Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2022, 2022:11
https://doi.org/10.28919/cmbn/6991
ISSN: 2052-2541

STRUCTURAL EQUATION MODELING OF THE COVID-19 INCIDENCE RATE ASSOCIATED WITH THE DEATH RATE AND THE IMPACT OF SOCIOECONOMIC FACTORS IN ASEAN COUNTRIES

I.G.N.M. JAYA[1,*], A. CHADIDJAH[2], N. SUNENGSIH[2], R. ARISANTI[2]

[1]Department Statistics, Universitas Padjadjaran, Indonesia and Faculty of Spatial Sciences, Groningen University, The Netherlands

[2]Department Statistics, Universitas Padjadjaran, Indonesia

Abstract: The pandemic crisis caused by COVID-19 jeopardizes our progressive social, ecological, and economic development. Many studies have focused on its impact on sustainable development in emerging economies and rapid regional development. We attempted to create a study from a distinct point of view. We investigate how socioeconomic factors contribute to regional variation in the incidence and death rate of COVID-19 risk and attempt to develop statistical modeling with a focus on the Association of Southeast Asian Nations (ASEAN). We conducted quantitative research using structural equation modeling based on structure variance, also known as partial least square path modeling. Our findings suggest that socioeconomic factors have a significant impact on the incidence and mortality rate of COVID-19. COVID-19 infection rates are highest in developed countries in general. As discovered in this study, socioeconomic factors have a positive effect on the incidence rate but a negative effect on the death rate. The findings provide policy recommendations for controlling and preventing COVID-19 transmission and lowering the death rate in the ASEAN community. In general, the paper adds to our limited understanding of the socioeconomic consequences of the COVID-19 pandemic.

Keywords: COVID-19; partial least square; structural equation modeling; incidence and death rate; ASEAN.

2010 AMS Subject Classification: 93A30.

---

[*]Corresponding author
E-mail address: mindra@unpad.ac.id
Received November 13, 2021

1. INTRODUCTION

The Coronavirus disease (COVID-19) pandemic, which was discovered in Wuhan, China, in December 2019, continues to pose serious health and economic consequences worldwide [1-3]. By early August 2021, globally, over 220 million cases had been announced by WHO, resulting in over 4.5 million deaths [4]. Despite extraordinary public health interventions such as social distancing, contact tracing, and full and partial lockdown, the disease burden has risen, with significant variation between countries and regions and many countries experiencing multiple waves [1, 5]. The COVID-19 crisis exposed significant inequalities between countries worldwide [6]. As of 8 August 2021, the ASEAN countries had reported over 9.5 million cases of infection and over 212,730 deaths [7].

Understanding how pandemics differ in different countries will benefit in future pandemic response. Current research from high- and middle-income nations reveals that demographic characteristics, comorbidities, healthcare resources, and response rigor all contribute significantly to the incidence of COVID-19-related infections [1, 8, 9]. The effects of socioeconomic factors, on the other hand, remain unknown [10].

Liu et al. (2020) [10] and Zhang et al. (2021) [1] discovered an unexpectedly positive effect of the HDI on the confirmed rate and death rate. Higher HDI countries have more people with multiple chronic diseases, lower smoking rates, and a higher annual gross salary.

Despite the fact that the most developed countries have high numbers of cases, mortality rates in these countries are low. This could be a result of major geographical variations in the availability and accessibility of health care resources [8].

We developed a method for modeling structural equations (SEM) based on structure variance, also known as partial least squares path modeling (PLSPM). SEM is a type of statistical model that can be used to simultaneously model latent variables with complex relationships [11]. The SEM methodology is proposed in this study due to the high correlation between socioeconomic indicators such as income per capita, human development index, nurses and midwives, hospital bed availability, and population density. The socioeconomic indicators were obtained from the United Nation's report [12].

The paper is divided into four sections: Section 2 discusses the materials and techniques, Section 3 discusses the empirical findings, and Section 4 concludes with a discussion.

2. Material and Method

2.1 Material

Study area

The Association of Southeast Asian Nations, or ASEAN, is comprised of eleven member countries: Indonesia, Myanmar, Laos, Malaysia, the Philippines, Singapore, Thailand, Vietnam, Timor-Leste, Brunei, and Cambodia. ASEAN is home to approximately 8.5% of the global population [13]. Figure 1 shows the ASEAN country regions.

Figure 1. ASEAN country regions

COVID-19 Data

COVID-19 data were obtained from the Johns Hopkins University Center for Systems Science and Engineering (2021). The data can be found at https://github.com/datasets/COVID-19.git. The data was collected between 22 January 2020 and 25 August 2021. The information includes the number of confirmed cases and deaths. Figure 2(A-B) show the cumulative incidence and death rates for COVID-19 in ASEAN countries.

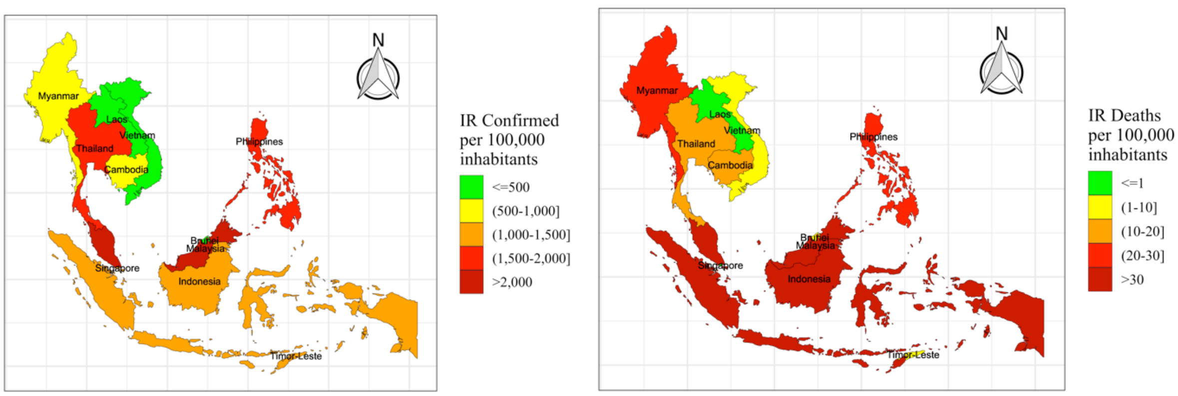

Figure 2. The cumulative incidence and death rates of COVID-19 in ASEAN countries

Figure 2 shows the cumulative incidence and death rates for COVID-19 in ASEAN countries. The spatial pattern of the death rate seems to not correspond to the incidence rate. To avoid the problem of underreporting, the data is transformed into weekly data. The data consists of the number of confirmed cases and deaths.

Socioeconomic Data

The United Nations Report 2020 provided socioeconomic data such as the number of people, population density, per capita income, human development index (HDI), nurses, midwives, and hospital beds. The data used in this study is presented in Table 1. The spatial patterns of the socioeconomic variables are presented in Figure 3.

Table 1 Socioeconomic data

Country	Incidence rate	Death rate	Density (/km^2)	Income per-capita (USD)	HDI	Nurses and midwives (/10,000 people)	Hospital bed (/10,000 people)
Myanmar	700.02	27.09	79	7,220	0.583	10	9
Laos	191.17	0.15	30	8,684	0.613	10	15
Malaysia	4,993.65	45.78	96	34,567	0.810	41	19
Philippines	1,718.44	29.65	356	10,094	0.718	2	10
Singapore	1,142.02	0.89	8,005	105,689	0.938	72	24
Thailand	1,579.32	14.45	135	21,361	0.777	30	21
Vietnam	391.79	9.60	288	8,677	0.704	14	26
Indonesia	1,472.21	47.27	141	14,841	0.718	21	12
Timor-Leste	1,160.15	3.94	85	5,321	0.606	17	59
Brunei	493.51	1.37	74	85,011	0.838	66	27
Cambodia	541.51	10.98	90	5,044	0.594	10	8

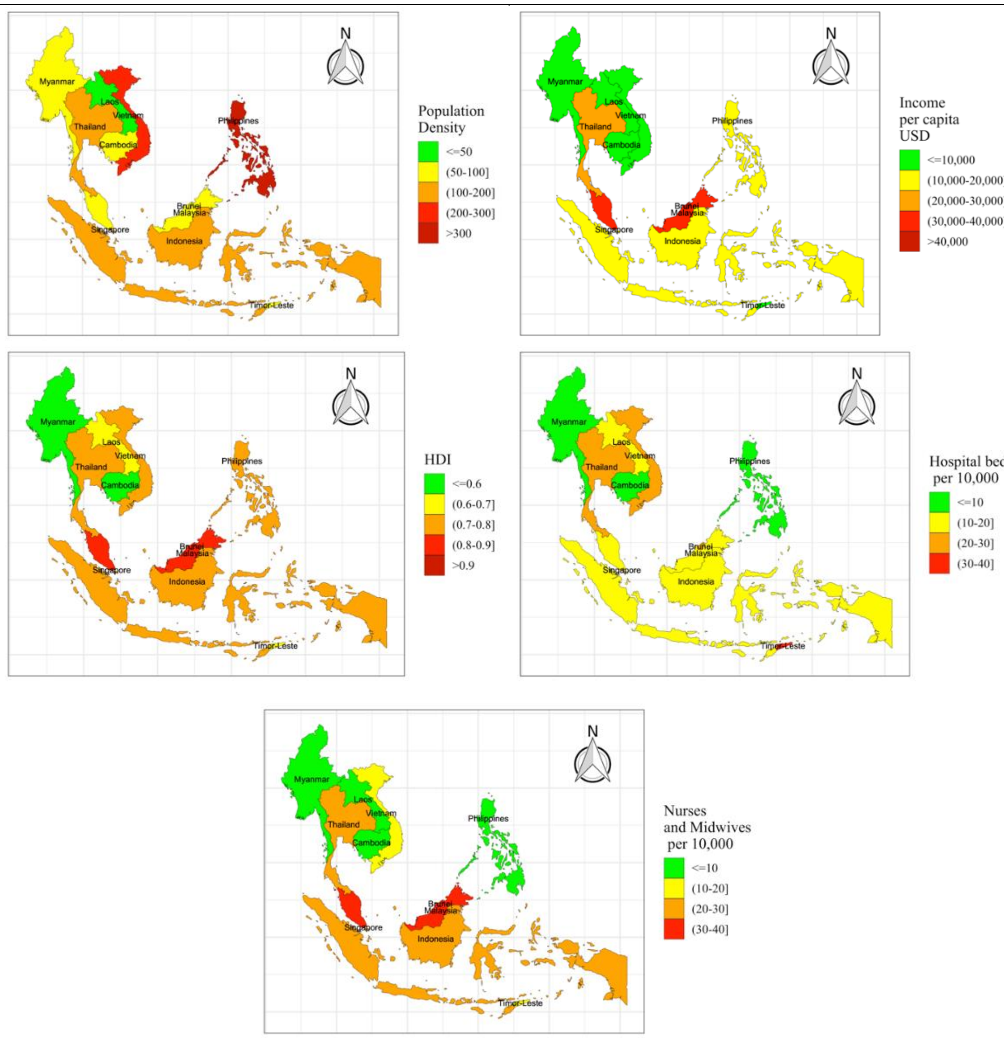

Figure 3. Spatial pattern of socioeconomic variables

2.2 Method

To model the socioeconomic factor on incidence and death rates at the same time, we use structural equation modeling (SEM) based on structural variance. Partial least square SEM (PLSPM) is a term used to describe the structure variance of SEM [14-16]. PLSPM is a strategy for causal modeling that prioritizes maximizing the variance explained by the dependent latent construct over generating a theoretical covariance matrix [15, 16]. PLSPM can be used to estimate the complex relationship between exogenous and endogenous variables without imposing stringent assumptions about normality or sample size [17]. When working with small samples and asymmetric distributions, PLSPM is particularly effective because it evaluates latent variable scores directly. It can also make factor identification easier by introducing a flexible residual covariance structure. It can also make accurate predictions in the presence of interdependent observations [16]. PLSPM consists of two models: (i) a measurement model (outer model) and (ii) a structural model (inner model) [17]. PLSPM specifically follows a four-stage process. The first stage is to define the structural model. The second stage is the selection and specification of measurement models. At the third stage, the model is estimated the parameters model [15, 18, 19].

Stage 1: Specifying the structural model

The following structure Figure 4 illustrates the research hypothesis:

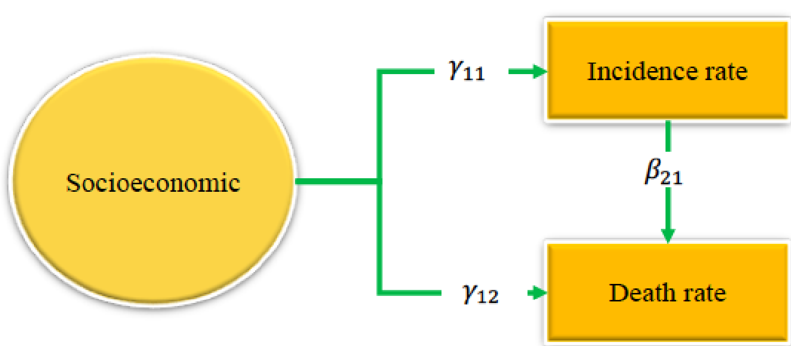

Figure 4. Path diagram of research hypotheses

Figure 4 can be written in statistical models as follow:

$$\eta_1 = \gamma_{11}\xi_1 + \varsigma_1$$
$$\eta_2 = \gamma_{12}\xi_1 + \beta_{21}\eta_1 + \varsigma_2 \qquad (1)$$

where η_1 denotes incidence rate, η_2 death rate, ξ_1 is socioeconomic factor with γ_{11} and γ_{12} denote the effect of socioeconomic factor on incidence and death rates respectively with β_{21} is the effect of incidence rate to death rate.

Stage 2: Selecting and specifying the measurement models

Socioeconomic factor consists of five indicators including income per capita (USD) (x_1), human development index (x_2), nurses and midwives (nurse/10,000 people) (x_3), hospital bed (bed/10,000 people) (x_4), and population density (inhabitants/km^2) (x_5). The path diagram of measurement model is presented in Figure 5.

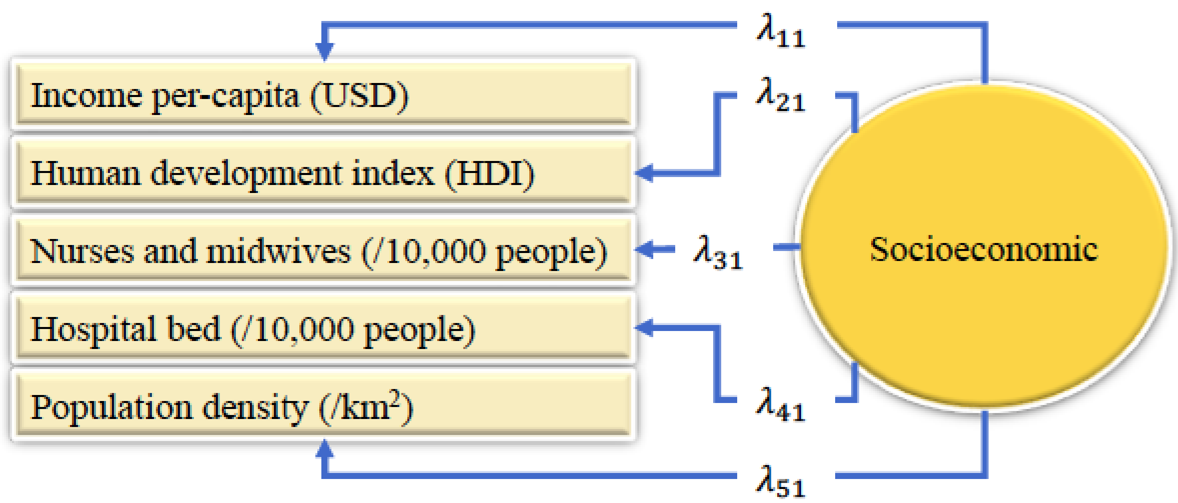

Figure 5. Path diagram of measurement models

The measurement models can be written as:

$$x_1 = \lambda_{11}\xi_1 + \varepsilon_1;\ x_2 = \lambda_{21}\xi_1 + \varepsilon_2;\ x_3 = \lambda_{31}\xi_1 + \varepsilon_3;\ x_4 = \lambda_{41}\xi_1 + \varepsilon_4 \text{ and}$$
$$x_5 = \lambda_{51}\xi_1 + \varepsilon_5 \qquad (2)$$

where $\lambda_{11}, \dots, \lambda_{51}$ denotes the standardized loading factors of $x_1,\dots,x_5$ respectively.

Stage 3: Model estimation and validation

The ordinary least square method is used to estimate the standardized loading factors and the regression coefficients as follows:

Reflective measurement:

$$X = \xi\lambda' + \varepsilon$$
$$\lambda' = (\xi'\xi)^{-1}\xi' X \qquad (3)$$

where $\boldsymbol{X} = (x_1, x_2, x_3, x_4, x_5)'$ and $\boldsymbol{\lambda} = (\lambda_{11}, \lambda_{21}, \lambda_{31}, \lambda_{41}, \lambda_{51})'$

Model structural

$$\beta = (\xi'\xi)^{-1}\xi'\eta$$

$$\eta = (\eta_1, \eta_2)'$$

$$\beta = (\gamma_{11}, \gamma_{21}, \beta_{21})' \tag{4}$$

Model validation includes item validity and reliability [20]. Convergent validity is typically used to determine item validity, while composite reliability and average variance extraction are typically used to determine reliability.

Convergent validity

Convergent validity quantifies the degree to which an indicator and its latent factor are correlated. If the indicator's loading factor is greater than 0.50, it has convergent validity [21].

Composite reliability ρ_c

Composite reliability (occasionally referred to as "construct reliability") is an approximate measure of internal consistency, similar to Cronbach's alpha [22]. It can be thought of as the ratio of the total variance in true scores to the total variance in scale scores [23].

$$\rho_c = \frac{(\sum_{k=1}^{K} \lambda_{k1})^2}{(\sum_{k=1}^{K} \lambda_{k1})^2 + \sum_{k=1}^{K}(1 - \lambda_{k1}^2)}$$

where K denotes the number of indicators that are used to measure the construct, with λ_{k1} denotes the kth standardized loading factors of first construct. Internal consistency is satisfied when the composite reliability of the items is greater than 0.70 [24].

Average variance extracted (AVE)

AVE is designed to measure the total variance of indicators that can be accounted by its latent factor (construct). The AVE is formulated as :

$$AVE = \frac{\sum_{k=1}^{K} \lambda_{k1}^2}{K}$$

where K denotes the number of indicators that are used to measure the construct, with λ_{k1} denotes the kth standardized loading factors of first construct. AVE is also used to measure internal reliability and discriminant validity [25]. Items meeting the A given construct should have a minimum reliability of 0.5 in order to be considered reliable [21]. The square root of AVE should be larger than the correlation between the constructs to satisfy the discriminant validity. We perform PLS-PM using the `plspm` package in R [26].

3. RESULT

3.1. Descriptive Statistics

Table 2 summarizes the descriptive statistics for the variables of the interest.

Table 2. Descriptive statistic of the main variables

Variables	Mean	STD	Minimum	Median	Maximum
Incidence COVID-19	870,862.000	1,246,999.000	2,159.000	380,879.000	4,026,837.000
Deaths COVID-19	19,339.000	37,799.000	6.000	9,349.000	129,293.000
Population at risk	60,783,622.000	80,563,937.000	437,479.000	32,365,999.000	273,523,615.000
Population density (/km^2)	852.636	2,374.197	30.000	96.000	8,005.000
Income per-capita (USD)	27,864.455	34,765.253	5,044.000	10,094.000	105,689.000
Human development index (HDI)	0.718	0.115	0.583	0.718	0.938
Nurses and midwives (/10,000 people)	26.636	23.551	2.000	17.000	72.000
Hospital bed (/10,000 people)	20.909	14.384	8.000	19.000	59.000

Multicollinearity checking

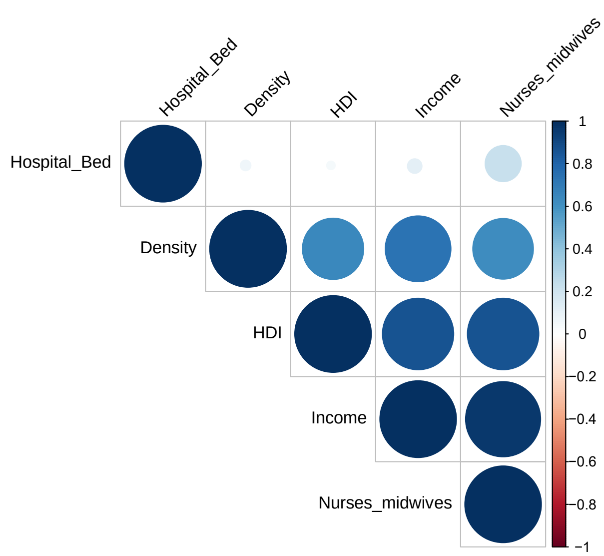

Figure 6. Pearson correlation matrix

The Pearson correlation coefficients for covariates other than hospital bed were greater than 0.70 (Figure 6), indicating a strong linear relationship between the socioeconomic indicators. As a result, several covariates have a VIF greater than 10 (Table 3), indicating a significant collinearity issue. We employ structural equation modeling to address this condition by developing socioeconomic factors that account for socioeconomic indicators.

Table 3. VIF of the covariates

Covariate	VIF
Population density (/km^2)	2.840
Income per-capita (USD)	24.634
Human development index (HDI)	4.782
Nurses and midwives (/10,000 people)	22.869
Hospital bed (/10,000 people)	1.347

3.2 Structural equation modeling of socioeconomic factor against incidence and death rates

Structural equation modeling is used to investigate the impact of socioeconomic factors on incidence and mortality. Five reflective indicators make up a socioeconomic factor. These indicators include the income per capita (USD), the human development index (HDI), the number of nurses and midwives (per 10,000 inhabitants), the number of hospital beds (per 10,000 inhabitants), and the population density (inhabitants/km^2).

Reliability and validity tests

Two approaches are used to evaluate the reliability and validity of the reflective indicators. First, determine the composite reliability [21], which must exceed 0.70 (or at least not be less than 0.60). Additionally, we present another measure of reliability, namely average variance extracted (AVE). The AVE must be greater than or equal to 0.500. AVE is also frequently used to assess discriminant validity. Second, examine the standardized loading factor to determine the item's validity for each indicator. It must be greater than or equal to 0.70 [21]. If one or more items are identified with a standardized loading factor of less than 0.6, the composite reliability is first evaluated. If the composite reliability value for the item is less than 0.600, the item should be excluded from the analysis. However, if the composite reliability value remains greater than 0.700, items with a value of less than 0.400 may be retained. As shown in Table 4, the reflective construction exhibits a relatively high level of internal reliability, with a composite reliability value of 0.897, an AVE value of 0.781, and standardized loading factors ranging from 0.404 to 0.948.

Table 4. Validity and reliability socioeconomic indicators

Indicator	Loading	Communality	Standard Error	q(0.025)	q(0.925)
Income per-capita (USD)	0.936	0.876	0.157	0.345	0.991
Human development index (HDI)	0.871	0.759	0.198	0.216	0.986
Nurses and midwives (/10,000 people)	0.948	0.899	0.181	0.288	0.990
Hospital bed (/10,000 people)	0.404	0.163	0.381	-0.515	0.942
Population density (/km^2)	0.744	0.554	0.506	-0.761	0.960
Composite reliability			0.897		
AVE			0.781		

Based on these results, all items are declared reliable and valid.

Table 5. Socioeconomic impact on incidence and death rates

	Estimate	Std. Error	t-value	p-value
Socioeconomic → Incidence rate	0.147	0.330	0.446	0.666
Socioeconomic → Death rate	-0.441	0.212	-2.080	0.071
Incidence → Death rate	0.741	0.212	3.490	0.008

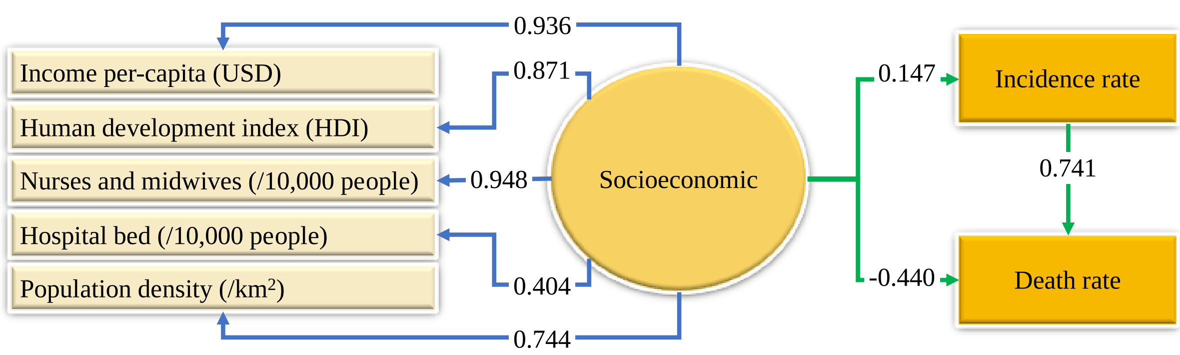

Figure 7. Socioeconomic factor on incidence and death rates

The estimation of structural equation model parameters is shown in Table 5, and a detailed path diagram with parameter estimation is shown in Figure 7.

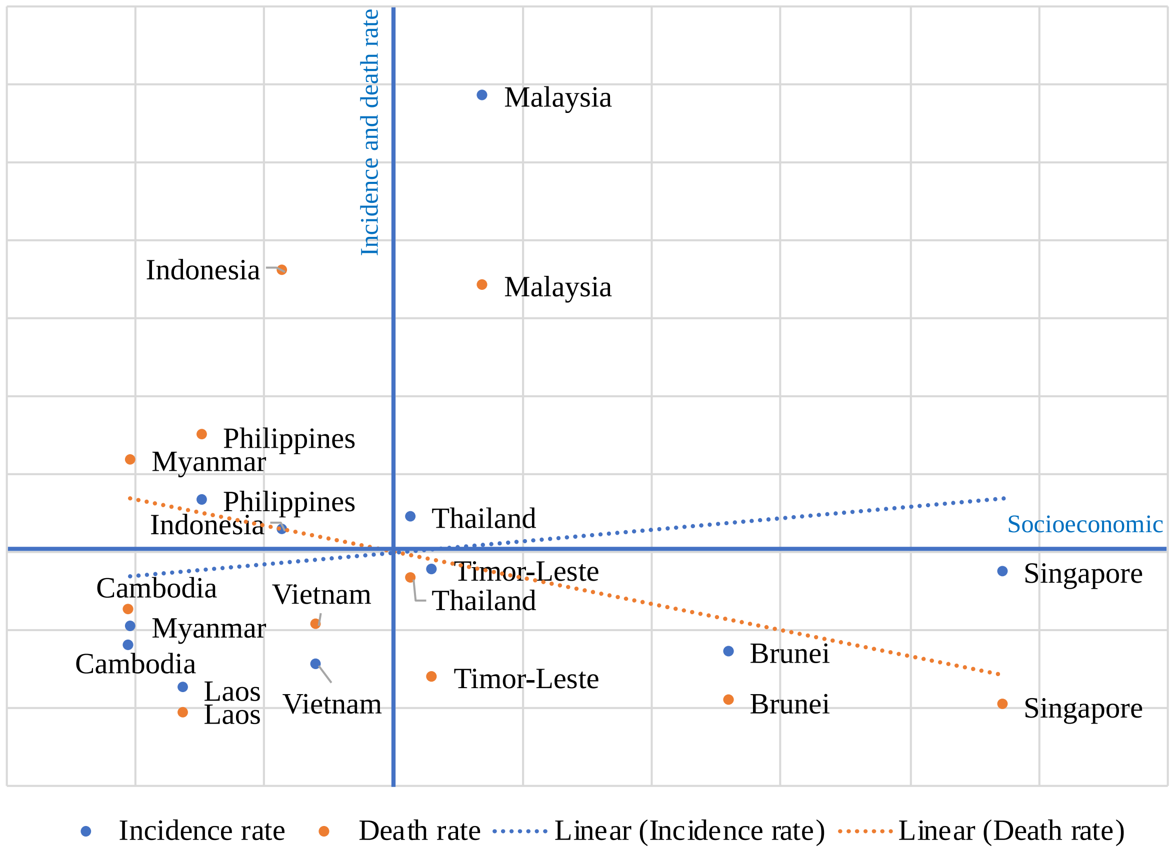

Figure 8. Socioeconomic factor against incidence and death rates by 08 August 2021

Given the information in Table 5, Figures 7-8, we know that socioeconomic status has a negative effect on death rates and a positive effect on incidence rates. The scatter plot demonstrates the existence of several distinct groups of countries. Vietnam, Laos, and Cambodia are low-income countries with low incidence and death rates. Indonesia and the Philippines are two countries with low socioeconomic status and high rates of disease and death. Myanmar is socioeconomically impoverished, with a low incidence rate but a high death rate. Timor-Leste, Brunei, and Singapore are the countries with the highest socioeconomic status and the lowest rates of disease and death. These countries' success in controlling incidences and deaths from COVID-19 is attributed to their superior health infrastructure and economic conditions. Malaysia is the ASEAN country with the highest socioeconomic score, as well as the highest incidence and death rate. Malaysia, despite its strong socio-economic conditions, has not been fully able to control COVID-19. Thailand also scores highly in terms of socioeconomic status, with a low death rate but a high incidence rate.

Finally, our findings demonstrate that socioeconomic factors have a positive impact on the incidence rate of COVID-19 while having a negative impact on the death rate of COVID-19. The confirmed rate has an impact on the death rate in a positive way. Additionally, we discovered that the most important indicators of socioeconomic factors are nurses and midwives, income per capita (USD), and the human development index (HDI). The findings of this study indicate that developed countries have a higher confirmed rate than developing countries or low-middle-income countries. However, developed countries have significant resources to prevent high mortality rates by providing very adequate health facilities.

4. Discussion and Conclusion

The COVID-19 pandemic is the century's most extraordinary health problem and the greatest threat to humanity since the Second World War [27]. It has a significant adverse effect on public health and has socioeconomic consequences. It is rapidly spreading throughout the world, as Wuhan, China reported an outbreak on 30 December 2019 [28]. On 1 March 2020, the WHO declared COVID-19 to be a pandemic. As of 20 August 2021, the coronavirus COVID-19 has been transmitted into over 200 countries and territories worldwide (Worldometer 2021). COVID-19 has had a negative impact on the economic structure of every country worldwide, including Southeast Asian countries. The slowdown in economic growth occurred in the second and third quarters of 2020. Almost all sectors of the economy experienced a slowdown in growth. The debt of countries is increasing as a result of the enormous socio-economic consequences of COVID-19. Developed countries, in general, have the highest rates of COVID-19 infection. This study discovered that socioeconomic factors have a positive effect on the incidence rate but a negative effect on the death rate. Income per capita and the number of nurses and midwives are the two most important indicators of socioeconomic factors. This is because developed countries have higher population densities and mobility, which allow for faster transmission of the casting virus. However, they have adequate health care facilities, which helps to reduce the death rate from COVID-19.

ACKNOWLEDGMENT

The authors would like to express their appreciation for the financial support provided by Universitas Padjadjaran. Additionally, we appreciate the substantial and constructive recommendation to enhance the paper's quality.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

[1] F. Zhang, H. Karamagi, N. Nsenga, et al. Predictors of COVID-19 epidemics in countries of the World Health Organization African Region, Nat. Med. 27 (2021), 2041–2047.

[2] E. Marhamah, I.G.N.M. Jaya, Modeling positive COVID-19 cases in Bandung City by means geographically weighted regression, Commun. Math. Biol. Neurosci. 2020 (2020), Article ID 77.

[3] R. Fitriani, I.G.N.M. Jaya, Spatial modeling of confirmed COVID-19 pandemic in East Java province by geographically weighted negative binomial regression, Commun. Math. Biol. Neurosci. 2020 (2020), Article ID 58

[4] WHO, Coronavirus disease (COVID-19) pandemic, WHO, 20 April 2021. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019?gclid=Cj0KCQiA4b2MBhD2ARIsAIrcB-TOsv8ghorom7tVqcaUH1PTpxsLd5IxCbxYBMCbM8IdwxIXeEKWzcoaAhSFEALw_wcB. [Accessed 02 May 2021].

[5] S. Hsiang, D. Allen, S. Annan-Phan, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic, Nature. 584 (2020), 262–267.

[6] I. Goldin, R. Muggah, COVID-19 is increasing multiple kinds of inequality. Here’s what we can do about it, 10 October 2020. [Online]. Available: https://www.weforum.org/agenda/2020/10/covid-19-is-increasing-multiple-kinds-of-inequality-here-s-what-we-can-do-about-it/. [Accessed 12 October 2021].

[7] CSSE, COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University, CSSE, 28 March 2021. [Online]. Available: https://github.com/CSSEGISandData/COVID-19. [Accessed 02 May 2021].

[8] Y. Ji, Z. Ma, M.P. Peppelenbosch, Q. Pan, Potential association between COVID-19 mortality and health-care resource availability, Lancet Glob. Health. 8 (2020), e480.

[9] Z. Zheng, F. Peng, B. Xu, et al. Risk factors of critical & mortal COVID-19 cases: A systematic literature review and meta-analysis, J. Infect. 81 (2020), e16–e25.

[10] K. Liu, M. He, Z. Zhuang, D. He, H. Li, Unexpected positive correlation between human development index and risk of infections and deaths of COVID-19 in Italy, One Health. 10 (2020), 100174.

[11] R. kline, Principles and practice of structural equation modeling, The Guilford Press, London, 2016.

[12] U. Nation, Policy brief: the impact of COVID-19 on South-East Asia, United Nation, Geneva, 2020.

[13] ASEAN, ASEAN, 09 October 2021. [Online]. Available: https://asean.org/about-us. [Accessed 09 October 2021].

[14] Y. Fan, J. Chen, G. Shirkey, R. John, S.R. Wu, H. Park, C. Shao, Applications of structural equation modeling (SEM) in ecological studies: an updated review, Ecol. Process. 5 (2016), 19.

[15] J.F. Hair, G.T.M. Hult, C.M. Ringle, M. Sarstedt, A primer on partial least squares structural equation modeling (PLS-SEM), SAGE Publication, Thousand Oaks, USA, 2017.

[16] N. Zeng, Y. Liu, P. Gong, M. Hertogh, M. König, Do right PLS and do PLS right: A critical review of the application of PLS-SEM in construction management research, Front. Eng. Manag. 8 (2021), 356–369.

[17] W.W. Chin, R.A. Peterson, S.P. Brown, Structural equation modeling in marketing: some practical reminders, J. Market. Theory Pract. 16 (2008), 287–298.

[18] A. Monecke, F. Leisch, semPLS: structural equation modeling using partial least squares, J. Stat. Softw. 48 (2012), 1-32.

[19] K.K.-K. Wong, Partial least squares structural equation modeling (PLS-SEM) techniques using SmartPLS, Market. Bull. 24 (2003), 1-32.

[20] D. Achjari, Partial least squares: another method of structural equation modeling analysis, J. Ekon. Bisnis Indonesia, 19 (2004), 238-248.

[21] C. Fornell, D. Larcker, Evaluating structural equation models with unobservable variables and measurement

error, J. Market. Res. 18 (1981), 39-50.

[22] R. Netemeyer, W. Bearden, S. Sharma, Scaling procedures issues and applications, SAGE Publication, London, 2013.

[23] M. Brunner, H. Sub, Analyzing the reliability of multidimensional measures: an example from intelligence research, Educ. Psychol. Measure. 65 (2005), 227-240.

[24] D. Barclay, C. Higgins, R. Thompson, The partial least squares (PLS) approach to causal modeling: personal computer adoption and use as an illustration, Technol. Stud. 2 (1995), 285-309.

[25] C. Fornell, F. L. Bookstein, Two structural equation models: LISREL and PLS applied to consumer exit-voice theory, J. Market. Res. 19 (1982), 440-452.

[26] G. Sanchez, PLS path modeling with R, CRAN, Berkeley, 2013.

[27] I. Chakraborty, P. Maity, COVID-19 outbreak: Migration, effects on society, global environment and prevention, Sci. Total Environ. 728 (2020), 138882.

[28] D. Kang, H. Choi, J.-H. Kim, J. Choi, Spatial epidemic dynamics of the COVID-19 outbreak in China, Int. J. Infect. Dis. 94 (2020), 96-102.