ECONOMETRICS
EKONOMETRIA
Advances in Applied Data Analysis

sciendo

Econometrics. Ekonometria.
Advances in Applied Data Analysis
Year 2023, Vol. 27, No. 3
ISSN 2449-9994

Evaluation of the Labour Market Situation of Young People in EU Countries – The Multiple Regression Approach

Magdalena Kawecka

University of Economics in Katowice, Katowice, Poland

e-mail: magdalena.kawecka@edu.uekat.pl

ORCID: 0000-0001-7671-8340

Quote as: Kawecka, M. (2023). Evaluation of the Labour Market Situation of Young People in the European Union Countries – The Multiple Regression Approach. *Econometrics. Ekonometria. Advances in Applied Data Analysis*, *27*(3).

DOI 10.15611/eada.2023.3.03

JEL Classification: C30, C39, J24, F16

Abstract: The article considers the problems of young people aged 20-24 on the labour market affected by unemployment in European Union countries. Unemployment is one of the most important economic and social problems, which at the same time constitutes one of the biggest measures characterising the condition of the economy. The diversity of the economic situation in EU countries directly affects young people, an individual group of people entering the labour market and have little or no professional experience. At the same time, they are ready to start work, facing great difficulties in entering the market, influenced by socio-economic as well as demographic factors which directly and indirectly affect employment. Considering the above premise, the aim of the article was to identify the determinants of unemployment of young people aged 20-24 in the EU. The study used data from two years: 2010 and 2020, and applied multiple regression. Statistical data were taken from Eurostat databases. The study allowed to examine the dependence of the influence of individual socio-economic as well as demographic factors on youth unemployment. The study found that the multivariate regression showed that factors related to young people's participation in education and training (including the NEET rate) relative to labour market status, as well as social inclusion, had a significant impact on the unemployment studied. Over the decade, a decrease was seen in unemployment in most EU member states, in as many as 19 countries, while the remaining eight countries showed an increase.

Keywords: labour market, European Union, young people, unemployment, multiple regression

1. Introduction

The labour market is the subject of much comparative research. Researchers focus on selected areas or the whole economy of an international, local, or specifically selected location. This research identified differences between countries, finding existing similarities. In these cases, it is necessary to consider a number of characteristics of the studied objects, which occur in their environment in relations and dependencies between each other in a direct or indirect way. The links between them are determined by macroeconomic and demographic characteristics of the labour market situation. Dagnelie (1975, p. 362) noted that "Multivariate statistical analysis should be regarded as a set of statistical methods whose purpose is to study the relationships existing between a number of dependent or interdependent variables". Thus, multivariate analysis focuses on multiple variables that characterise the same objects or phenomena, and the variables under study are dependent or interdependent (Balicki, 2013).

Multivariate statistical methods supply a tool for the verification of identified relations to detect unrecognised correlations between the characteristics under study. This has a significant impact on the correct understanding of the relations that occur between selected characteristics. The research subject of this article is the labour market of young people in the European Union (EU) member states. Based on substantive and formal-statistical criteria (such as universality, variability, degree of correlation, validity), twelve macroeconomic and demographic variables representing the labour market situation of young people aged 20-24 from two years: 2010 and 2020. In addition to the substantive criterion, the availability, comparability, and validity of statistical data were also considered. The theory of correlation and multiple regression allowed to accurately determine to what extent the studied characteristics are related to each other and whether there is any relation between them, and above all, whether the independent variables influence the dependent variable, i.e. the studied characteristic (Cormack, 1971; Mardia et al., 1979).

The aim of the research was to identify the determinants of youth unemployment, and to assess the labour market status of young people in EU countries. The research was conducted considering factors that influence the level of employment and unemployment of young people using data obtained from the databases of the European Statistical Office Eurostat for 2010 and 2020, where the harmonised unemployment rate is the result of Eurostat's standardised method of determining this indicator for each of the European Union countries. The data were calculated by Eurostat based on the quarterly results of the Labour Force Survey (LFS), using the International Labour Organization (ILO) definition, and monthly registered unemployment data (GUS, 2022). The study focuses on a narrow group of young people who have completed education, regardless of the level and belong to the age group of 20-24. The article continues the research on the labour market for young people, examining factors affecting unemployment, but most importantly shows how important a social group young people are (Trzpiot and Kawecka, 2021a, 2021b).

In the last few years the labour market has undergone many changes, which have caused a lot of instability for many age groups. One of the groups most affected by these changes was the group of young people who were on the labour market for some time or were just starting out (Nowak-Brzezińska, 2018). The problems that young people face in the labour market can be presented in the form of the following questions: "What is the current situation of young people in 2020, and what was it like in 2010? What factors influence the development of unemployment among young people? Can certain trends/patterns in this regard be discerned?". These questions constitute an additional aim of the paper, which is to verify them based on the conducted literature review and statistical research.

2. Literature review and research gap

The basis for conducting the research was a review of the literature, as well as the results already obtained by others in this area of research. The key point was to clarify the concept of the labour market and the role it plays in the national economy of each country. The employer in the labour market is regarded as a buyer, while the employees are regarded as sellers (see Kryńska and Kwiatkowski, 2013). As noted by Kryńska and Kwiatkowski (2013, p. 11), the general labour market can be defined as "the place where transactions of exchange of labour services between employees and employers take place and where the size of the said transactions and their conditions are determined, especially the price of these services, i.e. wages." This market is therefore based on people, and thus active participation in the market is defined as employment. Human capital is the basis of the labour market, and thus people with their skills, experience, abilities, and qualifications. The labour market, like other markets, is based on demand and supply regulated by changing prices, in the case of the labour market it is the wage (understood as salary) based on a contract between the employer and the employee. There are many factors that have a significant impact on the functioning of the labour market, such as the number of people available in the market, wage flexibility, labour mobility, development, among others (Abbott, 2013).

Wawrzyniak et al. (2020, pp. 416-429) conducted a study on the similarity of EU countries in terms of the structure of the unemployed. Knowledge of unemployment structures is essential when designing strategies to combat unemployment, especially among people in special situations on the labour market. The survey showed the similarity of the structures of the unemployed according to selected categories (gender, age, education, and duration of unemployment) in most countries. However, it is possible to distinguish countries in which these structures differ significantly from each other, which may result from the specific conditions of the labour markets of individual countries (e.g. in terms of legal regulations). Fergusson and Yeates (2021, pp. 22-37) focused their attention on social inequalities and the exclusion of young people in terms of forms of employment, pointing out that precarious employment leads to an increase in unemployment. Youth unemployment, which is a social problem in many European countries, has increased, especially during the financial and economic crisis since 2007 (Nienaber et al., 2020, pp. 5-27). The mobility of young people contributes to employability. The experience of mobility, i.e. adaptability, professional identity, and social and human capital, is more effective in increasing the employability of young people. The authors (pp. 5-27) drew attention to the fact that mobility also contributes to the stabilisation of employment. Lambovska et al. (2021, pp. 55-63) emphasised that even the smallest changes in the structure of the economy, but above all, rapid changes lead to a sudden increase in unemployment among young people (an example of which is the introduction of restrictions due to the COVID-19 pandemic). In addition, any economic recession affects the persistence of long-term youth unemployment. Sulich et al. (2020, pp. 1-10), observed that in both Poland and Belgium around 15% of young people find their first employment in the green jobs sector, but in the Czech Republic the proportion is far lower (1.83%). Therefore, the increasing focus on building a green economy supplies excellent employment opportunities for young people looking for their first job.

The youth unemployment rate is much higher than the total unemployment rate in every European country. Oesingmann (2015, pp. 52-55) showed that in countries such as Greece, Spain, Croatia, Italy, Slovakia and France, the most striking unemployment rates among young people aged 15 to 24 can be seen. Despite the huge discrepancies in Europe with respect to youth unemployment, the author noted that the phenomenon of unemployment itself is not a new phenomenon that should be entirely attributed to the Great Recession. Most European countries are struggling with persistent unemployment.

Young people are most often considered to be persons aged between 15 and 24 (World Youth Report..., 2004), where the same definition is adopted by the International Labour Organisation and the European Commission (ILO, 2012).

Young people as a specific social group are constantly changing, and its dynamics and specificity undergo many transformations along with the generational changes taking place in society (Głód and Miotła, 2014, pp. 323-327). Many determinants directly and indirectly affect the situation of young people, among which: education (according to ISCED 2011 levels), participation in education (including NEET indicators), social inclusion (e.g. material deprivation, risk of poverty and social exclusion). Young people are a group of people eager to work who want to pursue self-development. One of the main problems in taking up a job is the lack of professional experience, which affects employment by the employer, but also the conditions and stability on the labour market, business cycles, related to GDP fluctuations, investment outlays, population income and company profits (Burda and Wyplosz, 2000). Employers, when considering the costs of hiring staff, training and implementation into the system prevailing at work, are less inclined to employ young people on permanent employment contracts. More often, these are fixed-term contracts, part-time contracts or commission contracts, contracts for specific work, as well as internship contracts (European Commission, n.d.). In the best case, these contracts are renewed, but due to high fluctuations in the labour market, youth unemployment is increasing. Youth unemployment varies across the EU, as shown in Table 1.

The table shows the unemployment rate in terms of the passage of time, and compares the ratio of the unemployment rate in 2020 to that in 2010. Thus, out of 27 Member States, eight recorded an increase in the unemployment rate of young people aged 20-24 in 2020 compared to 2010. The highest increase can be seen in Luxembourg (62.79%), followed by Austria (22.5%), Cyprus (12.82%). In the remaining five countries, the increase was below 8.5%.

Table 1. Ratio of the size of the unemployment rate in 2020 to 2010

GEO/TIME	2010	2020	Chain index (2010 = 100%)
Austria	8.5	10.4	22.35
Belgium	20.4	14.3	-29.90
Bulgaria	19.1	12.1	-36.65
Croatia	27.9	15.2	-45.52
Cyprus	15.6	17.6	12.82
Czechia	15.9	6.5	-59.12
Denmark	12.7	9.4	-25.98
Estonia	29.7	15.5	-47.81
Finland	17.0	17.6	3.53
France	20.8	18.7	-10.10
Germany	9.5	6.2	-34.74
Greece	32.1	34.6	7.79
Hungary	24.9	11.4	-54.22
Ireland	25.6	12.9	-49.61
Italy	24.8	26.7	7.66
Latvia	32.8	13.1	-60.06
Lithuania	33.6	19.1	-43.15
Luxembourg	12.9	21.0	62.79
Malta	8.4	9.1	8.33
The Netherlands	9.4	7.7	-18.09
Poland	22.8	10.3	-54.82
Portugal	20.4	21.2	3.92
Romania	21.1	14.3	-32.23
Slovakia	30.5	16.8	-44.92
Slovenia	14.5	13.6	-6.21
Spain	36.9	35.1	-4.88
Sweden	19.8	17.5	-11.62

Source: authors' work based on Eurostat database (2010 and 2020).

In the remaining nineteen countries, the unemployment rate in 2020 decreased compared to 2010. The highest decreases in the range of 40% to 60% were observed in Lithuania (43.15%), Slovakia (44.15%), Croatia (45.52%), Estonia (47.81%), Ireland (49.61%), Hungary (53.22%), Poland (54.82%), Czechia (59.12%), and Latvia (60.06%). As previously mentioned, the youth unemployment rate was higher than the unemployment rate. Bell and Blanchflower (2011, pp. 241-267) found that young people are more vulnerable to the effects of a recession than older generations. In many European countries, the youth unemployment rate has risen significantly faster than the general unemployment rate. In addition, for young people who lose their jobs, the difficulty in finding new employment may last longer than older workers. O'Higgins (2012, pp. 395-412) came to similar conclusions, noting that compared to earlier crises, the Great Depression had more diverse effects on young people in different European countries. In some countries, young people were disproportionately affected by the crisis, while in others the differences between young people and the general population were less pronounced. Eichhorst and Neder (2013, pp. 230-235) noted that youth unemployment in Europe is a complex problem that is not only due to the economic cycle. It is often the result of structural problems such as skills mismatches, rigid labour markets and the weak link between the education system and the labour market. They also proposed various policy measures, such as the promotion of apprenticeships and greater support for young people entering the labour market. It should be noted that the period from 2010 to 2020 was extremely important for the labour market, given the period 2010-2013, when the effects of the economic crisis continued for many European countries, leading to high unemployment rates, especially among young people. In some countries, such as Greece and Spain, the youth unemployment rate exceeded 50% (García-Pérez et al., 2018, pp. 1-52). The EU average was also high, but with significant differences between countries (Eichhorst et al., 2013, pp. 230-235).

Mussida and Sciulli (2023) note that both poverty and the NEET (not in employment, education or training) rate in EU countries (21 countries were included in the study) were characterised by a significant country-specific real dependence. In addition, the NEET rate itself has a negative impact on unemployment, while for young people living outside the family, the NEET situation is not detrimental to poverty, provided that adequate support is provided to young people. Pennoni and Bal-Domańska (2022) showed in their study of young people that Italy was the worst performing country in terms of both NEET and youth unemployment rates, while the Czech Republic was the best performing country in terms of NEET development, and Poland and Slovakia were the best in terms of youth unemployment development. Szluz (2017), based on data from Eurostat, pointed out that in Poland, the main reason why young people live in the family home is the scarcity of available housing, as well as the lack of money, followed by unemployment, as Piszczatowska-Oleksiewicz (2017) also observed.

The situation of young people on the labour market is constantly changing. The problem of unemployment is complex, and there are many definitions and analyses of this phenomenon in the literature itself, which consider its various aspects and approaches. Attention should be paid to the heterogeneity of the labour markets of the EU Member States. The structure of the labour market, institutions, trade union traditions and even employment protection systems affect the diversity of markets. Boeri and Garibaldi (2007, pp. 357-385) analysed the differences in employment protection in European countries and the introduction of two-level reforms, and found that two-stage labour market reforms have a transient 'honeymoon' effect that creates jobs. On the other hand, a dynamic model of demand for labour in conditions of uncertainty, because of reform, in addition to employment growth, there should be a reduction of 'employment idleness' and average and cross--cutting variability of labour productivity. Eichhorst and Marx (2011, pp. 73-87), referring to the recurring concept of dualism in the analysis of the political economy of labour market reforms in European countries, stated that the growing availability of non-standard forms of work increases the pressure on key workers to accept greater flexibility (Germany being the analysed case). The results of the study showed that labour market reforms were indeed aimed at outsiders and continued with small and sometimes contradictory steps. The direction of change was determined by the pressure of

socio-economic problems of a given period, but independent of the composition of the government. However, while insiders were fundamentally opposed to such marginal flexibility – after the reforms – they responded with wage restraint and other instruments, thus strengthening their competitiveness vis-à-vis flexible workers. Another study (Bentolila et al., 2012, pp. 155-187) analysed two-tier labour markets during the Great Recession, using the example of France and Spain due to their similar labour market institutions, where the unemployment rate was around 8% just before the Great Recession, but then rose to 10% in France and to 23% in Spain, respectively. The study found that Spain could avoid around a 45% increase in unemployment if it adopted France's employment protection legislation. Boeri and van Ours (2008) provided a detailed overview of imperfect labour markets in Europe, including the impact of different institutions on labour market balances. They noted that labour markets are experiencing strong seismic changes caused by, among others, growing self--employment, temporary employment, as well as the development of the sharing economy. The research directions also included structural and cyclical unemployment, e.g. in relation to the impact of labour market fluctuations. The authors (Elsby et al., 2015, pp. 64-82) emphasised that labour force flows account for one-third of unemployment fluctuations, and countercyclical attachment to the labour force among the unemployed is a key explanation. In addition, within the sheer complexity of work, one can mention e.g. indicators of difficulties in matching employers with employees (Cahuc et al., 2006, pp. 323-364; Gregg and Petrongolo, 2005, pp. 1981-2005;); social policy, including labour market rigidity (Boeri, 2002; Blanchard and Wolfers, 2000, pp. C1-C33); Human Capital and Education (Becker, 1964; Ciccone and Papaioannou, 2009, pp. 66-82; Psacharopoulos and Patrinos, 2004, pp. 111-134), social and cultural effects (Algan and Cahuc, 2009, pp. 111-145; Guiso et al., 2006, pp. 23-48; Kahanec and Zimmermann, 2016), and many other aspects. The high level of unemployment means that young people not only feel excluded from society, but also experience material deprivation and the threat of poverty. This creates a reluctance to seek active employment and education. This situation is unfavourable, as it evokes a certain regularity – "the younger and better educated the group, the worse its situation on the labour market" (Piecuch and Piecuch, 2014).

3. Research methodology

The subject of the study was the labour market of young people in EU countries. The author used data from the European Statistical Office (Eurostat) according to the international classification ISCED11, the European Union Labour Force Survey (EU LFS), EU-SILC and Euro SDMX metadata (ESMS). The research was conducted using Statistica software (version 13), and covered a total of twenty-seven member states; the data concerned young people aged 20-24 by gender in the European Union in two years 2010 and 2020. The choice was dictated by the dynamics of economic change that affect youth employment. The selected variables came from several categories, i.e. economy, labour market, education, social inclusion.

Following the existing scientific output in the field of quantitative methods, the set of variables characterising the labour markets of selected European Union countries, constituting an economic aggregate, was assessed according to statistical criteria (Maxwell, 1997; Panek, 2009; Strahl, 1998). As a result of the verification of the examined variables, twelve socio-economic and demographic variables ($X_1, X_2, \dots, X_k$) were included in the study, where dependent variable Y (explanatory variable) was the level of unemployment. In addition, the variables determining the level of unemployment of young people aged 20-24 were divided into gender groups. For the unification of variables, standardisation of variables was used, namely the standardisation of diagnostic characteristics was carried out, performed separately for 2010 and then for 2020. The full names of the variables, their abbreviations and units are presented in Table 2. The variables selected for the analysis were divided into three groups of factors: participation in education, social inclusion, type of employment (Trzpiot, 2015).

Table 2. List of variables used in employment and unemployment surveys

Data	Unit	Variables	Variable name
Unemployment	%	Y	Youth unemployment by gender, age
Participation in education[a]	%	X_1	Participation rate employed persons in education and training
		X_2	Participation rate not employed persons in education and training
		X_3	Participation rate not employed persons in education and training (incl. NEET rates) [b]
Social inclusion	%	X_4	People living in households with very low work intensity (population aged 0 to 59 years)
		X_5	Overcrowding rate by poverty status
		X_6	Severe material deprivation rate
		X_7	People at risk of poverty or social exclusion
		X_8	Housing cost overburden rate by poverty status
Type of employment	%	X_9	Self-employment by educational attainment level
		X_{10}	Temporary employees as percentage of the total number of employees
		X_{11}	Part-time employment as percentage of the total employment
		X_{12}	Involuntary part-time employment as percentage of the total part-time employment

[a] Education that is institutionalised, intentional and planned by an education provider. The defining characteristic of non--formal education is that it is an addition, alternative and/or a complement to formal education within the process of the lifelong learning of individuals. It is often provided to guarantee the right of access to education for all. It caters for people of all ages but does not necessarily apply a continuous pathway-structure; it may be short in duration and/or low intensity, and it is typically provided in the form of short courses, workshops, or seminars. Non-formal education mostly leads to qualifications that are not recognized as formal qualifications by the relevant national educational authorities or to no qualifications at all. Non-formal education can cover programmes contributing to adult and youth literacy and education for out-of-school children, as well as programmes on life skills, work skills, and social or cultural development (UNESCO UIS, 2013); [b] NEET (young people neither in employment nor in education or training) (Eurostat, n.d.).

Source: authors' work based on Eurostat database (2010 and 2020).

Regression analysis is a statistical method to determine the quantitative dependence of a variable on one or more independent variables. The more independent variables X_k, the better the explanation for dependent variable Y (Trzpiot, 2015). The multiple regression model is mostly used for:

1) "[…] recognition of the size and type of influence of one variable on another variable,
2) prediction of unknown values of one variable based on the known or assumed value of the other variable (prediction of a dependent variable),
3) explaining the variability of one variable by using the variability of the other variable,
4) predicting the value of a dependent variable" (Aczel and Sounderpandian, 2017, pp. 514-517).

The population regression model of dependent variable (explained) *Y* on a set of *k* independent (explanatory) variables $X_1, X_2, \dots, X_k$ was defined by (Aczel and Sounderpandian, 2017, pp. 514-520; Morrison, 1990; Stanisz, 2007, pp. 59-70):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon, \tag{1}$$

where: β_0 is the Y intercept of the regression surface and each β_i, $i = 1, \dots, k$, the slope of the regression surface sometimes called the response surface with respect to variable X_i, and ε is a random component.

Model assumptions: First, "for each observation, error term ε is normally distributed with mean zero and standard deviation σ and is independent of the error terms associated with all other observations,"[1] this is independent of other errors, i.e.:

$$\varepsilon_j \sim N(0, \sigma^2) \text{ for all } j = 1,2, \dots, n. \tag{2}$$

Second, "in the context of regression analysis, variables X_j are considered *fixed quantities*, although in the context of correlational analysis, they are random variables. In any case, X_j *are independent of the error term* ε. When we assume that X_j are fixed quantities, we are assuming that we have realizations of k variables X_j and that the only randomness in Y comes from the error term ε." (Aczel and Sounderpandian, 2017, pp. 514-520).

The estimated regression equation is as follows:

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2 + \cdots + b_i X_k, \tag{3}$$

where: $\hat{Y}$ – value of Y, the value lying *on* the estimated regression surface. The terms b_i, $i = 0, \dots, k$, are the least-squares estimates of the population regression parameters β_i (Aczel and Sounderpandian, 2017, pp. 517-519).

The F-test as applied to a multiple regression model was the first test needed to perform and to evaluate a multiple regression model. This test answered the basic question derived from the hypothetical regression equation: "whether there is a linear regression relationship between the explained variable and any of the independent variables X_i." The answer to this question was given by a test of the statistical hypothesis that there is a linear relationship between variable Y and any of variables X_i (Aczel and Sounderpandian, 2017, pp. 522-527):

$$\begin{gathered} H_0: \beta_1 = \beta_2 = \beta_3 = \cdots = \beta_k = 0 \\ H_1: Not\ all\ the\ \beta_i (i = 1, \dots, k)\ are\ zero \end{gathered}. \tag{4}$$

If "the null hypothesis is true, then there is no linear relationship between variable Y and any of the explanatory variables present in the hypothetical regression equation" which suggests the completion of the regression analysis. In contrast, "rejection of the null hypothesis means that there is a statistical basis for assuming that there is a linear relationship between variable Y and at least one explanatory variable X_i present in the hypothetical model" (Aczel and Sounderpandian, 2000, pp. 522-524). To carry out the important test in equation 4, an analysis of variance was performed. The ANOVA had k independent variables instead of only one as in simple linear regression, thus the F test of the analysis of variance is not equivalent to the t test for the significance of the slope parameter. Since in multiple regression there are k slope parameters, we have k different t tests to follow the ANOVA (Aczel and Sounderpandian, 2017, pp. 522-527; Francuz and Mackiewicz, 2005). By verifying the normality of the distribution of the variables under study, two tests were applied: Lilliefors and Shapiro-Wilk.

The Lilliefors test is a normality test that is a modification of the Kolmogorov-Smirnov test for cases where the distribution parameters (mean and variance) are not known and must be estimated from the sample. This test is particularly useful for small samples where the traditional K-S test might not be suitable due to parameter estimation. The hypotheses of the test assume (Conover, 1999):

[1] The multiple regression model is valid under less restrictive assumptions than these. The assumptions of normality of the errors allowed to perform t tests and F tests of model validity. All we needed was that the errors be uncorrelated with one another. However, normal distribution + noncorrelation = independence (Aczel and Sounderpandian, 2017, pp. 514-520).

$$H_0: The\ data\ are\ normally\ distributed$$
$$H_1: The\ data\ are\ not\ normally\ distributed\ .$$

The interpretation is as follows: if value p < the significance level (assumed p = 0.05), we reject H_0 and assume that the data do not come from a normal distribution. If p-value > the significance level, there is no basis for rejecting H_0 and it is assumed that the data can come from a normal distribution.

In contrast, the Shapiro-Wilk test is one of the tests used to assess the normality of the distribution of a sample of data. This test is often preferred in the context of small samples because it is more powerful than some other normality tests, such as the Kolmogorov-Smirnov test. The hypotheses of the test assume (Hanusz and Tarasińska, 2012; Royston, 1992, pp. 117-119; Shapiro and Wilk, 1965, pp. 591-611):

$$H_0: The\ data\ are\ normally\ distributed$$
$$H_1: The\ data\ are\ not\ normally\ distributed\ '$$

where *W*-values range from 0 to 1:

- W = 1 indicates an ideal normal distribution.
- W < 1 suggests deviations from normality. The more the value of W deviates from 1, the greater the deviation from normality.

The interpretation of p is the same as that of the Lilliefors test.

Studies assume a significance level of p = 0.05, considered as the standard significance level, with the level adopted being treated as a guideline rather than a strict criterion (Fisher, 1970).

4. Results

4.1. Model I – by participation of young people in education (NEETs)

The participation of young people in education is extremely important because it raises qualifications that have a key impact on employment, while at the same time making the potential candidate more attractive for employment. The report of the Voivodeship Labour Office in Szczecin "Survey of the profile of the desired employee – employer" from 2006 (p. 40) shows that for employers, attractive employee features are as follows: diligence (71%), honesty (58%) and loyalty (61%), followed by education, communication skills and commitment. Nowadays, equally important qualities for an employer are the desire to learn, self-development, as mentioned by Arntz et al. (2020, pp. 1-21), focusing on the impact of digitalisation on the labour market of young people. The impact of education and vocational education on employment prospects is also noted by Hanushek et al. (2020, pp. 241--279). The regression analysis was conducted by gender (female, male) and by year (2010 and 2020).

4.1.1. Unemployment of women by participation in education

The box-plot graphs allowed to visualise the distribution of women's unemployment by participation in education for the two periods studied, specifically to examine the median, quartile, minimum and maximum values. In 2020, an increase in variable X_1 (participation rate employed persons in education and training) and X_2 (participation rate not employed persons in education and training) can be seen compared to 2010 (Figure 1). In the case of variable X_3 (participation rate not employed persons in education and training (including NEET rates)), a decrease was seen.

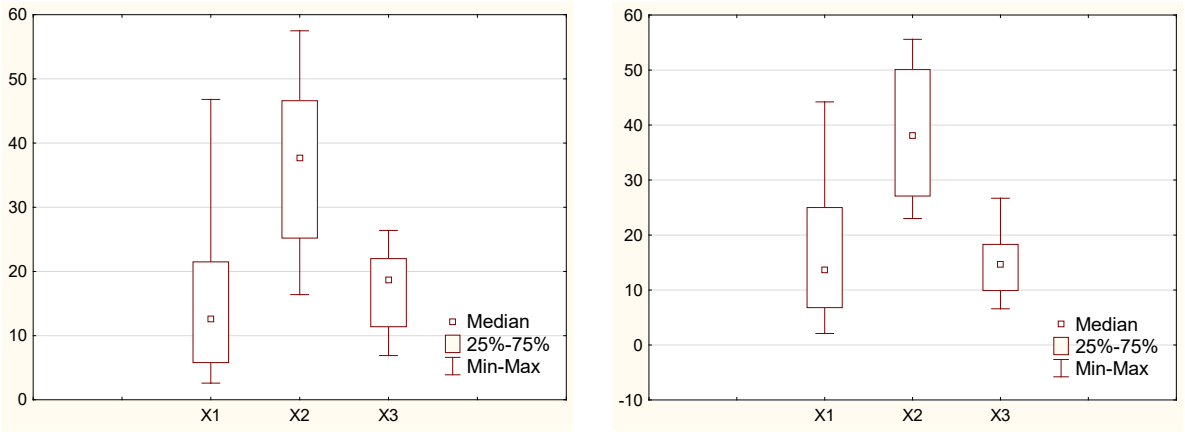

Fig. 1. Box-plot chart for women by participation in education (including NEET) in 2010 left and 2020 right (in %)

Source: authors' work based on Eurostat database (2010 and 2020).

Based on the estimated values of the parameters $(5 \text{ and } 6)$ coefficients, the relation between unemployment (dependent variable) and participation in education of women (independent variables) aged 20-24 in 2010 and 2020 can be described by the equation:

$$Y_{2010} = \underset{(4.397)}{-6.744} + \underset{(0.102)}{0.257X_2} + \underset{(0.176)}{0.984X_3}, \tag{5}$$

$$R^2 = 0.635;\ F(2,24) = 23.637;\ p < 0.0000, SE_e = 5.327 \tag{5.1}$$

$$Y_{2020} = \underset{(13.295)}{-31.263} + \underset{(0.223)}{0.476X_1} + \underset{(0.173)}{0.619X_2} + \underset{(0.345)}{0.965X_3}, \tag{6}$$

$$R^2 = 0.464;\ F(3,23) = 8.517;\ p < 0.00055, SE_e = 5.924, \tag{6.1}$$

where: Y_{year} – youth unemployment; $X_1, \dots, X_k$ –ependent variables; adjusted R^2 – is a special form of R^2, the coefficient of determination; $F_{(k,n-(k+1))}$ – *F*-test of ANOVA analysis of variance; p – *p*-value, probability value; assumption of significance level $\alpha = 0.05$.

In 2020, an increase can be seen in the index of variable X_2 (by 0.36), i.e. participation rate not employed women in education and training. In contrast, a decrease was seen for variable X_3, namely "participation rate not employed women in education and training (including NEET rates)" (by 0.019). The X_1 variable for 2010 is statistically insignificant. The model is well fitted, both for 2010 (5) and 2020 (6), as evidenced by the small standard error of estimation: $SE_{e_{2010}} = 5.33, SE_{e_{2020}} = 5.92$ (according to the assumption it should be as small as possible). The value of the coefficient of determination for 2010 (5.1) is $R^2 = 0.635$, and for 2020 (6.1) $R^2 = 0.464$, which means that 63.5% and 46.4% of the total variation in the dependent variable is explained by the model and the results of the F test of the analysis of variance: $F_{2010} = 23.637\ (5.1), F_{2020} = 8.517\ (6.1)$. According to the test of the statistical hypothesis that a linear relationship exists between the variable Y and any of variables $\mathrm{X_k}$ (formula (4)). Since the *p*-value is less than alpha ($p \leq \alpha$) (respectively $p_{2010;2020} < 0.001\ (5.1; 6.1)$), the author rejected the hypothesis that both directional coefficients β_1 and β_2 are zero, in favour of the alternative hypothesis that the directional coefficients are non-zero. Based on this, it was concluded that there are grounds for assuming that there is a regression linear relationship between unemployment and at least one of the explanatory variables. The obtained coefficients were interpreted as follows: if the participation rate not employed women in education and training (varible X_2) increases by 1 percentage point, unemployment will increase by about 0.476pp in 2020 if

the other variables do not change. This interpretation was applied to any variable that has a significant effect on the independent variable under study.

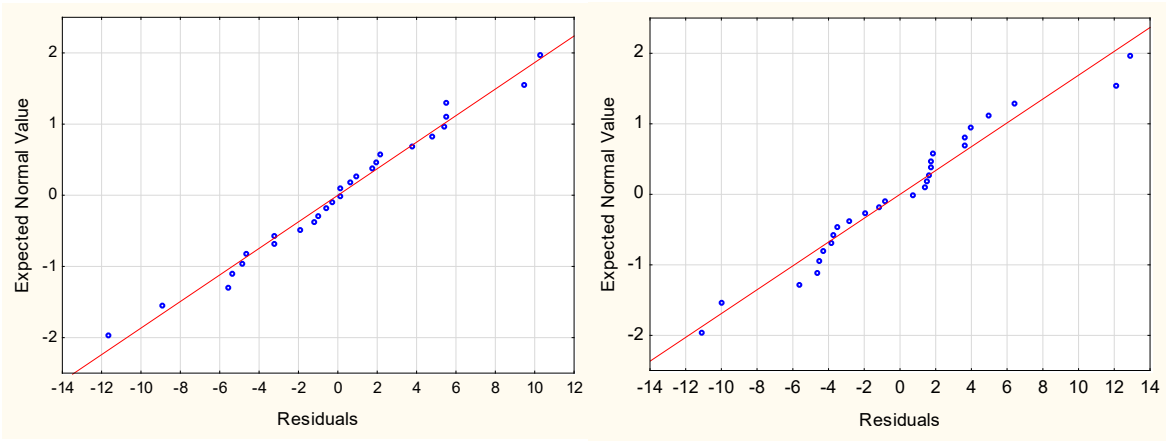

Fig. 2. Normal Probability Plot of Residuals for women by participation in education in 2010 (left) and 2020 (right)

Source: authors' work based on Eurostat database (2010 and 2020).

The resulting graph (Figure 2) provides a visual assessment of the correspondence of the residuals to the normal distribution of the EU countries studied – if the residuals do not have a normal distribution, the points will deviate from the straight line, and if the points form a clear shape around the straight line, this suggests the application of some transformation. In the case studied, the points lined up along a straight line, confirming the normality of the distribution of the residuals.

In turn, the Shapiro-Wilk and Lilliefors test was used to confirm the normality of the variables under study (see Table 3).

Table 3. Test results for normality of decomposition

Variables	Shapiro-Wilk test		Lilliefors test
Specification	Statistics	*p*-value*	*p*-value*
X_1	*W* = 0.87248	*p* = 0.00332	*p* < 0.10
X_2	*W* = 0.94234	*p* = 0.13916	*p* < 0.10
X_3	*W* = 0.94159	*p* = 0.13339	*p* < 0.10

* Assuming significance level *p* = 0.05.

Source: authors' work.

According to the test statistics, the variables have a normal distribution.

4.1.2. Unemployment of men by participation in education

In 2010, participation in formal and informal education (Figure 3) was highest for non-working men (variable X_2), and lowest for the independent variable X_1. In 2020, some changes could be seen: for variables X_1 and X_2 there was an increase, while for variable X_3 a decrease.

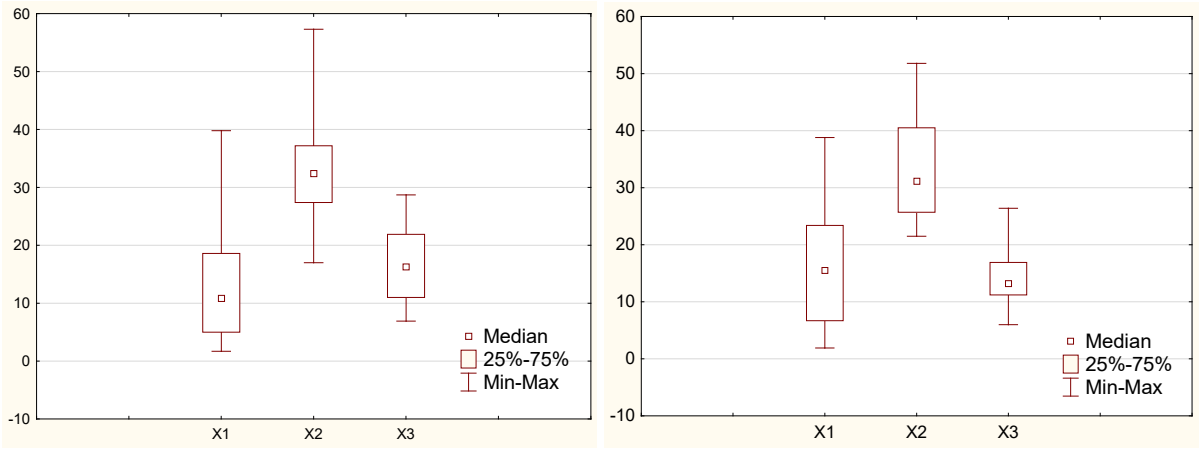

Fig. 3. Box-plot chart for "men by participation in education (including NEET)" in 2010 (left) and 2020 (right) (in thousands)

Source: authors' work based on Eurostat database (2010 and 2020).

Based on the estimated values of the parameter's coefficients (7 and 8), the relation between unemployment and education participation for men aged 20-24 in 2010 and 2020 can be described by the equation:

$$\underset{(3.749)\ (0.092)\ (0.120)}{Y_{2010} = -4.826 + 0.245X_2 + 1.117X_3}, \tag{7}$$

$$R^2 = 0.775;\ F(2,24) = 45.781;\ p < 0.0000, SE_e = 3.986, \tag{7.1}$$

$$\underset{(5.753)\ (0.101)\ (0.115)\ (0.190)}{Y_{2020} = -18.565 + 0.245X_1 + 0.430X_2 + 1.191X_3}, \tag{8}$$

$$R^2 = 0.707;\ F(3,23) = 21.942;\ p < 0.00000, SE_e = 3.837, \tag{8.1}$$

For the year 2010 (7), statistically significant variables at the significance level of 0.05 were variables X_2 and X_3, while for 2020 (8) they were all independent variables. Comparing the two models, one can see an increase in coefficients for variables X_2 (by about 0.18), X_3 (by about 0.8). The X_1 variable for 2010 is statistically insignificant. The model is well fitted, both for 2010 and 2020, as evidenced by the small standard error of the estimation: $SE_{e_{2020}} = 3.99$ and $SE_{e_{2020}} = 3.84$. The value of the coefficient of determination for 2010 is $R^2 = 0.775$ (7.1), while for 2020 $R^2 = 0.707$ (8.1), which means that 77.5% and 70.7% of the total dependent variability is explained by the model. The results of the F-test of the analysis of variance: $F_{2010} = 45.781$ (7.1), $F_{2020} = 21.942$ (8.1). According to the test, the value of $p \leq \alpha$ (respectively $p_{2010;2020} < 0.001$ (7.1; 8.1)), the author rejected the null hypothesis in favour of the alternative hypothesis that they are not equal to 0, and based on this, concluded that there are grounds for assuming that there is a regression line relationship between unemployment and at least one of the explanatory variables (the interpretation in subsection 4.2.1).

It can be seen (Figure 4) that the points lined up along a straight line, confirming the normality of the residuals of the EU countries studied.

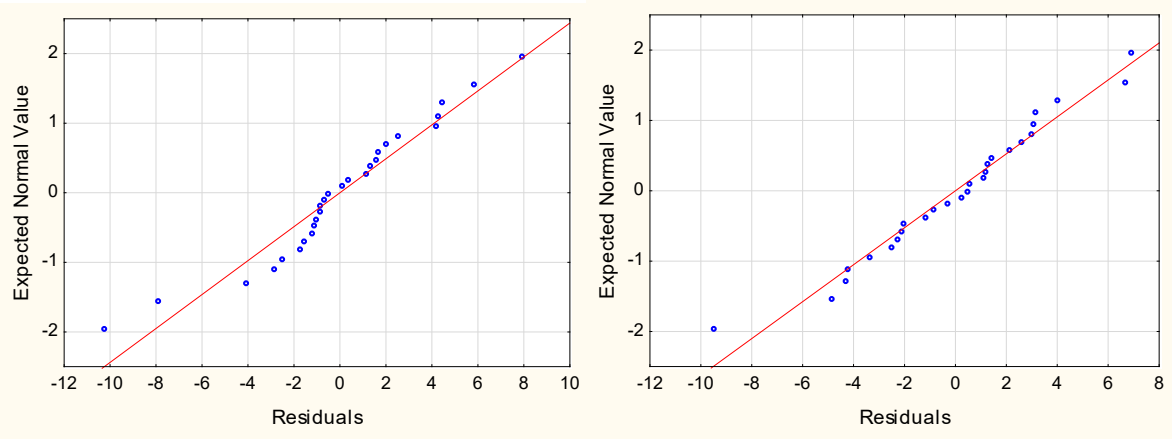

Fig. 4. Normal Probability Plot of Residuals for "men by participation in education" in 2010 (left) and 2020 (right)

Source: authors' work based on Eurostat database (2010 and 2020).

Table 4 shows the results of the normality tests of the variables studied.

Table 4. Test results for normality of decomposition

Variables	Shapiro-Wilk test		Lilliefors test
Specification	Statistics	*p*-value*	*p*-value*
X_1	*W* = 0.89175	*p* = 0.0664	*p* < 0.06
X_2	*W* = 0.94801	*p* = 0.19177	*p* < 0.20
X_3	*W* = 0.96624	*p* = 0.50619	*p* < 0.20

* Assuming significance level *p* = 0.05

Source: authors' work.

According to the test statistics, the variables have a normal distribution.

4.2. Model II – the social integration of young people

Social inclusion refers to the process by which people from different social groups take part fully in society and have equal opportunities to develop. Employment is one of the most important determinants of social inclusion, as the availability of labour and stable employment allow people to build financial stability and improve their self-esteem. Social inclusion and employment are linked, as employment allows participation in social and economic life and, at the same time, social inclusion allows better access to the labour market. The unit of measurement of variables is the number of people (in %) in households. The regression analysis was conducted based on gender breakdown.

4.2.1. Unemployment of women by social inclusion

The boxplot (Figure 5) allows to see the distribution of the variables for 2010 and 2020. For 2010, variable X_2 reached the highest values, while X_4 the lowest. The same is true for 2020, but a noticeable decrease can be seen for variable X_6.

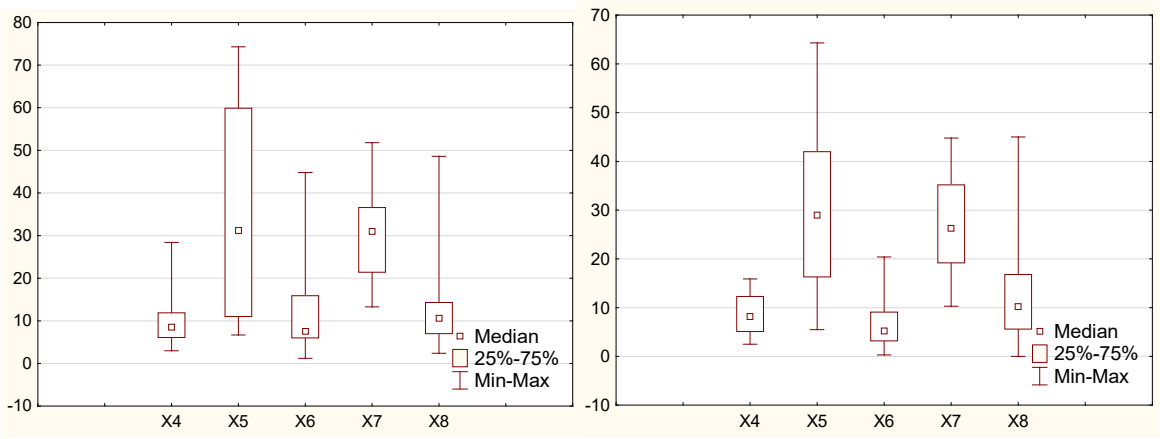

Fig. 5. Box-plot chart for "women by social inclusion" in 2010 (left and 2020 right (in thousand)

Source: authors' work based on Eurostat database (2010 and 2020).

Based on the estimated values of the parameters (9) coefficients of the relationship between unemployment and factors related to social integration of women aged 20-24, it was shown that the model for 2010 lacks statistically significant independent variables at the significance level of 0.05. In contrast, for the year 2020, a model was estimated which showed that X_6 (severe material deprivation for women) was a statistically significant variable:

$$\underset{(2.448)\ (0.295)}{Y_{2020} = 11.713 + 0.606X_6}, \tag{9}$$

$$R^2 = 0.110; F(1,25) = 4.207;\ p < 0.0001, SE_e = 2.448. \tag{9.1}$$

The model is a poor match for 2020 (9), as shown by the low value of the coefficient of determination $R^2 = 0.111$ (9.1), which means that 11.1% of the total dependent variation is explained by the model. Standard error of estimation equals $SE_{e_{2020}} = 2.448$. The low value of the R^2 coefficient of determination is described by Aczel and Sounderpandian: "when R^2 is below 0.5, the regression explains only less than 50% of the variation in Y; predictions may not be accurate; if one just wants to understand the relationships between variables, lower values of R^2 are acceptable" (Aczel and Sounderpandian, 2020, pp. 637-638). The authors noted that it is important to remember that these models then explain less of the period under study. Such a result, however, prompts further research to verify whether social inclusion among women has a significant impact on employment rates.

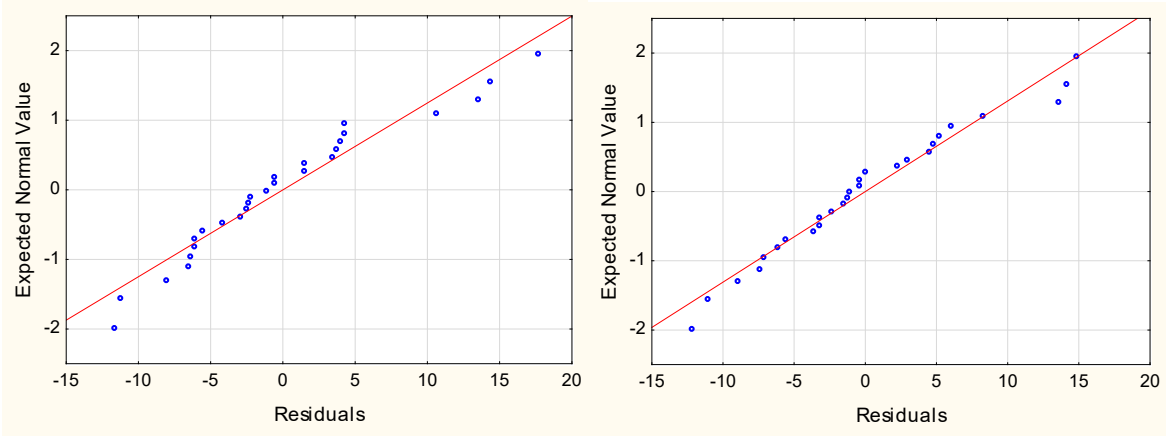

Fig. 6. Normal Probability Plot of Residuals for women by "social inclusion" in 2010 (left) and 2020 (right)

Source: authors' work based on Eurostat database (2010 and 2020).

It can be seen (Figure 6) that the points aligned along a straight line more strongly for the year 2020. However, the graph shows outlier EU countries studied, deviating from the line, which affects the normality of the residual values.

Table 5 shows the results of the normality tests of the variables studied.

Table 5. Test results for normality of decomposition

Variables	Shapiro-Wilk test		Lilliefors test
Specification	Statistics	*p*-value*	*p*-value*
X_4	*W* = 0.92970	*p* = 0.06796	*p* < 0.05
X_5	*W* = 0.94641	*p* = 0.17518	*p* < 0.20
X_6	*W* = 0.88493	*p* = 0.06155	*p* < 0.05
X_7	*W* = 0.97412	*p* = 0.71264	*p* < 0.20
X_8	*W* = 0.87063	*p* = 0.05030	*p* < 0.05

* Assuming significance level *p* = 0.05

Source: authors' work.

According to the test statistics, the variables have a normal distribution.

4.2.2. Unemployment of men by social inclusion

The distribution of the variables for 2010 shows that variable X_5 had the largest values, while X_4 the smallest. For 2020, variable X_5 had the largest value, whereas variables X_4 and X_5 the smallest.

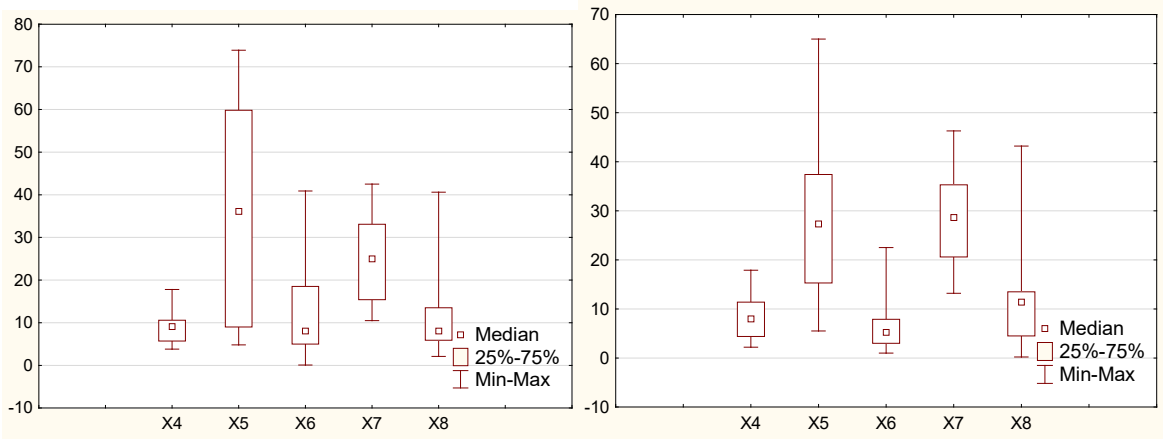

Fig. 7. Box-plot chart for men by "social inclusion" in 2010 (left) and 2020 (right) (in thousands)

Source: authors' work based on Eurostat database (2010 and 2020).

Based on the estimated values of the parameters (10 and 11) coefficients. For the 2010 model, variables X_6 and X_7 are statistically insignificant. For the 2020 model, only X_6 is statistically significant. The relation between unemployment and factors related to the social integration of men aged 20-24 in 2010 and 2020 can be described by the equation:

$$\underset{(4.113)\ (0.413)\ (0.054)\ (0.164)}{Y_{2010} = 4.914 + 1.602X_4 + 0.212X_5 - 0.416X_8}, \tag{10}$$

$$R^2 = 0.451;\ F(3,23) = 8.123\ ;\ p < 0.00072, SE_e = 6.225, \tag{10.1}$$

$$Y_{2020} = \underset{(2.090)}{12.532} + \underset{(0.232)}{0.497X_6}, \tag{11}$$

$$R^2 = 0.122;\ F(1,25) = 4.597;\ p < 0.04193, SE_e = 6.648. \tag{11.1}$$

For 2010 (10), the statistically significant variables at the 0.05 significance level are X_4 (men living in households with very low work intensity), X_5 (overcrowding rate by poverty status) and X_8 (housing cost overburden rate by poverty status), while for 2020 (11) it is X_6 (severe material deprivation rate). The model is well fitted, both for 2010 and 2020, as evidenced by the small standard error of estimation: $SE_{e_{2010}} = 6.34; SE_{e_{2020}} = 6.65$. The value of the coefficient of determination for 2010 was $R^2 = 0.451$ (10.1), while for 2020 $R^2 = 0.122$ (11.1), which means that 45.1% and 12.2% of the total dependent variability is explained by the model. The result of the F-test of the analysis of variance: $F_{2010} = 8.123$ (10.1); $F_{2020} = 4.597$ (11.1). According to the test, $p \leq \alpha$ value (respectively: $p_{2010} < 0.001$ (10.1), $p_{2020} < 0.04193$ (11.1)). For 2010, the null hypothesis was rejected in favour of the alternative hypothesis that they are not equal to 0, concluded that there are grounds for assuming that there is a regressive linear relationship between unemployment and at least one of the explanatory variables in 2010 (the interpretation in subsection 4.2.1).

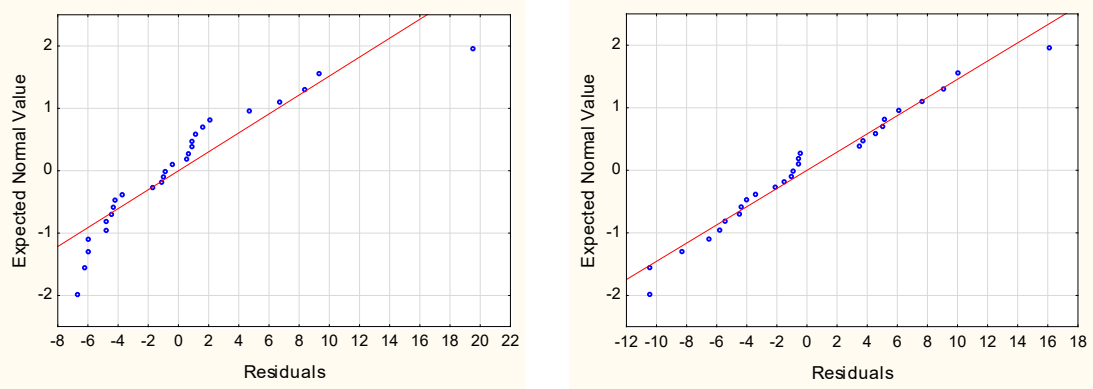

Fig. 8. Normal Probability Plot of Residuals for men by social inclusion in 2010 left and 2020 right

Source: authors' work based on Eurostat database (2010 and 2020).

It can be seen (Figure 8) that the points aligned along a straight line more strongly for the year 2020, however the graph shows outlier EU countries studied deviating from the line, which affected the normality of the residual values.

Table 6 shows the results of the normality tests of the variables studied.

Table 6. Test results for normality of decomposition

Variables	Shapiro-Wilk test		Lilliefors test
Specification	Statistics	*p*-value*	*p*-value*
X_4	*W* = 0.93472	*p* = 0.09027	*p* < 0.20
X_5	*W* = 0.90813	*p* = 0.05121	*p* < 0.05
X_6	*W* = 0.81760	*p* = 0.05811	*p* < 0.05
X_7	*W* = 0.97361	*p* = 0.69884	*p* < 0.20
X_8	*W* = 0.87741	*p* = 0.05719	*p* < 0.05

* Assuming significance level *p* = 0.05

Source: authors' work.

According to the test statistics, the variables have a normal distribution.

4.3. Model III – depending on the type of employment of young people

The most stable form of employment is the establishment of an employment contract between the employer and the employee. Due to the complexity and dynamics of the changing labour market, new forms of employment appear, and employers more willingly accept them, wanting to set up other contracts, e.g. civil law contracts, or propose self-employment. Concluding such contracts with young people may have both a positive dimension which may be desirable and acceptable for both parties (employer and employee), and a negative one which may lead to an increase in unemployment when employing a young person only and exclusively based on, among others, commissioned contracts, temporary contracts, part-time employment, self-employment (i.e. setting up a one-person business to reduce one's own costs). Types of employment can also include part-time employment for young people who would like to work full-time but could not find such work on the labour market, also known as involuntary part-time employment.

4.3.1. Unemployment of women by "type of employment"

In 2010, variable X_{12} (involuntary part-time employment as a percentage of the total part-time employment for young people) had the highest level, while in 2020 there was a decrease in variable X_{12} and an increase in variable X_{10} (young temporary employees as a percentage of the total number of employees) can be observed (Figure 9).

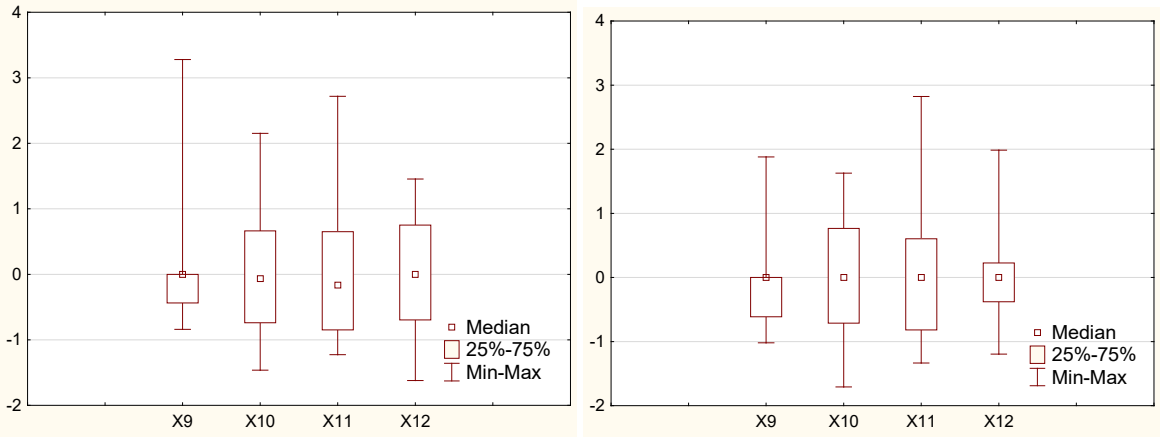

Fig. 9. Box-plot chart for women by type of employment in 2010 (left) and 2020 (right) (in thousands)

Source: authors' work based on Eurostat database (2010 and 2020).

Based on the estimated values of the parameters (12 and 13) coefficients, the relation between unemployment and type of employment for women aged 20-24 in 2010 and 2020 can be described by the equation:

$$\underset{(0.161)\ (0.192)}{Y_{2010} = -0.0000 + 0.662X_{12}}, \tag{12}$$

$$R^2 = 0.293; F(1,25) = 11.788;\ p < 0.00209, SE_e = 0.840, \tag{12.1}$$

$$\underset{(0.142)\ (0.197)}{Y_{2020} = 0.0000 + 0.936X_{12}}, \tag{13}$$

$$R^2 = 0.452;\ F(1,25) = 22.416;\ p < 0.00007, SE_e = 0.740, \tag{13.1}$$

For 2010 (12) and 2020 (13), the statistically significant independent variables at the significance level of 0.05 is variable X_{12} and thus involuntary part-time employment as a percentage of total part-time employment of young people. Comparing the two models, one can see an increase in the coefficient for variable X_{12} (by about 0. 274) from 2010 to 2020. The model is well fitted, both for 2010 and 2020, as evidenced by the small standard error of the estimation: $SE_{e_{2010}} = 0.840$ and $SE_{e_{2020}} = 0.740$. The value of the coefficient of determination for 2010 is $R^2 = 0.293\ (12.1)$, while for 2020 $R^2 = 0.452\ (13.1)$, which means that 29.3% and 45.2% of the total dependent variability is explained by the model. The result of the F -test of the analysis of variance: $F_{2010} = 11.788\ (12.1)$, $F_{2020} = 22.416\ (13.1)$. According to the test, the value of $p \leq \alpha$ (respectively $p_{2010} < 0.00209\ (12.1)$, $p_{2020} < 0.001\ (13.1)$), the null hypothesis was rejected in favour of the alternative hypothesis that they are not equal to 0, concluding that there are grounds for assuming that there is a regression linear relationship between unemployment and at least one of the explanatory variables (the interpretation in subsection 4.2.1).

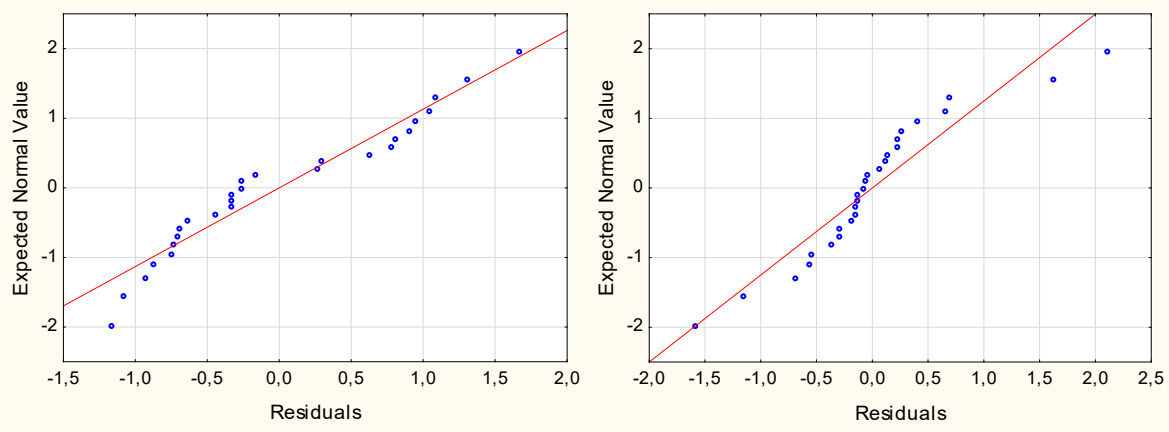

Fig. 10. Normal Probability Plot of Residuals for women by "type of employment" in 2010 (left) and 2020 (right)

Source: authors' work based on Eurostat database (2010 and 2020).

It can be seen (Figure 10) that the points aligned along a straight line, confirming the normality of the EU countries studied. Outlier EU countries are visible in the figure, but they are slightly deviated from the line, which does not significantly affect the normality of the residuals.

Table 7 shows the results of the normality tests of the variables studied.

Table 7. Test results for normality of decomposition

Variables	Shapiro-Wilk test		Lilliefors test
Specification	Statistics	*p*-value*	*p*-value*
X_9	*W* = 0.86727	*p* = 0.06897	*p* < 0.05
X_{10}	*W* = 0.97219	*p* = 0.66035	*p* < 0.20
X_{11}	*W* = 0.93183	*p* = 0.07664	*p* < 0.20
X_{12}	*W* = 0.89210	*p* = 0.05533	*p* < 0.05

* Assuming significance level *p* = 0.05

Source: authors' work.

According to the test statistics, the variables have a normal distribution.

4.3.2. Unemployment of men by "type of employment"

For both 2010 and 2020 it can be seen (Figure 11) that variable X_2 was the highest, while variable X_1 was the lowest for 2010 and variable X_4 for 2020.

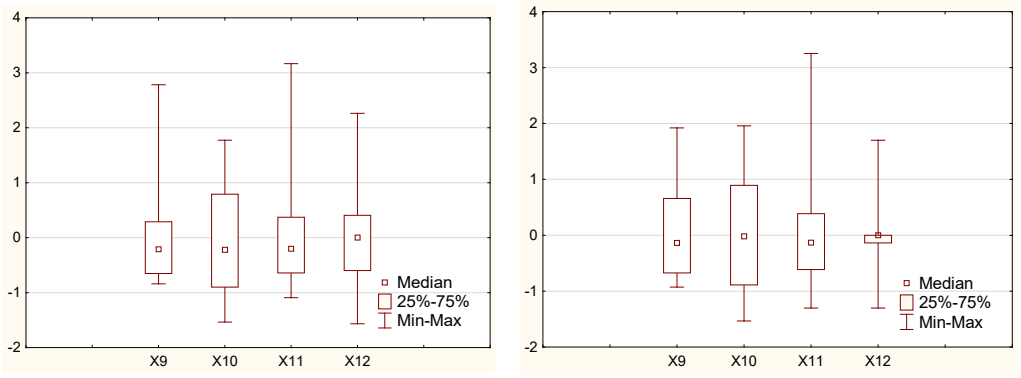

Fig. 11. Box-plot chart for men by "type of employment" in 2010 (left) and 2020 (right) (in thousands)

Source: authors' work based on Eurostat database (2010 and 2020).

Based on the estimated values of the parameters (14) coefficients of the relation between unemployment and type of employment for men aged 20-24, it was shown that the model for 2010 lacks statistically significant independent variables at the significance level of 0.05. In contrast, for the 2020, the model was estimated, which showed that X_9, X_{10} and X_{12} was a statistically significant variable:

$$Y_{2020} = \underset{(0.158)}{-0.0000} \underset{(0.219)}{- 0.543X_9} \underset{(0.186)}{+ 0.475X_{10}} \underset{(0.262)}{+ 0.879X_{12}}, \tag{14}$$

$$R^2 = 0.319; F(3,23) = 5.0609;\ p < 0.00775, SE_e = 0.825. \tag{14.1}$$

The model is well fitted for 2020 (14), as shown by the small standard error of estimation ($SE_{e_{2020}} = 0.825$). The value of the coefficient of determination is $R^2 = 0.319$ (14.1) which means that 31.9% of the total dependent variability is explained by the model. The results of the F-test of the analysis of variance is $F_{2020} = 5.0609$ (14.1). According to the test, $p \leq \alpha$ value ($p_{2020} < 0.007$(14.1)), so for 2020 the null hypothesis was rejected in favour of the alternative hypothesis that it is not null, concluding that there are grounds for assuming that there is a regression linear relationship between unemployment and at least one of the explanatory variables (the interpretation of the coefficients is the same as in subsection 4.2.1 for women).

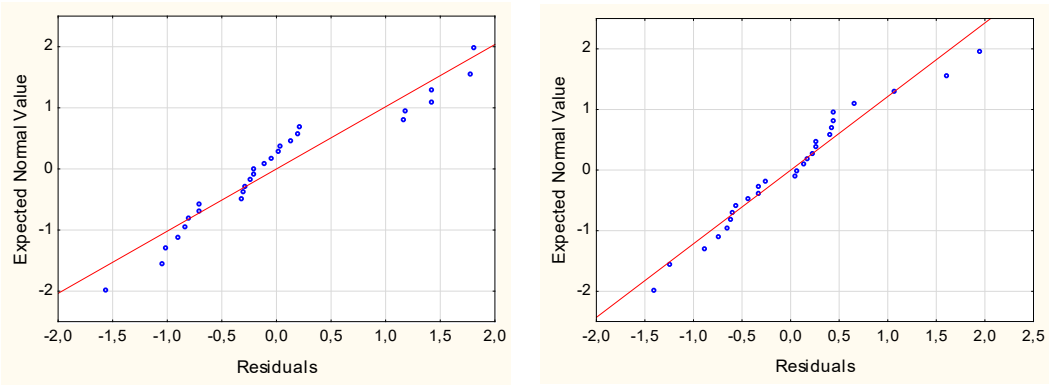

Fig. 12. Normal Probability Plot of Residuals for men by "type of employment" in 2010 (left) and 2020 (right)

Source: authors' work based on Eurostat database (2010 and 2020).

It can be seen (Figure 12) that the points aligned along a straight line, confirming the normality of the EU countries studied. Outlier EU countries are visible in the graph, but they are slightly deviated from the line, which does not significantly affect the normality of the residuals.

Table 8 shows the results of the normality tests of the variables studied.

Table 8. Test results for normality of decomposition

Variables	Shapiro-Wilk test		Lilliefors test
Specification	Statistics	*p*-value*	*p*-value*
X_9	*W* = 0.85020	*p* = 0.05116	*p* < 0.05
X_{10}	*W* = 0.96711	*p* = 0.52748	*p* < 0.20
X_{11}	*W* = 0.88645	*p* = 0. 06644	*p* < 0.10
X_{12}	*W* = 0.85815	*p* = 0.05114	*p* < 0.05

* Assuming significance level *p* = 0.05

Source: authors' work.

According to the test statistics, the variables have a normal distribution.

5. Conclusion

Unemployment is one of the most important economic and social problems which also constitutes one of the largest measures characterising the condition of the economy. The diversity and dynamics of changes in the labour market affect young people, who in turn make up an individual group of people entering the labour market. Unemployment is influenced by factors of a socio-economic and demographic nature, among others. The variables studied were of socio-economic as well as demographic in nature. The age range 20-24 was considered, as well as the gender breakdown (male, female) and by category. The dependent variable in the study was the level of youth unemployment.

Unemployment in terms of participation in education was influenced by variable X_2 (participation rate "not employed persons in education and training") and variable X_3 (participation rate "not employed persons in education and training (including NEET rates)") in 2010 and 2020. However, it can be noted that in 2020 variable X_1, i.e. participation rate "employed persons in education and training" significantly influenced the level of youth unemployment. In the case of social inclusion of women and men, the severe material deprivation index significantly influenced the unemployment rate in 2020. It should be noted, however, that for women in 2010 no variable had a significant impact, while for men these were as many as three variables X_4 (people living in households with very low work intensity), X_5 (overcrowding rate by poverty status) and X_8 (housing cost overburden rate by poverty status). For the type of employment of young women, variable X_{12} (involuntary part-time employment as percentage of the total part-time employment) was the significantly influential variable in 2010 and 2020, where for men, none of the variables examined had an impact in 2010. This situation changed in 2020, when variables X_9 (self-employment by gender, age, and educational attainment level), X_{10} (temporary employees as percentage of the total number of employees) and X_{12} affect male unemployment. Based on the research, it can be noted that in Model I for 2010 women, only one variable, X_1 (participation rate "employed persons in education and training") did not significantly affect unemployment, and the value of the coefficient of determination alone accounted for 69.4% of the total variation of the dependent. The model for 2020 showed 52.6% of the total variation in the dependent, but all the variables affected unemployment for young women in terms of participation in education. The same was true for men, where the model for 2010 showed 81.4% of the total dependent variation (X_1 also did not significantly affect it), while the model for 2020 showed 74.1% (all the variables significantly affected unemployment). It can be seen that education has an extremely

significant impact on youth unemployment, with both men and women experiencing an increase in unemployment (all the coefficients were positive, according to which, if independent variable X increased by 1, dependent variable Y increased by a given value on average). Regarding the integration of young women in 2010, none of the studied variables significantly affected unemployment, but for 2020, variable X_6, i.e. "severe material deprivation" rate (the coefficient of determination explains only in 20% the total variation of the dependent). For men, the situation was different. In 2010, variables X_4 (men living in "households with very low work intensity"), X_5 (overcrowding rate by poverty status), X_8 ("housing cost overburden" rate by poverty status) had a significant impact with about 54% being the total variation of the dependent, where only variable X_8 indicates that if its value increases by one unit then unemployment will fall. In contrast, in 2020, only variable X_6 ("severe material deprivation" rate) affected youth unemployment, where the coefficient of determination makes up 38.3% of the total dependent variable. Hence, one can say that an increase in poverty promotes an increase in unemployment. Finally, Model III for both periods studied, the variable significantly affecting unemployment was X_{12}, namely "involuntary part-time employment" as a percentage of the total part-time employment (the coefficient of determination was, respectively, for 2010 – 35.3%, for 2020 – 57.2%), where an increase in the variable by one unit caused an increase in unemployment. For men, only the model for 2020 (in around 45%), where the significant variables were X_9 (self-employment by "educational attainment" level), X_{10} (temporary employees as percentage of the total number of employees) and X_{12} caused an increase in unemployment, whereas an increase in variable X_9 by one unit made unemployment decrease. Based on this, it can be concluded that participation in education is the most principal factor affecting youth unemployment. The lack of participation in education increases the risk of social exclusion, and of material deprivation, which leads to poverty among young people. The research in terms of education and other factors needs to be deepened to be able to draw detailed conclusions to identify the cause of the increase in unemployment.

Answering the research questions, one can state that in the majority of European Union countries the unemployment rate has decreased, it is worth noting that out of the 27 countries included in the study, a decrease was recorded in 19 countries, while in the remaining 8, an increase, namely in Greece, Italy, Portugal, Luxembourg, Finland, Cyprus, Austria, and Malta (the highest was in Luxembourg, by 8.1 percentage points). All the changes depend on the diversity, but also on the dynamics of these changes in the given economic markets. The reasons for this state of affairs include, first of all, the fact that young people enter the market for the first time or move precariously on the labour market being employed on the basis of, among others, civil law contracts or simply lose their jobs quickly due to constant changes on the labour market. It is worth noting that the Baltic states in 2020 experienced the largest decrease in the unemployment rate (Estonia (by 14.2 percentage points), Lithuania (by 14.5 percentage points) and Latvia (by 19.7 percentage points). It should also be noted that there was a tendency for higher unemployment in South-Eastern Europe than in North-Western Europe, both in 2010 and 2020.

Given this and the precariousness of employment as shown by the last model, the only way to break this situation is through employment, which however is not so easy. Lack of employment at the beginning of a professional career results in lower chances of finding a job and adequate wages in acquiring professional competences in later life. Young people are at the centre of social attention because they are members of households. Therefore, this is a vulnerable social group to which all projects related to self-development, education but also family-friendly policy instruments are addressed. The greatest risk is faced by the group of people from the NEET category, i.e. young people not in education, training or employment, who are or may be in the near future the subjects of social programmes, which involves receiving social benefits, which will then burden the state budget and thus lead to fluctuations on economic markets. Hence, special attention should be paid to the economic and material situation of young people, who are affected by poverty, suffer from social exclusion and, above all, experience the negative effects of material deprivation, where the last group proved to have a significant impact in 2020 on both women and men. Exclusion from the labour market,

education system, social integration, as well as the type of employment of young people, entails negative consequences not only related to unused human capital (potential), but also related to unwillingness, and the indifference of young people. Special attention should also be paid to limiting possibilities of self-realisation, increasing, or extending professional qualifications, which relates to the lack of possibilities to earn money. All these negative aspects translate into the everyday life of young people, and negative effects on health, in particular, their mental health. Social exclusion can lead to anti-social behaviour or social migration. The study also showed that involuntary part-time employment results in increased levels of unemployment, if young people cannot find a job with an employment contract. Such a system of work lowers their self-satisfaction, causing an increased reluctance to continue looking for a job or to pursue personal and professional self-development. The above summary answers the research questions posed in the introduction.

Youth unemployment results in numerous economic and social consequences: the failure to exploit one's potential, the desire for self-development and opportunities, reflected in a lower economic potential, which also means the failure to use available human capital resources. This is why it is so important to understand the determinants affecting unemployment levels. Further in-depth analyses in this area and the identification of the factors influencing employment and unemployment of young people will provide more detailed knowledge of the behaviour of this social group on the labour market.

References

Bell, D. N., and Blanchflower, D. G. (2011). Young People and the Great Recession. *Oxford Review of Economic Policy*, *27*(2), 241-267.

Abbott, L. F. (2013). Theories of the Labour Market and Employment: A Review. *Industrial Systems Research*, 1-2.

Aczel, A. D., and Sounderpandian, J. (2017). *Statystyka w zarządzaniu*. Warszawa: Wydawnictwo Naukowe PWN.

Aczel, A. D., and Sounderpandian, J. (2020). *Statystyka w zarządzaniu*. Warszawa: Wydawnictwo Naukowe PWN.

Algan, Y., and Cahuc, P. (2009). Civic Culture and Labor Market Institutions. *American Economic Journal: Macroeconomics*, *1*(1), 111-145.

Arntz, M., Gregory, T., and Zierahn, U. (2019). Digitalization and the Future of Work: Macroeconomic Consequences. *IZA World of Labor*, 1-21.

Balicki, A. (2013). *Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*. Chicago: University of Chicago Press.

Bentolila, S., Cahuc, P., Dolado, J. J., and Le Barbanchon, T. (2012). Two-Tier Labour Markets in the Great Recession: France vs. Spain. *Economic Journal*, *122*(562), 155-187.

Blanchard, O., and Wolfers, J. (2000). The Role of Shocks and Institutions in the Rise of European Unemployment: The Aggregate Evidence. *The Economic Journal*, *110*(462), C1-C33.

Boeri, T. (2002). *Let Social Policy Models Compete and Europe Will Win. Conference on The Role of Unions in the Twenty-First Century: A Debate between Tony Blair and Gerhard Schröder*.

Boeri, T., and Garibaldi, P. (2007). Two Tier Reforms of Employment Protection: A Honeymoon Effect? *Economic Journal*, *117*(521), 357-385.

Boeri, T., and van Ours, J. (2008). *The Economics of Imperfect Labor Markets*. Princeton University Press.

Burda, M., and Wyplosz, C. (2000). *Makroekonomia – podręcznik europejski*. Warszawa: Wydawnictwo Naukowe PWN.

Cahuc, P., Postel-Vinay, F., and Robin, J. M. (2006). Wage Bargaining with on-the-job Search: Theory and Evidence. *Econometrica*, *74*(2), 323-364.

Ciccone, A., and Papaioannou, E. (2009). Human Capital, the Structure of Production, and Growth. *The Review of Economics and Statistics*, *91*(1), 66-82.

Conover, W. J. (1999). *Practical Nonparametric Statistics* (3rd ed.). Wiley: New York.

Cormack, R. M. (1971). Journal of the Royal Statistical Society. Series A (General). *A Review of Classification*, 134, *3*(1971), 321-367. https://doi.org/10.2307/2344237

Dagnelie, P. (1975). *Analyse statistique ŕ plusieres variables*. Gembloux: Les Presse Agronomique.

Eichhorst, W., Hinte, H., and Rinne, U. (2013). Youth Unemployment in Europe: What to Do about It? *Intereconomics*, *48*(4), 230-235. Retrieved September 15, 2023 from https://docs.iza.org/pp65.pdf

Eichhorst, W., and Marx, P. (2011). Reforming German Labour Market Institutions: A Dual Path to Flexibility. *Journal of European Social Policy*, *21*(1), 73-87.

Elsby, M. W., Hobijn, B., and Şahin, A. (2015). On the Importance of the Participation Margin for Labor Market Fluctuations. *Journal of Monetary Economics*, *72*, 64-82.

European Commission. (n.d.). *Europejski semestr – zestawienie informacji tematycznych. Zatrudnienie ludzi młodych*. Retrieved August 20, 2021 from https://ec.europa.eu/info/sites/info/files/file_import/european-semester_thematic-factsheet_youth_employment_pl.pdf

European Statistical Office Eurostat. (n.d.). Retrieved August 20, 2021 from https://ec.europa.eu/eurostat/data/database

Eurostat. (n.d.). *Statistics Explained, Glossary: Young People Neither in Employment Nor in Education and Training (NEET)*. Retrieved August 20, 2021 from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Young_people_neither_in_employment_nor_in_education_and_training_(NEET)

Fergusson, R., and Yeates, N. (2021). Endemic Youth Unemployment: A Social Policy Issue. *Sociology, Social Policy and Education*, 22-37.

Fisher, R. A. (1970). Statistical Methods for Research Workers. In Kotz, S., Johnson, N.L. (Eds.), *Breakthroughs in Statistics*. Springer Series in Statistics. New York: Springer. https://doi.org/10.1007/978-1-4612-4380-9_6, 66-70

Francuz, P., and Mackiewicz, R. (2005). *Liczby nie wiedzą, skąd pochodzą*. Lublin: Wydawnictwo KUL.

García-Pérez, J. I., Marinescu, I., Vall Castelló, J., and Hernanz, V. (2018). *Can Fixed-Term Contracts Put Low Skilled Youth on a Better Career Path? Evidence from Spain* (NBER Working Paper No. 24842, 1-52). Retrieved September 15, 2023 from extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.nber.org/system/files/working_papers/w22048/w22048.pdf

Głód, S., and Miotła, P. (2014). Młodzież jako grupa społeczna w ujęciu definicji naukowych. *Trendy ve vzdělávání*, *7*(1), 323-327.

Główny Urząd Statystyczny (Statistics Poland). (n.d.). *Metainformacje (Słownik pojęć).* Warszawa: GUS.

Gregg, P., and Petrongolo, B. (2005). Stock-Flow Matching and the Performance of the Labor Market. *European Economic Review*, *49*(8), 1981-2005.

Guiso, L., Sapienza, P., and Zingales, L. (2006). Does Culture Affect Economic Outcomes? *The Journal of Economic Perspectives*, *20*(2), 23-48.

Hanushek, E. A., Woessmann, L., and Zhang, L. (2020). General Education, Vocational Education, and Labor-Market Outcomes over the Life Cycle. *Journal of Human Resources*, *55*(1), 241-279.

Hanusz, Z., and Tarasińska, J. (2012). *O testach wielowymiarowej normalności opartych na statystyce Shapiro-Wilka*. Uniwersytet Przyrodniczy w Lublinie: Katedra Zastosowań Matematyki i Informatyki.

ILO. (May, 2012). *Global Employment Trends for Youth 2012*. Geneva: International Labour Organization (ILO).

Kahanec, M., and Zimmermann, K. F. (Ed.). (2016). *Labor Migration, EU Enlargement, and the Great Recession*. Springer.

Kryńska, E., and Kwiatkowski, E. (2013). *Podstawy wiedzy o rynku pracy*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Lambovska, M., Sardinha, B., and Belas, J. (2021). Impact Of the Covid-19 Pandemic on Youth Unemployment in The European Union. Žilinska univerzita v Žiline, Fakulta prevádzky an ekonomiky dopravy a spojov, Katedra ekonomiky. *Ekonomicko--manazerske spektrum*, *15*(1), 55-63.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979). *Multivariate Analysis*. London: Academic Press. https://doi.org/10.1002/bimj.4710240520

Maxwell, A. E. (1997). *Multivariate Analysis in Behavioural Research*. London: Chapman & Hall.

Morrison, D. M. (1990). *Wielowymiarowa analiza statystyczna*. Warszawa: PWN.

Mussida, CH., and Sciulli, D. (2023). Being Poor and being NEET in Europe: Are These Two Sides of the Same Coin? *The Journal of Economic Inequality*. Received: 7 September 2022/ Accepted: 9 December 2022. Springer Science+Business Media, LLC, part of Springer Nature 2023. https://doi.org/10.1007/s10888-022-09561-7.

Nienaber, B., Manafi, I., Vysotskaya, V., Roman, M., and Marinescu, D. (2020). Challenging Youth Unemployment Through International Mobility. *Journal of Social and Economic Statistics*, *9*(1), 5-27. doi: 10.2478/jses-2020-0002

Nowak-Brzezińska, A. (2018). Feature Selection Approach for Rule-Based Knowledge Bases. In *Advances in Feature Selection for Data and Pattern Recognition*. Cham: Springer.

Oesingmann, K. (2017). *Youth Unemployment in Europe* (CESifo DICE Report, *15*(1), 52-55).

O'Higgins, N. (2012). This Time It's Different? Youth Labour Markets During 'The Great Recession'. *Comparative Economic Studies*, *54*(2), 395-412.

Panek, T. (2009). *Statystyczne metody wielowymiarowej analizy porównawczej*. Warszawa: Wydawnictwo Szkoły Głównej Handlowej w Warszawie. Retrieved from https://www.researchgate.net/publication/267241683_Statystyczne_metody_wielowymiarowej_analizy_porownawczej_Teoria_i_zastosowania

Pennoni, F., and Bal-Domańska, B. (2022). NEETs and Youth Unemployment: A Longitudinal Comparison Across European Countries. *Social Indicators Research*, *162*, 739-761. https://doi.org/10.1007/s11205-021-02813-5

Piecuch, T., and Piecuch, M. (2014). Analiza sytuacji młodych ludzi na rynku pracy – rozważania teoretyczne i badania empiryczne. *Modern Management Review*, XIX, *21*(4), 175-186.

Pilch, T. (2004). *Encyklopedia pedagogiczna XXI wieku*. T. III. Warszawa: Wydawnictwo Akademickie Żak.

Piszczatowska-Oleksiewicz, M. (2017). Odraczanie wyprowadzki z rodzinnego gniazda – konieczność czy strategia? Analiza zjawiska w Polsce. *Soc. Communitas*, *2*, 177-188.

Psacharopoulos, G., and Patrinos, H. A. (2004). Returns to Investment in Education: A Further Update. *Education Economics*, *12*(2), 111-134.

Royston, P. (1992), Approximating the Shapiro-Wilk W Test for Non-Normality. *Statistics and Computing*, *2*, 117-119.

Shapiro, S. S., and Wilk, M. B. (1965). An Analysis of Variance Test for Normality. *Biometrika*, *52*(3/4), 591-611.

Stanisz, A. (2007). *Przystępny kurs statystyki z zastosowaniem Statistica PL na przykładach z medycyny*. T.1-3. Kraków: StatSoft Polska.

Statistica software (version 13). (n.d.). *Licence Provided by the University of Economics in Katowice*. Retrieved from https://www.statsoft.pl/Programy/Ogolna-charakterystyka/Platforma-Statistica/

Strahl, D. (Ed.). (1998). *Taksonomia struktur w badaniach regionalnych*. Wrocław: Wydawnictwo Akademii Ekonomicznej we Wrocławiu.

Sulich, A., Rutkowska, M., and Popławski, Ł. (2020). Green Jobs, Definitional Issues, and the Employment of Young People: An Analysis of Three European Union Countries. *Journal of Environmental Management*, *Science Direct*, *262*, 110314, 1-10.

Szluz, B. (2017). O dorosłych dzieciach mieszkających z rodzicami (na przykładzie wybranych państw). *Soc. Communitas*, *2*, 73-88.

Trzpiot, G., and Kawecka, M. (2021a). Description of the Labour Market Status of Young People in Selected Countries of the European Union – The Taxonomic Approach. *Econometrics. Ekonometria. Advances in Applied Data Analysis*, *25*(4), 17-39.

Trzpiot, G., and Kawecka, M. (2021b). Evaluation of the Labor Market Status of Young People in Selected Countries of the European Union – The Multiple Regression Approach. In G. Trzpiot (Ed.), *Modeling of Complex Data Sets and Risk Analysis* (pp. 46-73). Katowice: Publishing House of the University of Economics in Katowice.

Trzpiot, G. (Ed.). (2015). *Modelowanie wielowymiarowych struktur danych i analiza ryzyka*. Katowice: Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach.

UNESCO UIS. (2013). *Global*. Retrieved August, 20, 2021 from http://uis.unesco.org/en/glossary-term/non-formal-education

Wawrzyniak, K., Bak, I., Cheba, K., and Oesterreich, M. (2020). The Similarity of European Union Countries in Terms of the Structure of the Unemployed. *European Research Studies Journal*, *23*(4), 416-429.

World Youth Report 2003. (2004). *The Global Situation of Young People*. New York: Department of Economic and Social Affairs. United Nations.

Ocena sytuacji młodych ludzi na rynku pracy w krajach Unii Europejskiej – podejście regresji wielorakiej

Streszczenie: W artykule podjęto rozważania dotyczące problemów młodych ludzi w wieku 20-24 lata dotkniętych bezrobociem na rynku pracy w krajach Unii Europejskiej. Bezrobocie jest jednym z najważniejszych problemów gospodarczo-społecznych, który jednocześnie stanowi jeden z największych mierników charakteryzujących kondycję gospodarki. Zróżnicowanie sytuacji gospodarczej krajów UE wpływa bezpośrednio na młodych ludzi, stanowiących indywidualną grupę osób, które wchodzą na rynek pracy. Posiadają oni małe doświadczenie zawodowe lub zupełnie im go brakuje. Jednocześnie są to osoby gotowe do podjęcia pracy, borykające się z dużymi trudnościami z wejściem na rynek. Wpływają na to czynniki społeczno-gospodarcze, a także demograficzne. Celem artykułu jest określenie czynników determinujących bezrobocie młodych ludzi w wieku 20-24 lata w krajach Unii Europejskiej. Badanie zostało przeprowadzone w latach 2010 i 2020 z wykorzystaniem regresji wielorakiej. Dane statystyczne zaczerpnięto z baz danych Eurostatu. Z badania wynika, że regresja wieloraka wykazała istotny wpływ na badane bezrobocie czynników dotyczących uczestnictwa młodych ludzi w edukacji i szkoleniu się (w tym również wskaźnik NEET) względem statusu na rynku pracy, a także integracji społecznej. W ciągu dekady zaobserwowano spadek bezrobocia w większości krajów członkowskich – aż w 19 krajach. W pozostałych 8 krajach widoczny jest wzrost.

Słowa kluczowe: rynek pracy, Unia Europejska, młodzi ludzie, bezrobocie, regresja wieloraka.