

VULNERABILITY AND RESILIENCE IN HEALTH CRISES. EVIDENCE FROM EUROPEAN COUNTRIES

Gina Cristina DIMIAN ¹, Simona Andreea APOSTU ^{1,2},
Maria Denisa VASILESCU ^{1,3}, Mirela Ionela ACELEANU ^{4*},
Josef JABLONSKY ⁵

¹*Department of Statistics and Econometrics, Bucharest University of Economic Studies,
Bucharest, Romania*

²*Institute of National Economy – Romanian Academy, Bucharest, Romania*

³*National Scientific Research Institute for Labour and Social Protection, Bucharest, Romania*

⁴*Department of Economics and Economic Policy, Bucharest University of Economic Studies,
Bucharest, Romania*

⁵*Department of Econometrics, University of Economics Prague, Prague, Czech Republic*

Received 20 May 2020; accepted 11 January 2021

Abstract. The purpose of this paper is to draw the first lessons from the on-going coronavirus crisis and to identify viable solutions for what should become the goal of any country: transforming their own economies into sensitive and responsive economies regarding public health problems. The originality of our approach is given by its objective as well as the strategy employed for verifying research hypotheses. The objective is twofold: detecting the indicators that may constitute signals for the vulnerability of countries in times of health crisis and highlighting the underlying factors of the resilience capacity. Many indicators have been considered: six indicators concerning Covid-19 pandemic and 27 socio-economic indicators. Three main hypotheses have been formulated and tested using various statistical methods. Our findings provide deep insights for understanding how Covid-19 crisis is correlated to specific economic (urbanization, sectorial employment, health system) and demographic factors (aging, mortality). The study has succeeded in identifying the pattern of a country with greater resilience and better ability to cope with a health crisis. Our results could be useful when forecasting the spread of another pandemic wave, its impact on people's health and estimating how some markets will be reshaped.

Keywords: health crisis, Covid-19, vulnerability, resilient capacity, Principal Component Analysis, TwoStep Cluster Analysis.

JEL Classification: I12, I11, I18.

Online supplementary material: Supporting information for this paper is available as online supplementary material at <https://doi.org/10.3846/tede.2021.14753>

*Corresponding author. E-mail: aceleanu_mirela@yahoo.com

Introduction

Recently, a number of events have dramatically increased the perception of instability, insecurity and uncertainty around the globe (Christopherson et al., 2010). Today's world faces multiple threats: violence, wars, natural disasters, famine, economic crises, terrorism, political instability, and, most recently, the Covid-19 pandemic.

Estimating the economic impact of a pandemic disease is very difficult, because even a disease with a relatively small impact on health can have a major economic effect. Kilbourne (2004) said that the main effect of a flu-like disease is not mortality, but morbidity – which is enormous – leading to schools closing, absenteeism on the labour market, declining production and crowded emergency hospital wards. This is the immediate effect after the pandemic sets in, but on the long term there are serious consequences affecting all sectors.

The health crisis caused by Covid-19 affects the entire economy; due to pandemic, many companies have closed or reduced their activities, letting people unemployed, in danger of losing their jobs or with diminished salaries. At the same time, the working processes have changed, due to the switching to online work, resulting in new challenges for the employees in balancing work-life activities. In many countries severe blockages have been observed, leading to consumption decrease and production discontinuities. Also, consumers have changed their consumption patterns leading to shortages for several goods in supermarkets around the world. Tourism and recreational activities were reduced or even completely blocked, resulting in huge financial losses. Global financial markets have fallen sharply and it is estimated that this health crisis will lead to a deep financial and economic crisis.

Obviously, during this period, the interest given to health must prevail, the measures for saving lives being essential. But, at the same time, some measures must be taken to mitigate the economic shock. Politicians face a trade-off between the safety of citizens which involves strict containment measures and permitting certain economic activities, decisions that are not easy to make, especially given that not much information is yet known about this disease. Decisions are of course taken depending on the situation of each country: number of cases, number of deaths, mortality rate, performance of the country's health system, demographic and health characteristics, specific economic and social issues.

In this context, the aim of this paper is to draw the first lessons from the on-going coronavirus crisis and to identify viable solutions for what should become the goal of any country: transforming their own economies into sensitive and responsive economies regarding public health problems. Facing a pandemic, vulnerability or resilience may be displayed. A country that knows its weaknesses in the event of a health crisis and already has the measures to alleviate them will have a significant advantage. Also, the resilience, the ability to take rapid countermeasures and have clear and effective policies to overcome such crises, is essential. Therefore, our objective is twofold: 1) detecting the indicators that may constitute signals for the vulnerability of countries in times of health crisis and 2) highlighting the underlying factors of the resilience capacity. In order to address these objectives, we organized the rest of the article as follows: Section 1 contains the literature review, Section 2 is dedicated to describing the data and methods used, Section 3 presents the results obtained and the last section focuses on highlighting the main conclusions.

1. Literature review

In December 2019 in Wuhan, China, a new type of coronavirus was identified, which caused many people to become ill. The number of cases rapidly increased, spreading globally. On March 11, 2020, the World Health Organization declared Covid-19 a pandemic. The effects of the pandemic became visible very quickly, causing problems in many societies, disrupting economic activity and affecting the lives of millions of people.

Measures to limit the impact of pandemic diseases on health and the economy can sometimes be in opposition (Yoldascan et al., 2010). Analyses of pandemics and economic effects highlighted how an economy can be affected by a severe flu outbreak: reduced consumption for tourism, increased transportation costs and retail disruption; increased work absence; school closure; and higher demands for medical services. The four types of economic shocks caused by influenza pandemic, identified by Verikios et al. (2011), are: (1) an increased demand for hospital health care and other medical services; (2) a temporary increase in sick leave and school closures requiring the withdrawal of parents from the workforce; (3) mortality among the active population, associated with permanent reduction of the workforce; and (4) temporary reductions in tourism and international business travel. In such moments of crisis, the activity sectors in which remote work is easy to implement must be stimulated, for example the IT sector, especially since the positive impact of ICT activities on economic growth is proven (Cioaca et al., 2020).

In the XXth century the population faced 3 influenza pandemics (1918, 1957 and 1968) (Patriarca & Cox, 1997).

Karlsson et al. (2014) studied the impact of the 1918 influenza pandemic on economic performance in Sweden, the results indicating that capital returns were negatively affected, but the pandemic did not significantly influenced revenues. Most studies have shown that the 1918 flu pandemic had economic effects on short term. Many businesses, especially in the services and entertainment industry, have lost revenue; in contrast, healthcare companies have registered an increase in revenue. Academic research has shown that the 1918 flu pandemic caused a shortage of labour, leading to higher wages for workers.

During the 1957 flu pandemic, 25% of population became infected with the new pandemic virus (Henderson et al., 2009). In Singapore, between 10% and 30% absenteeism in commercial firms significantly influenced the economy in 1957–1958 (Lee et al., 2008). A model of transmission dynamics was studied using data from 1957, the results showing that school closure depended on contact patterns in the population, the interaction of different age groups and the moment when the measure was implemented. The authors concluded that the closure of schools and nurseries reduced the size of the epidemic “by only a small amount” (Vynnycky & Edmunds, 2008).

In 2003 the severe acute respiratory syndrome (SARS) has caused illnesses and deaths, disrupting economic activity in many Asian economies. The Asian Development Bank estimated an economic impact of \$ 18 billion in East Asia (Yoldascan et al., 2010) and \$ 40 billion around the world (Maffioli, 2020). The results of the pandemic were a sharp drop in tourism and travel, a decrease in consumer confidence, a decline in domestic demand and pessimism about business prospects. Caballero-Anthony (2005) highlighted the conse-

quences on national and regional stability in East and Southeast Asia, the results showing that infectious diseases affect state security and undermine the survival and well-being of communities.

On May 2009 in Mexico appeared a new virus, H1N1 (“swine flu”), within five weeks spreading in 74 different countries and infecting hundreds of millions of people (Presanis et al., 2009). The regions with greater economic integration in the world (through international trade) were more affected by the pandemic (Verikios et al., 2011).

In 2014 appeared the West African Ebola virus disease (EVD), its economic impact being estimated at \$25.2 billion (Maffioli, 2020). Kieny et al. (2014) analysed the problems caused by the absence of a strong health system in case of EVD, the results indicating that economic growth, government expenditure on health and investment in health systems all decreased.

Keogh-Brown et al. (2010) examined the potential economic cost of a modern pandemic in the United Kingdom, France, Belgium and the Netherlands and the results suggested 0.5–2% decreases in GDP due to the disease, but school closure and prophylactic absenteeism tripled these effects. The short-term impact of a pandemic on economy will be reflected in a significant reduction in trade and services, and on the long term, potential economic growth will decrease and poverty will increase (Bloom et al., 2005).

In only a few months, Covid-19 pandemic has affected the global economy, having serious socio-economic implications. Economically, in the first quarter of 2020 the markets registered a decrease, some indices reaching a lower level than the market collapse in 1992. Some businesses were closed, people lost their jobs, with negative consequences even after the end of the pandemic (Chohan, 2020).

Furthermore, the short-term share of credit by foreign banks, a key indicator of external vulnerability, grew; aviation industry suffered major losses, as a result of travel restrictions and drastic declines in air travel demand; the global manufacturing decreased and is expected to further decline; the tourism sector faced an unprecedented crisis, with declines of 60% to 80% in international tourist arrivals; disruptions to global supply chains in agriculture; and growing existing income inequalities show the impact of Covid-19 in the world economy (UNSD, 2020).

Fernandes (2020) studied the economic impact of the Covid-19 health crisis on 30 countries industries, estimating the GDP growth. The optimistic scenario implies a median decline in GDP in 2020 of –2.8%, in other scenarios GDP may fall by more than 10%, and in some countries, more than 15%. Service-oriented economies, dependent on foreign trade, will be affected and many people will lose their jobs.

Covid-19 has profoundly affected healthcare system, especially in low and middle income countries, traffic border restrictions being implemented in order to “break the chain” of infection and control its transmission (Mitra et al., 2020).

Liu et al. (2020) studied Covid-19 disease in six representative cities in China and the results showed that social contact explained the transmission patterns, as well as the associated risks. The most infected people are those between 23–44 and 45–64 years old, because are more often involved in social activities than the others. The link between migration and the spread of the virus was also considered in the literature, the results indicating that countries would have been better prepared to react if they had monitored migrant flows, which are a

good predictor for the spread of Covid-19, especially in regions known for labour mobility, such as the European Union (Sirkeci & Yucesahin, 2020; Şerban et al., 2020).

Peng et al. (2020) studied the management and treatment of Covid-19 in China. The results showed that drastic quarantine measures prevented the spread of SARS-CoV-2, the earlier the implementation, the less likely resources will be depleted.

In the case of a pandemic, similar measures can have different results, the impact of the pandemic depending on many factors. Investigating these typical effects is particularly important at the country level, in order to identify specific vulnerabilities.

Sarewitz et al. (2003) defines vulnerability as an inherent characteristic of individuals, communities, networks, infrastructure, and systems, capable of producing potential effects, regardless the risk of a particular shock, such as economic crises, natural and man-made disasters, or even terrorist attacks.

Analysing possible sources of vulnerability in the event of a pandemic, O'Sullivan and Bourgoin (2010) concluded that interdependencies in modern society are a disadvantage in this context. Hospitals and clinics depend on utilities and telecommunications, pharmaceuticals are closely related to transportation, the production of certain essential goods is relocated to other countries.

Moore et al. (2016) constructed an infectious disease vulnerability index, taking into account variables from several key areas: demography, health, disease dynamics, political environment and economic development. Their results pointed out that the most vulnerable countries to an infectious disease are in Africa.

Gilbert et al. (2020) assessed the readiness and vulnerability of African countries to their risk of importing Covid-19, in relation to the healthcare capacity of a country. The results showed that the countries with the highest import risk have a moderate or high capacity to respond to outbreaks, and the countries with moderate risk have a variable capacity and a high vulnerability.

When dealing with powerful shocks, such as the Covid-19 pandemic, a very important feature is resilience. According to the literature, three main definitions of resilience were identified: i) the ability to recover from a shock, ii) the ability to resist to a shock and iii) the ability to adapt after a shock or to develop new growth paths (Boschma, 2015). Modica et al. (2018) identified two major characteristics of resilience: the ability to recover from shocks and the degree of preparedness.

Resilient societies and economies are those capable of absorbing different types of shocks. There is a need for a balance between economic, social and political stability on the one hand and dynamism and creativity on the other, necessary to identify optimal solutions to overcome crises (European Commission, 2020).

Human health is important for the health of the economy. The uncertainty and fears that people feel about their lives influence demand and supply. Because of that, the efficiency of economic measures depends, primarily and mainly, on the success of medical measures and pandemic control (Marinescu, 2020).

Hynes et al. (2020) highlighted the importance of resilience at several levels. At the national level, the strategic need is to support economic survival and recovery in the Covid-19 context, taking into account the local demand and the possibility of local/regional production

and distribution. Another level is the company resilience, with a focus on preventing layoffs and bankruptcies. The third level is the household resilience, targeting the individual household's needs – necessary goods, services and social support.

Smith-Bingham and Hariharan (2020) stated that the businesses that invest in financial and operational resilience procedures in order to manage global economic risks have a competitive advantage, being better prepared to respond to this crisis and find recovery strategies.

Oppenheim et al. (2019) designed an Epidemic Preparedness Index, considering that a country's level of preparedness for an epidemic/pandemic is essential in order to respond effectively. Their index comprises five areas: economic resources, public health communications network, public health system, infrastructure, and institutional capacity. The results indicated that the most prepared countries are found in Europe and North America.

For the USA, a guide to pandemic resilience was recently developed and the main strategies identified as being effective were: innovative methodologies for testing a large number of people, ensuring the appropriate medical supply and infrastructure, development of public health facilities, increase in medical staff, and clear mechanisms, norms and procedures (Edmond J. Safra Center for Ethics, 2020).

2. Data and methods

Deriving from the objective, our research hypotheses are the following:

H1: Economic development grounded on connectivity (with risk areas) increases countries vulnerability to pandemics (cases, deaths and case fatality rate).

H2: Advanced age and associated diseases that raise the overall mortality rate increase countries vulnerability to pandemics (case fatality rate).

H3: The performance of the health system contributes to diminishing the negative effects of a pandemic.

We used in our research a number of 6 indicators of the Covid-19 pandemic and 27 socio-economic indicators. These latter were grouped in 3 categories: *population health-related indicators* (life_expectancy, age_65_plus, median_age, obesity, smoking, pm2.5, basic_sanitation, health_expenditure, physicians, hospital_beds), *development and connectivity indicators* (GDP, agr, ind, serv, FDI, exp, imp, international_tourism_expenditure, international_tourism_receipts, hotel_and_restaurants_households_expenditure, pop_density, urban_population) and *mortality indicators* (deaths_total, crude_death_rate, mortality_households and pollution, mortality_severe_diseases, mortality_sanitation).

Among Covid-19 indicators, 5 variables (*case_fatality_rate*, *deaths_rate*, *cases_rate*, *days_peak* and *days_first_death*) were computed based on the data retrieved from European Union databases – the European Center for Disease Prevention and Control, whereas the *tests_rate* was computed based on the data retrieved from the Statista database. Socio-economic indicators were retrieved from the World Bank database – World Development Indicators and the Eurostat databases.

Covid-19 indicators were calculated as rates, expressed in percentages for case fatality rate and per 100,000 populations for the rest of the indicators. The reference day for the Covid-19

indicators was 13/04/2020. Socio economic indicators were computed as averages of the most recent three years for which data were available for the sake of the stability of the estimation. All the indicators are presented in Supplementary information 1.

In order to test the 3 hypotheses of our research, we used a 5-step methodology: i) correlation analysis (parametric and nonparametric methods); ii) normality testing and outliers detection, iii) Principal Component Analysis, iv) TwoStep Cluster Analysis and v) Mann-Whitney U Test.

i) First, we have used correlation analysis with the purpose of identifying the association or the relationships between the two categories of variables included in our research: socio economic indicators and Covid-19 indicators. For the sake of the stability of the results, taking into consideration that the number of observations is not very large, we have used both a parametric (*Pearson linear correlation coefficient-PC*) and a non-parametric method (*Kendall rank correlation coefficient-KC*).

ii) Second, we have evaluated the data set in terms of the normality of the distribution of the variables that proved to correlate with the Covid-19 pandemic outcomes, as a prerequisite for the application of the Principal Component Analysis. The normality of the distribution was tested using the *Shapiro-Wilk test*, which is more appropriate when the sample is not too large.

In addition, in order to respond to the third hypothesis of the research, namely identifying the factors that reflect the resilience capacity, we detected the countries that qualify as outliers for certain variables and analysed their situation in terms of how they dealt with the Covid-19 pandemic. With this purpose, we used in the analysis the box plots graphs in which the length of the boxes is a measure of the interquartile range whereas the outliers are represented by the scores that are greater than 1.5 times the interquartile range (Ghasemi & Zahediasl, 2012). The kernel density estimation allowed us to identify groups of countries with common features in terms of the factors of interest for our analysis.

iii) Third, we have used *Principal Component Analysis* (PCA) with the main purpose of synthesizing in a smaller number of new variables as much as possible of the variation of the initial variables, so that we can obtain fundamental factors that can explain differences between countries in terms of the response to the Covid-19 pandemic. In this respect, the scores of the principal components obtained in this stage were used in the cluster analysis for grouping the countries.

The PCA algorithm begins with the vector of the p initial variables:

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} \tag{1}$$

with population variance-covariance matrix:

$$\sigma^2(X) = \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12}^2 & \dots & \sigma_{1p}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \dots & \sigma_{2p}^2 \\ \vdots & \vdots & \dots & \vdots \\ \sigma_{p1}^2 & \sigma_{p2}^2 & \dots & \sigma_p^2 \end{bmatrix}. \tag{2}$$

The first principal component is the linear combination of the elements of X , which has the maximum variance.

$$C_1 = \alpha_{11}X_1 + \alpha_{12}X_2 + \dots + \alpha_{1p}X_p, \quad (3)$$

with the vector of the coefficients:

$$\alpha_1 = \begin{pmatrix} \alpha_{11} \\ \alpha_{12} \\ \vdots \\ \alpha_{1p} \end{pmatrix} \quad (4)$$

chosen such as to maximize:

$$\text{var}(C_1) = \alpha_1' \sum \alpha_1 \quad (5)$$

subject to the following constraint:

$$\alpha_1' \alpha_1 = \sum_{j=1}^p \alpha_{1j}^2 = 1. \quad (6)$$

The second principal component retains maximum from the remaining variance and it is uncorrelated with the first principal component:

$$C_2 = \alpha_{21}X_1 + \alpha_{22}X_2 + \dots + \alpha_{2p}X_p \quad (7)$$

with the vector of the coefficients:

$$\alpha_2 = \begin{pmatrix} \alpha_{21} \\ \alpha_{22} \\ \vdots \\ \alpha_{2p} \end{pmatrix} \quad (8)$$

chosen such as to maximize:

$$\text{var}(C_2) = \alpha_2' \sum \alpha_2 \quad (9)$$

subject to the next two constraints:

$$\alpha_2' \alpha_2 = \sum_{j=1}^p \alpha_{2j}^2 = 1 \quad (=10)$$

and

$$\text{cov}(C_1 C_2) = \alpha_1' \sum \alpha_2 = 0. \quad (11)$$

The vector which explains the most variance of the data is the first eigenvector. Likewise the second vector which explains most of the remaining variance is the second eigenvector, etc.

iv) *TwoStep Cluster Analysis* (TSCA) has been chosen due to some advantages that it offers compared to the other classical methods, K-Means and Hierarchical clustering: it can handle both continuous and categorical variables, rapidly analyses large data sets and the number of clusters is automatically selected.

This method implies a two stages procedure: a pre-clustering stage which finalizes with the creation of the Cluster Features Tree (CF) and obtaining clusters centres; and a clustering stage that employs an agglomerative clustering procedure in order to obtain a range of solutions.

In the first stage, the procedure involves computing two types of variance (Bacher et al., 2004):

a) the variance of the continuous variables:

$$-n_v \sum_{j=1}^p \frac{1}{2} \log(\hat{\sigma}_{vj}^2 + \hat{\sigma}_j^2), \tag{12}$$

where n_v represents the number of the observations of cluster v , p is the number of continuous variables, $\hat{\sigma}_j^2$ represents the estimated total variance of the variable j and $\hat{\sigma}_{vj}^2$ is the estimated cluster variance of variable j .

b) the entropy or the variance of the categorical variables:

$$-n_v \sum_{j=1}^q \sum_{l=1}^{m_j} \frac{n_{vjl}}{n_v} \log \frac{n_{vjl}}{n_v}, \tag{13}$$

where n_v represents the number of the observations of cluster v , q denotes the number of categorical variables, n_{vjl} represents the number of observations from cluster v that takes the l category.

The above mentioned variances represent the two parts of the dispersion (ξ_v) within a given cluster v . This type of dispersion is used to compute the log-likelihood distance between any two clusters (i and s):

$$d(i,s) = \xi_i + \xi_s - \xi_{\langle i,s \rangle}. \tag{14}$$

Based on this distance, at each step, any two clusters that are at a minimum distance are merged, so that at the stage with k clusters the log-likelihood function is (Bacher et al., 2004):

$$l_k = \sum_{v=1}^k \xi_v. \tag{15}$$

In the second stage of the TSCA procedure, the maximum number of clusters is estimated by means of the ratio changes of an information criterion (Schwarz, Bayesian Criterion-BIC or Akaike Information Criterion-AIC), while the optimal number of clusters is given by the ratio change in distance for k clusters.

v) We have applied *Mann-Whitney U Test* (M-WUT) with the purpose to compare the pairs of clusters obtained from TSCA, assumed to be unrelated or independent with respect to the Covid-19 indicators. This non parametric method has been chosen for its advantage: it can be applied to populations that are not normal distributed and in the presence of outliers.

3. Results

The 6 variables related to the Covid-19 pandemic and the 27 socio-economic indicators have been analysed by means of the 5-step methodology and the results demonstrate the research hypotheses. Correlation analysis highlighted the socio-economic factors that can be related to the Covid-19 pandemic outcomes. The analysis of normality pointed to some countries that qualify as outliers. Principal Component Analysis allowed us to synthesize in two factors the most relevant variables that could explain the differences between countries in terms of how they responded to the challenges posed by the health crisis. The cluster analysis highlighted the groups of countries with common features, and the Mann-Whitney U test allowed us to validate the hypotheses related to the factors that could explain the vulnerability of some countries to pandemics.

3.1. Correlation analysis between Covid-19 indicators and socio-economic indicators

In order to measure the correlation between Covid-19 indicators and socio-economic indicators, we applied a parametric method (Pearson linear correlation coefficient-PC) and a nonparametric method (Kendall rank correlation coefficient-KC). We ground our analysis on the outcomes of the last test. The results of the parametric method were used only to confirm the nonparametric method results. (Corder & Foreman, 2014). We took into consideration the fact that Pearson coefficient assumes the existence of linear relationships between the variables, and that the results could be biased in the presence of significant outliers.

The three indicators (*case fatality rate_Covid-19*, *deaths rate_Covid-19* and *cases rate_Covid 19*), computed to measure the spread of the epidemic and the intensity of mortality due to Covid-19, were correlated one by one with the three categories of socio-economic indicators: population health related indicators, development and connectivity indicators and mortality related to other causes indicators.

3.1.1. Correlation with population health-related indicators

Case fatality rate_Covid-19 was found to be positively correlated with the following indicators: life expectancy (KC = 0.25), health expenditure (KC = 0.23) and median age (KC = 0.21). The relationships were significant at the 0.05 level. *Deaths_rate_Covid-19* correlated positively with indicators such as: life expectancy (KC = 0.64), basic sanitation (KC = 0.48), health expenditure (KC = 0.32), median age (KC = 0.25) and negatively with PM2.5 (KC = -0.42). *Cases rate_Covid 19* correlated with the same indicators as *deaths rate_Covid 19*: positively with life expectancy (KC = 0.69), basic sanitation (KC = 0.51), health expenditure (KC = 0.32) median age (KC = 0.21) and negatively with PM2.5 (KC = -0.48). All the correlation coefficients were significant at the 0.01 level, except for the correlation with the median age, significant at 0.05 level.

In conclusion, 3 socio-economic indicators were found to correlate significantly with all 3 variables related to the epidemic with Covid-19: *life expectancy, health expenditure and median age*. In addition, for most of these indicators, the results of the nonparametric method were also supported by the results of the Pearson correlation coefficients (Table 1).

However, the direction of the relationship between the variables related to Covid-19 and the three socio-economic indicators mentioned above may seem surprising at first glance, except for the relation with the median age. We assumed from the beginning that age is one of the indicators that can explain the differences between countries regarding their response to the epidemic. In other words, countries with an aging population are more vulnerable to such pandemics. Though, the results of the correlation analysis show that vulnerable are also the countries where life expectancy is high and so the health expenditure. In fact, these indicators seem to point also to the countries with an aging population, affected by specific diseases.

For example, for France, the country that obtained the highest *case fatality rate_Covid-19* (15.1%) in our sample, the three selected socio-economic indicators are presented as follows: life expectancy = 82.5 (years), health expenditure = 11.4 (%) and median age = 41.6 (years). In contrast, Slovak Republic, the country experiencing the lowest case fatality rate (0.3%), is characterized by: life expectancy = 77.2 (years), health expenditure = 6.9 (%) and median age = 40.2 (years).

Table 1. Correlation analysis between population health-related indicators and Covid-19 indicators (source: authors' computation using IBM SPSS Statistics 21)

Correlation Coefficients	Case fatality rate_Covid-19		Deaths rate_Covid-19		Cases rate_Covid-19		N
	PC	KC	PC	KC	PC	KC	
<i>life_expectancy</i>	0.36*	0.25*	0.82**	0.64**	0.87**	0.69**	47
<i>age_65_plus</i>	0.35*	0.17	0.19	0.2	0.12	0.17	44
<i>median_age</i>	0.31*	0.21*	0.32*	0.25*	0.24	0.21*	47
<i>obesity</i>	0.04	0.03	0.08	0.02	0.06	0.02	44
<i>smoking</i>	-0.08	0.01	-0.1	-0.13	-0.21	-0.19	44
<i>PM 2.5</i>	0	-0.1	-0.19	-0.42**	-0.42**	-0.48**	45
<i>basic_sanitation</i>	0.22	0.17	0.26	0.48**	0.32*	0.51**	47
<i>health_expenditure</i>	0.38*	0.23*	0.18	0.32**	0.17	0.30**	44
<i>physicians</i>	0.03	-0.06	0.31*	0.07	0.28	0.12	47
<i>hospital_beds</i>	-0.13	-0.09	-0.16	-0.2	-0.24	-0.2	40

Note: PC – Pearson Correlation Coefficient; KC – Kendall's tau_b Coefficient. * Significant at the 0.05 level. ** Significant at the 0.01 level.

3.1.2. Correlation with development and connectivity indicators

According to the Kendall rank correlation coefficient, in the sample of the 48 European countries, *case fatality rate_Covid-19* is significantly correlated only with the urbanization indicator: urban population (KC = 0.38). The relationship is direct and significant at 0.01 level.

Deaths Rate_Covid-19 is positively correlated at a 0.01 level of significance both with the indicators that measure the level of development and employment structure, as well as with the indicators that show the degree of agglomeration: GDP_per capita (KC = 0.52), services (KC = 0.43), urban population (KC = 0.36), population density (KC = 0.26). The exception is the indicator of the population employment in the agriculture sector which is negatively correlated with Covid-19 indicator (KC = -0.42).

Cases rate_Covid-19 correlates significantly with all the indicators with which *death rate_Covid-19* was correlated: GDP_per capita (KC = 0.60), services employment (KC = 0.50), urban population (KC = 0.35), population density (KC = 0.20) and agriculture employment (KC = -0.45). In addition, it correlates with the indicator that measure households expenditure on hotels and restaurants (KC = 0.25). This relationship is significant at 0.05 level (Table 2).

We can conclude that all three indicators Covid-19 are positively correlated with the degree of urbanization, which draws attention to the fact that the share of the urban population can be an indicator used to signal the vulnerability of countries in times of health crisis. To this we can also add, as signal indicator, the share of the population employed in the services sector, which demonstrates the degree of connectivity of employment.

According to the correlation analysis, it seems that the level of development of the country and the degree of connectivity with other economies have influenced the degree of dissemination of the epidemic of Covid-19, while the agglomeration (urbanization) has also influenced the intensity of mortality.

Table 2. Correlation analysis between development and connectivity indicators and Covid-19 indicators (source: authors' computation using IBM SPSS Statistics 21)

Correlation Coefficients	Case fatality rate_Covid-19		Deaths rate_Covid-19		Cases rate_Covid-19		N
	PC	KC	PC	KC	PC	KC	
<i>GDP_per capita</i>	0.21	0.15	0.36*	0.52**	0.63**	0.60**	46
<i>agr</i>	-0.25	-0.12	-0.33*	-0.42**	-0.42**	-0.45**	44
<i>ind</i>	-0.03	0.05	-0.03	-0.03	-0.25	-0.11	44
<i>serv</i>	0.30*	0.12	0.44**	0.43**	0.61**	0.50**	44
<i>FDI</i>	-0.11	-0.05	-0.04	-0.02	0.12	0.02	46
<i>exp</i>	-0.2	-0.06	-0.05	0.09	0.19	0.13	45
<i>imp</i>	-0.15	-0.08	-0.03	-0.05	0.24	0.01	45
<i>international tourism_expenditure</i>	-0.21	-0.04	-0.25	-0.02	-0.17	-0.04	44
<i>international tourism_receipts</i>	-0.11	-0.08	-0.15	-0.15	-0.05	-0.13	44
<i>hotel and restaurants_households expenditure</i>	-0.06	0	0.09	0.15	0.19	0.25*	33
<i>pop_density</i>	0.2	0.28**	0.48**	0.26**	0.36*	0.20*	48
<i>urban_population</i>	0.38**	0.2	0.50**	0.36**	0.44**	0.35**	47

Note: PC – Pearson Correlation Coefficient; KC – Kendall's tau_b Coefficient. * Significant at the 0.05 level. ** Significant at the 0.01 level.

Still, there are also exceptions to this rule, such as Iceland, the second country least affected by Covid-19 mortality whose development and connectivity indicators are comparable with those of Italy, the second most affected in terms of case fatality rate: employment in services = 79.25% (Iceland) compared to 70.3% (Italy) and urban population = 93.8% (Iceland) compared to 70.1% (Italy). The first explanation of the difference between these countries with regard to the case fatality rate may be the population density but we must not neglect the fact that Iceland has applied the policy of mass testing of the population for Covid-19 infection.

3.1.3. Correlation with mortality (related to other causes) indicators

The third category of variables that we assumed will be correlated with Covid-19 indicators is the intensity of mortality from other causes. We expected that countries in which mortality is in general high to experience also high mortality caused by Covid-19.

According to the Kendall rank correlation coefficient, *case fatality rate_Covid-19* significantly and positively correlates with the following mortality indicators: death (total) (KC = 0.29) and mortality_sanitation (mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene) (KC = 0.26). A significant but negative relationship can be noticed between *case fatality rate_Covid-19* and mortality_diseases (mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70) (KC = -0.22).

Table 3. Correlation analysis between mortality indicators and Covid-19 indicators (source: authors' computation using IBM SPSS Statistics 21)

Correlation Coefficients	Case fatality rate_Covid-19		Deaths rate_Covid-19		Cases rate_Covid-19		N
	PC	KC	PC	KC	PC	KC	
<i>death_total</i>	0.45**	0.29**	0.11	0.05	-0.07	-0.1	47
<i>crude_death_rate</i>	-0.09	-0.06	-0.29*	-0.32**	-0.49**	-0.41**	48
<i>mortality_severe_diseases</i>	-0.39**	-0.22*	-0.52**	-0.63**	-0.66**	-0.69**	44
<i>mortality_households and pollution</i>	-0.34*	-0.15	-0.47**	-0.51**	-0.60**	-0.59**	44
<i>mortality_sanitation</i>	0.12	0.26*	0.34*	0.19	0.23	0.06	44

Note: PC – Pearson Correlation Coefficient; KC – Kendall's tau_b Coefficient. * Significant at the 0.05 level. ** Significant at the 0.01 level.

Deaths rate_Covid-19 turned out to be significantly but negatively correlated with the following mortality indicators: *mortality_diseases* (KC = -0.63), *mortality_households and pollution* (mortality rate attributed to household and ambient air pollution, age-standardized) (KC = -0.51), *crude_death_rate* (KC = -0.32). The relationships are significant at 0.01 level.

Cases rate_Covid-19 proved to be correlated with the same indicators as *deaths rate_Covid-19*: *mortality_diseases* (KC = -0.69), *mortality_households and pollution* (KC = -0.59), *crude_death_rate* (KC = -0.41) (Table 3).

The negative relationship between Covid-19 indicators and mortality (related to other causes) indicators is, at first glance, surprising. It would seem that countries that had high rates of mortality from different causes in the past are currently experiencing low mortality rates caused by Covid-19. An explanation for these negative relationships is that the virus has spread faster in the developed countries where it has done more victims compared to the countries where mortality from causes that may be related to the low level of development is, in general, quite high.

Another explanation can be found in the differences between countries with respect to the accuracy of the records of the deaths caused by Covid-19, related also to the numbers of the tests performed.

To shed a light on these differences, we analysed the correlation between Covid-19 indicators. As we expected *case fatality rate_Covid-19* correlates positively both with *cases rate_Covid-19* and with the *tests rate_Covid-19* and negatively with the indicator *days_first death* (number of days from the first confirmed case to the first death) (Table 4).

Basically, it seems that countries that have managed to slow the spread of the infection have increased the number of days from the first case to the first death and have reduced case fatality rates. However, there are also some exceptions, countries where the number of days is high because for a long time from the first case the population has not been tested, governments not considering there could be a pandemic.

Table 4. Correlation analysis between Covid-19 indicators (source: authors' computation using IBM SPSS Statistics 21)

Correlation Coefficients	Case fatality rate_Covid-19		Deaths rate_Covid-19		Cases rate_Covid-19		N
	PC	KC	PC	KC	PC	KC	
case fatality rate_covid-19	1	1	0.60**	0.48**	0.29*	0.22*	48
deaths rate_covid-19	0.60**	0.48**	1	1	0.83**	0.74**	48
cases rate_covid-19	0.29*	0.22*	0.83**	0.74**	1	1	48
tests rate_covid-19	-0.19	0.29**	-0.04	0.05	0.55**	-0.1	41
days_peak	0.45**	0.13	0.06	0.04	-0.12	-0.05	48
days_first death	-0.13	-0.23*	-0.16	-0.08	-0.15	-0.02	48

Note: PC – Pearson Correlation Coefficient; KC – Kendall's tau_b Coefficient. * Significant at the 0.05 level. ** Significant at the 0.01 level.

3.2. The analysis of normality and outliers identification

The analysis of the normality of the distribution (more precisely Shapiro-Wilk Test) showed first that all Covid-19 indicators fail to pass the null hypothesis of normality except for the indicator *days_first death*.

Regarding *case fatality rate_Covid-19*, as the box-plot (Figure 1) and Kernel Density graphs (Supplementary information 2) demonstrate, a large number of countries qualify as outliers: France, Italy, United Kingdom, Belgium, Netherlands, Spain and San Marino. They form the core of the group of countries with a case fatality rate caused by Covid-19 of over 8%, while the median value of the sample is only 2.5%. At the opposite end of the hierarchy there are countries such as Slovak Republic, Iceland, Latvia, Malta and Russian Federation, with the case fatality rate below 1%.

About the same remains the group of outliers in the case of the variables *deaths rate_Covid-19* and *cases rate_Covid-19*, the only difference being the very high rates recorded by San Marino. Kernel Density Estimation points to the existence of more than two groups of countries.

The distribution of countries in terms of the *tests rate_Covid-19* draws attention to Iceland, which had as a strategy in the fight against the virus large-scale population testing. The same strategy seems to have been followed by Luxembourg and Malta, a successful strategy if we take into consideration the case fatality rates, which range from 0.5% in Island to 2% in Luxembourg. It should be kept in mind, however, the fact that these are countries with smaller number of inhabitants, and the strategy of applying a large number of tests was possible.

The number of days from the first case to the epidemic peak (*days_peak*) varies between 14 days in Liechtenstein and 72 days in the United Kingdom and the Russian Federation. The median value in this case is 33 days. The group of outliers (United Kingdom, Russian Federation, Sweden, France, Germany and Spain) seems to consist mainly of countries that began large-scale population testing with some delay. Except for the Russian Federation and Germany, all the other countries from the group experienced high rates of mortality caused by Covid-19 (*deaths rates_Covid 19* range from 36.2% in Spain to 8.8% in Sweden).

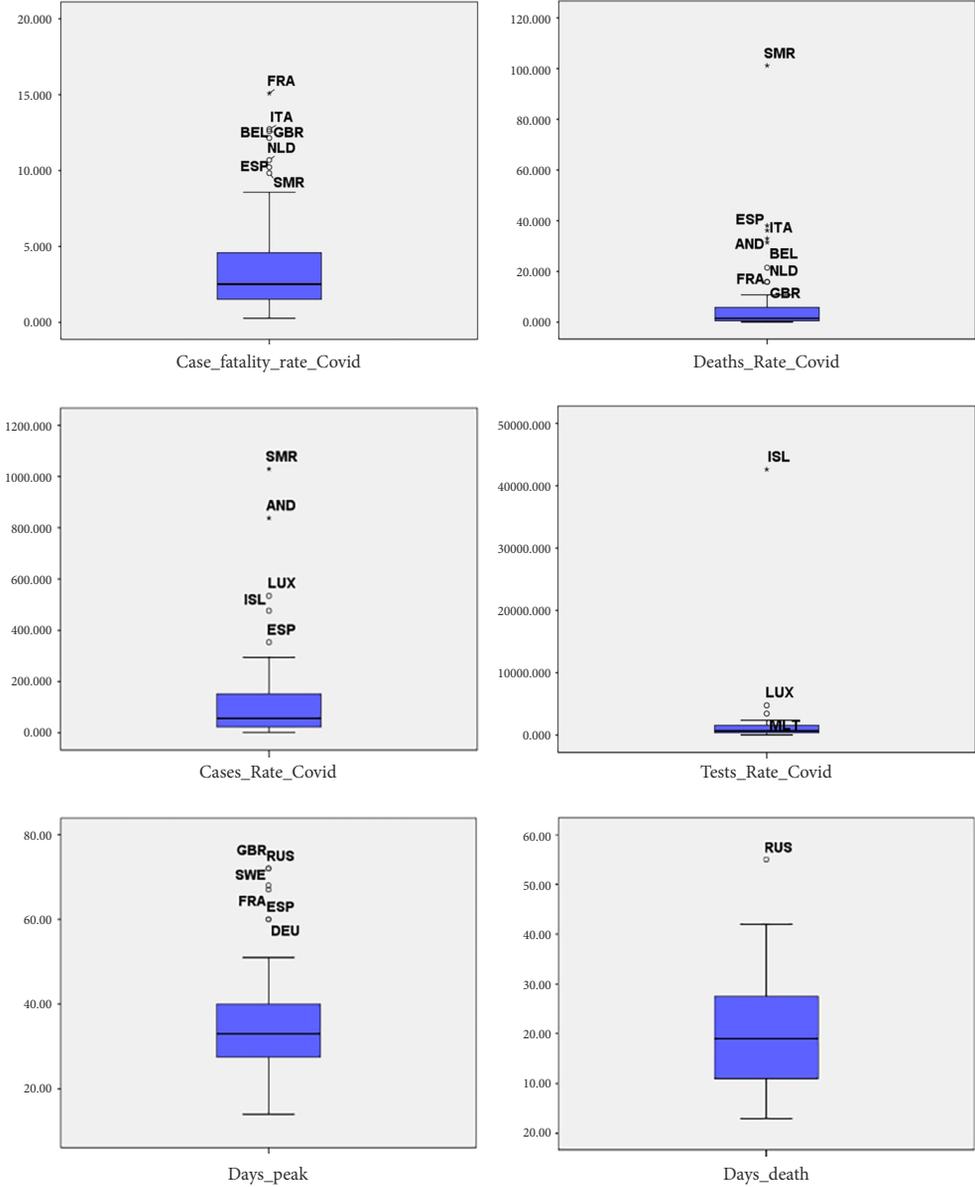


Figure 1. Box-plots for Covid-19 Indicators (source: authors' computation using IBM SPSS Statistics 21)

The variable *days_first death* is the only one of the category of Covid-19 indicators that meets the normality hypothesis. The median is much lower (19 days) compared to the previous analysed indicator, but the difference between the highest and the smallest value is larger (52 days compared to 49 days in the case of *days_peak* variable). The smallest value (3 days) was recorded by Albania, while the largest (55 days) by the Russian Federation (Figure 1).

Among the indicators that measure the mortality incidence, only *crude_death_rate* meets the hypothesis of normality (Shapiro-Wilk Statistic = 0.980, sig. = 0.568). The lowest crude_

death_rates were recorded by countries such as Andorra, Turkey and Azerbaijan (less than 6 deaths per 1,000 population) whilst the highest rates were reported by countries like Lithuania, Serbia, Ukraine, Latvia and Bulgaria (over 14 deaths per 1,000 population). It is worth noting that many of the countries strongly affected by the Covid-19 epidemic have registered in the last years a general crude_death_rate near the median of the sample of 10 deaths per 1,000 people (Italy, Belgium, France, United Kingdom and Spain).

Moreover, the same group of countries is situated below the median of the sample (15.8%) in regards to the variable *mortality_severe_diseases* (mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70) as well as regards the variable *mortality_households_and_pollution* (mortality rate attributed to household and ambient air pollution, age-standardized, per 100,000 population). In the case of the latter indicator, the Nordic countries (Sweden, Finland, Norway and Iceland) performed the best in the last years, having less than 9 deaths per 100,000 persons.

As the Kernel Density graphs show, in terms of mortality rates, European countries tend to separate into two groups. Developed countries form the group with low mortality rates, although nowadays they are most affected by Covid-19 infection (Supplementary information 2).

Only in the case of the variable *mortality_sanitation* (mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene, per 100,000 population), countries such as Germany, Belgium, France, United Kingdom, the Netherlands, Spain and Sweden are situated above the median of the sample, a situation that explains the positive correlation between *case_fatality_rate_Covid-19* and *mortality_sanitation* (Figure 2).

Among the population health-related indicators, only *health_expenditure* (% of GDP) proved to meet normality hypothesis (Shapiro-Wilk Statistic = 0.978, sig. = 0.550). Life expectancy ranges from 71.6 years in Ukraine to 85.6 years in San Marino, Kernel Density graph highlighting the existence of two groups of countries, formed on either side of the median (79 years) (Supplementary information 2). Except for a small group of countries whose *median age* is under 32 years (Kosovo, Turkey and Azerbaijan), the distribution of European countries is close to the normal one in terms of the above mentioned variable. The highest median age is registered in Italy and Germany, of over 46 years. Many of the developing countries qualify as outliers with respect to the variable *basic_sanitation* (people using at least basic sanitation services, % of population): Moldova, Romania, Bulgaria, Russian Federation, Georgia, Ireland and Latvia being at the lower end of the distribution (Figure 3).

Among the indicators related to the level of development and connectivity of the countries in the sample, *serv* (employment in services, % of total employment) and *urban_population* (% of total population) proved to have a normal distribution (Shapiro-Wilk = 0.957 for the first variable and 0.950 for the second, with p-value = 0.102 for *serv* and 0.070 for *urban_population*). The same shape of the distribution is outlined by the graphs in the case of GDP per capita, if Luxembourg is removed from the sample.

Concerning the indicator *agr* (employment in agriculture, % of total employment), several countries appear as outliers (Armenia, Azerbaijan, Moldova, Albania and Georgia), with a share of the population employed in this sector above 30%. Countries such as Malta, Belgium, Great Britain, Luxembourg and Germany have only 1% of the population employed in the primary sector.

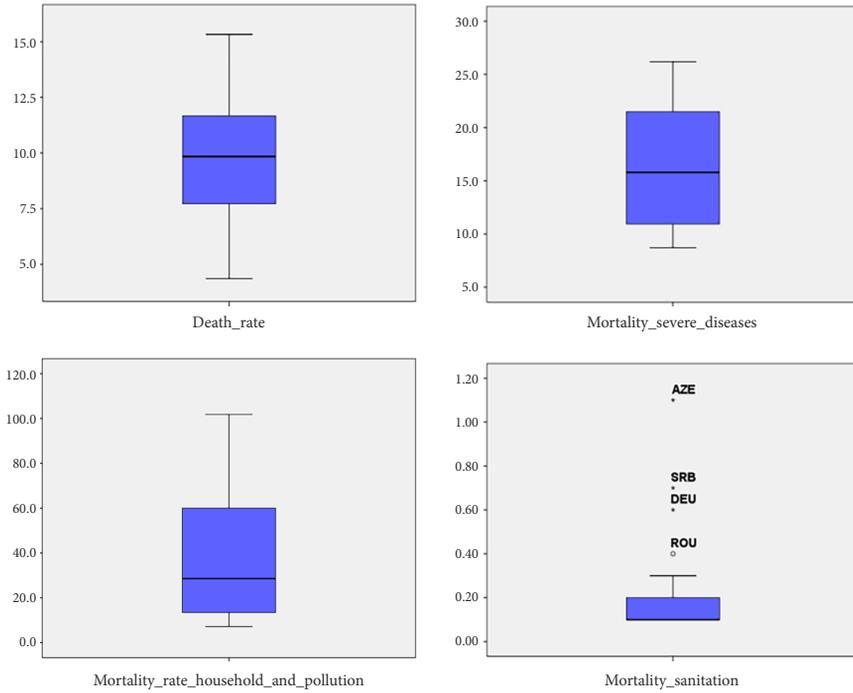


Figure 2. Box-plots for mortality indicators (source: authors' computation using IBM SPSS Statistics 21)

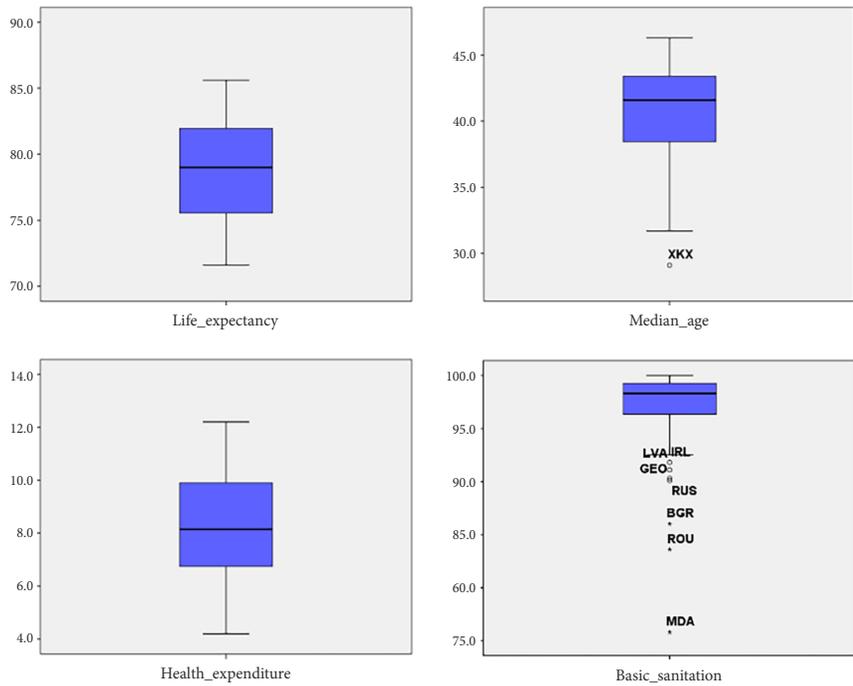


Figure 3. Box-plots for population health-related indicators (source: authors' computation using IBM SPSS Statistics 21)

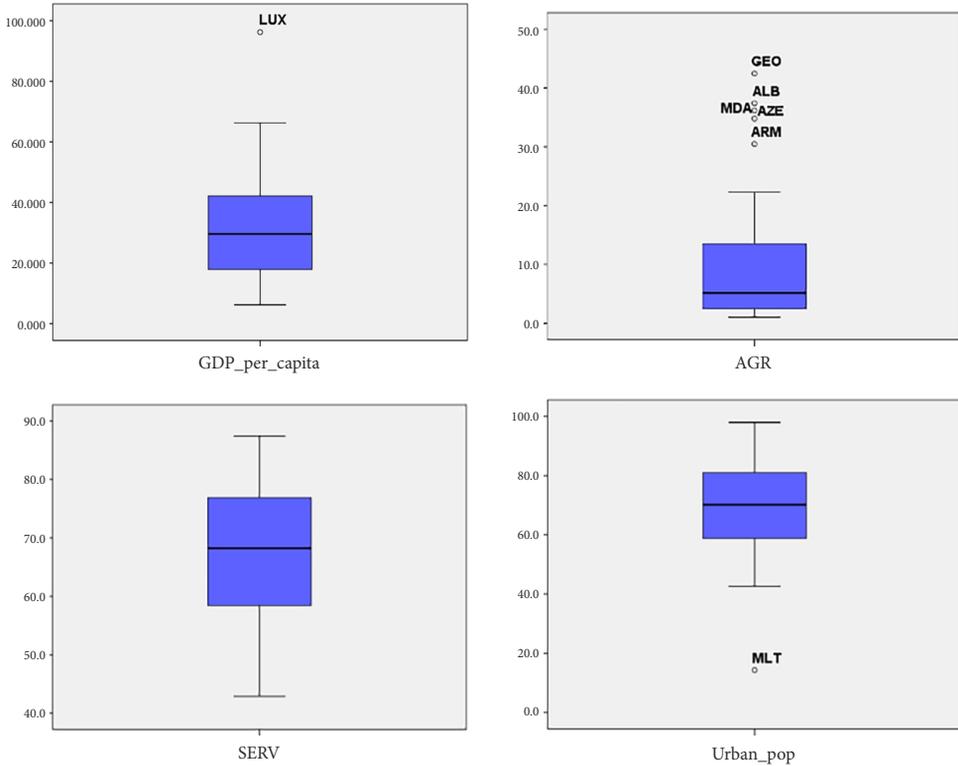


Figure 4. Box-plots for development and connectivity indicators (source: authors' computation using IBM SPSS Statistics 21)

Even for the *serv* and *urban population* variables, the graphs show large differences between the countries that are at the two tails of the distribution: Albania and Georgia with less than 45% population employed in services compared to Netherlands and Luxemburg with over 81%; Malta with 14.3% urban population compared to Belgium, San Marino, Luxemburg, Iceland, Netherlands and Lithuania with over 90% urban population (Figure 4).

3.3. Dimension reduction. Principal Component Analysis (PCA)

The first two principal components explain 74% of the variance of the 6 variables included in the PCA. The eigenvalue for the first principal component (after rotation) is 2.85. This principal component explains 47% of the variance of the initial variables. The eigenvalue for the second principal component (after rotation) is 1.58. This component explains 26% of the variance of the initial variables.

The KMO (Kaiser-Meyer-Olkin Measure of Sampling Adequacy) value, which is greater than 0.5, demonstrates that the solution with the two components is acceptable, although by including other variables in the PCA a better solution could be found.

Bartlett's Test of Sphericity result proves that the null hypothesis which states that the correlation matrix is an identity matrix can be rejected at 0.01 level of significance. More precisely the variables are suitable to be used for factor detection (Table 5).

The communalities, as an expression of the common variance, range from 0.360 for *urban population* to 0.931 for the variable *crude death rate*. More precisely, the two extracted principal components explain 36% of the variance of the variable *urban population* and over 93% of the variance of the variable *crude death rate*.

The component loadings, which measure the correlation between principal components and initial variables, show that variables such as: *serv*, *GDP per capita*, *health expenditure* and *urban population* strongly correlate with the first principal component while median age and *crude death rate* with the second principal component. *Crude death rate* negatively correlates with the first component, whereas *GDP per capita* and *serv* with the second component (Table 6).

We named the first principal component *development and connectivity* and the second principal component *population aging and mortality*.

Table 5. Total variance explained in the 2-component PCA (source: authors’ computation using IBM SPSS Statistics 21)

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.86	47.70	47.70	2.86	47.70	47.70	2.85	47.44	47.44
2	1.56	25.99	73.70	1.56	25.99	73.70	1.58	26.26	73.70
3	0.80	13.33	87.02						
4	0.47	7.83	94.86						
5	0.19	3.11	97.97						
6	0.12	2.03	100.00						
Kaiser-Meyer-Olkin Measure of Sampling Adequacy							0.602		
Bartlett’s Test of Sphericity (Approx. Chi-Square)							118.109**		

Note: Extraction Method: Principal Component Analysis. ** Significant at the 0.01 level.

Table 6. Communalities and component loadings in the 2-component PCA (source: authors’ computation using IBM SPSS Statistics 21)

Variables	Communalities		Component Matrix ^a		Rotated Component Matrix ^b	
	Initial	Extraction	1	2	1	2
<i>median_age</i>	1.000	0.878	0.302	0.887	0.398	0.848
<i>health_expenditure</i>	1.000	0.575	0.746	0.137	0.756	0.054
<i>GDP_per_capita</i>	1.000	0.843	0.907	-0.143	0.886	-0.242
<i>serv</i>	1.000	0.835	0.912	0.066	0.913	-0.035
<i>urban_population</i>	1.000	0.360	0.572	0.182	0.589	0.118
<i>crude_death_rate</i>	1.000	0.931	-0.484	0.834	-0.390	0.882

Note: Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization; a. 2 components extracted; b. Rotation converged in 3 iterations.

3.4. Cluster analysis. TwoStep Cluster Analysis (TSCA)

We applied TSCA for each of the principal component extracted from the initial variables and the 40 countries from the sample grouped in two clusters (Luxembourg was excluded, being an outlier in terms of GDP per capita, whereas Albania, Andorra, Bosnia and Herzegovina, Liechtenstein, Kosovo, San Marino, due to missing data).

The two clusters based on the first principal component, *development and connectivity*, exhibit the following features:

Cluster 1

- Includes 22 countries: Armenia, Azerbaijan, Belarus, Bulgaria, Croatia, Czech Republic, Estonia, Georgia, Hungary, Latvia, Lithuania, Malta, Moldova, North Macedonia, Poland, Romania, Russian Federation, Serbia, Slovak Republic, Slovenia, Turkey and Ukraine (Figure 5).
- Is negatively correlated with *development and connectivity* component: mean score = -0.78 , median score = -0.70 , while overall median = -0.10 .
- For all the socio economic indicators the overall median (OM) is over the group median (GM): health expenditure = 6.6% (OM = 8.2%); GDP per capita = 24079\$ (OM = 30010\$); serv = 60.2% (OM = 68.2%); urban population = 61.8% (OM = 70.7%).
- As regards the indicators related to Covid-19 pandemic, the overall median (OM) is again over the group median (GM) for the following variables: case fatality rate = 2.1% (OM = 2.5%); crude death rate = 0.6 (OM = 1.6); cases rate = 33.7 (OM = 57.7); tests rate = 505.8 (OM = 772); days_peak = 33 (OM = 36), except for the indicator days_death = 23.5 (OM = 21).

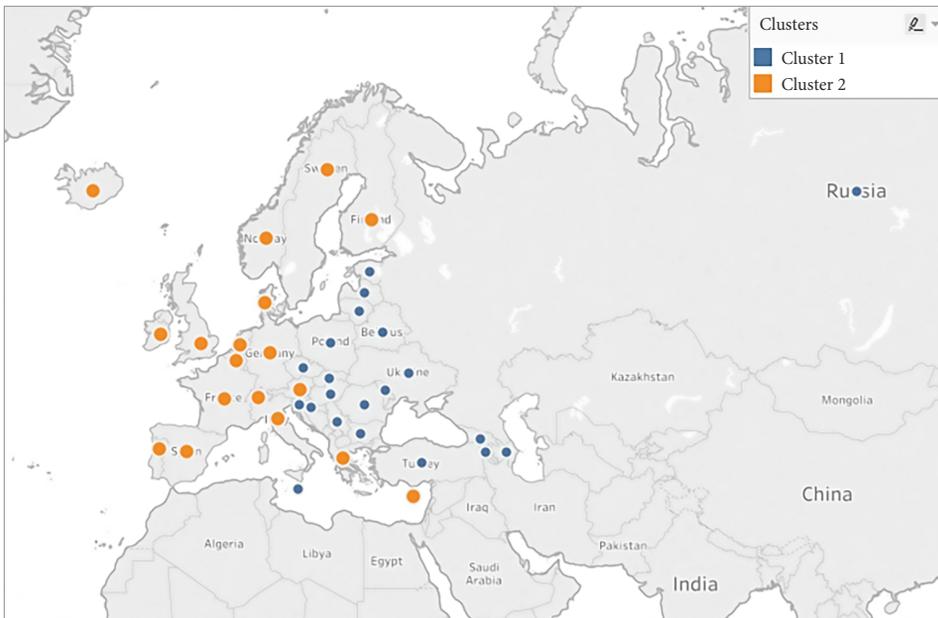


Figure 5. The two clusters based on the first principal component, *development and connectivity* (source: authors' computation using Tableau Public 10.4)

Cluster 2

- Includes 18 countries: Austria, Belgium, Cyprus, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and United Kingdom (Figure 5).
- Is positively correlated with *development and connectivity* component: mean score = 0.96, median score = 0.97, while overall median = -0.10.
- For all the socio economic indicators the overall median (OM) is well below the group median (GM): health expenditure = 10% (OM = 8.16%); GDP per capita = 44273\$ (OM = 30010\$); serv = 77% (OM = 68.2%); urban population = 80.1% (OM = 70.7%).
- With respect to the indicators related to Covid-19 pandemic, the overall median (OM) is again well below the group median (GM): case fatality rate = 4% (OM = 2.52%); crude death rate = 5.9 (OM = 1.6); cases rate = 148 (OM = 57.7); days_peak = 39.5 (OM = 36), with exception of the indicators tests rate = 1118 (OM = 772); and days_death = 17.1 (OM = 21).

TSCA also led to the formation of two clusters in the case of the component *population aging and mortality*.

Cluster 1

- Includes 8 countries: Armenia, Azerbaijan, Cyprus, Iceland, Ireland, Malta, Norway and Turkey (Figure 6).
- Is negatively correlated with *population aging and mortality* component: mean score = -1.56, median score = -1.42, while overall median = 0.10.

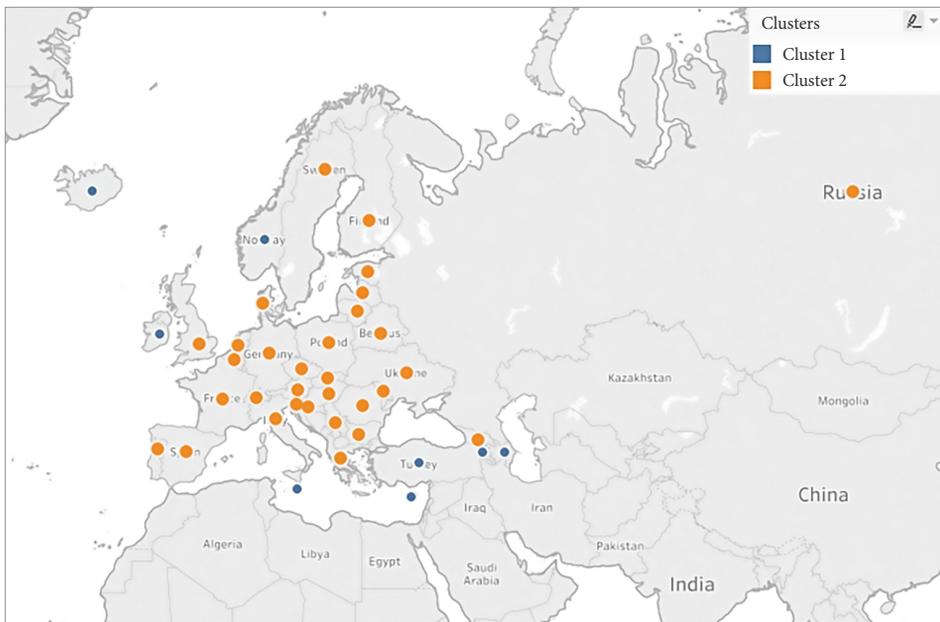


Figure 6. The two clusters in the case of the component, *population aging and mortality* (source: authors' computation using Tableau Public 10.4)

- For the social indicators the overall median (OM) is over the group median (GM): median age = 36.8 (OM = 41.7); crude death rate = 6.8 (OM = 10).
- With respect to the indicators related to Covid-19 pandemic, the overall median (OM) is again over the group median (GM) for the next variables: case fatality rate = 1.5% (OM = 2.5%); days_peak = 32 (OM = 36), days_death = 17.6 (OM = 21) except for the indicators crude death rate=1.7 (OM = 1.6); cases rate = 92 (OM = 57.7); tests rate = 1091.5 (OM = 772).

Cluster 2

- Includes 32 countries: Austria, Belarus, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Moldova, Netherlands, North Macedonia, Poland, Portugal, Romania, Russian Federation, Serbia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, United Kingdom and Ukraine (Figure 6).
- Is positively correlated with *population aging and mortality* component: mean score = 0.39, median score = 0.22, while overall median = 0.10.
- For the social indicators the overall median (OM) is below the group median (GM): median age = 42.4 (OM = 41.7); crude death rate = 10.7 (OM = 10).
- With regard to the indicators related to Covid-19 pandemic, the overall median (OM) is again below the group median (GM) for the next variables: case fatality rate = 3% (OM = 2.5%); days peak = 36.6 (OM = 36), days_death = 21.6 (OM = 21) except for the indicators crude death rate = 1.5 (OM = 1.6); cases rate = 53.6 (OM = 57.7); tests rate = 665.5 (OM = 772).

3.5. Independence testing. Mann-Whitney U Test

We tested the first two research hypotheses by means of a nonparametric method *Mann-Whitney U Test (M-WUT)*. The results of the *M-WUT*, significant at a 0.01 level proved that there are significant differences between countries from the two clusters formed in terms of *development and connectivity* regarding the *case fatality rate_Covid-19*. The effect size, or the degree of associations between the two variables, is a medium large one (ES = 0.45). Moreover, the first cluster obtained a mean rank of 15.73 which appear to be at great distance from the mean rank of the second cluster, of 26.33.

The same result, but even more conclusive, was obtained in the case of two other indicators related to the Covid-19 pandemic: *cases rate* and *deaths rate*. The results of the *M-WUT* were significant at a 0.01 level and demonstrated that there are significant differences between the two clusters as regards the above mentioned variables. In this case, the effect size, or the degree of association is large (ES = 0.75). The mean rank obtained by the countries in the first cluster represents less than half of the average rank of the second cluster.

There are significant differences between the countries from the two clusters also in terms of the *tests rate_Covid-19*. This fact leads us to the idea that this indicator can be an explanation of the differences between the two groups regarding the response to Covid-19 pandemic, but also a factor that can signalize the resilience capacity of a country (Table 7).

Table 7. The results of the Mann-Whitney U Test (source: authors’ computation using IBM SPSS Statistics 21)

Covid-19 Indicators	Mann-Whitney U	Z	ES	Mann-Whitney U	Z	ES
	development and connectivity			population aging and mortality		
<i>case fatality rate_covid-19</i>	93	-2.855**	0.45	65	-2.130*	0.34
<i>deaths rate_covid-19</i>	24	-4.730**	0.75	124	-0.135	0.02
<i>cases rate_covid-19</i>	23	-4.758**	0.75	96	-1.082	0.17
<i>tests rate_covid-19</i>	98	-2.218*	0.35	80	-1.328	0.21
<i>days_peak</i>	128.5	-1.892	0.30	91.5	-1.236	0.20
<i>days_first death</i>	176	-0.599	0.09	103.5	-0.829	0.13

Note: $ES = |z|/\sqrt{n}$ is the Effect Size. **significant at 0.01 level. * significant at 0.05 level.

The *M-WUT* results showed that there are significant differences between clusters formed based on the *population aging and mortality* factor only in terms of *case fatality rate_Covid-19*. The effect size, or the degree of associations between the two variables, is a medium one (ES = 0.34). In the case of the other variables related to Covid-19 pandemic, there were no significant differences between the two clusters.

Discussions and concluding remarks

The objective of this research was to identify socio-economic indicators that can signal countries’ vulnerability to health crises and indicators that can measure the resilience capacity. We intended to draw some lessons from the countries common experience to the Covid-19 pandemic and to find what is to be learned from the particular cases.

We studied a number of 48 European countries and investigated a set of 6 indicators that depict the Covid-19 pandemic situation in these countries as well as 27 economic and social indicators grouped into three categories: population health-related indicators, development and connectivity indicators and mortality indicators. This research was written during the mid of the first wave of Covid-19 pandemic crisis using available official data, and some of them could be reviewed. We have assumed this limitation for a higher purpose: better understanding the economic resilience.

The methodology was selected so as to ensure the verification of the research hypotheses. We assumed that certain categories of economic and social factors can signal the vulnerability of a country in the case of a pandemic, such as economic development based on connectivity. In addition, population characteristics, advanced age and associated diseases and high mortality rates may be other indicators of vulnerability. Young countries with low mortality rates are likely to have a greater resilience capacity as well as countries with performing health systems.

In a first stage, the correlation analysis was employed with the aim of identifying the factors that correlate significantly with the indicators of the Covid-19 pandemic. From the category of factors that can influence the population’s health, the most relevant in relation

to the Covid-19 pandemic are: the median age of the population, life expectancy, health expenditures and access to basic sanitation. The results obtained are very interesting: there is a very strong direct relationship between Covid-19 mortality rate and life expectancy, indicating the major effects of this pandemic in well-developed countries. This conclusion is also supported by the direct relationship between the mortality rate and the median age obtained from the analysis. Moreover, a particularly interesting aspect is the statistically significant direct relationship between case-fatality rate and health expenditure. Countries with a well-organized health system, which provides good living conditions for the elderly, are a model of development in today's world. But the analysis indicated the vulnerability of these countries to the Covid-19 pandemic, their citizens being the most affected. Old age was, in the first wave of the pandemic, the main influencing factor of Covid-19 mortality. Even high-performing health systems have failed to mitigate the serious effects of the pandemic. These results represent the first step in drawing the conclusion that one of the indicators that can signal the vulnerability of countries in the face of pandemics is the age of the population, more precisely the advanced age. The size of the health expenditures could be an indicator of the resilience capacity after such pandemics, but only to the extent that they are directed towards preventing the occurrence of diseases associated with age.

Regarding the factors that reflect the level of development and connectivity, the indicators of Covid-19 pandemic were positively correlated with GDP per capita, employment in services and the degree of urbanization. Once again, developed countries have proven to be more vulnerable. We saw that characteristics that define economic progress and efficiency (GDP per capita, development of the services sector, urbanization) are associated with a greater impact of the Covid-19 pandemic. The degree of urbanization is an indicator of vulnerability because large urban agglomerations can facilitate the spread of the virus, while the high share of employment in services indicates an increased level of connectivity which in the pandemic can be considered a disadvantage.

As expected, Covid-19 indicators were significantly correlated with mortality indicators, but the direction of the relation appeared surprising at first glance: countries with high overall mortality rates have low mortality rates due to the infection with Covid-19. But the countries where the general mortality rate is higher are less developed, consequently the health system is less efficient, the life expectancy is lower, so the median age is lower. As the elderly population was most affected by the first wave of the pandemic, countries with a younger population recorded relatively lower losses.

In the second stage of the analysis, the previously selected indicators were investigated in terms of the shape of the distribution. We identified the countries that qualify as outliers and analysed their particular situation in order to find the factors that made them vulnerable or those that determined their resilience capacity. Among the latter there are Iceland and Germany, countries with a high-performance health system. Iceland is one of the countries with the lowest mortality rates, and in the fight against the Covid-19 epidemic it has managed to apply the strategy of population large-scale testing. Germany, although vulnerable in terms of many social indicators (population age, mortality rate, urban agglomeration and rather late onset of the epidemic) has managed to keep case fatality rates at very low levels compared to France, Italy, Spain or the United Kingdom.

In order to test the research hypotheses and to verify if there are significant differences between the European countries in terms of the response to the Covid-19 pandemic, we applied Principal Component Analysis for factor identification and then TwoStep Cluster Analysis for grouping the countries. The cluster analysis based on the development and connectivity component divided the European countries into two main groups: Eastern and Western Europe. Western countries, better developed and with good connectivity, have suffered more from the pandemic compared to Eastern European countries. The Mann-Whitney U Test showed that indeed factors such as development and connectivity, respectively age and mortality significantly influence the country's ability to cope with health crises.

All methods used in this analysis led to the same general conclusions. Old age is a particularly important factor in the Covid-19 pandemic, the most affected countries being those with an aging population. Also, although developed countries usually have stronger resilience mechanisms and are able to quickly implement shock reduction policies, they have been the most vulnerable to the pandemic. We aim to analyse in a future study the medium-term impact of the pandemic and how countries have really managed to adapt to the new dynamics imposed by the pandemic, both in terms of health systems, as well as of economic recovery. Last but not least, the high degree of connectivity also meant a high level of dependence, an important lesson being related to outsourcing versus the production of basic goods and services in the country.

We deduced from these results that the economies of the future will have several characteristics that will ensure their success. First of all, "young" economies are a desideratum, so effective policies are needed to increase the birth rate, to support young people, to ensure the work-family balance, and to attract or maintain young people in the country. Also, the return to an upward trend in the economy needs a balanced and flexible business environment, based on technology and digitalization, which should include creative and innovative activities, environmentally friendly and sensitive to the health needs of the population. Digitalisation of work is already becoming a common practice, so new skills will be needed, and the education system must adapt quickly to new requirements.

Health systems need to be reorganized, rethought so that a pandemic can be better dealt with. Information technology will play an increasingly important role in the development of health systems and will allow an efficiency of the system, both from the point of view of the doctors' activity, and from the point of view of the patient's safety. In particular, the performance of the health systems could be improved through quantitative investments (number of hospitals and medical centres, state-of-the-art facilities, a sufficient number of medical staff) as well as qualitative investments: effective policies to reduce demographic aging, health education, improved sanitary conditions and access to healthcare. There is also a clear need to continue and strengthen national programs targeting modern chronic diseases as well as disease prevention programs.

Disclosure statement

Authors do not have any competing financial, professional, or personal interests from other parties.

References

- Bacher, J., Wenzig, K., & Vogler, M. (2004). *SPSS TwoStep Cluster – a first evaluation*. Social Science Open Access Repository. <https://nbn-resolving.org/urn:nbn:de:0168-ssaoar-327153>
- Bloom, E., de Wit, V., & Carangal-San Jose, M. J. (2005). *Potential economic impact of an avian flu pandemic on Asia* (ERD Policy Brief No. 42). Asian Development Bank. <http://hdl.handle.net/11540/2165>
- Boschma, R. (2015). Towards an evolutionary perspective on regional resilience. *Regional Studies*, 49(5), 733–751. <https://doi.org/10.1080/00343404.2014.959481>
- Caballero-Anthony, M. (2005). SARS in Asia: Crisis, vulnerabilities, and regional responses. *Asian Survey*, 45(3), 475–495. <https://doi.org/10.1525/as.2005.45.3.475>
- Chohan, U. W. (2020). *A post-coronavirus world: 7 Points of discussion for a new political economy* (CASS Working Papers on Economics & National Affairs No. EC015UC). <https://doi.org/10.2139/ssrn.3557738>
- Christopherson, S., Michie, J., & Tyler, P. (2010). Regional resilience: Theoretical and empirical perspectives. *Cambridge Journal of Regions, Economy and Society*, 3(1), 3–10. <https://doi.org/10.1093/cjres/rsq004>
- Cioaca, S. I., Cristache, S. E., Vuta, M., Marin, E., & Vuta, M. (2020). Assessing the impact of ICT sector on sustainable development in the European Union: An empirical analysis using panel data. *Sustainability*, 12(2), 592. <https://doi.org/10.3390/su12020592>
- Corder, G. W., & Foreman, D. I. (2014). *Nonparametric statistics: A step-by-step approach* (2nd ed.). Wiley.
- Edmond J. Safra Center for Ethics. (2020). *Roadmap to pandemic resilience*. Harvard University. https://ethics.harvard.edu/files/center-for-ethics/files/roadmaptopandemicresilience_updated_4.20.20.pdf
- European Commission. (2020). *The challenge of resilience in a globalised world*. Joint Research Centre. https://ec.europa.eu/jrc/sites/jrcsh/files/jrc-resilience-in-a-globalised-world_en.pdf
- Fernandes, N. (2020). *Economic effects of coronavirus outbreak (COVID-19) on the world economy* (IESE Business School Working Paper No. WP-1240-E). <https://doi.org/10.2139/ssrn.3557504>
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3693611/>
- Gilbert, M., Pullano, G., Pinotti, F., Valdano, E., Poletto, C., Boelle, P. Y., D'Ortenzio, E., Yazdanpanah, Y., Eholie, S. P., Altmann, M., Gutierrez, B., Kraemer, M. G., & Colizza, V. (2020). Preparedness and vulnerability of African countries against importations of Covid-19: a modelling study. *The Lancet*, 395(10227), 871–877. [https://doi.org/10.1016/S0140-6736\(20\)30411-6](https://doi.org/10.1016/S0140-6736(20)30411-6)
- Henderson, D. A., Courtney, B., Inglesby, T. V., Toner, E., & Nuzzo, J. B. (2009). Public health and medical responses to the 1957-58 influenza pandemic. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, 7(3), 265–273. <https://doi.org/10.1089/bsp.2009.0729>
- Hynes, W., Linkov, I., & Trump, B. (2020). *A systemic approach to dealing with Covid-19 and future shocks*. New Approaches to Economic Challenges (NAEC). OECD. http://www.oecd.org/naec/projects/resilience/NAEC_Resilience_and_Covid19.pdf
- Karlsson, M., Nilsson, T., & Pichler, S. (2014). The impact of the 1918 Spanish flu epidemic on economic performance in Sweden: An investigation into the consequences of an extraordinary mortality shock. *Journal of Health Economics*, 36, 1–19. <https://doi.org/10.1016/j.jhealeco.2014.03.005>
- Keogh-Brown, M. R., Smith, R. D., Edmunds, J. W., & Beutels, P. (2010) The macroeconomic impact of pandemic influenza: Estimates from models of the United Kingdom, France, Belgium and The Netherlands. *The European Journal of Health Economics*, 11, 543–554. <https://doi.org/10.1007/s10198-009-0210-1>

- Kieny, M., Evans, D. B., Schmets, G., & Kadandale, S. (2014). Health-system resilience reflections on the Ebola crisis in western Africa. *Bulletin of the World Health Organization*, 92(12), 850. <https://doi.org/10.2471/BLT.14.149278>
- Kilbourne, E. D. (2004). Influenza pandemics: Can we prepare for the unpredictable? *Viral Immunology*, 17(3), 350–357. <https://doi.org/10.1089/vim.2004.17.350>
- Lee, V. J, Wong, C. S., & Tambyah, P. A. (2008). Twentieth century pandemics in Singapore. *Annals of the Academy of Medicine*, 37(6), 470–476.
- Liu, Y., Gu, Z., Xia, S., Shi, B., Zhou, X. N., Shi, Y., & Liu, J. (2020). What are the underlying transmission patterns of Covid-19 outbreak? An age-specific social contact characterization. *EClinicalMedicine*, 22, 100354. <https://doi.org/10.1016/j.eclinm.2020.100354>
- Maffioli, E. M. (2020). How is the world responding to the Novel Coronavirus disease (COVID-19) compared with the 2014 West African Ebola epidemic? The importance of China as a player in the global economy. *The American Journal of Tropical Medicine and Hygiene*, 102(5), 924–925. <https://doi.org/10.4269/ajtmh.20-0135>
- Marinescu, C. (2020). *Corona-economy: From Crisis to rebuilding trust*. <https://en.cosmin-marinescu.ro/corona-economy-crisis-rebuilding-trust/>
- Mitra, P., Misra, S., & Sharma, P. (2020). COVID-19 pandemic in India: What lies ahead. *Indian Journal of Clinical Biochemistry*, 35, 257–259. <https://doi.org/10.1007/s12291-020-00886-6>
- Modica, M., Reggiani, A., & Nijkamp, P. (2018). *Vulnerability, resilience and exposure: Methodological aspects and an empirical applications to shocks* (SEEDS Working Paper 13/2018). <http://www.sustainability-seeds.org/papers/RePec/srt/wpaper/1318.pdf>
- Moore, M., Gelfeld, B., Okunogbe, A. T., & Paul, C. (2016). *Identifying future disease hot spots: Infectious disease vulnerability index*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR1605.html
- O'Sullivan, T., & Bourgoin, M. (2010). *Vulnerability in an influenza pandemic: Looking beyond medical risk*. Public Health Agency of Canada. https://homelesshub.ca/sites/default/files/Lit%20Review%20-%20Vulnerability%20in%20Pandemic_FINAL.pdf
- Oppenheim, B., Gallivan, M., Madhav, N., Brown, N., Serhiyenko, V., Wolfe, N., & Ayscue, P. (2019). Assessing global preparedness for the next pandemic: Development and application of an Epidemic Preparedness Index. *BMJ Global Health*, 4(1). <https://doi.org/10.1136/bmjgh-2018-001157>
- Patriarca, P. A., & Cox, N. J. (1997). Influenza pandemic preparedness plan for the United States. *The Journal of Infectious Diseases*, 176(Suppl. 1), S4–S7. <https://doi.org/10.1086/514174>
- Peng, F., Tu, L., Yang, Y., Hu, P., Wang, R., Hu, Q., Cao, F., Jiang, T., Sun, J., Xu, G., & Chang, C. (2020). Management and treatment of COVID-19: The Chinese experience. *The Canadian Journal of Cardiology*, 36, 915–930. <https://doi.org/10.1016/j.cjca.2020.04.010>
- Presanis, A. M., De Angelis, D., Hagy, A., Reed, C., Riley, S., Cooper, B. S., Finelli, L., Biedrzycki, P., & Lipsitch, M. (2009). The severity of pandemic H1N1 influenza in the United States, from April to July 2009: A Bayesian analysis. *PLoS Medicine*, 6(12), e1000207. <https://doi.org/10.1371/journal.pmed.1000207>
- Sarewitz, D., Pielke, R., & Keykhah, M. (2003). Vulnerability and risk: some thoughts from a political and policy perspective. *Risk Analysis*, 23(4), 805–810. <https://doi.org/10.1111/1539-6924.00357>
- Șerban, A. C., Aceleanu, M. I., Dospinescu, A. S., Țircă, D. M., & Novo-Corti, I. (2020). The impact of EU immigration on economic growth through the skill composition channel. *Technological and Economic Development of Economy*, 26(2), 479–503. <https://doi.org/10.3846/tede.2020.11954>
- Sirkeci, I., & Yucesahin, M. (2020). Coronavirus and migration: Analysis of human mobility and the spread of Covid-19. *Migration Letters*, 17(2), 379–398. <https://doi.org/10.33182/ml.v17i2.935>

- Smith-Bingham, R., & Hariharan, K. (2020). *This is the impact of the Coronavirus on business*. World Economic Forum. <https://www.weforum.org/agenda/2020/02/why-is-coronavirus-a-global-business-risk/>
- UNSD. (2020). *How Covid-19 is changing the world: A statistical perspective*. Committee for the Coordination of Statistical Activities. <https://unstats.un.org/unsd/ccsa/documents/covid19-report-ccsa.pdf>
- Verikios, G., Sullivan, M., Stojanovski, P., Giesecke, J., & Woo, G. (2011). *The global economic effects of pandemic influenza* (Working Papers G-224). Victoria University, Centre of Policy Studies/IMPACT Centre. <http://www.copsmodels.com/elecpr/g-224.htm>
- Vynnycky, E., & Edmunds, W. J. (2008). Analyses of the 1957 (Asian) pandemic in the United Kingdom and the impact of school closures. *Epidemiology & Infection*, *136*(2), 166–179. <https://doi.org/10.1017/S0950268807008369>
- Yoldascan, E., Kurtaran, B., Koyuncu, M., & Koyuncu, E. (2010). Modeling the economic impact of pandemic influenza: A case study in Turkey. *Journal of Medical Systems*, *34*(2), 139–145. <https://doi.org/10.1007/s10916-008-9225-x>