Discrete & Continuous Models & Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2023, 31 (3) 260–272
http://journals.rudn.ru/miph

Research article

UDC 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-3-260-272
EDN: KIJKGU

Identification of COVID-19 spread factors in Europe based on causal analysis of medical interventions and socio-economic data

Kouame A. Brou

Peoples' Friendship University of Russia (RUDN University)
6, Miklukho-Maklaya str., Moscow, 117198, Russian Federation

(received: May 30, 2023; revised: July 25, 2023; accepted: September 8, 2023)

Abstract. Since the appearance of COVID-19, a huge amount of data has been obtained to help understand how the virus evolved and spread. The analysis of such data can provide new insights which are needed to control the progress of the epidemic and provide decision-makers with the tools to take effective measures to contain the epidemic and minimize the social consequences. Analysing the impact of medical treatments and socioeconomic factors on coronavirus transmission has been given considerable attention. In this work, we apply panel autoregressive distributed lag modelling (ARDL) to European Union data to identify COVID-19 transmission factors in Europe. Our analysis showed that non-medicinal measures were successful in reducing mortality, while strict isolation virus testing policies and protection mechanisms for the elderly have had a positive effect in containing the epidemic. Results of Dumitrescu–Hurlin paired-cause tests show that a bi-directional causal relationship exists for all EU countries causal relationship between new deaths and pharmacological interventions factors and that, on the other hand, some socioeconomic factors cause new deaths when the reverse is not true.

Key words and phrases: causality analysis, COVID-19, socio-economic, Dumitrescu–Hurlin' panel

1. Introduction

In January 2020, the SARS-CoV-2 coronavirus from 2019 made its way to Europe. As a result, the European Union and the majority of European nations had documented their first case. It should be observed, nevertheless, that the infection spread unevenly. At the end of April, there were more than three million confirmed cases of the severe acute respiratory syndrome coronavirus (COVID-19) worldwide (CSS, 2020), (SARS-CoV-2). The first human instance of the coronavirus was discovered in Wuhan, China, in late 2019 despite the fact that its origins are still unknown. One way the coronavirus is spread from person to person is through respiratory droplets created when infected people cough or sneeze in front of others [1].

Air travel is one of the factors contributing to the coronavirus outbreak in Europe. Late January or early February saw the confirmation of the first instances. Human contacts after the virus's introduction to Europe helped it spread quickly. Social contact is crucial for the spread of all viruses, including COVID-19, according to research [2]. Human behavior is frequently viewed as a crucial safeguard for stopping the COVID-19 pandemic [3]. Globally, policymakers and health professionals are urging people to exercise social responsibility by limiting social interaction, adhering to stringent cleanliness and distancing guidelines, and being vaccinated. 1 Politicians are advising their constituents to weigh the social costs of their individual acts in terms of economics. In order to counteract COVID-19, official strategies heavily rely on this method of using social capital. The significance of social capital to controlling COVID-19 and preserving population health, however, is not well supported by systematic studies. According to what we know, this study is the first to rigorously analyze the dynamic link between social capital and health outcomes, as determined by COVID-19 instances and excess mortality. We systematically demonstrate that social capital has a causal and beneficial impact on pandemic-related health outcomes based on different analyses for seven European nations: Austria, Germany, Great Britain, Italy, the Netherlands, Sweden, and Switzerland. Personal hygiene habits and non-pharmaceutical interventions are the only ways to stop the spread of COVID-19 in the absence of vaccines and medications.

The development of a broad framework for the causal analysis of COVID-19 in Europe is the goal of this research. As response variables, the number of new cases and fatalities attributable to COVID-19 are used. Potential causative variables include intervention factors and measures.

2. Related works

Several studies have used various approaches and linked data from the WHO and other COVID-19 data sources to study the pandemic's spread or serve as a guide for developing measures. Using the COVID-19 government response tracker data from the University of Oxford, employed Nonlinear Additive Noise Models for Bivariate Causal Discovery to determine the causative effect of a factor or an intervention measure on the number of new cases or an intervention measure. Reference [4] used data from the pandemic that affected 31 provinces and regions in China from January 20, 2020, to February 24, 2021, and the directed acyclic graph to demonstrate the causal link between influencing factor and daily cases. Using information from the official reports of the Robert Koch Institute, [5] studied the spread of the virus in Germany and the causative influence of restriction measures. In order to estimate the total causal effects based on directed acyclic graph analysis by negative binomial regression, collected data for 401 German districts between 15 February and 8 July 2020 from publicly accessible sources in Germany (e.g., the Robert Koch Institute, Germany's National Meteorological Service, Google). The most commonly used statistical methods for analysing epidemiological factors of COVID-19 and evaluating intervention measures include correlation, regression, logistic regression and a dynamic model coupled with a linear model. Yet, if particular structures are considered, statistical methods like regression can only be regarded as instruments for causal analysis because they only allow a measure of causal dependence to be defined for these structures. On the basis of natural hypotheses, a procedure that is more effective than those now in use can be developed. Based on association analysis, this technique is known as dependency analysis. The statistical examination of the impacts of influencing factors and health interventions on the dissemination of COVID-19 has used association analysis as a reference. Yet, it is still challenging to comprehend the COVID-19 transmission pathway based on association analysis. The data were taken from the GlobalEconomy.com website used Pearson correlation analysis and multivariate linear regression to uncover economic and socio-political aspects that could fuel the coronavirus's expansion.

3. Materials and methods

3.1. Data Description

The analysis includes data for European economies from February 1st, 2020, through November 27th, 2022. Based on the statistics that are available, the era and the group of nations are chosen. The University of Oxford's COVID-19 government response was where the information came from. The Government Response Index can be created using the data in this set, which also includes a stringency index, a containment and health index, and an economic support index (see table 1).

Table 1

Definition of variables

Variables	Definition
NEW_DEATHS	News recorded deaths of COVID 19
STRINGENCY	Stringency Index
CONTAINMENT	Containment Health Index
ECONOMIC_SUP	Economic Support Index
VACCINATION	Vaccination policy
TESTING	Availability of detection
PROTECT_ELD	Care policy for the elderly population

The stringency index collects data on social segregation measures, coded from eight indicators: stay-at-home regulations, workplace closures, public event cancellations, gathering size restrictions, closures to public transportation, and travel restrictions both domestically and internationally.

Three indices that stand for public awareness efforts, testing regulations, and contact tracing make up the containment and health index. The index stands for the government's emergency health system policies, including the coronavirus testing program.

The government's income support program for citizens in times of crisis is reflected in the economic support index, which consists of two indicators:

household anticipated debt alleviation and government income assistance. Each of these three metrics is expressed as a simple sum of the values for the underlying metrics, scaled to a range between 0 and 100. These indexes are provided for comparison and shouldn't be used as a judgment on the suitability or efficacy of a nation's approach. The WHO is the source of the daily total of new cases. The time frame for the study is from January 1, 2020, to December 4, 2022, and it includes 230 different nations.

Table 2 displays a statistical breakdown of the key variables. The greatest value is 1623, the minimum value is 1918, and the average value is 42.27578, using the daily number of new deaths as an example. The number of new deaths is chosen as the explanatory variable since all efforts implemented by different governments around the world aim to prevent mortality, and reducing the number of cases will likely result in a decrease in deaths. So, the analysis will show us which measures not new instances as was noted in earlier literature really had an impact on pandemic related deaths.

Table 2

Descriptive and Summary Statistics

Variables	Mean	Standard Deviation	Minimum	Maximum
NEW_DEATHS	42.27578	104.4784	-1918.000	1623.000
STRINGENCY	43.11964	23.07557	0	96.30000
CONTAINMENT	49.82720	17.54525	0	90.00000
ECONOMIC_SUP	57.41835	34.87956	0	100.0000
VACCINATION	2.998873	2.247651	0	5.000000
TESTING	2.355943	0.799513	0	3.000000
PROTECT_ELD	1.588960	1.006744	0	3.000000

3.2. Methodology

In our empirical research, we examined how health interventions and socioeconomic observational data contributed to the global spread of COVID-19. Using this method, we may assess how health measures have affected the spread of COVID-19. In order to accomplish our goal, we used in this study a linear function that incorporates socioeconomic observational data and health treatments as an extra variable to control factors that are equivalent to COVID-19. As suggested by Pesaran and Shin, the equation is calculated using a time series autoregressive distributed lag model (ARDL). The advantage of the ARDL framework is that it can differentiate between short- and long-term impacts, which enhances earlier material. We may also predict a consistent short-term cross-sectional influence (short term coefficient of nations) due to our extensive sample size. Due to its distinction between short- and long-term impacts, the ARDL methodology aids in addressing the shortcomings of earlier work.

Using both time and cross-sectional dimensions increases the overall number of data and their variability in our panel estimation. A panel estimation also reduces the noise that results from a single time-series estimation, leading to more trustworthy inference.

3.2.1. Panel unit root tests

The determination of the order of integration of variables serves as the foundation for estimating any econometric model. It is required to verify that the variables in the regression are either integrated of order zero $I(0)$ or at most integrated at order one I during the ARDL model estimate procedure $I(1)$. Reference [6] was used to check the integration of the variables in the proper sequence. The ADF regression for panel data serves as the foundation for these two tests and is described as follows in (1):

$$\Delta y_{i,t} = \omega_i y_{i,t-j} + \sum_{j=1}^{p} \phi_j \Delta y_{i,t-j} + \epsilon_{i,t}, \text{ where } \omega_i = \rho_i - 1. \tag{1}$$

Both tests evaluate the zero-unit root H_0: $\omega_i = 0$ $(\rho = 1)$ with respect to the stationarity alternative H_1: $\omega_i < 0$ $(\rho_i < 1)$. The LLC test assumes that the tested parameters are the same in all panels, i.e., $\rho_i = \rho$ for all countries in the panel. The IPS test, which averages the ADF statistic and enables the parameters to vary across panels, is less constrictive than the LLC test. However, because they do not take into consideration the cross-section dependency issue that could arise as a result of macroeconomic linkages, unexplained residual independence, and unobserved common factors, both the IPS and LLC tests are regarded as first-generation unit root tests. In order to determine whether the variables in the model for this study have any cross-sectional dependence, second-generation unit root tests are run. Then, the Pesaran [7] proposed cross-section dependence (CD) test is conducted. When N is more than T, the CD test can be used to determine whether there is any cross-sectional dependency among the variables. The pair correction coefficients of OLS residual regressions are averaged to form the basis of the CD test. After the CD test has confirmed whether cross-sectional dependence exists, the cross-sectional Augmented Dickey-Fuller (CADF) test is carried out by Pesaran [8]. In order to test the null hypothesis of cross-sectional dependency among a panel of nations, the CADF test considers cross-section dependence among the variables. This is done to verify that the variables are still either $I(0)$ or I even if there is cross-sectional dependency among the group of countries $I(1)$.

3.2.2. Panel cointegration tests

After the order of integration is established, the next step in the study is to test for evidence of long-run cointegration between NEW DEATHS and the independent variables using the panel cointegration tests from Pedroni, Kao and Westlund may be used for samples smaller than 100 in number. Based on the panel-data model for an $I(1)$ dependent variable y, the Pedroni and Kao tests compare the cointegration alternative to the null hypothesis of no cointegration (see (2)):

$$y_{i,t} = x'_{it}\beta_i + z'_{it}\tau_i + \epsilon_{i,t}, \tag{2}$$

where both tests demand that the covariates not be integrated among themselves for each panel I the variables in $x(i,t)$ are an $I(1)$ series. The Kao test constrains $\beta_i = \beta$ by assuming that all of the nations in the panel share a common cointegration vector. There are some distinctions between the two tests even though they both use the identical null and alternative hypotheses. In reality, the Pedroni test differs from the Kao test in that it accepts panel-specific cointegrating vectors.

3.2.3. Autoregressive distributed Lag model

The ARDL model is estimated via unit root and cointegration tests. The ARDL model can be employed with confidence for short sample periods and distinguishes between short- and long-run coefficients. In fact, [8] shows that the long-run parameters are super-consistent even with a small sample size, whereas the short-run values are $\sqrt{T}$ consistent. A panel ARDL (p, q_1, q_2, q_3, q_4, q_5) Equation is used to express the connection as a result, where p stands for the lags of the dependent variable and q for the lags of the independent variables. In (3), we can see a representation of the panel ARDL equation:

$$\Delta\mathrm{NEW\_DEATHS}_{i,t} = \\ = \alpha_0 + \sum_{j=1}^{p} \alpha_{i,j}^{1} \Delta\mathrm{NEW\_DEATHS}_{t-i} + \sum_{k=1}^{6} \sum_{j=0}^{qk} \alpha_{i,j}^{k+1} X_{t-j}^{k} + \epsilon_{i,t}, \quad (3)$$

where $i = 1, 2, 3, \dots, N$ and $t = 1, 2, 3, \dots, T$, α_i represents the fixed effects, X^k $\alpha_{i,j}^k$, $k = 1, 2, \dots, 9$ are the lagged coefficients of the independent variables (Stringency Index, Containment Health Index, Economic Support, Vaccination policy, Testing policy, Protection to elderly) and the regressors and $\epsilon_{i,t}$ is the error term which is assumed to be white noise and varies across countries and time. In a panel error correction (ECM) representation equation (4) is formulated as follows:

$$\Delta\mathrm{NEW\_DEATHS}_{i,t} = \alpha_i + \sum_{j=1}^{p} \alpha_{i,j}^{1} \Delta\mathrm{NEW\_DEATHS}_{t-i} + \\ + \sum_{k=1}^{6} \sum_{j=0}^{qk} \alpha_{i,j}^{k+1} X_{t-j}^{k} + \sum_{k=1}^{6} \beta_k X_{t-i}^{k} + \epsilon_{i,t}, \quad (4)$$

where Δ is the first difference of variables. Also, α_1–α_7 are the short-run coefficients. While β_1–β_7 are the long-run coefficients of stringency index, containment health index, economic support index, vaccination policy, testing policy and protection of elderly respectively. In order to estimate the short-term equation, Hendry's [9] suggestion that after establishment of long-run relationship between the dependent and independent variables, the panel error correction Model (ECM) model is expressed in equation (5) as follows:

$$\Delta\text{NEW\_DEATHS}_{i,t} = \alpha_0 + \sum_{j=1}^{p} \alpha_{i,j}^{1} \Delta\text{NEW\_DEATHS}_{t-i} + \\ + \sum_{k=1}^{6} \sum_{j=0}^{qk} \alpha_{i,j}^{k+1} X_{t-j}^{k} + \Theta\text{ECM}_t - 1 + \epsilon_t, \quad (5)$$

where Θ is the ECM coefficient, which gauges the rate at which the economy adjusts each year in the direction of long-run equilibrium. The Akaike's lag selection criteria are used to establish the ECM model's ideal lag length. All the nations in the sample are considered while estimating the panel ECM.

This offers a broad overview as well as a platform of the connections between health interventions, socioeconomic observational data, and news coronavirus mortality throughout European member states. The COVID-19 death, however, is dependent on a number of factors, including the stringency index, containment health index, economic support index, vaccination policy, testing policy, and the protection of elderly people, as emphasized in the research study. Using the pooled mean group (PMG) method, the panel ARDL regression is estimated. Reference [10] describes an estimation method that combines pooling and averaging of coefficients. The intercepts, short-run coefficients, and error variances can vary freely between groups using this panel approach. The likelihood-based PMG estimator, meanwhile, imposes the restriction that the long-run coefficients be constant across groups. When homogeneity restriction is in fact true, this results in consistent estimates. The PGM estimator is also less susceptible to outliers in situations where the cross-sectional (N) is very small, as it is in our study, and it may simultaneously fix the serial autocorrelation issue. Furthermore, by selecting the proper lag structure for both the dependent and independent variables, this likelihood-based estimation resolves the issue with endogenous regressors.

3.2.4. Panel causality test

Testing for bidirectional causality between the public announcement of COVID-19's death and health treatments and socioeconomic observational data is the last step in our empirical research. Reference [11] creates a technique for analyzing the causal link between time series in a major study. The Granger representation theorem shows that there must be at least a unidirectional causality between two time series if they are cointegrated. By extending this methodology, Dumitrescu and Hurlin make it possible to identify causality in panel data. To ascertain if there is unidirectional or bidirectional causation between the two variables, the Dumitrescu and Hurlin causality test is used [12]. This two-way Granger test is used to look into the direction of causality (see equations (6), (7)):

$$\Delta\text{NEW\_DEATHS}_{i,t} = \\ = \alpha_i + \sum_{i=1}^{p} \delta_{i,k} \Delta\text{NEW\_DEATHS}_{t-k} + \sum_{k=1}^{k} \pi_{i,k} \Delta X_{t-k} + \epsilon_t, \quad (6)$$

$$\Delta X_{i,t} = \alpha_i + \sum_{i=1}^{p} \delta_{i,k} \Delta X_{t-k} + \sum_{i=1}^{p} \pi_{i,k} \Delta \text{NEW\_DEATHS}_{t-k} + \epsilon_t \tag{7}$$

with $i = 1, \dots, N$ and $t = 1, \dots, T$, where $X_{1,t}$ are the observations of independent variables used previously for country i in period t. In essence, equations (4) and (5) examine the significance of X's impacts on the present values of *confirmed cases* and X's effects on the present values of *confirmed cases*, respectively. Hence, the alternative is: $H_0 : \pi_{i,1} = \dots = \pi_{i,k} = 0$ $\forall i = 1, \dots, N$ which is similar to the fact that there is no proof of causality for any of the panel's countries. The possibility of causality for each of the panel's countries, but not necessarily for all of them, is another crucial premise of this test.

4. Results

4.1. Panel unit root and cointegration tests

As it's crucial to make sure that the order of integration is either zero or one for ARDL modeling, the empirical analysis should begin with the execution of the unit root test. To look for signs of stationarity, the Levin Lin Chu (LLC) first-generation unit root tests are used. Overall, the findings show that the panel's order of integration for the variables included in the analysis is $I(0)$ or $I(1)$, allowing for their use in the estimation of an ARDL model. The second stage of the study is to test for cointegration between the dependent variable and the six regressors given the strong support of Integration order in all the variables throughout our panel. The possibility that there is no cointegration in the panel is investigated using the Pedroni and Kao residual-based cointegration tests. The null hypothesis that there is no cointegration in the three panels is substantially rejected by cointegration tests. Consequently, for all three panels, there is proof of a long-term link between the dependent and explanatory factors. This implies that findings from an estimation of the Error Correction Model (ECM) will be trustworthy in both the short- and long-term.

4.2. Panel ARDL estimation

The next step is to estimate the panel ARDL regression as indicated in the ECM equation using a Pooled Mean Group (PMG) estimation. This is done after checking that the five variables are not integrated of an order equal or larger than $I(2)$ and that the series are co-integrated. Based on the AIC lag selection criterion, the appropriate lag duration is chosen.

Table 3 presents the empirical results on COVID-19 new deaths and intervention variables for the panel of 27 EU member states and for the full sample period, February 1st, 2020, to November 27, 2022.

The next step is to estimate the panel ARDL regression using a Pooled Mean Group (PMG) estimation as stated in the ECM equation. This is done after making sure the series are co-integrated and that none of the five variables are integrated to an order equal to or greater than $I(2)$. The suitable lag time is selected using the AIC lag selection criterion.

Table 3

Panel Error Correction Model estimation (Long-Run Coefficients)

Variables	Pooled Mean Estimator	
	Coefficient	Standard Error
STRINGENCY	-0.098759***	0.023735
CONTAINMENT	0.184387***	0.034941
ECONOMIC_SUP	-0.004844	0.004357
VACCINATION	-0.116884	0.075845
TESTING	-0.678345***	0.231617
PROTECTION	-0.712513***	0.200808

Note: *,**,*** indicates statistical significance at the 10%, 5%, and 1% level.

The empirical results on the relationship between public debt and economic growth are presented in the table 3 for the panel of 27 EU member states and for the entire sample period, from February 1st, 2020, to November 27th, 2022, subject to other explanatory variables. In other words, the greater the measure, the stronger the control over the spread of the virus will be since variables are highly negative. The reappearance of new deaths is not much impacted by economic assistance or immunization policies. Only containment can be thought to have a 10% chance of having a major long-term impact on news death. The responses to COVID-19 have a long-term impact on lowering the number of new diseases brought on by the pandemic.

4.3. Causality

The few empirical studies that have examined the relationship between COVID-19 new deaths, healthcare interventions, and socioeconomic observational data have produced conflicting findings. In actuality, the outcomes differ depending on the nations and epochs studied in these studies. For this reason, a panel Granger causality test is carried out in the analysis's concluding section. The Granger causality test requires that the two-time series have a long-run association, or be cointegrated, in order for it to be valid. It was established in earlier phases of the analysis that there is a long-term association between new COVID-19 fatalities and health treatments and socioeconomic observational data across all panels through panel cointegration tests. This demonstrates that the relationship between COVID-19 death and other variables must at least have a unidirectional cause (see the table 4).

The paired Dumitrescu and Hurlin Panel causality test [12] is used to determine the direction of causality. The test compares a possible alternative demonstrating causality for at least one cross-sectional unit of the panel with the null hypothesis that there is no homogenous Granger causality.

Table 5 displays the outcomes of the pairwise Dumitrescu–Hurlin panel causality tests.

Table 4

Dumitrescu and Hurlin panel causality test

Null hypothesis	W-Stat	Zbar-Stat	p-value
STRINGENCY does not Granger Cause NEW_DEATHS	18.8074	43.4847	0.0000
NEW_DEATHS does not Granger Cause STRINGENCY	1.81496	-0.48906	0.6248
CONTAINMENT does not Granger Cause NEW_DEATHS	19.0667	44.1558	0.0000
NEW_DEATHS does not Granger Cause CONTAINMENT	2.33235	0.84984	0.3954
ECONOMIC_SUP does not Granger Cause NEW_DEATHS	8.10025	15.7761	0.0000
NEW_DEATHS does not Granger Cause ECONOMIC_SUP	2.12838	0.32200	0.7475
VACCINATION does not Granger Cause NEW_DEATHS	8.48059	16.7605	0.0000
NEW_DEATHS does not Granger Cause VACCINATION	4.70284	6.98432	3.E-12
TESTING does not Granger Cause NEW_DEATHS	7.71457	14.7781	0.0000
NEW_DEATHS does not Granger Cause TESTING	1.25976	-1.92584	0.0541
PROTECT_ELD does not homogeneously cause NEW_DEATHS	7.55908	14.3758	0.0000
NEW_DEATHS does not homogeneously cause PROTECT_ELD	1.44830	-1.43794	0.1505

The results reveal that for the full group of countries, there is a bidirectional causality between new deaths and vaccination policy at a 95% confidence level. At 90% confidence level we can also consider that there is a bidirectional causality between new deaths of COVID-19 and testing policy. We can also notice that stringency, containment, economic support and help to elderly people cause new death when the contrary is not true. The sense of the causality is given in ARDL model result (coefficients of stringency and economic support to elderly are negatives meaning that more the variable increase and less will be the number of recorded cases of deaths due to COVID-19). In the other hand, more containment should lead to more deaths.

Table 5

Panel Error Correction Model estimation (Short-Run Coefficients)

Variables	Pooled Mean Estimator	
	Coefficient	Standard Error
ECT(-1)	-0.204139***	0.051513
D(NEW_DEATHS(-1))	-0.446762***	0.050073
D(NEW_DEATHS(-2))	-0.366532***	0.031000
D(NEW_DEATHS(-3))	-0.171823***	0.015298
D(CONTAINMENT)	-2.888073*	1.621810

5. Conclusion

In order to analyze the impact of health and socioeconomic interventions, we used data on European Union countries from Oxford University and WHO. We also addressed the challenges of identifying causal risk factors and evaluating the causal effects of risk factors and intervention measures on COVID-19. Overall, the pandemic preventive strategies have been successful in lowering the number of new fatalities, according to the study's findings. The Panel Autoregressive Distributed Lag (ARDL) modeling approach provided us with a way to give policy-makers some means to adopt the best containment measures in order to stop the spread and maximize the societal impact. Containment measures are the sole component that has an impact immediately. The pairwise Dumitrescu–Hurlin panel causality tests, on the other hand, show that there is a bidirectional causality between new deaths and pharmaceutical intervention factors for the entire group of countries, and that, conversely, socioeconomic intervention factors cause new deaths when the converse is not true.

References

[1] S. L. Priyadarsini and M. Suresh, "Factors influencing the epidemiological characteristics of pandemic COVID 19: a TISM approach," *International Journal of Healthcare Management*, vol. 13, no. 2, pp. 89–98, 2020. DOI: 10.1080/20479700.2020.1755804.

[2] A. Farseev, Y.-Y. Chu-Farseeva, Q. Yang, and D. B. Loo, "Understanding economic and health factors impacting the spread of covid-19 disease," *medRxiv*, 2020. DOI: 10.1101/2020.04.10.20058222. eprint: https://www.medrxiv.org/content/early/2020/06/09/2020.04.10.20058222.full.pdf.

[3] K. Tantrakarnapa, B. Bhopdhornangkul, and K. Nakhaapakorn, "Influencing factors of COVID-19 spreading: a case study of Thailand," *Journal of Public Health*, vol. 30, pp. 621–627, 2020. DOI: 10.1007/s10389-020-01329-5.

[4] W.-X. Tang, H. Li, M. Hai, and Y. Zhang, "Causal Analysis of Impact Factors of COVID-19 in China," *Procedia Computer Science*, vol. 199, no. 10229, pp. 1483–1489, 2022. DOI: 10.1016/j.procs.2022.01.189.

[5] R. Chaudhry, G. Dranitsaris, T. Mubashir, J. Bartoszko, and S. Riazi, "A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes," *eClinical Medicine*, vol. 25, p. 100 464, 2020. DOI: 10.1016/j.eclinm.2020.100464.

[6] A. Levin, C.-F. Lin, and C.-S. J. Chu, "Unit root tests in panel data: asymptotic and finite-sample properties.," *Journal of Econometrics*, vol. 108 (1), pp. 1–24, 2002. DOI: 10.1016/S0304-4076(01)00098-7.

[7] M. H. Pesaran, T. Schuermann, and S. M. Weiner, "Modeling regional interdependencies using a global error-correcting macroeconometric model," *Journal of Business & Economic Statistics*, vol. 22, pp. 129–162, 2004. DOI: 10.1198/073500104000000019.

[8] M. Pesaran, Y. Shin, and R. J. Smith, "Bound Testing Approaches to the Analysis of Level Relationship.," *Journal of Applied Econometrics. Special Issue: In Memory of John Denis Sargan 1924–1996: Studies in Empirical Macroeconometrics*, vol. 16, no. 3, pp. 289–326, 2001. DOI: 10.1002/JAE.616.

[9] M. P. Clements and D. F. Hendry, "Forecasting in cointegrated systems," *Applied Econometrics*, vol. 10, pp. 127–146, 1995. DOI: 10.1002/jae.3950100204.

[10] S. Johansen, "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control*, vol. 12, no. 2–3, pp. 231–254, 1988. DOI: 10.1016/0165-1889(88)90041-3.

[11] R. Engle and C. Granger, " Cointegration and Error Correction: Representation, Estimation and Testing.," *Econometrica*, vol. 55, no. 2, pp. 251–276, 1987. DOI: 10.2307/1913236.

[12] E.-I. Dumitrescu and C. Hurlin, "Testing for Granger non-causality in heterogeneous panels," *Economic Modelling*, vol. 29, pp. 1450–1460, 2012. DOI: 10.1016/j.econmod.2012.02.014.

For citation:

K. A. Brou, Identification of COVID-19 spread factors in Europe based on causal analysis of medical interventions and socio-economic data, Discrete and Continuous Models and Applied Computational Science 31 (3) (2023) 260–272. DOI: 10.22363/2658-4670-2023-31-3-260-272.

Information about the authors:

Brou, Kouame A. — PhD student of Information Technology Department of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: broureino@gmail.com, ORCID: https://orcid.org/0000-0003-1996-577X)

УДК 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-3-260-272
EDN: KIJKGU

Выявление факторов распространения COVID-19 в Европе на основе причинно-следственного анализа медицинских вмешательств и социально-экономических данных

К. А. Бру

Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия

Аннотация. С момента появления COVID-19 было получено огромное количество данных, помогающих понять, как развивался и распространялся вирус. Анализ таких данных помогает получить новые знания, необходимые для контроля за развитием эпидемии и предоставить лицам, принимающим решения, инструменты для принятия эффективных мер по сдерживанию эпидемии и минимизации социальных последствий. Анализу влияния медицинских методов лечения и социально-экономических факторов на передачу коронавируса было уделено много внимания. В этой работе мы применяем панельное авторегрессионное моделирование с распределённым запаздыванием (ARDL) к данным Европейского союза для выявления факторов распространения COVID-19 в Европе. Наш анализ показал, что немедикаментозные меры были успешными в снижении смертности, а строгость изоляции, политика тестирования на вирус и механизмы защиты пожилых людей оказывают положительное влияние на сдерживание эпидемии. Результаты панельных тестов попарной причинно-следственной связи Думитреску–Херлина показывают, что для всех стран Евросоюза существует двунаправленная причинно-следственная связь между новыми смертями и факторами фармакологического вмешательства и что, с другой стороны, некоторые социально-экономические факторы вызывают новые смерти, когда обратное неверно.

Ключевые слова: анализ причинно-следственных связей, COVID-19, социально-экономические факторы, группа Думитреску–Херлина