

# Relationship between Coronavirus (COVID-19) and Social Factors: Evidence from 24 Countries Over the World During 88-Day Period from Jan 31 to April 27, 2020

Lien Nguyen-Phuong<sup>1</sup>, Hang Thu Thi Dang<sup>2</sup>, Nhung Phuong Thi Nguyen<sup>3</sup>

<sup>1</sup>A PhD and lecturer, Business Administration Department, Faculty of Business Administration, Van Lang University, Viet Nam. Address: 45 Nguyen Khac Nhu Street, Co Giang Ward, District 1, Ho Chi Minh City-70000, <sup>2</sup>A MBA and lecturer, Accounting-Finance Department, Faculty of Economics and Business, Hoa Sen University, Ho Chi Minh City, Viet Nam, <sup>3</sup>A PhD and lecturer, Business Administration Department, Faculty of Economics and Business, Hoa Sen University, Ho Chi Minh City, Viet Nam

## Abstract

The paper investigates the relationship among social and demographic factors and a new case of spreading coronavirus in 24 countries over 88-day period from Jan 31, 2020 to April 27, 2020. This study collected second data from World Health Organisation (WHO)'s and World Bank's database, applied Co-integration test and Causality test for a panel data set.

The research provides crucial finding: there is causality between new cases of spreading coronavirus and time to close school too late. Moreover, Asian region, where the studied countries are being increases, new case of spreading coronavirus decrease. Furthermore, classification of income of the research countries is statistically insignificant to new case equivalent to absent association between the changes in the new case of spreading coronavirus and classification of income of the research countries.

The spreading COVID – 19 can happen everywhere from high income, middle income to low-income countries. Another finding is that late closing school will enhance spreading COVID – 19 goes more than one thousand cases.

**Keywords:** COVID – 19, causality relationship and Asian region.

## Introduction

For years, people around the world have been facing many devastating disease relating to immunodeficiency syndrome such as HIV/AIDS, severe acute respiratory

syndrome coronavirus SARS-COV and coronavirus (COVID – 19)... Wherever the epidemic strikes, human and society are affected. In December of 2019, some cases of unknown diseases are reported in Wuhan City, Hubei Province of China. From December 31, 2019 to January 3, 2020, there was 44 cases of patients with pneumonia of unknown etiology<sup>1</sup>. Until now, updated to 29 April 2020, in global scale, it is detected exceeded 3 million cases. Specifically, European and American regions are both reported with highest infectious cases: 1.4 million cases and 1.2 million cases respectively. With the urgent cases of COVID – 19, there are many studies investigate the impact and treatment of this virus. Researchers have found that COVID – 19 has the same origin with former SARS – COV, MERS, specifically  $R_0$  – refers to the basic reproductive number of new cases for SARS in the late 2000is in the range between

---

### Corresponding Author:

#### Lien Nguyen-Phuong

A PhD and Lecturer, Business Administration Department, Faculty of Business Administration, Van Lang University, Viet Nam. Address: 45 Nguyen Khac Nhu Street, Co Giang Ward, District 1, Ho Chi Minh City-70000

e-mail: [nguyenphuonglien@vanlanguni.edu.vn](mailto:nguyenphuonglien@vanlanguni.edu.vn)

Phone: +84918604066

2–5, that is resemble with the range of the mean  $R_0$  for COVID – 19<sup>2</sup>. These papers focus on the medical nature, origin of COVID – 19 and how to treat COVID – 19 medically. The outbreak of COVID – 19 is an urgent issue therefore the research tends to investigate the linkage among new cases of COVID – 19 and social, demographic factors. Our paper primarily aims to trace the relationship of new case of spreading coronavirus and social & demographic factors such as time to close school too late, classification of income in 24 countries in four areas including European, Asian, American and African regions from January 31, 2020 to 27 April, 2020.

**Literature Review:** Contact networks models likely refers to interpersonal contacts then easily lead to diffusion diseases in a community. Specifically, contact networks may happen within schools, workplaces, hospitals, households, therefore people in community can easily transmit the diseases if they are infectious. A group of previous authors generated a stimulated contact networks for the small city in Vancouver, British Columbia for the dataset approximately 1000 households equivalent to 2600 people<sup>3</sup>. In general, the contact networks model shows the crucial of contact patterns to the pandemic, especially, with different contact networks may share resemble the basic reproductive number of new cases ( $R_0$ ), may have different epidemiological outbreaks results.

Regarding to medical geography, it is considered a spatial analytic method to evaluate the linkage between geographic determinant and illness people, in addition, can be used to identify the geographic processess affect pandemic dispersion. Traced back to the former cases to get involve in the sustantial role of spatial diffusion of deases, if we consider the impact of disease on global scale, the impacts of pandemic outbreaks in one part of the world will rapidly spread to the rest of the world. SARS outbreaks in Hong Kong and mainland China, then transmit rapidly to Toronto, the rest of the world. Gatrell<sup>4</sup> emphasized that diseases occur in specific location then by transportation or travel will enhance the diffusion. Whereby, group authors identified that if infectious people are barred at borders and if the Hong Kong as well as other continents did not leave their nation, there will not be a significant outcome as recognized until now<sup>5</sup>.

At the beginning with the influenza pandemic, there is not available vaccine or antiviral medicine to treat diseases, therefore, social distance with a community

is considered the vital method to protect itself. Some researchers indicated that this designed strategy is acknowledged effectively mitigate the diffusion of pandemic influenza without the use of vaccine<sup>6</sup>. Specifically, these authors at the first stage described how social contact network-focused mitigation can be constructed, then a stimulation model for the influenza outbreak, hence, built a model in which there is a population of 10.000 of a small US town, the result showed that children and adolescents are subject of the flu therefore keeping them at home will reduce the attack rate (90%), moreover, with higher infectious strains, tailored to specific communities such as quarantine, closing schools, social contact network, working from home, etc. are suitable local defenses against pandemic in the absence of vaccine or antiviral medicine<sup>6</sup>. Gumel et al.<sup>7</sup> indicated that reducing the contact rate between susceptible and diseased individuals by isolating, is critical importance in order to reduce the diffusion of SARS.

With respect of temperature, the outbreaks of SARS in the 2000s and the spread of COVID – 19 recently are caused by impact of weather. However, there are few studies investigate relation between diseases with region and why diseases refer outbreaks in countries with low temperatures. A few researchers used ecological analysis to demonstrate the significant relation between daily numbers of SARS infected people with environmental temperature, whereby, temperature in the range 16 – 28 Celsius degree is likely associated with SARS cases, which may increase the virus growth. It is proved that during the outbreak of SARS in Guangzhou, Hong Kong, Beijing and Taiyuan, the temperature subsequently approximately from 16.5°C to 25°C<sup>8</sup>.

## Material and Methodology

The data was extracted from World Health Organisation (WHO)'s database for the new cases variable and World Bank database for three remain variables: Time to close school to late, Income and Region from 24 countries in European, American, African and Asian regions. The study collects the panel data with the long time series characteristic. There are four key variables:

- New case of spreading coronavirus (NEWCASE)
- Time to close school too late (TCSCH). This is a dummy variable (early close (closing Schools before March, 2000) = 0 and too late (closing from

1<sup>st</sup> March) = 1). This indicator represents the level of social distance of a country.

- Classification of income of the research countries (INCOME). This is dummy variable (high income = 1 and the other = 0), which shows the one of demographic variables.
- Region where the studied country are being (REG). This is dummy variable (Asian region = 1, the other region = 0). This vector confirms the different climate, cultural and society in cross-country.

The reason for choosing this period because although the disease appeared at Wuhan, Hubei, China ever since December 2019 but the disease pandemic started on global scale from January 2020 to April 2020. Furthermore, the research period provides evidence of the long-data for investigating causality link between COVID – 19 and social-demographic factors in long-term. Besides, the reason for selecting these 24 countries due to available data daily. The table 1 also notes that the panel data has outlier so the paper has to deal with this phenomenon.

The research applies four – step method to analyze the relationships of Coronavirus and social & demographic factors as follows:

First, due to outlier data so we check the Gini coefficient to ensure the equality value of the frequency distribution data. Owing to the outlier characteristic of NEWCASE (see table 1) so the research conducts the

Gini coefficient correlation check that is appropriate with outlier phenomenon. Second, the research data is a panel with large T and small L, with outlier data so that we apply the Levin-Lin-Chu(LLC)<sup>9</sup> test (see table 3), which assumes that all panels have the same autoregressive parameter, that is, that  $\rho_i = \rho$  for all i. Then the alternative hypothesis is simply that  $\rho < 1$ . The hypotheses (H) are that  $H_0$ : Panels contain unit roots and  $H_a$ : Panels are stationary. The LLC test requires that the panels be strongly balanced. Third, we apply the Westerlund and Edgerton<sup>10</sup> cointegration test with hypotheses that  $H_0$ : No cointegration and  $H_a$ : Some panels are cointegrated. Rejection of  $H_0$  is thus substantial to validate an existence of co-integration given the entire the panel<sup>10</sup>. Finally, the study applies the causality test among all of the four variables to identify the direction of the relationship among COVID – 19 and other social factors following Ka and Mussard<sup>11</sup> for fixing outlier issue.

### Results

Before conducting the main results, we carry out the descriptive and correlation analyses. Table 1 reports the results of descriptive statistics. There are 2112 observations during 88- days research period from January 31, 2020 to April 27 2020. The mean of NEWCASE is 989.70 cases while the standard deviation is too large. The new case of spreading coronavirus gets a peak when the United States of America was on the epidemic top of 38509 case in April 26, 2020.

**Table 1: Description variables**

Variables	Observation	Mean	Std. Dev.	Min	Max
New Case	2112	989.70	3675.29	0	38509 (United States of America April 26)
TCSCH	2112	0.52	0.50	0	1
Income	2112	0.58	0.49	0	1
REG	2112	0.58	0.49	0	1

Source: WHO and World Bank’s database.

**Table 2: Unit root test results**

t-statistic	New Case		TCSCH		Income		REG	
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
Unadjusted t	-8.37	0.00***	-6.16	0.00***	-7.08	0.39	-8.50	0.00***
Adjusted t*	-4.46		-2.96		-0.29		-3.82	

Note: t statistics in parentheses: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

We check and know that the research data is a strongly balanced panel. The Gini coefficient correlation check shows that among independent variables (TCSCH, INCOME, REG) exist equality correlation excepts NEWCASE and among independent variable does not exist the correlation.

Table 2 confirms that almost variables stay, only variable “INCOME” is non-stationary. Next, we apply

the co-integration test following the Westerlund and Edgerton<sup>10</sup> test to check the co-integration between dependent stationary variable and INCOME variable that does not stay. Table 3 below confirm the appearance of co-integration between these two variables. The results indicate that to reduce the new case of Coronavirus, the Government should focus on the effects of social factors: time to close the school, the high-income countries, and the area, whether the country belongs to.

**Table 3: Results of cointegration test**

	New Case – Income with outlier		New Case – Income without outlier	
	Statistic	P-value	Statistic	P-value
Variance ratio	1.51	0.07*	3.21	0.00***

Before running the causality test we check the outlier and get the results for exception of outlier during running regression for the lest bias later.

Having examined the long-run equilibrium relationship, we turn to the causality test among all

considered variables, as shown in Table 4. The study conducts OLS and Gini regression with and without outlier to check robust of the model and get the results as seen as below.

**Table 4: Results of causality**

Variables	Gini with outlier		OLS with outlier		Gini without outlier		OLS without outlier	
	Coef Std. Err	z P>z	Coef Std. Err	t P> t	Coef Std. Err	z P>z	Coef Std. Err	t P> t
TCSCH	1984.51 (151.02)	13.14 (0.00***)	2552.05 (139.55)	18.29 (0.00***)	1051.29 (51.33)	20.48 (0.00***)	1252.34 (49.53)	25.29 (0.00***)
Income	371.53 (180.53)	2.06 (0.04**)	332.34 (1048.08)	0.32 (0.75)	215.20 (60.78)	3.54 (0.00***)	203.10 (329.33)	0.62 (0.54)
REG	-1834.38 (182.13)	-10.07 (0.00***)	-1933.04 (1048.32)	-10.07 (0.07*)	-803.95 (61.85)	-13.00 (0.00***)	-834.75 (329.55)	-2.53 (0.01***)
Constant	805.68 (207.33)	3.89 (0.00***)	589.42 (1159.49)	0.51 (0.61)	353.70 (69.86)	5.06 (0.00***)	273.57 (364.43)	0.75 (0.45)
No. Obs.	2112		2112		2073		2073	
R <sup>2</sup> /GR	0.80		0.13		0.75		0.24	
Gamma YYhat	0.73				0.76			
Gamma YhatY	0.58				0.57			
MSE			3429.3				1151.9	
Wald chi2			338.69 (0.00***)				647.07 (0.00***)	

Table 4 notes that both Gini and OLS regression with and without outlier points to the existence of direct effects on the new case of spreading coronavirus associated with time to close school too late. However, Gini regression

also provides evidence of the positive effect of INCOME on the NEWCASE. This means the INCOME causes a new case of spreading coronavirus, the COVID – 19 has a positive link with the high-income countries. Besides,

their negative causality between REG and NEWCASE. The difference of region except ASIA will lead to a new case of spreading coronavirus. The model of Gini without outlier provides the better robust than OLS and Gini with outlier<sup>11</sup>. All P-value and coefficients are statistically significant at 1% level except for INCOME variable. A positive coefficient of TCSCH indicates that the time to close school too late increases one day, the new case of spreading coronavirus goes up more than one thousand cases. Moreover, a negative coefficient of REG suggests that Asian region where the studied country are being increases, the new case of spreading corona virus decrease around eight hundred cases daily in Asia. This means the COVID – 19 pandemic is different across countries and social distance. Furthermore, INCOME is statistically insignificant to NEWCASE which means that there is no association between the changes in the new case of spreading coronavirus and classification of income of the research countries. Actually, the spreading COVID–19 can happen everywhere especially high income countries such as Italy, USA, Singapore or even low-income country such as Nepal. The result also supports and confirms the previous findings.<sup>6,7</sup>

### Discussion

The pandemic COVID–19 virus outbreaks recently need to be discussed urgently for its effect on every society and economy globally. Applying both co-integration test and causality test, the research result indicates that there is a bi-directional causality between new cases of spreading coronavirus and time to close the schools too late as well as region where the studied country are being. Close schools are important factor to avoid infection and protect the young generation. Furthermore, with the studied countries, located in Asian region, the new cases of COVID – 19 can be easily down to around eight hundred cases daily. Next, the classification of income of the research countries is statistically insignificant to new case of spreading coronavirus. Actually, from January 31, 2020 to April 27, 2020, the coronavirus has been spreading in global scale and in all classification of economics from high income to middle and low–income countries. Therefore, all governments should take highly consideration in social isolation in preventing the COVID – 19 in the absence of vaccine or antiviral medicine.

**Source of Funding:** The researchers do not get any research funding.

**Conflict of Interest:** We hereby confirm that the manuscript has no any actual or potential conflict of interest with any parties, including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work. Our paper has not been published previously, it is not under consideration for publication elsewhere.

**Ethical Clearance:** Taken research from secondary data through World Bank and WHO’s database following the COPE’s ethical committee.

### References

1. World Health Organization (WHO): Novel Coronavirus Situation Report - 1 21 January 2020; 1–7.
2. Wilder-Smith, A., Chiew, C. J., & Lee, V. J: Can we contain the COVID-19 outbreak with the same measures as for SARS? 2020;102–107.
3. Meyers, L. A., Pourbohloul, B., Newman, M. E. J., Skowronski, D. M., & Brunham, R. C: Network theory and SARS: Predicting outbreak diversity. 2005;71–81.
4. Gatrell, A. C: Complexity theory and geographies of health: A critical assessment. 2005;2661–2671.
5. Ruan, S., Wang, W., & Levin, S. A: The effect of global travel on the spread of SARS. 2006;205–218.
6. Glass, R. J., Glass, L. M., Beyeler, W. E., & Min, H. J: Targeted social distancing design for pandemic influenza. 2006;1671–1681.
7. Gumel, A. B., Ruan, S., Day, T., Watmough, J., Brauer, F., Van Den Driessche, P., ... Sahai, B. M: Modelling strategies for controlling SARS outbreaks. 2004;2223–2232.
8. Tan, J., Mu, L., Huang, J., Yu, S., Chen, B., & Yin, J: An initial investigation of the association between the SARS outbreak and weather: With the view of the environmental temperature and its variation. 2005;186–192.
9. Levin, A., C.-F. Lin, and C.-S. J. Chu: Unit root tests in panel data: Asymptotic and finite-sample properties. 2002;108, 1-24.
10. Westerlund, J., & Edgerton, D. L: Simple tests for cointegration in panels with structural breaks. 2008;665–704.
11. Ka, N., & Mussard, S:  $\ell_1$  regressions: Gini estimators for fixed effects panel data. 2016;1436–1446.