The long-term dynamic relationship between communicable disease spread, economic prosperity, greenhouse gas emissions, and government health expenditures: preparing for COVID-19-like pandemics

Muhammad Jawad Sajid1 · Syed Abdul Rehman Khan1,2 · Yubo Sun1 · Zhang Yu2,3

Received: 11 August 2022 / Accepted: 31 October 2022 / Published online: 10 November 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
The spread of communicable diseases, such as COVID-19, has a detrimental effect on our socio-economic structure. In a dynamic log-run world, socio-economic and environmental factors interact to spread communicable diseases. We investigated the long-term interdependence of communicable disease spread, economic prosperity, greenhouse gas emissions, and government health expenditures in India’s densely populated economy using a variance error correction (VEC) approach. The VEC model was validated using stationarity, cointegration, autocorrelation, heteroscedasticity, and normality tests. Our impulse response and variance decomposition analyses revealed that economic prosperity (GNI) significantly impacts the spread of communicable diseases, greenhouse gas emissions, government health expenditures, and GNI. Current health expenditures can reduce the need for future increases, and the spread of communicable diseases is detrimental to economic growth. Developing economies should prioritize economic growth and health spending to combat pandemics. Simultaneously, the adverse effects of economic prosperity on environmental degradation should be mitigated through policy incentives.

Keywords Economy · Environment · Regression analysis · Time series analysis · Sustainability · Socio-economic factor

Introduction
The spread of contagious illnesses across a population is a natural geographical and temporal phenomenon that is critical to contemporary civilization (Perez and Dragicevic 2009). COVID-19 is one such ongoing contagious viral disease. The first case of coronavirus disease 2019 (COVID-19), caused by a new coronavirus (SARS-CoV-2), was reported in Wuhan, China, on December 1, 2019 (Liu et al. 2020a). The World Health Organization (WHO) proclaimed the COVID-19 outbreak a public health emergency of international concern (PHEIC) on January 30, 2020, and categorized COVID-19 as a pandemic disease on March 11, 2020 (Liu et al. 2020a). Now standing, the latest (June 20, 2022) World health organization figures indicate more than 539 million confirmed cases and 6.3 million mortalities from the COVID-19 pandemic. Besides the human health impact, the COVID-19 pandemic substantially negatively impacts global economies (Singh et al. 2020; Sajid and Gonzalez 2021).

Considering the immense global impact of COVID-19, many researchers have focused on understanding the different aspects of the pandemic. Specifically, the socio-economic literature on COVID-19 can be divided into two main categories. The first category of related literature focuses on the COVID-19 impact on the economy and environment, i.e., under this category, COVID-19 is an independent variable. In this regard, on the economic-side COVID-19 impact on logistics and supply chains (Inoue and Todo 2020; Singh et al. 2020; Khan and Ponce 2021), education (Sajid 2020), industry (Sajid 2021), financial markets (Zhang et al. 2020), etc., have been extensively studied. Moreover, the ecological...
effects of COVID-19 on sustainability (Barbier and Burgess 2020; Ibn-Mohammed et al. 2021), CO2 emissions (Sajid and Gonzalez 2021; Sajid et al. 2022), water consumption (Kalbush et al. 2020), air pollution (Wang et al. 2020; Acharya et al. 2021b; Shen et al. 2021; Gao et al. 2022), public awareness (Rousseau and Deschacht 2020), climate change (Forster et al. 2020), waste management (Vananalli et al. 2021), and energy consumption (Aktar et al. 2021; Han et al. 2021) are also thoroughly investigated.

The second category focuses on the effects of different socio-economic factors on the COVID-19 pandemic. That implies that under the second category, COVID-19 is a dependent variable. Table 1 shows an overview of the latest works on the socio-economic drivers of COVID-19. Meanwhile, Fig. 1 depicts the categorization of socio-economic research on the COVID-19 pandemic.

Despite current attention in contagious virus-related literature on the many socio-economic components of COVID-19, the long-term inter-association of contiguous viruses with socio-economic determinants is rarely studied. Furthermore, most relevant publications have not considered Sims’ (1980) critique of the endogeneity of all socio-economic variables, which states that the values of all socio-economic variables inside a model are interdependent in an economy with forward-looking economic actors. Aside from the abovementioned COVID-19 literature, Sims’ critique of economic variables endogeneity is usually overlooked by the general literature on the effects of socio-economic factors on various environmental domains (Khan et al. 2022; Sharif et al. 2017a, 2017b, 2019, 2020a, b; Khan et al. 2019; Suki et al. 2020; Khan et al. 2021; Godil et al. 2021). Finally, India, the world’s most densely and widely inhabited, large GHG emitter (India is the world’s third largest carbon emitter (Sajid et al. 2020)), often lacks a comparison analysis of the infectious virus’s socio-economic aspects.

Using a variance error correction (VEC) model, we predict the long-run dynamic inter-relationship between communicable and other disease-related fatalities (disease spread),1 economic prosperity (GNI),2 GHG emissions, and government health policy (expenditure). Also, we argue that, in an economy with forward-thinking agents, the dynamic long-term relationship between the spread of communicable diseases and other explanatory variables means that all variables must be endogenous. Our idea is directly derived from Sims’ critique (Sims 1980), which states that no variable can be considered exogenous in a world containing rational, forward-thinking agents. According to Sims, the prevalent economic models made substantial assumptions about the dynamic nature of the interaction between macro-economic factors. They are also largely inconsistent with the belief that economic agents consider the impact of today’s decisions on tomorrow’s utility. Sims offered “vector autoregression models (VARs)” as a solution that allowed one to model macroeconomic data in an informative manner while imposing few constraints. Finally, India’s climatic and socio-economic factors make it an excellent location for propagating and nourishing hazardous viruses. By examining the dynamic relationship between communicable disease spread, GHG emissions, economic prosperity (GNI), and health policy in an important developing country like India, we can be much better prepared to control and mitigate potential future contagious pandemics like COVID-19.

Conceptual framework of the proposed endogeneity model

Figure 2 depicts a conceptual framework of the proposed endogeneity model. Taking the lead from Sims’ critique, we propose that no socio-economic or environmental variable can be considered exogenous during the spread of contagious disease and related mortalities. Environmental, economic, and health decisions are made with the future in mind by forward-thinking economic agents.

- In particular, factor like economic progress and prosperity is considerably impacted by GHG emissions (Deke et al. 2001; Khan et al. 2021a; Batten 2018), health expenditure (Suhrcke et al. 2006; Oni 2014; Khan et al. 2022a; Piabuo and Tieguhong 2017; Yang 2020), and communicable diseases (like the COVID-19 pandemic) (Bagchi et al. 2020; Song and Zhou 2020; Mohsin et al. 2021; Priya et al. 2021).
- Simultaneously, GHG emissions are impacted by economic progress (Girod and De 2010; Enzler and Diekmann 2019; Baležentis et al. 2020; Bjelle et al. 2021), health expenditure (Wang et al. 2019; Bilgili et al. 2021; Khan et al. 2021b; Ganda 2021; Pervaiz et al. 2021), and communicable disease spread (Liu et al. 2020b; Sajid 2020; Sajid and Gonzalez 2021; Sajid et al. 2022).
- Similarly, the government’s health spending is influenced by communicable disease spread (Eissa 2020), GHG emissions (Ullah et al. 2019; Wang et al. 2019; Alimi et al. 2020), and economic progress and prosperity (Di 2005; Haseeb et al. 2019).

---

1 The term “communicable disease spread” is used in this article as abbreviation for the indicator “cause of death, by communicable diseases and maternal, prenatal, and nutrition conditions (% of total).”

2 While GDP is one of the most commonly used economic indicators, gross national income (GNI) may be a better indicator of a country’s overall economic progress if it includes significant foreign investments (Maverick 2018). As a result, for a developing country like India with significant foreign direct investment, GNI provides a more accurate estimate of economic progress than GDP.
| Study region | Socio-economic factors | Prime results | Reference |
|-------------|------------------------|---------------|-----------|
| Germany     | “Average age, the population density, and the share of people employed in elderly care, the share of schoolchildren, children in daycare, and physician density.” | Case and death statistics, for example, are strongly positively linked with early cases from the outbreak, average age, population density, and the proportion of individuals involved in senior care. In contrast, they are considerably inversely related to the ratio of school-aged children and children in daycare and physician density. | (Ehler 2021) |
| United States | “% of population aged older than 65 years, % of population Black, % of the uninsured population aged younger than 65 years, chronic kidney disease prevalence, COPD prevalence, heart disease prevalence, diabetes prevalence, obesity prevalence.” | More significant rates of COVID-19 illnesses and deaths are closely related to lower education levels and a higher proportion of the black population. | (Hawkins et al. 2020) |
| Global      | “Air quality, hazardous waste management, wastewater, health budget, industrial manufacturing, job loss, and unemployment.” | Although there has been a significant decrease in air pollutants (NO₂ and PM2.5) globally, high-air-polluted cities have shown substantial mortality in COVID-19-afflicted regions. The use of health safety equipment prevented transportation, and the work-from-home policy significantly influenced the amount of solid and hazardous waste disposal services. Wastewater was discovered to be another mechanism of enteric transmission of SARS-CoV-2. The epidemic had a significant socio-economic effect, exacerbating the nation’s economic burden. | (Suthar et al. 2021) |
| Italy       | “Aging Index, population density, density inside housing, % sports associations/resident population, unemployment rate, employment rate, % of active companies/resident population, local public transport use per capita, % of private long-term care beds, % of private acute care hospital beds, % of anesthesiologists/resident population, general practitioners rate, dentists rate, nursing staff rate, acute care hospital beds rate, long-term care hospital beds rate, emergency medical service rate.” | Demographic and socio-economic parameters and healthcare organization variables were found to be related to a substantial variation in the rate of COVID-19 dissemination in 36 Northern Italian regions. An elderly population seems to have a natural concentration of social relationships. The availability and coordination of healthcare services may have a significant role in the spread of illness. | (Buja et al. 2020) |
|             | “Distance from the first outbreak, value-added per employee, the intensity of export relationships, number of frost days in a year, mortality from infectious diseases, PM10, employment rate, percentage of employment in manufacturing, average family members, percentage of employment in agriculture.” | The findings of the initial estimate revealed the dominance of parameters linked to the intensity and interconnections of economic activity. The importance of these factors was significantly reduced in the second estimate, implying that the pandemic will be mitigated after the economic lockdown. Finally, by looking at cases at the start of the “second wave,” the investigators found that the epidemic’s geographical spread is linked to economic conditions. | (Bloise and Tancioni 2021) |
| Global      | “Population, density, age, GDP per capita, GDP nominal per capita, Gini, human development index, total fertility rate, Flights, Flights per capita.” | The correlation study revealed substantial connections between COVID-19 incidence and various measures, including GDP per capita and the number of flights per capita. Still, mortality is mainly associated with population age. | (Gangemi et al. 2020) |
| Global      | “Temperature, population, diseases, stringency, air pollution, income level, smoking testing.” | While ambient PM2.5, nitrogen dioxide, ozone, pressure, dew, wind gust, and wind speed accelerate the spread of COVID-19, high relative humidity and ambient temperature mitigate COVID-19, reducing the number of confirmed cases. Furthermore, the study’s findings revealed that various variables, including underlying health problems, meteorology, air pollution, health system quality, socio-economics, and demography, promote COVID-19 across nations. | (Sarkodie and Owusu 2021) |
| Study region       | Socio-economic factors                                                                                                                                                                                                                                                                                                                                 | Prime results                                                                                                                                                                                                                                                                                                                                 | Reference                                                                                                                                   |
|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Brazil            | “Proportion of people with low income, percentage of rented households, percentage of families with the social program, Gini index, running water”                                                                                                                                                                                                                           | The strongest explanatory power factors for the rise in COVID-19 infections were the percentage of persons with low income, the proportion of leased housing, the proportion of families participating in social programs, the Gini index, and access to running water. | (Silva et al. 2022)                                                                                                                            |
| Stockholm, Sweden | “Age, sex, living arrangements (nursing homes versus all others), and Mosaic groups”                                                                                                                                                                                                                                                                                                                                 | The elderly, nursing home residents and those from lower socio-economic classes are disproportionately affected by COVID-19 deaths. As a result of the epidemic, the number of dying nursing home residents admitted to acute hospitals has increased. | (Strang et al. 2020)                                                                                                                             |
| Global            | “Temperature, humidity, air quality, population density, GDP (PPP) per capita, per capita health expenditure, life expectancy and total testing (for COVID-19)”                                                                                                                                                                                                                                                               | The higher population density was an essential factor in the rapid spread of COVID-19, as maintaining social distancing criteria is more difficult in urban areas. Furthermore, while GDP (PPP) and PM2.5 are associated with fewer COVID-19 instances, PM10 and the total number of tests are significantly associated with increased COVID-19 case counts. | (Ahmed et al. 2021)                                                                                                                             |
| Italy             | “Atmospheric pollution”                                                                                                                                                                                                                                                                                                                                                                                               | The authors offered data that persons living in polluted areas are more likely to acquire chronic respiratory diseases susceptible to any infective agent. Furthermore, even in young and healthy people, persistent exposure to air pollution causes chronic inflammatory stimulation. | (Conticini et al. 2020)                                                                                                                           |
| Dublin, Ireland   | “Percentage of population aged 0–4, percentage of the population aged 5–14, percentage of the population aged 65 and over, percentage of a single population, percentage of house-share household, percentage with higher education degrees, percentage of professional social class, percentage of the unemployed population”                                                                                                                                               | According to the study, high-risk clusters had the lowest proportions of the elderly population, a high percentage of work and private rental income, and a high percentage of people aged 25 to 44. They also have the highest percentage of home ownership. | (Ghahramani and Pilla 2021)                                                                                                                      |
Finally, factors like economic development and prosperity (Suhrcke et al. 2011; Gangemi et al. 2020; Khan et al. 2022b; Acharya et al. 2021a; Silva et al. 2022), GHG emissions (Ahmed et al. 2021; Bloise and Tancioni 2021; Khan et al. 2022c; Sarkodie and Owusu 2021; Suthar et al. 2021), and health expenditure (Buja et al. 2020; Coccia 2021) affect communicable disease spread and related deaths.

Data sources and methods

Data sources

The data for all variables came from the World Bank’s “World Development Indicators.” The World Bank provides data on development indicators from 1960 to 2021 (The World Bank 2022). However, our study uses the time series data from 1981 to 2021 due to significant missing data. Furthermore, the missing data were filled using cubic spline interpolation. The cubic spline is one of the most widely used interpolation techniques for filling in missing data values (Prasad et al. 2018).

Methods

Simple multiple linear regressions demonstrating a linear relationship in levels between economic progress (GNI), GHG emissions, government health expenditure, and communicable and other disease spread (in terms of deaths), with all explanatory variables being exogenous, can be presented as:

\[ m_i = \alpha + \alpha_mc_i + \alpha_fh_i + \alpha_dd_i + \varepsilon_{m,i}, \]  
\[ c_i = \beta + \beta_mm_i + \beta_fh_i + \beta_dd_i + \varepsilon_{c,i}, \]  
\[ h_i = \gamma + \gamma_cm_i + \gamma_fh_i + \gamma_dd_i + \varepsilon_{h,i}, \]  
\[ d_i = \delta + \delta_cm_i + \delta_fh_i + \delta_dd_i + \varepsilon_{d,i}, \]

where \( m_i, c_i, h_i, \) and \( d_i \) respectively present the estimated values of economic progress, GHG (carbon) emissions, government health expenditure, and communicable disease spread for time series of sample size \( i(i = 1, 2, 3, \ldots, t) \) number of observations. \( \alpha, \beta, \gamma, \) and \( \delta \) represent the \( y \)-intercepts (constants) for respective equations. \( \alpha_m, \beta_m, \gamma_m, \) and \( \delta_m(k = m, c, h, d) \) coefficients of regression for Eqs. 1–4, respectively. And \( \varepsilon_{m,i}, \varepsilon_{c,i}, \varepsilon_{h,i}, \) and \( \varepsilon_{d,i} \) demonstrate the white noise errors associated with Eqs. 1–4. After developing the main structure of the linear relationships between the variables of concern under Eqs. 1–4, the exogenous variables (for forward-looking economic agents as per (Sims 1980)) can be endogenized using a VAR approach in the following manner:
A VEC model with endogenous variables for economic prospects of non-stationary I(1) series exist. The VECM model exerts a VAR of p lags (i.e., I(p)). The above VAR is only reliable if the relationship is cointegrated. As shown in Table 2, the null hypothesis of the unit root is rejected for all variables in the “none, intercept, and intercept plus trend” accounting equations under both the ADF and PP tests (except for the PP test for disease spread with no intercept and trend). Similarly, the null hypothesis of stationarity is rejected at various significance levels for all variables using the KPSS test. Therefore, all variables (in levels) are non-stationary.

### Cointegration analysis

A VAR for non-stationary variables is only possible if the I(1) or trend stationary series data. For the entire sample period from 1981 to 2022, India’s communicable and other disease-related deaths have been steadily declining. GNI has been steadily increasing since 2005. From 1981 to 2000, the government’s healthcare spending increased dramatically. However, the growth rate in India’s health expenditure begins to decline from 2001 until the end of the sample in 2022. India’s GHG emissions have rapidly increased over the last two decades. From 2019 onwards, there was some relief in the upward trend of India’s GHG emissions, possibly because of COVID-19-related economic activity restrictions. The descriptive statistics of the variables considered in our study are shown in Supplementary Table S1.

### Stationarity of the time series

For a VAR analysis, all variables must be stationary. The “Augmented Dickey-Fuller (ADF)” and “Phillips-Perron (PP) tests” are two of the main unit root tests that test a time series’ non-stationarity. Meanwhile, the “Kwiatkowski-Phillips-Schmidt-Shin (KPSS)” test is well-known for validating a time series’ stationarity assumption. Table 2 shows the results of the ADF, PP, and KPSS tests (in levels) on the variables Income, GHG emissions, health expenditure, and communicable disease spread. As shown in Table 2, the null hypothesis of the unit root is not rejected for all variables in the “none, intercept, and intercept plus trend” accounting equations under both the ADF and PP tests (except for the PP test for disease spread with no intercept and trend). Similarly, the null hypothesis of stationarity is rejected at various significance levels for all variables using the KPSS test. Therefore, all variables (in levels) are non-stationary.

### Trend analysis

Figure 3 depicts a temporal analysis of the spread of communicable diseases, income, health expenditure, and GHG emissions. As shown in Fig. 3, all variables show a clear trend, implying the possibility of non-stationarity in the time series data. For the entire sample period from 1981 to 2022, India’s communicable and other disease-related deaths have been steadily declining. GNI has been steadily increasing since 2005. From 1981 to 2000, the government’s healthcare spending increased dramatically. However, the growth rate in India’s health expenditure begins to decline from 2001 until the end of the sample in 2022. India’s GHG emissions have rapidly increased over the last two decades. From 2019 onwards, there was some relief in the upward trend of India’s GHG emissions, possibly because of COVID-19-related economic activity restrictions. The descriptive statistics of the variables considered in our study are shown in Supplementary Table S1.

### Results

For a VAR analysis, all variables must be stationary. The “Augmented Dickey-Fuller (ADF)” and “Phillips-Perron (PP) tests” are two of the main unit root tests that test a time series’ non-stationarity. Meanwhile, the “Kwiatkowski-Phillips-Schmidt-Shin (KPSS)” test is well-known for validating a time series’ stationarity assumption. Table 2 shows the results of the ADF, PP, and KPSS tests (in levels) on the variables Income, GHG emissions, health expenditure, and communicable disease spread. As shown in Table 2, the null hypothesis of the unit root is not rejected for all variables in the “none, intercept, and intercept plus trend” accounting equations under both the ADF and PP tests (except for the PP test for disease spread with no intercept and trend). Similarly, the null hypothesis of stationarity is rejected at various significance levels for all variables using the KPSS test. Therefore, all variables (in levels) are non-stationary.
running the cointegration test, one must first perform a reduced-form VAR analysis. Following a trial, a test basis four lags structure was chosen for the initial VAR analysis. Then, using three lags and the most appropriate trend assumption of “no deterministic trend,” Johansen’s cointegration test revealed a maximum of three cointegrating equations. Tables 3 and 4 detail Johansen’s cointegration test.

![Deaths by communicable and other diseases (% of total)](image1)

![GNI (current US$)](image2)

![Health expenditure per capita (current international $)](image3)

![Total GHG emissions (kt of CO2 equivalent)](image4)

**Fig. 3** A temporal analysis of the spread of communicable diseases, income, health expenditures, and greenhouse gas emissions. Here, y-axes represent values, while x-axes represent years

**Table 2** The results of the ADF, PP, and KPSS tests (in levels) on the variables Income, GHG emissions, health expenditure, and communicable disease spread

| Items | P-values |
|-------|----------|
| **H$_0$: Series has a unit root** | | **H$_0$: Series is stationary** |
| **ADF tests** | **PP tests** | **KPSS** |
| None | Intercept | Intercept and trend | None | Intercept | Intercept and trend | Intercept |
| Disease | 0.9927 | 0.9999 | 1.0000 | 0.0000 | 0.9817 | 0.6966 | 0.0000 | 0.05 < $P < 0.10$ |
| Health | 0.7001 | 0.1361 | 0.9920 | 0.7607 | 0.0595 | 0.9897 | 0.7607 | 0.05 < $P < 0.10$ |
| GHG | 0.9699 | 0.8888 | 0.0836 | 0.0000 | 0.9433 | 0.7727 | 0.7727 | 0.05 < $P < 0.10$ |
| GNI | 1.0000 | 1.0000 | 0.9998 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.01 < $P < 0.10$ |

**Residuals autocorrelation, normality, and heteroskedasticity test**

The autocorrelation LM test for our VEC model is provided in Table 5. Meanwhile, supplementary figure S1 depicts the residual autocorrelation correlograms. Based on Table 5, we can conclude that the null hypothesis of no serial correlation for lags 1 to 2 can be rejected at $p = 0.05$ using the
LM test. However, the null of no serial correlation cannot be dismissed beyond three lags. As a result, we will stick with the VEC with three lags. The Jarque–Bera residual normality test for our VEC model is depicted in Table 6. Supplementary tables 2 and 3 show the VECM’s Skewness and Kurtosis residuals normality tests. Table 6 shows that the null hypothesis of the VEC’s residuals’ normal distribution is not violated at $p = 0.1999$, implying that our VEC’s combined residuals are normally distributed. Furthermore, the first three factors’ individual probability distributions do not reject the null hypothesis at 99% confidence, whereas the factor accepts the null only at 90%. Finally, Table 7 shows the individual and joint heteroskedasticity tests for the residuals of our VEC model. Table 7 shows that the null hypothesis of residual homoscedasticity is not rejected at the individual and joint levels. As a result, we can state that our VECM is not affected by residual heteroskedasticity.

**Impulse response analysis**

The results of our impulse response analysis are depicted in Fig. 4. Meanwhile, the results of the VAR and VEC analyses are shown in supplementary tables S4 and 5. We are interested in the row-by-row analysis of impulse responses in this case. The first row depicts the impulse response of communicable disease spread to its own and other factors innovations. As can be seen, the most significant impulse response to disease spread is GNI shocks. The shocks in the values of communicable disease spread also had a considerable effect on itself, stabilizing towards the end. Meanwhile, Fig. 4’s second row depicts the response of GHG emissions to various variable shocks. The greatest response to GHG emissions was against GNI shocks. A minor downward trend in the impulse response of GHG emissions to disease-spread shocks was observed, which could be attributed

| Table 3 Results of the unrestricted cointegration rank test (trace) |
|---------------------------------------------------------------|
| Hypothesized no. of CE (s) | Eigenvalue | Trace statistic | 0.05 critical value | Probability** |
| None *           | 0.710816  | 86.84202       | 40.17493            | 0             |
| At most 1 *      | 0.439919  | 39.69575       | 24.27596            | 0.0003        |
| At most 2 *      | 0.37131   | 17.66813       | 12.3209             | 0.0058        |
| At most 3        | 0.000833  | 0.031666       | 4.129906            | 0.8843        |

The trace test indicates three cointegrating equations at the 0.05 significance level. The * denotes rejection of the hypothesis at the 0.05 significance level. The ** denotes MacKinnon-Haug-Michelis’s (1999) $P$-values.

| Table 4 Results of the unrestricted cointegration rank test (maximum eigenvalue) |
|---------------------------------------------------------------------------------|
| Hypothesized no. of CE (s) | Eigenvalue | Max-Eigen statistic | 0.05 critical value | Probability** |
| None *           | 0.710816  | 47.14627           | 24.15921            | 0             |
| At most 1 *      | 0.439919  | 22.02763           | 17.7973             | 0.0109        |
| At most 2 *      | 0.37131   | 17.63646           | 11.2248             | 0.0033        |
| At most 3        | 0.000833  | 0.031666           | 4.129906            | 0.8843        |

The Max-eigenvalue test indicates three cointegrating equations at the 0.05 significance level. The * denotes rejection of the hypothesis at the 0.05 significance level. The ** denotes MacKinnon-Haug-Michelis’s (1999) $P$-values.

| Table 5 The results of the VEC residual serial correlation LM test |
|---------------------------------------------------------------|
| Lag* | LRE** stat | df | Probability | Rao F-stat | df | Probability |
| 1    | 37.46276  | 16 | 0.0018      | 2.91013    | (16, 49.5) | 0.002 |
| 2    | 28.21427  | 16 | 0.0298      | 2.003672   | (16, 49.5) | 0.0318 |
| 3    | 23.78391  | 16 | 0.0943      | 1.619269   | (16, 49.5) | 0.0986 |
| 4    | 23.31606  | 16 | 0.1055      | 1.580407   | (16, 49.5) | 0.11  |

*Sample: 1981 2022; Included observations: 38
**Edgeworth expansion corrected likelihood ratio statistic
to the adverse effects of communicable disease spread on economic activity and population (as in the case of the COVID-19 pandemic). A slight upward impulse response trend was observed between the Indian government’s health expenditure and the innovations from GHG emissions. And a downward impulse response trend (that stabilized near the end) was observed between the Indian government’s health expenditure and health spending-related innovations. Finally, GNI only showed a relatively apparent upward impulse trend in response to communicable disease spread shocks.

Variance decomposition analysis

The variance decomposition analysis (VDA) can assist us in determining the sources (variables) that account for most of the variance in the variable of interest. Figure 5 depicts the VDA for communicable disease spread, GHG emissions, government health spending, and GNI. At shorter time horizons, most of the variation in the predicted value of communicable disease spread was caused by itself. However, as time passes, the GNI causes an increasing variation in communicable disease spread. The GNI is responsible for the greatest variance in communicable disease spread at lag 10. Again, most of the variation in the value of GHG emissions over a shorter time horizon was caused by itself. However, at the tenth lag, the majority of the variation in GHG emissions was due to variations in GNI. The largest variation in health expenditure was caused by itself. Other variables’ contributions increased as time horizons lengthened, but health expenditure remained the primary variance source. Finally, the most significant source of GNI variation was itself. It was followed by the variance contribution from communicable disease spread.

Policy implications

According to our trend analysis, India’s communicable disease spread rate is constantly decreasing while the GNI is increasing, which is a positive sign. Since 2000, the Indian government’s per capita health expenditure has decreased. Despite the COVID-19-related respite (from 2019 to 2022), India’s total GHG emissions have risen dramatically. The Indian government should improve on its already encouraging communicable disease spread and GNI figures. Furthermore, to better combat any future COVID-19-like pandemics, the government should urgently increase its declining share of per capita health expenditure. Simultaneously, the alarming rate of increase in GHG emissions should be

Table 6 The results of VEC analysis’s residual normality test

| Component     | Jarque–Beraa | df  | Probability |
|---------------|--------------|-----|-------------|
| Δlog(dₜ)      | 1.357173     | 2   | 0.5073      |
| Δlog(eₜ)      | 0.054317     | 2   | 0.9732      |
| Δlog(hₜ)      | 2.48966      | 2   | 0.288       |
| Δlog(mₜ)      | 7.146469     | 2   | 0.0281      |
| Joint         | 11.04762     | 8   | 0.199       |

aNull hypothesis: Residuals are multivariate normal. Orthogonalization: Cholesky (Lutkepohl)

Table 7 The results of VEC analysis’s residual heteroskedasticity tests (levels and squares)

|                      | R-squared | F(30.7) | Probability | Chi-sq(30) | Probability |
|----------------------|-----------|---------|-------------|------------|-------------|
| Individual components |           |         |             |            |             |
| res₁*res₁            | 0.966952  | 6.827043| 0.0068      | 36.74417   | 0.1847      |
| res₂*res₂            | 0.848677  | 1.308623| 0.3787      | 32.24973   | 0.356       |
| res₃*res₃            | 0.902565  | 2.161414| 0.1465      | 34.29745   | 0.2691      |
| res₄*res₄            | 0.706929  | 0.562834| 0.871       | 26.86332   | 0.6304      |
| res₂*res₁            | 0.856973  | 1.398064| 0.3403      | 32.56499   | 0.3417      |
| res₃*res₁            | 0.864836  | 1.492966| 0.3043      | 32.86377   | 0.3285      |
| res₃*res₂            | 0.893242  | 1.952303| 0.182       | 33.94321   | 0.2831      |
| res₄*res₁            | 0.704832  | 0.557176| 0.8749      | 26.7836    | 0.6346      |
| res₄*res₂            | 0.86343   | 1.475186| 0.3107      | 32.81032   | 0.3308      |
| res₄*res₃            | 0.872814  | 1.601251| 0.2684      | 33.16693   | 0.3153      |
| Joint test:          |           |         |             |            |             |
| Chi-sq              | 313.5572  |         |             |            |             |
| df                  | 300       |         |             |            |             |
| Probability         | 0.2835    |         |             |            |             |

The following residuals above are associated with the following variables from the VEC. Δlog(dₜ) = res₁, Δlog(eₜ) = res₂, Δlog(hₜ) = res₃, and Δlog(mₜ) = res₄
Fig. 4  The results from VECM-based impulse response analysis

Fig. 5  A variance decomposition analysis of communicable disease spread, GHG emissions, government health spending, and GNI
slowed by implementing a sticks and carrots approach, i.e., rewarding cleaner production while discouraging carbon-intensive production through carbon taxes and markets.

Secondly, our VEC-based impulse response analysis revealed that:

1) The spread of communicable diseases is highly sensitive to GNI shocks. That is, increased economic prosperity will have a negative impact on communicable disease spread (and related mortalities) in developing countries in general, specifically in India. Not only that, but economic prosperity can also help reduce disease spread and related mortality intensities from COVID-19-like contagious future pandemics.

2) But there is a catch to economic prosperity. Economic activity growth has a significant positive impact on GHG emissions. According to our impulse response analysis, total GHG emissions are highly positively sensitive to GNI innovations. As a result, economic progress must be environmentally sustainable. To achieve climate-resilient growth, developing-world governments, particularly the Indian government, should take all innovative measures, such as shifting from linear to circular economies, developing and importing climate-friendly technology, carbon capture and storage, and so on. That can ensure that the positive socio-economic benefits (such as the reduction in communicable disease spread) of economic growth can be realized without imposing additional climate-related costs.

3) In response to shocks in other variables, health expenditure did not show distinct upward or downward slopes. However, a slight upward trend in health expenditure was observed in response to GHG emissions innovations. Furthermore, a downward trend was observed between health expenditure and its shocks, which stabilized near the end. That implies that reducing GHG emissions may ease the government’s healthcare spending. Furthermore, increasing the government’s current health expenditure may (logically) necessitate less health spending in the future. This is because current health spending means more health facilities to serve a growing population, less disease spread, better maintained current health infrastructure, and so on.

4) Finally, except for communicable disease spread, GNI showed no clear upward or downward impulse trend in response to different variable shocks. As shown in Fig. 2, India’s communicable disease spread has decreased since 1981, while its GNI has increased. As a result, the positive response of GNI to shocks in communicable disease spread implies that a reduction in disease spread can positively impact an economy’s economic prosperity. As a result, the abovementioned preventive measures should be implemented immediately to reduce the massive negative economic impacts of COVID-19-like pandemics.

Aside from the impulse response analysis, which relied heavily on relatively distinct trends, variance decomposition analysis revealed the significant sources of variance in the values of various variables. Aside from the above policy-related inferences from impulse response analysis, variance decomposition revealed that most of the GNI’s variance was caused by itself. That makes sense because it implies that past income significantly impacts current income. In other words, today’s economic prosperity affects tomorrow’s economic well-being. Furthermore, while GNI innovations did not show a clear response trend in health expenditure, according to variance decomposition analysis, it was the second most influential factor in health expenditure variance. In conclusion, the role of economic prosperity (GNI) in health policy, disease prevention, and GHG emission reduction cannot be overlooked.

Conclusions

Economic prosperity (GNI), health, climate change (GHG emissions), and the spread of contagious diseases (such as the COVID-19 pandemic) all have an impact on the social fabric of human life. We developed the theory of endogeneity of all these critical socio-economic and environmental factors using the Sims’ critique (1980) of forward-looking economic agents. We used VAR and VEC models to investigate the interconnected effects of income, health, GHG emissions, and communicable disease in one of the world’s most carbon-emitting, densely and widely populated, rapidly developing economies, India. Our findings revealed that the number of Indian communicable and disease-related deaths has historically decreased. GNI and GHG emissions have increased overall, while government health spending initially increased and then declined from 2000 onwards. The unit root and stationarity tests (including intercept, intercept, and trend, and none) revealed non-stationarity (with stochastic trends), indicating that VAR was inappropriate. The cointegration test revealed three cointegrating equations at the 0.05 level. Hence, we used VECM analysis in our study. The residuals of VEC were checked for no autocorrelation, normality, and homoscedasticity using the “serial correlation LM test (no serial correlation for lags 1 to 2 rejected at \(p = 0.05\))”, Jarque–Bera Normality Test \((p = 0.1999)\), and “heteroskedasticity” tests (homoskedasticity not rejected at \(p = 0.2835\)). According to the impulse response analysis, GNI shocks had the greatest impact on impulse responses from communicable disease spread and GHG emissions. That was supported by the Variance decomposition
analysis, which found that GNI was primarily responsible for the variance in the values of these two variables at longer time horizons. Only a relatively distinct impulse response trend was observed for GNI against the innovations of communicable disease-related deaths. Furthermore, while there was no clear upward or downward impulse trend for GNI and health expenditure in response to GNI shocks, the GNI significantly contributed to the variances in health expenditure and GNI. Based on our findings, we propose that developing countries such as India focus on improving economic prosperity, which significantly impacts health policy and communicable disease spread. However, various policy mechanisms should mitigate the significant impact of increased economic prosperity on GHG emissions.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s11356-022-23984-9.

**Author contribution** Muhammad Jawad Sajid contributed to the study conception, design, software, formal analysis, investigation, supervision, administration, and writing—original draft. Material preparation, data collection, and validation were performed by Syed Abdul Rehman Khan. Yubo Sun and Zhang Yu worked on the writing—reviewing and editing. All authors read and approved the final manuscript.

**Funding** This research is supported by the National Natural Science Foundation of China (Grant number: 72250410375); the 2020 Major Basic Research Project of the Natural Science Foundation of the Jiangsu Higher Education Institutions, Grant number: 20KJA120003.

**Data availability** The socioeconomic data related to various factors is available freely from the World Bank’s data bank (https://databank.worldbank.org/source/world-development-indicators).

**Declarations**

**Ethical approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

**References**

Acharya A, Wolfson C, Matta S et al (2021a) The role of public health expenditures in COVID-19 control: evidence from local governments in England. SSM Popul Health 15:100861. https://doi.org/10.1016/j.ssmph.2021.100861

Acharya P, Barik G, Gayen BK et al (2021b) Revisiting the levels of Aerosol Optical Depth in south-southeast Asia, Europe and USA amid the COVID-19 pandemic using satellite observations. Environ Res 193:110514. https://doi.org/10.1016/j.envres.2020.110514

Ahmed J, Jaman MdH, Saha G, Ghosh P (2021) Effect of environmental and socio-economic factors on the spreading of COVID-19 at 70 cities/provinces. Heliyon 7:e06979. https://doi.org/10.1016/j.heliyon.2021.e06979

Aktar A, Alam MdM, Al-Amin AQ (2021) Global economic crisis, energy use, CO2 emissions, and policy roadmap amid COVID-19. Sustain Prod Consum 26:770–781. https://doi.org/10.1016/j.spc.2020.12.029

Alimi OY, Ajide KB, Isola WA (2020) Environmental quality and health expenditure in ECOWAS. Environ Dev Sustain 22:5105–5127. https://doi.org/10.1007/s11356-019-00416-2

Bagchi B, Chatterjee S, Ghosh R, Dandapat D (2020) Impact of COVID-19 on global economy. In: Coronavirus outbreak and the great lockdown. SpringerBriefs in Economics. Springer, Singapore, pp 15–26. https://doi.org/10.1007/978-981-15-7782-6_3

Baležentis T, Liobikienė G, Štreimikienė D, Sun K (2020) The impact of income inequality on consumption-based greenhouse gas emissions at the global level: a partially linear approach. J Environ Manage 267:110635. https://doi.org/10.1016/j.jenvman.2020.110635

Barbier EB, Burgess JC (2020) Sustainability and development after COVID-19. World Dev 135:105082. https://doi.org/10.1016/j.worlddev.2020.105082

Batten S (2018) Climate change and the macro-economy: a critical review. Staff working paper No. 706 Bank of England. https://www.bankofengland.co.uk/working-paper/2018/climate-change-and-the-macro-economy-a-criticalreview. Accessed 20 Sept 2022

Bilgili F, Kuşkaya S, Khan M et al (2021) The roles of economic growth and health expenditure on CO2 emissions in selected Asian countries: a quantile regression model approach. Environ Sci Pollut Res 28:44949–44972. https://doi.org/10.1007/s11356-021-13639-6

Bjelle EL, Wiebe KS, Tøbben J et al (2021) Future changes in consumption: the income effect on greenhouse gas emissions. Energy Econ 95:105114. https://doi.org/10.1016/j.eneco.2021.105114

Bloise F, Tancioni M (2021) Predicting the spread of COVID-19 in Italy using machine learning: do socio-economic factors matter? Struct Chang Econ Dyn 56:310–329. https://doi.org/10.1016/j.strueco.2021.01.001

Buja A, Paganini M, Coccio S et al (2020) Demographic and socioeconomic factors, and healthcare resource indicators associated with the rapid spread of COVID-19 in Northern Italy: an ecological study. PLoS One. https://doi.org/10.1371/journal.pone.0244535

Coccia M (2021) High health expenditures and low exposure of population to air pollution as critical factors that can reduce fatality rate in COVID-19 pandemic crisis: a global analysis. Environ Res 199:111339. https://doi.org/10.1016/j.envres.2021.111339

Conticini E, Frediani B, Caro D (2020) Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? Environ Pollut 216:114465. https://doi.org/10.1016/j.envpol.2020.114465

da Silva CFA, Silva MC, dos Santos AM et al (2022) Spatial analysis of socio-economic factors and their relationship with the cases of COVID-19 in Pernambuco, Brazil. Tropical Med Int Health 27:397–407. https://doi.org/10.1111/tmi.13731

Deke O, Hooss KG, Kasten C, et al (2001) Economic impact of climate change: simulations with a regionalized climate-economy model. Kiel working paper No. 1065. Kiel institute for the world economy (IFW Kiel). https://econpapers.repec.org/paper/zbwif/wkwp/1065.htm. Accessed 21 Sept 2022

Di ML (2005) The macro determinants of health expenditure in the United States and Canada: assessing the impact of income, age distribution and time. Health Policy (New York) 71:23–42. https://doi.org/10.1016/j.healthpol.2004.05.007

Ehler A (2021) The socio-economic determinants of COVID-19: a spatial analysis of German county level data. Socioecon Plann Sci 78:101083. https://doi.org/10.1016/j.seps.2021.101083
