THE ROLE OF SOCIAL CONNECTEDNESS IN THE CORONAVIRUS DISEASE (COVID-19) PANDEMIC OUTCOME

Viktor Stojkoski
Faculty of Economics - Skopje, Ss. Cyril and Methodius University
vstojkoski@eccf.ukim.edu.mk

Dragan Tevdovski
Faculty of Economics - Skopje, Ss. Cyril and Methodius University
dragan@eccf.ukim.edu.mk

ABSTRACT
The magnitude of the coronavirus disease (COVID-19) pandemic has an enormous impact on the social life and the economic activities in almost every country in the world. Besides the biological and epidemiological factors, a multitude of social and economic criteria also govern the extent of the coronavirus disease spread in the population. Consequently, there is an active debate regarding the critical socio-economic factors that contribute to the resulting pandemic. In this paper, we contribute towards the resolution of the debate by examining the role of an individual's social connectedness in the extent to which the coronavirus spread. To measure social connectedness we create a network in which nodes represent countries and the interactions between pairs of countries are given by the number of shared Facebook contacts, whereas the coronavirus outcome is simply quantified as the number of registered cases per million population in the country. We find that there exists a robust and stable relationship between the level social importance of a country as well as the degree to which its people mix, and the outcome of the coronavirus health crisis. Countries which take a more central role in the network of social connections are also more susceptible to the coronavirus, whereas countries where there is less social mixing are less affected by the induced disease. Our results are an empirical verification for standard theories, which suggest that social network structures play a critical role in disease spreading processes. More importantly, they serve as a validation that social distancing measures introduced by governments are essential policies for preventing a fatal coronavirus outcome, and can be implemented for developing appropriate social distancing measures.

Keywords: *social connectedness, COVID-19, economic development, networks*

JEL classification: *C31, I15, I18*

1. INTRODUCTION

A multitude of social and economic criteria have been attributed as potential determinants for the magnitude of the coronavirus disease (COVID-19) pandemic, for a recent discussion see Stojkoski et al. (2020). However, despite an abundance of research, the potential impact of cross-country social interactions has been neglected. In this paper we aim to provide a valuable insight on the role of this phenomena and examine the impact of the social interactions on the coronavirus disease spreading process by examining country level data. To measure social connectedness we create a network in which nodes represent countries and the interactions between pairs of countries are given by the number of shared Facebook contacts, whereas the coronavirus outcome is simply quantified as the number of registered cases per million population in the country. We find that there exists a robust and stable relationship between the level social importance of a country as well as the degree to which its people mix, and the outcome of the coronavirus health crisis. Countries which take a more

http://doi.org/10.47063/EBTSF.2020.0022
http://hdl.handle.net/20.500.12188/9686

central role in the network of social connections are also more susceptible to the coronavirus. In turn, countries where there is less social mixing are less affected by the induced disease. This relationship is robust against both alternate specification of the model and the inclusion of other potentially significant contributors to the coronavirus outcome. As such, our findings serve as a verification for standard theories which emphasize the role of social networks in economic interactions.

The rest of the paper is organized as follows. Section 2 gives a short overview of the literature which motivated our research. In Section 3 we present the methods and data used to quantify the relationship between the coronavirus outcome and the social connectedness of a country. In Section 3 we present our findings, and in Section 4 we discuss their implications.

2. LITERATURE REVIEW

The primary contribution of this paper is empirical. It provides new information on the socio-economic determinants which helped shape the magnitude of the health crisis induced by the COVID-19 outbreak. As suchit joins the growing literature on the role of social networks of the COVID-19 outbreak 2020 (Bailey et al., 2018; Kuschler et al. 2020; Klepac et al. 2020). It also joins the literature on the socio-economic determinants which governed the coronavirus outbreak (Stojkoski et al. 2020; Jinjarak et al. 2020; Ehlert, 2020; Szulczyk and Cheema 2020).

From a methodological perspective the paper contributes to the literature which investigates the impact of the shape of an empirical social network structure on a disease-spreading process, especially those that take into account the process behind COVID-19 (Basnarkov, 2020; Pastor and Vespignani, 2001). Usually such researchrelies on micro scale longitudinal data, which cannot be easily obtained. This paper presents a way to use a reliable existing macro data source and thus allows to easily expand the analysis done here in the future.

3. METHODS AND DATA

We follow the methodology implemented in Stojkoski et al. (2020) and assume that the log of the registered COVID-19 cases per million population (p.m.p.) is a result of a disease spreading process. Obviously, the extent to which a disease spreads depends on its various natural characteristics, such as its infectivity or the duration of infectiousness, and the social distancing measures imposed by the government. Also, it depends on a plethora on socio-economic factors that govern the behavioral interactions within a population. Here we focus on one specific factor – the social connectivity – which quantifies the intensity of online social interactions within a community and study its relationship with the coronavirus outcome within the linear regression framework. The linear regression framework is the simplest tool used for quantifying the relationship be- tween a given outcome and a set of potential determinants. Its advantage lies in the efficient and unbiased analytical inference of the strength of the linear relationship. In addition it allows us to use powerful statistical techniques to determine the explanatory power of each independent variable. As such it has been widely used in modeling of epidemiological phenomena (Fogli and Veldkamp, 2012).

Our model is specified as

$$log(Cases_i) = b_0 + b_1 log(SocialConnectedness_i) + b_2 log(GDP_i) + b_3 Controls + u_i,$$

where the dependent variable is the log of the accumulated number of registered COVID-19 cases p.m.p. in country$i$since the first observed case, up until the threshold chosen as an end date in each country. The threshold, which will be discussed in the following, is different for each country and represents an estimate for the date at which the critical health condition caused by the first wave of COVID-19 ended.

To test our hypothesis for the potential effect of the social connectedness in the same country we use the PageRank of the Social Connectedness Index (SCI). SCI measures the strength of connectedness between two geographic areas as represented by Facebook friendship ties.

Formally it is quantified as$SCI_{ij} = \frac{FBConnections_{ij}}{FBUsers_i \times FBUsers_j}$,

where$FBConnecions_{ij}$is the total number of Facebook connections between i and j and $FBUsers_l$is the number of Facebook users in country $l$that were observed at the end of 2019. Thus, we end up with a weighted social network in which the countries represent nodes, and the edges are given by the number of Facebook friendships between people belonging to a pair of countries. The PageRank measures the centrality of a country in social interactions, with larger value implying that the country is socially more important (central) and has far more social mixing in comparison to other countries.

Since the social connectedness alone is not enough to explain the variations in the COVID-19 cases among the countries in each regression we include the level of economic development in the country. We measure this observable simply via the GDP per capita of the country in 2019 and expect it to have a significant impact on the coronavirus spread. Moreover, we include two additional control variables in the model. The first one is an aggregate stringency index for the country imposed social distancing measures during the coronavirus crisis. It is quantified as a weighted average from Oxford's daily COVID-19 stringency index, where the average is estimated from the day of the first registered coronavirus case in the country, up to the last date in which the daily index was at its maximum. We give earlier dates a larger weight in the overall stringency index. This is because, evidently, countries which imposed more stringent social distancing measures and at earlier dates, are also expected to have less cases p.m.p.. The second measure simply quantifies the duration of the crisis, and is estimated as the number of days since the first observed case in the country up to the last day at which the data for the number of COVID-19 cases p.m.p. in the country was gathered. This date coincides to the last day at which the daily stringency index was at its maximum. The threshold is chosen as a means to capture the moment when a country gains the ability to control and stabilize the propagation of the disease.

The data for the dependent variable are taken from Our World in Data coronavirus tracker[1]. The tracker offers daily coverage of country coronavirus statistics, by collecting data mainly from the European Centre for Disease Prevention and Control. As pointed out, the stringency data are taken from Oxford's COVID-19 tracker[2], and the SCI data is gathered from the Data For Good database[3]. To reduce the noise from the data we restrict to using only countries with population above 1 million thus ending up with a total of 106 observations. Table 1 Gives the summary statistics for all four variables whereas the list of countries can be found in (Stojkoski et al., 2020). In general, we observe that there is great variation between the countries in both the COVID-19 Cases and Social connectedness. This is also holds for the summary statistics of the control variables.

[1] www.ourworldindata.org/coronavirus

[2] www.covidtracker.bsg.ox.ac.uk

[3] www.dataforgood.fb.com,

Table 1: Summary statistics.

Variable	Mean	Std. Dev.	Max.	Min.
log(Cases)	5.27	2.06	9.61	0.75
log(SocialConnectedness)	-0.79	1.30	1.84	-4.05
log(Stringency)	-1.10	0.56	-0.13	-2.74
Days since 1st Case	70.91	25.15	145.00	29.00
log(GDPpc)	9.54	1.09	11.53	6.97

(Source: Authors' calculations)

4. RESULTS

The results from the regression analysis are reported in Table 2. In order for adequate comparison between the estimated coefficients, each variable was first transformed to its z-score and afterwards the regression was estimated.

Column (I) gives the results form the baseline regression analysis in which the log of COVID-19 cases p.m.p. is regressed on the log of the Social Connectedness of the country, the log of the GDP per capita and the control variables. We observe that both the social connectedness and the level of economic development have a significant positive impact on the extent to which the coronavirus disease spreads within a country. An increase of one standard deviation in the PageRank of SCI is associated with an increase in the number of observed cases by 0.22%, whereas a same increase in the GDP per capita is associated with an increase of 0.68%. The magnitude of the marginal effect of social connectedness is only 3 times less than the impact of the level of economic development, thus suggesting that a such networked structure may play an important role in the disease outcome.

In order to validate our results in Columns (II) and (III) of the same table we conduct two robustness checks. With the first robustness check we examine whether the specification of our model is correct by adding an additional explanatory variable in it. In particular, we include a dummy variable characterizing whether the country includes a population which is predominantly Catholic or not (over 60% of the population is catholic). We opted for this variable since recent observations suggest that countries with dominant catholic religion were also the ones that were most hit by the initial wave of the coronavirus disease. In addition, it is widely acknowledged that religion drives a person's attitudes towards cooperation, government, legal rules, markets, and thriftiness, see Guiso et al. (2003). The results, as given in Column (II) suggest that indeed catholic countries had more COVID-19 cases p.m.p.. More importantly, the statistical significance of the social connectedness and economic development persist, and their marginal effect remains of similar size.

Lastly, in Column (III), we substitute the log of GDP per capita with an alternate measure for the economic development of the country – the log of the average government expenditure in health per capita between 2010 and 2019. Clearly, more developed countries are also the ones which are able to spend more money on improving the national health. In the case, the Social connectedness remains a significant explanatory variable of the coronavirus outcome, though with a slightly smaller marginal effect.

Table 2: Regression results.

Dependent Variable: log(Cases)	**(I)**	**(II)**	**(III)**
log(SocialConnectedness)	0.22*	0.23*	0.16*
	(0.08)	(0.07)	(0.07)
log(GDP)	0.68*	0.62*	
	(0.08)	(0.07)	
IsCatholic		0.31*	
		(0.07)	
log(GovExpenditure)			0.69*
			(0.07)
R^2	0.41	0.51	0.45
Observations	106	106	106

Notes: Standard error in brackets; denotes significance at 5% level;All regressions include the two control variables described in Section 2.*

(Source: Authors' calculations)

5. DISCUSSION

Our results suggest there exists a robust and stable relationship between the level social importance of a country as well as the degree to which its people mix, and the outcome of the coronavirus health crisis. Countries which take a more central role in the network of social connections are also more susceptible to the coronavirus, whereas countries where there is less social mixing are less affected by the induced disease. A plentiful of reasons can be used as a possible interpretation for these results. For instance, it is known that in structured populations, the degree of epidemic spread scales inversely with the level of social mixing, Draief et al. (2006). This is because, everything else considered, in more sparse populations it is easier to identify and target the critical individuals that are susceptible to the disease as described in Kitsak et al. (2010). It often turns out, that these are exactly the individuals which are more socially connected, Pastor and Vespignani (2001).

In short summary, the presented findings are an empirical verification for standard theories, which suggest that social network structures play a critical role in disease spreading processes. These theories have been widely used, both theoretically and empirically as a means to develop coherent social distancing measures for preventing a fatal epidemic outcome.

REFERENCES

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., & Wong, A. (2018), “Social connectedness: Measurement, determinants, and effects”. *Journal of Economic Perspectives*, 32(3), 259-80.

Basnarkov, L. (2020), “Epidemic spreading model of COVID-19”. arXiv preprint arXiv:2005.11815.

Draief, M., Ganesh, A., & Massoulié, L. (2006, October), "Thresholds for virus spread on networks", In Proceedings of the 1st international conference on *Performance evaluation methodolgies and tools* (pp. 51-es).

Ehlert, A. (2020),“The socioeconomic determinants of COVID-19: A spatial analysis of German county level data”. medRxiv.
Fogli, A., & Veldkamp, L. (2012), “Germs, social networks and growth”, Working paper No. w18470, National Bureau of Economic Research.
Guiso, L., Sapienza, P., & Zingales, L. (2003), "People's opium? Religion and economic attitudes", *Journal of monetary economics*, 50(1), 225-282.
Jinjarak, Y., Ahmed, R., Nair-Desai, S., Xin, W., & Aizenman, J. (2020), “Accounting for Global COVID-19 Diffusion Patterns, January-April 2020". National Bureau of Economic Research.
Kitsak, M., Gallos, L. K., Havlin, S., Liljeros, F., Muchnik, L., Stanley, H. E., & Makse, H. A. (2010), "Identification of influential spreaders in complex networks", *Nature physics*, 6(11), 888-893.
Klepac, P., Kucharski, A. J., Conlan, A. J., Kissler, S., Tang, M., Fry, H., & Gog, J. R. (2020)., “Contacts in context: large-scale setting-specific social mixing matrices from the BBC Pandemic project.” medRxiv.
Kuchler, T., Russel, D., & Stroebel, J. (2020), “The geographic spread of COVID-19 correlates with structure of social networks as measured by Facebook”. National Bureau of Economic Research.
Pastor-Satorras, R., & Vespignani, A. (2001), "Epidemic spreading in scale-free networks", *Physical review letters*, 86(14), 3200.
Stojkoski, V., Utkovski, Z., Jolakoski, P., Tevdovski, D., & Kocarev, L. (2020), "The socio-economic determinants of the coronavirus disease (COVID-19) pandemic", arXiv preprint arXiv:2004.07947.
Szulczyk, K. R., & Cheema, M. A. (2020), “COVID-19: The impact of socio-economic characteristics on the fatality rate”, SSRN Electronic Journal, DOI: 10.2139/ssrn.3591551