

LC-IRT models with covariates in Polish job satisfaction analysis

Genge, Ewa

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Genge, E. (2020). LC-IRT models with covariates in Polish job satisfaction analysis. *Argumenta Oeconomica*, 44(1), 207-226. <https://doi.org/10.15611/aoe.2020.1.09>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by-nc-nd/3.0/deed.de>

Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:

<https://creativecommons.org/licenses/by-nc-nd/3.0>

*Ewa Genge**

LC-IRT MODELS WITH COVARIATES IN POLISH JOB SATISFACTION ANALYSIS

Employment is at the heart of European Union (EU) policies as it is the basis for wealth creation. Knowing how satisfied EU residents are with their occupation is very important, since losing one's job may undermine one's life satisfaction and its overall meaning (European Commission 2015). According to the most recent Eurostat data (European Commission 2017), Poland reported an average job satisfaction well above the EU mean, ranked 8th (behind Denmark, Iceland, Austria, Finland, Norway, Switzerland and Sweden). Thus, it is interesting to present an analysis focused on the job satisfaction of workers in Poland – a country of emigration, with the highest percentage of temporary contracts in Europe (European Commission 2016).

The main aim of our study is understanding how the different socio-economic features affect the groups of workers with similar job satisfaction levels in Poland. Most of the Polish job satisfaction studies are focused on selected professional groups, in selected regions of Poland. This article presents another, the latent variable models approach to the heterogeneous data set for different subgroups of workers in all the regions of Poland. The combination of the two latent variable models enables to find homogeneous classes of individuals characterized by the similar latent ability levels, and at the same time, the item characteristics analysis (usually identified as discrimination indices and difficulty parameters) as well. Latent Class Item Response Theory (LC-IRT) models are more flexible in comparison with traditional formulations of Item Response Theory (IRT) models, often based on restrictive assumptions, such as normality of latent trait (explicitly introduced). Moreover, the authors also apply the extended latent variable models under the discrete assumption of the latent trait including individual socio-demographic features, such as age, sex, education, marital status or current financial situation.

The article analyzes data collected as part of the International Social Survey Programme 2015 using R software. The results may help policymakers tailor their employment policies as well as to create and deliver services focused on special socio-economic groups of the Polish society.

Keywords: job satisfaction, Polish workers

JEL Classifications: C1, C35, J81

DOI: 10.15611/aoe.2020.1.09

* University of Economics in Katowice.

1. INTRODUCTION

Job satisfaction results from the assessment made by employees about their job in terms of meeting their wants, desires or needs (Lambert and Hogan, 2009). It can be considered in a variety of ways and in many various studies it has been defined differently (Lam, Zhang and Baum, 2001).

One of the most widely used definitions of job satisfaction is “a pleasurable or positive emotional state resulting from the appraisal of one’s job or job experiences” (Locke, 1976). An individual’s satisfaction is derived from a process of comparing an aspect of the current job with a frame of reference (Smith et al., 1969).

Locke (1969) defined job satisfaction as “the pleasurable emotional state resulting from the appraisal of one’s job as achieving or facilitating the achievement of one’s job values”. It is also suggested that job satisfaction reflects the feelings of employees toward their work (Babakus et al., 2003; Eliyana, Yusuf and Prabowo, 2012). Similarly, Spector (1997) indicated that job satisfaction is reflected in how people feel about their work.

The studies also showed that job satisfaction is one of the most important predictors of overall life satisfaction (Judge and Watanabe, 1993; Rain, 1991).

Furthermore, job satisfaction, or more precisely job dissatisfaction, may have adverse impacts on individuals’ physical and mental health and attitudes. The consequences of job dissatisfaction may also include absenteeism and grievances (Locke, 1976). Some studies suggest that the residual of the job satisfaction equation represents individuals’ propensity to quit their present job position (Lévy-Garboua et al., 2007). This implies that the analysis of the job satisfaction variable may help to gain insight into workers’ behavior.

Several job satisfaction measurement tools have been developed for use in high-income countries (Brayfield and Rothe, 1951; Weiss et al., 1967; Smith et al., 1969, Judge et al., 2003; Bruno et al., 2013). These commonly measure an individual’s satisfaction with different aspects of their job, as well as overall job satisfaction (Bruck et al., 2002). However, the existing literature on the subject of job satisfaction – especially in the context of Eastern European countries – remains scant.

The job satisfaction of Russian workers has been studied in papers such as Linz (2003), Linz and Semykina (2012), and Bartolucci et al. (2015). As far as the Polish society is concerned, Wilczyńska et al. (2016) applied two separated logit models for the group of knowledge workers and the group of other workers. The presented results showed that knowledge workers are

more likely to attain job satisfaction working on their own account instead of having a permanent contract. This finding proved that, despite the common view, flexible employment is not necessarily worse than a permanent contract.

Gaszynska et al. (2014) assessed the level of life and job satisfaction of Polish anesthesiologists using the nonparametric Mann-Whitney U test and principal component analysis with varimax rotation. Kowal and Roztocki (2015) examined the effects of gender difference on the job satisfaction of information technology professionals in Poland. Statistical methods in this work comprised mainly descriptive statistics, analysis of variance and factor analysis.

Grobelna et al. (2016) examined selected antecedents of hotel employees' job satisfaction and its critical consequences which are crucial for both individuals and hotel organizations as a whole using structural equation modeling.

Brayer et al. (2017) analyzed the professional satisfaction of Polish nurses on the basis of factor analysis with the use of varimax rotation. The aim of this work was to determine factors contributing to job satisfaction among Master's-degree nurses in central and eastern Poland.

The remainder of the paper is organized as follows. The next section provides the justification for tackling Polish job satisfaction analysis. Section 3 presents the theoretical framework. The dataset and the empirical analysis are presented in Sections 4 and 5. The final remarks are presented in the last section of the article.

2. BACKGROUND

Most Polish job satisfaction studies are focused on selected professional groups, in selected regions of Poland. Additionally, the assumption of many of the applied multivariate statistical methods is the homogeneity of the data set of observations (explicitly introduced). The problem of heterogeneity of the analyzed data is very important (especially when responses are observed on each subject, i.e. for different subgroups in all of the regions of Poland). Then it is often simply assumed (without prior testing) that the analyzed data set is homogeneous and statistical analyses are carried out. Furthermore, the questionnaire analyses in the presented studies are based mainly on Cronbach's alpha coefficient, the most common measure of the scale reliability. It is worth emphasizing that the methods built on classical test theory require multiple assumptions, which are very often not tested. In turn, the latent class item response theory (LC-IRT) models (under the assumption

of discreteness of the latent trait¹) are used to clustering observational units into a – preferably small – set of homogenous latent classes and, at the same time, to properly study item characteristics (Bacci et al., 2014; Bartolucci et al., 2014).

Moreover, the authors realize that job satisfaction is influenced by a whole set of factors going beyond the factual aspects of employment. The measurable level of job satisfaction may not include factors which an employee could directly acknowledge as connected with work processes (Butlewski and Tytyk, 2008). Therefore, this article presents another – the latent variable models approach – to Polish job satisfaction analysis. This approach combines two types of latent variable models: item response theory model and latent class model. In the latent class analysis the latent variable is discrete and denotes the number of classes in the population. In the IRT model the continuous latent variable is used to represent the trait of interest (i.e. ability in certain subjects, the level of job satisfaction). The combination of the two latent variable models concerns the assumption that the population under study is composed of homogeneous classes of individuals who have very similar latent ability levels. The main part of this study aimed at understanding how the different socio-economic features affect the groups of workers with the similar job satisfaction levels in Poland.

3. POLYTOMOUS LC-IRT MODELS WITH COVARIATES

Let n denote the number of individuals to j ($j = 1, \dots, m$) items of the questionnaire, X_j denote the response variable for the j -th item. This variable has l_j categories indexed from 0 to $l_j - 1$. The authors also suppose that the items, measure a latent trait Θ , i.e. the saving skills, general satisfaction, ability in the educational context². A crucial assumption characterizing the LC-IRT models concerns the discreteness of the distribution of the latent trait which corresponds to the s ($s = 1, \dots, u$) latent classes of subjects with the same ability level (support point) ξ_s .

¹ Alternative item response theory (IRT) is a model-based theory, established on the idea that the responses to each test item depend on some person and item characteristics, according to specific probabilistic relations. Traditional IRT models assume the continuous (usually normal) distribution for the latent trait.

² The authors consider that the questionnaire is aimed to measure, for each individual i the level of Θ of the certain latent trait.

Additionally, the probability to belong to a given latent class $\pi_s = p(\Theta = \zeta_s)$ may be influenced by some individual level covariates, then \mathbf{z} denotes the column vector of the covariates corresponding to the sample unit. The covariates are considered as fixed and known. In this way, there is not a set of weight π_s common to all subjects in the sample, but weights $\pi_s(\mathbf{z})$ which are subject-specific, i.e. there is $\pi_s(\mathbf{z}) = p(\mathbf{Z} = \mathbf{z})$ for each individual i ($i = 1, \dots, n$) and latent class s ($s = 1, \dots, u$).

Then, the rule to obtain the manifest distribution of the response vector $\mathbf{X} = X_1, \dots, X_m$ is extended as follows (see Bacci et al., 2014):

$$p(\mathbf{x}) = p(\mathbf{X} = \mathbf{x}) = \sum_{s=1}^u p(\mathbf{x} | \zeta_s) \pi_s(\mathbf{z}). \quad (1)$$

The posterior distribution is defined as:

$$p(\zeta_s | \mathbf{x}) = \frac{p(\mathbf{x} | \zeta_s) \pi_s(\mathbf{z})}{p(\mathbf{x})}, \quad (2)$$

and the conditional probabilities $p(X_j = x | \Theta = \zeta_s)$ depend on the nature of the response variables. Then each of the known item parameterizations may be obtained in latent class IRT approach. In the case of the binary variables, the two-parameter logistic (2PL) specification (Birnbbaum, 1968) or Rasch type model (Rasch, 1960) may be adopted.

There are many different classifications of IRT models for polytomous items (Van der Linden and Hambleton, 1997; Nering and Ostini, 2010). Based on the specification of the link function and the adopted constraints on the item parameters (α_j and \mathcal{G}_{jx} are item parameters which are usually identified as discrimination indices and difficulty levels³ and on which suitable constraints may be assumed), different IRT models for polytomous responses results.

Among the most well-known models for polytomous responses are: the Graded Response Model (GRM; Semejima, 1969) and the Generalized Partial Credit Model (GPCM; Muraki, 1992), models with free both item parameters, global and local logit link function, respectively, as well as the

³ The item difficulty parameter for ordinal response variable τ_x is also called the threshold (difficulty) parameter and denotes the difficulty of the response category x to all items.

Partial Credit Model⁴ (PCM; Masters, 1982) and the Rating Scale Model⁵ (RSM; Andrich, 1978).

The RS-GRM model (with free α_j and constrained \mathcal{G}_{jx} , global logit link function) is the rating scale version of the GRM introduced by Muraki (1990), RS-GPCM (with free α_j and constrained \mathcal{G}_{jx} , local logit link function) are the rating scale versions of GPCM (Muraki, 1997); 1P-GRM⁶ and 1P-RS-GRM⁷ (Van der Ark, 2001) are the equally discriminating versions of GRM and RS-GRM, respectively (see Bartolucci et. al., 2016b, p. 127; Genge, 2016; Genge, 2017 for more details).

The influence of the covariates on the probabilities $\pi_s(\mathbf{z})$ is based on the logit function model. The resulting types of models belong to the class of concomitant variable latent class models (Dayton and Macready, 1988; Formann, 2007). The most general parameterization is based on reference-category logits:

$$\log \frac{\pi_s(\mathbf{z})}{\pi_1(\mathbf{z})} = \delta_{0s} + \mathbf{z}'\boldsymbol{\delta}_{hs}, \quad s = 2, \dots, u, \quad (3)$$

where: δ_{0s} is the intercept specific for class s and $\boldsymbol{\delta}_{hs}$ denotes the effect of covariate Z_h on the logit of $\pi_s(\mathbf{z})$ with respect to $\pi_1(\mathbf{z})$.

However, the interpretation of the multinomial parameterization may be difficult with heterogeneous data set (when a few latent classes are analyzed). Then, there is a set of parameters referred to the comparison between each latent class (apart from the first) and the first. This implies that, for each covariate, there are as many covariates parameters as latent classes minus 1. In particular, it may be difficult to understand if the ability level tends to increase or to decrease with each covariate in \mathbf{z} (especially for covariates with more than two categories).

Another, more parsimonious parameterization supposing the increasingly ordered ability levels ($\xi_1 < \dots < \xi_s$) based on global logits, is defined as:

$$\log \frac{\pi_s(\mathbf{z}) + \dots + \pi_u(\mathbf{z})}{\pi_1(\mathbf{z}) + \dots + \pi_{s-1}(\mathbf{z})} = \delta_{0s} + \mathbf{z}'\boldsymbol{\delta}_h, \quad s = 2, \dots, u, \quad (4)$$

⁴ Model with constrained α_j and free \mathcal{G}_{jx} – local logit link function.

⁵ Model with constrained α_j and constrained \mathcal{G}_{jx} – local logit link function.

⁶ Model with constrained α_j and free \mathcal{G}_{jx} – global logit link function.

⁷ Model with constrained α_j and constrained \mathcal{G}_{jx} – global logit link function.

The covariate parameters collected in δ_h are common to each latent class and are of simple interpretation. If one coefficient in δ_h is positive, this implies that the corresponding covariate in \mathbf{z} has a positive effect on the ability level and as the covariate increases, the ability level also increases. On the contrary, if this element is negative, then the corresponding covariate has a negative effect on the ability level (see Bartolucci et al., 2016b, p. 243).

As far as the estimation method for LC-IRT models with covariates is concerned, the authors applied (in the empirical part of this article) the discrete marginal maximum likelihood using `MultiLCIRT` package of R. Typically, the estimates are obtained through the EM algorithm (Bartolucci, 2007; Dempster et al., 1977). LC-IRT models with different parameterization are compared on the basis of log-likelihood ratio (LR) test as well as information criterion such as the Bayesian Information Criterion (BIC; Schwarz, 1978) and Akaike Information Criterion (AIC; Akaike, 1974).

4. DATA PRESENTATION

The analyses presented below are based on $n = 756$ currently working⁸ interviewers⁹ who participated in the International Social Survey Programme (ISSP)¹⁰ in 2015. The data collected on these interviewees included responses to questionnaire items about job satisfaction (the public data set, available at www.diagnoza.com, see also (Diagnoza Społeczna 2015)).

The computations and graphics in this paper were done mainly in `MultiLCIRT` (Bartolucci et al., 2016a) package of R. The authors considered the items concerning the different aspects of job satisfaction. The original questions concern the assessment of the main job:

X_1 (*HSW12_1*) – My job is secure,

X_2 (*HSW12_2*) – My income is high,

X_3 (*HSW12_3*) – My opportunities for advancement are high,

X_4 (*HSW12_4*) – My job is interesting,

X_5 (*HSW12_5*) – I can work independently,

⁸ ‘Currently working’ refers to both self-employed, employees, and include persons on leave if they are in an employment relationship.

⁹ The authors dropped records with at least one missing response and “can’t choose” category.

¹⁰ International Social Survey Programme is a continuing annual programme of cross-national collaboration on surveys covering topics important for social science research and nationally representative. The authors analyzed the section about the job characteristics and the social dimension of the survey conducted by the Institute for Social Studies (ISS), University of Warsaw.

X_6 (HSW12_6) – In my job I can help other people,

X_7 (HSW12_7) – My job is useful to society,

X_8 (HSW12_8) – In my job, I have personal contact with other people.

All of the analyzed items have five ordered response categories: strongly agree (1), agree (2), neither agree nor disagree (3), disagree (4), strongly disagree (5). However, in order to have a clearer interpretation of the results, the response categories were arranged in increasing order (from absolutely unsatisfied to absolutely satisfied). Table 1 shows the distribution of item responses among the five categories.

Table 1
Frequency distribution of each category (x) of the response variable X_j

Item/Category	$x = 1$	$x = 2$	$x = 3$	$x = 4$	$x = 5$
X_1	0.0503	0.2751	0.2685	0.3439	0.0622
X_2	0.1892	0.4921	0.2222	0.0741	0.0225
X_3	0.1825	0.4921	0.2275	0.0780	0.0198
X_4	0.0648	0.1627	0.3399	0.3690	0.0635
X_5	0.0807	0.3452	0.2130	0.3003	0.0608
X_6	0.0754	0.2566	0.2447	0.3677	0.0556
X_7	0.0317	0.1177	0.2619	0.4907	0.0979
X_8	0.0212	0.0847	0.1085	0.5966	0.1892

Source: own calculations in R.

Altogether, the responses are mainly concentrated in categories 2, 3 and 4, whereas category 5 which denotes the highest level of satisfaction, is selected less than 10% of the times for each item (with the exception of the last item, i.e. 18.92%).

In addition, the analysis considered the information provided by an individual level covariates:

Z_1 – age,

Z_2 – sex: 1=male, 2=female,

Z_3 – education: 1=pre-secondary education, 2=some unfinished secondary education, 3=secondary education or secondary professional education, 4=postgraduate education,

Z_4 – marital status: 1=living single, 2=living as a couple,

Z_5 – current financial situation: 1=very bad, 2=bad, 3=neither good nor bad, 4=good, 5=very good.

5. EMPIRICAL ANALYSIS

This section first deals with the problem of the model selection, regarding in particular the optimal number of latent classes and item parameterization (scale parameters). Then with the analysis of the ability distribution and identification of the influential covariates.

Concerning the issue of the number of the latent classes the study employed standard LC models that differ by the number of latent classes and, following McLachlan and Peel (2000), to choose the optimal s as the one corresponding to the smallest value of the BIC and AIC indices. The results of the selection procedure for the LC-IRT models without the covariates are reported in Table 2. On the basis of the adopted criteria the authors decided to choose four latent classes as the most suitable number for this data.

Table 2
The information criteria (BIC, AIC), log-likelihood (LL) values,
number of parameters (npar)

s	LL	npar	BIC	AIC
1	-8030.20	32	16272.50	16124.40
2	-7657.55	65	15745.92	15445.10
3	-7411.39	98	15472.33	15018.78
4	-7202.36	131	15273.00	14666.72
5	-7170.99	164	15428.99	14669.99
6	-7143.74	197	15593.20	14681.48
7	-7134.11	230	15792.67	14728.22
8	-7116.60	263	15976.38	14759.21

Source: own calculations in R.

In the second step of the model selection¹¹, i.e. the choice of the best logit link function, the authors compared the graded response type model and a partial credit type model¹², assuming $s = 4$, free item discriminating and difficulties parameters and a general multidimensional structure for the data (see Bartolucci et al., 2014).

¹¹ The authors followed the four consecutive ordered steps (Bacci et al. 2014; Bartolucci et al. 2014) to choose the optimal number of latent classes and item parameterization.

¹² The continuation ratio logit link function is not suitable in the context of the empirical study because the item response process does not consist of the sequence of successive steps (Bacci et al. 2014).

Table 3
Graded response and partial credit type models with $s = 4$

Logit	LL	npar	BIC	AIC
Global	-7486.968	45	15272.198	15063.936
Local	-7526.356	45	15350.974	15142.712

Source: own calculations in R.

As the global logit link has to be preferred to a local logit link function (the smaller BIC and AIC values, see Table 3) the authors fitted different types of the LC graded response type models with free and constraint discriminating index as well as free and constrained threshold difficulty parameters, for each item. This implies a comparison among four models (LC-GRM, LC-RS-GRM, LC-1P-GRM, LC-1P-RS-GRM models in Table 4). For the sake of completeness, the results for local logit link function are also included in Table 4 (see LC-GPCM, LC-RS-GPCM, LC-PCM, LC-RSM).

Table 4
Log-likelihood, AIC and BIC results for different polytomous LC-IRT models

Model	LL	AIC	BIC
LC-GRM	-7486.97	15063.94	15272.2
LC-RS-GRM	-7582.82	15213.64	15324.71
LC-1P-GRM	-7527.25	15130.49	15306.36
LC-1P-RS-GRM	-7612.77	15259.54	15338.22
LC-GPCM	-7526.36	15142.71	15350.97
LC-RS-GPCM	-7639.02	15326.04	15437.11
LC-PCM	-7570.83	15217.65	15393.52
LC-RSM	-7654.55	15343.1	15421.78

Source: own calculations in R.

Since the compared models are nested, the parameterization is selected on the basis of an LR test and BIC, AIC criteria as well. The results presented in Table 4 show that LC-GRM model has to be preferred among all of the LC-IRT models considered, that is the LC graded response model with free discriminant and difficulty parameters.

The estimates of the item parameters (discriminant and difficulty parameters) for the selected model are given in Table 5. The most difficult (the highest threshold parameters τ_{jx}^{13}) is the second and third item, which

¹³ τ_{jx} is known as the threshold or cutoff point between categories.

concern income and opportunities for advancement, as confirmed by the frequency distribution (previously reported in Table1). The easiest item is the eighth one (personal contact with other people). As far as the discrimination parameters (α_j) are concerned interesting job and personal contact with other people were considered to have the highest and the lowest discriminating power, respectively.

Table 5
The item parameter estimates for LC-GRM

Item	LC-GRM model				
	τ_1	τ_2	τ_3	τ_4	α_j
1	-3.368	-0.856	0.438	3.125	0.991
2	-1.294	0.675	1.972	3.376	1.468
3	-1.454	0.677	2.072	3.634	1.303
4	-1.914	-0.912	0.209	<i>1.953</i>	2.042
5	-2.363	-0.270	0.594	2.698	1.239
6	-2.258	-0.654	0.277	2.566	1.409
7	-3.863	-2.018	-0.441	2.604	0.982
8	<i>-5.410</i>	<i>-3.086</i>	<i>-1.902</i>	2.124	<i>0.755</i>

Source: own calculations in R.

The main part of this study aimed at understanding how the covariates affect the latent trait of the model, starting with the LC-GRM specification under the multinomial logit parameterization (see Eq. 3) and the distribution of the latent variable based on four ordered latent classes analysis. The estimated support points $\hat{\xi}_s$ and weights $\hat{\pi}_s(\mathbf{z})$ are given in Table 6.

Table 6
The estimated support points and prior probabilities
for LC-GRM model with covariates

Parameter	Cluster 1	Cluster 2	Cluster 3	Cluster 4
$\hat{\xi}_s$	0.413	2.800	4.178	6.468
$\hat{\pi}_s(\mathbf{z})$	0.041	0.515	0.406	0.038

Source: own calculations in R.

The authors observed that most of the subjects belongs to classes 2 (51.5%) and 3 (40.6%), characterized by an intermediate level of satisfaction. There are over 4% of subjects in class 1 and they tend to have the lowest level of satisfaction, whereas the remaining 3.8% of individuals belong to class 4, which is characterized by the highest level of job satisfaction.

Moreover, on the basis of the conditional probabilities analysis, the study confirmed that the probabilities of answering with a high response category (denoting a high level of satisfaction) increase from class 1 to 4, whereas the probabilities of answering with a low response category (denoting a low level of satisfaction) decrease from class 1 to class 4.

In further analysis the authors ran the test for significance of the covariates coefficients. The estimates of the significant covariates coefficients for the LC-GRM model based on the multinomial logit parametrization are given in Table 7. This parameterization implies that, for each covariate, there are as many covariate parameters as latent classes minus 1. As the first latent class (corresponding to the lowest level of job satisfaction) is the reference class, each covariate parameter denoted as $\hat{\delta}_{h2}$ concerns the effect of the corresponding covariate (*finance*, *education*, *sex*) on the logarithm of the probability of being in class 2 with respect to class 1. Then, each covariate parameters in Table 7 denoted as $\hat{\delta}_{h3}$ or $\hat{\delta}_{h4}$ compares the probability of being in classes 3 or 4 (corresponding to the upper-intermediate and the highest satisfaction levels) respectively, with that of class 1.

The authors observed that the significant covariates are *finance*, *education* and *sex*. Moreover, the effects of the education covariate on the logarithm of the probability of being in class 2 with respect to class 1 equals $\hat{\delta}_{h2} = 0.127$. Since the latent classes of respondents are ordered from that with the lowest to that with the highest level of satisfaction and three estimates (for $s=2$, $s=3$ and $s=4$) of the parameters for this covariate are positive (0.127, 0.721, 0.942), as the educational level increases, the level of job satisfaction also increases. One can also see in Figure 1 that the probability to belong to the first class (characterized by the lowest level of the latent trait) is the lowest for those with the highest educational level (postgraduate education) but is clearly higher for those with pre-secondary education. However, the prior probabilities to belong to the latent classes characterized by the upper-intermediate ($\hat{\xi}_3 = 4.178$) and the highest level of the latent trait ($\hat{\xi}_4 = 6.468$) is the highest for those with postgraduate education.

Similarly, all the estimates of the parameters for the *finance* covariate are also positive. Moreover, this covariate is the most significant. Respondents with the higher perceived financial situation are more satisfied with their job (characterized by the highest level of the latent trait, i.e. $\hat{\xi}_4 = 6.468$). Furthermore,

the analysis of the prior probabilities at varying levels of finance covariate (see Figure 2) leads to the similar conclusions. The prior probability to belong to classes 1 and 2 (corresponding to those with the lowest and intermediate satisfaction levels) are the highest and especially for those who perceive their current financial situation respectively as very bad or bad, whilst those who perceive their financial situation as very good are more prone to belong to class 4 (characterized by the highest level of the latent trait).

Table 7

The estimates of the covariates coefficients for LC-GRM model with $s=4$ based on the multinomial logit parameterization

Covariate	$\hat{\delta}_{h2}$	$\hat{\delta}_{h3}$	$\hat{\delta}_{h4}$
finance	1.495	2.921	4.161
education	0.127	0.721	0.942
sex	-0.413	-0.814	-1.102

Source: own calculations in R.

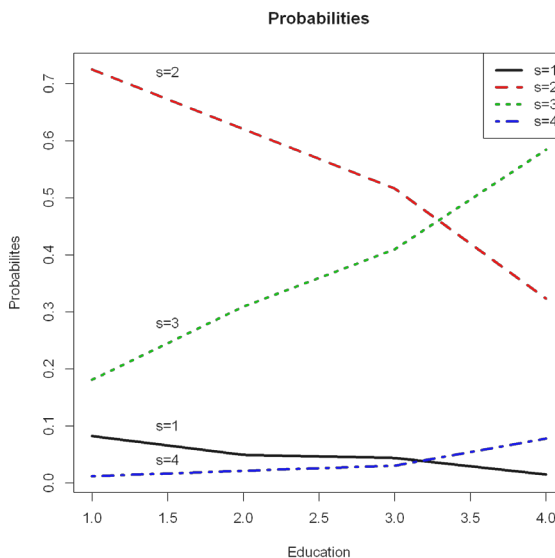


Fig. 1. Prior probabilities of latent class ($s=1, \dots, 4$) membership at varying levels of education (1=pre-secondary education, 2=some unfinished secondary education, 3=secondary education or secondary professional education, 4=postgraduate education)

Source: own calculations in R.

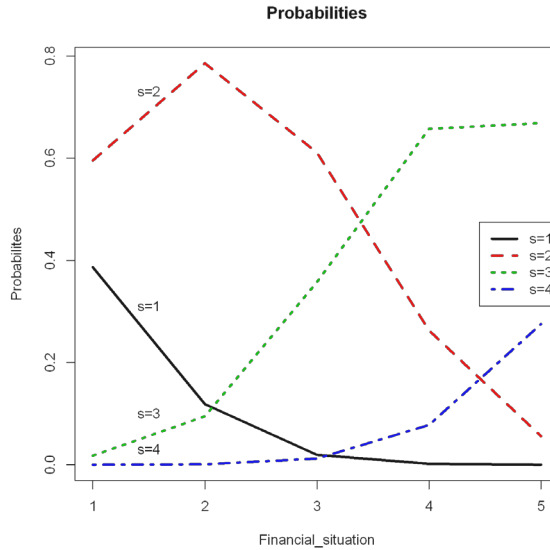


Fig. 2. Prior probabilities of latent class ($s = 1, \dots, 4$) membership at varying levels of financial situation (1=very bad, 2=bad, 3=neither good nor bad, 4=good, 5=very good)

Source: own calculations in R.

As far as sex covariate is concerned, one can observe the negative regression parameters which denotes the decreasing tendency of the job satisfaction levels for women compared to men. Similar findings for the socio-demographic features were obtained in the research of the job satisfaction (based mainly on the descriptive statistics) among all of the EU workers (European Commission 2015).

Finally, the study also considered the LC-GRM model with covariates based on the global logit parameterization (Eq. 4). The model based on global logit parameterization is more parsimonious with respect to the model based on multinomial logit parameterization (see Table 8) and the interpretation of the effect of covariates is easier (there is a single regression parameter of each covariate)¹⁴.

The estimates of the significant covariates coefficients for the LC-GRM model based on the global logit parameterization are given in Table 9. The significant covariates are again finance, education and sex. The finance and education covariates have a significant, positive effect on the job satisfaction

¹⁴ The interpretation of the conditional response probabilities and support points are similar and not presented because of limitations of space.

level, whereas one can observe the negative regression parameter for the sex covariate which can be interpreted as the negative effect on the level of the latent trait for women (women are less satisfied than men who were assumed as the reference category).

Table 8
LC-GRM models with covariates based on multinomial and global parametrization

Logit	LL	BIC	AIC
Multinomial	-7329.73	15056.75	14779.46
Global	-7331.57	14994.21	14763.14

Source: own calculations in R.

Table 9
The estimates of the covariates coefficients for the LC-GRM model with s=4 based on the global logit parameterization

Covariates	$\hat{\delta}_h$
finance	1.610
education	0.529
sex	-0.393

Source: own calculations in R.

Obviously, on the basis of the estimated model the probabilities of belonging to the latent classes ($\hat{\pi}_{si}, s = 1, 2, 3, 4$) for different patterns of covariates (see Table 10) may be assessed.

Table 10
The estimates of individual weights ($\hat{\pi}_{si}, s = 1, 2, 3, 4$) for different values of covariates, for the LC-GRM model with s=4 based on the multinomial and global logit parameterization

Covariate			Multinomial logit				Global logit			
Sex	Education	Finance	$\hat{\pi}_{1i}$	$\hat{\pi}_{2i}$	$\hat{\pi}_{3i}$	$\hat{\pi}_{4i}$	$\hat{\pi}_{1i}$	$\hat{\pi}_{2i}$	$\hat{\pi}_{3i}$	$\hat{\pi}_{4i}$
1	1	1	0.3473	0.6362	0.0165	0.0000	0.3814	0.6017	0.0165	0.0004
1	3	3	0.0131	0.5384	0.4328	0.0157	0.0143	0.5646	0.4052	0.0159
1	4	5	0.0000	0.0256	0.5842	0.3902	0.0003	0.0317	0.5665	0.4015
2	2	1	0.4613	0.5308	0.0079	0.0000	0.4862	0.5028	0.0108	0.0002
2	3	3	0.0233	0.6310	0.3363	0.0094	0.0211	0.6500	0.3182	0.0108
2	4	5	0.0001	0.0399	0.6963	0.2636	0.0005	0.0448	0.6360	0.3187

Source: own calculations in R.

The estimated $\hat{\pi}_s(\mathbf{z})$ allow for several types of prediction. For instance, men with the lowest level of education and the worst financial situation have a very small (1.65%) probability to be assigned to class 3 and are not able to achieve the highest level of satisfaction (to be assigned to class 4). The same probability increases to 43.28% and 1.57%, respectively for the same employee with secondary or secondary vocational education and neither a good nor bad financial situation. The probability to achieve the highest satisfaction level is the highest for men with postgraduate education and a very good financial situation.

Moreover, the probability to achieve the highest satisfaction level ranges from 0 to 39.02% for male employees and from 0 to 26.36% for female employees. The authors observed that the predicted values are slightly different for global logit parameterization but generally lead to the same conclusion.

CONCLUSIONS

Most of the job satisfaction and questionnaires analyses are based on homogeneity and normality of the latent trait assumptions (explicitly introduced). These assumptions may be restrictive to fulfill especially in socio-economic data analyses.

The authors applied the extended latent class Item Response Theory models for ordinal polytomously-scored items allowing for covariates that influence the weights of the latent classes, then compared the results of different kinds of LC-IRT models and confirmed that the analyzed item responses can be explained by four latent class graded response-type model. The study analyzed the item parameters of the selected model and showed the influence of covariates such as sex, education and finance situation based both on multinomial and global logit parameterization.

This article discusses which group of Polish workers is most satisfied with their job and which class of employees should improve the level of job satisfaction and increase employee engagement as well. The results may help policymakers tailor their employment policies as well as to create and deliver services focused on special socio-economic groups of Polish society. Furthermore, most organizations strive for employee satisfaction. If employees feel better about their jobs, they are less likely to quit or even seek employment in a foreign country. Therefore the detailed analysis of the estimated parameters may be also of benefit in maintaining productivity,

profitability as well as keeping their workforce engaged (especially some specific segments of the workforce) and satisfied with their jobs.

In *future research*, it would be worthwhile to extend the questionnaire items analysis using LC-IRT models to the data at different points of time.

REFERENCES

- Akaike, H., *A new look at statistical model identification*, "IEEE Transactions on Automatic Control", 19, pp. 716-723, 1974.
- Andrich, D., *A rating formulation for ordered response categories*, "Psychometrika", 43, pp. 561-573, 1978.
- Babakus, E., Yavas, U., Karatepe, O. M., Avci, T., *The Effect of Management Commitment to Service Quality on Employees' Affective and Performance Outcomes*, "Journal of The Academy of Marketing Science", 31 (3), pp. 272-286, doi: 10.1177/0092070303253525, 2003.
- Bacci, S., Bartolucci, F., Gnaldi, M., *A class of Multidimensional Latent Class IRT models for ordinal polytomous item responses*, "Communication in Statistics – Theory and Methods", 43, pp. 787-800, 2014.
- Bartolucci, F., *A class of multidimensional IRT models for testing unidimensionality and clustering items*, "Psychometrika", 72, pp. 141-157, 2007.
- Bartolucci, F., Bacci, S., Gnaldi, M., *MultiLCIRT: An R package for multidimensional latent class item response models*, "Computational Statistics and Data Analysis", 71, pp. 971-985, 2014.
- Bartolucci, F., Baschina, A., Giovanni, S. F., Demidova B. O., Signorelli M., *Determinants of Job Satisfaction in Young Russian Workers*, Discussion Papers 7, 2015, CRISEI, University of Naples "Parthenope", Italy, 2015.
- Bartolucci, F., Bacci, S., Gnaldi, M., *MultiLCIRT: Multidimensional latent class Item Response Theory models. R package version 2.10*, URL <http://CRAN.R-project.org/package=MultiLCIRT>, 2016a.
- Bartolucci, F., Bacci, S., Gnaldi, M., *Statistical Analysis of Questionnaires. A Unified Approach Based on R and Stata*. Boca Raton: Chapman & Hall/CRC, 2016b.
- Birnbaum, A., *Some latent trait models and their use in inferring an examinee's ability* [in:] Lord, M. F., Novick, R. M. (eds.), *Statistical Theories of Mental Test Scores*, pp. 395-479. MA: Addison-Wesley, Reading, 1968.
- Brayer, A., Foley, M., Doroszkiewicz, H., Jamiokowski, J., Marcinowicz, L., *Job satisfaction among masters in nursing in Central and East Poland: a preliminary study*, "Family Medicine & Primary Care Review", 1, pp. 7-11, 2017.
- Brayfield, A. H., Rothe H. F., *An index of job satisfaction*, "Journal of Applied Psychology", 35, pp. 307-311. doi: 10.1037/h0055617, 1951.
- Bruck, C.S., Allen, T.D., Spector, P.E., *The relation between work-family conflict and job satisfaction: A finer-grained analysis*, "Journal of Vocational Behavior", 60, 336-353. doi: 10.1006/jvbe.2001.1836, 2002.

- Bruno, G. S. F., Caroleo F. E., Dessy O., *Stepping stones versus dead end jobs: Exits from temporary contracts in Italy after the 2003 reform*, "Rivista Internazionale per le Scienze Sociali", 1, pp. 31-62, 2013.
- Butlewski, M., Tytyk, E., *The method of matching ergonomic nonpowered hand tools to maintenance tasks for the handicapped* [in:] Karwowski, W., Salvendy, G. (eds), (CD ROM) 2nd International Conference on Applied Human Factors and Ergonomics, Las Vegas, Nevada, USA. Conference Proceedings, 2008.
- Eliyana, A., Yusuf, R. M., Prabowo, K., *The Influence of Employee's Job Satisfaction Factors on Organizational Commitment*, "American Journal of Economics", Special Issue, pp. 141–144. doi: 10.5923/j.economics.20120001.31, 2012.
- European Commission (2015). *Eurostat – Quality of Life in Europe – Facts and Views – Employment*. http://ec.europa.eu/eurostat/statistics-explained/index.php/Quality_of_life_in_Europe_-_facts_and_views_-_employment#Job_satisfaction. Accessed September 16, 2017.
- European Commission (2016). *Eurostat – Labour force survey: Temporary employees as percentage of the total number of employees, by sex and age (%)*. http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=lfsa_etpga&lang=en. Accessed September 13, 2017.
- European Commission (2017). *Eurostat: Average rating of satisfaction by domain, sex, age and educational attainment level*. http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_pw01&lang=en. Accessed September 13, 2017.
- Dayton, C.M., Macready, G.B., *Concomitant-variable latent-class models*, "Journal of the American Statistical Association", 83, pp. 173-178, 1988.
- Dempster, A.P., Laird, N.P., Rubin D.B., *Maximum likelihood for incomplete data via the EM algorithm (with discussion)*, "Journal of the Royal Statistical Society", 39, ser. B, pp. 1-38, 1977.
- Diagnoza społeczna 2015. Warunki i jakość życia Polaków (raport) [Social Diagnosis. Objective and Subjective Quality of Life in Poland]*, Czapiński J., Panek T. (eds.), Warszawa, Rada Monitoringu Społecznego (5.01.2017). doi: www.diagnoza.com/index-en.html
- Formann, A.K., *(Almost) equivalence between conditional and mixture maximum likelihood estimates for some models of the Rasch type* [in:] Davier, von M., Carstensen, C. (eds.), *Multivariate and Mixture Distribution Rasch Models*, pp. 177-189. Springer-Verlag, New York 2007.
- Gaszynska, E., Stankiewicz-Rudnicki, M., Szatko, F., Wieczorek, A., Gaszynski, T., *Life satisfaction and work-related satisfaction among anesthesiologists in Poland*, "Scientific World Journal", 2014, doi: 10.1155/2014/601865, 2014.
- Genge, E., *Teoria reakcji na pozycję w podejściu modelowym w taksonomii [Item response theory in model-based clustering]*, „Ekonometria”, Wrocław University of Economics and Business, 1 (51), pp. 9-19, 2016. doi:10.15611/ekt.2017.4.04. 10.15611/ekt.2016.1.01.
- Genge, E., *IRT and LC-IRT Models for Items with Ordinal Polytomous Responses*, „Ekonometria”, Wrocław University of Economics and Business, 4 (58), pp. 62-76, 2017. doi:10.15611/ekt.2017.4.04.

- Grobelna, A., Sidorkiewicz, M., Tokarz-Kocik, A., *Job satisfaction among hotel employees: analyzing selected antecedents and job outcomes. A case study from Poland*, "Argumenta Oeconomica", 2 (37), pp. 281-310, 2016.
- Judge, T.A., Erez, A., Bono, J.E., Thoresen, C.J., *The core self-evaluations scale: Development of a measure*, "Personnel Psychology Journal", 56, pp. 303-331. doi: 10.1111/j.1744-6570.2003.tb00152.x.2003.
- Judge, T. A., Watanabe, S., *Another look at the job satisfaction-life satisfaction relationship*, "Journal of Applied Psychology", 78(6), 939-948, 1993.
- Kowal, J., Roztocki, N., *Job Satisfaction of IT Professionals in Poland: Does Business Competence Matter?*, "Journal of Business Economics and Management", 16 (5), pp. 995-1012, 2015.
- Lam, T., Zhang H., Baum T., *An Investigation of Employees' Job Satisfaction: The Case of Hotels in Hong Kong*, "Tourism Management", 22(2), 157-165. DOI: 10.1016/S0261-5177(00)00039-X.2001.
- Lambert, E., Hogan, N., *Creating a Positive Workplace Experience: The Issue of Support from Supervisors and Management in Shaping Job Stress, Job Satisfaction, and Organizational Commitment of Private Correctional Staff*, "Journal of Applied Security Research", 4(4), pp. 462-482. DOI: 10.1080/19361610903176286.2009.
- Lévy-Garboua, L., Montmarquette, C., Simonnet, V., *Job satisfaction and quits*, "Labour Economics", 14(2), pp. 251-268, 2007.
- Linz, S. J., *Job satisfaction among Russian workers*, "International Journal of Manpower", 24, pp. 626-652, 2003.
- Linz, S. J., Semykina A., *What makes workers happy? Anticipated rewards and job satisfaction*, "Industrial Relations", 51, pp. 811-844, 2012.
- Locke, E.A., *The Nature and Causes of Job Satisfaction* [in:] Dunette, D. M. (ed.), *Handbook of industrial and organizational psychology*, pp. 1295-1349. Rand McNally, Chicago, 1976.
- McLachlan, G., Peel, D., *Finite Mixture Models*. John Wiley & Sons, New York, 2000.
- Masters, G., *A Rasch model for partial credit scoring*, "Psychometrika", 47, pp. 149-174, 1982.
- Muraki, E., *Fitting a polytomous item response model to Likert-type data*, "Applied Psychological Measurement", 14, pp. 59-71, 1990.
- Muraki, E., *A generalized partial credit model: application of an EM algorithm*, "Applied Psychological Measurement", 16, pp. 159-176, 1992.
- Muraki, E., *A generalized partial credit model* [in:] Van der Linden, W., Hambleton, K. R. (eds.), *Handbook of modern item response theory*, pp. 153-164. Springer, New York, 1997.
- Nering, M. L., Ostini, R., *Handbook of polytomous item response theory models*. Taylor and Francis, New York, 2010.
- Rain, J. S., *A current look at the job satisfaction/life satisfaction relationship: Review and future considerations*, "Human Relations", 44(3), pp. 287-307, 1991.
- Rasch, G., *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research, Copenhagen 1960.
- Schwarz, G., *Estimating the dimension of a model*, "Annals of Statistics", 6, pp. 461-464, 1978.

- Samejima, F., *Estimation of latent ability using a response pattern of graded scores*, "Psychometrika Monograph", 17, 1969.
- Smith, P.C., Kendall, L.M., Hulin C.L., *The measurement of satisfaction in work and retirement*. Rand McNally, Chicago 1969.
- Spector, P. E., *Job Satisfaction: Application, Assessment, Causes and Consequences*. Sage Publications, Thousand Oaks, CA, 1997.
- Weiss, D.J., Dawis, R.V., England, G.W., Lofquist L.H., *Manual for the Minnesota Satisfaction Questionnaire Minnesota*, "Studies in Vocational Rehabilitation", 22, 1967.
- Wilczyńska, A., Batorski D., Torrent-Sallens J., *Employment Flexibility and Job Security as Determinants of Job Satisfaction: The Case of Polish Knowledge Workers*, Social Indicators Research 126(2), pp. 633-656, doi: 10.1007/s11205-015-0909-6.2016.
- Van Der Ark, L.A., *Relationships and properties of polytomous Item Response Theory models*, "Applied Psychological Measurement", 25, pp. 273-282, 2001.
- Van der Linden, W., Hambleton, R.K., *Handbook of modern item response theory*. Springer, 1997.

Received: November 2017, revised: July 2018

Acknowledgment: *I acknowledge the research grant (SONATA 12, UMO-2016/23/D/HS4/00989, "Latent variable models in the identification of homogenous structures in socio-economic longitudinal data") of the National Science Centre, Poland.*