



The Connectedness of the Coronavirus Disease Pandemic in the World: A Study Based on Complex Network Analysis

Sha Zhu¹, Meng Kou², Fujun Lai^{3*}, Qingxiang Feng^{4*} and Guorong Du⁵

¹School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming, China, ²School of Aerospace Engineering, Beijing Institute of Astronautical System Engineering, Beijing, China, ³School of Finance, Yunnan University of Finance and Economics, Kunming, China, ⁴Institute of Guangdong Hong Kong and Macao Development Studies, Sun Yat-sen University, Guangzhou, China, ⁵Chinese Academy of Finance and Development, Central University of Finance and Economics, Beijing, China

OPEN ACCESS

Edited by:

Hui-Jia Li,
Beijing University of Posts and
Telecommunications (BUPT), China

Reviewed by:

Ju Xiang,
Changsha Medical University, China
Chengyi Xia,
Tianjin University of Technology, China

*Correspondence:

Fujun Lai
fjlj999@163.com
Qingxiang Feng
fengqx8@mail.sysu.edu.cn

Specialty section:

This article was submitted to
Social Physics,
a section of the journal
Frontiers in Physics

Received: 02 September 2020

Accepted: 07 December 2020

Published: 25 January 2021

Citation:

Zhu S, Kou M, Lai F, Feng Q and Du G
(2021) The Connectedness of the
Coronavirus Disease Pandemic in the
World: A Study Based on Complex
Network Analysis.
Front. Phys. 8:602075.
doi: 10.3389/fphy.2020.602075

Since the coronavirus disease (COVID-19) pandemic started at the beginning of 2020, it has seriously affected various countries' economic and social development and accelerated the economic recession worldwide. Therefore, the connectedness of the global COVID-19 network across countries is studied in this article. Based on COVID-19 correlations in 122 countries, we construct a complex network of COVID-19 from January 19, 2020, to August 15, 2020. We then deconstruct the overall global network connectedness and analyze the connectedness characteristics. Moreover, we empirically investigate the network connectedness influencing factors by using various countries' macroeconomic and social data. We find that the global COVID-19 pandemic network has some prominent complex network properties, such as low path length, high clustering, and good community structure. Furthermore, population density, economic size, trade, government spending, and quality of medical treatment are significant macrofactors affecting COVID-19 connectedness in different countries.

Keywords: COVID-19, pandemic, connectedness, complex network, economic and social factors

INTRODUCTION

The COVID-19 global pandemic has changed the world remarkably. In just a few months, from limited local transmission in several countries, COVID-19 evolved into a multicountry spread and raged in more than 100 countries and regions across five continents. The COVID-19 pandemic has disrupted and changed the lifestyle of millions of people around the world and has had a profound impact on international relations and the economy. At present, increasing attention is being paid to the prevention and control of the COVID-19 pandemic. The pandemic is well under control in some countries, such as China and Germany, but most countries, including the United States, still do not have effective ways to control COVID-19. The fight against the COVID-19 pandemic has become a top priority for governments in many countries and regions, and the public health sector has paid unprecedented attention to protecting people's lives.

The rapid spread of the COVID-19 pandemic is more than just a medical problem. It is also linked to an infection network among countries worldwide. The complex network science of infectious diseases can reveal necessary information about disease transmission behavior, as many studies have shown. First, the research has revealed the typical characteristics of disease transmission behavior, such as the critical transmission threshold. The virus can spread in the network only when the transmission probability is more significant than a specific value called the propagation threshold of

the network (Moore and Newman [1]). Moreover, nonrandom networks, such as small-world and community structure networks, are critical to disease transmission. Many real networks have small-world characteristics in that they tend to have small average distances and large clustering coefficients. In such a network, diseases more easily infect others. In addition, scholars [2] have found that a complex network with a community structure can promote or effectively inhibit the spread of diseases.

Furthermore, the key nodes in a network have a profound impact on disease transmission. Outbreaks of infectious diseases often spread rapidly from a small number of nodes to the entire network. Scientists propose indicators such as page rank, node degree, and *k*-shell to evaluate the network science characteristic [3]. In disease research, the information of these key nodes can reveal the path of disease transmission. Finally, network science information can be formulated to control the spread of diseases and reduce harmful transmission processes. In recent research, Wang et al. [4] use a two-layer network model method to find that people who received positive prevention information had a lower risk of disease infection than those who received negative prevention information.

Therefore, this study establishes a complex network of COVID-19 infections worldwide and investigates the influencing factors of the pandemic's connectedness across countries. We integrate three methods, i.e., complex networks, visualization principles, and regression models, to examine the global infection path of the COVID-19 pandemic and explore possible causes for the rapid COVID-19 spread. First, we calculate the correlation of COVID-19 infection growth rates between each pair of countries and obtain a global correlation matrix for COVID-19 infections using the sample from January 19, 2020, to August 15, 2020. Second, we establish a global infection complex network using a global correlation matrix to outline global COVID-19 pandemic relations and study the global infection complex network properties. Third, based on the properties of our complex network, we further investigate how economic and social factors (such as economic size, population density, international trade, government expenditure, inflation, and medical level) affect the network connectedness of the global COVID-19 infection complex network, shedding light on the prevention and control of the COVID-19 pandemic for policymakers.

The main conclusions are as follows. First, the complex network of the global COVID-19 pandemic has some prominent complex network properties, such as small-world network and community structure characteristics. In our network, there are a total of 958 edges in 122 countries around the world. The average path length between any two countries is approximately 3, the clustering coefficient is 0.58, and the network can be divided into several modularity classes. Second, in terms of intercountry connectedness, we further examine important topological characteristics in our network, e.g., degree, triangles, centrality, modularity class, hub, and authority. We use visual figures to demonstrate these features. The deeper the color is, the greater the connectedness in the visual figure is. This suggests that our network reflects the actual situation of the COVID-19 pandemic. Moreover, based on

similarity, we divide the complex network into ten modularity classes. On the one hand, the community classification shows that the United States, Russia, and Colombia, which are suffering severely from the COVID-19 pandemic, are in one community, and China and Japan, which have better epidemic control, are in another community. On the other hand, the community classification is closely related to each country's geographical location. Finally, the analysis of influencing factors reveals that countries with considerable economic size, high population density, and high government expenditure have a higher network connectedness value. In contrast, countries with adequate import capacity and better medical levels can reduce the connectedness. In short, these empirical results can clearly explain the reasons for the rapid global spread of the COVID-19 pandemic.

The contributions of this study are as follows. First, the global COVID-19 pandemic network is constructed based on the number of infected people, and the topological information and connectedness characteristics of the network are obtained, to some extent explaining the spread of the COVID-19 pandemic from the perspective of a complex network. The global COVID-19 pandemic network is a complex network with the characteristics of short average distance and large clustering coefficient and has a good community structure division. Second, integrating the visualization method into the global COVID-19 pandemic network provides visual results of network connectedness, helping us determine the potentially dangerous countries. Finally, we identify economic and social factors affecting COVID-19 network connectedness, including population density, GDP, exports, government spending, and life expectancy. To reduce network connectedness, the government should implement measures to reduce population density, increase medical equipment imports, improve the medical treatment level, and effectively allocate government expenditures.

LITERATURE REVIEW

Vespignani, a professor who is famous for modeling the global spread of epidemics, proposes that there are inherent limitations to predictions in complex sociotechnical systems. However, mathematical and computational models have successfully forecast the size of epidemics and have been used to communicate the risks of uncurbed infectious disease outbreaks in recent years [5]. Since the beginning of the COVID-19 outbreak, many mathematical models have been proposed to describe the pandemic's spread. Some of them forecast the future dynamics of COVID-19. Perc et al. [6] forecast COVID-19 by using a simple iteration method that needs only the daily values of confirmed cases as input. Other researchers focus on diagnosing the spread of COVID-19 and analyzing the factors of disease transmission. Liu et al. [7] propose a susceptible-asymptomatic-infected-removed (SAIR) model on social networks to describe the spread of COVID-19 and analyze the outbreak based on epidemic data from Wuhan. Kraemer et al. [8] show that China's drastic control measures substantially

mitigated the spread of COVID-19. Tsiotas and Magafas [9] build on the visibility graph algorithm to study the Greek COVID-19 infection curve as a complex network. These studies consider the spread of COVID-19 to be complicated, and diagnosing the scale and characteristics of the COVID-19 pandemic is conducive to governments effectively blocking the spread of this epidemic in their nations. Hence, we believe that it is very effective and important to prevent and control the further spread of COVID-19 by studying the connectedness characteristics of the pandemic complex network.

Studying the connectedness characteristics in a complex network through theoretical analysis and numerical simulation is helpful in understanding the microscopic mechanism of the spread process of an epidemic to discover practical ways to control the spread. Previous research has shown that a real social network is a typical complex network with small-world, scale-free, and high clustering characteristics. Small-world characteristics accelerate the spread of an epidemic, scale-free features facilitate the spread of an epidemic, and high clustering characteristics cause community clustering in epidemic transmission [5, 10–17].

Hence, some scholars have combined complex network theory with epidemiology. Ni [18] builds a dynamic model of an epidemic and shows that infectious diseases spread much faster in public networks than in in-home networks. In the beginning, the number of infected people in the network increases exponentially, and the spread speed obviously increases with increasing local structural scale. Chen et al. [19] simulate the propagation of the H1N1 virus based on complex networks. The simulation results show that the targeted immunization strategy can effectively inhibit the spread of diseases. Wang et al. [3] deeply investigates influential spreaders' identification in complex networks based on various centrality indices. Wang et al. [4] use a two-layer network model method to find that people who received positive prevention information had a lower risk of disease infection than those who received negative prevention information.

All the abovementioned studies try to provide feasible ways to curb the spread of the epidemic by using mathematical models or complex network tools. They can help us to work with limited data to solve problems in a constantly changing environment. However, at present, there are few studies that examine the global network of COVID-19. Therefore, determining the complicated connectedness of the COVID-19 complex network can fill gaps in the literature and effectively help the government identify good policies to prevent the COVID-19 pandemic in a timely manner. This study attempts to show the global situation of the COVID-19 pandemic based on a complex network. It profoundly analyzes the significant economic and social factors that affect the connectedness of the COVID-19 network to powerfully supplement the above literature.

DATA

The growth rate of COVID-19-infected people in each country is calculated using the COVID-19 infection sample from January

19, 2020, to August 15, 2020. The data came from the CSMAR database. To analyze the influencing factors of global COVID-19 infection network connectedness, we use each country's economic and social macrodata in 2018 as samples. Economic and social variables include GDP, total exports, total imports, inflation rate, government expenses, population density, and life expectancy. All these data are from the World Bank database. From them, we take logarithms of GDP, total exports, total imports, government expenses, and population density to reduce the absolute values.

METHODOLOGY

Construction of COVID-19 Pandemic Network

In this study, we build a COVID-19 infection complex network among 122 countries. Any node in our network is a specific country. Moreover, the possible connection between any two countries is called an edge in the network. These nodes and connected edges constitute the entire network of global COVID-19 infections. The network may rely more on certain countries, as many paths pass through them. Once this kind of country is identified and excluded, the network may collapse. Therefore, identifying these import infection nodes can enable governments to better fight the pandemic.

The key step in constructing the COVID-19 pandemic network is measuring the correlation of the growth rate of COVID-19-infected people among countries. The specific construction process of the COVID-19 pandemic network is as follows.

First, we exclude countries with fewer than 2,000 infected people and then calculate the growth rate of infected people in the remaining 122 countries individually.

$$g_{i,t} = \left(\frac{\text{infected}_{i,t}}{\text{infected}_{i,t-1}} - 1 \right) \times 100\% \quad i = 1, 2, \dots, 122, \quad (1)$$

where $g_{i,t}$ is the infected persons growth rate of country i in month t , $\text{infected}_{i,t}$ is the number of infected persons of country i in month t , and $\text{infected}_{i,t-1}$ is the number of infected persons in country i in month $t-1$.

Second, we compute the correlation coefficients of the infected persons' growth rate between any two countries. The correlation coefficient is formulated as follows:

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X]\text{Var}[Y]}}, \quad (2)$$

where $\text{Cov}(X, Y)$ is the covariance of X and Y , $\text{Var}[X]$ is the variance of X , and $\text{Var}[Y]$ is the variance of Y .

Third, we need to set a threshold for correlation coefficients to define the existence of the possible connection. To fulfill this purpose, we take half of the mean of each country's maximum correlation coefficient (approximately 0.4) as the threshold value for the existence of connectedness. If the correlation coefficient between the two countries is higher than the threshold, a line is drawn in

the network to connect these two countries, indicating a connection between them.

This method can test the network connectedness between all countries, draw undirected connecting lines in the network, and finally create a COVID-19 pandemic infection network.

Complex Network Characteristics Centrality Analysis

Centrality is an important indicator of each node’s status and role in the network, and it can be used to judge whether an individual node is influential [3]. The higher a node’s centrality is, the more significant its influence in the network is, suggesting that it can influence other nodes more prominently. In complex network analysis, there are many ways to characterize the centrality of a node. The most important are betweenness centrality, closeness centrality, eigenvector centrality, and page rank.

1) Betweenness centrality

In a network, a certain node’s degree may be very small, but it may be the intermediate node between any two other nonadjacent nodes. If this node is removed, the connectivity between the nonadjacent nodes will be interrupted. Therefore, the node plays a vital role in the network. For such a node, we can define betweenness.

The core idea is that the interaction between two nonadjacent nodes depends on other nodes in the network, especially those on the shortest path. They exert a certain control over the two nonadjacent nodes. For example, if node A is located on multiple shortest paths between other nonadjacent nodes, node A plays a greater role and has greater betweenness centrality.

Freeman [20] suggests that the shortest path between nonadjacent nodes v_j and v_l will pass some nodes in a network. If v_i is passed by many other shortest paths, then it is important in the network. Moreover, the importance or influence of v_i for v_j and v_l can be represented by B_i , defined as follows:

$$B_i = \sum_{1 \leq j \leq l \leq N} \left[\frac{n_{jl}(i)}{n_{jl}} \right], \tag{3}$$

where n_{jl} is the number of shortest paths between nodes v_j and v_l ; $n_{jl}(i)$ is the shortest path between nodes v_j and v_l , which passes through node v_i ; and N is the total number of nodes in the network.

Then, the betweenness centrality of v_i is defined as follows:

$$C_B(v_i) = \frac{2B_i}{[(N - 1)(N - 2)]}. \tag{4}$$

2) Closeness centrality

Sabidussi [21] proposes that the closeness centrality of v_i measures the extent of proximity to other nodes through the shortest path. The closer the node is, the more important it is in the network. The closeness centrality value is the reciprocal of the average distance between node

v_i and all other nodes that v_i can reach. Specifically, closeness centrality is defined as follows:

$$C_i^c = \frac{1}{L_i} = \frac{n - 1}{\sum_{j \in \Gamma_i, j \neq i} d_{ij}}, \tag{5}$$

where Γ is the set of nodes that v_i can reach, L_i is the average distance to all other reachable nodes of v_i , and d_{ij} is the distance between v_i and v_j . The closeness indicator describes the difficulty of reaching other nodes in the network through a certain node v_i . It reflects the ability of node v_i to exert influence on other nodes in the network. Therefore, the closeness centrality can effectively reflect the global structure of the network.

3) Eigenvector centrality

Bonacich [22] suggests that eigenvector centrality is another important centrality measurement method for nodes. For node v_i , let the centrality score x_i be proportional to the sum of the centrality scores of all nodes connected to it.

$$x_i = \frac{1}{\lambda} \sum_{j=1}^N a_{ij}x_j, \tag{6}$$

where N is the total number of nodes and λ is the maximum eigenvalue. Each score of the eigenvector must be positive.

4) Page rank

Page rank [23] can be understood as a random walk on a graph. It can be defined as follows:

$$x_k = \sum_{i=1}^N a_{ik}x_i = (1 - d) + d \sum_{i: b_{ik}=1} \frac{x_i}{r_i}, \tag{7}$$

where x_k is the page rank value of node k ; d is the model parameter, usually $d = 0.85$; and a_{ik} is the probability of transition from page i to page j .

Small-World Network

The small-world property is an important topological feature of the network. Watts and Strogatz [10] propose the small-world network model (WS small-world network). They argue that some network systems are large in size but have a relatively small distance between any two nodes. A typical small-world network has a shorter average distance and a larger clustering coefficient. For example, there are no more than six people between any two strangers in a small-world social network.

1) Clustering coefficient

In a WS small-world network, the clustering coefficient quantifies the tendency of the nodes of a complex graph to cluster. In terms of a certain node v in the WS small-world network, v_i and v_j are two neighboring nodes of node v . If we reconnect v_i and v_j by random probability p , then the probability of v_i and v_j still being neighboring nodes is $(1 - p)$, and $(1 - p)^3$ is the probability that there is still an edge between v_i and v_j . Watts

and Strogatz [10] and Barrat and Weigt [24] argue that the WS clustering coefficient can be calculated as follows:

$$C_{ws}(p) = \frac{3(k-1)}{2(2k-1)}(1-p)^3. \quad (8)$$

The initial network is a one-dimensional lattice of N sites, with periodic boundary conditions (i.e., a ring), each vertex being connected to its $2k$ nearest neighbors. The vertices are then visited one after the other; each link connecting a vertex to one of its k nearest neighbors in the clockwise sense is left in place with probability $(1-p)$ and with probability p is reconnected to a randomly chosen other vertex.

2) Average path length

The average distance between any two nodes would also show interesting properties in a given complex network. The diameter D of a complex network is defined as the maximum distance between any pair of nodes.

Watts and Strogatz [10] argue that in terms of random networks, almost all networks have the same diameter. This means that the N th-order random network with probability P has a tiny change in diameter. Watts [25] suggests that the diameter D is usually concentrated as follows:

$$D = \frac{\ln N}{\ln \langle k \rangle} \approx \frac{\ln N}{\ln(pN)}. \quad (9)$$

Supposing L_{ER} is the average path length of the ER random network, for a randomly selected node in the ER random network, there are approximately $\langle k \rangle^{L_{ER}}$ other nodes in the network with a distance from the selected node that is equal or very close to L_{ER} . Therefore, $L_{ER} \propto \frac{\ln N}{\ln \langle k \rangle}$. It is a typical small-world property that the average path length is proportional to the logarithmic growth function. Since $\ln N$ grows slowly with N , a large random network may have small average path lengths.

Modularity Class

Newman [26, 27] proposes modularity, which enables a clear evaluation indicator to measure the segmentation of network communities. In other words, modularity refers to the difference between an ordinary network and a random network under a specific community division criterion. Because the random network does not have a community structure, the larger the difference is, the better the community division will be.

Newman [27] defines modularity as follows:

$$Q = \frac{1}{4m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) (s_i s_j + 1) = \frac{1}{4m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) s_i s_j, \quad (10)$$

where m is the total number of edges in the network, A is the corresponding adjacency matrix of the network, and $A_{ij} = 1$ represents the existence of an edge between node i and node j . Otherwise, there is no edge connection. k_i is the degree of node i , s_i is the label of node i belonging to a community, and $\frac{1}{2}(s_i s_j + 1) = 1$ if and only if $s_i = s_j$; otherwise, $\frac{1}{2}(s_i s_j + 1) = 0$.

EMPIRICAL RESULTS AND DISCUSSION

COVID-19 Pandemic Network

In this section, we present our COVID-19 pandemic network. We first calculate the daily growth rate of the infected population in 122 countries individually and then calculate the correlation coefficient (ρ_{ij}) of the growth rate between any two countries. Furthermore, the threshold value for the existence of connectedness is set as 0.4 since the half mean of each country's maximum correlation coefficient is 0.38. If the correlation coefficient between country i and country j is higher than the threshold value, i.e., $\rho_{ij} > 0.4$, then the element $\{a_{ij}\}$ in the connectedness matrix is 1; otherwise, it is 0. If $a_{ij} = 1$, there is an edge between country i and country j , and if $a_{ij} = 0$, there is no edge between country i and country j . In addition, the value of the element on the diagonal of the connectedness matrix is set as 0, which indicates that there is no link between any country and itself. We obtain our COVID-19 pandemic connectedness matrix through the above calculation method (as the number of countries involved is 122, we cannot present the 122×122 matrix here). Because a zero growth rate will affect the similarity calculation, we excluded 91 countries and regions with fewer than 2,000 infected people.

Finally, the COVID-19 pandemic network obtained in this study is a complex network without considering the direction, and it is an undirected network with 122 nodes and 958 connected edges.

Using the software Gephi, we generate a visual graph of our network. The results are shown in **Figure 1**. Taking China, the United States, and Brazil as examples, China has connectedness with Singapore, Japan, and Thailand; the United States has connectedness with Australia, Kenya, Colombia, Pakistan, Dominica, Turkey, and Luxembourg; and Brazil has connectedness with 16 countries, including Austria, Portugal, Denmark, Saudi Arabia, and France.

Then, we use Gephi to obtain the network topology characteristic values, which is the basis for the next step of the analysis. **Figures 2 and 3** show the COVID-19 pandemic network parameters. **Figure 2** shows the degree distribution in our COVID-19 pandemic network. The overall distribution of the graph shows that the COVID-19 pandemic network is relatively dispersed. **Figure 3** shows the clustering coefficient graph. The larger the clustering coefficient is, the higher the extent of the connection between neighbors is. Our COVID-19 pandemic network has a clustering coefficient of 0.58. This large clustering coefficient indicates that the COVID-19 pandemic network is a very close world.

Other overall characteristics of the COVID-19 pandemic network are shown in the following table.

Table 1 shows that the COVID-19 pandemic network has 122 nodes, and the number of edges between nodes is 985 with an average degree of 7.85, average path length of 3.104, clustering coefficient of 0.58, and diameter of 8. Moreover, the clustering coefficient is large, so the network has close-world network characteristics.

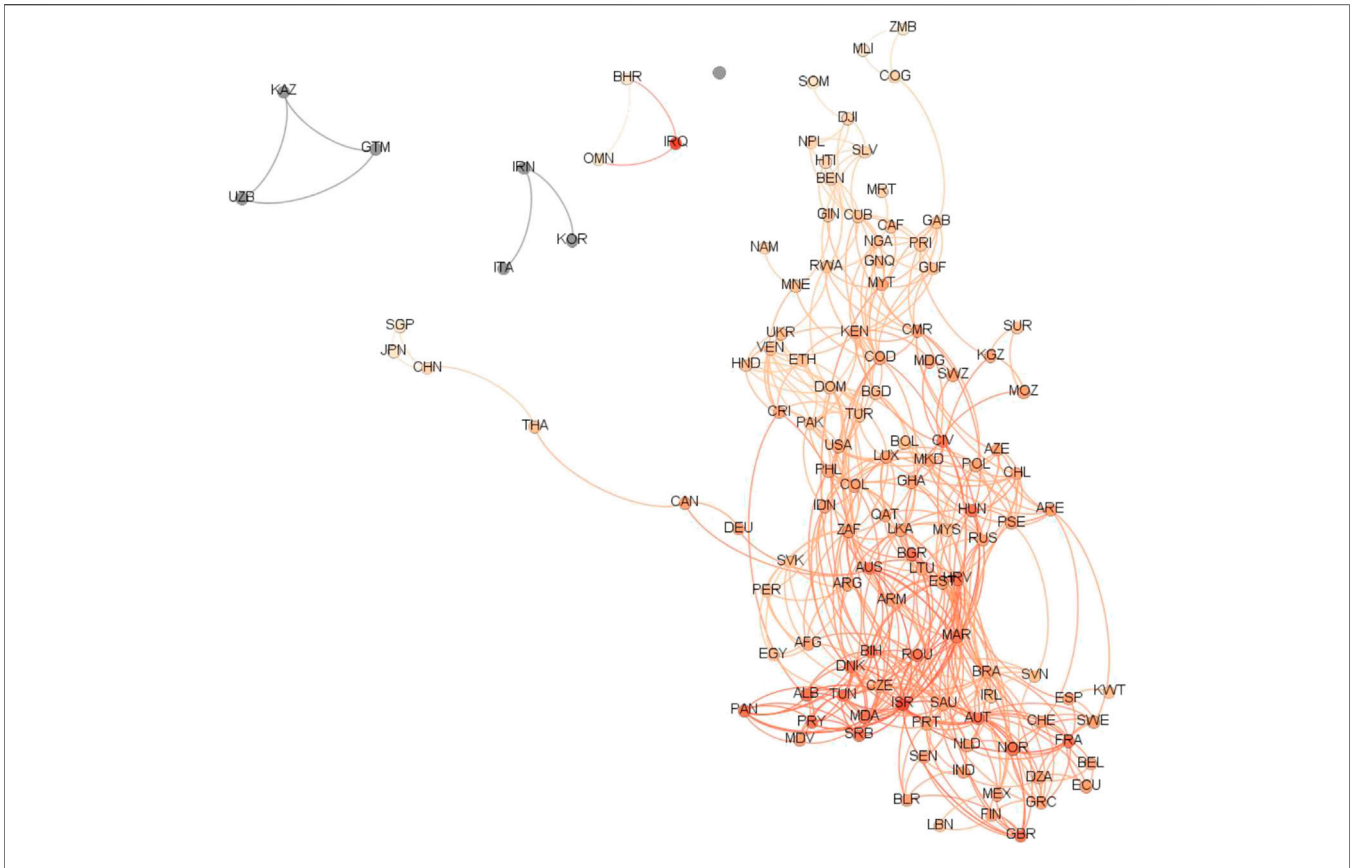


FIGURE 1 | The global COVID-19 pandemic network. Note: This is the global COVID-19 pandemic network. It excludes countries with fewer than 2,000 infected people. When the correlation of the growth rate of infected people is larger than 0.4, there is a connected edge between countries.

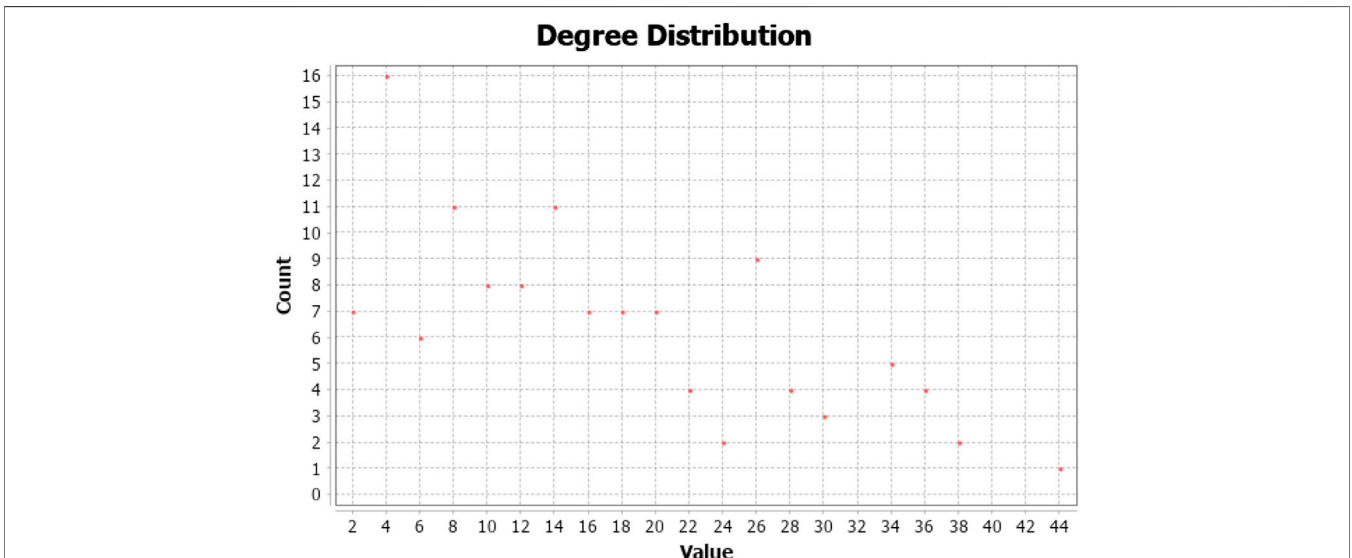


FIGURE 2 | The degree distribution of the global network. Note: **Figure 2** represents the degree distribution. The vertical and horizontal axes of each subplot are node counts and parameter distribution values, respectively. The horizontal axis represents the degree value of each country. The vertical axis represents the number of countries. For example, value = 2 and count = 7 means that there are seven countries, and they all have a degree of two.

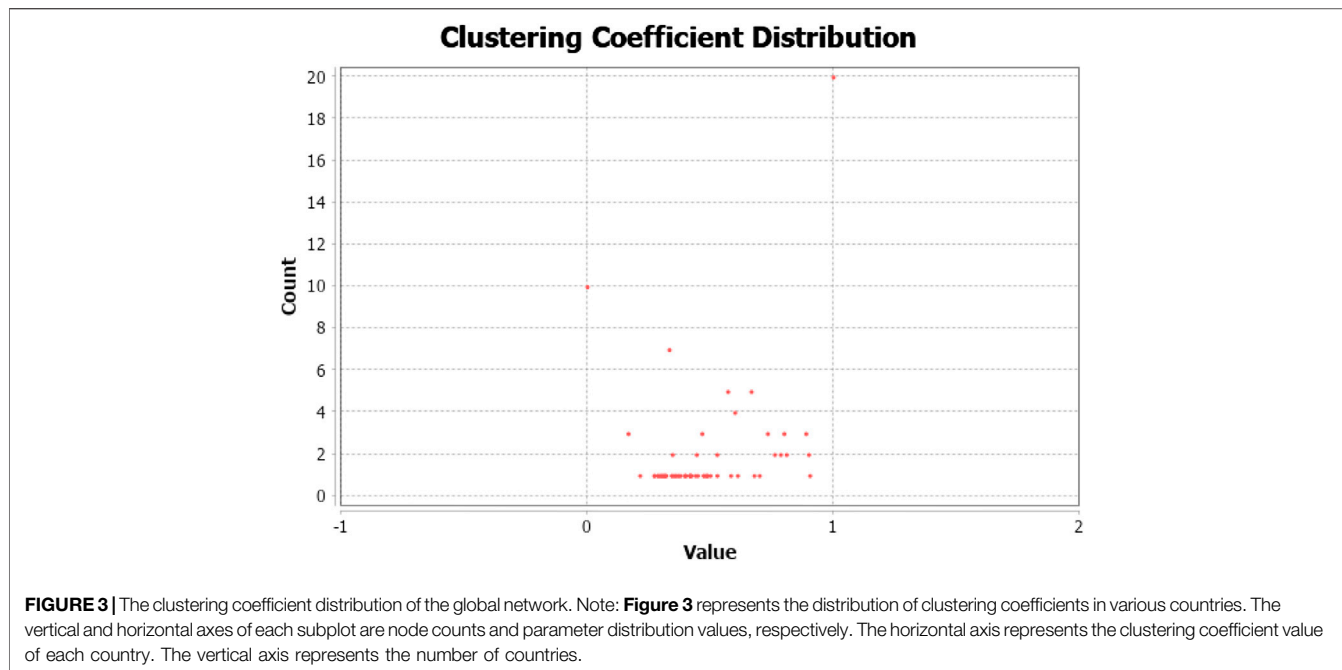


TABLE 1 | Topological characteristics of the COVID-19 pandemic network.

Network	N	E	<d>	K	C	D
COVID-19 pandemic	122	958	7.85	3.104	0.58	8

Table 1 presents the topological information of the COVID-19 pandemic network. *N* is the number of nodes. *E* is the number of edges. *<d>* is the average degree of the whole network. *K* represents the average path length of the network. *C* is the clustering coefficient of the network. *D* is the diameter of the network.

In addition, the small-world network has the characteristics of a small path length and a high clustering coefficient between nodes. Our global COVID-19 pandemic network’s average path length is 3.104, and its clustering coefficient is 0.58, indicating that the global COVID-19 pandemic network has some characteristics of a small-world network.

Connectedness Characteristics of the Global COVID-19 Pandemic Network

In this section, we further analyze some important characteristics of the global COVID-19 pandemic network. Specifically, we focus on the degree, triangles, centrality, modularity class, and authority and then provide the visual results.

Figure 4 represents the degree of each country in the global COVID-19 pandemic network. The degree of country *i* is the total number of edges directly connected to that country. The deeper a country’s color is, the greater its degree is. As shown in **Figure 4**, Brazil has the deepest color on the American continents; in Europe, Austria has the greatest degree; in Asia, Israel and Sri Lanka have the highest number of edges; in Oceania, the color of Australia is the deepest; and in Africa, the countries with the greatest degree are South Africa and Morocco. The above results

indicate that these countries have the greatest degrees with other countries and a prominent ability to spread COVID-19 to other countries. Thus, more attention should be paid to their capacity and measures for COVID-19 pandemic prevention and control.

Figures 5–8 show the centrality analysis results, reflecting the relative importance of each node in the network. **Figure 5** shows betweenness centrality. The countries with the deepest colors are the Democratic Republic of the Congo, Australia, and Croatia, which are located in Africa, Oceania, and Europe, respectively. This finding indicates that these three countries play an important betweenness role in the COVID-19 pandemic network. Additionally, Canada and Cuba in North America are also significant. We may more easily halt the global spread of COVID-19 if we control the pandemic in these three countries first.

Figure 6 shows eigenvector centrality. The three countries with the deepest colors are Moldova, Sri Lanka, and Austria, located in Europe, Asia, and Europe, respectively. This finding shows that the three countries have the highest centrality scores and are the nodes with the greatest importance in the COVID-19 pandemic network. Again, Austria shows its importance in the COVID-19 pandemic network, indicating that the effective control of COVID-19 in Austria has a significant role in halting the global spread of the COVID-19 pandemic.

Figure 7 represents closeness centrality. The three countries with the deepest colors are Kazakhstan, Iraq, and Uzbekistan, all located in Asia. This means that Kazakhstan, Iraq, and Uzbekistan may be important for Asian COVID-19 control. **Figure 8** shows harmonic closeness centrality. The empirical results are the same as those shown in **Figure 7**.

Figure 9 shows the analysis results of authority, reflecting the influence of nodes in the network. The three countries with the deepest colors are Moldova, Israel, and Austria, located in

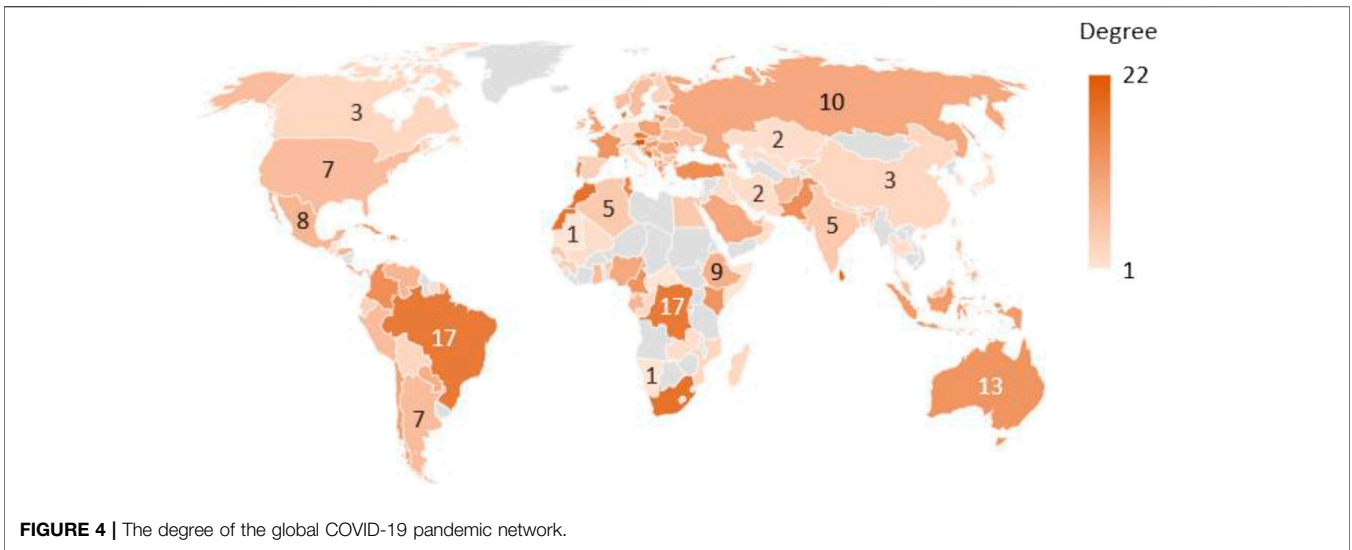


FIGURE 4 | The degree of the global COVID-19 pandemic network.

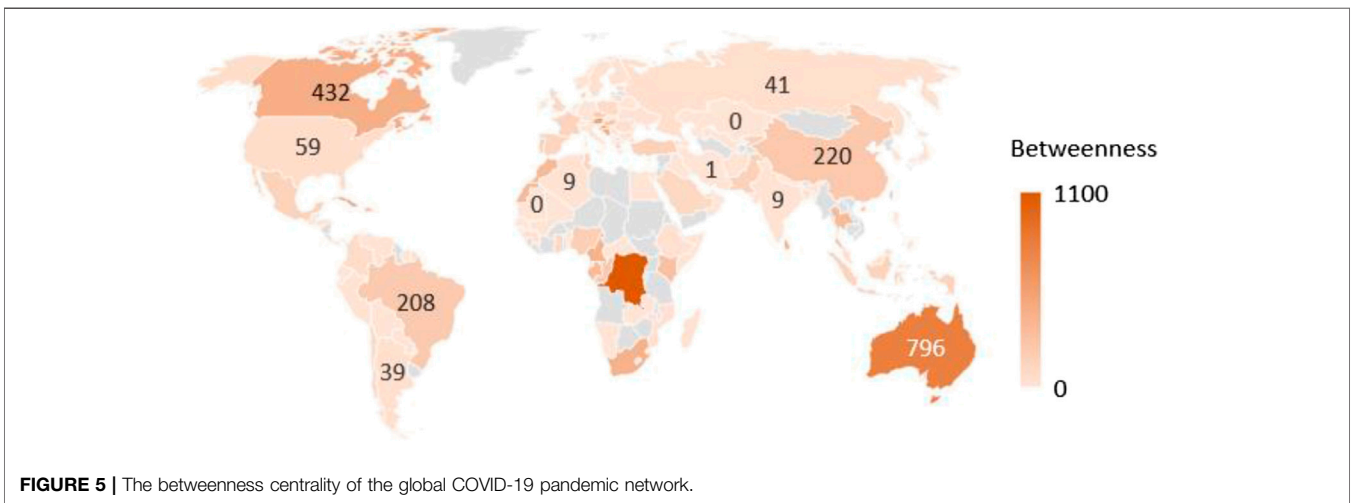


FIGURE 5 | The betweenness centrality of the global COVID-19 pandemic network.

Europe, Asia, and Europe, respectively. This indicates that the infections in Moldova, Israel, and Austria are highly dangerous, as these three countries may have high virus spread capacity.

Figure 10 shows modularity-class analysis results, describing the community situation in the network. The results show that the modularity coefficient is 0.526, and the COVID-19 pandemic network is divided into 10 communities, representing community structures. Community 0 contains 8 countries, including Palestine, Poland, and the United Arab Emirates. Community 1 contains Cuba, the Democratic Republic of the Congo, and 23 other countries. Community 2 contains Mexico, Moldova, Sri Lanka, Israel, and so on. Community 3 contains 13 countries, including China, Canada, Britain, India, Germany, Australia, and Japan. Community 4 contains 24 countries, including France, Brazil, Austria, and South Africa. Community 5 contains Bahrain, Iraq, and Oman. Community 6 contains 11 countries, including Bangladesh, Kyrgyzstan, and Ghana. The countries in Community 7 are Uzbekistan, Kazakhstan, and Guatemala.

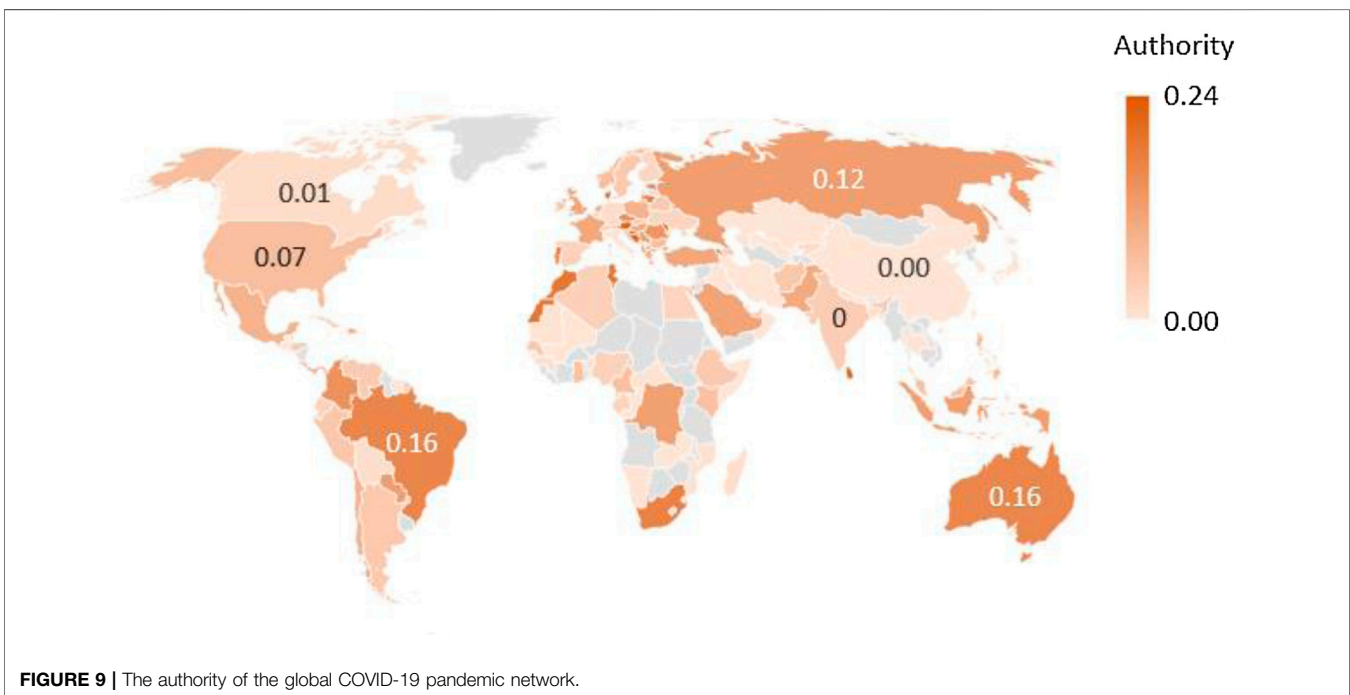
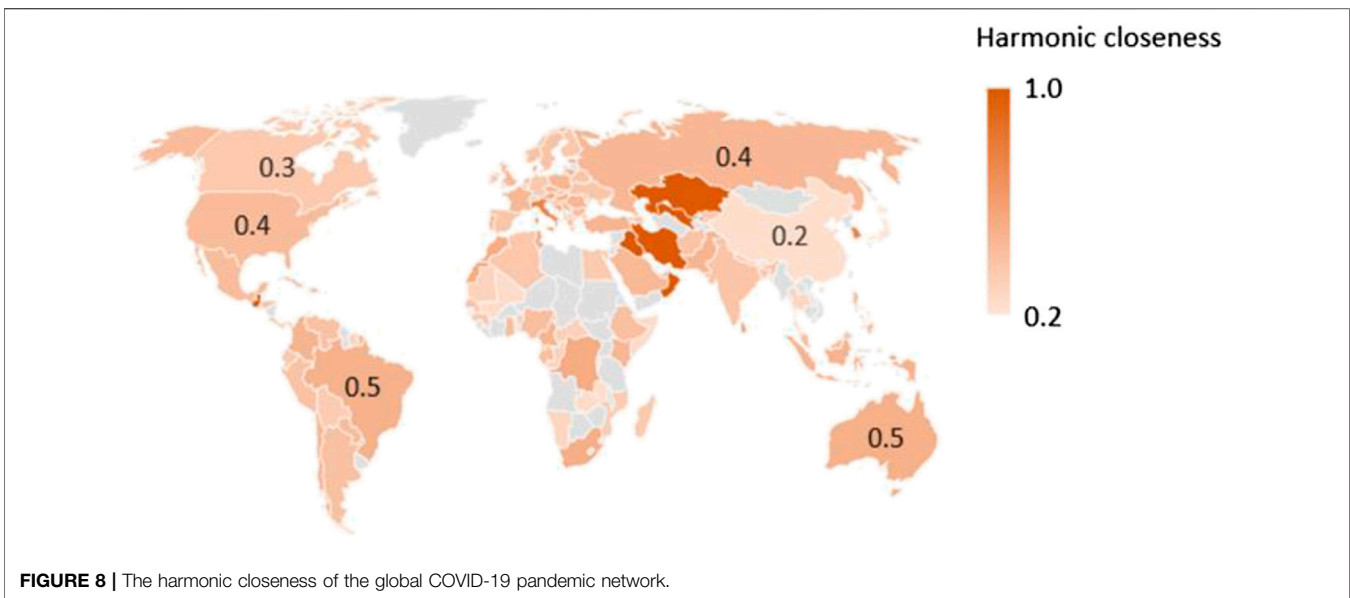
The countries in Community 8 are Iran, South Korea, and Italy. Community 9 contains 16 countries, including the United States, the Russian Federation, and Colombia.

The Determinants of Connectedness

According to the above analysis, the connectedness of the global COVID-19 infection network varies across different countries. Therefore, what factors influence the network connectedness of COVID-19 infection? Hereafter, this study further explores how economic factors, population density, medical level, and other social factors influence the connectedness of COVID-19 by using a quantitative regression method to provide a theoretical basis and guidance for policymakers' prevention and control measures.

We build the econometric model as follows:

$$\begin{aligned} \text{connectedness}_i = & \beta_0 + \beta_1 \text{LnDensity} + \beta_2 \text{LnGDP} + \beta_3 \text{LnExport} \\ & + \beta_4 \text{LnImport} + \beta_5 \text{Inflation} + \beta_5 \text{Expense} \\ & + \beta_6 \text{LifeExpectancy} + \varepsilon_i, \end{aligned}$$



increase page rank. However, LnImport is significantly negative at 5%, indicating that imports significantly reduce page rank. In columns (7) and (8), Expense has a significantly positive coefficient at 1%, meaning that the amount of government expenditure will significantly increase authority and hub. In column (9), LnGDP and Expense are significantly positive at 5%, indicating that the size of the economy and government spending significantly increase triangles.

Overall, the economic development level and government spending show significant positive effects on COVID-19 infection connectedness. If policymakers hope to curb the spread of COVID-19 effectively, they should give priority to

the COVID-19 pandemic, appropriately reduce economic activities, and avoid crowd gatherings to reduce the risk of infection. Moreover, government expenditures need to be allocated more effectively. Controlling and preventing the spread of COVID-19 should be the first goal of government expenditures. More government resources should be used to purchase medical devices and improve medical staff protection and treatment. The United States is a case in point. As the United States has prioritized the economy in its COVID-19 prevention policy, a large amount of fiscal spending has been used to subsidize economic activities rather than medical resources. As a result, the number of

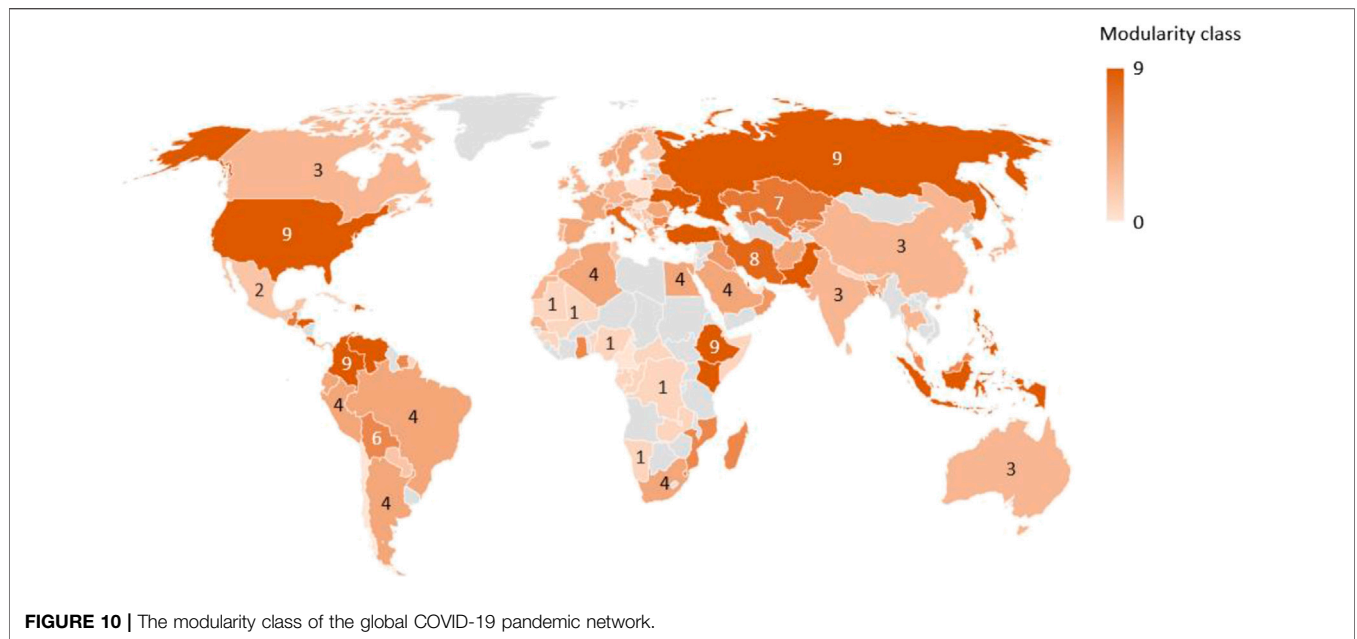


FIGURE 10 | The modularity class of the global COVID-19 pandemic network.

TABLE 2 | Determinants of connectedness: global.

Variables	(1) Degree	(2) Eigen	(3) Betweenness	(4) Closeness	(5) Harmonic	(6) Page rank	(7) Authority	(8) Hub	(9) Triangles
LnDensity	0.7450 (1.46)	0.0217 (0.75)	-13.9320 (-0.53)	-0.0020 (-0.12)	0.0024 (0.14)	0.0007** (2.03)	0.0036 (0.55)	0.0036 (0.55)	2.1949 (1.50)
LnGDP	2.9796** (2.28)	0.1207* (1.76)	72.1867** (2.39)	0.0440 (1.60)	0.0551* (1.81)	0.0025*** (3.51)	0.0249 (1.63)	0.0249 (1.63)	8.1021** (2.08)
LnExport	0.2037 (0.10)	-0.0009 (-0.01)	-10.2764 (-0.20)	-0.0181 (-0.52)	-0.0111 (-0.27)	0.0004 (0.31)	-0.0014 (-0.06)	-0.0014 (-0.06)	0.3559 (0.05)
LnImport	-4.1171 (-1.50)	-0.1654 (-1.12)	-73.2675 (-1.06)	-0.0328 (-0.80)	-0.0537 (-1.05)	-0.0036** (-2.36)	-0.0341 (-1.03)	-0.0341 (-1.03)	-12.4868 (-1.48)
Inflation	-0.0531 (-0.78)	-0.0037 (-1.03)	-6.4161** (-2.39)	-0.0000 (-0.02)	-0.0001 (-0.02)	-0.0001* (-1.78)	-0.0010 (-1.29)	-0.0010 (-1.29)	-0.0766 (-0.33)
Expense	0.2523*** (3.53)	0.0126*** (3.47)	2.4979 (1.06)	-0.0001 (-0.03)	0.0012 (0.55)	0.0001*** (3.39)	0.0029*** (3.48)	0.0029*** (3.48)	0.7041*** (3.25)
Life	0.0605 (0.54)	0.0067 (1.12)	0.6182 (0.14)	0.0031 (1.28)	0.0033 (1.29)	0.0000 (0.46)	0.0019 (1.38)	0.0019 (1.38)	0.5413 (1.47)
Constant	7.0192 (1.00)	0.0728 (0.21)	217.2144 (0.92)	0.2184 (1.48)	0.2472 (1.45)	0.0080* (1.95)	-0.0029 (-0.04)	-0.0029 (-0.04)	8.3716 (0.41)
Observations	79	79	79	79	79	79	79	79	79
R-squared	0.228	0.212	0.102	0.039	0.057	0.232	0.218	0.218	0.232

*This table constructs regression models for global determinants of connectedness. Dependent variables are network connectedness characteristics. Independent variables are economic and social factors. Values in parentheses are the T statistical values; ***, **, and * indicate significance at the levels of 1, 5, and 10%, respectively.*

confirmed COVID-19 cases in the United States is approaching 9 million.

The Determinants of Connectedness: Asia and Europe

To further analyze how economic and social factors affect Asian and European countries' connectedness characteristics differently, this study divides the samples into Asian and European countries.

Table 3 shows the empirical results in Asia. In Asia, LnDensity, LnGDP, and LnExport are significantly positive at the 10% confidence level for degree, eigenvector, authority, hub, and triangles, indicating that population density, economic activities, and export activities show significantly positive effects in the Asian COVID-19 pandemic network. However, LnImport is persistently significantly negative at the 10 or 5% confidence level for degree, eigenvector, page rank, authority, hub, and triangles, showing that imports can effectively reduce the characteristics of the COVID-19 pandemic network in Asia.

TABLE 3 | Determinants of connectedness: Asia.

Variables	(1) Degree	(2) Eigen	(3) Betweenness	(4) Closeness	(5) Harmonic	(6) Page rank	(7) Authority	(8) Hub	(9) Triangles
LnDensity	3.1811* (2.28)	0.1962** (2.87)	21.1112 (0.69)	-0.0711 (-1.01)	-0.0586 (-0.88)	0.0013* (1.92)	0.0438** (2.87)	0.0438** (2.87)	10.8935** (2.80)
LnGDP	6.5357* (2.00)	0.3872* (2.30)	62.9644 (0.79)	0.0091 (0.09)	0.0281 (0.26)	0.0038* (1.96)	0.0874* (2.29)	0.0874* (2.29)	21.9928* (2.24)
LnExport	10.9008 (1.78)	0.6387* (2.16)	50.8330 (0.37)	0.0305 (0.16)	0.0631 (0.31)	0.0059 (1.36)	0.1464* (2.19)	0.1464* (2.19)	35.9940* (1.96)
LnImport	-21.0506** (-2.43)	-1.2356** (-3.04)	-149.3194 (-0.74)	-0.0607 (-0.25)	-0.1177 (-0.46)	-0.0118* (-1.92)	-0.2827** (-3.06)	-0.2827** (-3.06)	-70.3091** (-2.81)
Inflation	0.0684 (0.19)	-0.0058 (-0.34)	-7.8613 (-0.82)	0.0080 (0.52)	0.0072 (0.48)	0.0000 (0.19)	-0.0025 (-0.65)	-0.0025 (-0.65)	0.4325 (0.41)
Expense	0.1327 (0.46)	0.0092 (0.67)	3.9181 (0.74)	-0.0034 (-0.35)	-0.0019 (-0.20)	0.0000 (0.22)	0.0023 (0.74)	0.0023 (0.74)	0.3640 (0.43)
Life	-0.3162 (-0.69)	-0.0248 (-1.13)	-2.2921 (-0.22)	0.0116 (0.56)	0.0087 (0.42)	0.0001 (0.44)	-0.0051 (-1.05)	-0.0051 (-1.05)	-0.6678 (-0.50)
Constant	66.9493* (1.88)	4.1986** (2.42)	655.5296 (0.83)	0.2487 (0.16)	0.5011 (0.29)	0.0241 (1.08)	0.9375** (2.39)	0.9375** (2.39)	188.7532 (1.77)
Observations	16	16	16	16	16	16	16	16	16
R-squared	0.545	0.632	0.250	0.311	0.292	0.520	0.640	0.640	0.611

Note: This table constructs regression models for the determinants of connectedness in Asia. Dependent variables are network connectedness characteristics. Independent variables are economic and social factors. Values in parentheses are the T statistical values; ***, **, and * represent significance at the levels of 1, 5, and 10%, respectively.

TABLE 4 | Determinants of connectedness: Europe.

Variables	(1) Degree	(2) Eigen	(3) Betweenness	(4) Closeness	(5) Harmonic	(6) Page rank	(7) Authority	(8) Hub	(9) Triangles
LnDensity	0.9020 (0.60)	0.0689 (0.89)	28.5923 (0.72)	0.0345 (1.67)	0.0392 (1.68)	0.0007 (0.69)	0.0157 (0.87)	0.0157 (0.87)	3.5322 (0.74)
LnGDP	0.2223 (0.09)	0.0478 (0.36)	60.4216 (0.73)	0.0681 (1.37)	0.0740 (1.39)	0.0009 (0.56)	0.0140 (0.45)	0.0140 (0.45)	-0.1817 (-0.02)
LnExport	0.5186 (0.07)	0.1156 (0.30)	12.5290 (0.06)	0.0634 (0.89)	0.0832 (0.95)	-0.0007 (-0.14)	0.0137 (0.15)	0.0137 (0.15)	-7.0843 (-0.28)
LnImport	-2.1137 (-0.21)	-0.2598 (-0.49)	-113.1319 (-0.36)	-0.1503 (-1.25)	-0.1819 (-1.30)	-0.0010 (-0.14)	-0.0503 (-0.41)	-0.0503 (-0.41)	2.1277 (0.06)
Inflation	-1.0538** (-2.37)	-0.0673*** (-2.85)	-25.0360* (-1.85)	-0.0101** (-2.22)	-0.0120** (-2.15)	-0.0007** (-2.26)	-0.0168*** (-3.01)	-0.0168*** (-3.01)	-4.0368** (-2.49)
Expense	0.2490** (2.26)	0.0105* (1.83)	6.4767 (1.54)	0.0034* (1.99)	0.0045** (2.37)	0.0002** (2.60)	0.0023 (1.69)	0.0023 (1.69)	0.6057* (1.73)
Life	-0.5094* (-1.90)	-0.0381** (-2.62)	-8.2352 (-1.06)	-0.0066** (-2.61)	-0.0073** (-2.34)	-0.0003 (-1.48)	-0.0088** (-2.57)	-0.0088** (-2.57)	-1.7204* (-1.77)
Constant	62.6958** (2.78)	4.5161*** (3.93)	1,092.8542 (1.45)	0.9114*** (4.58)	1.0772*** (4.53)	0.0359** (2.30)	1.0523*** (3.94)	1.0523*** (3.94)	217.3826** (2.71)
Observations	33	33	33	33	33	33	33	33	33
R-squared	0.326	0.398	0.232	0.301	0.335	0.330	0.410	0.410	0.409

Note: This table constructs regression models for the determinants of connectedness in Europe. Dependent variables are network connectedness characteristics. Independent variables are economic and social factors. Values in parentheses are the T statistical values; ***, **, and * indicate significance at the levels of 1, 5, and 10%, respectively.

In summary, for Asian countries, population density and economic activities significantly increase the connectedness of the COVID-19 pandemic network, which is consistent with the full sample results. As most Asian countries have large populations and high residential density, it is very important and influential for them to decrease economic activities and strengthen quarantine measures appropriately. In addition, more substantial imports would reduce the connectedness of

the COVID-19 pandemic network in Asian countries. We suspect that one of the reasons is that most Asian countries' medical level is still relatively low. More imports of materials and medical equipment can effectively reduce domestic resource pressure, thus reducing connectedness.

Table 4 shows the empirical results for Europe. In Europe, the coefficients of Expense for degree, eigenvector, closeness, harmonic closeness, and page rank are significantly positive at

10 or 5%, meaning that government expenses in Europe may be a burden and that governments do not effectively use their resources to control the COVID-19 pandemic. The estimated coefficients of inflation for all characteristics are all significantly negative except for betweenness and page rank. All the estimated coefficients of Life are significantly negative at 10 or 5%, indicating that as countries with higher life expectancy are more likely to have a better medical level; life expectancy is negatively associated with the connectedness of the COVID-19 pandemic network.

Compared with Asian countries, economic activity, population density, and government expenditure do not significantly increase the connectedness of the COVID-19 pandemic network in Europe. However, life expectancy can effectively reduce network connections. As European countries have more capital intensive and high value-added industries compared with Asia, the infection possibility of COVID-19 is relatively low in those industries. We believe that cutting down economic activities can not effectively prevent and control the spread of COVID-19 pandemic in Europe. Moreover, as the population density in Europe is much lower than that in Asia, it is difficult for the government of European countries to lower their population density furthermore. Therefore, population density is not significantly associated with the connectedness of the COVID-19 pandemic in Europe. Compared with Asian countries, most European countries have a higher economic development level, so their government expenditure is generally used for medical expenditure. Government expenditure can effectively strengthen a country's medical resources. Therefore, life expectancy in Europe is higher, and life expectancy is negatively associated with the connectedness of the COVID-19 network, suggesting that the European medical level is useful for preventing and controlling COVID-19.

Due to differences in economic and social characteristics, policymakers should use different prevention and control measures for different countries. Each country should implement prevention policy measures according to its local conditions. Only in this way can we effectively control the further spread of COVID-19.

CONCLUSION

Based on the growth pattern of the number of COVID-19 infections worldwide, this study constructed a complex network among 122 countries. Complex network analysis was used to deconstruct the overall connectedness of COVID-19 infection networks, visualize the connectedness characteristics among countries in the network, and further explore the influencing factors of the global COVID-19 infection network based on macrodata such as the economy, population, and life expectancy.

Our research shows that, first, in terms of overall connectedness, the global COVID-19 pandemic network has

small-world network, close-world network, and community structure features. There are a total of 958 edges in 122 countries worldwide, which are divided into ten communities. This network has a small average path length (3.14) and a high clustering coefficient (0.58), and its diameter is 8. Second, in terms of intercountry connectedness, some countries have a significant influence on the global COVID-19 pandemic network, such as Austria (Europe), Croatia (Europe), Moldova (Europe), Brazil (South America), Canada (North America), Iran (Asia), Kazakhstan (Asia), Uzbekistan (Asia), Republic of the Congo (Africa), and Austria (Oceania). More attention should be paid to these countries for further control of COVID-19.

In addition, we divide the network into ten communities. The United States, Russia, and Colombia, which are suffering more severely from COVID-19, are in a community, while China and Japan, which have better control of COVID-19, are in another community.

Finally, in the influencing factor analysis, our study finds that population density, economic size, exports, and government expenditure significantly increased the connectedness of the COVID-19 infection network. However, imports and life expectancy significantly reduced the connectedness of the COVID-19 infection network. This result indicates that countries with large economic size, high population density, and high government expenditure more easily suffer from and spread COVID-19. In contrast, countries with adequate import capacity and a better medical level can reduce the network connectedness.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

SZ performed the coding, data analysis, and original draft. MK gave formal writing-review and editing. FL proposed the conceptualization and methodology and modified final draft. QF performed the funding acquisition, original draft preparation, and data preparation. GD performed the data collection and original arrangement.

FUNDING

This study was funded by the Scientific Research Fund Project of Yunnan Education Department (No.2021J0586), the Fundamental Research Funds for the Central Universities. (No.20wkpy85), and the China Postdoctoral Science Foundation (No. 2019M663356).

REFERENCES

- Moore C, Newman ME. Epidemics and percolation in small-world networks. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* (2000) 61(5): 5678–82. doi:10.1103/physreve.61.5678
- Stegehuis C, Van Der Hofstad R, Van Leeuwen JS. Epidemic spreading on complex networks with community structures. *Sci Rep* (2016) 6(1):29748–7. doi:10.1038/srep29748
- Wang J, Li C, Xia C. Improved centrality indicators to characterize the nodal spreading capability in complex networks. *Appl Math Comput* (2018) 334: 388–400. doi:10.1016/j.amc.2018.04.028
- Wang Z, Xia C, Chen Z, Chen G. Epidemic propagation with positive and negative preventive information in multiplex networks. *IEEE Trans Cybern* (2020) [Epub ahead of print]. doi:10.1109/TCYB.2019.2960605
- Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. Epidemic processes in complex networks. *Rev Mod Phys* (2015) 87(3):925. doi:10.1103/RevModPhys.87.925
- Perc M, Gorišek Miksić N, Slavinec M, Stožer A. Forecasting Covid-19. *Front Phys* (2020) 8:127. doi:10.3389/fphy.2020.00127
- Liu C, Wu X, Niu R, Wu X, Fan R. A new SAIIR model on complex networks for analysing the 2019 novel coronavirus (COVID-19). *Nonlinear Dynam* (2020) 1–11. doi:10.1007/s11071-020-05704-5
- Kraemer MUG, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* (2020) 368(6490):493–7. doi:10.1126/science.abb4218
- Tsiotas D, Magafas L. *The effect of anti-COVID-19 policies to the evolution of the disease: a complex network analysis to the successful case of Greece* (2020) arXiv preprint arXiv:2004.06536.
- Watts DJ, Strogatz SH. Collective dynamics of ‘small-world’ networks. *Nature* (1998) 393(6684):440–2. doi:10.1038/30918
- Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. *Phys Rev Lett* (2001) 86(14):3200–3. doi:10.1103/PhysRevLett.86.3200
- Newman ME. Communities, modules and large-scale structure in networks. *Nat Phys* (2012) 8(1):25–31. doi:10.1038/nphys2162
- Wang P, Tian C, Lu J-a. Identifying influential spreaders in artificial complex networks. *J Syst Sci Complex* (2014) 27(4):650–665. doi:10.1007/s11424-014-2236-4
- Estrada E, Hatano N, Benzi M. The physics of communicability in complex networks. *Phys Rep* (2012) 514(3):89–119. doi:10.1016/j.physrep.2012.01.006
- Liu Y-Y, Barabási A-L. Control principles of complex systems. *Rev Mod Phys* (2016) 88(3):035006. doi:10.1103/RevModPhys.88.035006
- Liao H, Mariani MS, Medo M, Zhang Y-C. Ranking in evolving complex networks. *Phys Rep* (2017) 689:1–54. doi:10.1016/j.physrep.2017.05.001
- Hu J, Chengyi X, Huijia L, Peican Z, Wenjun X. Properties and structural analyses of USA’s regional electricity market: a visibility graph network approach. *Appl Math Comput* (2020) 385:125434. doi:10.1016/j.amc.2020.125434
- Ni S. *Research on modeling of infectious disease spreading based on complex Network theory*. Beijing, China: School of Aeronautics and Astronautics, Tsinghua University (2009).
- Chen D, Huang S, Shang M. Complex network model and application in epidemic spreading and controlling. *Comput Sci* (2011) 38(6):118–21. doi:10.3969/j.issn.1002-137X.2011.06.028
- Freeman LC. A set of measures of centrality based on betweenness. *Sociometry* (1977) 40:35–41. doi:10.2307/3033543
- Sabidussi G. The centrality of a graph. *Psychometrika* (1966) 31(4):581–603. doi:10.1007/BF02289527
- Bonacich P. Some unique properties of eigenvector centrality. *Soc Network* (2007) 29(4):555–64. doi:10.1016/j.socnet.2007.04.002
- Page L, Brin S, Motwani R, Winograd T. The PageRank citation ranking: bringing order to the web. Stanford InfoLab (1999) 9(1):1–14.
- Barrat A, Weigt M. On the properties of small-world network models. *Eur Phys J B* (2000) 13(3):547–60. doi:10.1007/s100510050067
- Watts DJ. Networks, dynamics, and the small-world phenomenon. *Am J Sociol* (1999) 105(2):493–527. doi:10.1086/210318
- Newman ME. Fast algorithm for detecting community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys* (2004) 69(6):066133. doi:10.1103/PhysRevE.69.066133
- Newman MEJ. Modularity and community structure in networks. *Proc Natl Acad Sci USA* (2006) 103(23):8577–82. doi:10.1073/pnas.0601602103

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Zhu, Kou, Lai, Feng and Du. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.