreme factor scores should be selected for a deeper investigation.
Keywords: Factor model; Job satisfaction; Multilevel model; Ordinal variable.

1. External effectiveness at Florence University

Nowadays it is relevant for the Universities to improve their efficiency and effectiveness, in order to ensure a good allocation of public funds, guarantee the rights of the students and their families to have good services and educational programmes, and, nonetheless, state the relevance of the University as a cultural, social and economic institution.

With this aim, the University of Florence has developed an evaluation system in the last years (Chiandotto *et al.*, 2004). External effectiveness is evaluated with respect to the employment results, such as the employment rate, the time span to the first job, the probability to find a job consistent with the acquired skills. The analysis of job satisfaction is a relevant part of the University evaluation. In the Italian context this issue is treated, among the others, by Santoro & Pisati (1996), Bini (1999), Mazzolli (2000), Bartolozzi (2001).

The main goal of the paper is to analyse and summarise the aspects of job satisfaction by means of a multilevel factor model (Goldstein & McDonald, 1988; Longford & Muthén, 1992), decomposing the factor structure into the graduate and study programme components. To this end, the data are taken from a survey conducted on the 1998 graduates of the University of Florence, interviewed about two years after the degree. Due to the ordinal scale of the response variables, a multilevel factor model for ordinal variables (Skrondal & Rabe-Hesketh, 2004; Grilli & Rampichini, 2006) is specified.

The structure of the paper is as follows. In Section 2 the model is defined, while in Section 3 the results of the analysis of job satisfaction of the 1998 graduates of the University of Florence, taken from a telephone survey conducted, about two years after the degree, are presented. Section 4 concludes our paper.

2. The statistical model

Let $Y_{ij}^{(h)}$ be the h-th ordinal variable (h=1, …, H) observed for the i-th subject (i=1, …, n_j) belonging to the j-th cluster (j=1, …, J). In the following, the subject level will be referred to also with the term 'within' and the cluster level with the term 'between'. In the application presented in Section 3 the clusters are the study programmes, the subjects are the graduates and the ordinal variables are the ratings on 5 items of the questionnaire (H=5).

A two-level factor model for ordinal variables can be set up by defining two components, namely:

- a threshold model which relates a set of continuous latent variables $\tilde{Y}_{ij}^{(h)}$ to the observed ordinal counterparts $Y_{ij}^{(h)}$;
- a two-level factor model for the set of continuous latent variables $\tilde{Y}_{ij}^{(h)}$.

As for the threshold model, let assume that each of the observed responses $Y_{ij}^{(h)}$, which take values in {1, 2, …, C_h}, is generated by a latent continuous variable $\tilde{Y}_{ij}^{(h)}$ through the following relationship:

$$\left\{Y_{ij}^{(h)} = c^{(h)}\right\} \quad \Leftrightarrow \quad \left\{\gamma_{c^{(h)}-1}^{(h)} < \tilde{Y}_{ij}^{(h)} \leq \gamma_{c^{(h)}}^{(h)}\right\}, \tag{1}$$

where the thresholds satisfy the inequality

$$-\infty = \gamma_0^{(h)} \leq \gamma_1^{(h)} \leq \ldots \leq \gamma_{C_h-1}^{(h)} \leq \gamma_{C_h}^{(h)} = +\infty .$$

The factor model can now be defined on the set of latent variables. A general formulation is (Goldstein & McDonald 1988; Longford & Muthén 1992):

$$\tilde{Y}_{ij}^{(h)} = \mu^{(h)} + \left[\sum_{m=1}^{M_u} \lambda_{u,m}^{(h)} u_{mj} + \delta_j^{(h)}\right] + \left[\sum_{m=1}^{M_v} \lambda_{v,m}^{(h)} v_{mij} + \varepsilon_{ij}^{(h)}\right]. \tag{2}$$

In this model the cluster level has M_u factors with corresponding loadings $\lambda_{u,m}^{(h)}$, while the subject level has M_v factors with corresponding loadings $\lambda_{v,m}^{(h)}$. Note that even if $M_u = M_v$ the factor loadings are generally different, the factors may have different interpretations.

Now it is convenient to express the general two-level model (2) for the latent variables in matrix notation:

$$\tilde{\mathbf{Y}}_{ij} = \boldsymbol{\mu} + \left[\boldsymbol{\Lambda}_u \mathbf{u}_j + \boldsymbol{\delta}_j\right] + \left[\boldsymbol{\Lambda}_v \mathbf{v}_{ij} + \boldsymbol{\varepsilon}_{ij}\right], \tag{3}$$

where

- $\tilde{\mathbf{Y}}_{ij} = (\tilde{Y}_{ij}^{(1)}, \cdots, \tilde{Y}_{ij}^{(H)})'$ is the vector of response variables
- $\boldsymbol{\mu} = (\mu^{(1)}, \cdots, \mu^{(H)})'$ is the vector of the means
- $\boldsymbol{\delta}_j = (\delta_j^{(1)}, \cdots, \delta_j^{(H)})'$ are the specific errors at cluster level
- $\mathbf{u}_j = (u_{1j}, \cdots, u_{M_u j})'$ are the common factors at cluster level
- $\boldsymbol{\varepsilon}_{ij} = (\varepsilon_{ij}^{(1)}, \cdots, \varepsilon_{ij}^{(H)})'$ are the specific errors at subject level
- $\mathbf{v}_{ij} = (v_{1ij}, \cdots, v_{M_v ij})'$ are the common factors at subject level
- $\boldsymbol{\Lambda}_u$ is the matrix of factor loadings at cluster level with h-th row $(\lambda_{u,1}^{(h)}, \cdots, \lambda_{u,M_u}^{(h)})$
- $\boldsymbol{\Lambda}_v$ is the matrix of factor loadings at subject level with h-th row $(\lambda_{v,1}^{(h)}, \cdots, \lambda_{v,M_v}^{(h)})$.

The standard assumptions on the item specific errors of model (3) are:

$$\boldsymbol{\delta}_j \overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Psi}_\delta), \qquad \text{where } \boldsymbol{\Psi}_\delta = diag\{(\psi_\delta^{(h)})^2\},$$

$$\boldsymbol{\varepsilon}_{ij} \overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Psi}_\varepsilon), \qquad \text{where } \boldsymbol{\Psi}_\varepsilon = diag\{(\psi_\varepsilon^{(h)})^2\},$$

while for the factors it is assumed that

$$\mathbf{u}_j \overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Sigma}_u), \qquad \mathbf{v}_{ij} \overset{iid}{\sim} N(\mathbf{0}, \boldsymbol{\Sigma}_v),$$

where the covariance matrices $\boldsymbol{\Sigma}_u$ and $\boldsymbol{\Sigma}_v$ are, in principle, unconstrained, but in the following we assume they are diagonal. Moreover, all the errors and factors are assumed mutually independent, so model (3) is equivalent to the following variance decomposition

$$Var(\tilde{\mathbf{Y}}_{ij}) = \left[\boldsymbol{\Lambda}_u \boldsymbol{\Sigma}_u \boldsymbol{\Lambda}_u' + \boldsymbol{\Psi}_\delta\right] + \left[\boldsymbol{\Lambda}_v \boldsymbol{\Sigma}_v \boldsymbol{\Lambda}_v' + \boldsymbol{\Psi}_\varepsilon\right]. \tag{4}$$

This amounts to a couple of factor models, one for the between covariance matrix and the other for the within covariance matrix (Muthén, 1994).

The outlined factor model (3) raises several identification issues, related to the two components: *(i)* the threshold model which relates the continuous latent variables $\tilde{Y}_{ij}^{(h)}$ to their observed ordinal counterparts $Y_{ij}^{(h)}$; and *(ii)* the two-level factor model for the continuous latent variables $\tilde{Y}_{ij}^{(h)}$.

The total relative communality for the h-th item can be computed as

$$\frac{\sum_{m=1}^{M_v}\left(\lambda_{v,m}^{(h)}\right)^2\sigma_{v,m}^2+\sum_{m=1}^{M_u}\left(\lambda_{u,m}^{(h)}\right)^2\sigma_{u,m}^2}{Var_T\left(\tilde{Y}_{ij}^{(h)}\right)}. \tag{5}$$

The proportion of total variance (relative communality) of the *h*-th item explained by the *k*-th subject-level factor, can be computed as

$$\frac{\left(\lambda_{v,k}^{(h)}\right)^2\sigma_{v,k}^2}{Var_T\left(\tilde{Y}_{ij}^{(h)}\right)}, \tag{6}$$

while the proportion of total variance of the h-th item explained by a cluster-level factor can be computed similarly. Note that, since the covariance between the *h*-th latent variable and *k*-th subject-level factor is $\lambda_{v,k}^{(h)}\sigma_{v,k}^2$, the corresponding correlation equals the square root of the proportion of total variance (6).

Finally, it is to be stressed that, even if all the estimable quantities are expressed in terms of the item-specific subject-level standard deviations $\psi_{\varepsilon}^{(h)}$ (Grilli & Rampichini, 2006), the interpretable quantities just described are unaffected by the item scale, since they are ratios of parameters within the same item.

3. Results

We used the ordinal multilevel factor model to analyse five items on job satisfaction of employed Florentine graduates.

Altogether, the considered data set includes 2,432 graduates from 36 study programmes, with a highly unbalanced structure: the minimum, median and maximum number of employed graduates by programme are: 3.0, 31.5 and 495, respectively.

Table 1. Univariate distributions of job satisfaction items. 1998 graduates, University of Florence.

Items	*Levels of satisfaction*					Total	
	1	2	3	4	5	%	N
a. Earning	7.8	23.9	38.1	20.5	9.7	100.0	2421
b. Career	11.0	28.2	32.6	18.0	10.2	100.0	2393
c. Consistency	24.5	27.5	24.2	12.5	11.3	100.0	2427
d. Professionalism	26.0	40.3	22.8	7.7	3.2	100.0	2420
e. Interests	21.5	32.7	28.2	10.8	6.8	100.0	2419

The question "*How much are you satisfied with the following aspects of your present job*?" required a response on a five point scale: *1*. absolutely satisfied, *2*. very satisfied, *3*. satisfied, *4*. unsatisfied, *5*. very unsatisfied. The five considered items are: *a*. earning, *b*. career's opportunities, *c*. consistency of job with degree programme curriculum, *d*. acquisition of competences (professionalism), *e*. correspondence with own cultural interests. The distributions of the items are reported in Table 1. Note that the number of responses for each item is different, due to partial non-response.

The main aim of the analysis is to describe and summarise the aspects of satisfaction measured by the five considered items, separately for the graduate and degree programme levels. The two-level factor model for ordinal variables defined in Section 2 is a useful tool to achieve this goal. In our application, the model is fitted by maximum likelihood with adaptive numerical integration, as implemented in the *GLLAMM* procedure of Stata (Rabe-Hesketh *et al*., 2004). Since the model fitting process is very time-consuming, it is useful to follow a step-by-step procedure:

1. *Univariate two-level models*. As a first step, it is advisable to fit a set of univariate ordinal probit variance component models, one for each item, using standard multilevel software. The estimated proportions of between variance ($\mathrm{ICC}^{(h)}$) allow us to evaluate if a two-level analysis is worthwhile, while a comparison of the thresholds among the items should give some hints about the restrictions to be imposed in the multivariate model.
2. *Exploratory non-hierarchical factor analysis*. In order to shade some light upon the covariance structure of the data, it is useful to estimate the matrix of product-moment correlations among the latent variables, i.e. the polychoric correlation matrix of the items, and to use this matrix to perform an exploratory non-hierarchical (i.e. single-level) factor analysis by means of standard software.
3. *Exploratory between and within factor analyses*. More specific suggestions for the two-level model specification can be obtained from separate exploratory factor analyses on the estimated between and within correlation matrices of the latent variables. The results of this two-stage proce-

dure are expected to be similar to that obtained from the full two-level analysis, as in the continuous case (Longford & Muthén 1992).

4. *Confirmatory two-level factor analysis.* The results of the exploratory two-stage factor analysis, as outlined in point 1, are used to specify one or more confirmatory two-level ordinal factor models as defined by equation (2) of Section 2. These models can be fitted with likelihood or Bayesian methods, and compared with reference to appropriate indicators. The exploratory two-stage factor analysis of point 1 provides fine initial values for the chosen estimation procedure, which may allow a substantial gain in computational time.

3.1 Univariate two-level models

The analysis begins by fitting five univariate ordinal *probit* variance component models. The results, obtained with GLLAMM, are reported in Table 2.

The between proportion of variance, expressed by the ICC, is significantly different from zero for all items. Note that the ICC value for the first three items is about 6-7%, which is measurable in a framework with categorical variables and indicates that a non-negligible part of variance can be explained by degree programme factors.

In a factor model for ordinal variables the thresholds can be left free, while fixing the item means and standard deviations. However, when all the items are on the same scale (and thus $C_h=C$ for each h) a more parsimonious specification can be achieved by assuming that the thresholds differ among the items only by a linear transformation, i.e. $(\gamma_c - \mu^{(h)})/\psi_{\varepsilon}^{(h)}$, where $\gamma_1,\ldots,\gamma_{c-1}$ is a set of thresholds common to all the items (Grilli & Rampichini, 2006).

In this application, the linear restriction on the thresholds is supported by the entries of Table 2. In fact, the differences between adjacent thresholds among the items are similar, except for the third one, which has smaller differences. This suggests that the third item has a higher variability, as also confirmed by the variances calculated after item scoring (Table 5).

Table 2. Univariate ordinal *probit* variance component models: estimated ICC and thresholds. 1998 graduates, University of Florence.

Items	*ICC (%)*	*Thresholds*			
		γ_1	γ_2	γ_3	γ_4
a. Earning	6.0	-1.53	-0.55	0.47	1.27
b. Career	7.4	-1.37	-0.38	0.52	1.24
c. Consistency	6.8	-0.69	0.05	0.71	1.21
d. Professionalism	2.2	-0.64	0.44	1.26	1.88
e. Interests	2.4	-0.77	0.14	0.98	1.54

In light of these remarks, the two-level confirmatory factor model of Section 3.4 will include only one set of thresholds, $\gamma_1,\ldots,\gamma_4$ while allowing the item means $\mu^{(h)}$ and item standard errors $\psi_{\varepsilon}^{(h)}$ to be freely estimated (with the exception of a reference item).

3.2 Exploratory non-hierarchical factor analysis

The second step requires the estimation of the matrix of product-moment correlations among the latent variables, i.e. the polychoric correlation matrix (see Table 3), whose entries are all significant.

We performed an exploratory maximum likelihood factor analysis on this matrix. The results of this analysis (Table 4) suggest the presence of two factors: a *cultural* factor (labelled *Factor 1*), that explains chiefly the *Consistency-Professionalism-Interests* correlations, and a *status* factor (labelled *Factor 2*), explaining mainly the *Earning-Career* correlation.

Given the low proportions of between variance (ICC of Table 2), this structure is expected to be quite similar to the within structure, thought it may be very different from the between structure.

Table 3. Polychoric correlation matrix of the items. 1998 graduates, University of Florence.

Item	*a*	*b*	*c*	*d*	*E*
a. Earning	1.00				
b. Career	0.54	1.00			
c. Consistency	0.11	0.25	1.00		
d. Professionalism	0.28	0.45	0.54	1.00	
e. Interests	0.16	0.33	0.61	0.58	1.00

Table 4. Exploratory factor analysis on the polychoric correlation matrix of the items: varimax rotated factors and communalities. 1998 graduates, University of Florence.

Item	Factor pattern		*Communality*
	Factor 1	Factor 2	
a. Earning	0.08	0.65	0.43
b. Career	0.26	0.80	0.70
c. Consistency	0.77	0.07	0.60
d. Professionalism	0.68	0.34	0.58
e. Interests	0.78	0.16	0.63

3.3 Exploratory between and within factor analyses

The third step of analysis calls for the decomposition of the overall correlation matrix of the latent variables into the between and within components. This task would require the fitting of a two-level multivariate ordinal model with five random effects for each level, which takes too long to be fitted with numerical integration. Therefore, an approximate procedure is adopted, assigning a score to each item category. Various sophisticated scoring systems could be applied (Fielding, 1999), but given the preliminary nature of this step, the simplest scoring system is applied, assigning the rank value to each category. After scoring, the within and between covariance matrices can be estimated by fitting a multivariate two-level model for continuous responses. To this end, the MLwiN software with RIGLS algorithm (Goldstein *et al.*, 1998) is used, yielding restricted maximum likelihood estimates, which are better for the estimation of variance-covariance parameters than unrestricted ones.

The results are shown in Tables 5 and 6. As for Table 5, note the following points:

- it is clear from the last row of Table 5 that the third item (*Consistency*) has the higher variability, as yet noted in the univariate analysis (Table 2);
- the between proportions of variance are in line with ICC of Table 2;
- the between proportions tend to be higher for covariances than for variances.

As for Table 6, note the following points:

- the total correlation matrix, which is obtained from the between and within components, is similar to the polychoric correlation matrix (Table 3), with a moderate attenuation;
- the structures of the between and within correlation matrices are quite different. Particularly, the between correlations are always higher than the within correlations: this means that the factor model, which explains the correlations, is suitable for the between level even more than might be appreciated by simply looking at the total correlation matrix;

Table 5. Two-level multivariate model on item scores: between variance-covariance percentage and total variance of items. 1998 graduates, University of Florence.

Item	*a*	*b*	*c*	*d*	*e*
a. Earning	5.90				
b. Career	12.96	8.63			
c. Consistency	21.02	13.09	7.37		
d. Professionalism	10.55	7.63	6.62	2.30	
e. Interests	9.57	4.75	6.07	2.76	2.36
Total variance	1.15	1.31	1.68	1.04	1.31

Table 6. Two-level multivariate model on item scores: correlation matrices. 1998 graduates, University of Florence.

ITEM	*a*	*b*	*c*	*d*	*e*
Between					
a. Earning	1.00				
b. Career	0.89	1.00			
c. Consistency	0.36	0.40	1.00		
d. Professionalism	0.72	0.69	0.79	1.00	
e. Interests	0.39	0.32	0.81	0.62	1.00
Within					
a. Earning	1.00				
b. Career	0.46	1.00			
c. Consistency	0.10	0.23	1.00		
d. Professionalism	0.23	0.39	0.48	1.00	
e. Interests	0.14	0.31	0.55	0.52	1.00
Total					
a. Earning	1.00				
b. Career	0.49	1.00			
c. Consistency	0.11	0.24	1.00		
d. Professionalism	0.25	0.40	0.49	1.00	
e. Interests	0.15	0.30	0.55	0.53	1.00

- the within correlation matrix is similar to the total correlation matrix, due to the low proportion of between variances and covariances.

The results of the exploratory maximum likelihood factor analyses performed on the within and between correlation matrices of Table 6 are reported in Tables 7 and 8, respectively.

As for the within structure (Table 7), Bartlett's test indicates that two factors are sufficient (*p*-value=0.5082). The factor patterns are similar to those found in the non-hierarchical analysis (Table 4).

Table 7. Exploratory maximum likelihood factor analysis on the within correlation matrix: varimax rotated factor loads and communalities. 1998 graduates, University of Florence.

Item	*Factor pattern*		*Communality*
	Factor 1	*Factor 2*	
a. Earning	0.07	0.59	0.35
b. Career	0.25	0.75	0.63
c. Consistency	0.72	0.07	0.53
d. Professionalism	0.64	0.32	0.50
e. Interests	0.74	0.16	0.58

Table 8. Exploratory maximum likelihood factor analysis on the between correlation matrix: factor loads and communalities. 1998 graduates, University of Florence.

Item	Factor pattern		Communality
	Factor 1	Factor 2	
a. Earning	0.00	1.00	1.00
b. Career	0.08	0.89	0.80
c. Consistency	0.93	0.36	1.00
d. Professionalism	0.57	0.72	0.84
e. Interests	0.71	0.39	0.66

As for the between structure (Table 8), while one factor is not enough, the estimation with two or more factors encounters a Heywood case. We decided to retain two factors, forcing the specificities to be non-negative. The second factor loads all items, while the first factor presents relevant loadings only for the last three items.

3.4 Confirmatory two-level factor analysis

Finally, in the light of the results of the preliminary analysis of Section 3.3, a two-level confirmatory factor analysis is performed using model (2). The model is fitted with *GLLAMM*, via adaptive numerical integration with five quadrature points. This is a flexible procedure, but as the complexity of the random part of the model increases, the computational time becomes very long. Since we are not particularly interested in decomposing item specificities, in order to reduce the computational effort the between error terms $\delta_j^{(h)}$ are omitted, so the variances of the remaining item-specific errors $\varepsilon_{ij}^{(h)}$ represent total specificities[1].

The within and between structures emerging from the exploratory analyses are not equally reliable: the within part is estimated on a large number of observations and Bartlett's test clearly indicates the presence of two factors, while the between part is estimated on only 36 degree programmes and the estimation is complicated by the presence of an Heywood case.

Therefore, for the within part of the model the two-factor structure suggested by the exploratory within factor analysis (Table 7) is retained, constraining to zero the loadings that were close to zero, that is the loading of *Earning* in the first factor and the loadings of *Consistency* and *Interests* in the second. As for the between structure, since the hints from the exploratory analysis are less clear, two configurations at this level have been tried:

[1] Grilli & Rampichini (2006) discuss the consequences of this choice.

(*i*) a one-factor unconstrained structure (model M1); and (*ii*) a two-factor structure (model M2), with unconstrained loadings in the first factor and two loadings equal to zero in the second factor (*Earning* and *Career*, see Table 8).

Models M1 and M2 are fitted by means of GLLAMM, using maximum likelihood with five-point adaptive quadrature. In both cases, convergence is achieved after a few iterations but the computational times are in terms of several days.

The likelihood ratio test comparing the models M1 and M2 clearly indicates that the second is better (LR statistic=95.6, *df*=3). The preferred model M2 has 27 estimable parameters: 4 item means $\mu^{(h)}$, 4 common thresholds γ_c, 4 specificities $\psi_{\varepsilon}^{(h)}$, 5 factor loadings $\lambda_{v,m}^{(h)}$ and 2 factor variances $\sigma_{v,m}^{(2)}$ at the student level (*m*=1,2), 6 factor loadings $\lambda_{u,m}^{(h)}$ and 2 factor variances $\sigma_{u,m}^{(2)}$ at the degree program level (*m*=1,2). The parameter estimates are reported in Table 9.

The interesting feature of the model is the covariance structure at both levels, which does not depend on the item means and thresholds and can be summarized by the communalities (Table 10). These values are obtained as suitable transformations of model parameters: specifically, the factor '*%Communalities*' are computed from formulae such as (6), the '*Total %Communality*' is obtained by summing the row values FW1, FW2, FB1 and FB2 (see equation (5)), while the last column of the Table is the percentage of total communality due to the between level. The following points should be noted:

- for the first three items the between component is greater for the communality (last column of Table 10) than for the total variance (ICC of Table 2);
- the last two items, *Professionalism and Interests,* are poorly explained by the factors at degree programme level;
- the first factor at the degree programme level, FB1, is interpretable as a status factor, while the second one, FB2, is essentially related to *Consistency*.

Table 9. Confirmatory two-level factor analysis: model M2 parameter estimates. 1998 graduates, University of Florence.

Item	*Loadings*				$\psi_{\varepsilon}^{(h)}$	*Mean*
	Within		*Between*			
	$\lambda_{v,1}^{(h)}$	$\lambda_{v,2}^{(h)}$	$\lambda_{u,1}^{(h)}$	$\lambda_{u,2}^{(h)}$		
a. Earning	-	0.80	0.70	-	1.24	0.21
b. Career	$1^{(*)}$	$1^{(*)}$	$1^{(*)}$	-	$1^{(*)}$	$0^{(*)}$
c. Consistency	3.45	-	0.24	$1^{(*)}$	1.55	-0.79
d. Professionalism	2.34	0.36	0.27	0.19	1.16	-1.42
e. Interests	3.09	-	0.01	0.30	1.15	-0.94
Factor variance	0.26	1.66	0.31	0.34		
Thresholds: $\gamma_1 = -2.71$, $\gamma_2 = -0.69$, $\gamma_3 = 1.13$, $\gamma_4 = 2.52$						

The symbol (*) denotes a fixed value.

Table 10. Confirmatory two-level factor analysis: communalities. 1998 graduates, University of Florence.

Item	*% Communality*					*% Between on Total*
	Within		*Between*		*Total*	
	FW1	*FW2*	*FB1*	*FB2*		
a. Earning	-	38.9	5.5	-	44.4	12.4
b. Career	8.2	51.4	9.5	-	69.1	13.7
c. Consistency	53.2	-	0.3	5.8	59.4	10.3
d. Professionalism	47.5	7.2	0.7	0.4	55.8	2.0
e. Interests	65.1	-	0.0	0.8	65.9	1.2

The factor scores of degree programmes are represented in Figure 1, where the labels concern the degree programmes having at least one score greater than 0.5 or less than -0.5. The points on the right side of the diagram indicate a high satisfaction on *Earning* and *Career*, while the points at the top denote a

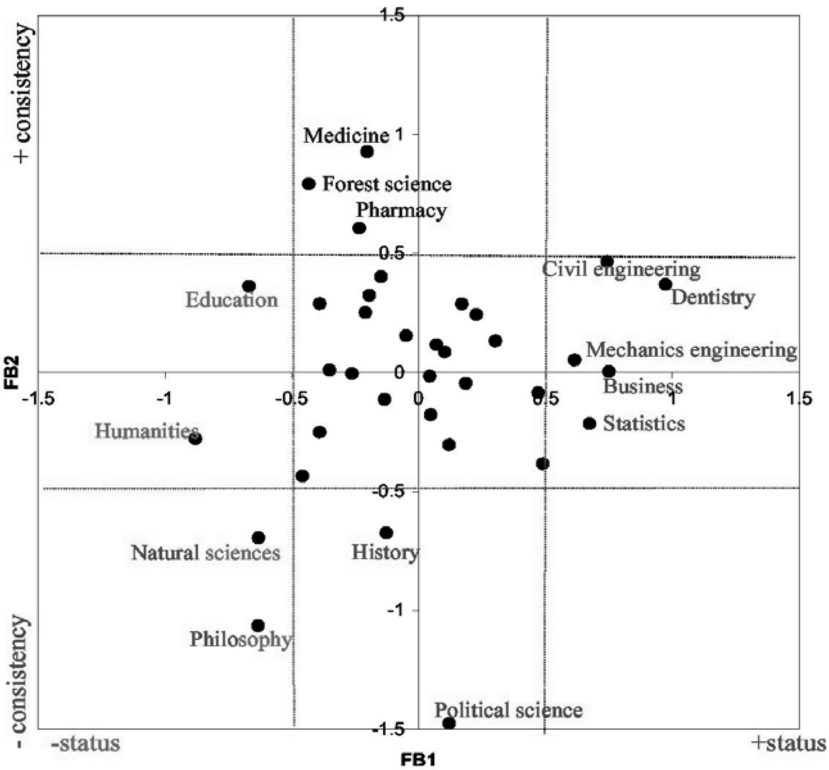

Figure 1. Estimated factor scores for the degree programmes

high satisfaction on *Consistency*. Note that some degree programmes are low only on one dimension (as Humanities on FB1 and Political Science on FB2), while there are two degree programmes lying in the left-down corner (Philosophy and Natural Sciences) with low satisfaction on both dimensions.

4. Concluding remarks

Our analysis showed that there are two relevant factors at both graduate and study programme level. The *status* factors, FW2 and FB1 on Table 9, essentially determine the same variables, *Earning* and *Career*. The other factors, FW1 and FB2 on Table 9, both have a high loading of *Consistency*, but the loadings of *Professionalism* and *Interests* are relevant only at the graduate level. This is in line with the subjective nature of such aspects of the job.

Looking at the estimated factor scores at the degree programme level (Figure 1), extreme cases should be selected for further investigation.

The analysis could be deepened by adding individual-level covariates. This extension is straightforward and does not require a significant additional computational effort.

At present, the major obstacle to a wide use of multilevel factor models for ordinal variables is due to software limitations. The GLLAMM procedure of STATA is very flexible, but it was extremely slow in the present application. Alternative software for fitting such models is M*plus* (Muthén & Muthén, 2003). Currently M*plus* cannot fit exactly the same model used in our analysis, but a trial with a slightly different version of the model shows that the modified EM algorithm implemented in M*plus* is considerably faster, achieving convergence in a few hours.

Anyway, even if fast estimation algorithms are available, it is advisable, especially in the case of ordinal response variables, to fit the multilevel factor model as the final step of the analysis, after having explored the data with simpler techniques.

References

BARTOLOZZI M. (2001) Analisi della soddisfazione sul lavoro tramite un modello ad equazioni strutturali, *Studi e note di economia*, **3**: 163-179.

BINI M. (1999) *Valutazione della Efficacia dell'Istruzione Universitaria rispetto al Mercato del Lavoro,* Rapporto di Ricerca, Osservatorio per la Valutazione del Sistema Universitario - Ministero dell'Università e della Ricerca Scientifica e Tecnologica, Roma.

CHIANDOTTO B., BACCI S., BERTACCINI B. (2004) *Profilo e sbocchi occupazionali dei laureati e diplomati dell'Ateneo fiorentino nell'anno 2000,* University of Florence. http://valmon.ds.unifi.it.

FIELDING A. (1999) Why use arbitrary point scores? Ordered categories in models of educational progress, *Journal of the Royal Statistical Society (A)*, **162**: 303-328.

GOLDSTEIN H., MCDONALD R. P. (1988) A general model for the analysis of multilevel data, *Psychometrika*, **53**: 455-467.

GOLDSTEIN H., RASBASH J., PLEWIS I., DRAPER D., BROWNE W., YANG M., WOODHOUSE G., HEALY M.J.R. (1998) *A User's Guide to MLwiN*, Institute of Education, London.

GRILLI L., RAMPICHINI C. (2006) Multilevel factor models for ordinal variables, *Structural Equation Modeling* (in press).

LONGFORD N., MUTHÉN B. (1992) Factor analysis for clustered observations, *Psychometrika*, **57**: 581-597.

MAZZOLLI B. (2000) *Valutazione dell'efficacia dell'ateneo fiorentino tramite un modello multilivello ad equazioni strutturali*, Ph.d. Thesis, Department of Statistics "G. Parenti", University of Florence.

MUTHÉN B. (1994) Multilevel covariance structure analysis, *Sociological Methods & Research*, **22**: 376-398.

MUTHÉN L.K., MUTHÉN B.O. (2003) M*plus User's Guide* (version 3). Muthén & Muthén, Los Angeles, CA.

RABE-HESKETH S., SKRONDAL A., PICKLES A. (2004) *GLLAMM Manual*. U.C. Berkeley Division of Biostatistics Working Paper Series, Working Paper 160.

SKRONDAL A., RABE-HESKETH S. (2004) *Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models*, Chapman & Hall/ CRC Press, Boca Raton.

SANTORO M., PISATI M. (1996) *Dopo la laurea. Status, sfide, strategie*, Il Mulino, Bologna.