Romanian Statistical Review
Revista Română de Statistică

THE JOURNAL OF NATIONAL INSTITUTE OF STATISTICS

4 /2023

www.revistadestatistica.ro

CONTENTS 4/2023

Labor market outcomes of migrants at the times of the covid-19 pandemic and geopolitical uncertainty

Prof. Dr. Graţiela Georgiana NOJA (gratiela.noja@e-uvt.ro)
West University of Timisoara, Faculty of Economics and Business Administration, Department of Marketing and International Economic Relations, Timisoara, Romania

Prof. Dr. Mirela CRISTEA[1] (simona.cristea@edu.ucv.ro)
University of Craiova, Faculty of Economics and Business Administration, Department of Finance, Banking and Economic Analysis, Craiova, Romania

Prof. Dr. Ciprian PÂNZARU (ciprian.panzaru@e-uvt.ro)
West University of Timisoara, Faculty of Sociology and Psychology, Department of Sociology, Timisoara, Romania

Dr. Sorin BELEA (sorin.belea@timis.insse.ro)
National Statistics Institute

Dr. Laura Maria JIANU (maria.jianu@timis.insse.ro)
National Statistics Institute

Ph.D. Candidate Irina-Maria GRECU (irina.grecu96@e-uvt.ro)
Doctoral School of Economics and Business Administration, West University of Timisoara, Timisoara, Romania,

ABSTRACT

This research explores the impact of the Covid-19 pandemic and geopolitical threats on migrants' labour market outcomes across the European Union and the additional pressure exerted upon the European labour markets accordingly. The methodology embeds network analysis through Gaussian graphical models, applied to a newly compiled dataset for 2021. Main results entail the adverse impact of the Covid-19 pandemic and geopolitical risks on the labour market outcomes of all workers, disproportioned and more profound in the initial stages and for sectors where migrants have a strong overrepresentation and that were most affected by the pandemic and less keen in EU countries that implemented substantial job retention schemes during

1. corresponding author

the pandemic and beyond. The findings mainly highlight some of the most important effects induced by the Covid-19 pandemic as regards the severe disruption of migrants' access to employment opportunities. Several migration policies and strategic guidelines are presented and largely discussed within the paper.
Keywords: *migrants, Covid-19 pandemic, geopolitics, unemployment, econometric modelling.*
JEL Classification: *F22, F63, R23.*

1. INTRODUCTION

The international migration of human capital, specifically labour migration, constitutes a multifaceted phenomenon, exerting substantial economic repercussions on both the countries of origin and destination, as well as on the migrants themselves. Variations in the volume of migrants, along with shifts in the underlying rationales, motivations, and perspectives influencing the decision to emigrate, present causes for concern at the host-country level. These disparities manifest in detrimental effects on the resident population, state finances, logistics, medical infrastructure, and the deployment of human resources engaged in the accommodation and assimilation of newcomers into the host nation. This involvement, both in terms of quantity and quality, is of considerable import (Franc et al., 2019).

Furthermore, there is a crucial emphasis on measures directed at ensuring and facilitating immigrants' entry into the labour market, supported by initiatives such as skill development programs, language courses, and collaborative activities that foster interaction between newly arrived individuals and native citizens (Gabrielli & Impicciatore, 2022).

A paramount contemporary challenge confronting European states, as hosts to millions of foreign residents, arises from the surges of migrants, with the geographic concentration of these flows showing a marked and escalating trend in recent years (Libanova, 2019). These large-scale population movements have been catalysed by an array of external events, with reverberating impacts on both source and destination countries. Notable among these catalysts are the Covid-19 pandemic and the ongoing conflict between Russia and Ukraine. The host country's economic landscape has been substantially affected, with millions of jobs vanishing in the wake of the war, rendering numerous sectors exceptionally scarce or obsolete (Dluhopolskyi et al., 2019).

Consequently, a pivotal imperative lies in retraining the remaining Ukrainian workforce, streamlining their transition into sectors anticipated to be more active in the post-war period, such as construction, engineering, healthcare, and information technology (Economic and Policy Research Center, 2022). Concurrently, the migrant crisis is compounded by the multifaceted

needs and care requirements of immigrants, including many who seek political asylum. Following their entry into the host country, the provision of primary care and rapid response to emergencies, often necessitating the engagement of all available competent forces within the host nation, is imperative. This assistance frequently extends to the provision of psychological support for the immigrants and their families. Consequently, the pressure exerted by waves of immigrants on host countries is substantial and resource-intensive. The financial commitments required for providing temporary housing, sustenance, sanitary products, education, language acquisition, and vocational training, along with administrative and human resources, often surpass the state's capacity to cope (Mulska et al., 2021).

The effective operation of life and the labour market within host states, particularly within the context of the crisis situations delineated, necessitates the formulation and implementation of well-crafted policies and practices. These should be designed to foster the integration of immigrants and their equitable integration into the labour market, without undermining the prospects of both migrants and native citizens in accessing specific professions and employment opportunities.

In this complex framework, current research aims to dive deep into the labour market outcomes of migrants across the European Union (EU) and the additional pressure exerted upon the European labour markets, accordingly. Using a newly compiled dataset with recent indicators about the labour market insertion and performance of migrants in EU member states extracted from Eurostat (European Commission, 2023), network analysis is employed through Gaussian graphical models (GGMs) estimated by the Extended Bayesian information criteria with graphical least absolute shrinkage and selection operator (EBICglasso), as well as by the partial correlations (PCOR) methods. Different from other studies, current research provides new empirical evidence on the linkages and inferences between international labour migration and the economic development/welfare and labour market performance in EU countries, by accounting for migrants' labour market outcomes in the challenging geopolitical and Covid-19 pandemic context.

The remaining parts of the paper comprise distinctive sections that cover, firstly, a brief literature review on the subject of international migration flows, shaped by the Covid-19 pandemic and geopolitical crises. Therewith, the data and methodology section details the indicators considered for the analysis with their descriptive statistics, along with the methodology applied, namely, GGMs. The results and discussion follow next which leads to representative conclusions in the last section of the paper. Additional information that adds consistency to our results is presented in the Appendix section.

2. BRIEF LITERATURE REVIEW

Migration is a movement of human capital, based on a series of reasons and motivations, which can be proactive and reactive (Richmond, 1993).

Proactive motivation depends on each individual, on what a person wants to achieve in terms of education, training, employment, standard of living, passions and talent, seeking and choosing those host countries that can meet these criteria and provide individuals with everything they need to be satisfied with their livelihood and status (Hajro et al., 2019). The criteria mentioned, the main ones in the case of an emigration decision, together with other criteria of varying importance for each person, are common to most individuals who decide to look abroad for everything that their country of origin cannot offer them. Also, the attractiveness for certain states and their specificity lead to significant mobilizations of foreign citizens within them over time.

Globalisation and regional integration have ensured four freedoms of movement of key importance, including the free movement of human capital, mainly labour, which has made it even easier to travel between EU Member States, where the mere presentation of an identity card allows citizens access to other countries. The attractiveness of a state is also given by its standard of living, by the possibility of easy and rapid adaptation of newly arrived citizens in terms of mentality, religion, ease with which language can be learned, types of work activities that can be carried out and quality of life overall.

On the other hand, reactive motivation appears as a population constraint and not as an expression of one's own will (Richmond, 1993). It is generated by a major, armed conflict (such as the current conflict between Russia and Ukraine) or by a natural cataclysm, or by the propagation and manifestation of a disease (Covid-19), basically forcing citizens to emigrate to free zones, safe zones, to save their lives and resume income-generating activities, which in their country of origin could no longer be possible (Bosilj, 2022). In such situations, the decision to emigrate becomes imperative in terms of the safety of oneself and family, without being analyzed very much beforehand, being a response to a direct threat and not a decision based on one's own motivations and desires.

In recent years, waves of migrants have formed and moved in Europe and not only as a reactive response to everyday threats and challenges, thus there are significant fluctuations and variations in the structure of immigrants, mainly describing the quantity, number and, to a lesser extent, the quality of arrivals (Léonard & Kaunert, 2022). Immigrants' qualifications and training

in certain fields, key criteria of the job offer in host countries, have become increasingly difficult to meet, especially as immigrants no longer choose host countries because of these available job offers, but choose them from the point of view of the safety and security they can provide to citizens and the reduced risk of those states becoming parties to armed conflicts (Brücker et al., 2021; Platt et al., 2022). The large number of migrants arriving in or transiting through a country deepens the crises that migration in general can cause, creating high pressure on authorities and citizens in host countries.

The crisis of migrants from Europe and migration for employment was further deepened when citizens of Ukraine began to arrive following the war, either to transit through certain countries or to settle there for a certain period and obtain stable employment (Libanova, 2019; Mulska et al., 2021). In contrast, migrants from the Middle East and Africa were accommodated in refugee camps with minimal living conditions, access to medical services being rudimentary for them (Brazinskaite, & Goldsmith, 2022). EU Member States should, in view of the above, develop and implement a series of policies to ensure equal, non-discriminatory treatment of immigrants arriving from other Member States or not, regardless of where they come from and what purpose they chose that host country for (refugees from armed conflicts, displaced by natural cataclysms, or jobseekers) (Bosilj, 2022).

Given the contemporary migration phenomenon, which can propagate in waves, whose impact on the labour market of the host country can be very high, some states have decided to implement a system of granting the right to work in that country for citizens of other nationalities, who have expressed a strong desire to enter the country and obtain residency for the purpose of carrying out legal work activities and, respectively, the desire to work in that country. Work permits were granted to people who wanted to live in the host country while working, as well as to people who had to commute daily from their home country to the host country and vice versa. Such permits have easily been granted in huge numbers between countries such as Poland and Ukraine, with the number of Ukrainian migrants to Poland exploding with the outbreak of armed conflict in which Ukraine is involved (Ociepa-Kicińska & Gorzałczyńska-Koczkodaj, 2022). According to UNHCR (2023), about 1.5 million refugees from Ukraine have arrived in Poland since the outbreak of the war, with costs for Ukrainian refugees exceeding 6 billion euros, while one million chose Germany and another four hundred thousand Czechia. About 62,000 refugees from Ukraine arrived in Lithuania during 2022, accommodated with Lithuanian families who volunteered to do so, received free specialized medical treatment, as well as psychological counseling, as well as access to numerous courses and events to accelerate and facilitate adaptation in the new

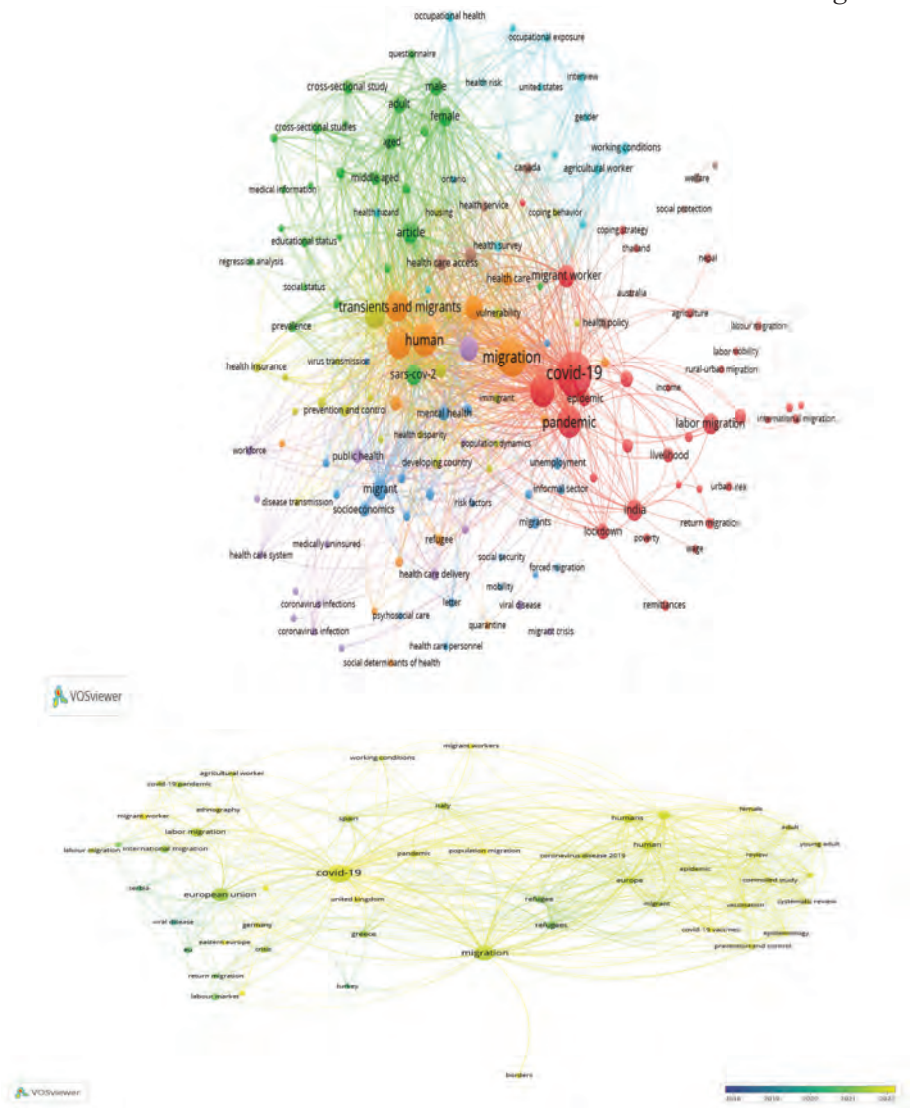
country. About 108,000 refugees arrived in Romania, many of whom later chose other states. Attracting and retaining skilled foreign workers can be key to ensuring economic growth in industrialised countries, especially during periods of population ageing and decline or stagnation of the native workforce (Noja et al., 2018; Cristea et al., 2022). To meet the high demand for such qualities and skills, there are various tools and policies. These range from admitting cross-border commuters or seasonal workers, to granting temporary or permanent residency or even citizenship. Whereas giving more rights to foreign workers may attract a much better trained and higher productivity workforce, but may be perceived as more costly and controversial. Thus, a crucial question for a decision-maker would be how far they should go in this direction. The years before the two major crises that affected and continue to affect the population, namely the crisis caused by the Covid-19 pandemic and the one triggered by the war situation, described the Ukrainian labour market as characterized by instability and disparities, inherent in the transition and crisis periods of economic development (Levytska & Mulska, 2018). The sphere of social and labour relations presents various forms of asymmetry in the case of Ukrainian citizens, as well as inconsistencies in the structure of employment, wages and labour productivity, with uneven income distribution, social vulnerability of poor or poverty-prone people, high unemployment and asymmetric processes of adaptation to constantly changing labour market conditions being maintained and propagated. The expansion of informal employment, which defies and bypasses legal norms and the increasing rate of the shadow workforce, leads to a large number of labour market processes that are uncontrollable and unregulated in the case of Ukraine (Lehmann & Pignatti, 2018). Moreover, major imbalances between workforce training and labour market needs in Ukraine lead to inconsistencies and discrepancies between the level of qualification of workers and the jobs they occupy, thus leading to a deterioration in the quantitative and qualitative structure of the labour market (Lukianenko & Oliskevych, 2015).

Therefore, the recent literature review has focused on international migration flows widely shaped by the caveats of the Covid-19 pandemic and geopolitical crises, particularly at the level of the EU. To summarize and capture the main theoretical guidelines and research directions associated with this topical subject, we further performed a bibliometric analysis of current migration studies. Hence, we extracted a sample of 247 scientific articles published during 2018-2022 and indexed in Scopus, following several keywords resulting from the systematic literature review and in line with our research objective. A large amount of data was, therefore, extracted from Scopus and processed in WOSviewer to capture the co-occurrences of

all keywords related to migration, labour market, Covid-19 pandemic, and geopolitics. The graphical mapping of co-occurrences resulting from this process is presented in Figure 1 below.

Co-occurrence, links and clustering of terms/keywords approached in recent migration literature

Figure 1



Source: designed by authors in VOSviewer, using Scopus indexed articles

Besides the usual terms like migration, migrants/refugees and Covid-19 crisis, in the sample of recent studies analyzed, we accounted for specific keywords like labour migration, migrant workers, working conditions, and the labour market, thus entailing that the labour market outcomes of migrants during the Covid-19 pandemic and geopolitical crisis are a major concern among researchers and policy-makers. From a geographical perspective, most studies focused on the EU, and, particularly, Eastern Europe in terms of emigration, respectively Germany, the United Kingdom, Italy, and Spain, as regards immigration.

In summary, the scientific literature dives deep into the shaping factors and economic consequences of international migration in the present geopolitical context and considering the Covid-19 pandemic effects. However, to the best of our knowledge, there is a gap in comprehensive analyses that provide robust estimates on the labour market outcomes and performance of migrants while also considering the spatial inferences and differentials between the migrant sending and receiving countries, as well as the challenges and opportunities brought by the globalized digital economy.

3. DATA AND METHODOLOGY

3.1. Data, indicators, descriptive statistics

The dataset compiled for this research comprises data extracted from Eurostat (European Commission, 2023) for all EU-27 member states. To capture the effects of the Covid-19 pandemic and the inferences caused by the military conflict between Russia and Ukraine, we configured a cross-sectional dataset at the level of 2021, focusing on the economic, labour market and migration-specific indicators, as follows:

- *Economic development/welfare indicators*: "Gross domestic product (GDP) per capita" (euro) (*GDP_C*), "At-risk-of-poverty rate" (%) (*PR*), "Tertiary educational attainment" (25-34 years, %) (*TEA*), "Average monthly earnings" (Euro) (*AME*);
- *Migration indicators*: "Total emigrants" (number) (*EM*), "Total immigrants" (number) (*IM*);
- *Labour market indicators (including proxies for migrants' labor market outcomes):* "Unemployment rate, % population in the labour force 15-74" (*UR_T*), "Employment rate, %, 20-64" (*ER_T*), "Employed foreign-born by change in skill level from last job before migrating to current job", total 15-74 years (*EFB_SK_T*), and increase, 15-74 years (*EFB_SK_I*), "Foreign-born population by time required to find the first paid job in the host country",

thousand persons 15-74 (*FBP_TFJ*), "Foreign-born population by level of current skills in the main host country language", thousand persons 15-74, total (*FBP_CSK*) and proficient skills (*FBO_PSK*), "Employment by job satisfaction, migration status", thousand persons 15-74, total (*ER_JS_MS*) and high job satisfaction (*ER_JS_MS_H*).

Summary statistics are presented below in Table 1 and provide detailed information about the dataset covering all indicators compiled for 2021 at the level of EU-27.

Descriptive statistics

Table 1

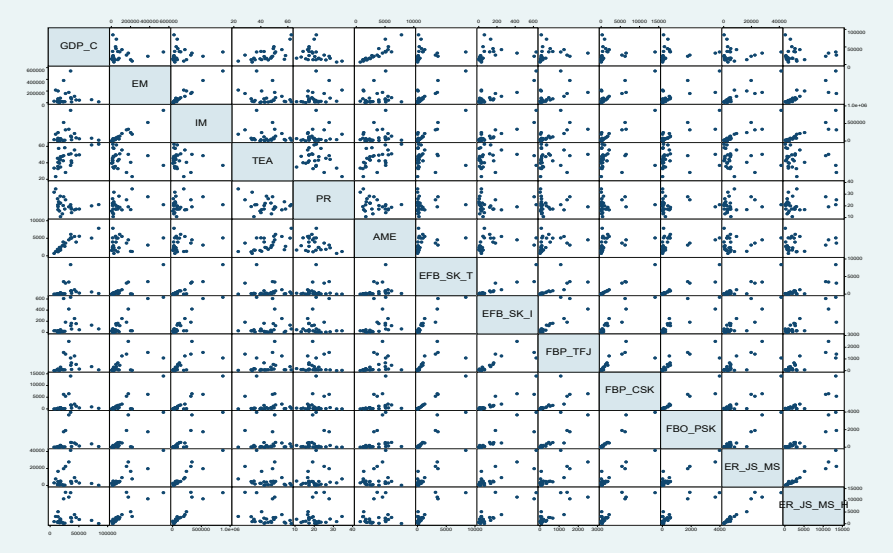
	Median	Mean	Std. Deviation	Skewness	Kurtosis	Shapiro-Wilk	P-value Shapiro-Wilk	Min	Max
GDP_C	22750.00	28331.481	18543.389	1.582	2.625	0.846	< .001	6950.00	84490.00
EM	44008.00	91940.667	124631.585	2.459	6.519	0.677	< .001	3395.00	543162.00
IM	63489.00	138676.926	192120.195	2.660	8.035	0.663	< .001	5733.00	874367.00
TEA	44.20	44.578	9.869	-0.063	-0.304	0.986	0.967	23.300	62.600
PR	20.00	20.693	5.519	0.729	0.513	0.959	0.342	10.800	34.500
AME	2889.53	3336.147	1863.359	0.562	-0.493	0.936	0.098	697.630	7860.330
UR_T	6.20	6.637	2.895	1.560	3.091	0.852	0.001	2.800	14.800
ER_T	75.60	74.644	5.156	-1.010	0.445	0.901	0.014	62.600	81.700
EFB_SK_T	242.90	949.789	1811.301	3.123	10.985	0.551	< .001	9.700	8396.100
EFB_SK_I	39.80	116.678	172.571	2.189	4.183	0.657	< .001	3.500	630.800
FBP_TFJ	123.70	439.250	623.289	1.997	3.840	0.696	< .001	7.500	2501.500
FBP_CSK	358.10	1622.030	3102.482	2.977	9.800	0.556	< .001	17.300	14055.400
FBO_PSK	80.60	549.270	1059.058	2.553	5.876	0.556	< .001	3.400	3974.300
ER_JS_MS	3923.00	7321.796	9888.637	2.195	4.780	0.691	< .001	268.800	41323.400
ER_JS_MS_H	1819.20	3208.059	4030.191	1.799	2.020	0.695	< .001	105.700	13434.900

Source: authors research in JASP

The summary statistics reveal notable differentials among the EU member states as regards the economic welfare levels, migration flows and labour market outcomes. To also capture the degree of correlation between specific indicators, across the EU-27 MS, in 2021, we have performed the correlation matrix, which is graphically represented in Figure 2.

Correlation graph matrix of the indicators used in the empirical (network) analysis

Figure 2

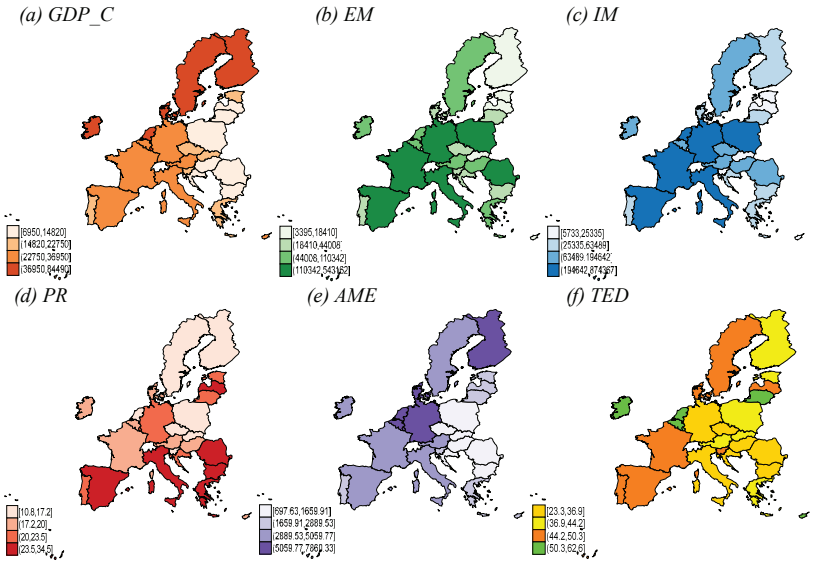


Source: configured by authors in Stata 16

As regards the countries' situation, there is a notable gap in terms of economic welfare levels between the new EU member states (MS) that adhered to the EU in 2004, 2007, and 2013 and the old EU MS. These significant differences are reflected through the GDP per capita (GDP_C) levels (Figure 3a), as well as the average monthly earnings levels (AVE) (Figure 3e) and poverty rates (PR) (Figure 3d). Luxembourg, Ireland, Denmark, the Netherlands, Sweeden and Finland have registered the highest GDP per capita in 2021, while, on the opposite side, Bulgaria, Romania, Croatia, Hungary, Poland and the Baltic states have accounted for the lowest GDP per capita levels in 2021. These credentials become relevant milestones in the migration decision for emigrants in Central and Eastern Europe who, mainly, seek higher income through better jobs and improved living standards. In 2021, Romania and Poland continued to register a high number of emigrants (Figure 3b), despite the pandemic, while a significant number of immigrants (Figure 3c) can be found in Germany, France, Italy, Spain, but also Poland because of the military conflict between Ukraine and Russia. On these lines, the migrants' selection process is enforced according to their educational background and linked to professional profiles, considering that tertiary educational attainment levels vary significantly across the EU countries (Figure 3f).

EU-27 mapping of main economic and migration indicators employed in current research, 2021

Figure 3

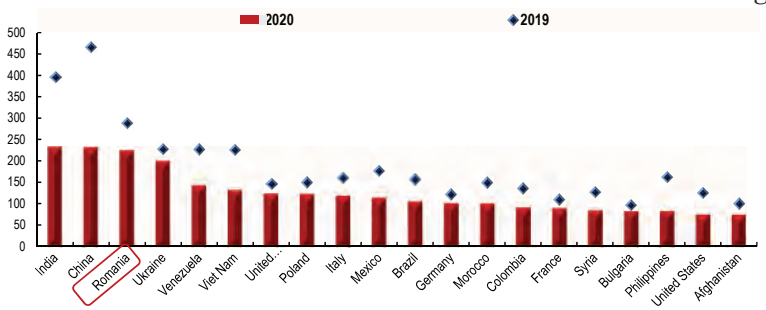


Source: configured by authors in Stata 16

The Covid-19 pandemic has widely shaped the international migration flows across the globe. In Europe, the number of new immigrants was reduced in 2020 compared to 2019, with Romania and Ukraine being in the top 20 countries of origin of the new migrants to the OECD (Figure 4).

Top 20 countries of origin of new immigrants to the OECD, 2019-2020

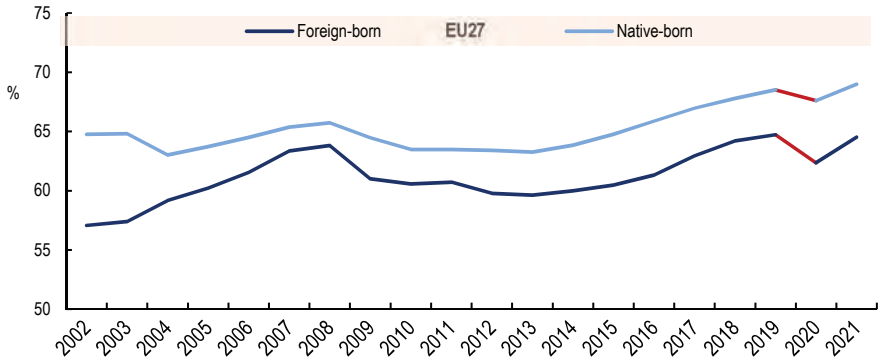
Figure 4



Source: OECD (2023)

Evolution of the employment rates of the foreign-born and native-born population during 2002-2021, at the level of EU-27

Figure 5

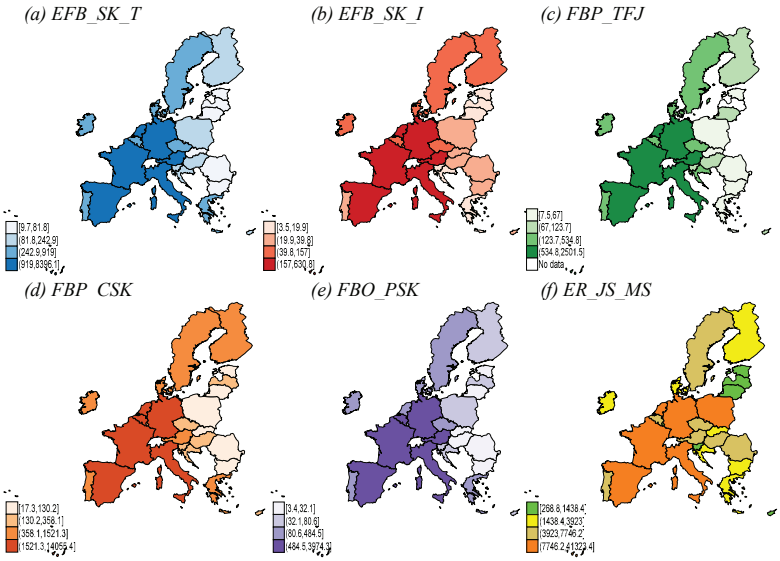


Source: OECD (2023)

At the same time, if we look at the labour market outcomes of migrants and natives during 2002-2021, at the level of the EU, we note that there still remains a notable gap in employment between the foreign-born and the native-born population (Figure 5), that has been emphasized during the Covid-19 pandemic, migrants being most affected by lockdowns and job losses. Labour market outcomes of migrants in EU-27 MS in 2021 are shown in Figure 6, where it can be noted that migrants have good employment paths in main host economies (Greece, France, Spain, Italy, and also Poland in the context of the Ukrainian crisis).

EU-27 mapping of main migrant labour market indicators used in current research, 2021

Figure 6



Source: configured by authors in Stata 16

On this complex frame of facts and challenges, current research aims to examine the labour market outcomes of migrants at the times of the Covid-19 pandemic and geopolitical uncertainty. Shaping and deterring factors, as well as migration caveats, are targeted by this study, therefore the methodological rationale is based on network analysis performed through Gaussian graphical models (GGMs), estimated through two modern techniques, namely the "Extended Bayesian Information Criteria with graphical least absolute shrinkage and selection operator" (EBICglasso) and "partial correlation" (PCOR).

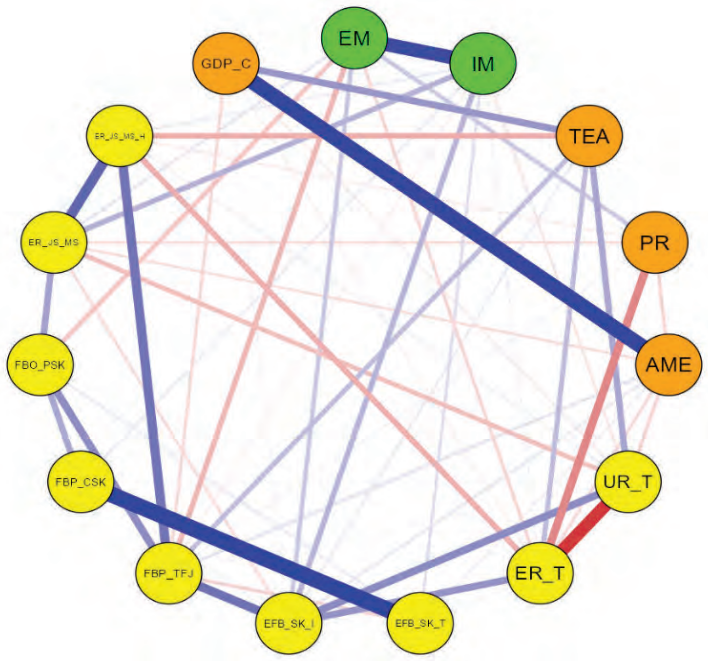
The GGM represents a network of conditional associations that allows capturing the inferences and linkages between specific economic welfare, labour market, and migration-specific credentials. It presents the advantage of avoiding spurious correlations and identifies the relationship between two variables while controlling for the indirect influence of other variables. The model consists of a series of nodes and edges across them that entail the positive (blue) and negative (red) correlations between the variables (nodes of the network), while the thickness and width of the edges show the intensity of the connection, with no edge implying that there is no relation between the variables.

4. RESULTS AND DISCUSSION

The first GGM model was configured through the EBICglasso technique, and the network/results obtained are presented in Figure 7, while the details of the weights matrix are captured in Table A1 in the Appendix. The centrality and clustering plots (Figure 8) accompany the GGM model. For the centrality characteristics of the network, we capture four measures, namely betweenness, closeness, degree and expected influence (Estrada & Rodríguez-Velázquez, 2005), while for the clustering coefficients, we capture the Barrat, Onnela, WS and Zhang indices (Barrat et al., 2004; Onnela et al., 2005; Watts & Strogatz, 1998; Zhang & Horvath, 2005).

GGM model, EU-27, Extended Bayesian information criteria (EBICglasso) estimation method

Figure 7



Source: configured by authors in JASP

Analyzing the results, we noted that there is a strong positive connection between the educational attainment and skills level of migrants and their labour market insertion (employment) (FBP_CSK and EFB_SK_T),

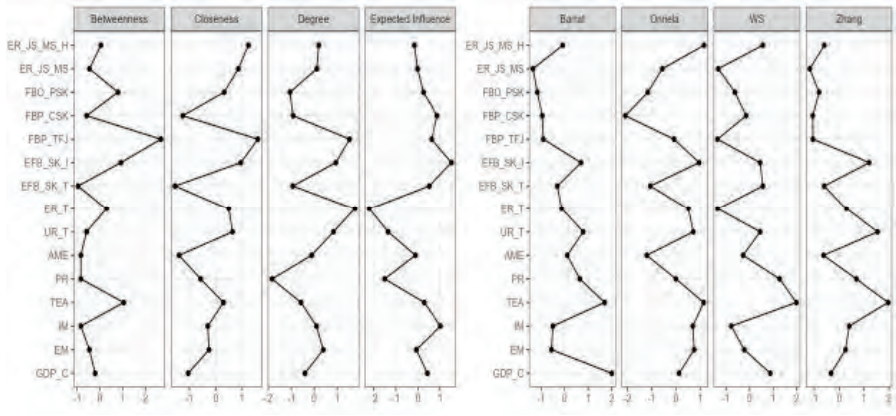
but also between the time required to find the first paid jobs in the host country (FBP_TFJ) and the increase in skills level from last job before migrating to the current job in the host country (EFB_SK_I) according to high professional status (ER_JS_MS_H). Strong positive linkages are also captured between the level of GPD per capita (GDP_C) and average monthly earnings (AME). Emigrants (EM) have negative connections with their proficient skills (FBO_PSK) and the time required to find the first paid jobs in the host country (FBP_TFJ), while immigrants (IM) have positive connections with job satisfaction (ER_JS_MS) and the skills level from last job before migrating to the current job in the host country (EFB_SK_I), thus showing that an increase in the number of immigrants leads to an increase in the employment of foreign-born if the skill levels of migrants improve from last job before migrating to the current job in the host EU country, job satisfaction also playing an important role in the labour market performance of immigrants.

As expected, both migration variables are strongly and positively interrelated. Inverse correlations are noted also between the poverty risk (PR) and total employment rate (ER_T), but also between the tertiary educational attainment (TEA) and the employment by high levels of job satisfaction and professional status (ER_JS_MS_H). Negative connections but of low intensity are registered between average monthly earnings (AME) and key labour market coordinates (ER_T, UR_T and ER_JS_MS), thus entailing that the earnings levels are notably shaped by the labour market performance, particularly the employment levels by migration and professional status, issues also examined by Pirtea et al. (2022) and Noja et al. (2020).

The degree of centrality measures the immediate influence, compared to the long-term effect in the network (Estrada & Rodríguez-Velázquez, 2005), respectively the direct connection between nodes and the risk of infection/degree of influence upon the other nodes in the network that are related. The betweenness centrality is based on the number of times the shorted path between two nodes goes through another node whose centrality is being measured, thus showing how influential this node is in the network. These values (Figure 8, left) are very high for the time required to find the first paid jobs in the host country (FBP_TFJ), but also for immigrants' proficient skills (FBO_PSK), skills level from last job before migrating to the current job in the host country (EFB_SK_I) and the tertiary educational attainment (TEA).

Centrality plot (left) and clustering plot (right) for GGM EBICglasso

Figure 8



Source: configured by authors in JASP

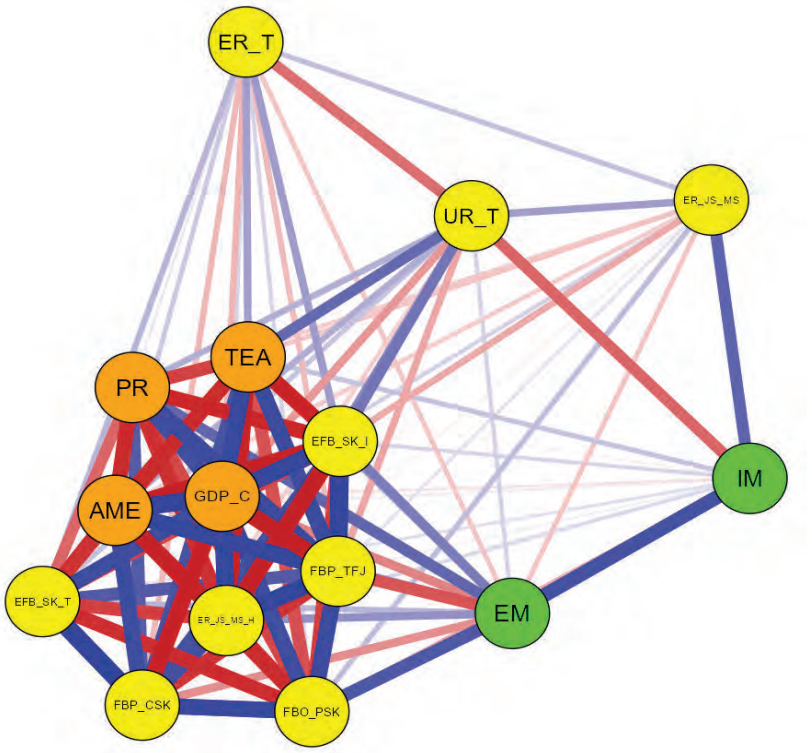
The clustering coefficients (Barrat et al., 2004; Onnela et al., 2005; Watts & Strogatz, 1998; Zhang & Horvath, 2005) quantify the degree to which the nodes of the network cluster together. High values are registered for tertiary educational attainment (TEA), but also the unemployment rate (UR_T), skills level from last job before migrating to the current job in the host country (EFB_SK_I), and also immigrants (IM) and emigrants (EM) in case of Onnela and Zhang values (Figure 8, right).

To add robustness to the results obtained, we also designed and processed another GGM model estimated through the partial correlations method. Main results and the graphical model/network are presented in Figures 9-10, and the details of the weights matrix are captured in Table A2 in the Appendix.

In this case, additional connections that are keener have been identified between all variables considered in our study. In this particular setting, the total employment rate and total unemployment rate are placed separately in the network together with the number of emigrants and immigrants within the EU-27 countries. Even more, the job satisfaction of migrants (ER_JS_MS) is being placed in this setting together with total employment and unemployment rates (ER_T, UR_T), IM and EM and it has strong positive connections with IM, and negative (but with much lower intensity) with EM. EM has also negative connections with the time required to find the first paid jobs in the host country (FBP_TFJ) and the level of current skills in the main host country (FBP_CSK) and very strong positive linkages with their proficient skills (FBO_PSK), skills level from last job before migrating to the current job in the host country (EFB_SK_I), the high levels of job satisfaction (ER_JS_MS_H) and poverty rate (PR).

GGM model, EU-27, partial correlation (PCOR) estimation method

Figure 9



Source: configured by authors in JASP

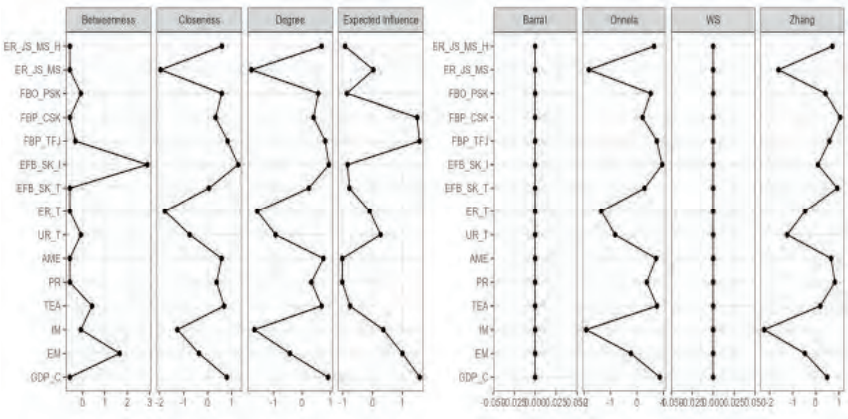
Strong positive connections are also entailed between GDP_C (placed at the core of the network) and tertiary educational attainment (TEA), earnings (AME), skills level from the last job before migrating (EFB_SK_I) and the high levels of job satisfaction (ER_JS_MS_H), and negative with the time required to find the first paid jobs in the host country (FBP_TFJ). This network is more complex and encompasses additional layers in a complex setting, showing economic ties across multiple environments where the labour market performance of migrants (both immigrants and emigrants) plays a significant role.

Analyzing the centrality and clustering coefficients of the second GGM model (Figure 10), we noted that the skills level from last job before

migrating (EFB_SK_I) has the largest betweenness centrality, followed by EM, these two being among the most influential nodes of the network.

Centrality plot (left) and clustering plot (right) for GGM PCOR

Figure 10



Source: configured by authors in JASP

Large values for closeness, degree, and expected influence centrality are also accounted for by GDP_C, but also the first paid jobs in the host country (FBP_TFJ) and the level of current skills in the main host country (FBP_CSK). The Barrat and WS clustering coefficients have the same value for all variables, indicating it is a regular network, while Onnela and Zhang indices show significant variations and quantify the abundance of connected triangles in the network (Masuda et al., 2018) and entail additional layers and transitivity in the network.

5. CONCLUSIONS

Contemporary society grapples with an increasingly complex array of challenges that, while often inescapable, necessitate prudent mitigation strategies. In this context, governments bear the responsibility of formulating and executing policies with the overarching objective of safeguarding the well-being of their citizenry, the broader societal framework, and the integrity of the labor market. Recent years have witnessed discernible fluctuations within the labor market, marked by the obsolescence of certain occupational roles, a renewed emphasis on the significance of others, and the emergence of novel vocations that yield remuneration and foster individual betterment. These

transformations bear significance not only for the quality of life of individuals but also exert discernible effects on state finances.

In this vein, the main findings of our research show that the skills levels and educational background of migrants play an essential role in their labour market insertion and performance in the Covid-19 context and geopolitical uncertainty (we identified strong positive connections between the educational attainment and skills level of migrants and their labour market insertion (employment)). Higher skill levels ensure a high professional status and a reduced time until finding the first paid job in the host country. Working conditions and increased job satisfaction also positively correlate with high employment rates of migrants.

As regards reactive migration, migrants who have left conflict areas are likely to stay permanently or for a long time in host countries, with further options highlighting the possibility of switching to new countries and, to a very small extent, returning to their country of origin, even after the danger has been removed. It is very important for host countries to get involved in supporting migrants and to make significant efforts to ensure their accommodation and adaptation to the specificities and norms of the country, regardless of the criteria and mentality they have been accustomed to, as well as efforts to ensure access to the labor market for as many people as possible, without restricting the rights of its own citizens and without manifesting racist and discriminatory beliefs towards immigrants. As regards the distinct treatment applied by the authorities of host countries to immigrants arriving from different parts of the world, it may not be based on reasoning relating to ethnicity and race, but precisely on a series of affinities and similarities between nationals of the host countries and those of the countries of origin. Thus, there is a tendency to accept more easily immigrant citizens, but coming from a neighboring country, or with the same religion, common origins, as in the case of people from Latin countries, or from a country with which there are certain agreements, regional policies, or countries that are part of the EU. The movement of people for jobs in other countries will resume its usual course, after periods of crisis, characterized by a certain linearity and predictability previously identified, but each new wave of migrants will leave its mark on host countries and present a new picture and structure of the labour market.

Current research faces some caveats due to the limited availability of data on a larger timespan for the indicators capturing the labour market insertion and performance of migrants in host economies. Estimating the network only for one year may yield insufficient observations for capturing the amplitude of the migration process and therefore may hide relevant connections among relevant variables. In this respect, future research targets a longer lapse of time

for the empirical analysis that also captures the dynamics of migration flows and associated economic impacts on both migrants sending and receiving countries.

6. REFERENCES

1. Barrat, A., Barthelemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, *101*(11), 3747-3752. https://doi.org/10.1073/pnas.0400087101
2. Bosilj, K. (2022). *European welcome for Ukrainian migrants – double standards or justified favouritism?,* Human Security for the 21st Century Collection of students' papers, Belgrad, 23-32. OSCE Mission to Serbia, Public Policy Research Center
3. Brazinskaite, R., & Goldsmith D. (2022). *Lithuanian Refugee Crises: Comparing Middle Eastern/African and Ukrainian Migrants*, Environmental Health Perspectives.
4. Brücker, H., Glitz, A., Lerche, A., & Romiti, A. (2021). Occupational recognition and immigrant labor market outcomes. *Journal of Labor Economics*, *39*(2), 497-525.
5. Buechel, B., Gangl, S., & Huber, M. (2023). How residence permits affect the labor market attachment of foreign workers: Evidence from a migration lottery in Liechtenstein. *European Economic Review*, 152.
6. Cristea, M. S., Pirtea, M. G., Suciu, M. C., & Noja, G. G. (2022). Workforce participation, ageing, and economic welfare: New empirical evidence on complex patterns across the European Union. *Complexity*, *2022*, 1-13. https://doi.org/10.1155/2022/7313452
7. Dluhopolskyi, O., Zatonatska, T., Lvova, I., & Klapkiv, Y. (2019). Regulations for Returning Labour Migrants to Ukraine: International Background and National Limitations. *Comparative Economic Research*, *22*(3), 45-64. http://doi.org/10.2478/cer-2019-0022
8. Estrada, E., & Rodríguez-Velázquez J. A. (2005). Subgraph centrality in complex networks. Phys. Rev. E, 71(5), 056103. 10.1103/PhysRevE.71.056103
9. European Commission. (2023). Eurostat database. Available at: https://ec.europa.eu/eurostat/data/database
10. Franc, S., Časni, A.Č., & Barišić, A. (2019). Determinants of migration following the EU enlargement: a panel data analysis. *The South East European Journal of Economics and Business*, *14*(2), 13-22. http://journal.efsa.unsa.ba/index.php/see/article/view/1056/289
11. Gabrielli, G., & Impicciatore, R. (2022). Breaking down the barriers: educational paths, labour market outcomes and wellbeing of children of immigrants. *Journal of Ethnic and Migration Studies*, *48*(10), 2305-2323. https://doi.org/10.1080/1369183X.2021.1935655
12. Hajro, A., Stahl, G.K., Clegg, C.C. and Lazarova, M.B., 2019. Acculturation, coping, and integration success of international skilled migrants: An integrative review and multilevel framework. *Human Resource Management Journal*, *29*(3), pp.328-352. https://doi.org/10.1111/1748-8583.12233
13. Lehmann, H., & Pignatti, N. (2018). Informal employment relationships and the labor market: Is there segmentation in Ukraine?. *Journal of Comparative Economics*, *46*(3), 838-857. https://doi.org/10.1016/j.jce.2018.07.011
14. Léonard, S., & Kaunert, C. (2022). The securitisation of migration in the European Union: Frontex and its evolving security practices. *Journal of Ethnic and Migration Studies*, *48*(6), 1417-1429.
15. Levytska, O., & Mulska, O. (2018). Emerging trends in the labour market of Ukraine in terms of global transformations. *Perspectives: Journal on Economic Issues*, *2*, 53-70.

16. Libanova, E. (2019). Labour migration from Ukraine: key features, drivers and impact. *Economics and Sociology*, *12*(1), 313-328.
17. Lukianenko, I., & Oliskevych, M. (2015). The Effects of Shocks on the Ukrainian Labor Market: SVEC Modeling*, Procedia Economics and Finance*, 27, 311-322.
18. Masuda, N., Sakaki, M., Ezaki T., & Watanabe, T. (2018). Clustering Coefficients for Correlation Networks, *Frontiers in Neuroinformatics*, 12(7). https://doi.org/10.3389/fninf.2018.00007
19. Mulska, O., Levytska, O., Zaychenko, V., Vasyltsiv, T., & Ilyash, O. (2021). Pull environment of migration in the EU countries: Migration vector from Ukraine. *Problems and Perspectives in Management*, *19*(4), 283-300.
20. Noja, G.G., Cristea, M., Stefea, P., & Panzaru, C. (2020), "Challenges and Perspectives of International Migration in Europe within the Brexit Framework" (chapter) in Book Series *Emerald Studies in Finance, Insurance, and Risk Management*, vol. 104, pp. 1-19, Emerald Publishing Limited. doi:10.1108/978-1-80043-095-220201002
21. Noja, G. G., Cristea, S. M., Yüksel, A., Pânzaru, C., & Drăcea, R. M. (2018). Migrants' role in enhancing the economic development of host countries: Empirical evidence from Europe. *Sustainability*, *10*(3), 894. https://doi.org/10.3390/su10030894
22. Ociepa-Kicińska, E., & Gorzałczyńska-Koczkodaj, M. (2022). Forms of aid provided to refugees of the 2022 Russia–Ukraine war: the case of Poland. *International Journal of Environmental Research and Public Health*, *19*(12), 7085. https://doi.org/10.3390/ijerph19127085
23. OECD. (2023). International Migration Database. *OECD International Migration Statistics* (database). https://doi.org/10.1787/data-00342-en
24. Onnela, J. P., Saramaki, J., Kertesz, J., & Kaski, K. (2005). Intensity and coherence of motifs in weighted complex networks. *Physical Review E*, 71(6), 065103.
25. Platt, L., Polavieja, J., & Radl, J. (2022). Which integration policies work? The heterogeneous impact of national institutions on immigrants' labor market attainment in Europe. *International Migration Review*, *56*(2), 344-375. https://doi.org/10.1177/01979183211032677
26. Pirtea, M.G., Noja, G. G., Cristea, M., & Grecu, I. (2022), „Labor Market Outcomes of Romanian Migrants under the Covid-19 Pandemic and Digitalization Challenges: a Spatial Bootstrap Approach", (chapter) in Book Series *Contemporary Studies in Economic and Financial Analysis – The New Digital Era: Digitalization and Emerging Risks and Opportunities*, vol. 109A, pp. 287-304, eISBN: 978-1-80382-979-1. https://www.emerald.com/insight/content/doi/10.1108/S1569-37592022000109A017/full/html
27. Richmond, A. H. (1993). Reactive migration: Sociological perspectives on refugee movements. *Journal of Refugee Studies*, *6*(1), 7-24. https://doi.org/10.1093/jrs/6.1.7
28. UNHCR. (2023). Ukraine Refugee Situation. Available at: https://data.unhcr.org/en/situations/ukraine
29. Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, 393(6684), 440-442.
30. Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. Statistical Applications in Genetics and Molecular Biology, 4(1). https://doi.org/10.2202/1544-6115.1128

Table A1. GGM 1 – PCOR - Weights matrix

	Network														
Variable	**EM**	**IM**	**TEA**	**PR**	**AME**	**UR_T**	**ER_T**	**EFB_SK_T**	**EFB_SK_I**	**FBP_TFJ**	**FBP_CSK**	**FBO_PSK**	**ER_JS_MS**	**ER_JS_MS_H**	**GDP_C**
EM	0.000	0.762	0.332	0.589	0.479	0.171	-0.216	0.283	0.489	-0.649	-0.432	0.691	-0.220	0.416	-0.471
IM	0.762	0.000	0.227	-0.050	0.077	-0.582	-0.040	0.125	0.161	0.064	0.008	-0.327	0.598	0.150	-0.102
TEA	0.332	0.227	0.000	-0.721	-0.813	0.562	0.363	-0.715	-0.853	0.834	0.737	-0.670	-0.221	-0.881	0.887
PR	0.589	-0.050	-0.721	0.000	-0.811	0.304	0.088	-0.588	-0.824	0.851	0.659	-0.707	-0.200	-0.789	0.796
AME	0.479	0.077	-0.813	-0.811	0.000	0.313	0.206	-0.778	-0.894	0.906	0.828	-0.805	-0.198	-0.882	0.974
UR_T	0.171	-0.582	0.562	0.304	0.313	0.000	-0.549	0.227	0.517	-0.341	-0.184	0.041	0.372	0.456	-0.396
ER_T	-0.216	-0.040	0.363	0.088	0.206	-0.549	0.000	0.278	0.375	-0.258	-0.273	0.169	0.239	0.162	-0.289
EFB_SK_T	0.283	0.125	-0.715	-0.588	-0.778	0.227	0.278	0.000	-0.709	0.678	0.984	-0.796	-0.049	-0.728	0.822
EFB_SK_I	0.489	0.161	-0.853	-0.824	-0.894	0.517	0.375	-0.709	0.000	0.951	0.760	-0.740	-0.323	-0.905	0.922
FBP_TFJ	-0.649	0.064	0.834	0.851	0.906	-0.341	-0.258	0.678	0.951	0.000	-0.765	0.855	0.120	0.929	-0.919
FBP_CSK	-0.432	0.008	0.737	0.659	0.828	-0.184	-0.273	0.984	0.760	-0.765	0.000	0.884	-0.005	0.775	-0.864
FBO_PSK	0.691	-0.327	-0.670	-0.707	-0.805	0.041	0.169	-0.796	-0.740	0.855	0.884	0.000	0.249	-0.785	0.822
ER_JS_MS	-0.220	0.598	-0.221	-0.200	-0.198	0.372	0.239	-0.049	-0.323	0.120	-0.005	0.249	0.000	-0.055	0.173
ER_JS_MS_H	0.416	0.150	-0.881	-0.789	-0.882	0.456	0.162	-0.728	-0.905	0.929	0.775	-0.785	-0.055	0.000	0.912
GDP_C	-0.471	-0.102	0.887	0.796	0.974	-0.396	-0.289	0.822	0.922	-0.919	-0.864	0.822	0.173	0.912	0.000

Source: authors research in JASP.

Table A2. GGM 2 – EBICglasso - Weights matrix

Variable	Network														
	EM	**IM**	**TEA**	**PR**	**AME**	**UR_T**	**ER_T**	**EFB_ SK_T**	**EFB_ SK_I**	**FBP_ TFJ**	**FBP_ CSK**	**FBO_ PSK**	**ER_ JS_MS**	**ER_JS _MS_H**	**GDP _C**
EM	0.000	0.708	0.000	0.175	-0.028	0.000	-0.125	0.000	0.178	-0.228	0.000	-0.184	0.086	0.000	-0.018
IM	0.708	0.000	0.000	0.000	0.000	-0.082	0.053	0.108	0.236	-0.055	0.072	0.000	0.247	0.067	0.000
TEA	0.000	0.000	0.000	-0.021	0.015	0.278	0.208	-0.005	0.068	0.206	0.000	0.000	0.000	-0.259	0.330
PR	0.175	0.000	-0.021	0.000	-0.128	0.024	-0.378	0.000	0.000	0.046	0.000	0.000	-0.105	-0.067	0.000
AME	-0.028	0.000	0.015	-0.128	0.000	-0.135	-0.117	0.056	0.079	0.094	0.017	0.004	-0.100	0.000	0.779
UR_T	0.000	-0.082	0.278	0.024	-0.135	0.000	-0.631	0.000	0.364	0.029	0.022	0.000	-0.202	0.056	-0.066
ER_T	-0.125	0.053	0.208	-0.378	-0.117	-0.631	0.000	0.033	0.337	0.000	0.000	-0.029	0.000	-0.255	-0.052
EFB_SK_T	0.000	0.108	-0.005	0.000	0.056	0.000	0.033	0.000	0.035	-0.136	0.820	0.071	0.000	0.000	0.000
EFB_SK_I	0.178	0.236	0.068	0.000	0.079	0.364	0.337	0.035	0.000	0.436	0.043	0.000	-0.107	0.000	0.042
FBP_TFJ	-0.228	-0.055	0.206	0.046	0.094	0.029	0.000	-0.136	0.436	0.000	0.000	0.356	0.000	0.433	-0.117
FBP_CSK	0.000	0.072	0.000	0.000	0.017	0.022	0.000	0.820	0.043	0.000	0.000	0.281	0.001	0.000	-0.011
FBO_PSK	-0.184	0.000	0.000	0.000	0.004	0.000	-0.029	0.071	0.000	0.356	0.281	0.000	0.295	0.000	0.000
ER_JS_MS	0.086	0.247	0.000	-0.105	-0.100	-0.202	0.000	0.000	-0.107	0.000	0.001	0.295	0.000	0.489	0.000
ER_JS_MS_H	0.000	0.067	-0.259	-0.067	0.000	0.056	-0.255	0.000	0.000	0.433	0.000	0.000	0.489	0.000	0.035
GDP_C	-0.018	0.000	0.330	0.000	0.779	-0.066	-0.052	0.000	0.042	-0.117	-0.011	0.000	0.000	0.035	0.000

Source: authors research in JASP.

Impact of ageing on economic growth at regional level in the Czech Republic

Jan POKORNÝ (pokorny@fek.zcu.cz)
Department of Finance and Accounting, Faculty of Economics, University of West Bohemia

ABSTRACT

The process of population aging represents an important phenomenon in society. Therefore, this phenomenon is given wide attention. Population aging has many significant impacts. The purpose of this study is to analyse the impact on economic growth at the regional level in the Czech Republic. First, there is a literature search on the issue of population aging and economic growth. It is possible to use the view of the Cobb-Douglas production function or product creation as a component of the gross domestic product. Based on the studies, it is possible to expect a negative relationship, that a larger share of the elderly population leads to decreased economic growth. The study examines this relationship at the level of regions in the Czech Republic. The ordinary least squares model is used for this analysis. Based on the results of the study, it is possible to state that the aging process is present in all regions except the capital city of Prague. Economic growth was the highest in the Moravian-Silesian region. Thus, these two regions represent outliers. Based only on the comparison of the two extreme values of the time 2000-2021, it is possible to state that there is a noticeable decrease in economic growth in connection with the aging of the population. In time series analysis, this statement cannot be declared. That is why five-year periods were used, which partially eliminated the problem of being influenced by the cycle. However, the result still needs to be confirmed. The limitation of the research is the limited time series and the variables used. Further research can focus on a more detailed analysis involving other parameters.

Keywords: *ageing, economic growth, regions, Czech Republic*
JEL Classification: *O18, O47, R11*

1. INTRODUCTION

Population ageing is a crucial process in current society. This process influences many different areas as health (Yu, 2021), long-term care (McCormack, McNally & O'Shea, 2021), active ageing (Ortega, 2021), pension systems (Bazzana, 2020; Mesa-Lago, Moreno & Kay, 2022) and connected income inequality (Hwang, Choe & Choi, 2021).

In this paper, the impact of ageing is examined with regard to economic growth. This topic is very connected with the pension system, respectively, with the public pension scheme in the Czech Republic.

This paper develops the issue on which the author focuses. This issue is Age management as a process of responding to the aging of the population. Age management issues from various angles are then dealt with Fabisiak & Prokurat (2012) or Gorzeń-Mitka, Sipa & Skibiński (2017). Effects of the aging process thus play an important role that needs to be addressed.

Based on the research intent, the aim could be defined. The paper aims to analyse the ageing effect on economic growth.

The literature review is used for analysing the issue ageing and economic growth in this paper. Then the ordinary least squares (OLS model) is used to quantitatively analyse both parameters. Lastly, the synthesis method is used for summaries.

This paper is divided into six parts. Section 2 Literature review describing the issue ageing connected with economic growth based on the description. Section 3 defines the data used for the processing of this paper. Then the results are described in Section 4. Lastly, Section 5 summarizes the paper in the Conclusion.

2. LITERATURE REVIEW

The aging of the population is a significant process that affects all societies. However, this effect is particularly noticeable in the countries of Europe and Japan, and the United States of America. This situation can be illustrated using the so-called old-age dependency ratio, i.e., the ratio of people over 65 and the age group from 15 to 64. While globally this ratio was 13% in 2015, it is expected to increase to 38% in 2100. The higher rate will be in the mentioned countries, which will burden public finances (Conesa & Kehoe, 2018).

Nwakeze (2014) describes her summary of research on the relationship between macroeconomics and demography in Nigeria:

"The major findings indicate that population size affects gross domestic product positively. The age dependency ratio was also found to have a negative effect which reflects the Nigerian situation as a country with a youthful population and unemployment crisis."

Given this, primary analysis of population and age cohort changes is appropriate because each country can have different effects of this process.

While Nigeria is a country with a young population, European countries have an older population. However, society is gradually reacting to this and creating its mechanisms, which also means financing replacement income, i.e., pensions.

More than ten years ago, Herrmann (2012) came to a conclusion that, despite the aging of the population, there is no shortage of labour. This could be given mainly in the context of the given time.

This is also helped by the focus on higher added value and changes within the production process (digitalization and automation). This can be doubted in some countries today, given the very low unemployment rate and the overheated labour market.

This paper uses economic growth as the parameter for evaluating ageing outputs. Thus, the gross domestic product is an important indicator, and the views could be utilized. Firstly, the Cobb-Douglas production function is part of the overlapping generations (OLG) model and affects the output:

$$Y_t = A_t \, K_t^{\alpha} L_t^{\,1-\alpha} \tag{1}$$

The real economic output Y_t is composed of a scaling variable A_t, capital K_t and labour L_t. The exponent represents constant return in both factors, which means capital and labour (Lindh & Malmberg, 2009).

However, on the real economic output could use the second view based on final goods output, then:

$$Y_t = C_t + G_t + I_t^n \tag{2}$$

where is variable for consumption C_t, government expenditures G_t and net investment I_t^n (Fougère & Mérette, 1999).

Both equations could be connected, then the output of equation (1) and (2) is:

$$A_t \, K_t^{\alpha} L_t^{\,1-\alpha} = C_t + G_t + I_t^n \tag{3}$$

Then the parameters could be defined with regard to dependency on the population and related indicators. On the left side equation (3) labour is highly dependent on the population because labour force is affected by the size population and population distribution between age cohorts. Labour force is simplified and composed of age cohorts from 15 to 64 years. The lower limit is also limited by legislation.

However, it is possible to define this age group further more closely, as entry into the labour market is rather delayed. Thus, this age cohort could be defined from 20 years. This could lead to the definition of partial age cohorts participating in the labour market.

At the same time, people aged 65 and over can participate. However, the such representation will be lower. Therefore, it is possible to define the main working group between 20 and 64. Due to the data, the range mentioned above of 15-64 years is more often used, which will continue to be used.

Capital is affected by savings (Lindh & Malmberg, 2009). Thus, the population has a significant effect on this variable because their net income is divided into consumption and savings. The income can be a wage for a worker or a social benefit. For retirees, this is typically a pension, which can be from public sources (from insurance premiums or taxation) or a private source based on the previous appreciation of savings. As a result, when private resources are drawn down, the accumulated capital decreases. Similarly, when drawing down earlier savings, which can affect the entire population, however, it is more evident in the case of pensioners.

On the right side of equation (3), there is a different breakdown of what is mentioned here. Consumption is defined by the size of the population and its tendency to consume or to save.

Consumption is then defined by the purchase of a product that depends on a sub-age group, i.e., again within a more detailed definition, e.g., in five-year intervals. A different level of consumption can be expected, even with regard to the shopping basket, which is differentiated not only according to income but also according to age.

Net savings are then affected by the residual value between income and consumption. This indicator has already been described above. Again, there is a strong dependence on age.

Government spending is then affected by income and any additional resources. In contrast, it was providing goods to a given population. It also fulfils a redistributive function, which is an important part, as it is represented by taxes (reducing possible consumption and savings) and transfers (increasing possible consumption, it is impossible to think about savings since social transfers are aimed more at low-income people).

Thus, defining the outcome of economic growth within the aging process is complex. On the one hand, there is a negative effect, as there is a decrease in the labour force, which stimulates higher economic growth with regard to the level of potential output. At the same time, there is a possible increase in savings, i.e., levies (insurance and taxation), which can then be used as part of government spending without the influence of other funds.

On the other hand, savings are drawn, which are reallocated from investments to consumption which positively affects the product as part of multiplier effects.

The author expects negative effects associated with population aging on the final product. However, it is necessary to consider partial positive and negative effects that depend on other conditions of the given economy, especially within the given technological development or use of capital and its further development.

Maestas, Mullern & Powell (2023) add that, based on empirical data from the United States of America in the period 1980-2010, the aging process leads to a decrease in gross domestic product growth per capita. Specifically, with a 10% increase in the population aged 60 and over, GDP per capita should decrease by 5.5%. Based on the impact analysis, the authors conclude that 2/3 of the impact is due to lower growth in labour productivity. At the same time, the remaining part is affected by the low growth of employment per capita.

A similar result is reached by Gagnon, Johannsen & López-Salido (2021), again using the example of the USA.

3. DATA AND METHODS

The paper aims to analyse the ageing effect on economic growth. To achieve this goal, the OLS model is used to find the relationship between the demographic structure and economic growth.

The OLS model is composed of the dependent variable and independent variables, also known as explanatory variables. For this paper is used:

$$GDP_{ij} = \beta_0 + \beta_1 Y_{ij} + \beta_1 M_{ij} + \beta_1 O_{ij} + \varepsilon_{ij} \quad (4)$$

The dependent variable is gross domestic product GDP_{ij}, which the demographic structure affects. Three parts of the population describe this, which could be defined as young, middle, and old. Firstly, the young age is under 15 years. This group has not representation in the labour market. Then, middle age is the main group in the labour market. The employment rate influences it. However, the younger part of the second group (students) could have an effect on the development of economic growth.

Lastly, old age is associated with retirees. The part of this group can be active in the labour market. On the other hand. The main part of the group is inactive in the labour market.

Based on this description as well as based on the Section 2 Literature Review, the author expected a positive effect on economic growth with the

increasing share of middle age. On the other hand, a negative effect is expected with the increasing number of other groups (young and old age).

Data used in this paper are described in Table 1. Firstly, the time horizon is from 2000 to 2021. This time is affected based on a statistical database that is used. It means Czech Statistical Office.

Data Description

Table 1

Variable	Description	Source
i	Time from 2000 to 2021	X
j	Regions of the Czech Republic	X
GDP_{ij}	Gross domestic product at the time and in the region	Czech Statistical Office (2023a)
Y_{ij}	Population under 15 years at the time and in the region	Czech Statistical Office (2023b)
M_{ij}	Population from 15 to 65 years years at the time and in the region	
O_{ij}	Population above 65 years at the time and in the region	

Source: own processing, 2023

The regions of the Czech Republic are used for analysis in this paper. The Czech Republic has 14 regions, 13 regions, and the capital city, respectively. It is about Prague as the capital city and regions: Central Bohemian, South Bohemian, Pilsen, Karlovy Vary, Ústí nad Labem, Liberec, Hradec Králové, Pardubice, Vysočina, South Moravian, Olomouc, Zlín, and Moravian-Silesian Region.

Share of Gross Domestic Product at Regional Level (differences against 2000)

Table 2

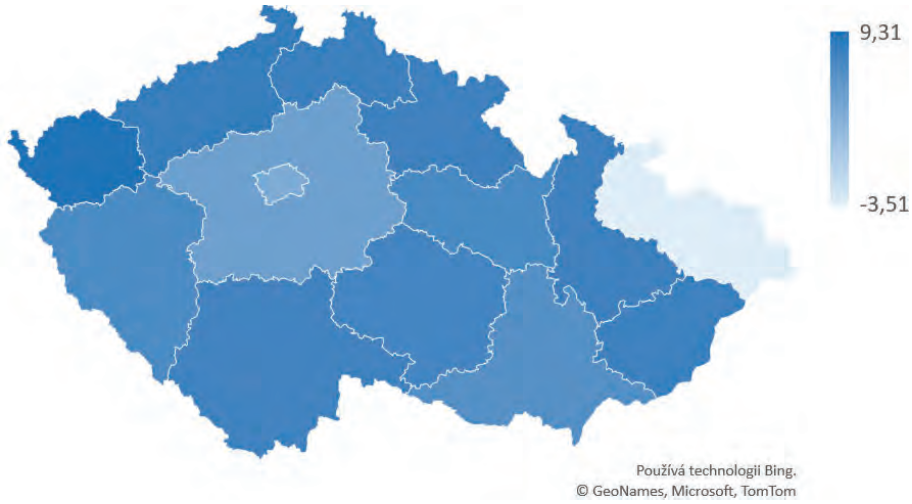
Region	2000	2005	2010	2015	2020	2021
Prague	22.7	2.6	1.7	-0.4	0.7	4.7
Central Bohemian	10.8	-0.4	0.1	0.6	0.2	0.5
South Bohemian	5.7	-0.3	-0.5	-0.1	0.1	-1.0
Pilsen	5.0	0.0	-0.1	0.0	-0.2	-0.1
Karlovy Vary	2.5	-0.3	-0.2	-0.2	-0.2	-0.9
Ústí nad Labem	6.7	-0.2	-0.3	-0.3	-0.5	-1.4
Liberec	3.8	-0.3	-0.3	0.0	-0.1	-0.8
Hradec Králové	5.0	-0.4	-0.2	0.0	0.2	-0.3
Pardubice	4.2	-0.3	0.0	0.0	0.1	-0.3
Vysočina	4.2	-0.1	-0.2	0.0	0.1	-0.3
South Moravian	10.0	-0.2	0.4	0.3	0.3	1.0
Olomouc	5.0	-0.4	-0.1	0.1	0.1	-0.3
Zlín	4.7	-0.1	0.0	0.1	-0.1	-0.1
Moravian-Silesian	9.6	0.4	-0.5	-0.2	-0.8	-0.7

Source: own computations based on CZSO (2023a), 2023

Table 2 shows the share of gross domestic product at the regional level. In 2000, Prague had the highest share (22.7 %) in other years. Since 2005, Prague has created over a quarter of the gross domestic product. It meant an increase of about 4.7 percentage points in 2021.

Differences of Old Age Share between 2000 and 2021 (percentage points)

Figure 1



Source: own computation based on CZSO (2023b), 2023

Other important regions are Central Bohemian and South Moravian. Both regions created more than 10 % of the gross domestic product in the monitored period.

In this contrast, the increasing share in 2021 against 2000 has Prague, Central Bohemian, and South Moravian regions. Other regions have to decrease their shares from -0.1 to -1.4 percentage points.

The next variable is the share-based age. The population increased by about 0.25 million people from 2000 to 2021. The old-age cohort is incising about 52 % from 1.42 to 2.17 million people. Middle age is decreasing from 7.18 to 6.65 million people, about -7.31 percentage points. The share of young is relatively similar in 2000 and 2021, 16.21 % of the population, respectively 16.1 %.

Figure 1 shows the changes in the share of retirees in the population from 2000 to 2021. The greatest change is 9.31 percentage points in Karlovy Vary Region. On the other hand, the lowest change is -3.51 percentage points in Moravian-Silesian Region. Thus, this region is the only part of the

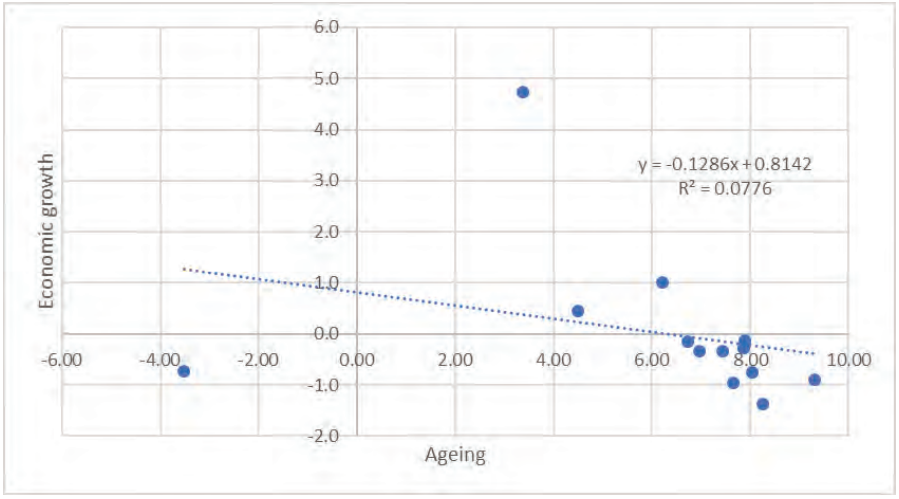
Czech Republic where the share decreased. The second lowest value is 3.38 percentage points in Prague.

4. RESULTS

Before proceeding to the analysis based on the model, it is possible to mention a certain similarity due to Section 3. Based on Table 2 and Figure 1, the differences in the index approach have certain similarities. Figure 2 shows regression between these parameters.

Index of Ageing and Economic Growth

Figure 2



Source: own computation based on CZSO (2023a; 2023b), 2023

Figure 2 uses the index between 2000 and 2021 with respect to the previously described data. The regression using 14 samples, it means very few observations. However, the changes do not show clearly that higher ageing corresponds with lower economic growth. The adjusted coefficient of determination has a low declarative value. This lower value is mainly due to two outliers: Prague and the Moravian-Silesian region.

Based on this, outliers were removed, which differ from other regions by a lower rate of aging, thus limiting the indicative value.

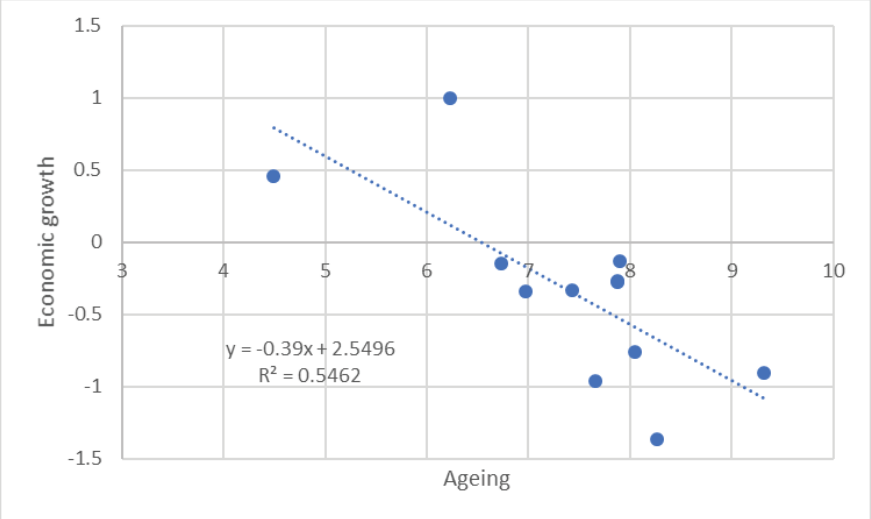
Due to the adjustment, there was an increase in the adjusted coefficient of determination, which increased from 0.07 to 0.55, i.e., to a level of relatively strong dependence. In the second regression in Fig. 3, the relationship between

population aging and lower economic development at the regional level is more evident. A study from the USA confirms this.

Pearson correlation could be used, and 14 variables show a correlation of -0.279, which means – there is a weak negative relationship. The correlation is very positive without outliers (Prague and Moravian-Silesian Region), with 0.997.

Index of Ageing and Economic Growth without outliers

Figure 3



Source: own computation, 2023

Based on the mentioned changes, there is a noticeable difference within the regions mentioned, which deviate from these assumptions. The OLS model is also used to improve the estimate, which is applied to the time horizon 2000-2021, i.e., with respect to the available data set.

Table 3 shows OLS models for every region in the Czech Republic. Table 4 reflected the issue outliers and thus showed Czechia regions without Prague and Moravian-Silesian regions.

Ordinary Least Squares Model

Table 3

	Model 1	Model 2	Model 3	Model 4
HDP_{ij}	Basic	Ln	I	Δln
Constant	-14 338 (-6.29)***	-3.89 (-8.44)***	0.89 (3.35)***	0.05 (14.44)***
Y_{ij}	-4.43 (-5.40)***	-1.11 (-5.19)***	-0.39 (-2.64)***	-0.39 (-2.70)***
M_{ij}	0.55 (3.50)***	0.83 (5.22)***	0.79 (3.72)***	0.74 3.64***
O_{ij}	5.36 (11.21)***	1.57 (18.28)***	-0.24 (-1.66)*	-0.22 (-1.56)
Adj R	0.64	0.86	0.11	0.11
Sample	308	308	294	294

Source: own computation, 2023

In both models, the OLS model is created on the basic data (models 1 and 5) and on adjustments, which are the logarithm of the original data (models 2 and 6). In addition, indices (models 3 and 7) and differences in logarithms (models 4 and 8) are created. Due to the modification, 308 data are used for the first part, or 264 data without two regions. In the second part, the number of data is reduced to 294 (252).

In table 3, models 1 and 2 show a similarly adjusted coefficient of determination level of 0.95, which is a very high value. In contrast, for models 3 and 4, the value is at the level of 0.1, and the explained part of the model is very low. The statistical significance of the variables is in all cases only in model 2. In other cases, the problematic population is under 14 years.

Model 2 finds a negative relationship between the young and older part of the population, which is due to the fact that the given part of the population will enter the labour market only in the future. In other cases, model 2 shows a positive relationship, which, however, is higher for the aging part of the population, which is surprising on two levels: 1. the assumption based on foreign studies is not confirmed, and 2. the value is higher than in the case of the working part. Based on this, it would be possible to assume a high increase in consumption, which compensates for other decreases.

Models 3 and 4, however, show the expected reduction in both the young and the elderly population. However, in the case of the younger part of the population, there is no statistical significance. At the same time, the model explains only 10 %.

Ordinary Least Squares Model without Prague and the Moravian-Silesian region

Table 4

	Model 5	Model 6	Model 7	Model 8
HDP_{ij}	Basic	Ln	I	Δln
Constant	-46 245.70 (-10.23)***	-1.45 (-5.35)***	1.21 (3.16)***	0.05 (9.42)***
Y_{ij}	0.26 (-1.22)	-0.41 (-3.87)***	-0.18 (-0.89)	-0.18 (-0.93)
M_{ij}	-0.14 (-3.46)***	0.20 (2.24)**	0.63 (2.47)**	0.59 (2.41)**
O_{ij}	2.76 (27.76)***	1.36 (33.16)***	-0.52 (-2.09)**	-0.50 (-2.03)**
Adj R	0.95	0.95	0.10	0.10
Sample	264	264	252	252

Source: own computation, 2023

Table 4 adjusts the version to fewer regions (without Prague and the Moravian-Silesian region). However, the results are similar. Therefore, the original assumption, when a change is compared to table 3 was expected, is not fulfilled. Considering this, it is not necessary to reflect on this part further.

Due to the previous results, there is an expansion of the table 5, where five-year periods are analysed, which limits the number of data. The aim is similar studies to those presented above, but drawn from different periods and with different explanatory variables. Due to the small differences between the results in tables 3 and 4, there is no use of differentiation for all regions and a narrower selection. Models 9-12 have a relatively high rate of explanation for the models. The results are similar to the above. The basic model and its logarithmic form (models 9 and 10) negatively correlate economic growth and the young generation. There is a positive trend in other parts. At the same time, the data are statistically significant. This changes in models 11 and 12, where there is a negative relationship between the young and older generations with economic growth. However, there is no statistical significance in the data.

Ordinary Least Squares Model 5-years

Table 5

	Model 9	Model 10	Model 11	Model 12
HDP_{ij}	Basic	Ln	I	Δln
Constant	-133 778 (-3.06)***	-3.92 (-4.37)***	0.81 (3.12)***	0.20 (14.13)***
Y_{ij}	-4.62 (-3.04)***	-1.30 (-3.74)***	-0.14 (-0.51)	-0.26 (-1.58)
M_{ij}	0.64 (2.12)**	0.99 (3.23)***	0.44 (4.30)***	1.70 (5.91)***
O_{ij}	5.07 (5.74)***	1.58 (9.96)***	-0.13 (-1.59)	-0.24 (-1.55)
Adj R	0.63	0.86	0.36	0.48
Sample	84	84	70	70

Source: own computation, 2023

5. CONCLUSION

In addition to the size of the population, the distribution of the population between individual age cohorts also matters. Population aging is an important phenomenon in today's world that affects various human aspects, including a country's economic performance.

When the population ages, it is possible to expect a reduction in the economic performance of countries, which can be observed through the Cobb-Douglas production function, which is composed of labor and capital. At the same time, capital is affected by savings.

On the other hand, the product as a gross domestic product is characterized, among other things, by consumption, which is increased due to the reduction in savings of the elderly population.

The influence on the economic product is thus two-sided and is influenced by the given conditions.

Based on the presented data, it is evident that the Czech Republic is also struggling with an aging population within the individual regions, except for Prague. However, a negative effect on economic growth cannot be confirmed because, on the contrary, there is a positive relationship between an older population and economic growth. This would mean outweighing this effect.

A limitation of the study is the limited data set, which also includes economic fluctuations. Another limitation is the limited number of variables included in this study. However, this study aimed to explain the relationship between economic growth and population ageing. An alternative could be to

expand the number of age cohorts and a more detailed analysis, which can be used for further research.

Further research can also be focused on other explanatory variables that will expand the model, which may lead to a better explanation. On the other hand, in this case, it is possible to pay attention to the model of overlapping generations and its versions, which better captures this issue about the prediction of further development over time.

ACKNOWLEDGMENT

This paper was created within the project SGS-2023-007 'Current Challenges and Problems of Modern Society from the Perspective of Finances and Accounting' at the University of West Bohemia, Faculty of Economics.

REFERENCES

1. Bazzana, D. (2020). Ageing population and pension system sustainability: reforms and redistributive implications. *Economia Politica, 37*(3), 971-992. https://doi.org/10.1007/s40888-020-00183-8
2. Conesa, J. C., & Kehoe, T. J. (2018). An introduction to the macroeconomics of aging. *The Journal of the Economics of Ageing*, *11*, 1-5. https://doi.org/10.1016/j.jeoa.2018.03.002
3. Czech Statistical Office (2023a). *Regional accounts.* Retrieved from https://vdb.czso.cz/vdbvo2/faces/en/index.jsf?page=vystup-objekt&z=T&f=TABULKA&skupId=4809&katalog=30832&pvo=NUCD05&pvo=NUCD05&str=v324
4. Czech Statistical Office (2023b). *Distribution of the population by age as at 31.12. (basic group) - territory selection*. Retrieved from https://vdb.czso.cz/vdbvo2/faces/en/index.jsf?page=vystup-objekt&z=T&f=TABULKA&katalog=30845&pvo=DEMZU04&u=v8__VUZEMI__100__3018
5. Fabisiak, J., & Prokurat, S. (2012). Age Management as a Tool for the Demographic Decline in the 21st Century:An Overview of its Characteristics. *Journal of Entrepreneurship, Management and Innovation, 8*(4), 83-96. https://doi.org/10.7341/2012846
6. Fougère, M., & Mérette, M. (1999). Population ageing and economic growth in seven OECD countries. *Economic Modelling, 16*(3), 411-427. https://doi.org/10.1016/S0264-9993(99)00008-5
7. Gagnon, E., Johannsen, B. K., & López-Salido, D. (2021). Understanding the New Normal: The Role of Demographics. *IMF Economic Review, 69*, 357–390. https://doi.org/10.1057/s41308-021-00138-4
8. Gorzeń-Mitka I., Sipa M., Skibiński A. (2017). Multifaceted Character of the Issues of Age Management. *Polish Journal of Management Studies, 16*(2), 110-121. https://doi.org/10.17512/pjms.2017.16.2.10
9. Herrmann, M. (2012). Population Aging and Economic Development: Anxieties and Policy Responses. Popu*lation Ageing, 5*, 23–46. https://doi.org/10.1007/s12062-011-9053-5
10. Lindh, T., & Malmberg, B. (2009). European Union economic growth and the age structure of the population. *Economic Change and Restructuring* 42, 159-187. https://doi.org/10.1007/s10644-008-9057-1

11. Maestas, N., Mullen, K. J., & Powell, D. (2023). The effect of population aging on economic growth, the labor force, and productivity. *American Economic Journal: Macroeconomics, 15*(2), 306-332. https://doi.org/10.1257/mac.20190196
12. McCormack, R., McNally, S., & O'Shea, D. (2021). AGEING POPULATION: THE IMPACT ON ADMISSIONS TO LONG-TERM CARE. *Age and Ageing, 50*(3), ii9–ii41. https://doi.org/10.1093/ageing/afab219.97
13. Mesa-Lago, C., Moreno, C., & Kay, S. J. (2022). The financial sustainability of public pensions in Cuba: The impact of ageing, structural reforms and the economic crisis. *International Social Security Review, 75*(2), 25-46. https://doi.org/10.1111/issr.12293
14. Nwakeze, N. M. (2014). The Nexus Between Macroeconomics and Demographics: Implications for Sustainable Development. *The Economic and Social Review, 45*(2), 285–298
15. Ortega, J. A. (2021). Is Active Ageing Coping with Population Ageing? Draft Submitted to Journal of Population Aging. *Journal of Population Ageing, 14*(1), 37-52. https://doi.org/10.1007/s12062-020-09265-8
16. Seokchae, H., Choe, C., & Choi, K. (2021). Population ageing and income inequality. *The Journal of the Economics of Ageing*, 20, 100345. https://doi.org/10.1016/j.jeoa.2021.100345
17. Yu, Y. (2021). Healthy ageing in urban China: Governing the ageing population. *Geographical Journal, 187*(1), 28-38. https://doi.org/10.1111/geoj.12372

The perception of European workers on the adoption of telework during and after the Covid-19 Pandemic: The use of supervised and unsupervised learning techniques

Giani Ionel GRĂDINARU[1](giani.gradinaru@csie.ase.ro)
The Bucharest University of Economic Studies, Institute of National Economy - Romanian Academy, Romania

Diana TIMIȘ
The Bucharest University of Economic Studies, Romania

Cătălin-Laurențiu ROTARU
The Bucharest University of Economic Studies, Romania

ABSTRACT

Globally, 2020 began with unprecedented changes. With the worldwide declaration of the COVID-19 pandemic, the transition to teleworking took place instantly, becoming one of the largest historical experiments. The objectives of this study are to identify a limited number of characteristics of European workers regarding telework trends, to find typologies of groups of European countries where workers similarly perceive the forced adoption of telework imposed by the COVID-19 pandemic, and to find common patterns between Romanian and other European workers. At the same time, the research aims to predict the behaviour of European workers after the COVID-19 pandemic and how the attrition towards working from home can be improved. The Eurofund 2020 survey was used to conduct quantitative research. To achieve the first objective, unsupervised learning techniques (principal component analysis) were used to highlight the types of European workers. and how they have been affected by telecommuting. The results showed groups of European workers who were deeply affected in terms of isolation, personal life or work satisfaction, but also groups of European workers who can claim to be gainers as a result of remote work. For the second and

1. corresponding author

third objectives, cluster analysis was selected as the method. The similarities and differences between the perceptions of European workers regarding the adoption of teleworking were assessed, with Romanian citizens having the same concepts as Poles and Irish. For the objective of predicting the high attrition of home work, following the results from the first objectives of the analysis, workers from European countries were divided into home workers and office workers and based on these, using supervised learning (logistic regression) one can predict which component should be improved in order to have a greater attrition regarding home workers.

Keywords: *telework, European workers, principal component analysis, clustering, unsupervised learning, COVID-19*

JEL Classification: *A130, C150, J440*

INTRODUCTION

Telework is defined as the use of communication technologies to perform work outside the employer's premises (Vega et al., 2014). In recent years, information technologies have developed extremely rapid, and with them, also the ways of communication (World at work, 2011). In this context, telework has become a strategic field as it has the opportunity to develop very quickly, has a very wide potential, can influence the lives of workers for the better and most importantly, it becomes one of the main components of new scientific and technological revolutions that the society pursues (Bailey and Kurland, 2002), (Baruch, 2002).

Contrary to all the positive aspects, telework can generate marginalization and isolation (O'Neill et al. 2014), can increase stress for workers and can favour their exploitation (Loukidou et al., 2009). Although it can improve working methods, it can strengthen inequalities in the labour market in certain social groups (Belzunegui-Eraso and Erro-Garcés, 2020).

The year 2020 produced unprecedented changes in the global economy, on March 11 the World Health Organization (WHO) declared the new global pandemic coronavirus urging all governments around the world to take the necessary measures to protect themselves, so the first drastic measure was to block borders at the level of each country. This measure was followed by restrictions such as staying at home which directly influenced telework and all that it represented up to that point. In a scenario such as the COVID-19 pandemic, working from home was a very important aspect of ensuring business continuity, while normally the main benefits of teleworking would be reduced commuting time (Duxbury, Higgins and Neufeld, 1998) and the possibility of a better balance between family and professional life (Daniels et al., 2000). Under these conditions, the transition to telework was made instantly, without accommodation time, becoming one of the largest experiments in history.

The aim of this paper is to find common patterns in the perception of European workers regarding the adoption of remote work imposed by the COVID-19 pandemic, using unsupervised learning techniques. After applying unsupervised learning, supervised learning is applied to see which components can be improved considering telecommuting behavior during the COVID-19 pandemic in order to increase post-Pandemic adoption of home working among selected European countries in the analysis. The research is based on data from the Eurofound Selective Survey (2020) and provides an overview of what remote working has meant for European workers.

The research begins with the analysis of European workers most severely affected by the adaptation to telework imposed by the COVID-19 pandemic. The use of unsupervised learning techniques creates the premise of highlighting common patterns in the perception of European workers. At the level of each pattern of European workers, we aim to find typologies of groups of European countries where workers similarly perceive the forced adoption of telecommuting imposed by the COVID-19 pandemic and how affected or unaffected they were by it. Finally, Romanian workers are placed in a common pattern with other European workers. Supervised learning helps to predict the behaviour already defined by unsupervised learning and which components of the components defined by unsupervised learning it should improve to frame the groups in the adoption of work from home.

1. SCIENTIFIC LITERATURE REVIEW

1.1 Benefits and limitations of telework (aspects regarding productivity and well-being)

The literature reports numerous benefits of telework for both workers and organizations. The main organizational benefits include savings on real estate costs, savings on electricity and savings on parking spaces (Alizadeh, 2013). Employee benefits include better time management, better work-life balance, and the elimination of transportation costs for office travel (Baker and Avery, 2007). Job satisfaction and productivity are positively associated with telework (Pyoria, 2011). Contrary to these benefits, there are many limitations associated with teleworking. An example would be the limitations in terms of organizational management that include difficulties in managing tasks (Ruth, 2009). Previous studies indicate that social exclusion is a very important concern that has a negative impact on the productivity of teleworkers (Shadler, 2016). This, together with limited interaction with colleagues, can affect the long-term promotion and career progression of certain patterns of workers.

Telework could increase organizational and individual productivity (Silva, 2007). However, expectations for higher productivity from teleworkers are often created without a careful analysis of how managers perceive productivity. At the same time, the lack of an adequate measurement tool to effectively monitor productivity can often lead to inconclusive or even erroneous results from managers (Troup and Rose, 2012). According to the study (Westfall, 2004) four key factors are proposed that should be taken into account when measuring productivity: labour intensity, workload, work efficiency and organizational costs. In addition to these essential factors, others that have a positive impact on individual productivity should be included, such as: frequency of social interactions, completion of tasks, and managers' view of individual productivity (Warr, 2015).

Teleworking contributes significantly to the balance of workers in terms of work, personal life and family (Shadler, 2016). Worker well-being can be affected by work-related characteristics such as: level of difficulty and stress, task requirements, level of autonomy and social support (Alizadeh, 2013). These findings suggest that individual well-being is positively influenced by job perceptions and is consistent with flexibility in the workplace. Some studies on this concept focus on topics such as: work-life balance (Baker, 2007), teleworking trends, well-being, productivity (Vega, 2015), and the impact of culture on telework (Westfall, 2004).

1.2 The impact of telework upon professional activity

With the spread of COVID-19 worldwide, the impact of the pandemic on the labour market has become increasingly widespread, with the imposition of restrictions affecting most workers (Bick, 2020). For many of them, the introduction of physical distance measures has had major consequences such as reduced wages, increased working hours, introduction of additional sessions and even job losses (Reisenwitz, 2020). For a significant segment of the active population, the pandemic has led to a sharp increase in workload, as well as considerable changes in working conditions. Although a large proportion of employees have been able to continue working through telework, a very large number of work-related issues should be considered to ensure the safety and health of employees (Mullen, 2020). The European Framework Agreement on Telework provides guidance on the organization of personal work, which indicates that workload and performance standards must be the same for both teleworkers and employers' workers (ETUC, 2002). Teleworkers enjoy the same legal protection and have the same responsibilities regarding data protection, confidentiality, and access to training, safety, and work organization.

Telework research has repeatedly shown that employees working from home tend to work more hours than when working at the employer's premises (Mullen, 2020). The main reasons would be to replace travel time with work-related activities and to change the routine regarding the boundary between professional and personal life (John, 2019). Specific research on the working hours of employees working from home due to the COVID-19 pandemic has shown that 38% of them are willing to work overtime in the evenings or on weekends (McCulley, 2020). Employees who have children or other careers need to find extra time on their working day to do their homework, either starting work much earlier or working late into the evening. Also, the work schedule during the day may suffer an interspersed segmentation with short breaks dedicated to childcare or household chores (Kolakowski, 2019).

With no previous teleworking experience or only very limited experience, companies that worked from home in 2020 experienced a lack of clarity about the priorities and tasks they need to perform (McCulley, 2020). . Many organizations, where teleworking was non-existent, had to instantly adapt to an unknown work system (Eurasia Review, 2020). Isolation creates uncertainty among workers, leading to specific issues such as shyness in approaching other colleagues or shyness in asking for specific support.

2. RESEARCH METHODOLOGY

For the analysis of workers who have been severely affected by the adaptation to telework, data on European workers was used (Eurofound, 2020). The Eurofound selective research (2020) collected data in two stages, during the lockdown period and immediately after the lifting of the lockdown to discover the living and working conditions during the pandemic. The analysis includes two important aspects of the telework period, professional life, and personal life. There were questions about employment and quality of life, looking at the level of telework satisfaction from the workers' point of view. Eurofound (2020) identifies the most important areas of work affected by the COVID-19 pandemic, providing data on material aspects.

Information on teleworking and how it has been perceived by European workers is numerous and often contradictory. Techniques specific to unsupervised learning are intended to reduce chaotic informational content by finding common features (Chapelle et al., 2006). Unsupervised learning is a machine learning technique in which the presence of users during the process is not required (Pedrycz and Reformat, 2006). This allows the model to work on its own to discover patterns and pieces of information that were not previously detected (Girolami, 2002). Specific algorithms for unsupervised

learning are clustering, anomaly detection, neural networks, and principal component analysis. These allow the performance of very complex processing tasks compared to algorithms specific to supervised learning (Karpathy et al., 2014).

One of the main techniques used in the study is principal component analysis. This technique is relevant because it significantly reduces the number of dimensions without losing important information (Smith, 2002). It applies to a single set of variables that form different independent subsets. Variables that are correlated with each other (but are largely independent of other sets of variables) are combined and given the name of factors (Gorunescu, 2006).

The second unsupervised learning technique used in the study is clustering. It mainly deals with finding a structure or patterns from an uncategorized data collection. Clustering-specific algorithms process data and find the groups that can be formed.

Supervised learning methods are also used in the analysis Logistic Regression, defined as one of the most powerful supervised learning algorithms, being an extension of the general regression model. Logistic regression identifies the possibility that a new observation belongs to classes 0 or 1. Logistic regression works on the basis of probability: if the probability value is greater than 0.5 then it will fall into class 1, if it is less than 0.5, then it is observed, will fit into box 0. (Uddin S. 2019).

The research challenges referred to:

P1: Identify a small number of characteristics of European workers with regard to forced adoption of telework.

P2: Finding typologies of groups of European countries in which workers similarly perceive the forced adoption of telework imposed by the COVID-19 pandemic.

P3: Placing Romanian workers in a common pattern with other European workers.

P4: Deciding the components that need to be improved for some European countries that did not fall under the favour of working from home in order to favour it.

To meet the first challenge, the principle component analysis method was considered the most appropriate, as it reduced the number of variables and obtained a small number of key factors based on which were able to identify the characteristics of European workers.

Regarding the principal component analysis, the table specific to the descriptive analysis was obtained, respectively the mean and the standard deviation. For the whole data set, the covariance matrix was generated in order to illustrate the correlations useful for the study. To get principal components, the table of eigenvalues and the scree plot graph were analysed.

To achieve the next two challenges, the hierarchical clustering method was used. This offers the possibility to find certain typologies of groups of European countries in which workers similarly perceive the forced adoption of telework imposed by the COVID-19 pandemic, as well as the placement of Romanian workers in a common pattern with other European workers. The hierarchical clustering method is achieved by running the clustering algorithm and interpreting the results.

For the last point, after obtaining the main components from the principal components analysis, using the descriptive analysis together with the clustering on the resulting components, the components will be divided into countries with affected components and unaffected components. Clustering will be applied again on the new recoded database and a new column will be created with home workers and office workers, the column created again with the help of clustering. At the end, based on new columns and the data that built the column with the help of logistic regression, it will be predicted for certain countries that have not entered the class of home workers which component should be improved in order to favour home work in the future

3. RESULTS AND DISCUSSION

3.1. Statistical analysis regarding the influence of COVID-19 pandemic on telework

Most European countries have been negatively affected by the unemployment rate. During the pandemic period imposed by COVID-19, 28.1% of workers were negatively affected by the pandemic, ending up losing their job permanently or temporarily. Looking individually, at each country, it is observed that most workers who became inactive on the labour market during the pandemic are found in Greece (46.1%), followed by Cyprus (43.5%) and Slovenia (42%). There are also countries such as Denmark (10%), Sweden (11.4%) and the Netherlands (13%) that have managed differently the quarantine period of spring 2020 suffered globally, and which have managed to keep the unemployment rate lower by comparison with other European countries.

European lower-educated workers had a 4% higher job loss rate than those with higher education. This finding can be partly explained by analysing the type of employment contract; for example, 81% of workers have an indefinite (permanent) term, 16% a limited contract and 3% an agency or apprenticeship contract. According to data from the first quarter of 2020 taken from Eurostat, it is observed that among young people, 15-24 years, 45.6% have a temporary employment contract, while among those older than 55 years or over, 5 ,1% fall into the same category.

During the COVID-19 pandemic, most employees underwent changes in terms of workload. Eurofound (2020) data on the lockdown period show that 49% of all workers indicated a decrease in working hours. This situation has partially improved after lockdown was lifted where 37% reported that the number of working hours decreased. Moreover, a large number of workers (26%) stated that their work schedule had increased significantly. Official statistics show a decrease in working hours, according to Eurostat (2020) the number of working hours decreased by 3.7% in the euro areas in the first quarter of 2020, compared to the previous quarter.

States such as Italy (-9.7%), Slovakia (-8.7%), Greece (-7.9%) and Austria (-7.8%) recorded a decrease in the index of hours worked in the workplace in terms of comparison between the last quarter of 2019 and the first quarter of 2020, but there are also countries such as Finland that have seen an increase in hours worked during this period (+ 3%) (Labour Force Statistics, 2020).

In order to be able to observe the influences produced by telework during the COVID-19 pandemic on European workers, 121 indicators with a large influence on telework were analysed (Eurofound, 2020). These indicators refer to changes in working hours, fatigue felt after teleworking, the power of concentration of workers working from home, job insecurities, the type of contract held, and the emotional balance felt after telework. These indicators can be split into three groups, the first reveals the situation of the workers in terms of financial security felt, the second the balance between personal and professional life, and the third the changes in workload and performance. All these indicators are very relevant for the analysis, but the goal is to get the most accurate conclusions, so the factor analysis was used to reduce the number of indicators. Factor analysis uses a wide range of statistical techniques to represent a small number of variables.

The 121 indicators are processed in SPSS (Statistical Package for the Social Sciences). Data are collected at the level of European countries and do not require standardization because they are in the form of percentages.

The first results of interest for the study are obtained from descriptive statistics. Being a large database, those indicators of high importance will be deepened for the results of the study. Analysing the answers of the workers at European level, it is observed that on average 40% of them are tired after work, avoiding fulfilling their duties at home. At the same time, 39% of workers said that working hours did not change due to teleworking at home. Another interesting result is that 50% of workers say “I never feel like family responsibilities keep me from working time”, and another 36% declare “I never find it difficult to focus on work because of the family”, showing that among the workers there are many

individuals who carry out their activity as before, without any impediments. Contrary to the above results, an average 36% of workers are emotionally overwhelmed by workload and 23% often feel isolated when working. From these statistical results, based on the descriptive statistics, the conclusion is that there were two categories of people in Europe: some workers successfully coped with the trials of teleworking, and the other part felt emotionally overwhelmed, failing to find a balance between personal and professional life.

Telework has many advantages (Mann and Varey, 2000) such as saving resources, freeing up traffic, flexibility (Fonner and Roloff, 2010), lower costs for employers (Wellman and Haythornthwaite, 2008) and increased productivity (Azarbouyeh and Naini, 2014), but following the COVID-19 pandemic, a pertinent observation is that the impact of telework can also be negative. Lack of contact with co-workers leads to isolation, which influences other issues such as dissatisfaction with work and loss of interest in work.

3.2. Common characteristics of European workers

In the next part of the study the focus will be on the analysis of correlations. The strongest correlation exists between the indicators *Never (Found it difficult to concentrate on your job because of your family responsibilities)* and *Never (Found that your family responsibilities prevented you from giving the time you should to your job)* with a correlation coefficient of 0.836 at a significance level of 5%. This result is expected, as many workers among the workers have adapted perfectly to the telework at home. Reasonable, direct and significant correlations were found between the indicators *Increased a little (Change in working hours)* and *Daily (How often do you prefer to work from home)* with a value of the correlation coefficient of 0.592 and the indicators *I agree (I am satisfied with the amount of work I managed to do)* and I agree *(Overall, I am satisfied with my experience of working from home)* with a correlation value of 0.667. These correlations illustrate details about those European workers who managed to increase their determination and concentration, becoming more productive than when they worked in the office. Another significant correlation coefficient 0.631 helps to notice that there is a direct and significant correlation between the indicators *Rarely (You feel emotionally drained by work)* and *Rarely (You feel physically exhausted at the end of the working day)*.

Due to very weak correlations between certain variables, a decision to continue the analysis with only some of these indicators was made, so the most significant results for each component to be obtained.

Based on the table of eigenvalues, it was observed that the first main factor has an eigenvalue of 774,419 and retains about 24% of the initial

variable variation. In order to develop relevant conclusions, the cumulative percentage of information retained by the components must exceed the 70% threshold. Analysing the table of eigenvalues, it has been noted that the first 6 components retain about 73% of the variation of the variables.

Principal components

Table no. 1

Component 1	➢ (-)Never(Found that your job prevented you from giving the time you wanted to your family) ➢ (-)Never(Found it difficult to concentrate on your job because of your family responsibilities) ➢ Sometimes(Found that your family responsibilities prevented you from giving the time you should to your job) ➢ Rarely(You feel emotionally drained by work) ➢ (-)Never(You feel isolated when working)
Component 2	➢ Increased a little (Change in working hours) ➢ Rarely(Your colleagues or peers help and support you) ➢ (-)Agree(I am satisfied with the amount of work I managed to do) ➢ (-)Agree(I am satisfied with the quality of my work) ➢ (-)Agree(With the equipment I have at home I could do my work properly) ➢ Rarely(Physical contact) ➢ Decreased a lot(Performance) ➢ Strongly disagree(Overall, I am satisfied with my experience of working from home)
Component 3	➢ Decreased a lot(Performance) ➢ Rarely(Felt too tired after work to do some of the household jobs which need to be done) ➢ Never(Found it difficult to concentrate on your job because of your family responsibilities) ➢ Rarely(Found that your family responsibilities prevented you from giving the time you should to your job) ➢ Rather unlikely (Do you think you might lose your job in the next 3 months?) ➢ Always(You have the feeling you are doing useful work) ➢ Strongly agree(Overall, I am satisfied with my experience of working from home)
Component 4	➢ Several times a month(Work from home preference) ➢ Most of the time(Physical contact) ➢ Not very well informed(Informed about COVID-19 prevention) ➢ Neither agree nor disagree(I am satisfied with the quality of my work)

Component 5	➢ Rarely(Found that your job prevented you from giving the time you wanted to your family) ➢ Once or twice a week(Over the last 2 weeks, how often have you worked in your free time to meet work demands?) ➢ Rather unlikely(Do you think you might lose your job in the next 3 months?) ➢ Sometimes(You feel physically exhausted at the end of the working day) ➢ Increased a lot (Performance)
Component 6	➢ Less often(Frequency of working from home before the outbreak) ➢ Agree(Overall, I am satisfied with my experience of working from home)

The first main component is determined in a negative way by the indicators Never *(Found that your job prevented you from giving the time you wanted to your family)*, Never *(Found it difficult to concentrate on your job because of your family responsibilities)* and *Never (You feel isolated when working)*. This component includes those European workers who have been affected from a personal point of view in terms of telework at home during the COVID-19 pandemic. They did not manage to provide enough time for the family, they often focused very hard on the work they had to do, feeling all this time that they are isolated from colleagues and the work environment. These workers were emotionally affected by telework. There is no evidence of their performance, but in terms of family and work balance, they felt overwhelmed. On the positive side, the first component is determined by the indicators *Sometimes (Found that your family responsibilities prevented you from giving the time you should to your job)* and *Rarely (You feel emotionally drained by work)* that support the idea described above, as a conclusion of the results obtained. The first component is generically called the **Emotional impact of telework.**

The second component is positively determined by the indicators *Increased a little (Change in working hours)*, *Rarely (Your colleagues or peers help and support you)," Rarely (Physical contact)*, *Decreased a lot (Performance*) and *Strongly disagree(Overall, I am satisfied with my experience of working from home)*. This time the results are that European workers are dissatisfied with the experience of telework, but the main reasons are different from the first component. They underwent changes in the work schedule, low interest from colleagues and isolation from others. Overall, this experience negatively influenced their job performance, causing them to be dissatisfied with the experience of working from home during the COVID-19 pandemic. In a negative sense, this component is determined by indicators that

reveal details about the fact that teleworking does not provide them with the equipment they want, which is why the quality of work decreases. This second component will be named **Labour productivity in telework.**

The third component is determined in a positive way by all indicators. European citizens who are part of the component can be analysed from two points of view. First, by the indicators *Rather unlikely (Do you think you might lose your job in the next 3 months?)*, *Always (You have the feeling you are doing useful work)*, *Strongly agree (Overall, I am satisfied with my experience of working from home)* and *Decreased a lot (Performance)* it is noticed that European workers feel safe in terms of service, working hours have decreased and the work they do considers significant. Overall, they are satisfied with their professional life during telework. Secondly, the observation is that personal life is also in a good balance with the indicators *Rarely (Felt too tired after work to do some of the household jobs which need to be done)*, *Never (Found it difficult to concentrate on your job because of your family responsibilities* and *Rarely (Found that your family responsibilities prevented you from giving the time you should to your job)* highlighting this idea. The conclusion is that this component contains those European workers who adapted to telework during the COVID-19 pandemic and can generically be called **High degree of adaptability to telework**.

The fourth main component is determined in a positive way by the following indicators: *Several times a month (Work from home preference), Most often (Physical contact)*, *Not very well informed (Informed about COVID-19 prevention)* and *Neither agree nor disagree (I am satisfied with the quality of my work)*. These European citizens are not very well informed about the pandemic and prefer to work from home several times a month. This component is determined by indifferent workers regarding telework during the COVID-19 pandemic and can generically be called **Indifferent workers**.

The fifth component reveals a new category of workers. This is determined in a positive way by the indicators *Rarely (Found that your job prevented you from giving the time you wanted to your family), Once or twice a week (Over the last 2 weeks, how often have you worked in your free time to meet work demands)*, *Rather unlikely (Do you think you might lose your job in the next 3 months?)*, *Sometimes (You feel physically exhausted at the end of the working day)* and *Increased a lot (Performance)*. This information lead to the conclusion that these workers used the pandemic period to improve professionally, so this component can generically be called **Winners of telework**.

The sixth component is determined in a positive way by only two components *Less often (Frequency of working from home before the outbreak)*

and *Agree (Overall, I am satisfied with my experience of working from home).* This component contains those European workers who discovered telework in the pandemic and easily adjusted, so it can generically be called **Workers who easily accepted the transition to telework.**

3.2 Typology of European countries

The clustering method within each component was chosen in order to illustrate the finding of typologies of groups of European countries in which workers similarly perceive the forced adoption of telework imposed by the COVID-19 pandemic and to place Romanian workers in a pattern common with other European workers.

An analysis of the similarities and differences between European citizens will be conducted and those countries where the effects of telework during the COVID-19 pandemic had an impact like that recorded on Romanian workers will be identified.

Based on the first component, the emotional impact of telework, two main clusters were identified. Citizens who have been affected to a lesser extent on a personal and family level are found in countries such as Finland, Belgium, Portugal and Spain. Romania forms a secondary cluster with the Czech Republic, Bulgaria and Slovakia, an expected result due to the fact that all these countries belonged to the former communist bloc, the citizens being trained and educated in the same direction. A surprising result is that the countries of northern Germany, the Netherlands, Austria and Sweden were severely affected emotionally, although these countries are highly administratively evolved and have implemented telework in many organizations before the COVID-19 pandemic. The citizens of Romania differ from the citizens of northern Europe, having more similarities with those in the southern and central areas (Italy, Slovakia).

For the second component, Labour productivity in telework, which highlights those citizens affected in terms of performance, two main clusters were identified. Romania is in the second cluster, among the most severely affected countries, along with countries such as Poland, the Czech Republic and France. The countries that have not been so severely affected are Germany, Finland, Austria and Estonia. Romanian citizens are in a state of decline in terms of productivity along with countries such as Slovenia, Hungary, Ireland and Portugal. The second main cluster consists of citizens of Denmark, the Netherlands, Belgium and Estonia, who manage to maintain their best performance compared to all other European countries.

Depending on the component High degree of adaptability to telework, other two main clusters were identified which are divided as follows: one

cluster contains the countries Italy, Portugal, Spain, Belgium, Poland and France, and the second cluster contains all the other remaining countries. The citizens of the countries belonging to the first cluster have easily adapted to telework during the pandemic and are totally satisfied with this experience. Romanian citizens belong to the cluster with those who have not adapted so easily to teleworking, being very similar to Irish citizens. One explanation for this result may be the large number of Romanians currently working in Ireland. The countries of the former communist bloc Romania, Lithuania, Bulgaria differ from the northern countries Sweden, the Netherlands, Austria, and Germany in terms of adaptability to teleworking.

The component named after the study Indifferent Workers is divided into two main clusters, one consisting of 6 countries and the other 19. The countries of Central Europe have formed a main cluster (Slovakia, Czech Republic, Croatia, Poland, Slovenia) being the countries that Following the telework at home during the COVID-19 pandemic, they did not feel much change in terms of either personal or professional life. Northern European countries such as Germany, Austria, Sweden, and Ireland have formed a secondary cluster proving that there are similarities between their citizens. It should be noted that Romanian workers are like those in central Europe.

Depending on the Winners of Telework component, three main clusters resulted. Citizens who have had a very professional career from the Czech Republic, Slovenia, Estonia, Croatia, Slovakia, and Luxembourg, have used the telework period in order to evolve, dedicating their personal time to develop in the service they owned. In terms of Winners, Romanian workers are very similar to those in France and Italy.

For the sixth component Workers who easily accepted the transition to telework, findings are that the countries of Latvia, the Netherlands and Sweden have formed a main cluster. These countries are similar in terms of teleworking, easily adapting to the whole process caused by the COVID-19 pandemic. With small differences, but still similar, noticed that the citizens of Finland, the Czech Republic, Denmark, Slovenia, and Luxembourg form the second main cluster. The citizens of Ireland, Italy and Poland are very similar to those in Romania forming a secondary cluster.

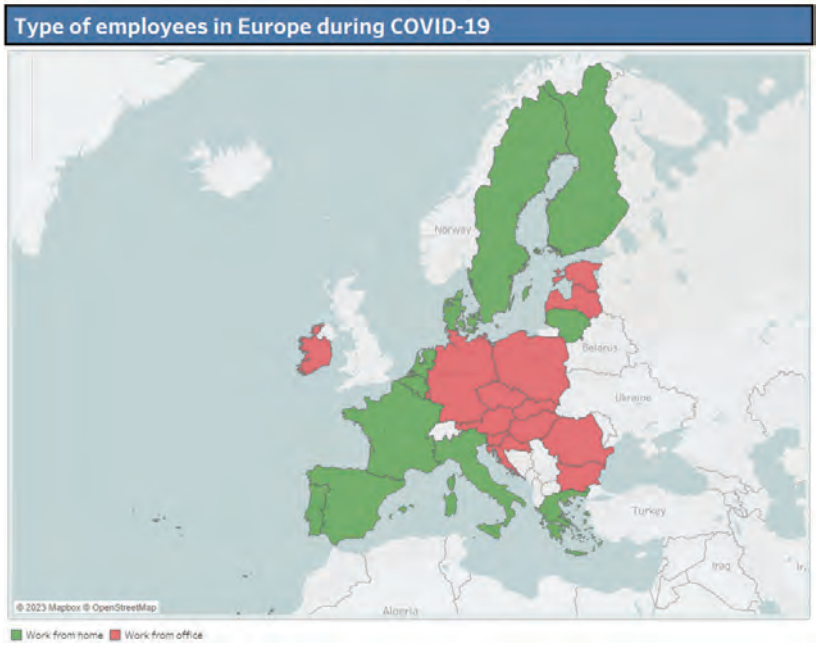
Regarding the placement of Romanian citizens in a common pattern with other citizens, Romania is often similar to Poland and Bulgaria on the emotional impact of telework and labour productivity in telework and Ireland on the high degree of adaptability to telework and workers who easily accepted the transition.

3.2 Logistic regression - prediction of workers' behaviour

After analysing the 6 components and using the clustering dendrogram together with descriptive statistics, for each component, the selected European countries will be divided into countries affected by the components and countries not affected by the component. For example, for the component – Emotional impact of Telework, Austria is classified as an affected country, because the employees of this country have a negative emotional impact regarding remote work during the pandemic.

Type of employees in Europe during Covid-19

Fig. 1



Most developed countries encourage working from home, countries such as France, Spain or Norway, but there are also very developed countries such as Germany that do not encourage working from home, they favour the type of office employee. Most of the eastern countries of the former communist bloc are classified as countries where employees prefer to go to the office.

Based on the new database obtained from the main components and on the new column with the type of worker, we will apply Logistic Regression. The algorithm will be applied in python with the help of the sklearn library, which is one of the most used machine learning libraries.

As the waves of Covid-19 started in 2020 and there is currently no deadline for them to stop, one must consider that working from home is a necessary thing and what people could do in the coming years, and how it would be in 2021, 2022, 2023, etc. to have the comfort of working from home. Thus, we will apply the logistic regression on 2 European countries, respectively on Germany, a country with a developed economy and on Romania, a European country from the former communist bloc with a developing economy, to find out which components must be unaffected.

For Germany, if at least the Emotional Impact of Telework or Labour Productivity in Telework were no longer negative, then Germany would also become a country in the future where working from home would be desirable. On the other hand, according to the logistic regression, even if they were Winners of Telework, they would not leave the office.

For Romanian employees to become employees from home, Labour Productivity in Telework should be improved, Romanians considering that working from home affects work productivity. At the same time, the High degree of adaptability to telework component needs to be improved, the employees in this country not being able to adapt very well to remote work.

CONCLUSIONS

The research developed a better understanding of what teleworking meant for European workers during the COVID-19 pandemic, highlighting both positive and negative aspects. Analysing the perceptions of European workers, it is discovered that teleworking had multiple and contradictory implications on teleworkers. Some of them declared themselves negatively influenced both personally and professionally, while at the opposite pole, the others said they were very satisfied.

Following the study, six main components were identified that illustrate different patterns of teleworkers. For the first component, the emotional impact of telework, it was discovered that the analysis is supported by the literature (Duxbury et al., 1998). The second component, labour productivity in telework, highlighted that group of European workers who were affected in this regard. Contrary to this component, Fonner and Roloff (2010) state that teleworking should increase productivity, but research has shown that there are also European workers who have been negatively influenced in terms of performance during the COVID-19 pandemic. The third component, high adaptability to telework, contains those European workers who are generally satisfied with the changes produced by telework, they do not professional excel, but are not affected in terms of family life. For the fourth component,

indifferent workers, it is found that this group of workers are uninformed about the pandemic and do not have conclusive opinions regarding telework. The fifth component reveals a new category of European workers, the winners of telework. Studies on them are numerous, most support telework and provide it as a solution for a better balance in terms of family and personal life (Azarbouyeh and Naini, 2014). The sixth component is generically called the workers who easily accepted the transition to telework and highlights those workers who discovered telework during the pandemic and adapted very easily.

Regarding the placement of Romanian workers in the context of other European workers, the cluster analysis was performed. As a result, it was found that in terms of the emotional impact of telework, Romanian workers resemble Bulgarians, Czechs and Slovaks, while in terms of labour productivity these are very similar to the Poles, Czechs and French. Highlighting the cluster analysis, Romanian workers are very similar to the Irish in terms of high adaptability and very different from workers in northern countries such as Sweden, the Netherlands, Austria and Germany. At the same time, Romanian workers are very different from Nordic workers in terms of indifferent workers, but similar to those in the countries of central Slovakia, the Czech Republic, Croatia, Poland and Slovenia. Analysing the winners of telework component, Romanian workers form the same cluster with French workers and Italian workers. Regarding the component of workers who easily accepted teleworking, Romanian workers form a secondary cluster with Irish, Italian and polish workers.

The limitations of the research are generated by the data on the perception of European workers whose reference time is the period July 2020. A useful comparison would be the analysis of the perceptions of teleworkers after a year of activity in relation to the results currently achieved. Obtaining new main factors may highlight other influences and effects (felt after a longer period of telework) or may support the conclusions gathered by strengthening the results assimilated from the study created.

References

1. Alizadeh, T. Planning Implications of Telework: A Policy Analysis of the Sydney Metropolitan Strategy.*Australian Planner* 50, no. 4(2013): 4-315
2. Azarbouyeh, A & Naini, S. (2014). A study on the effect of teleworking on quality of work life.Management Science Letters , 4(6), 1063-1068.
3. Bailey, D. E., & Kurland, N. B. (2002). A review of telework research: Findings, new directions and lessons for the study of modern work. Journal of Organizational Behavior, 23, 383–400. doi:10.1002/job.144 [Crossref], [Web of Science ®], [Google Scholar]
4. Baruch Y. Teleworking: benefits and pitfalls as perceived by professionals and

managers. New Tech. Work Employ. 2000;15(1):34–49. [Google Scholar] [Ref list]
5. Baker, E., G. C. Avery, and J. Crawford. Satisfaction and Perceived Productivity When Professionals Work from Home.*Research and Practice in Human Resource Management* 15, no 1(2007): 3-62
6. Belzunegui-Eraso, A.; Erro-Garcés, A. Teleworking in the Context of the Covid-19 Crisis. *Sustainability* 2020, *12*, 3662. https://doi.org/10.3390/su12093662
7. Chapelle, O., Schölkopf, B., and Zien, A. (Eds.) (2006). Semi-supervised learning, chapter Graph-Based Methods. The MIT Press.
8. Daniels, K., Lamond, D., & Standen, P. (2000). Managing telework: An introduction to the issues. In K. Daniels, D. Lamond, & P. Standen (Eds.), Managing telework: Perspectives from human resource management and work psychology (pp. 1–8). London: Thomas Learning. [Google Scholar]
9. Duxbury, L. E., Higgins, C. A., & Neufeld, D. (1998). Telework and the balance between work and family: Is telework part of the problem or part of the solution? In M. Igbariaand & M. Tan (Eds.), The virtual workplace (pp. 218–255). Hershey, PA: Idea Group. [Google Scholar]
10. Fonner, K. L., & Roloff, M. E. (2010). Why teleworkers are more satisfied with their jobs than are office-based workers: When less contact is beneficial. Journal of Applied Communication Research, 38(4), 336-361.
11. Girolami, M. 2002. Mercer kernel-based clustering in feature space. *IEEE Transactions on Neural Networks*, 13: 780–784
12. Gorunescu, F., 2006. Data mining. Concepts, Models and Techniques , (Cluj-Napoca, Romania)
13. Karpathy A, Toderici G, Shetty S, Leung T, Sukthankar R, Fei-Fei L (2014) Large-scale video classification with convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1725–1732
14. Loukidou, L., Loan-Clarke, J., & Daniels, K. (2009). Boredom in the workplace: More than monotonous tasks. International Journal of Management Reviews, 11, 381–405. doi:10.1111/j.1468-2370.2009.00267.x [Crossref], [Web of Science ®], [Google Scholar]
15. Mann, S., Varey, R. and Button, W. (2000). An exploration of the emotional impact of tele-working via computer-mediated communication, Journal of Managerial Psychology, Vol. 15 No. 7, pp. 668-690. https://doi.org/10.1108/02683940010378054
16. O'Neill, T. A., Hambley, L. A., & Bercovich, A. (2014). Prediction of cyberslacking when employees are working away from the office. Computers in Human Behavior, 34, 291–298. doi:10.1016/j.chb.2014.02.015 [Crossref], [Web of Science ®], [Google Scholar]
17. Pedrycz, A., and Reformat, M. 2006. Hierarchical FCM in a stepwise discovery of structure in data, *Soft Computing*, 10: 244–256
18. Pyoria, P. Managing Telework: Risk, Fears, and Rules. Management Research Review 34, no 4 (2011): 386-99
19. Ruth, S. Green It More than a Three Percent Solution?. Internet Computing IEEE 13, no 4 (2009): 74-8
20. Shadler, T. US Telecommuting Forecast 2009 to 2016- a Digital Home Report (2016)
21. Silva, C. From 60 to 0:Denver Group Halts Turnover at Aerea Businesses Through Telework.*Employee Benefit News* 21, no 8 (2007): p.44.
22. Smith, L., 2002. A tutorial on Principal Components Analysis, 12-13
23. Troup, C. and J. Rose. Working From Home: Do Formal or Informal Telework Arrangements Provide Better Work-Family Outcomes?. *Community, Work & Family 15*, no4, (2012): 471-86
24. Vega, R.P., Anderson, A.J. & Kaplan, S.A. A Within-Person Examination of the

Effects of Telework. J Bus Psychol 30, 313–323 (2015)
25. Warr, P. B. Decision Latitude, Job Demand, and Employee Well-Being. Work and Stress 4, no 4 (2015): 285-94
26. Wellman, B., & Haythornthwaite, C. (Eds.). (2008). The Internet in everyday life. John Wiley & Sons
27. Westfall, R. D. Does Telecommuting Really Increase Productivity?. Communications of the ACM 47, no 8 (2004): 93-6
28. World at Work. (2011). Telework 2011: A world at work special report. Washingtonfrom http://www.worldatwork.org/waw/adimLink?id=53034 [Google Scholar]
29. Bick, R. et al. 2020. A blueprint for remote working: Lessons from China. McKinsey & Company. 23 March.
30. Eurostat2020. Fall in hours worked more visible for women
31. Eurofound 2020. Fieldwork for EWCS 2020 halted due to COVID-19
32. Eurasia Review. 2020. Almost 90% Of Workers Would Be Willing To Continue Teleworking After The Pandemic. Eurasia Review. 4 May
33. ETUC-UNICE-EUAPME-CEEP (European Trade Union Confederation, Union of Industrial and Employers' Confederations of Europe, European Association of Craft, Small and MediumSized Enterprises and European Centre of Employers and Enterprises providing Public Services) 2002. European Framework Agreement on Telework. Brussels.
34. Jon C. (ed.). 2019. Telework in the 21st Century, An Evolutionary Perspective, ILO Future of Work series. Cheltenham, UK and Geneva: Edward Elgar and ILO
35. Kolakowski, N. 2020. COVID-19 Burnout Growing Among Remote Workers. 5 May.
36. Labour Force Statistics in OECD 2020. https://www.oecd.org/employment/oecd-labour-force-statistics-23083387.htm [accessed 22 January 2021]
37. McCulley, L. 2020. Lockdown: Homeworkers putting in extra hours - instant messaging up 1900%. 27 April
38. Mullen, C. 2020. Worker stress, resiliency a concern as company leaders consider return to office. 6 May.
39. Reisenwitz, C. 2020. How COVID-19 is impacting workers' calendars. Clockwise blog. 21 April.
40. Uddin S., Khan A., Hossain M.E., Moni M.A. - Comparing different supervised machine learning algorithms for disease prediction BMC Med. Inf. Decis. Making,. December 2019.

Urban Green Index estimation based on data collected by remote sensing for Romanian cities

Marian NECULA[1] (neculamarian18@stud.ase.ro)
Economic Cybernetics and Statistics Doctoral School,
The Bucharest University of Economic Studies

Tudorel ANDREI, Ph.D. (andreitudorel@csie.ase.ro)
Faculty of Cybernetics, Statistics and Economic Informatics,
The Bucharest University of Economic Studies

Bogdan OANCEA, Ph.D. (bogdan.oancea@faa.unibuc.ro)
The Faculty of Business and Administration, University of Bucharest

Mihaela PĂUN, Ph. D. (mihaela.paun@faa.unibuc.ro, mihaela.paun@incdsb.ro)
The Faculty of Business and Administration, University of Bucharest /
Bioinformatics Department, National Institute for R &D for Biological Sciences

ABSTRACT

The modernization of official statistics involves the use of new data sources, such as data collected through remote sensing. The document contains a description of how an urban green index, derived from the SDG 11.7 objective, was obtained for Romania's 41 county seat cities based on free data sets collected by remote sensing from the European and North American space agencies. The main result is represented by an estimate of the areas of surfaces covered with vegetation for the 40 county seat towns and the municipality of Bucharest, relative to the total surface. To estimate the area covered with vegetation, we used two data sets obtained by remote sensing, namely data provided by the MODIS mission, the TERRA satellite, and data provided by the Sentinel 2 mission from the Copernicus space program. Based on the results obtained, namely the surface area covered with vegetation, estimated in square kilometers, and the percentage of the total surface area or urban green index, we have created a national top of the county seat cities.

Keywords: *official statistics, experimental statistics, remote sensing, urban vegetation.*

JEL: *R140*

1. corresponding author.

1. INTRODUCTION

Land surface monitoring based on satellite imagery has been used since the 1960s and 1970s and has been a prolific source of important data for an extremely wide range of users. With initial applications in the military and meteorological fields, the monitoring capabilities have been widely expanded, for example the Sentinel missions within the Copernicus program (ESA, 2022a) generated in 2021 approximately 32 petabytes of data. In this context, there is an increased interest from the official statistics to explore and possibly incorporate this data source into the statistical production. Thus, within the ESSnet on Big Data II project, a pilot was carried out to explore the potential of remote sensing data (Work Package H, 2021) supported by several case studies, among which we list applications regarding the monitoring of surfaces, agricultural crops and natural vegetation, applications regarding the monitoring of environmental indicators included in sustainable development strategies, etc. Also, the the United Nations Statistics Division, through projects exploring Big Data data sources (Task Team on Earth Observations, 2017), estimates that this data source has a growing potential to be used in statistical production, either in complementing existing statistics or developing innovative statistics in response to the increased demand from data users.

Another argument for researching the potential of data sources collected through remote sensing derives from the modernization strategy of the National Institute of Statistics (INS, 2022) which considers the diversification of data sources and methods. The main limitation identified in the development of such projects is represented by the human resource that must have knowledge both in the field of official statistics and in the computational methods related to the exploitation of the new data sources.

Within the Sustainable Development Goal 11.7 regarding universal access to green public spaces, indicator 70 was proposed regarding the proportion of green spaces relative to the total area of urban settlements, in order to monitor the quality of life within human settlements classified as cities and to ensure equitable access to green infrastructure. The implementing of such an indicator is built on the assumption that a city whose green infrastructure is maintained and increased in relation to the total area, is translated into economic and social effort to ensure resilience to extreme climatic phenomena and improve the quality of life, e.g. reducing the effects of pollution, reducing crime, increasing residents' satisfaction, etc. (UN, 2022). At the European level, there are many initiatives regarding the sustainable development of cities and the conservation/expansion of green infrastructure and ecosystems, among which we mention the Urban Agenda (European Commission, 2016),

the European Strategy on Green Infrastructure (European Commission, 2013) or the decision of the European Parliament on the General Union Environment Action Program to 2020 (European Parliament 2013). Among the European initiatives for monitoring urban surfaces, the most important, from our point of view, is represented by the Land Surface Monitoring Service, within the Copernicus program of the European Commission developed in partnership with the European Space Agency (European Commission, 2022). The Copernicus program provides users, free of charge, with access to data obtained through remote sensing, through the Sentinel space missions, but also to data obtained through in situ collection (ESA, 2022c). Green urban spaces, according to the definition presented in the Mapping and Assessment of Ecosystems and their Services report of the European Commission (European Commission 2014), represent land surfaces fully or partially covered with vegetation, components of the green infrastructure of cities. The definition used does not specify what type of vegetation covers the respective surfaces and does not functionally address the type of use of the green space, i.e. whether it is public or private property, whether it is vegetation of natural origin or as a result of human development, whether the type of vegetation is interspersed between buildings and infrastructure or is continuous in nature, etc. Thus, two types of classifications of urban green spaces are presented that can contain the following atomic elements:

- green buildings (e.g. the roof is covered with vegetation);
- private property (alleys, singular trees, etc.);
- public gardens and parks;
- lands covered with vegetation intended for agriculture;
- lands covered with natural vegetation;
- areas of the region between water and land (shore, coastal areas, etc.).

Within the classification used in the Copernicus program (ESA, 2022b), of data collected in-situ through the CORINE Land Use Land Cover statistical research, green urban spaces are named green urban areas and can include: parks; ornamental gardens; private properties landscaped with vegetation; botanical and zoological gardens; public squares covered with vegetation; green spaces between glades; cemeteries; areas with vegetation for recreational purposes; etc.

This classification does not include: agricultural land included in the urban area, cemeteries outside the urban area and other possible types of surfaces covered with vegetation.

2. DATA AND METHODS

The present study uses two sources of remote sensing data, namely:

• Hyperspectral data Terra MODIS (Moderate resolution imaging spectroradiometer) were accessed through the Earthdata service (NASA, 2022a), an archive for accessing and distributing a large collection of remote sensing data. There are other web services through which MODIS data can be accessed and downloaded. The data are provided in the hierarchical data format – hdf5- (HDF Group, 2022), a scientific data transmission standard independent of the hardware or software architecture of computing machines. The MODIS hdf5 file contains the NDVI (Normalized Difference Vegetation Index) data and associated spatial attributes. The file contains the NDVI as well as the Enhanced Vegetation Index (EVI). More details about the NDVI vegetation index are presented later in this chapter. On average, the size of a MODIS hdf5 file for NDVI is between 150 and 250 megabytes (covering an area of about 5500 km^2). The period for which we downloaded the data is July 2022.

• Sentinel-2 multispectral data were accessed through the Open Access Hub (ESA, 2022d). The data is provided in SENTINEL-SAFE format (ESA, 2022e), a .zip archive containing metadata files (requirement of data preprocessing, data quality, etc.) and the actual image files in .jp2 (JPEG2000) format for those 13 spectral bands and other pre-calculated products (cloud mask, pixel value quality mask, surface classification mask, etc.). On average, the size of such an archive is about 1 gigabyte (referred to an area of 100km^2). The period for which the data was downloaded is between 01-Jun-2022 and 31-Jul-2022.

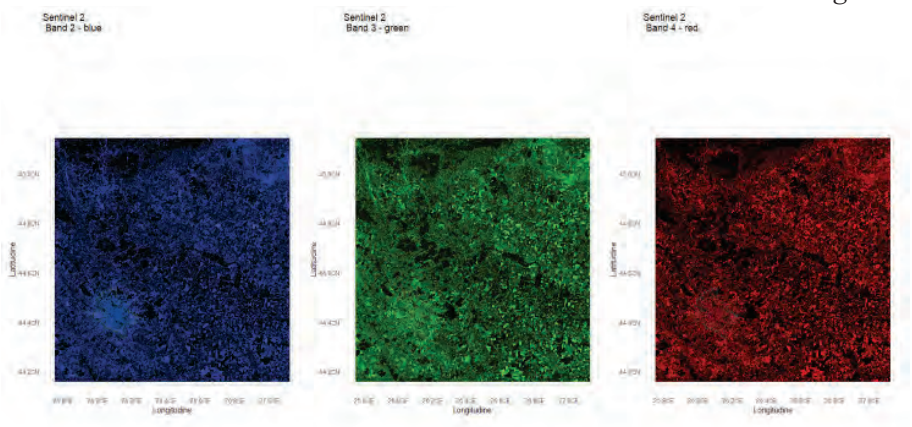
MODIS is one of the instruments for measuring the electromagnetic radiation reflected by the earth's surface, installed on the Terra satellite (NASA, 2022b) launched and operated by NASA in 1999. The objective of the mission is to monitor the earth's surface and atmosphere. The Terra satellite covers the entire Earth's surface on average every two days and through MODIS provides images at a spatial resolution of 250, 500 and 1000 meters. The instruments are capable of recording electromagnetic radiation (36 spectral bands) between 400 and 14400 nm. The data provided is pre-processed and prepared for use in scientific or other analyzes with minimal effort. MODIS datasets include, among others: vegetation indices; land surface temperature and temperature anomaly detection (fires); the reflectance of the earth's surface. Certain data sets are available daily, depending on the degree of pre-processing or spatial/ temporal coverage required by users.

Sentinel-2 is one of the space missions of the European Space Agency, with the objective of monitoring land surfaces (ESA, 2022a). The mission consists of two satellites launched between 2016 and 2017, Sentinel-2A and Sentinel-2B, with polar orbits, phase-shifted by 180 〖^∘〗, which have an average revisit time at the Equator of about 5 days. The satellites are equipped with a MultiSpectral Instrument (MSI), which can detect electromagnetic radiation between 400 and 2200 nm (13 spectral bands). The instrument can provide data at a spatial resolution between 10m and 60m. In Romania, the duration between two consecutive visits to the same area is about 10 days, on average, or about 3-4 times a month. The big disadvantage of the two instruments, in particular, the one installed on the Sentinel-2 satellites, is represented by the surfaces masked by clouds or extreme atmospheric effects. For both data sources, and/or other similar data sources, data sets are available, according to a convention, on pre-processing levels necessary to perform a certain type of analysis (Level 1A, Level 1B, Level 1C, Level 2A, Level 2B, etc.). In this case we have retrieved L2A (Level 2A) data sets, which have been corrected by the data provider, both geometrically and optically, and can be used directly in the analyses. Thus, the object of this analysis is represented by a dimensionless physical quantity called spectral (directional) reflectance, which represents the ratio between the radiance reflected by the surface of a material and the radiance incident on that surface and directly depends on the type/material of the surface (ISO, 1989).

In figure 1, the image obtained by the MSI Sentinel 2 instrument is represented graphically, by merging the spectral bands related to the reflectance in the visible range. When passing through the Sentinel 2 observation areas, it covers an area with an average width of 290 km.

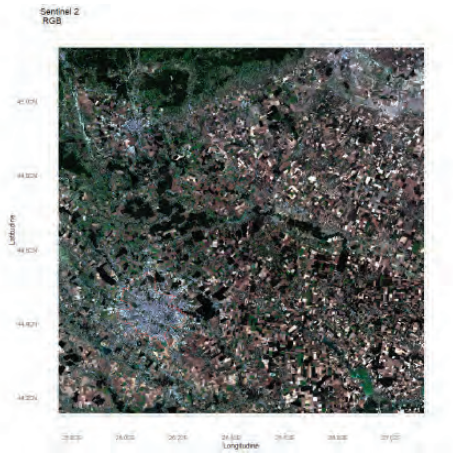
Sentinel 2 MSI related spectral bands in the visible range (RGB)

Figure 1



Composite RGB image for Sentinel 2
(June 2022, area T35TMK - Bucharest)

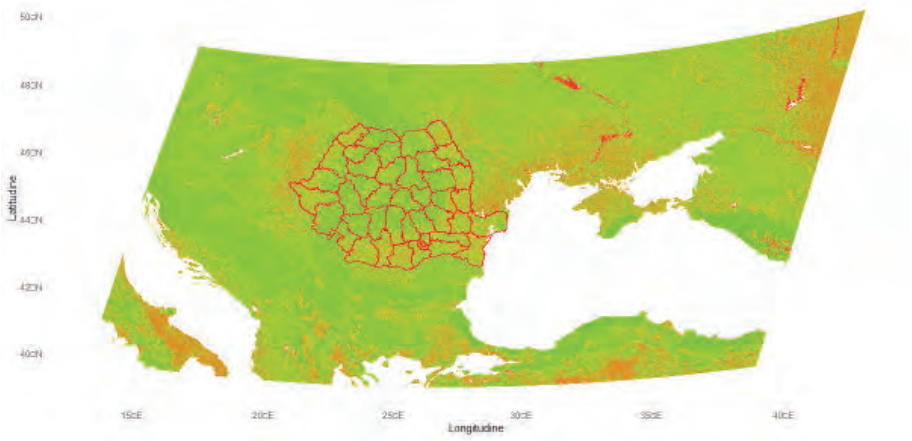
Figure 2



In the case of MODIS data, we downloaded the dataset containing the green index, represented in figure 3 for the entire surface of the country. The image was obtained by mosaic between area 194 and 204 related to the cartographic design mode used by MODIS. As a side note, MODIS covers an area 2330 km wide in a single pass.

MODIS image - normalized difference vegetation index
(June 2022, area 194 + area 204 Romania)

Figure 3



As auxiliary data, we used the cartographic delimitation of the county-seat municipalities, for all 41 counties and the municipality of Bucharest, graphically represented in figure 4. The delimitation used does not include the territorial administrative unit with the same name, and refers strictly to the urban, specific urban area component city of the county seat. For example, the territorial administrative unit Alba Iulia is composed of the city of Alba Iulia (within and outside the built-up area), Bărăbanț, Micești, Oarda și Pâclișa, from which we strictly choose the city of Alba Iulia.

The total data size is approximately 80 Gb for Sentinel 2 and 300 Mb for Modis. In the case of Sentinel 2, it was necessary to increase the temporal period of interest to 2 months, in order to identify data sets that do not contain the missing date or whose quality is not affected by the extreme presence of cloud cover of the surfaces of interest.

Romania. County residences and the municipality of Bucharest.

Figure 4

The normalized difference vegetation index is a popular measure in remote sensing data analysis applications for the identification/visualization of vegetated areas (Weier and Herring, 2000). The index is a dimensionless quantity, with values in the closed interval [-1, 1], quantifying the degree of vegetation coverage of the earth's surface. According to Huete, Justice and van Leeuwen (1999), the normalized difference vegetation index was created based on empirical observations of the interaction between electromagnetic waves, light in the red visible spectrum and near infrared spectrum, and vegetation.

Through the measurements, an increase in the intensity of radiation in the near infrared zone and the absorption of red light from the visible spectrum was observed. NDVI has some ability to discriminate between the types of vegetation covering an area (agricultural vegetation, forest, shrubs, etc.), as well as the quality of that vegetation (dry, green, etc.). The sensor-independent formula is:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

where
NIR = the associated near-infrared wavelength between 750 and 1400 nm
RED = the wavelength associated with red visible range between 620 and 750 nm.

For MODIS, the spectral bands associated with the two wavelengths are band 1, red visible range, and band 2, near infrared, and in the case of Sentinel 2, band 4, red visible range, and band 8, near infrared. However, there are differences between the calculation methods.

MODIS NDVI is already calculated based on the following algorithm:

Step 1: The NDVI index is estimated for each MODIS sensor record over an area of interest.

Step 2: The composite NDVI index for 16 days at a resolution of 250m per pixel is selected as the maximum of the NDVI value associated with the same pixel (spatial coordinates), and of course, the value is not affected by measurement errors, either due to the sensor or external factors.

The MODIS NDVI composite index is used in the paper. In the case of Sentinel 2, the index is calculated for a single record (calendar date) based on the previously described calculation formula and the availability of an image that is not affected by measurement errors, or excessive presence of clouds.

Some NDVI intervals for vegetation discrimination

Table 1

Category	Vegetation sub-category	STATCAN (Lantz, Grenier, Wang, 2021)	Weier and Herring, 2000 (sensor NOAA AVHRR)	ESA 2022f
Non-vegetation	N/A	[-1, 0.5]	[-1, 0]	<= 0.4
Vegetation	-	[0.5, 1]	(0, 1]	>0.4

In order to select an optimal threshold for the discrimination between areas covered with vegetation and those with non-vegetation, we performed manual comparisons between the images represented in natural (visible) colors provided by the Google Maps service (satellite image layer) and different intermediate discrimination thresholds starting from literature data, respectively the 0.5 to 0.7 NDVI threshold for MODIS, and the 0.3 to 0.6 NDVI threshold for Sentinel-2, both thresholds built with a step of 0.05. In table 1. The values for the discrimination threshold for each county seat city and the average used for the final calculation of the areas covered with vegetation are presented. Thus, for MODIS we obtained a value close to 0.58 and for Sentinel 2 we obtained 0.4.

From figure 5, we can see the sensitivity of the results to the selection of the discrimination threshold between the surfaces, the threshold being in an inversely proportional, almost linear relationship with the vegetation area detected by the green index.

Figure 6 shows the histogram of NDVI values for the Bucharest. From the graph it can be seen that the distribution of NDVI values is similar, but, on the one hand, there are differences derived from the different resolution of the two sensors and, on the other hand, differences between the parameters of the two distributions, e.g. MODIS values are concentrated around the 0.5 point, and Sentinel 2 values tend, roughly, to a point in the range [0.3-0.4]. These values are similar to the values resulting from the manual determination of the discrimination threshold used to estimate the green index.

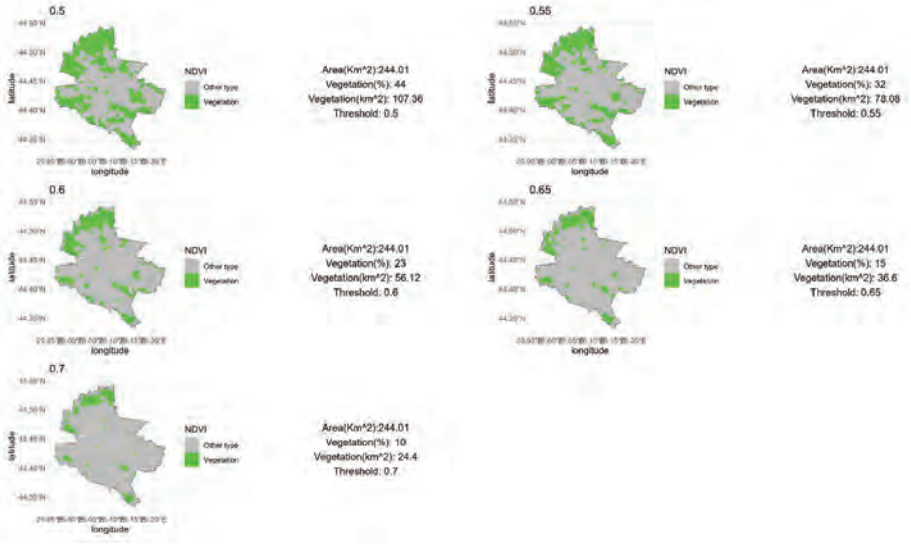
Vegetation discrimination threshold – manually estimated

Table 2

Residence	Modis	Sentinel 2
Alba Iulia	0.6	0.35
Alexandria	0.55	0.4
Arad	0.6	0.4
Bacau	0.65	0.45
Baia Mare	0.65	0.4
Bistrita	0.6	0.4
Botosani	0.5	0.35
Braila	0.6	0.45
Brasov	0.65	0.45
Bucuresti	0.7	0.45
Buzau	0.55	0.35
Calarasi	0.55	0.35
Cluj-Napoca	0.65	0.4
Constanta	0.55	0.4
Craiova	0.6	0.35
Deva	0.65	0.45
Drobeta-Turnu Severin	0.5	0.4
Focșani	0.55	0.4
Galati	0.6	0.35
Giurgiu	0.55	0.4
Iasi	0.6	0.4
Miercurea Ciuc	0.55	0.35
Oradea	0.65	0.35
Piatra-Neamt	0.65	0.45
Pitesti	0.6	0.5
Ploiesti	0.6	0.35
Ramnicu Valcea	0.6	0.4
Resita	0.5	0.45
Satu Mare	0.55	0.35
Sfantu Gheorghe	0.55	0.35
Sibiu	0.55	0.35
Slatina	0.55	0.45
Slobozia	0.6	0.45
Suceava	0.55	0.4
Targoviste	0.6	0.35
Targu Jiu	0.55	0.4
Targu Mures	0.5	0.35
Timisoara	0.5	0.4
Tulcea	0.55	0.35
Vaslui	0.5	0.45
Zalau	0.55	0.35
Mean	0.58	0.4

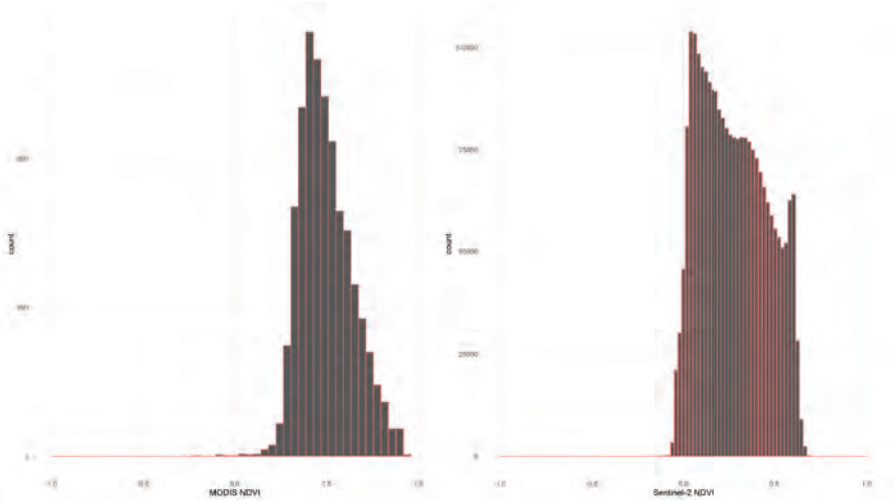
Vegetation discrimination threshold vs. non-vegetation – MODIS NDVI

Figure 5



Histogram of MODIS and Sentinel 2 NDVI values for Bucharest

Figure 6



a) Access methods

To access the data we developed two R procedures, one for MODIS and the other for Sentinel 2, using the getSpatialData packages (Schwalb-Willmann, 2022) and the sen2r package (Ranghetti et al, 2022). The download procedure requires a valid account on the EarthData data service, for MODIS, and on the Open Access Hub, for Sentinel 2. The procedure essentially involves selecting a product associated with a mission/sensor (i.e. MODIS TERRA), selecting a window temporal and a region of interest. The native map projection of the data was used for area calculation, sinusoidal projection for MODIS, respectively UTM/WGS84 for Sentinel 2 (it was not necessary to reproject the data for Sentinel 2, given that we did not encounter the problem of a city of residence being located at border region between two adjacent UTM zones).

b) Auxiliary data

As auxiliary data, we used the delimitation of the intra-urban areas of the county-seat cities in vector format, transformed to the native cartographic projection of MODIS and Sentinel data.

c) Data pre-processing.

In the case of MODIS, we extracted the regions of interest from the mosaic raster, applied the discrimination threshold and calculated the area covered by vegetation/the percentage of the total area covered by vegetation. For Sentinel 2 we created a pyramid formed by the spectral bands: B02, B03, B04 and B08, of which we used the first 3 associated with the visible light spectrum to create the "true" color image of the surface of interest, respectively B04 and B08 to calculate the NDVI index, then filtered with the associated discrimination threshold.

The R scripts used for computation are available at the following github address: https://github.com/MarianNecula/NDVI_experimental.git

3. RESULTS

Table 3 shows the results obtained after estimating the green index for MODIS and Sentinel 2 data. Between the two sensors, there are differences between the MODIS and Sentinel 2 green index values, either as a result of the spatial resolution, the processed MODIS data having the resolution of 250 meters per pixel, and Sentinel 2 data 10 meters per pixel. For example, in an area of a pixel (square) with a side of 250 meters or a pixel at the MODIS resolution, vegetation represents the majority class (> 50%) covering that

area, and the rest <= 50% is represented by other classes, such as built-up areas, then that pixel will be classified as belonging to a vegetated area. In the case of Sentinel 2, the observations being made much more granular, following the same example, the discrimination between a pixel associated with the vegetation class and a pixel associated with another class, is made at the level of a pixel (square) with a side of 10 meters.

Estimated area of the surfaces covered with vegetation for the county seat cities. -July 2022 Romania

Table 3

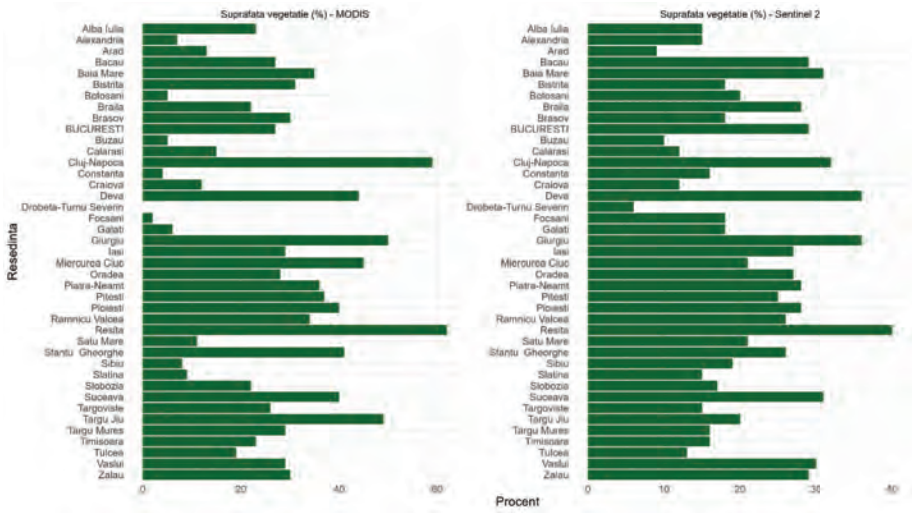
Name	Estimated Surface (km2)	The surface of the vegetation (%) - MODIS	The surface of the vegetation (%) - Sentinel 2	The surface of the vegetation (km2) - MODIS	The surface of the vegetation (km2) - Sentinel 2
Alba Iulia	15.52	23	15	3.57	2.33
Alexandria	9.62	7	15	0.67	1.44
Arad	45.57	13	9	5.92	4.1
Bacău	34.18	27	29	9.27	9.91
Baia Mare	31.75	35	31	11.12	9.85
Bistrița	16.7	31	18	5.18	3.01
Botoșani	15.37	5	20	0.77	3.07
Brăila	33.17	22	28	7.34	9.28
Brașov	37.55	30	18	11.3	6.76
București	244.51	27	29	66.23	70.86
Buzău	20.52	5	10	1.03	2.05
Călărași	19.41	15	12	2.93	2.33
Cluj-Napoca	93.09	59	32	54.95	29.8
Constanța	44.19	4	16	1.78	7.06
Craiova	44.28	12	12	5.31	5.31
Deva	12.43	44	36	5.47	4.47
Drobeta-Turnu Severin	12.98	0	6	0	0.78
Focșani	11.86	2	18	0.24	2.13
Galați	56.12	6	18	3.39	10.09
Giurgiu	26.39	50	36	13.23	9.49
Iași	46.42	29	27	13.53	12.52
Miercurea Ciuc	9.14	45	21	4.13	1.92
Oradea	54.69	28	27	15.3	14.75
Piatra-Neamț	18.29	36	28	6.61	5.12
Pitești	28.91	37	25	10.72	7.23
Ploiești	50.92	40	28	20.44	14.25
Râmnicu Vâlcea	10.11	34	26	3.44	2.63
Reșita	16.2	62	40	10.03	6.47
Satu Mare	21.81	11	21	2.4	4.58
Sfântu Gheorghe	10.95	41	26	4.5	2.84
Sibiu	25.57	8	19	2.05	4.86
Slatina	17.56	9	15	1.58	2.64
Slobozia	9.68	22	17	2.14	1.64

Suceava	29.98	40	31	12.03	9.29
Târgoviște	16.06	26	15	4.19	2.41
Târgu Jiu	25.46	49	20	12.48	5.09
Târgu Mures	23.6	29	16	6.85	3.78
Timișoara	68.87	23	16	15.82	11
Tulcea	17.39	19	13	3.33	2.26
Vaslui	9.57	29	30	2.79	2.87
Zalău	17.23	30	29	5.17	4.99

In Appendix we provide a few maps for the lowest (Drobeta Turnu Severin), highest (Reșița), ranked city, respectively the capital city in terms of vegetation coverage estimates.

Top county residences according to the percentage of the total area covered with vegetation

Figure 4



4. LIMITATIONS AND DISCUSSIONS

a) The low resolution of MODIS data can be an impediment, when estimates are needed for small areas, where vegetation is interspersed between buildings and is below the area covered by MODIS pixels (<250m). On the other hand, MODIS can provide time series starting from 2000, by comparison with data provided by the Sentinel 2 mission, starting in 2016. From the literature, we identified a series of algorithms that can be used to fuse MODIS data with those of Sentinel 2 to extend the time span while increasing the spatial resolution of the data.

b) Image quality is an important factor. The data can be affected by 2 problems: the complete or partial absence of data from a satellite pass through the area of interest, for example which can be triggered by errors at the sensor level, or the presence of massive clouds that prevent light radiation from reaching to the sensor. The latter type of problem can be partially countered with data coming from another type of sensor, the active monitoring of the earth's surface, of the radar type, which allows the detection of both landforms and natural or anthropogenic elements that cover the earth's surface. In this case, the complexity of data preprocessing increases accordingly, given the need to calibrate the radar signal for a multitude of factors.

c) The construction of a time series for this type of application is considered to identify if there have been substantial changes in the areas covered with vegetation, e.g. the uncontrolled expansion of residential areas within cities as a result of the boom in the real estate market over the last 20 years.

d) There are many other types of applications for green indices in agriculture or forestry that can be developed based on remote sensing data.

e) Within the project we are considering the design of a data dissemination application through a GIS application, with a spatiotemporal selection and comparison functionality between different areas and time periods.

CONCLUSIONS

This document describes a case study on the use of remote sensing data in official statistics. Starting from a similar project carried out by the Statistics Canada, we used data from the MODIS mission, and additionally data from the Sentinel-2 mission to estimate an urban green index and to make a classification of county residences in Romania according to of the percentage of the city's surface that contains a form of vegetation. At the NIS level (2021) there is a survey that collects and disseminates data on the area covered with vegetation at the level of cities (Verdure spots area in municipalities and towns by macro regions, development regions and counties - matrix GOS103B from the Tempo database), but there are differences at the level of definition of green space from cities. Thus, in the NIS survey, green spaces are considered: parks, public gardens, squares with vegetation, plots with trees and flowers, forests, cemeteries, sports fields and bases, by comparison with the present index that assimilates with vegetation all types of surfaces that contain vegetation, regardless of ownership or use. Using auxiliary data, models or algorithms can be created by which green spaces can be detected or reduced to those listed for the NIS survey to obtain an alternative to the current way of producing

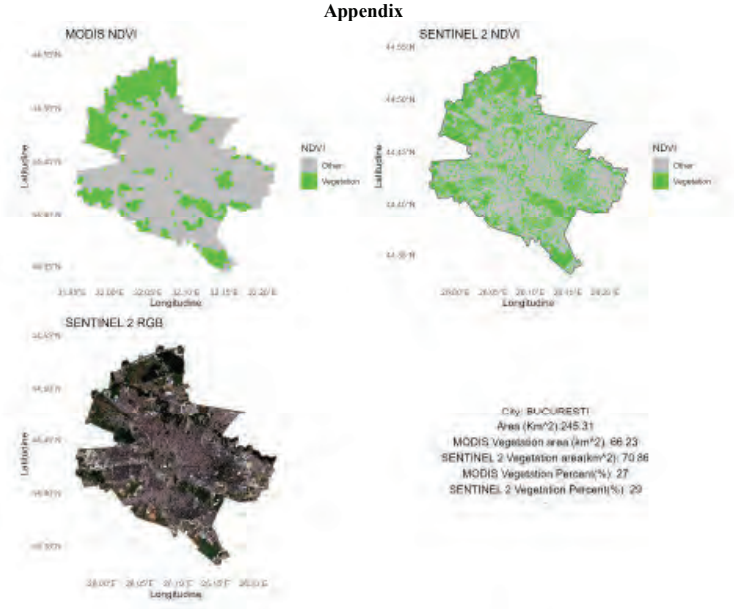
this type of statistic. Also, an argument in favor of the exploitation of this data source, the spatial and temporal granularity of the statistics can be considerably increased, from cities and municipalities to areas of arbitrary size, from an annual coverage to a monthly statistic or even with a bi-frequency monthly. A disadvantage is the considerable pre-processing effort, given the size and specificity of these data, while having a set of interdisciplinary knowledge (GIS, remote sensing, statistics). This classification can be used as a test in the experimental statistics section to check the potential interest of statistical data users relative to the advantages/disadvantages of the data source.

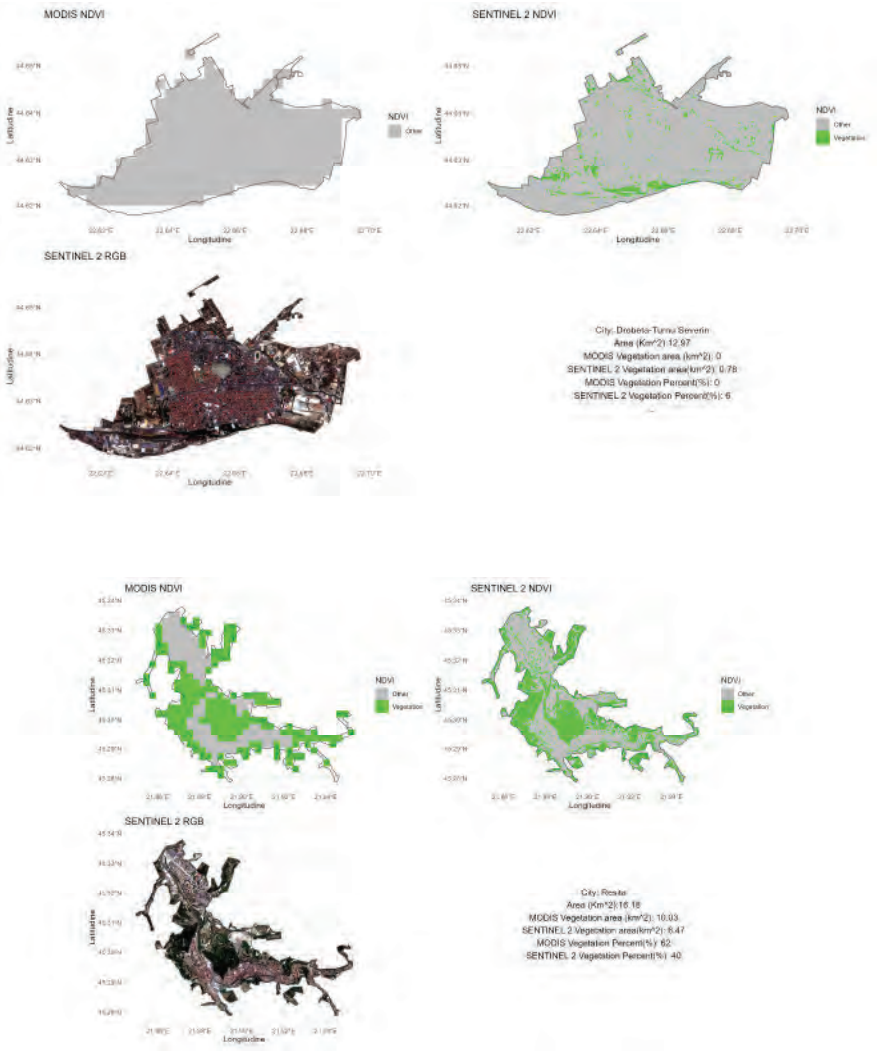
References

1. European Commission 2014. "Mapping and Assessment of Ecosystems and Their Services." https://ec.europa.eu/environment/nature/knowledge/ecosystem_assessment/pdf/102.pdf.
2. European Commission 2013. "The EU Strategy on Green Infrastructure." https://ec.europa.eu/environment/nature/ecosystems/strategy/index_en.htm.
3. ———. 2016 "The Urban Agenda for EU." https://futurium.ec.europa.eu/en/urban-agenda/pages/what-urban-agenda-eu.
4. ———. 2022 "Copernicus Earth Observation Programme." https://www.copernicus.eu/en/about-copernicus.
5. European Parliament 2013. "7th Environmental Action Programme." https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013D1386.
6. European Spatial Agency (ESA) 2022a. "2021 Copernicus Sentinel Data Access Annual Report." https://scihub.copernicus.eu/twiki/pub/SciHubWebPortal/AnnualReport2021/COPE-SERCO-RP-22-1312_-_Sentinel_Data_Access_Annual_Report_Y2021_merged_v1.0.pdf
7. ———. 2022b "Green Urban Areas Classification." https://land.copernicus.eu/.
8. ———. 2022c. "Land Monitoring Service." https://land.copernicus.eu/.
9. ———. 2022d. "Open Access Hub" https://scihub.copernicus.eu/dhus/#/home
10. ———. 2022e. "Data Formats" https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/data-formats
11. ———. 2022f. "Level-2A Algorithm Overview." https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm
12. Didan, K. Munoz, A. B. Solano, R. Huete, A. 2015 "MODIS Vegetation Index User's Guide". https://vip.arizona.edu/documents/MODIS/MODIS_VI_UsersGuide_June_2015_C6.pdf
13. HDF Group, 2022. "General HDF5 User's Guide". https://portal.hdfgroup.org/display/HDF5/HDF5+User+Guides
14. Huete, A. Justice, C. Van Leeuwen, W.J.D. 1999 "MODIS vegetation index (MOD13)" https://www.researchgate.net/publication/268745810_MODIS_vegetation_index_MOD13
15. ISO 1989. "Thermal Insulation - Heat Transfer by Radiation - Physical Quantities and Definitions." https://www.iso.org/standard/16943.html.
16. Institutul Național de Statistică, 2014. "Strategia de dezvoltare a Sistemului Statistic Național și a statisticii oficiale a României în perioada 2015-2020". https://insse.ro/cms/files/legislatie/programe%20si%20strategii/Stragtegie_2015-2020.pdf
17. Institutul Național de Statistică, 2021. Suprafața spațiilor verzi în municipii și orașe, pe macroregiuni, regiuni de dezvoltare și județețe. http://statistici.insse.ro:8077/tempo-online/#/pages/tables/insse-table

18. National Space Agency (NASA) 2022a. “EarthData” https://search.earthdata.nasa.gov/search
19. National Space Agency (NASA) 2022b. “Moderate Resolution Imaging Spectroradiometer.” https://modis.gsfc.nasa.gov/
20. ONU 2022. “Obiectivele Dezvoltării Durabile: Obiectivul 11.7 - Indicatorul 70.” https://indicators.report/targets/11-7/.
21. Ranghetti, L. Boschett, M. Nutini, F. Busetto, L. 2022. “sen2r: An R toolbox for automatically downloading and preprocessing Sentinel-2 satellite data” https://sen2r.ranghetti.info/
22. Schwalb-Willmann, J. 2022. “getSpatialData R package”. https://github.com/16EAGLE/getSpatialData
23. STATCAN, 2021. “Urban greenness, 2001, 2011 and 2019”. https://www150.statcan.gc.ca/n1/pub/16-002-x/2021001/article/00002-eng.htm
24. UNSD Task Team on Earth Observations (2017), “Earth Observations for Official Statistics. Satellite Imagery and Geospatial Data Task Team report”. https://unstats.un.org/bigdata/task-teams/earth-observation/UNGWG_Satellite_Task_Team_Report_WhiteCover.pdf
25. Weier, J. Herring, D. 2000 Measuring Vegetation (NDVI & EVI) https://earthobservatory.nasa.gov/features/MeasuringVegetation

Appendix




MODIS NDVI
SENTINEL 2 NDVI
SENTINEL 2 RGB
Latitudine
Longitudine
NDVI
Other
Vegetation
City: Drobeta-Turnu Severin
Area (Km^2):12.97
MODIS Vegetation area (km^2): 0
SENTINEL 2 Vegetation area(km^2): 0.78
MODIS Vegetation Percent(%): 0
SENTINEL 2 Vegetation Percent(%): 6
MODIS NDVI
SENTINEL 2 NDVI
SENTINEL 2 RGB
City: Resita
Area (Km^2):16.18
MODIS Vegetation area (km^2): 10.03
SENTINEL 2 Vegetation area(km^2): 6.47
MODIS Vegetation Percent(%): 62
SENTINEL 2 Vegetation Percent(%): 40