



Paula Gołębiowska

Mechanizm autoselekcji do próby w badaniu
wolnych miejsc pracy w Polsce

Self-selection mechanism on job vacancy
market in Poland

Bachelor's thesis

Thesis Supervisor: dr Maciej Beręsewicz

Date of submission:

Supervisor's signature

Field of study: Informatyka i ekonometria

Specialisation: Analityka gospodarcza

Poznań 2018

Contents

Introduction	2
1 Self-selection mechanism on job vacancy market	3
1.1 New data sources for statistics	3
1.2 Selection error and representativeness in statistics	4
1.2.1 Taxonomy of survey errors	4
1.2.2 Representativeness	4
1.3 Data sources on the job vacancies in Poland	5
1.3.1 Basic definitions used in official statistics	5
1.3.2 The Job Demand Survey	6
1.3.3 The Human Balance Capital – survey of employers	8
2 Latent class analysis	12
2.1 Introduction to Classic Latent Class	12
2.2 Classic Latent Class Models	14
2.2.1 The model	14
2.2.2 Latent class regression model	16
2.3 Concomitant-variable latent class model	17
2.3.1 Choice of the model and assessment of the goodness of fit	18
2.4 The polCA package	19
3 Empirical results of quantification of self-selection mechanism	21
3.1 Exploratory Analysis	21
3.1.1 Questionnaire	21
3.1.2 Description of manifest variables	22
3.2 Description of concomitant variables	23

3.3	Description of the model	26
3.3.1	Models	26
3.3.2	Self-selection Analysis	29
	Conclusions	32
	List of Tables	35
	List of Figures	36
	List of R codes	37
A	The survey of employers' questionnaire	38

Introduction

Motivation

Motivation for the thesis is twofold. First, we observe an exponential growth in creation and collection of data, mainly due to automated data collection (i.e. big data). Second, we observe a decrease in response rates in surveys as well as need for timely and detailed statistics. Therefore there is a need to look for alternative sources of information which could be used to supplement, extend or even replace currently conducted surveys.

An notable example is job vacancy market which is studied by Statistics Poland by means of the Demand for Labor survey – a probabilistic sample with response rate around 60%. On the other hand, employers look for employees via variety of channels including district labor offices (pol. powiatowe urzędy pracy) or the Internet – a non-probabilistic samples that are created by means of self-selection mechanism.

Thus, before data created by the employers can be used for statistics, it is crucial to identify the selection mechanism, which may be used to correct bias.

Objective of the study

The aim of the thesis is to (1) identify what are the most popular channels to post job offers and if these channels cluster and (2) describe the self-selection mechanism by the characteristics of companies that advertise job offers in specific channels.

Research methodology

In order to describe selection mechanism we used the survey of employers conducted during the Human Balance Capital 2010-2014 survey. For the initial analysis of the data, χ^2 tests and other descriptive statistics were applied and results are presented in graphs or in tables.

To verify to what extent channels cluster we used Latent Class Models implemented in the R language (R Core Team, 2017) in the polCA package (Linzer & Lewis, 2011). We used two types of models – with and without concomitant variables. Selection of models were based on information criteria such as Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC).

Structure of the thesis

The thesis consists of three chapters. The first focuses on representativeness, errors in surveys and data sources. The source of data is the Human Balance Capital (HBC) survey conducted by The Polish Agency for Enterprise Development (PARP) and Kantar MillwardBrown. The second chapter describes latent class analysis. Explains what the latent class model is, what it is used for and how to use it. The last chapter is devoted to empirical results and its interpretation. The thesis ends with conclusions.

Chapter 1

Self-selection mechanism on job vacancy market

1.1 New data sources for statistics

So far, statistics rely on data sources such as sample surveys, censuses or more recently registers (Gołata, 2018). However, due to increasing non-response problem in sample surveys and their high cost, statisticians need to find new alternative sources of collecting information such as big data and the Internet. The advantage of new data sources is that *they provide simple, cheap, and fast access to a large group of potential respondents (members of target population) provided they have access to the Internet* (Bethlehem, 2010).

One of the main problem of using new data sources is that their non-sampling errors and representativeness or bias have not been assessed yet. Moreover, they were not created for statistical purposed, which means that they need to be transformed to suit them (Beręsewicz, 2017). New data sources can be, for instance: data created from social networks, social media, blogs, pictures, comments etc.; social business systems such as transactions, making ordes, registering a customer, banking records; Internet of Thins (machine generated data) GPS, mobile phones, street monitoring (Beręsewicz, 2016, p. 13-14).

In this context Beręsewicz (2017) defines Internet data source is a self-selected (non-probabilistic) sample that is created through the Internet and maintained by entities external to National Statistical Institutes (NSIs) and administrative regulations.

This means that individuals, entities or companies make choice themselves whether to participate in collecting data process and provide information or not. Respondents can simply access the questionnaire / form via Internet and fill the questionnaire / form if they decide to. The researcher can not decide about the selection process. This is called self-selected mechanism (Bethlehem, 2010).

1.2 Selection error and representativeness in statistics

1.2.1 Taxonomy of survey errors

Bethlehem (2010) distinguish taxonomy of survey errors and provides examples based on web surveys. The proposed classification is generic, which means the taxonomy also apply to new data sources. Description of errors are presented in Table 1.1

Table 1.1. Taxonomy of survey errors

No.	Error	Description
1. Sampling errors		
1.1	estimation error	sample is based on a random selection process
1.2	specification error	true selection probabilities differ from the selection probabilities; specified in the sampling design
2. Non-sampling errors		
2.1 Observation errors		
2.1.1	measurement error	respondent does not give the true answer
2.1.2	over-coverage error	elements are included in the survey and that do not belong to the target population
2.1.3	processing error	made during data processing
2.2 Non-observation errors		
2.2.1	under-coverage error	elements of the target population do not have a corresponding entry in the sampling frame
2.2.2	non-response error	elements selected in the sample do not provide the required information

Source: Bethlehem (2010, p. 164).

1.2.2 Representativeness

There is no straightforward definition of representativeness. However, there is a list of denotations used in statistical and non-statistical literature:

1. general, unjustified acclaim for the data,
2. absence of selective forces,
3. mirror or miniature of the population,
4. typical or ideal case(s),
5. coverage of the population,
6. a vague term to be made precise,
7. representative sampling as a specific sampling method,
8. representative sampling as permitting good estimation,
9. representative sampling as good enough for a particular purpose.

According to the absence of selective sources denotation, it is assumed that *that a probabilistic sample has been drawn, but the final outcome is ultimately determined by the selection mechanism. Information about sampled units and their participation is required in order to assess the representativeness of the response* (Beręsewicz, 2017).

Mirror (or miniature) of the population is understood as such a selection of a sample that is as close and remarkable as possible to the population. All elements of the population have known probabilities of getting into the sample (Beręsewicz, 2017).

The larger the sample size, the better. However, the sample size itself does not guarantee that it is representative. It depends on the method of sample selection.

1.3 Data sources on the job vacancies in Poland

1.3.1 Basic definitions used in official statistics

The Statistics Poland defines the demand for labor as the number of vacancies offered by economy in specific social and economic conditions (Statistics Poland, 2014). Demand for labor is understood as:

1. occupied jobs, determined by the actual number of the employed,
2. vacancies.

The definition of vacancies, that we can find in official statistics in Poland, states that vacancies are positions or jobs unoccupied due to labor turnover or newly created ones that simultaneously meet the following three conditions:

1. positions and jobs were actually unoccupied on the survey day,
2. the employer made efforts to find people willing to take up the job,
3. if adequate candidates were found to occupy the vacancies, the employer would readily take them in.

On the other hand in statistical system we can find another definition — job offer. It is defined as declaration by the employer to the district labor office:

1. at least one vacant place of employment or other paid work,
2. place of occupational activation, adopted to implementation.

It may seem that definitions of vacancy and job offer are the same, however there is a significant difference. Job offers must be registered in the district labor office, whereas vacancies must meet mentioned above conditions.

1.3.2 The Job Demand Survey

1.3.2.1 The idea of the survey

The following description has been prepared based on the report (Statistics Poland, 2014). Statistics Poland conducts research on the demand for labor in Poland since 1995. Although its form has been changing several times during past years. Finally, the survey covers entities of all kinds of economic activity and with the number of 1 person employed or more. Research is conducted as a sample survey in accordance with the requirements of the Eurostat regarding the unification of the study in terms of content and methodology in all European Union countries.

The objective of the survey is to obtain information on satisfied and unsatisfied demand i.e. the employed, vacancies, and the newly created and liquidated jobs.

1.3.2.2 Sampling scheme of the survey

The survey on demand for labor in Poland is carried out as a probability sample survey. Due to applied method results are being generalized over the general population and can not be presented below the level of the voivodeship. The sample is selected separately for units employing more than 9 persons and units employing up to 9 persons. This results from separate study objectives for those parts of population.

- **Over 9 persons employed** – Objective of the survey is obtaining information about particular kinds of activity (by REGON sections) by voivodeship, 304 subpopulations were specified in this section. Sampling scheme: Assumed size of each population should be around 50,000. Allocation of the sample survey is conducted in a way to obtain approximately the same level of precision. Within each population the units are sorted in descending order by the number of employees. The largest units are included in the survey without a sampling. By using numerical optimization methods, the sample with the assumed size is allocated between 304 subpopulations. In each population is determined a threshold of number of the employed. The units above threshold are included in the survey without sampling. The remaining units are randomly sampled using a proportional random sampling method. The remaining parts of each subpopulation is divided into uniform and previously determined size strata with the use of the proportional method.
- **Units employing up to 9 persons** – Objective of the survey is obtaining precise information for particular kinds of activity (19 NACE sections). Sampling scheme: Therefore, the allocation of the assumed sample size (50,000 units) is made between the particular sections of the NACE in such a way that the expected precision for these sections is approximately the same. Inside sections, units are stratified by voivodships, then the sample is selected by using the method to the proportional sampling scheme.

1.3.3 The Human Balance Capital – survey of employers

1.3.3.1 The idea of the survey

The following description has been prepared based on the report (Antosz, 2014). The Human Balance Capital survey was conducted by Polish Agency for Enterprise Development (PARP) (pol. Polska Agencja Rozwoju Przedsiębiorczości) in cooperation with Jagiellonian University. The objective of the survey among employers was to determine the demand for competence in the Polish economy. The survey covered employers, i.e. the current labor market operators, who at the time of fieldwork employed at least one employee. The study was conducted in 2010-2014, using the Computer-Assisted Telephone Interviewing (CATI) method and supplementary Computer-Assisted Personal Interview (CAPI) method and Computer-Assisted Web Interview (CAWI) on a assumed sample of 16 000 employers each year.

Subject of the survey, which is human capital, is concentrated on chosen kinds of entities. As the result following entities were excluded from the study population:

1. Local entities
2. Entities from department number 94 activities of membership organizations by NACE sections in 2007
3. Entities from sections
 - section A - agriculture, hunting and forestry,
 - section O - public administration and defense, compulsory social security,
 - section T - activities of households and employers; undifferentiated goods- and services- producing activities of households for own use,
 - section U - activities of extraterritorial organizations and bodies.
4. Entities classified into one of these specific legal forms:
 - code 48 - foundations,
 - code 50 - Catholic Church,
 - code 51 - other churches and religious associations,
 - code 55 - associations not entered to the legal register of companies,

- code 60 - social organizations not mentioned separately, not entered to the legal register of companies
- code 70 - political parties,
- code 72 - trade unions,
- code 73 - organizations of employers,
- code 76 - economic local and professional governments not entered to the legal register of companies,
- code 85 - condominiums,
- code 90 - groups of agricultural producers.

Decision about exception of listed entities was determined by the idea of the study – connecting obtained results with purposes on spending funds dedicated on human capital development. Publics administration staffing needs are analyzed separately and independently from described survey. In addition, complete research of excepted companies and institutions would not be possible due to size of assumed sample.

1.3.3.2 Sampling scheme

List of units used for sampling came from National Business Registry Number (NACE) base verified by other available CSO information about those units. Using NACE database as a data source, allowed selecting sample consistent with the target population assumptions which provides proper quality of results. Sampling frame is weighted with error caused by data out dating, however error scale does not affect results precision. The research entity, disclosed and improved occurred errors to minimize necessity of survey modification.

Sample selection scheme and sample size is described by using data from 2013 report. The sample consists of units from panel sample and units randomly selected again which are cross-sectional sample.

Panel Sample Panel base consists of 9 990 units which is pro-founding sample of units hiring at least 100 employers. The sample consists of not only units of the individuals with whom the study was conducted in previous editions, but also those with which the

interview was not performed, and those who were established during the time interval between the study editions.

Verification removed 476 units:

- winded up/ liquidated or not existing during research;
- duplicated records (verification based on REGON database).

Furthermore, in order to update information about major enterprises (especially over 250 employees) researcher added more entities complying with requirements based on available information.

The effective panel sample (9 514 units) was the largest enterprises in terms of the number of employees employed:

- all units hiring 1000 an more employees (673);
- all units hiring from 250 to 999 employees (3 224);
- randomly selected units hiring from 100 to 249 employees (5 167).

Cross-sectional Sample Cross-sectional randomly selected sample with substitute sample included consists of 52 832 units. Sample was selected by using disproportional stratified sampling method. Stratas were voivodeships and number o employees.

Verification removed 9 420 units: - operating / running sole proprietorship;

- winded up/ liquidated or not existing during research;
- duplicated records (verification based on NBRN database).

The effective cross-sectional sample consists of 43412 unites including:

- units hiring from 2 to 9 employees (23 466);
- units hiring from 10 to 49 employees (17 434);
- units hiring from 50 to 99 employees (2 512).

1.3.3.3 Weighting scheme

Weights were implemented due to compensating unequal probabilities of including units into population. Therefore it results from applied sampling method and uneven return rate in stratas. The applied weights were calculated providing input into the sample which is a combination of sampling strata (voivodeship) and class of employees and 6 REGON sections, adequate to their participation in the sample frame. Two types of weights were calculated: population weight and normalized weigh. First one enables population size estimation in analyzing process, the other one brings the total number to population size.

For the purpose of estimating the number of demanded employees, it was assumed that cases with an extreme number of declared wanted employees will have population weights set at level 1. As the extremity criterion, the so-called upper limit was adopted. “Tukey’s hinge”, i.e. a value equal to the upper quartile, increased by twice the quadrant range (this is a well-known way of determining extreme values in statistics, used, among others, to prepare diagnostic box plots or stem-leaf charts); the top “Tukey hinge” was calculated separately for each layer of enterprises in terms of the number of employees.

The weights obtained in this way are characterized by a high variance in the global estimation at the level of the whole country. The variance of weights drops strongly in the case of analysis at the level of voivodships and in terms of the size of enterprises. Within these categories, the lower sample size is compensated to some extent by a smaller loss of precision due to variance of weights. In the case of national analysis, this mechanism works in the opposite direction.

In this way, the collected and presented data constitute a base that is pending for further research, for which latent class analysis was used.

Chapter 2

Latent class analysis

2.1 Introduction to Classic Latent Class

A latent variable is a variable that cannot be observed or measured directly, they are "hidden", however they have impact on manifest variables (MacCallum & Austin, 2000). For instance, intelligence, usability or human satisfaction. Whereas, manifest variables, that can be also called as observed variables, are possible to be measured or observed directly, for instance: height, gender, age. Thus, latent variables can be measured by using observed variables (Sauro, 2016). For this purpose, following methods can be used: factor analysis, cluster analysis and latent class analysis.

Methods that are used for the analysis of latent variables can be divided due to the character of the manifest variables and latent variables (Wołodźko, 2015, p. 106-123) which is shown in table 2.1.

Table 2.1. Methods for the analysis of latent variables

Manifest Variables	Latent continuous variables	Latent discrete variables
Continuous variables	Factor analysis	Latent profile analysis
Discrete variables	Latent trait analysis	Latent class analysis

Source: Wołodźko (2015, p. 107)

Latent class models (LCM), also known as latent class analysis (LCA) belong to latent variable models (Bąk, 2011, p. 204-222). In these models the latent variable is discrete. These models are an example of model-based clustering and use the idea of mixture of distributions known in multivariate statistics.

In contrast to classic cluster analysis methods using measures of distance as a basis

for classifying objects, in the model approach the parameters of the model are estimated and the probability of assigning objects to classes is calculated. An important advantage of the model approach in cluster analysis is the possibility to include in the description of classified objects variables measured on any scale (both metric and non-metric) (Bąk, 2011, p. 204-222).

Latent class analysis is a relatively new technique of multivariate analysis. The essence of modeling latent classes is the study of relations between the categories of nominal and ordinal variables (Brzezińska, 2015, p. 43-50). It uses the data contained in the contingency table (Genge, 2013). This method was introduced by Lazarsfeld (1950). In the following years, Goodman (1970, p. 226-256) made the LCA method possible for practical application by using the method of the maximum likelihood method for estimating the parameters of the model, whereas Haberman (1979) showed the connection between the LCA with the logarithmic-linear analysis for tables with missing data. This method is still very popular and is being developed among other scholars like Hagenaars and McCutcheon (2002), Vermunt (2010), Linzer and Lewis (2011), Genge (2013).

Latent class models have applications in many different fields, such as political science, economics, sociology, psychometry, psychiatry, educational and market research, genetics, astronomy, epidemiology and many more. Applications of LCA are similar to the usage of classification methods (cluster analysis, taxonomy). In economics they involve including research related to consumer segmentation, spatial analysis, comparative analysis and regional analysis (Bąk, 2011, p. 204-222).

Nevertheless, and despite the present availability of user-friendly-software with which latent class models can be easily and routinely applied, practicing social and behavioral (Wołodźko, 2015, p. 106-123) researches do not always consider latent class analysis a serious alternative for better-known techniques, such as factor analysis or linear structural equation modeling, even where it would be more appropriate means to address their questions (Hagenaars & McCutcheon, 2002). It is difficult to say what is the reason for the small - though growing - popularity of these methods. Their limited knowledge may contribute to this. However, in the literature, it is emphasized that latent variables do not provide any additional information over those that are contained in observable variables (Everitt & Dunn, 2001, p. 305).

2.2 Classic Latent Class Models

2.2.1 The model

It is assumed that in the sample there is a finite number of observation groups with similar properties (for example, companies with similar channels' preferences). On the other hand, there are significant differences between groups. These groups are not known a priori (they are discrete "latent"), because neither the membership of particular observations (companies) on specific segments nor the number of groups is unknown. Estimation of the latent class model bases on the estimation of the number and size of individual segments (Bağ, 2011, p. 204-222). The general form of the latent class model represents a mixture of multivariate multinomial distributions, shown in equation (2.1)

$$f(\mathbf{x}_i|\Theta) = \sum_{s=1}^u \pi_s f_s(\mathbf{x}_i|\Theta_s), \quad (2.1)$$

where: $f(\mathbf{x}_i|\Theta)$ is observation distribution function (for example companies preferences), u is number of latent classes, π_s is a priori probability that observation belongs to P_s class, $\sum_{s=1}^u \pi_s$ is unconditional probabilities of belongings to particular classes, distribution, $f(\mathbf{x}_i|\Theta_s)$ is conditional distribution function, $\Theta = (\pi_s, \Theta_s)$ is all the unknown model parameters; Θ_s is a vector of unknown parameters in s -th class (for example μ_s and σ_s for normal distribution).

For the π_s and Θ_s estimators a posterior probability of belonging of objects to individual classes can be calculated using the Bayes formula, shown in equation (2.2).

$$P(s|\mathbf{x}_i, \Theta) = \frac{\hat{\pi}_s f(\mathbf{x}_i, \hat{\Theta}_s)}{\sum_{q=1}^u \hat{\pi}_q f(\mathbf{x}_i, \hat{\Theta}_q)}. \quad (2.2)$$

Parameters π_s and Θ_s are estimated with maximum likelihood method based on a model (2.1). Maximum likelihood function for a sample consisting of u companies is described by following formula (form of the function is dependent of preferences distribution) (2.3)

$$L(\mathbf{x}_i|\Theta) = \prod_{i=1}^n f(\mathbf{x}_i|\Theta). \quad (2.3)$$

Maximum likelihood method which uses optimization algorithms (for example Newton-Raphson, Expectation–maximization) is being used to choosing function to data (Bağ, 2011, p. 204-222).

Latent class models have specific features which are relevant from their usage in measurement of self-selection mechanism on job vacancy market, which are:

- manifest variables are statistically independent,
- they allow identifying segments based on manifest variables or predictor variables,
- contain one categorical latent variable (number of categories equals number of segments),
- classifying objects to segments is estimated based on the probability model of belonging,
- manifest variables can be measured on different scales,
- concomitant variables and predictor variables can be included in the model.

There are following types of variables in latent class models:

- latent variables which can be measured in nominal or order scale, the model must contain at least one latent variable;
- manifest variables or dependent variables which can be measured on different scales, the model must contain at least one of these variables,
- concomitant variables (covariates) and predictor variables, these variables are not required in the model.

Latent class models, in contrast to cluster analysis methods, represent model approach, because classification of observations to classes are based on probability. Possibility of measuring manifest variables on different scales is one of the most crucial advantages of applying these methods (Bağ, 2011, page 204-222). The model can also be extended with covariates and predictor variables (which allows segmentation using information about geographic, demographic, socio-economic, psychological respondents).

2.2.2 Latent class regression model

Latent class model in its general form shows (2.1), however among latent classes regression models, two variants can be specified as in following equations: (2.4) and (2.5):

1. Models with predictor variables:

$$f(\mathbf{x}_i|\mathbf{y}_i) = \sum_{s=1}^u \pi_s f(\mathbf{y}_i|\pi_s, \mathbf{x}_i), \quad (2.4)$$

where: \mathbf{x}_i is observations vector, \mathbf{y}_i are predictor variables having an impact on \mathbf{x}_i , π_s is probability of membership to $s - th$ class or segment size (latent variables);

2. Models with predictor variables and covariates:

$$f(\mathbf{x}_i|\mathbf{y}_i, \mathbf{z}_i) = \sum_{s=1}^u (\pi_s|\mathbf{z}_i) f(\mathbf{x}_i|\pi_s, \mathbf{y}_i), \quad (2.5)$$

where: \mathbf{z}_i are covariates having an impact on latent variable (membership to latent classes).

Model with predictor variables consists only with manifest variables. Model with predictor variables and covariates additionally contains concomitant variables that affect the latent variable, belonging to the latent class (Bağ, 2011, page 204-222).

Latent class regression models have following properties (Wedel & Kamakura, 2000, page 17-29):

- have one manifest variable (dependent variable),
- there are predictor variables in the model,
- the model can also have concomitant variables.

There are also models in which there are many manifest variables and concomitant variables, affecting the classification of class observations (Bağ, 2011, pag 204-222). In the analysis, this model was used.

2.3 Concomitant-variable latent class model

Latent class models, besides manifest variables, can also include concomitant variables (also known as covariates) which can have impact on objects belonging in classes, impact on a priori probabilities (Dayton & Macready, 1988, page 173-178). Concomitant variables along with manifest variables are involved in the estimation of latent class model parameters, on the basis of which it will be possible to classify new objects without the participation of manifest variables. Concomitant variables are mostly used in marketing, economic and psychological research due to the high cost of obtaining manifest variables (Genge, 2013).

Parameters of concomitant variables are most often estimated by using one-step technique for estimating the effects of covariates. This is based on simultaneous estimation of concomitant variables along with other parameters of the latent class model. Alternative way for covariates estimation is three-step approach. In this approach, firstly parameters of classic latent class model are estimated (2.1). Then a posteriori probabilities are calculated (2.2). In the third step, regression model parameters are estimated. Those probabilities are treated as dependent variables, and the concomitant variables as explanatory variables.

Concomitant-variable latent class model, where covariates have an effect on the probabilities of belonging objects to classes π_s , model is shown in equation (2.6)

$$f(\mathbf{x}_i, \mathbf{z}_i | \Theta) = \sum_{s=1}^u \pi_s(\mathbf{z}_i, \alpha) f_s(\mathbf{x}_i | \Theta_s), \quad (2.6)$$

where: \mathbf{z}_i is covariates vector, $\mathbf{z}_i = [z_{i1}, \dots, z_{im_2}]^T$, α is a covariates parameters vector.

The following condition must be met

$$\pi_s(\mathbf{z}_i, \alpha) \geq 0 \wedge \sum_{s=1}^u \pi_s(\mathbf{z}_i, \alpha) = 1, \Theta_s \neq \Theta_l \forall s \neq l. \quad (2.7)$$

While estimating the parameters of the concomitant-variable latent class model, the first class is usually called the reference class. It is assumed that odds ratio of a priori probabilities for latent classes in comparison to that reference class is a linear function of concomitant variables. For m_2 covariates, vector of that parameters α_s has $m_2 + 1$ length (for every concomitant variable and constant term).

A priori probability shows the following formula (2.8).

$$\pi_{si} = \pi_s(\mathbf{z}_i; \boldsymbol{\alpha}) = \frac{e^{\mathbf{z}_i \boldsymbol{\alpha}_s}}{\sum_{q=1}^u e^{\mathbf{z}_i \boldsymbol{\alpha}_s}}, \quad (2.8)$$

In concomitant-variable latent class model, $u - 1$ $\boldsymbol{\alpha}_s$ vectors are estimated and also conditional probabilities of belonging objects to latent classes. A posteriori probability and belonging objects to classes is obtained by replacing π_s in formula (2.2) with function $\pi_s(\mathbf{z}_i; \boldsymbol{\alpha})$ from formula (2.8):

$$P(s|\mathbf{x}_i, \mathbf{z}_i) = \frac{\hat{\pi}_s(\mathbf{z}_i; \hat{\boldsymbol{\alpha}}) f(\mathbf{x}_i, \hat{\boldsymbol{\Theta}}_s)}{\sum_{q=1}^u \hat{\pi}_q(\mathbf{z}_i; \hat{\boldsymbol{\alpha}}) f(\mathbf{x}_i, \hat{\boldsymbol{\Theta}}_s)} \quad (2.9)$$

Number of estimated parameters in this latent class model equals: $u \sum_j (l_m - 1) + (s + 1)(u - 1)$, where: l is a category of manifest variable.

Latent class model estimation is based on number and size of several classes estimation. Parameters of LCM are estimated with usage of maximum likelihood method (2.6). Maximum likelihood function shows equation:

$$\log L = \sum_{i=1}^n \ln \sum_{c=1}^C \pi_c(\mathbf{z}_i; \boldsymbol{\alpha}) \prod_{j=1}^m \prod_{h=1}^{l_j} (\theta_{cjh})^{x_{ijh}}, \quad (2.10)$$

where: x_{ijh} is i -th observation of j -th variable and h -th category.

The most commonly used method of maximum likelihood estimation is EM algorithm or Newton-Raphson algorithm.

2.3.1 Choice of the model and assessment of the goodness of fit

There are several methods of selecting the model and assessing the quality of the match. Information criterion are the most commonly used, which are a compromise between the quality of the match and the complexity of the model. The most popular are Bayesian Information Criterion (BIC; Schwartz, 1978) and Akaike Information Criterion (AIC; Akaike, 1974). However, these criteria can give inconclusive indications about the latent class model's assessment (Genge, 2013). In the R package minimization criteria are the most often used, which can be shown in following formulas:

$$BIC_s = -2 \log P(\mathbf{x}_i | \hat{\boldsymbol{\Theta}}_s, M_s) + \log(n) v_s, \quad (2.11)$$

$$AIC_s = -2\log P(\mathbf{x}_i|\hat{\Theta}_s, M_s) + 2v_s, \quad (2.12)$$

where: $\log P(\mathbf{x}_i|\hat{\Theta}_s, M_s)$ is maximum likelihood function logarithm for vector of estimated model's parameters; M_s , v_s are number of model's parameters and n is number of observations.

First part of these equations is responsible for choosing the model of the best quality of the match and the other part rejects models with excessive amount of parameters. These models are used to compare models with different number of classes. The lower the value of criteria the better quality of the match (Genge, 2013).

2.4 The poLCA package

Package poLCA is acronym for Polytomous variable Latent Class Analysis was created by Linzer and Lewis (2011). This package is used to estimate latent class models and latent class regression models. They can be used to identify clusters of similar "types" of observations or individuals from multivariate categorical data. Characteristics of these groups are estimated and probability of belonging each observation o each group is returned. The classic latent class model stratifies the observed variables using theoretical latent categorical variable and attempts to eliminate any spurious between the manifest variables. The latent class regression model further estimates the effects of concomitant variables (covariates) on predicting membership to latent classes (Linzer & Lewis, 2011). Package poLCA uses Newton-Raphson and expectation-maximization algorithms to find maximum likelihood of the latent class and latent class regression models parameters.

```
poLCA(formula , data , nclass = 2 ,                                1
      maxiter = 1000, graphs = FALSE, tol = 1e-10,                 2
      na.rm = TRUE, probs.start = NULL,                             3
      nrep = 1, verbose = TRUE, calc.se = TRUE)                     4
```

Listing 2.1. poLCA function

Besides the arguments presented in Table 2.2, poLCA() package has some other options, that were not used in this paper, they allow for example to calculate the standard errors of the estimated class-conditional response probabilities and mixing proportions (Linzer & Lewis, 2011). Function `polca.reorder` is used to adjust the order of the latent

Table 2.2. Description of polCA() function arguments

Arguments	Description
formula	a formula expression of the form response ~ predictors
data	a data frame containing manifest variables in formula
nclass	the number of latent classes to selected model
graphs	logical, whether polCA should graphically display, the parameter estimates at the completion of the estimated; algorithm, the default is FALSE
nrep	number of times to estimate the model, using different value of probs.start, the default is 1
verbose	logical, indicating whether polCA should output to the screen, the results of the model
probs.start	a list of matrices of class-conditional response probabilities, to be used as the starting values for the estimation algorithm, the default is NULL

Source: Linzer and Lewis (2011)

classes by changing the order of the probs.start.

Chapter 3

Empirical results of quantification of self-selection mechanism

3.1 Exploratory Analysis

3.1.1 Questionnaire

The analysis was conducted on the HBC database from 2010 until 2014. The survey covered 80 018 companies. The aim of this study was to determine where employers put job offers and to study how representative are these data source. In other words we would like to verify differences between companies. It is assumed that companies do not use only one place to put job offers but multiple channels are used. Latent class analysis allows to describe the combinations of these channels.

The survey was conducted by using the CATI method. The question starting the interview was a request to talk with a person who has the best knowledge about company's personnel policy including recruitment, competences of employees and organization of trainings. The survey continued if the answer was yes. Then, the respondent was asked several questions from the part A on general knowledge about company. Such questions were for example: how many employees the company hires, what is the main profile of the company's activity (NACE sections) and the sector of the company (private, public, both public and private). The next part of the survey, the part Z, concerned employment, the personnels policy and already hired staff. Questions concerned whether the company is currently looking for employees and how many is needed. Also

important questions was whether education, experience, set of skills or possessed certifications and specific qualifications were important. The next part of the survey, the part P, concerns the past and the future of the company in terms of employment. The most important question from the point of view of this paper was question what kind of channels company uses (or used) in the finding employees process. The last section, the part D, concerned staff development, their training courses and qualifications. The questionnaire can be found in the Appendix in Chapter A.

3.1.2 Description of manifest variables

For the analysis, part of the aforementioned database was used. Selected manifest variables are shown in table 3.1.

Manifest variables were measured on a dichotomous scale and assume values 1 and 2 (2 is "yes" and 1 is "no" to the question whether companies where looking for employees in mentioned places).

Table 3.1. Channels for placing job offers in the Human Balance Capital survey

No.	Variable name	Where does your company sought for employees?	Yes	No
1	p10a	District labor office		
2	p10b	Job agency		
3	p10c	Headhunters		
4	p10d	Training courses		
5	p10e	Press		
6	p10f	The Internet		
7	p10g	Within company		
8	p10h	Recommendations		
9	p10i	Recruitment fairs		

Source: Own elaboration.

We are also interested in how the choice of places to post job offers by employers was changing through years 2010-2014. The line plot 3.1 shows that changes may not be considered as significant, however we can observe that advertisements in press are used less frequently.

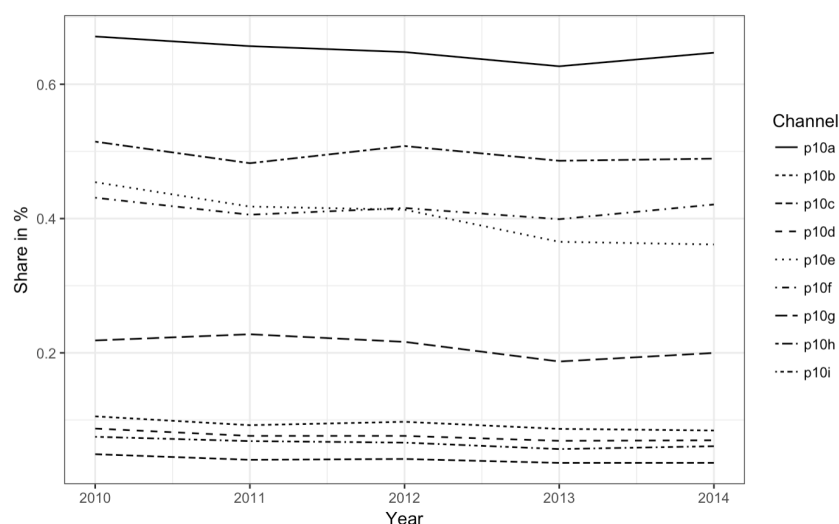


Figure 3.1. Changes of channels for placing job offers in years 2010-2014
Source: own elaboration.

3.2 Description of concomitant variables

The analysis was carried out taking into account the following concomitant variables shown in table 3.2. The following figure 3.2 show shares (in %) relations between concomitant variables and manifest variables.

Table 3.2. Selected concomitant variables

No.	Concomitant variables	Levels
1	the main profile of the company by NACE	industry and mining; education; construction and transport; trade, accomodation, gastronomy; health care and social asistance; specialized services
2	the voivodeship in which the company operates	all 16 voivodeships in Poland
3	the company size	1-9, 10-49, 50+
4	the sector of the company	private, public, both private and public

Source: Own elaboration.

To describe the association between variables, χ^2 test and Cramers-V statistics were calculated. The analysis of the results of χ^2 test shows that p-value values are smaller than 0,05 which suggests association between variables. However, analysis of Cramers-V statistics, the conclusion about strength of correlation differs. Cramers-V statistics values from the range (0,1-0,3) indicate moderate correlation between variables. Detailed statistics are shown in table 3.3. The strongest correlation is between the sector

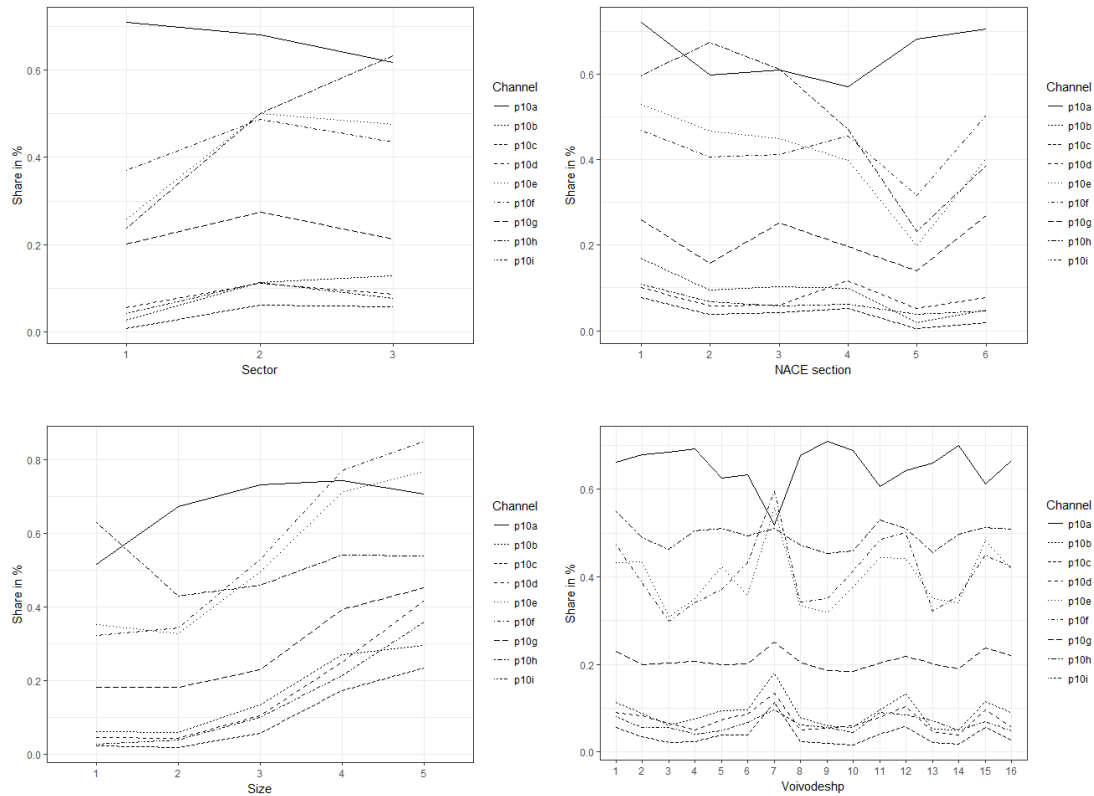


Figure 3.2. Relations between covariates and manifest variables

Source: Own elaboration.

operates and looking for the employees via recommendations.

The aforementioned variables p10a:p10i were included in latent class model as manifest variables and formed the basis for the segmentations of respondents. The latent class model was estimated using the package polCA.

Table 3.3. Selected Cramers-V statistics

Covariate	The biggest value	Manifest variable	The smallest value	Manifest variable
2010				
Sector	0,32	p10h	0,04	p10g
NACE section	0,28	p10h	0,09	p10d
Size	0,28	p10f	0,11	p10h
Voivodeship	0,18	p10f	0,06	p10g
2011				
Sector	0,36	p10h	0,02	p10g
NACE section	0,31	p10h	0,08	p10d
Size	0,29	p10f	0,16	p10a
Voivodeship	0,16	p10f	0,05	p10g
2012				
Sector	0,36	p10h	0,03	p10g
NACE section	0,32	p10h	0,08	p10d
Size	0,29	p10f	0,14	p10g
Voivodeship	0,16	p10f	0,06	p10h
2013				
Sector	0,39	p10h	0,02	p10g
NACE section	0,23	p10h	0,09	p10d
Size	0,26	p10f	0,14	p10g
Voivodeship	0,14	p10f	0,05	p10g
2014				
Sector	0,41	p10h	0,07	p10g
NACE section	0,34	p10h	0,09	p10i
Size	0,25	p10d	0,12	p10g
Voivodeship	0,15	p10f	0,03	p10g

Source: Own elaboration.

3.3 Description of the model

3.3.1 Models

In order to choose the best model, two groups of models were estimated that differ in the number of concomitant variables and number of latent classes.

1. **Model 1** with manifest variables p10a:p10e and number of latent classes from 1 to 3,
2. **Model 2** with manifest variables p10a:p10e, number of latent classes from 1 to 3, and 4 covariates: NACE section, size of the company, voivodeship and the sector of the company.

In order to choose the best model with the best goodness of fit and estimate the number of classes, the information criteria were used, both BIC and AIC. Algorithm was generated five times four 1, 2 and 3 classes. The analysis of the results indicates **Model 2.3** with 3 classes fits the data better. For this model values of both criterion are the lowest, which is presented in table 3.4. Basic characteristics about the conditional item responses for this model are presented in Listing 3.1.

Table 3.4. Information criteria to select the best models

Number of classes	Information criterion	
	AIC	BIC
Model 1		
1.1	669 638	669 722
1.2	632 974	633 151
1.3	626 436	626 706
Model 2		
2.1	669 558	669 641
2.2	618 219	618 637
2.3	603 223	603 975

Source: Own elaboration.

As a result of the latent class concomitant model estimation, **Model 2**, the Figure 3.3 shows probabilities of choosing a channel (the answer "yes") in the first, second and third class.

Conditional item **response** (column) probabilities,
by outcome **variable**, for each **class** (row)

1
2
3

\$p10a				4
		Pr(1)	Pr(2)	5
class 1:	0.3147	0.6853		6
class 2:	0.2551	0.7449		7
class 3:	0.4608	0.5392		8
\$p10b				9
		Pr(1)	Pr(2)	10
class 1:	0.9916	0.0084		11
class 2:	0.7505	0.2495		12
class 3:	0.9680	0.0320		13
\$p10c				14
		Pr(1)	Pr(2)	15
class 1:	0.9998	0.0002		16
class 2:	0.8793	0.1207		17
class 3:	0.9930	0.0070		18
\$p10d				19
		Pr(1)	Pr(2)	20
class 1:	0.9767	0.0233		21
class 2:	0.8019	0.1981		22
class 3:	0.9845	0.0155		23
\$p10e				24
		Pr(1)	Pr(2)	25
class 1:	0.8731	0.1269		26
class 2:	0.1737	0.8263		27
class 3:	0.7249	0.2751		28
\$p10f				29
		Pr(1)	Pr(2)	30
class 1:	0.7560	0.2440		31
class 2:	0.1192	0.8808		32
class 3:	0.8393	0.1607		33
\$p10g				34
		Pr(1)	Pr(2)	35
class 1:	0.8526	0.1474		36
class 2:	0.6873	0.3127		37
class 3:	0.8246	0.1754		38
\$p10h				39
		Pr(1)	Pr(2)	40
class 1:	0.8197	0.1803		41
class 2:	0.4372	0.5628		42
class 3:	0.2935	0.7065		43
\$p10i				44
		Pr(1)	Pr(2)	45
class 1:	0.9764	0.0236		46
class 2:	0.8360	0.1640		47
class 3:	0.9831	0.0169		48
Estimated class population shares				49
0.3139 0.3157 0.3704				50
Predicted class memberships (by modal posterior prob.)				51
0.2941 0.321 0.3849				52

Listing 3.1. Model 2.3 output – latent class probabilities

The first class consisted of 29.41% of the whole sample. The largest share (68%) in this group are enterprises that post their job offers in district labor offices. Around 24% of respondents said that they use the Internet to sought for employees. Recommenda-

tions were chosen by 18% of respondents. Looking for employees within the very same company was used by around 15% of employers that belong to this class. Around 13% of employers posted job offers in the press. Only 2.4% companies look for staff at recruitment fairs. Less than 1% of employers included in this group use channels such as: job agencies, headhunters or training courses.

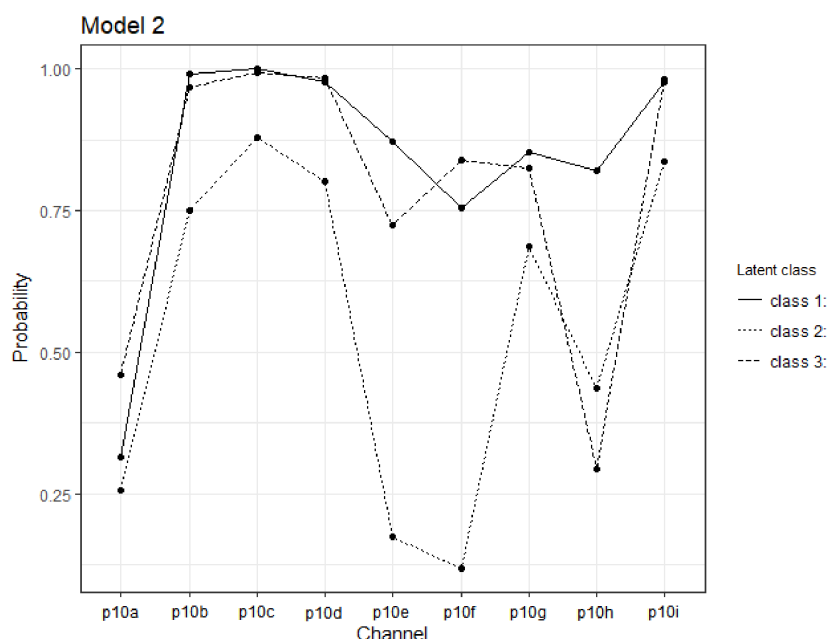


Figure 3.3. Probabilities of the answer "yes" where employees post job offers
Source: Own elaboration.

The second class contains a slightly larger share (32.1%) of respondents than the first class. The Internet and the press were chosen as channels to post job offers, both for 82% of companies. Not much less respondents were looking for employees via district office (74.5%) and recommendations (56.3%). Around one third of employers sought for new staff among already hired staff in their company. Job agencies were chosen by 25% of the enterprises in this group. Between 20% and 10% employers were looking for staff at training courses, recommendations and headhunters.

The third class, the last one and the biggest one consisted of 38,49% share of respondents. In this group, 70.7% of respondents used recommendations to find employees. Over half of respondents post job offers in labor district offices. Around 28% of companies advertise job offers in the press. Looking for employees within the very same company was used by around 17.5% and 16% used the Internet. Less than 3.5% respondents pointed that they look for staff by job agencies, at recruitment fairs, training courses and headhunters.

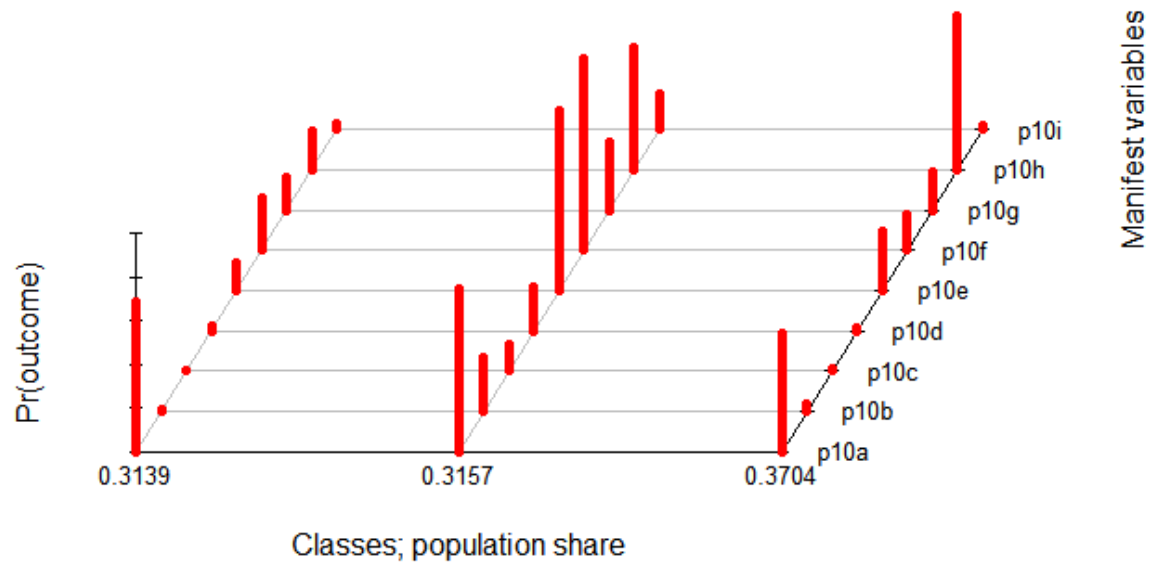


Figure 3.4. Visualisation of latent classes from the model 2.3

Source: Own elaboration.

Function `poLCA()` also creates a graph 3.4 in which the height of the bars corresponds to the probabilities of choosing channels in the chosen Model 2 (with 3 classes). The graph also shows the information about the size of particular classes (percentages).

3.3.2 Self-selection Analysis

Listing 3.2 presents estimated parameters for the model 2.3. In the summary of latent class analysis, main channels of posting job offers in every of 3 classes are variation about those four channels: district labor office, the Internet, recommendations and the press. Labor district office appears in the 3 most frequently chosen channels in every class.

- for the first class: district labor office, the Internet, recommendations,
- for the second class: the Internet, the press, district labor office,
- for the third class: recommendations, district labor office, the press.

```
=====
Fit for 3 latent classes:
=====
```

```
2 / 1
```

	Coefficient	Std. error	t value	Pr(> t)
(Intercept)	-0.51719	0.12414	-4.166	0.000
NACE 2	0.46469	0.10515	4.419	0.000
NACE 3	0.10922	0.08899	1.227	0.220
NACE 4	-0.34046	0.07066	-4.818	0.000
NACE 5	-1.19909	0.06793	-17.653	0.000
NACE 6	-0.06647	0.07457	-0.891	0.373
voivodeship 2	-0.49935	0.11486	-4.348	0.000
voivodeship 3	-1.14037	0.11388	-10.013	0.000
voivodeship 4	-0.53353	0.12451	-4.285	0.000
voivodeship 5	-0.71991	0.11573	-6.221	0.000
voivodeship 6	-0.80173	0.11318	-7.084	0.000
voivodeship 7	0.29557	0.10029	2.947	0.003
voivodeship 8	-0.69497	0.12164	-5.713	0.000
voivodeship 9	-1.28975	0.11391	-11.322	0.000
voivodeship 10	-0.75263	0.11349	-6.632	0.000
voivodeship 11	-0.03500	0.11219	-0.312	0.755
voivodeship 12	-0.31395	0.10486	-2.994	0.003
voivodeship 13	-1.01114	0.11876	-8.514	0.000
voivodeship 14	-0.68923	0.11566	-5.959	0.000
voivodeship 15	-0.37722	0.11709	-3.222	0.001
voivodeship 16	-0.26628	0.11676	-2.280	0.023
size 2	-0.42584	0.08174	-5.210	0.000
size 3	0.49135	0.08370	5.870	0.000
size 4	2.15341	0.10392	20.721	0.000
size 5	3.79188	0.24727	15.335	0.000
sector 2	1.25768	0.10446	12.040	0.000
sector 3	2.83770	0.07127	39.814	0.000

```
3 / 1
```

	Coefficient	Std. error	t value	Pr(> t)
(Intercept)	-1.73951	0.31226	-5.571	0.000
NACE 2	0.65246	0.11784	5.537	0.000
NACE 3	-0.08272	0.10121	-0.817	0.414
NACE 4	-1.03876	0.08942	-11.616	0.000
NACE 5	-1.60773	0.11211	-14.340	0.000
NACE 6	-0.08353	0.11701	-0.714	0.476
voivodeship 2	-0.33593	0.14358	-2.340	0.020
voivodeship 3	-0.59678	0.14227	-4.195	0.000
voivodeship 4	-0.13717	0.15393	-0.891	0.373
voivodeship 5	-0.42092	0.14423	-2.918	0.004
voivodeship 6	-0.61794	0.14074	-4.391	0.000
voivodeship 7	-0.08263	0.12848	-0.643	0.521
voivodeship 8	-0.26474	0.15144	-1.748	0.081
voivodeship 9	-0.86478	0.14194	-6.093	0.000
voivodeship 10	-0.84901	0.14303	-5.936	0.000
voivodeship 11	-0.24710	0.13986	-1.767	0.078
voivodeship 12	-0.28435	0.13285	-2.140	0.033
voivodeship 13	-0.66272	0.14749	-4.493	0.000
voivodeship 14	-0.10936	0.14488	-0.755	0.451
voivodeship 15	-0.24725	0.14465	-1.709	0.088
voivodeship 16	-0.22776	0.14593	-1.561	0.119
size 2	-1.19708	0.08955	-13.367	0.000
size 3	-1.86596	0.09724	-19.189	0.000
size 4	-1.36585	0.13622	-10.027	0.000
size 5	-0.86360	0.42466	-2.034	0.043
sector 2	3.84667	0.30691	12.534	0.000
sector 3	5.47293	0.28250	19.373	0.000

```
=====
number of observations: 79765
number of estimated parameters: 81
residual degrees of freedom: 430
maximum log-likelihood: -301530.5
=====
```

AIC(3): 603223
BIC(3): 603975.2
X²(3): 29042.2 (Chi-square goodness of fit)

69
70
71

Listing 3.2. Model 2.3 output - concomitant variables

The probability that the second-class companies can move to the first class is shown in above mentioned output. What kind of companies place job offers in the Internet and within the same company compared with those that use district labor offices. It concerns companies that operate in mazowieckie voivodeship (the probability grows) in NACE sections such as specialized services, education, health care and social services (the probability drops).

The aforementioned output also shows similar phenomenon but it concerns the probability that the third-class companies can move to the first class. What kind of companies use recommendations to find employees compared with those that mainly use labor district office on that purpose. t value indicates that this probability grows for companies that operate according to NACE section in building and construction, for the rest of NACE sections and for all voivodeships this probability drops, however it also drops when it comes to sectors such as public and both private and public sector.

Conclusions

So far, statisticians relied on data sources such as sample surveys, censuses, registers. However, increasing non-response started pursuit for new, alternative data sources adopted to current situation and development of technology. The Internet as a data source can be a solution. However, these data sources were not created for statistical purposes and the researcher can not decide about the sample selection process. Respondents make decision themselves whether to participate in data collection process or not. This is called a self-selection mechanism and it question sample representativeness. One example is demand for labor which could be either measured directly by survey of employers or indirectly by using electronic (online) trait that employers left.

The main aim of this thesis was to identify what channels do companies use to post job offers and how these companies cluster. The other goal of the study was to describe the self-selection mechanism by the characteristics of companies that advertise job offers in specific channels. To accomplish this subject, latent class analysis was used.

The research indicated that companies can be clustered into classes that vary with class membership and companies that operate in different NACE sections, voivodeships, sectors and size. Latent class analysis enabled segmentation of companies. The results obtained show that three classes were extracted that consisted of companies with similar preferences regarding the channels of posting job advertisements that the companies indicated in the questionnaire. The impact of the self-selection mechanism on their belonging to classes was also assessed.

The first class consisted of mainly companies that prefer to post job offers in the district labor office and in the Internet. The second class consisted of companies that sought for employees via the Internet and the press. The third class consisted in companies that use recommendations and district labor office.

Bibliography

- Akaike, H. (1974). *A new look at the statistical model identification*. IEEE.
- Antosz, P. (2014). Raport metodologiczny z badań realizowanych w 2014 roku w ramach V edycji balansu kapitału ludzkiego.
- Bąk, A. (2011). *Analiza danych jakościowych i symbolicznych z wykorzystaniem programu R*. Wydawnictwo C.H. Beck.
- Beręsewicz, M. (2016). *Internet data sources for real estate market statistics* (PhD thesis).
- Beręsewicz, M. (2017). A two-step procedure to measure representativeness of internet data sources. *International Statistical Review*, 85(3), 473–493.
- Bethlehem, J. (2010). Selection bias in web surveys.
- Brzezińska, J. (2015). Analiza klas ukrytych w badaniach sondażowych., 43–50.
- Dayton, C. M. & Macready, G. B. (1988). *A latent class covariate model with applications to criterion-referenced testing*. Springer.
- Everitt, B. & Dunn, G. (2001). *Applied multivariate data analysis*. John Wiley and Sons.
- Genge, E. (2013). Rola kobiet w polskim społeczeństwie - analiza empiryczna z wykorzystaniem klas ukrytych dla danych jakościowych.
- Gołata, E. (2018). *Koniec ery tradycyjnych spisów ludności*. Uniwersytet Ekonomiczny w Poznaniu.
- Goodman, L. (1970). *The multivariate analysis of qualitative data: interactions among multiple classification*. Journal of the American Statistical Association.
- Haberman, S. J. (1979). *Analysis of qualitative data*. Academic Press, New York.
- Hagenaars, J. & McCutcheon, A. (2002). *Applied latent class analysis*. Cambridge.
- Lazarsfeld, P. (1950). *The logical and mathematical foundations of latent structure analysis*. John Wiley and Sons.
- Linzer, D. A. & Lewis, J. B. (2011). poLCA: an R package for polytomous variable latent class analysis. *Journal of Statistical Software*, 42(10), 1–29.

- MacCallum & Austin. (2000). *Applications of structural equation modeling in psychological research*. Annual review of psychology.
- R Core Team. (2017). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria.
- Sauro, J. (2016). The difference between observed and latent variables.
- Schwartz, G. (1978). *Estimating the dimensions of a model*. Annals of Statistics.
- Statistics Poland. (2014). Popyt na pracę w 2014r.
- Vermunt, J. (2010). *Latent class modeling with covariates: two improved three-step approaches*. Political Analysis.
- Wedel, M. & Kamakura, W. A. (2000). *Market segmentation: conceptual and metodological foundations*. Springer.
- Wołodźko, T. (2015). *Modele cech ukrytych w badaniach edukacyjnych, psychologii i socjologii. teoria i zastosowania* instytut badań edukacyjnych. Instytut Badań Edukacyjnych.

List of Tables

1.1	Taxonomy of survey errors	4
2.1	Methods for the analysis of latent variables	12
2.2	Description of polCA() function arguments	20
3.1	Channels for placing job offers in the Human Balance Capital survey . .	22
3.2	Selected concomitant variables	23
3.3	Selected Cramers-V statistics	25
3.4	Information criteria to select the best models	26

List of Figures

- 3.1 Changes of channels for placing job offers in years 2010-2014 23
- 3.2 Relations between covariates and manifest variables 24
- 3.3 Probabilities of the answer "yes" where employees post job offers 28
- 3.4 Visualisation of latent classes from the model 2.3 29

List of software scripts

2.1	poLCA function	19
3.1	Model 2.3 output – latent class probabilities	26
3.2	Model 2.3 output - concomitant variables	30

Appendix A

The survey of employers' questionnaire

BILANS KAPITAŁU LUDZKIEGO W POLSCE

Badanie pracodawców

Dzień dobry! Nazywam się i jestem pracownikiem SMG/KRC. Prowadzimy badanie dla Uniwersytetu Jagiellońskiego oraz Polskiej Agencji Rozwoju Przedsiębiorczości, dotyczące sytuacji polskiego rynku pracy. Wyniki tego badania będą miały wpływ na kształt projektów finansowanych z Unii Europejskiej, dlatego zależy nam na uzyskaniu informacji od polskich pracodawców. Gwarantujemy, że wszelkie Pana(-i) wypowiedzi traktowane będą w pełni poufnie. Raport z badań będzie dostępny na stronie internetowej projektu. Jest to już druga edycja tego badania – pierwszy pomiar został przeprowadzony w 2010 roku. Jeżeli brali Państwo udział w badaniu zeszłorocznym, bardzo prosimy o ponowną odpowiedź na pytania ankiety. Jednym z celów badania jest opisanie zmian na rynku, dlatego Państwa ponowny udział w projekcie będzie bardzo cenny.

Uwaga: Tam, gdzie nie ma komentarza dla ankietera możliwa jest tylko jedna odpowiedź.

A0	<p>Czy mógłbym/mogłabym rozmawiać z osobą, która w Państwa firmie posiada największą wiedzę na temat polityki personalnej firmy, w tym działań rekrutacyjnych, kompetencji pracowników oraz organizacji szkoleń?</p>	<div> <div>1. tak.....</div> <div>2. nie → ustalić powód odmowy (zgodnie z kluczem) i zakończyć wywiad.....</div> </div> <div> <div>Jeśli tak:</div> <div>Ile osób zatrudnia Państwa firma?</div> <div> <div>Nie zatrudniam nikogo.....1→ zakończyć</div> <div>1-9.....2</div> <div>10-49.....3</div> <div>50-249.....4</div> <div>250-999.....5</div> <div>1000 lub więcej.....6</div> </div> </div>	<div>1</div> <div>2</div>
A1	<p>Jaką pozycję zajmuje Pan(i) w strukturze przedsiębiorstwa?</p> <p>ANK.: NIE odczytywać odpowiedzi na głos. Zaznaczyć odpowiedź na liście, lub zapisać dosłownie w kategorii „inna”.</p>	<div> <div>1. właścicielem / współwłaścicielem.....</div> <div>2. dyrektorem / prezesem / zastępcą dyr. / zastępcą prez.</div> <div>3. członkiem zarządu.....</div> <div>4. specjalistą ds. personalnych, kadr i plac.....</div> <div>5. kierownikiem działu, decydującym o zatrudnieniu w tym dziale.....</div> <div>6. specjalistą ds. księgowości, pracownikiem księgowości.....</div> <div>7. zajmuję inną pozycję – JAKĄ?.....</div> </div>	<div>1</div> <div>2</div> <div>3</div> <div>4</div> <div>5</div> <div>6</div> <div>7</div>
A2	<p>Czy Państwa placówka jest...?</p> <p>ANK.: Odczytać.</p>	<div> <div>1. samodzielna, niezależna firma.....</div> <div>2. siedzibą główną firmy wielooddziałowej.....</div> <div>3. jednym z oddziałów większej firmy.....</div> </div>	<div>1</div> <div>2</div> <div>3</div>
A6	<p>W jakich branżach działa Państwa firma? Jaki jest profil Państwa działalności?</p> <p>ANK.: Dopytać o wszystkie branże.</p> <p>ANK.: Jeśli respondent wskazał więcej niż jedną branżę, odczytać:</p> <p>Która z tych branż jest główną branżą Państwa działalności?</p> <p>KODER: NA PODSTAWIE OPISU ZAKODOWAĆ PKD</p>	<div> <div>1.....</div> <div>2.....</div> <div>3.....</div> <div>4.....</div> <div>5.....</div> </div>	<div> <div>Kod PKD</div> <div> <div>1.....</div> <div>2.....</div> <div>3.....</div> <div>4.....</div> <div>5.....</div> </div> </div>
A6.1	<p>GŁÓWNA DZIAŁALNOŚĆ: wpisać cyfrę z listy powyżej: ____</p>		

A7	Jaka jest forma własności Państwa przedsiębiorstwa?	<div>1. w całości państwowa.....</div> <div>2. częściowo państwowa, a częściowo prywatna.....</div> <div>3. w całości prywatna.....</div>	<div>1</div> <div>2</div> <div>3</div>
----	---	---	--

Część 2 – ZATRUDNIENIE

ANK.: Odczytać:
Chciał(a)bym teraz porozmawiać o polityce zatrudnienia i osobach zatrudnianych w Państwa przedsiębiorstwie. Jeżeli w PYT. A2 → 3. **Jednym z oddziałów większej firmy ANK.: Odczytać:**
Na wszystkie pytania proszę odpowiadać mając na myśli Państwa oddział, placówkę.

Z1	Czy wobec nowo przyjmowanych osób do pracy na stanowiska inne niż kierownicze zakłada się, że ...?	<div>1. posiadają pełne przygotowanie.....</div> <div>2. przejdą niewielkie doszkolenie.....</div> <div>3. przejdą większe doszkolenie.....</div> <div>4. przejdą pełne doszkolenie.....</div>	<div>1</div> <div>2</div> <div>3</div> <div>4</div>
Z2	Czy Pana(i) przedsiębiorstwo poszukuje obecnie osób do pracy na jakimś stanowisku?	<div>1. tak.....</div> <div>2. nie → PRZEJDŹ DO PYTANIA ZP2</div>	<div>1</div> <div>2</div>
	Z2.1 Na ile stanowisk szuka się obecnie pracowników?	LICZBA STANOWISK →	<div>1.....</div> <div>2.....</div> <div>3.....</div> <div>4.....</div>
	Z2.2 Ile osób potrzebują Państwo na te stanowiska?	LICZBA OSÓB →	<div>1.....</div> <div>2.....</div> <div>3.....</div> <div>4.....</div>

ANK.: Odczytać:
Proszę wymienić stanowiska, na które szukają Państwo pracowników, nazywając je możliwie precyzyjnie. Dla przykładu, jeśli szukają Państwo monterów instalacji elektrycznych, to proszę podać pełną nazwę / specjalizację tego stanowiska, a nie tylko jego funkcję – czyli po prostu montera.
Jeśli tych stanowisk jest więcej, to proszę zacząć od tych, na które poszukują Państwo największej liczby ludzi.

Z2.1.1	Jakie stanowisko wymienił(a)by Pan(i) w pierwszej kolejności?	KOD ISCO-08	Ile osób?
Z2.1.2	Jakie stanowisko wymienił(a)by Pan(i) w drugiej kolejności?	KOD ISCO-08	Ile osób?
Z2.1.3	Jakie stanowisko wymienił(a)by Pan(i) w trzeciej kolejności?	KOD ISCO-08	Ile osób?
Z2.1.4	Jakie stanowisko wymienił(a)by Pan(i) w czwartej kolejności?	KOD ISCO-08	Ile osób?
Z2.1.5	Jakie stanowisko wymienił(a)by Pan(i) w piątej kolejności?	KOD ISCO-08	Ile osób?
Z2.1.6	Jakie stanowisko wymienił(a)by Pan(i) w szóstej kolejności?	KOD ISCO-08	Ile osób?

→ Z PODANYCH STANOWISK WYLOSOWAĆ JEDNO I ZAPYTAĆ O NIE SZCZEGÓŁOWO:

Stanowisko nr: WPISZ ZGODNIE Z NUMERACJĄ POWYŻSZEGO PYTANIA		Nazwa stanowiska: PRZEPISZ Z POWYŻSZEGO PYTANIA		KOD ISCO-08 _ _ _ _ _ _ _	
ANK.: Odczytać: Szukają Państwo ludzi na stanowisko [ODCZYTAJ NAZWĘ Z WIERWSZA POWYŻEJ].					
Z3	Czy do tej pory ktoś wykonywał już w Państwa firmie tę pracę, czy też jest to całkiem nowe stanowisko?		1. już wykonywał 1 2. całkiem nowe 2		
Z3.1	Jaka będzie forma zatrudnienia tej osoby?		1. umowa na pełny etat na czas określony 1 2. umowa na pełny etat na czas nieokreślony 2 3. umowa na niepełny etat 3 4. umowa o dzieło 4 5. umowa zlecenie 5 6. bez formalnej umowy 6 9. NIE WIEM nie odczytywać 9		
ANK. Możliwość wielu odpowiedzi.					
Proszę pomyśleć o idealnej osobie do pracy na tym stanowisku.					
Z4	Czy w jej przypadku ważny będzie poziom posiadanego wykształcenia?		1. tak 1 2. nie → PRZEJDŹ DO PYT. Z5 2		
Z4.1	Jakie wykształcenie powinna mieć ta osoba? UWAGA: JEŻELI POJAWIA SIĘ WYKSZTAŁCENIE INŻYNIERSKIE, ŚREDNIE DOPIYTAĆ CZY CHODZI OSOBĘ Z TYTUŁEM MAGISTRA CZY LICENCIATA I ZANACZYĆ TO		1. podstawowe → PRZEJDŹ DO PYT. Z5 1 2. zasadnicze zawodowe → PRZEJDŹ DO PYT. Z4.4 2 3. średnie → PRZEJDŹ DO PYT. Z4.4 3 4. wyższe licencjackie 4 5. wyższe magisterskie 5		
Z4.4	A czy w jej przypadku będzie ważne, jaki ma wyuczony zawód?		1. tak 1 2. nie → PRZEJDŹ DO PYT. Z5 2		
Z5	A czy zatrudniliby Państwo na tym stanowisku osobę z wykształceniem: UWAGA: NIE ODCZYTYWAĆ POZIOMU WYKSZTAŁCENIA PODANEGO W PYT. Z4.1				
	1. podstawowym		tak		nie
	2. zasadniczym zawodowym		1		2
	3. średnim		1		2
	4. wyższym licencjackim		1		2
	5. wyższym magisterskim		1		2
Z6	Czy wymagają Państwo doświadczenia w pracy na podobnym stanowisku?		1. tak 1 2. nie → PRZEJDŹ DO PYT. Z7 2		
Z6.1	Ilu miesięcy bądź lat doświadczenia Państwo wymagają?		_ _ _ _ lat _ _ _ _ miesięcy		
Z7	A czy do pracy woleliby Państwo przyjąć...		1. zdecydowanie kobietę 1 2. raczej kobietę 2 3. raczej mężczyznę 3 4. zdecydowanie mężczyznę 4 8. JEST TO OBOJĘTNE 8		
Z8	W jakim przedziale wieku powinna być osoba na to stanowisko?		1. od _ _ _ _ lat do _ _ _ _ lat 1 2. JEST TO OBOJĘTNE 2		
Z9	Jakie trzy najważniejsze umiejętności powinna posiadać taka idealna osoba na to stanowisko?		1. KOD _ _ _ _ 2. KOD _ _ _ _ 3. KOD _ _ _ _		

Z10	Zależnie od stanowiska od pracownika oczekuje się określonego poziomu kompetencji. Pewne kompetencje mogą być w ogóle niepotrzebne, a inne są wymagane w stopniu podstawowym, średnim, wysokim bądź bardzo wysokim. Proszę powiedzieć, czy kompetencje, które Pan(i) wymienię są potrzebne w przypadku tego stanowiska? A jeżeli tak, to w jakim stopniu? Rotacja odczytywanych stwierdzeń				
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	1. Wyszukiwanie i analiza informacji oraz wyciąganie wniosków.				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim	wysokim	bardzo wysokim	1
	1	2	3	4	
	niepotrzebne → PRZECZYTAĆ NASTĘPNA				
	2				
	potstawowym	średnim</			

ANK.: Odczytać: W tym badaniu interesuje nas również dynamika zatrudnienia w gospodarce. Chcielibyśmy wiedzieć jak zmieniło się zatrudnienie w Państwa placówce bądź firmie w ciągu ostatnich 12 miesięcy.					
P4	Ile osób jest ogólnie zatrudnionych w Państwa placówce bądź firmie obecnie ? Proszę powiedzieć o wszystkich zatrudnionych, w tym na: umowy o pracę, umowy o dzieło, umowy zlecenia, umowy agencyjne, osobach pracujących na własny rachunek i zatrudnionych dorywczo.	JESIŃ NIE ZNA DOKŁADNEJ LICZBY, POPROŚC O PRZYBLIŻENIE. JESIŃ NIE WIE WPISAĆ 99999.	_ _ _ _ _ _ _ osób		
P5	A ile osób było ogólnie zatrudnionych w Państwa placówce bądź firmie 12 miesięcy temu ? Proszę powiedzieć o wszystkich zatrudnionych, w tym na: umowy o pracę, umowy o dzieło, umowy zlecenia, umowy agencyjne, osobach pracujących na własny rachunek i zatrudnionych dorywczo.	JESIŃ NIE ZNA DOKŁADNEJ LICZBY, POPROŚC O PRZYBLIŻENIE. JESIŃ NIE WIE WPISAĆ 99999.	_ _ _ _ _ _ _ osób		
P6	Czy w ciągu ostatnich 12 miesięcy na jakimś stanowisku nastąpił wzrost zatrudnienia bądź utworzono zupełnie nowe stanowisko?	1. tak 2. nie → PRZEJDZ DO PYT. P7	1 2		
P6.1	Na jakim stanowisku miał miejsce największy wzrost zatrudnienia?	KOD ISCO-08 _ _ _ _ _ _ _			
P6.2	O ile osób wzrosło zatrudnienie na tym stanowisku? [w tym pytaniu interesują nas wszelkie zmiany dotyczące zatrudnienia, również te dotyczące niepełnego etatu. W sytuacji kiedy zmniejszyli Państwo zatrudnienie, ale tylko o część etatu należy to zmiane zapisać w postaci ułamka dziesiętnego. np. zmniejszyli Państwo zatrudnienie o 3 osoby na danym stanowisku, każdą o 0,5 etatu co daje nam łącznie wzrost zatrudnienia o 1,5 etatu]	JESIŃ NIE ZNA DOKŁADNEJ LICZBY, POPROŚC O PRZYBLIŻENIE. JESIŃ NIE WIE WPISAĆ 9999.	_ _ _ _ _ _ _ osób/etatów		
P7	Czy w ciągu ostatnich 12 miesięcy na jakimś stanowisku nastąpił spadek zatrudnienia bądź zlikwidowano jakieś stanowisko?	1. tak 2. nie → PRZEJDZ DO PYT. P8	1 2		
P7.1	Na jakim stanowisku nastąpił największy spadek?	KOD ISCO-08 _ _ _ _ _ _ _			
P7.2	O ile osób zmalało zatrudnienie na tym stanowisku? [w tym pytaniu interesują nas wszelkie zmiany dotyczące zatrudnienia, również te dotyczące niepełnego etatu. W sytuacji kiedy zmniejszyli Państwo zatrudnienie, ale tylko o część etatu należy to zmianę zapisać w postaci ułamka dziesiętnego. np. zmniejszyli Państwo zatrudnienie o 3 osoby na danym stanowisku, każdą o 0,5 etatu co daje nam łącznie wzrost zatrudnienia o 1,5 etatu]	JESIŃ NIE ZNA DOKŁADNEJ LICZBY, POPROŚC O PRZYBLIŻENIE. JESIŃ NIE WIE WPISAĆ 9999.	_ _ _ _ _ _ _ osób/etatów		
P8	Jak Pan(i) uważa, czy w ciągu następnych 12 miesięcy liczba osób zatrudnionych w Państwa firmie ... ANK.: Odczytać. Przeznacz teraz Pan(u) różne utrudnienia ograniczające wzrost zatrudnienia. Przy każdym z nich proszę powiedzieć, czy dotyczy ono Państwa firmy. Rotadą odczytywania	1. zmniejszy się 2. zwiększy się 3. czy też pozostanie na tym samym poziomie..... 8. TRUDNO POWIEDZIEĆ	1 2 3 8		
P9	1. niestabilna sytuacja gospodarcza 2. trudny dostęp do kredytów dla firm	nie tak 2 1 2 1	NIE WIEM 9		

[illegible]

D1	Czy prowadzili Państwo jakieś kursy i szkolenia, poza obowiązkowymi, takimi jak np. BHP i przepisy przeciwpożarowe w ciągu poprzedniego roku kalendarzowego, 2012 ?	1. tak 2. nie → PRZEJDŹ DO PYT. D2 9. NIE WIEM.....	1 2 9
	1. tylko kursy wewnętrzne → PRZECZYTAJ PYT. D1.2 Tzn. przygotowane i prowadzone przez samo przedsiębiorstwo – na terenie lub poza terenem przedsiębiorstwa. 2. tylko kursy zewnętrzne → PRZEJDŹ DO PYT. D1.2 Tzn. przygotowane i prowadzone przez innego niż samo przedsiębiorstwo realizatora – na terenie lub poza terenem przedsiębiorstwa. 3. zarówno kursy wewnętrzne jak i zewnętrzne → PRZECZYTAJ PYT. D1.1.1		1 2 3
D1.1.1	Jaką część wszystkich realizowanych kursów stanowiły procentowo kursy wewnętrzne?	W PRZYPADKU ODPOWIEDZI „NIE WIEM” → WPISAC -1 %
D1.2	Biorąc pod uwagę wszystkie rodzaje kursów i szkoleń, to ilu Państwa pracowników uczestniczyło w takich zajęciach w 2012 roku Proszę pamiętać, że chodzi wyłącznie o kursy i szkolenia pozaobowiązkowe. Każda osoba powinna być liczona tylko raz, niezależnie od liczby kursów, w których uczestniczyła.	W PRZYPADKU ODPOWIEDZI „NIE WIEM” → WPISAC -1 osób
D1.3	Proszę teraz pomyśleć o kursach, szkoleniach – z wyłączeniem kursów i szkoleń obowiązkowych – w których uczestniczyła największa liczba pracowników. Czego one dotyczyły? Jaka była ich tematyka?		1. Kod 2. 3.
D1.4	Proszę się chwilę zastanowić nad efektami przeprowadzonych w ciągu poprzedniego roku kalendarzowego, 2012 szkoleń, kursów. Gdyby mogliby Państwo zdecydować raz jeszcze to :		1. przeprowadziliby Państwo w większości te same szkolenia czy kursy 1 2. wybraliby Państwo inne szkolenia czy kursy 2 3. ponieważ efektywność pracowników nie wzrosła istotnie, zainwestowalibyśmy te środki inaczej 3
D2	Czy dofinansowywali Państwo naukę swoich pracowników w szkołach zawodowych i średnich (zasadniczych szkołach zawodowych, technicach, liceach) w ciągu poprzedniego roku kalendarzowego, 2012 ?	1. tak 2. nie 9. NIE WIEM	1 2 9
D3	A czy dofinansowywali Państwo naukę swoich pracowników w szkołach wyższych (studia podyplomowe, MBA, studia zaoczne i inne) w ciągu poprzedniego roku kalendarzowego, 2012 ?	1. tak 2. nie 9. NIE WIEM	1 2 9
D4	Czy wysłali Państwo lub sami organizowali dla swoich pracowników konferencje, seminaria lub warsztaty) w ciągu poprzedniego roku kalendarzowego, 2012 ?	1. tak 2. nie 9. NIE WIEM	1 2 9
D5	Czy w ciągu poprzedniego roku kalendarzowego, 2012 stosowali Państwo system oceny kompetencji pracowników?	1. tak 2. nie 9. NIE WIEM	1 2 9
D6	Czy w ciągu poprzedniego roku kalendarzowego, 2012 stosowali Państwo indywidualne plany rozwoju pracowników?	1. tak 2. nie 9. NIE WIEM	1 2 9

D7	Czy dofinansowywali Państwo samokształcenie pracowników (np. poprzez zakup dla nich książek, prenumeratę czasopism, itp.) w ciągu poprzedniego roku kalendarzowego, 2012 ?	1. tak 2. nie 9. NIE WIEM	1 2 9
D9	Biorąc pod uwagę wszystkie działania jakie podejmowali Państwo w ciągu ostatniego, 2012 roku służące podnoszeniu kwalifikacji i kompetencji Państwa pracowników, proszę powiedzieć czy były one finansowane:	1. ze środków własnych przedsiębiorstwa/instytucji JEŻELI „NIE” → PRZEJDŹ DO PYT. D13 2. ze środków publicznych np. EFS 3. ze środków własnych pracowników Jakie były koszty własne Państwa firmy poniesione na podnoszenie kwalifikacji i kompetencji Państwa pracowników w ciągu ostatniego, 2011 roku ?	tak nie NIE WIEM
D10	ANK.: Po zapisaniu odpowiedzi przejdź do pyt. D11	JEŻELI NIE WIEM → WPISAC 9999999 PLN
D10.1	ANK.: Jeżeli respondent się waha/odmawia odpowiedzi, odczytać: Który z następujących przedziałów, odpowiada łącznym wydatkom na podnoszenie kwalifikacji i kompetencji Państwa pracowników w ciągu ostatniego, 2012 roku ?	1. do 1000 zł 2. 1001-2500 zł 3. 2501-5000 zł 4. 5001-10000 zł 5. 10001-25000 zł 6. powyżej 25000 zł	1 2 3 4 5 6
D11	Uwzględniając wszystkie wykorzystane źródła finansowania przeznaczane na podnoszenie kwalifikacji i kompetencji Państwa pracowników w ciągu ostatniego, 2012 roku , jaką część procentowo stanowiły środki własne przedsiębiorstwa?	JEŻELI NIE WIEM → WPISAC 999 %
D12	ANK.: Pytanie zadawane tylko osobom, które nie podejmowały żadnych działań z PYT. D1-D8 (w PYT. D1-D8 odpowiedzi „NIE”). Dlaczego nie zdecydowali się Państwo na przeprowadzenie kursów, szkoleń czy innych form dokształcania dla pracowników? Przeczytaj Panu(i) różne możliwe powody i proszę powiedzieć, czy miały one wpływ na rezygnację ze szkoleń i innych metod dokształcania zawodowego. Rotacja	 miało wpływ nie miało wpływu
	1. obecne kwalifikacje i umiejętności pracowników w pełni odpowiadają naszym potrzebom 2. trudno nam ocenić, jakie szkolenia byłyby najbardziej potrzebne naszym pracownikom 3. brak oferty dostosowanej do naszych potrzeb 4. szkolenia są dla nas zbyt drogie 5. przeprowadziliśmy już szkolenia w latach poprzednich i w tym roku nie było takiej potrzeby 6. pracownicy, ze względu na brzące obowiązk, nie mają czasu na udział w szkoleniach 7. obecnie kluczowe są dla nas inne inwestycje niż w podnoszenie kompetencji pracowników 8. inne powody		1 2 1 1 1 1 1 1
D13	Czy w ciągu najbliższych 12 miesięcy zamierzają się Państwo ubiegać o środki unijne na rozwój i doskonalenie pracowników, kursy czy szkolenia?	1. tak 2. nie 3. NIE WIEM	1 2 3
D14	Czy mają Państwo obecnie zakładowy fundusz szkoleniowy lub budżet na szkolenia?	1. tak 2. nie	1 2

		3. NIE WIEM.....	3
D15	Czy w ciągu najbliższych 12 miesięcy Państwa przedsiębiorstwo zamierza organizować kursy wewnętrzne lub zewnętrzne doskonalące umiejętności zawodowe pracowników?	1. tak..... 2. nie..... 3. NIE WIEM.....	1 2 3
D16	Czy w ciągu najbliższych 12 miesięcy Państwa przedsiębiorstwo zamierza organizować dla pracowników inne formy doskonalenia zawodowego?	1. tak..... 2. nie..... 3. NIE WIEM.....	1 2 3

Dziękujemy za rozmowę!