APAEA

# The Effect of Country-level Factors and Government Intervention on the Incidence of COVID-19

Anasuya Haldar[1] , Narayan Sethi[1] [a]

[1] Dept. of Humanities and Social Sciences, National Institute of Technology Rourkela, Odisha, India

## Asian Economics Letters

This study explores the effects of demographic, socio-economic, and public-response factors on the incidence of new COVID-19 cases for the 10 countries with the greatest number of confirmed cases. Results show that demographic factors and government policies are significant determinants of COVID-19. Socio-economic factors, such as GDP per-capita and the human development index, appear statistically insignificant. The findings are important for policymakers in their efforts to reduce the number of new cases.

## I. Introduction

In this study, we test whether demographic characteristics, socio-economic factors and public policy parameters are significant determinants of new COVID-19 cases. Our empirical framework follows a multivariate negative binomial regression model, covering the 10 countries having the highest number of confirmed cases of the virus per million people.

Several studies have shown that pre-existing health conditions and non-communicable diseases (NCDs) increase both the incidence of the virus and related mortality rate (Banik et al., 2020; Chan et al., 2020; Mathur & Rangamani, 2020). Singh and Misra (2020) show, in a meta-analysis from the pooled studies of China, USA, and Italy, that the severity of COVID-19 is associated with other non-communicable diseases.

Although COVID-19 affects all age groups, people above the age of 65 are more vulnerable to the infection than the younger population. However, obesity and smoking habits among younger population make them equally vulnerable to COVID-19 (WHO, 2020a). Regions with increased smoking habits are found to have increased COVID-19 cases even among younger population (Yu, 2020). Apart from the NCDs and age-related factors, government pharmaceutical and non-pharmaceutical policies also affect the incidence of COVID-19 (Hale et al., 2020; OECD, 2020). Allel *et al.* (2020) find that a delay in the government lockdown responses significantly affected the incidence rate ratios (IRR) of COVID-19 cases. Health care capacity of a country is another important determinant of pandemic preparedness (Chaudhry et al., 2020; Khan et al., 2020; Kraef et al., 2020; Mbunge, 2020). Health inequalities due to inadequate health capacity, directly affect vulnerable people (Bambra et al., 2020).

Inspired by the above findings, this study helps to understand the country-specific factors and government responses in the countries with the highest number of confirmed cases. The study period is divided into 2 parts: the first part covers March to May 2020 (complete/partial lockdown in most countries) and the second part covers June to September 2020 (lockdown relaxed in most countries). The study is carried out for top 10 countries which have the greatest number of active COVID-19 cases per million people, as of 30th September 2020. We examine the country-specific effects of social, demographic, and health-related risk factors, along with government measures to contain the spread of COVID-19, on the number of new COVID-19 cases, for the 2 different time periods distinguished by country lockdowns.

Our research focuses on the determinants of the incidence of COVID-19 and is inspired by the variation in the incidence of the virus, both within and across countries. The COVID-19 virus had spread to most parts of the world by March 2020. The World Health Organisation (WHO) declared Europe as the new epicentre of the virus on 13th March, 2020, and more than 1 million people were affected worldwide by 4th April 2020 (WHO, COVID-19 Dashboard, World Health Organization, 2020b). Persons with pre-existing health conditions were deemed to be the most vul-

nerable and have had the highest mortality rate due to COVID-19. This study provides an understanding of the gaps in government policies in the countries with the highest virus cases, by dividing the pandemic period into pre- and post-lockdown phases. Moreover, by emphasising the role of NCDs, age, population density, and human development index, we consider the demographic and socio-economic factors along with the government policies. This is one of our contributions to the literature particularly from a policy point of view given that policy design is dictated by understanding determinants of the virus over time. To-date, there is no study that covers all these aspects in the manner we do. We find that demographic factors and government policies influence the incidence of new cases while socio-economic factors have a limited role. In the next section, data and methodology are discussed, followed by empirical results and analysis in section III and, finally, conclusion and policy implications in section IV.

## II. Methodology and Data

### A. Methodology

The association between the country-level factors and the number of new COVID-19 cases can be studied by fitting multivariate negative binomial regression models for both the pre- and post-lockdown sub-samples. This helps us understand the most relevant factors in explaining the variation in the number of new cases. A negative binomial regression is used when the dependent variable is a count variable with non-negative integers. Negative binomial regression is a generalised Poisson regression that relaxes 'the variance equal to the mean' assumption made by the Poisson model.

In a negative binomial regression, the mean of the dependent variable is determined by the exposure time, $t$, and a set of regressor variables, and each regressor variable has the same length of observation time. The slope coefficients can be estimated with consistency provided that the conditional mean is specified correctly, and the standard errors obtained are robust to possible misspecification of the distribution. This is like linear regression, except that consistent estimates and robust inferences can be obtained even when the normality assumption does not hold (Cameron & Trivedi, 2013).

Due to large variation in the number of confirmed new COVID-19 cases per thousand people ($CC$), within countries and regions, the distribution of $CC$ is over-dispersed (OECD, 2020). In this case, the Poisson distribution is inadequate to fit the model as it has only one parameter, $\mu$, and also requires variance-mean equality assumption. Negative binomial regression is appropriate for over-dispersed count data, that is, when the conditional variance exceeds the conditional mean. The traditional negative binomial regression model is given by

$$\ln(\mu_{CC})_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_p x_{it} \ln(\mu_{CC})_i \quad (1)$$
$$= \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_p x_{it} + \alpha_i + \epsilon_{it}$$

where $\mu_{CC}$ is the conditional mean of the dependent variable $CC$; the predictor variables $x_{1t}, x_{2t}, \ldots, x_{it}$ are given; $\alpha_i$ is the individual-specific fixed effect term, $\epsilon_{it}$ is the

residual term and the population regression coefficients, $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ are to be estimated. The regression coefficients of a negative binomial distribution are interpreted as the difference between the logarithm of two consecutive expected counts, $x_0$ and $x_{0+1}$, and written as:

$$\beta = log(\mu_{x_{0+1}}) - log(\mu_{x_0}) \quad (2)$$

where $\beta$ is the regression coefficient, $\mu_{x_{0+1}}$ and $\mu_{x_0}$ are the expected count of the predictor variable, evaluated at $x_0$ and $x_{0+1}$, respectively; $x_{0+1}$ denotes an incremental change in the predictor variable, $x$. Alternatively, Equation (2) can be written as:

$$log(\mu_{x_{0+1}}) - log(\mu_{x_0}) = log(\mu_{x_{0+1}} / \mu_{x_0}) \quad (3)$$

This representation of the parameter estimates in Equation (3) gives the *IRR*, which is interpreted as the rate of change in the dependant variable for an incremental change in the predictor variable. The impact of an incremental change in the explanatory variable on the rate of occurrence of the dependant variable can be estimated by calculating the *IRR* for that variable as given in Equation (3). The sign of the coefficient shows whether the dependant variable increases or decreases with incremental changes in the predictor variable and the *IRR* shows the magnitude of the rate of change. We report both the estimated coefficients and the *IRR* values in order to infer the effect of country specific factors on the incidence of daily new COVID-19 cases.

### B. Data

In this study, the Government Stringency Index (*GSI*) and the number of testing done per 1000 people (*TEST*) proxy for government policy; the number of beds per 10,000 people (*BEDS*) and the number of physicians per 10,000 people (*PHY*) are used as proxies for health capacity; population density (*PD*), median age of the population (*POP*) and proportion of people with pre-existing health conditions (*NCDs*) are demographic factors; and GDP per capita (*GDP*) and Human Development Index (*HDI*) represent socioeconomic factors. These are the explanatory variables in this study. The dependent variable is the number of confirmed new COVID-19 cases (*CC*). Daily data for the variables are compiled from the John Hopkins University database and the Oxford University Online Repositories for the period March 15, 2020 to September 30, 2020. The countries analysed are India, the USA, Brazil, Argentina, France, Colombia, Russia, Israel, the UK, and Peru.

### III. Empirical Findings

The results of the coefficient estimate and the *IRR* values for the negative binomial regression, for both time periods, are given in Table 1 below.

Table 1 shows that the *IRR* estimates are statistically significant for all the variables in model 1. The results show that with a unit increase in *GSI, TEST, BEDS and PHY*, keeping other variables in the model constant, the *CC* would be expected to decrease by a factor of 0.673 (*GSI*), 0.184 (*TEST*), 0.782 (*BEDS*) and 0.175 (*PHY*).

Negative coefficient estimates for *GSI, TEST, BEDS and PHY* (Table 1) show that an increase in each of these vari-

**Table 1: Results of negative binomial regression for two different time periods**

| Independent Variables | Model 1 (March-15 to May-31, 2020) | | Model 2 (June-01 to September-30, 2020) | |
|---|---|---|---|---|
| | Coefficients | (SE) | Coefficients | (SE) |
| *GSI* | -0.024** | (0.38) | -0.086 | (0.35) |
| *TEST* | -0.012* | (0.11) | -0.124* | (0.11) |
| *BEDS* | -0.032* | (0.07) | -0.134* | (0.04) |
| *PHY* | -0.085* | (0.05) | -0.116* | (0.06) |
| *NCD* | 2.847* | (0.07) | 2.126** | (0.04) |
| *PD* | 0.125** | (0.16) | 0.126** | (0.05) |
| *AGE* | 3.132** | (0.15) | 3.076** | (0.11) |
| *HDI* | -0.005*** | (0.78) | -0.045 | (0.30) |
| *GDP* | -0.001*** | (0.14) | -0.008 | (0.14) |
| *Constant* | 8.764*** | (0.99) | 42.654*** | (4.59) |
| Pseudo R$^2$ | 7.3% | | 6.7% | |
| *Ln*(alpha), SE | 0.742** (0.08) | | 1.167*** (0.15) | |
| **Independent Variables** | **Model 1 (March-15 to May-31)** | | **Model 2 (June-01 to September-30)** | |
| | *IRR* | (SE) | *IRR* | (SE) |
| *GSI* | 0.673** | (0.39) | 0.867 | (0.36) |
| *TEST* | 0.184* | (0.10) | 0.168* | (0.11) |
| *BEDS* | 0.782* | (0.08) | 0.917* | (0.05) |
| *PHY* | 0.175* | (0.05) | 0.178* | (0.07) |
| *NCD* | 1.887* | (0.08) | 1.632** | (0.05) |
| *PD* | 1.145** | (0.15) | 1.750** | (0.05) |
| *AGE* | 7.896** | (0.16) | 1.076** | (0.10) |
| *HDI* | 0.009*** | (0.79) | 0.054 | (0.30) |
| *GDP* | 0.001*** | (0.13) | 0.080 | (0.15) |
| *Constant* | 9.876*** | (0.98) | 41.654*** | (4.60) |
| Pseudo R$^2$ | 3.8% | | 2.7% | |
| *Ln*(alpha), SE | 0.756** (0.08) | | 1.170*** (0.15) | |

This table shows the coefficient estimates and the *IRR* values for the negative binomial regression. The sign of the coefficient shows the direction of change and the value of *IRR* shows the magnitude of the rate of change in the dependent variable for incremental changes in the independent variables. The dependent variable is the number of confirmed new COVID-19 cases, *CC*. *IRR* is incidence rate ratio. All values in brackets are the standard errors (SE). Finally, *p < 0.1; **p < .05; ***p < .01 (two-tailed tests) show the level of significance at the 1%, 5% and 10% levels, respectively.

ables leads to fall in the rate of new COVID-19 cases; however, the magnitude of the decrease in the dependant variable, *CC*, for an incremental rise in the predictor variable is given by the *IRR* values. Similarly, for the demographic factors, for an incremental increase in *NCD*, *PD* and *AGE*, the dependant variable *CC* increases by 1.887, 1.145 and 7.896, respectively. The *IRR* estimates show that the median age of the population (*AGE*) is the largest contributor to the incidence of COVID-19. The *IRR* estimates for the socio-economic indicators, *GDP* and *HDI*, are significant at the 10% level, but the values are small. In Model 2, when lockdown restrictions were partially or completely relaxed, *GSI*, *HDI* and *GDP* became insignificant in both Models 1 and 2. This may be due to increased workforce participation in this period, especially those working in the informal sector. All other variables remain statistically significant.

## IV. Conclusion

This study aims to understand the determinants of COVID-19 confirmed cases using a range of government response and socio-economic variables. The study focusses on the 10 most affected countries. The findings of this study emphasise the role of demographic characteristics of the population as well as government stringency and testing policies as important factors in reducing the incidence of COVID-19. Consistent with previous studies, we also find support for dynamic government lockdown policies, with periodic lockdowns effective in controlling new cases. The study suggests that the best course of action for the countries with high COVID-19 cases will be to continue implementing periodic lockdowns, while increasing the number of COVID-19 testing. The study also suggests that these countries should also strengthen their healthcare capacity, proxied in this study by the number of physicians and the number of beds, in order to meet requirements of the vulnerable section of the population.

# References

Allel, K., Tapia-Muñoz, T., & Morris, W. (2020). *Country-level factors associated with the early spread of COVID-19 cases at 5, 10 and 15 days since the onset*. Global Public Health.

Bambra, C., Riordan, R., Ford, J., & Matthews, F. (2020). *The COVID-19 pandemic and health inequalities*. J Epidemiol Community Health.

Banik, A., Nag, T., Chowdhury, S. R., & Chatterjee, R. (2020). Why Do COVID-19 Fatality Rates Differ Across Countries? An Explorative Cross-country Study Based on Select Indicators. *Global Business Review*, *0972150920929897*.

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (Vol. 53). Cambridge university press. https://doi.org/10.1017/cbo9781139013567

Chan, E. Y. Y., Kim, J. H., Lo, E. S. K., Huang, Z., Hung, H., Hung, K. K. C., Wong, E. L. Y., Lee, E. K. P., Wong, M. C. S., & Wong, S. Y. S. (2020). What Happened to People with Non-Communicable Diseases during COVID-19: Implications of H-EDRM Policies. *International Journal of Environmental Research and Public Health*, *17*(15), 5588. https://doi.org/10.3390/ijerph17155588

Chaudhry, R., Dranitsaris, G., Mubashir, T., Bartoszko, J., & Riazi, S. (2020). A country level analysis measuring the impact of government actions, country preparedness and socioeconomic factors on COVID-19 mortality and related health outcomes. *EClinicalMedicine*, *25*, 100464. https://doi.org/10.1016/j.eclinm.2020.100464

Hale, T., Hale, A. J., Kira, B., Petherick, A., Phillips, T., Sridhar, D., & Angrist, N. (2020). Global Assessment of the Relationship between Government Response Measures and COVID-19 Deaths. *medRxiv*.

Khan, J. R., Awan, N., Islam, Md. M., & Muurlink, O. (2020). Healthcare Capacity, Health Expenditure, and Civil Society as Predictors of COVID-19 Case Fatalities: A Global Analysis. *Frontiers in Public Health*, *8*, 347. https://doi.org/10.3389/fpubh.2020.00347

Kraef, C., Juma, P., Kallestrup, P., Mucumbitsi, J., Ramaiya, K., & Yonga, G. (2020). The COVID-19 Pandemic and Non-communicable Diseases-A Wake-up Call for Primary Health Care System Strengthening in Sub-Saharan Africa. *Journal of Primary Care & Community Health*, *11*, 215013272094694. https://doi.org/10.1177/2150132720946948

Mathur, P., & Rangamani, S. (2020). COVID-19 and noncommunicable diseases: Identifying research priorities to strengthen public health response. *International Journal of Noncommunicable Diseases*, *5*(2), 76. https://doi.org/10.4103/jncd.jncd_33_20

Mbunge, E. (2020). Effects of COVID-19 in South African health system and society: An explanatory study. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*.

OECD. (2020). *The Territorial Impact of COVID-19: Managing the Crisis across Levels of Government*.

Singh, A. K., & Misra, A. (2020). Impact of COVID-19 and comorbidities on health and economics: Focus on developing countries and India. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*.

World Health Organization. (2020a). *Coronavirus disease (COVID-19)*.

World Health Organization. (2020b). *The impact of the COVID-19 pandemic on noncommunicable disease resources and services: Results of a rapid assessment*.

Yu, X. (2020). Risk interactions of coronavirus infection across age groups after the peak of COVID-19 epidemic. *medRxiv*.