Are High-Altitude Residents More Susceptible to Covid-19 in India? Findings and Potential Implications for Research and Policy

Sushmita Chakraborty[1], Upasak Das[2,3] (iD), Udayan Rathore[4] and Prasenjit Sarkhel[5] (iD)

Abstract

In this paper, we study the incidence of COVID-19 and the associated fatality with altitude using high frequency, district level data from India. To understand the implications of the nationwide lockdown after the outbreak, we use data for about four months- two from the lockdown period starting from March 25 till May 31, 2020 and about two months after unlocking was initiated (June 1-July 26, 2020). The multivariate regression result indicates slower growth in average rate of infection during the lockdown period in hilly regions, the gains of which attenuated after the unlocking was initiated. Despite these early gains, the rate of fatalities is significantly higher during the lockdown period in comparison to the plains. The findings remain robust to multiple alternative specifications and methods including one that accounts for confounding possibilities via unobservable and provides consistent estimates of bias adjusted treatment effects. The evidence supports the need for provisioning of public health services and infrastructure upgradation, especially maintenance of adequate stock of life support devices, in high altitude regions. It also underscores the necessity for strengthening and revising the existing Hill Areas Development Programme and integrating important aspects of public health as part of this policy.

Keywords

altitude, COVID-19, India, fatality, health services

The outbreak of the COVID-19 pandemic spread by the SARS-Cov2 virus has resulted in more than 5.5 million deaths (as of January 18, 2022) worldwide, with India being one of the worst affected countries. The threat of infection continues to escalate with the potential emergence of multiple waves after those driven by the Delta and the Omicron variants. Importantly, studies have noted significant inequalities in COVID-19 cases and mortality, reflecting a "social gradient" of health outcomes.[1] There is evidence of positive association between COVID-19 fatality and deprived neighborhood[2] as well as higher infection susceptibility of ethnic minorities.[3] One important aspect that conditions health inequality is the geographical context that subsumes both social relations and access to physical resources.[4] Specifically, adverse geographical features such as high altitude and steep terrains may constrain access to health amenities relative to plain land due to poor transportation and communication.[5] Income poverty, which is related to health status, is also higher in mountainous regions.[6] Thus, systemic health shocks such as COVID-19 are expected to further elevate health inequality in hilly regions and result in adverse health outcomes. In this article, we

examine the linkage between variations in altitude and differences in health outcome in the COVID-19 pandemic using district-level data from India.

Health inequality in India is deeply rooted.[7] Recent investigations into health performance indicators have found a sharp divide across gender, regions, and social and economic groups.[8–10] The disparity is more acute across geographical features. In particular, hilly regions have poorer infrastructure and often lag behind the plain land in terms of income and health.[11] In the state of Manipur in India, for instance, people residing in mountainous terrain are found to have

[1]Independent Research, Siliguri, India
[2]Global Development Institute, University of Manchester, Manchester, UK
[3]Centre for Social Norms and Behavioral Dynamics, University of Pennsylvania, Philadelphia, PA, USA
[4]Oxford Policy Management, New Delhi, India
[5]Department of Economics, University of Kalyani, Kalyani, India

Corresponding Author:
Upasak Das, Global Development Institute, University of Manchester, Manchester, UK; Centre for Social Norms and Behavioral Dynamics, University of Pennsylvania, USA.
Email: upasak.das@manchester.ac.uk

significantly lower utilization of maternal health care compared to those residing in valleys.[12] A study on the estimate of sectoral costs for health, education, and road infrastructure revealed that costs in hilly areas are almost threefold compared to the plains.[13] Thus, altitude is likely to be an important source of health inequalities in the Indian context, especially against the backdrop of the COVID-19 pandemic.

The global evidence on the relationship between variation in altitude and infection spread is mixed. Some studies note that the severity of infection reduces at high altitude, while others find no such association.[14–16] The same holds true for the relationship between altitude and mortality.[15,17] There are multiple pathways through which higher-altitude regions can experience systematically different infection spread and fatality. First, studies show that epidemiological factors, such as physiological adaptation to hypoxia, and environmental factors, such as thinner atmosphere and higher ultraviolet rays, negatively influence health outcomes.[18,19] However, high altitude may also reduce COVID-19 susceptibility.[14] One of the main pathophysiological considerations is that high-altitude residents are often exposed to chronic hypobaric hypoxia, which can lead to an increase in tissue oxygen delivery and improve oxygen utilization through genomic and non-genomic mechanisms.[20,21] Second, high-altitude areas are also characterized by a number of sociodemographic factors that inhibit infection spread. For example, these regions are often remote, and the difficulty in accessing their terrain limits mobility relative to plains. Remoteness and lower population density may also facilitate social distancing that keeps COVID-19 cases at bay.

Nevertheless, because of remoteness and difficult terrain in high-altitude regions, access to emergency health care facilities and expansion of existing health infrastructure, especially the availability of life-saving equipment such as ventilators, might be more challenging. Similarly, due to its rugged terrain, it is likely that testing and contact tracing are arduous and hence lower in high-altitude areas.[21] A combination of these factors can have a countervailing influence and raise both the rate of infection and fatalities in these regions.

In this article, using district-level data for India, we study the growth of infection and fatality in hilly areas during the initial period after the outbreak, controlling for a range of socioeconomic factors that may confound with the relationship of interest. In particular, we study this relationship over a period of about four months, which includes the two months of publicly mandated, nationwide lockdown (March 25 to May 31, 2020). This allows us to assess the growth after unlocking (starting May 31, 2020) and consider an equivalent two-month period for analysis (until July 26, 2020). This is pertinent because the implications of remoteness and difficulty in movement that characterize high-altitude areas may get amplified during the universally imposed lockdown. For example, restrictions on mobility

can have adverse effects in hilly regions in terms of access to health care facilities. Also, it is important to note that during the initial period after the outbreak, health facilities needed critical health infrastructure to ensure lower mortality. This included equipment such as ventilators, Personal Protective Equipment (PPE), and testing equipment, among others, which were in short supply across India. Because of the associated remoteness and restrictions in mobility, high-altitude regions may have been additionally constrained in this respect, as compared to plains, in ways that may have adversely affected both the infection and fatality rates. Such outcomes would also seem quite plausible in the context of the existing evidence that finds health facilities in hilly regions of India to be systematically poorly resourced, even before the occurrence of the COVID-19 pandemic.[22,23]

Our article has multiple vital contributions to make. First, and to the best of our knowledge, there are no studies in India that document the differential spread of COVID-19 cases and fatalities and the aggravated health inequalities across varying altitudes. This is despite the fact that hilly regions constitute a significant proportion of India's total landmass, with a large area positioned in the Himalayas, extending up to 2500 kilometers in length and 250 to 400 kilometers in breadth. In addition, areas classified as hill and mountain zones are distributed across 23 out of 28 Indian states.[24] The states with hilly regions include Uttarakhand, Himachal Pradesh, erstwhile Jammu and Kashmir, West Bengal, and some of the northeastern states, along with the southern and western states of Tamil Nadu, Kerala, Maharashtra, Goa, and Karnataka, among others. Thus, altitude effects in India, if present, are likely to extend over a large geographical habitat and would have significant implications for public health and pandemic response. Second, our work provides evidence of differential implications of mandatory lockdowns on spread and fatality from COVID-19 cases in hilly areas as against plains. Finally, the findings provide important policy lessons for further COVID-19 waves and future pandemics.

Methods

Data

We use the Development Data Lab's COVID-19 India database to gather day-level data on COVID-19 cases and deaths across districts. This data source has information on 684 of the 741 total districts in India. We use data from March 25, 2020, to July 26, 2020. Importantly, a nationwide lockdown was announced by the Government of India in four continuous phases starting on March 25 and continuing until May 31. This limited movement of the entire population across the country was a preventive measure against the pandemic. Thereafter, more freedom was given to the respective state governments to decide on further restrictions. For a fair

representation of the entire duration, we consider data that includes about two months of complete lockdown and two months of unlocking of the economy. Researchers previously have analyzed the lockdown policy in India with respect to COVID-19 incidence.[25]

To calculate district-level minimum and maximum elevation across the country, a high-resolution Digital Elevation Model has been used. More specifically, Shuttle Radar Topographic Mission Digital Elevation Model 2000, with 90-meter ground resolution, has been employed to calculate the zonal statistics on which relevant elevation measures are based. For this, multiple datasets have been combined into a single dataset to calculate zonal statistics using the Arc-GIS software. This method has been used to assess the relationship between altitude and mortality in India.[26]

To control for the influence of potential confounders, we use a range of indicators that account for the socioeconomic, demographic, and mobility characteristics in addition to those related to administrative efficiency, mobility, and hygiene. For this purpose, we use household information provided in the NFHS-4 that is representative at the district level. The survey was conducted in 2015-2016 by the Ministry of Health and Family Welfare through the Indian Institute of Population Science in Mumbai, India, and included 640 districts from 29 states and seven union territories. Apart from socioeconomic and demographic characteristics, the survey provides self-reported information on health indicators, hygiene practice, and basic amenities.

We now provide a brief description of the controls and motivation behind using these covariates in the analysis. Because education can serve as an important determinant affecting the spread of infection through awareness, we use the share of households in the district with at least one member with an educational attainment of higher secondary or more. In addition, we control for the proportion of men who read newspapers and watch television daily and the proportion of households possessing mobile phones. These variables are likely to be correlated with how informed district residents are about the compliance protocols for arresting transmission of COVID-19. Furthermore, we control for penetration of cell phones and for the number of short- and long-term migrants from the districts. These factors account for access to information on COVID-19 and potential infection spread emerging from returning of migrants, respectively. This also adjusts for the fact that even if migration is low in terms of share of population, it may be high enough to enable the rapid spatial spread of COVID-19 cases.

Evidence suggests higher vulnerability of the elderly to the infection.[27] Accordingly, we control for the proportion of households in a district with at least one elderly member. Studies have also indicated how urbanization is linked to the COVID-19 spread,[28] so we also control for share of households residing in urban areas. Living conditions and ability to isolate in these districts have been accounted for through two variables: average number of rooms used for sleeping and the share of families staying in owned houses. We also include other explicit measures of income and wealth as they are found to be related to health inequality. Since the publication of the Black Report in 1980,[29,30] studies have also associated systematic differences in health outcomes with social and economic disparities such as income and wealth.[6] Accordingly, we account for the economic profile of the districts in two ways. First, we take the average value of the standardized wealth index that NFHS-4 calculates for each surveyed household through possession of a set of durable assets. In addition, the proportion of households having a Below Poverty Line card is also considered.

Administrative efficiency is likely to be another correlate of disease transmission and has been controlled for through three indicators: share of households with at least one member having a bank or post office account; proportion of households with all members possessing an Aadhar, a unique, 12-digit identification number for every resident of India; and average distance of the district from the state capital, taken from Google Earth Pro. Apart from these variables, we control for mobility by taking into account the proportion of surveyed men who have stayed outside their residence for at least one month, as well as the total number of flights entering the district between April and July 2020 that had been taken from the Airport Authority of India. Both of these account for within-district variation in prevalence of migration and control for national and international travelers, who are key predictors of the spread of the virus.

Given the importance of hygiene in the context of the COVID-19 pandemic, we incorporate a set of related indicators into our analysis. First, we include the share of households at the district level having no toilet facilities along with the share with improved drinking water sources, as per World Health Organization guidelines. In addition, because studies have indicated how access to piped water and the type of cooking fuel have a direct implication for the health profile of household members, the share of households with piped water within the house premise or the yard, along with those using liquefied petroleum gas (LPG) as cooking fuel, have been included in the regression.[31,32] Further, we also control for the proportion of households where any members smoke inside the house and where soaps/detergents were found in handwashing locations. In addition, we controlled for district-level public health service using two indicators: share of households with at least one member covered by health insurance and share where members generally go to public health centers when sick.

Research has indicated that health disparity is a structural issue whereby the deprived population typically resides in underserved areas with inadequate provision of health services.[33] To further account for availability of physical and human health infrastructure, the number of hospitals per 10 000 households and doctors per 1 million households is considered. This has been taken from the Development Data

Lab's COVID-19 India database, which uses the Population Census and the Socio Economic Caste Census for 2011 and 2012 to retrieve the information.[34] We are able to map our main variable of interest, as well as the set of control variables, with information on COVID-19 cases and deaths for about 612 of the 719 (84%) districts from NFHS-4. Previously, similar data has been used to examine the effects of caste-based diversity on the infection spread in India.[35]

Empirical Models

We first estimate COVID-19 cases per 10 000 population from the commencement of the lockdown on March 25, 2020, until July 26, 2020, by dividing this time span into eight periods of about 15 days each. Note that the first four periods correspond to the four nationwide lockdown phases, with the next four representing the unlocking phases. In particular, we estimate the following regression equation through Ordinary Least Squares (OLS):

$$YP_{dst} = C + \beta_1.y_{ds0} + \beta_2.D_{ds(t-1)} + \beta_3 E_{ds}$$
$$+ \beta_4(E_{ds} * t) + \beta_5 X_d + \pi_s + t + u_{ds} \qquad (1)$$

Here, YP_{dst} denotes the number of cases per 10 000 population reported in district d located in state s at the end of the period $t.y_{ds0}$ is the logarithmic value of the number of COVID-19 cases of infection that had been reported before the announcement of the lockdown, and D_{dst-1} is the logarithmic value of the number of deaths reported due to the pandemic in the previous period $(t-1)$. The rationale behind controlling for pre-lockdown cases is to account for emergence of initial hotspots that may potentially be linked to geospatial variables and the outcome of interest. In addition, deaths in the previous period control for administrative responses undertaken by local authorities to curtail the growth in infection, which is likely to be more sensitive to fatalities in the previous period. To ensure that districts with no cases do not drop out selectively, leading to selection bias, we add one to each of the variables before logarithmic transformation.[36] E_{ds} takes the value of 1 if the mean the lowest and higher elevation point within the district is 1000 meters or above and 0 otherwise. The vector of the covariates at the district level is given by X_d. The state- and period-level fixed-effect vectors that control for heterogeneities across states and period of analysis are denoted by π_s and t, while u_{ds} is the error term. To assess the relationship between elevation and the changes in infections over time, we examine the coefficients, β_4. The standard errors in this model have been clustered at the district level.

To estimate the effects of elevation on the fatalities, we estimate the following model:

$$DP_{dst} = C + \beta_1 y_{ds0} + \beta_2 YP_{dst-1} + \beta_3 E_{ds}$$
$$+ \beta_4(E_{ds} * t) + \beta_5 X_d + \pi_s + t + u_{ds} \qquad (2)$$

Here, D_{dst} represents number of deaths per 10 000 population in district d located in state s at the end of the period t and is taken to be a function of pre-lockdown reported cases of infection as well as the number of cases per 10 000 population in the lagged period, $t-1$. This ensures that we normalize the growth of fatalities by the growth of infection that occurred in the previous period, especially as these are likely to be a critical determinant for future fatalities. The association between elevation and deaths over time is given by the vector of the coefficients, β_4.

To get closer to causal effects of elevation on growth in reported infection and fatalities over the eight periods, one needs to account for the potential endogeneity emanating from Omitted Variable Bias (OVB). We argue that reverse causality is less of a concern. Because of the unanticipated nature of the COVID-19-induced lockdown in India, the chances of shifting residential locations based on altitude with the outbreak would be highly unlikely. Studies in fact have shown low prevalence of cross-district or even cross-state migration in India.[37,38] Even NFHS-4 data indicates that about 86% of the sampled men reported having stayed in the surveyed house for at least 10 years.

However, it is possible that the residents of regions located at higher altitude are inherently different than those in lower elevation zones. We have controlled for an elaborate set of covariates that may capture some part of this difference, as mentioned earlier, in addition to using state-fixed effects to control for all the time-invariant characteristics that may be common to the states and affect infection and death rates. Nevertheless, to avoid bias that may crop up because of unobserved variables, we apply a method which under certain assumptions account for this possibility.[39–41] Here, a possible range of β_4 is derived through two parameters: δ and R_{max}. R_{max} is the R-squared value of a hypothetical regression that incorporates all the observable and unobserved components and hence is the maximum value of R-square it can achieve. δ is the extent to which unobservables are dependent on observables. For example, $\delta = 2$ implies that the extent of unobservables explaining the outcome variable is double of that explained by the observables. While there can be various options to compute the bound on R_{max}, using data from a wide sample of randomized experiments, it is proposed $R_{max} = 1.3 * R_0$, where R_0 is the R-squared value of the full model with observed control variables. δ can be assumed to lie between [-1,1]. The argument is if δ goes beyond 1 or falls below -1, the variation in the outcome explained by unobservable(s) has to be higher than the combined influence of all the observables put together to push the effect size of the coefficient of the variable of interest to zero. Accordingly, we check if $|\delta|$ for which the coefficient of interest turns 0 with $R_{max} = 1.3 * R_0$ exceeds a threshold of 1. A more detailed description on the method is given in the appendix explanation 1S.

Results

Descriptive Statistics and Preliminary Analysis

We first present the number of cases and deaths across each of the eight periods of our analysis (Figure 1). We observe that the total number of reported cases and deaths in each period increased somewhat moderately in the first four phases of the lockdown. As one would expect, the growth in these variables increased after the unlocking was initiated. A summary of period-wise cases and deaths per 10 000 of population for the duration of analysis is provided in Supplemental Material (Table 1S).

Summary statistics to show the number of cases and deaths per 10 000 population across the eight periods separately for districts with high and low average elevation are provided in Figure 1S of the Supplemental material. This categorization has been done on the basis of average elevation of a district being less than or greater than 1000 meters. This roughly constitutes of 87% and 13% of the total districts in our article, respectively. During the lockdown phase (Periods 1 to 4), we observe that the growth in infection and deaths was higher for the lower-altitude districts.

Digital Elevation Model

Next, we present a digital elevation model, which has been combined with period-wise number of reported cases per 10 000 population to represent its distribution in relation to altitude. Here, the landmass has been classified into seven altitude zones, starting from areas with elevation below 150 meters to those with more than 5000 meters from mean sea level. The data has been categorized into three groups that include low (no reported cases), moderate (less than one case per 10 000 population,) and high (one or more cases per 10 000 population).

Figure 2 presents the period-wise elevation model with the number of reported cases for Periods 1, 4, 5, and 8, with numbers for other periods given in the Supplemental material (maps for all periods are given in Supplemental Material, Figure 2S). As one can observe, the Himalayan districts, which include the northern (Himachal Pradesh and Uttarakhand) and northeastern states (Sikkim, Arunachal Pradesh, Manipur, Meghalaya, Nagaland, and Mizoram) having an average elevation varying from 1000 meters to 4000 meters, registered lower infection rates across the eight periods. On the other hand, districts located in low- and mid-altitude areas (coastal plain, Indo-Gangetic plain, central highlands, and Deccan plateau) are found to be more susceptible to COVID-19 infection. Nevertheless, some of the districts from the erstwhile state of Jammu and Kashmir, which have an average elevation of more than 1000 meters, reported high infection rates.

Figure 3 presents the superimposition of the number of deaths per 10 000 population over the elevation model over the Periods 1, 4, 5, and 8 (maps for all periods are given in
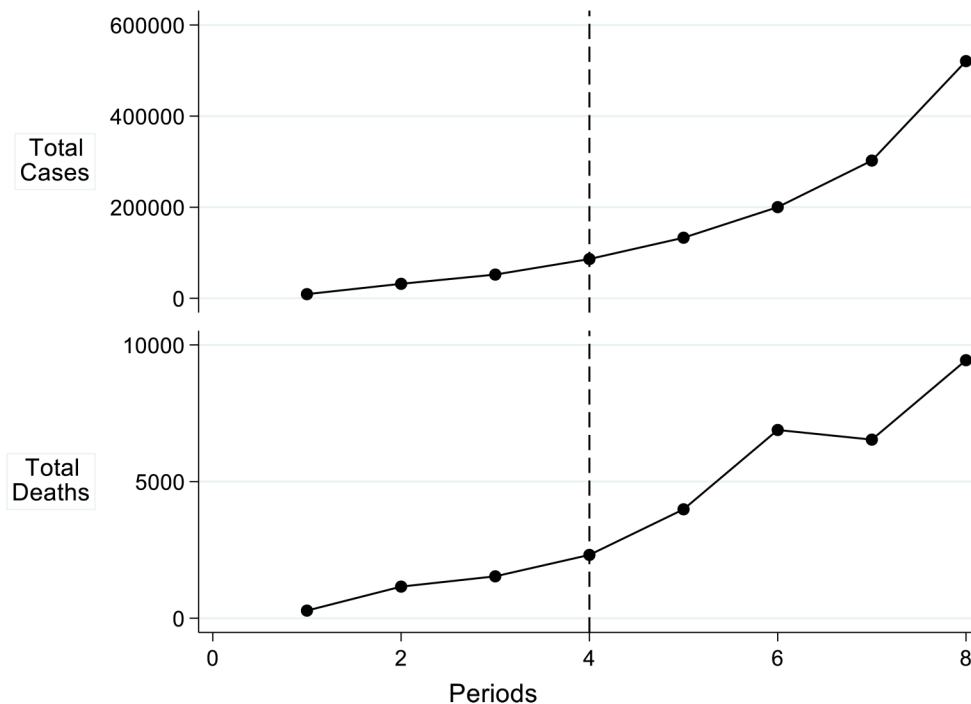

Figure 1. All-India total cases and deaths across each of the eight periods of lockdown and unlocking (march 25 to July 26, 2020). Source: COVID India Database.

Table 1. Summary Statistics.

	N	Mean	Mean If Average Elevation <1,000 m	Mean If Average Elevation ≥ 1,000 m
Distance between districts and state capital (100 km)	648	2.55	2.63	2.04*
Proportion of HHs with at least one member with higher secondary or above education	616	0.24	0.24	0.25
Proportion of HHs with at least one elderly member (60 years or above)	616	0.36	0.36	0.33*
Proportion of HHs with a bank/PO account	616	0.89	0.89	0.88
Proportion of HHs with members covered by health insurance	616	0.26	0.26	0.26
Proportion of HHs with Below Poverty Line cards	616	0.39	0.39	0.36
Proportion of HHs where members generally go to public health centers when sick	616	0.53	0.49	0.80*
Proportion of HHs where all members have Aadhar card	616	0.44	0.44	0.41
Proportion of HHs with improved drinking water facilities	616	0.87	0.88	0.84*
Proportion of HHs with piped water inside the house or yard	616	0.65	0.63	0.72*
Proportion of HHs with no toilets	616	0.38	0.42	0.14*
Proportion of HHs which use LPG as a cooking fuel	616	0.36	0.36	0.37
Proportion of HHs where members wash their hands with soap after defecation	616	0.61	0.59	0.70*
Standardized wealth index (0-1)	616	0.44	0.43	0.48
Average number of rooms used for sleeping	616	1.88	1.83	2.20*
Proportion of HHs where houses are owned	616	0.81	0.82	0.78*
Proportion of men who have been staying in their residence for at least 10 years	616	0.86	0.86	0.82*
Proportion of men who have stayed outside their residence for one month or above	616	0.17	0.16	0.22*
No. of HH in the district (*1000)	648	362.62	401.91	109.27*
Population density per square km	625	1232.83	1285.06	811.99
No. of allopathic hospitals per 10 000 HHs	648	0.60	0.50	1.27*
No. of allopathic doctors per million HHs	648	409.19	339.56	858.16*
Total flights between April and July (in 00 s)	648	1.28	1.38	0.62
Proportion of HHs staying in urban areas	616	0.28	0.29	0.23*
Proportion of men who read newspapers	616	0.28	0.29	0.22*
Proportion of men who watch television daily	616	0.57	0.57	0.58
Proportion of households with mobile phone	616	0.90	0.89	0.92*
Estimated no. of men migrant who migrated for 6 months (million)	616	14.4	15.5	7.68*
Estimated no. of men migrant who migrated for 1 months (million)	616	28.6	31.2	12.9*

Note: Of the 641 districts for which information on cases and deaths is available, 554 (86%) have average altitude of <1000 meters,*p value < 0.05.

Supplemental Material, Figure 3S). Initially, the figures indicate higher deaths per 10 000 population in some of the high-elevation areas. However, across subsequent periods, it becomes difficult to visually gauge this relationship as higher deaths are also observed in some of the lowland areas. This warrants the need to examine the relationship analytically through multivariate regression analysis that allows us to obtain an average estimate of the association with COVID-19 outcomes, after accounting for the possible confounders.

Regressions

First, we run pooled OLS regression as outlined in equation (1) to examine period-wise changes in reported infection rate at the district level in hilly regions. The summary measures of the control variables that have been used in the regression are given in Table 1.

Figure 4 presents the relevant marginal effects after accounting for the potential confounders, as discussed earlier. Interestingly, we observe that the rates of infection for the first three phases of the lockdown were significantly lower, on average, for higher-altitude regions. The effect size here varied from 81 to 94 lesser cases per 1 million population, on average, for the first three phases of the lockdown. Starting from the fourth period, which corresponds to one period before initiation of the unlocking, the gains in hilly regions appear to die down. During the unlocking period, we do not observe any significant gains in terms of a lower number of cases in these regions. Importantly, we are able
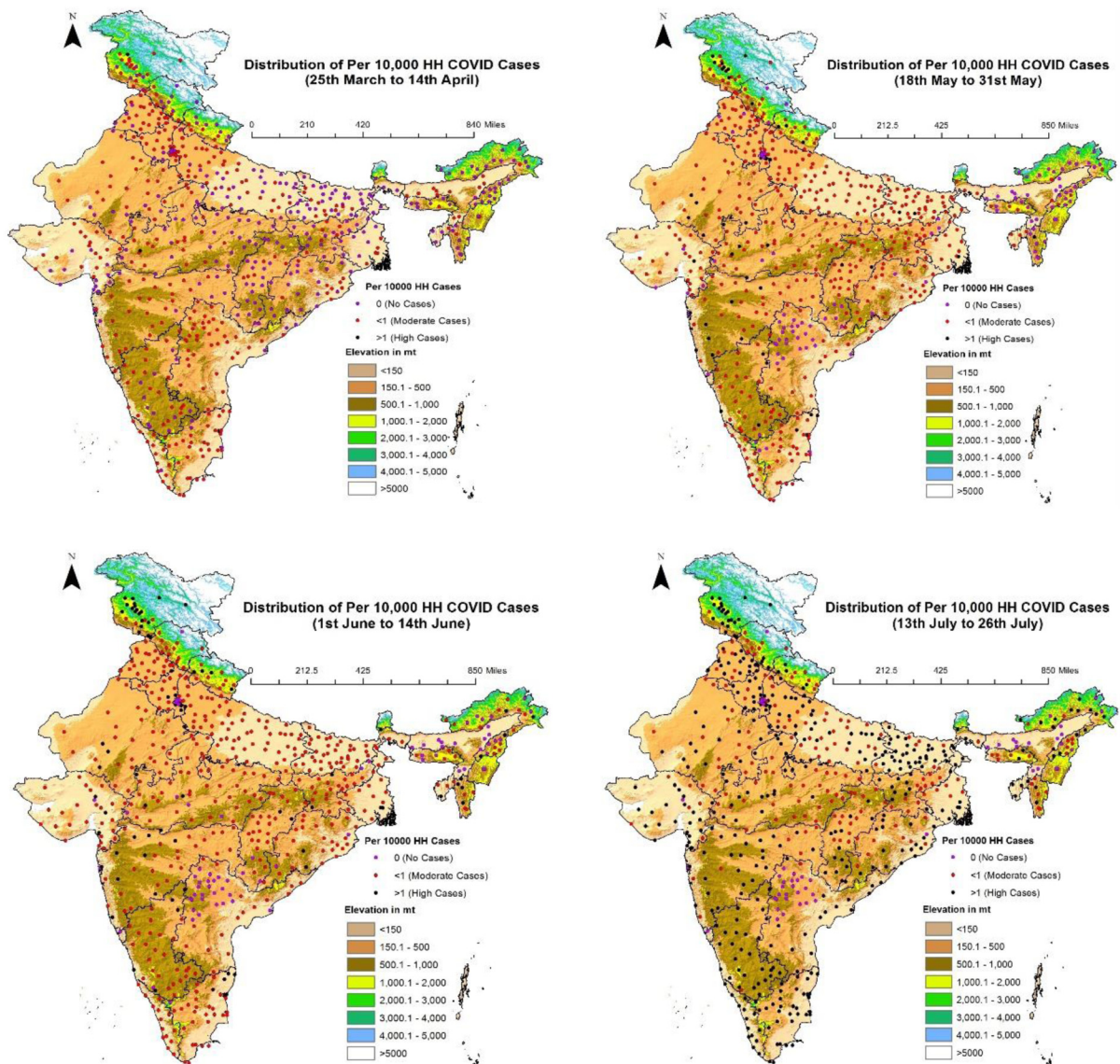
Figure 2. Elevation and cases per 10 000 population (periods 1, 4, 5, and 8).

to replicate these results using a continuous variable elevation measured by the mean of minimum and maximum elevation (Supplemental Material, Figure 4S). To check for any reporting errors, we also consider three-day and daily rolling averages of COVID-19 cases for outcome, and our findings remain statistically robust to these alternate specifications (Supplementary Table 2S). In addition, we check whether these results are being driven by districts that constitute the state capitals by validating our results for non-capital districts. We are able to replicate in this case as well, and the results can be provided on request.

We run similar regressions for the number of fatalities per 10 000 population as elucidated in equation (2). Figure 5 presents the associated marginal effects for districts with an elevation of 1000 meters or more. Notably, we find that for each

of the four periods that correspond to the nationwide lockdown in India, the rate of fatalities is found to be significantly higher in hilly districts. We find statistically similar results for regions with higher average elevation (Supplemental Material, Figure 5S) as well as results using three-day rolling and daily COVID-19-related deaths per 10 000 people as outcome variable (Supplementary Table 3S).

For an additional robustness check, we run pooled OLS regressions taking the daily number of reported cases per 10 000 population at the district level and replicate our findings. Instead of period dummies, as were taken in the earlier model, we include daily dummies and assess the changes of the coefficients of the relevant interaction term with the dummy variable that captures hilly regions over the days starting from March 25. The findings from Figures 6 and 7
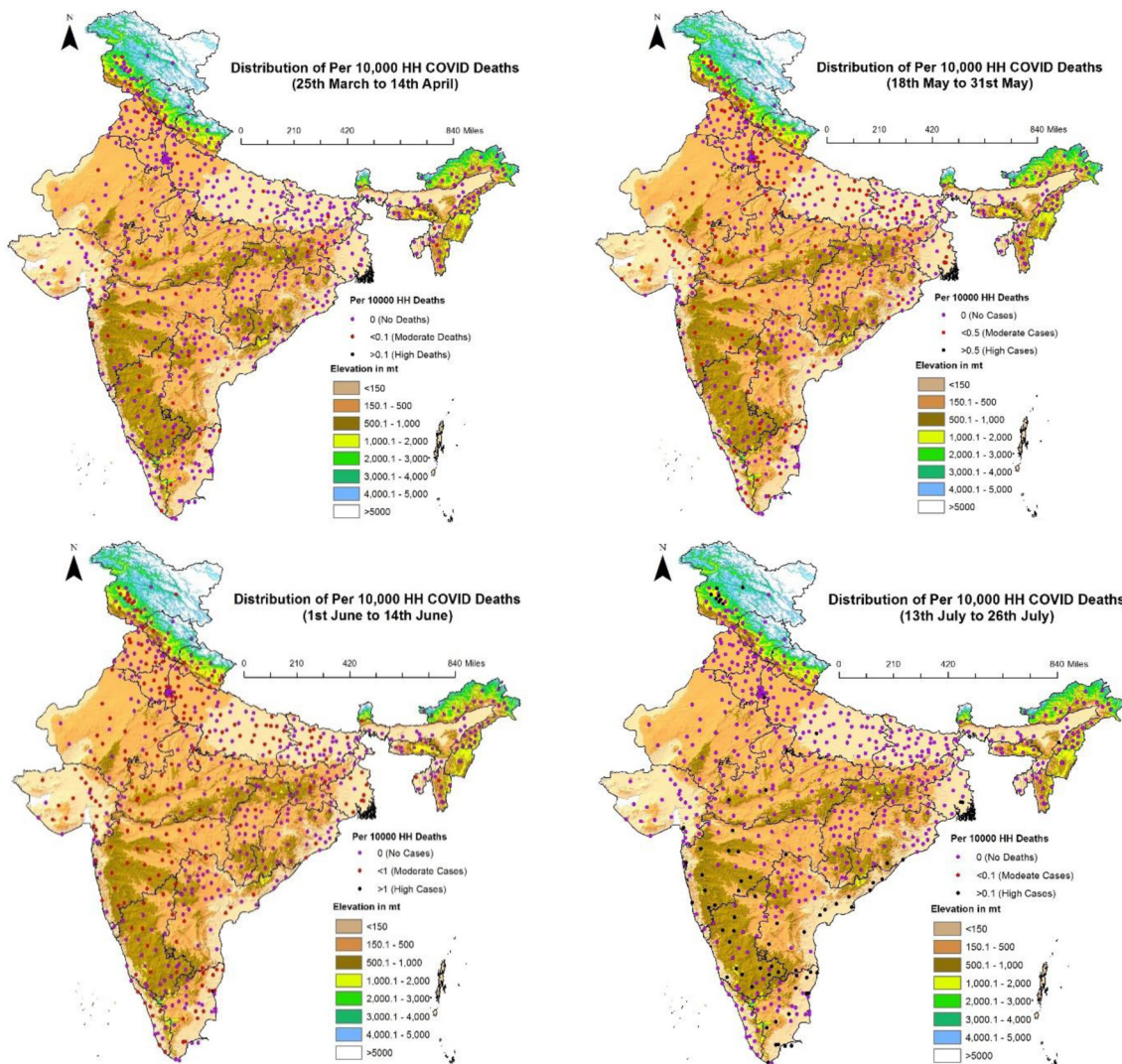
Figure 3. Elevation and number of deaths per 10 000 population (periods 1, 4, 5, and 8).

that present the marginal effects from the regression indicate similar results of average cases per 10 000 being lower in high-altitude regions during the lockdown, but average deaths per 10 000 being higher, ceteris paribus. Similar daily figures using a continuous measure of average elevation for reported cases and deaths per 10 000 population are given in Supplemental Material, Figures 6S and 7S.

In this article, controlling for various confounding possibilities, we attempt to isolate the average association between COVID-19 infection and fatality rates with variation in average altitude. For this, we control for a range of correlates, as discussed previously. However, it is still plausible that these results stem from OVB, which are correlated with both the variable of interest and the outcome. As discussed in Section 3, we use a novel method that provides consistent estimates of bias-adjusted treatment effects that may stem from OVB.[39,40] For periods where the coefficient

of interest was significant, modulus of δ statistic is found to be well above unity (Table 2), suggesting that the variation in outcome explained by the unobservables would have to be well over that explained by all the correlates used in our analysis to push these effects to zero. Given the fact that we use an extensive specification that accounts for a range of confounding possibilities and explains about half of the variation in the outcomes of interest, it is likely that our results would remain statistically significant.

It is possible that high-altitude regions may also have systematic and asymmetrically differential reporting of cases and deaths that may influence the results we observe. A recent article indicates that the likelihood of underreporting, while being substantially high during the second wave (starting from April 2021), was of limited concern during the first wave, on which our study is based.[42] Nevertheless, we consider the possible channels and argue that our estimates are
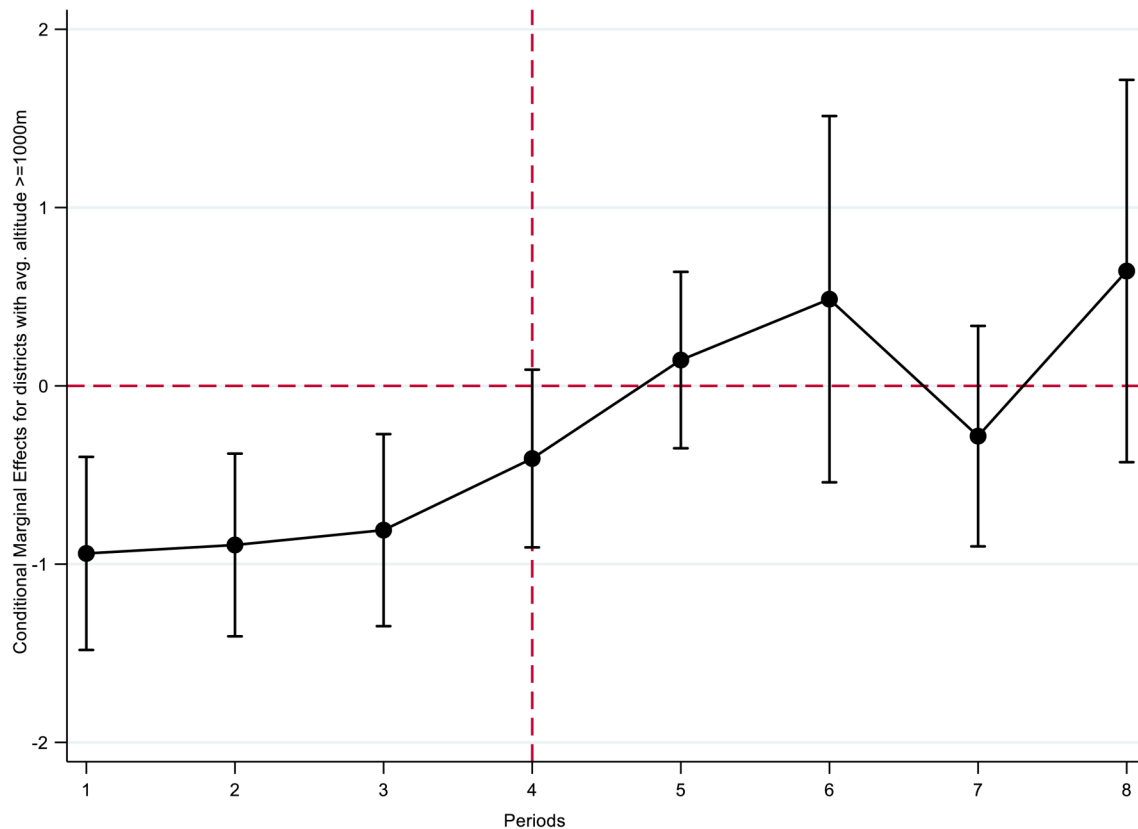
Figure 4. Conditional marginal effects for districts with average elevation over 1000 meters 1000 m on rate of infection of COVID-19 cases (per 10 000 population) across four periods of lockdown and four periods of unlocking, until July 26, 2020 (OLS regression). *Note: 95% Confidence Intervals calculated by clustering standard errors at the district level are given along with the marginal effects. The dependent variable is number of reported cases per 10 000 population. The period left of the red vertical line (Periods 1 to 4) denotes the lockdown period and that toward the right is when the unlocking procedure was initiated.*

unlikely to biased because of underreporting. The first channel might be differential testing in these regions, which may also lead to a lower number of cases. In the absence of district-level data on testing, we utilize the period-wise, state-level number of tests conducted and incorporate that in the main regression model as an additional control variable. Please note that public health in India comes under the state level, and state governments reserve the power to promulgate laws on and allocate resources to it. Therefore, it is likely that daily testing in a district would be influenced by state-level policies. The results from both sets of regressions indicate similar results: The growth rate in COVID-19 cases is significantly lower in high-altitude areas while growth in the death rate is significantly higher (Supplementary Figures 8S and 9S). Please note that the testing data is fully available during the period of analysis for only 26 states and Union Territories and is partially available for six states and Union Territories. Hence, we do not use it as our main specification and employ it as a robustness check exercise.

Please note that as control variables in the regression, we have accounted for the state of health infrastructure through the number of allopathic hospitals and doctors in the district. Also, we have controlled for the district-level dependence of households on public health centers, which are often responsible for the extent of testing. Furthermore, we have adjusted for the level of awareness through the proportion of males who watch television daily and read newspapers and share of households possessing mobile phones, in addition to a number of economic indicators. These variables can potentially be correlated with underreporting of cases and deaths and hence, to some extent, we account for it in the regression. In addition, the bias-adjusted treatment effect model also indicates these unobservables leading to underreporting are unlikely to confound our estimates and hence lends credence to our inferences. Nevertheless, these are proxy measures, and the inability to account for the actual extent of district-level underreporting because of paucity of data remains a limitation of the study.

With this caveat, to sum up, we find that the growth of infection had been significantly lower in the hilly areas as compared to the lowlands during the lockdown period. This could probably be because of better air quality, lower economic activity, and lower mobility in these areas, which
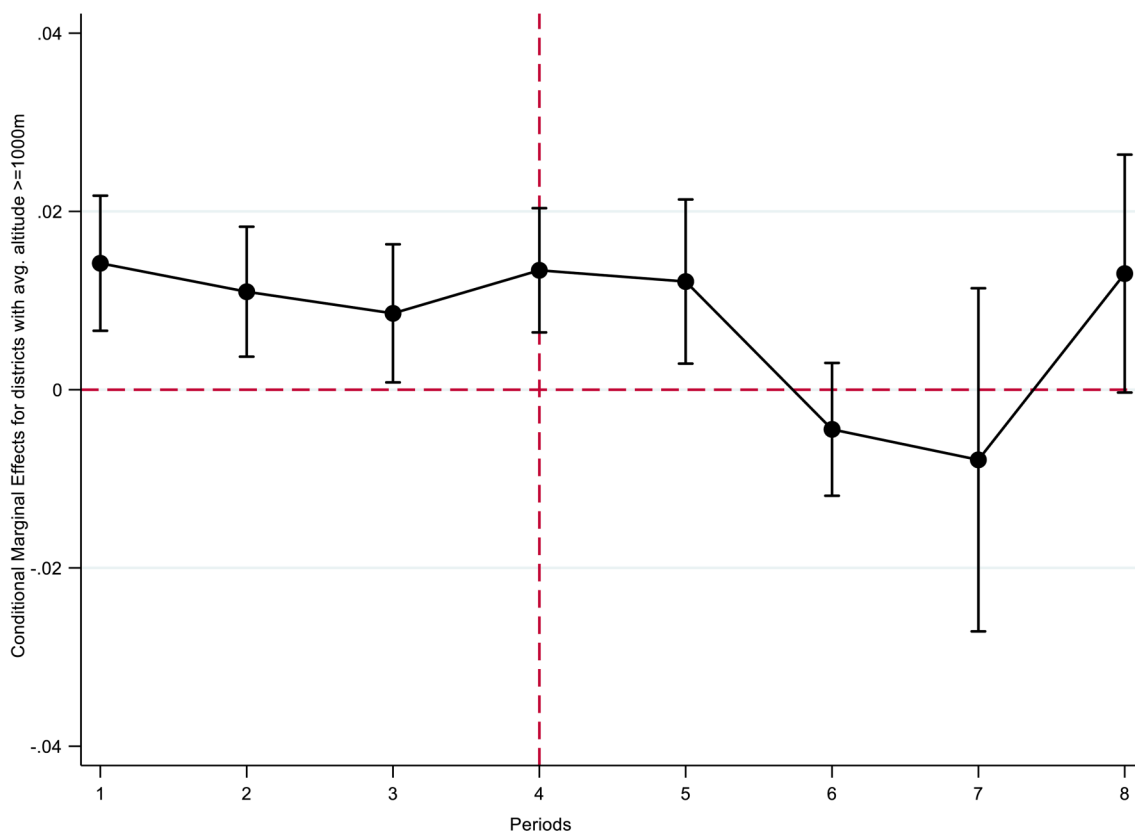
Figure 5. Conditional marginal effects for districts with average elevation over 1000 meters on deaths per 10 000 population from COVID-19 across four periods of lockdown and four periods of unlocking, until July 26, 2020 (OLS regression). *Note: 95% Confidence Intervals calculated by clustering standard errors at the district level are given along with the marginal effects. The dependent variable is number of reported deaths per 10 000 population. The period left of the red vertical line (Periods 1 to 4) denotes the lockdown period and that toward the right is when the unlocking procedure was initiated.*

were amplified during the lockdown period. However, we also observe these districts to experience significantly higher fatalities during the lockdown period, which eased after the unlocking procedure was initiated. While we are unable to single out specific mechanisms behind this finding, available evidence suggests that our results plausibly hint at the poorly equipped health care facilities in the isolated hilly regions of the country.[43] At a time when most of India experienced a shortfall in life-saving health infrastructure such as ventilators and PPE kits, this problem might have been even more acute for hilly regions as against plains.

Because of potential supply chain disruptions in provisioning necessary health care equipment during the sudden declaration of the lockdown, it is possible the number of deaths would have risen in these regions, which are generally remote. Importantly, during the time of the lockdown, mobility was highly restricted, which makes inhabitants from hilly areas more susceptible to fatalities. Post-lockdown, easing of mobility restrictions and improvement in provisioning of scarce health care infrastructure are possible reasons why growth in deaths would converge with those in low-lying areas.

Policy Discussions

Our results suggest that geographical predictors such as altitude can be a significant predictor of the disparate impacts of health shocks such as COVID-19. In high-altitude areas, potentially because of lower population density and remoteness, we find a temporary lowering of the infection spread. In fact, the low infection rate in high-altitude areas could be a possible reflection of the strengthening of the remoteness effect via the mandatory lockdown imposed by the government. The effect eroded once the mobility restrictions were withdrawn. Plausibly, economic hardships during the lockdown phases would have been higher in these regions that often depend on tourism and trade for livelihoods. With easing of the restrictions, the likelihood of violation of COVID-19 compliance norms that included social distancing would have been higher due to the intense economic needs. Of note is the fact that, during this period, the return of migration workers from urban centers and metropolitan areas was associated with higher infection growth.[28] Given the fact that the proportion of men staying outside their residence for one month or longer is higher in hilly areas (Table 1), more
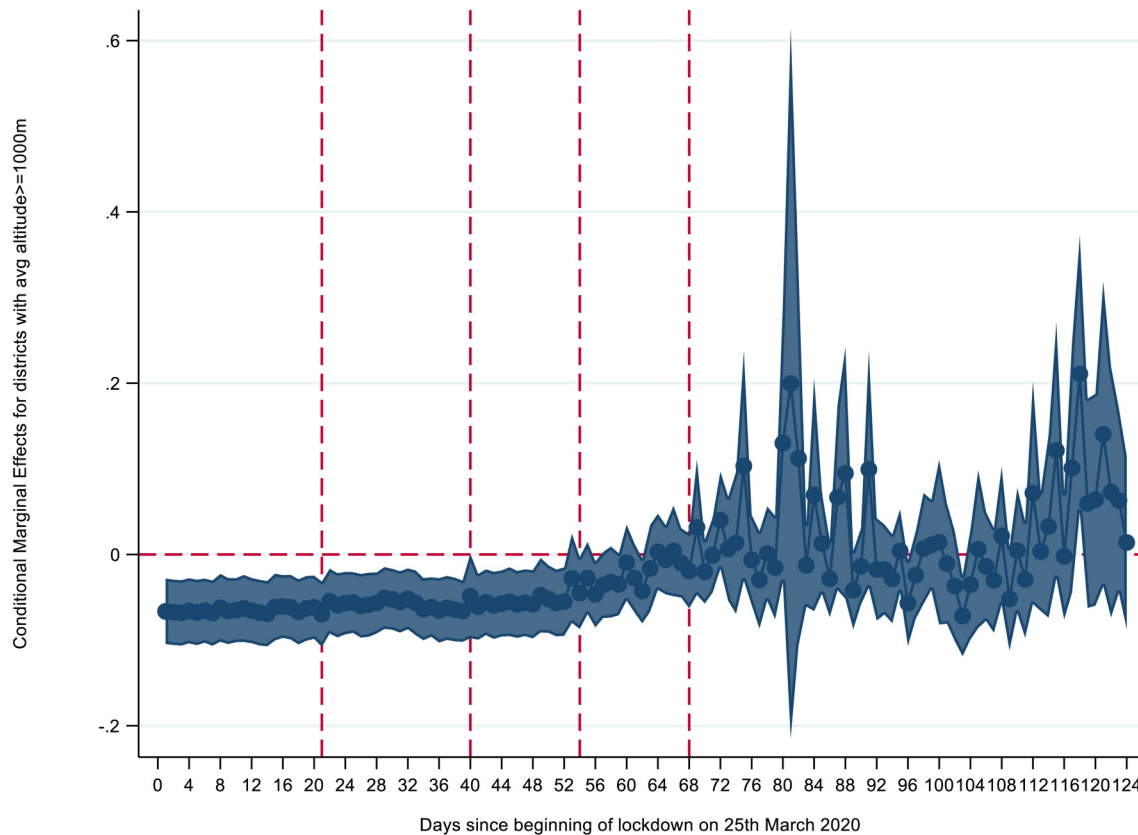
Figure 6. Coefficients of variables of interest from OLS regression for COVID-19 cases per 10 000 population for each day since initiation of lockdown for higher-altitude districts with average altitude over 1000 meters (until end of July 26, 2020). *Note: 95% Confidence Intervals calculated by clustering standard errors at the district level are given along with the marginal effects. The dependent variable is number of reported cases per 10 000 population. Vertical lines signify the end of each of the four phases of lockdown.*

attention needs to be paid to framing guidelines that ensure health surveillance of drifting workers.

The findings of higher COVID-19 fatality during the lockdown periods in hilly regions potentially emphasize the lack of timely access to health facilities and poor health infrastructure, especially the shortages of life-saving equipment such as ventilators and PPEs in high-altitude regions.[44] Access to health care in high-altitude terrains is often constrained due to the skewed distribution of health institutions and poor transport conditions and could be a potential source of unequal health outcomes. Studies have shown that the average radial distance to reach the primary health center in the hilly states is significantly higher than the national average.[45] Costly access to health care facilities is also seen to result in a reduction of utilization of health services, such as maternal and child health services, in hilly areas.[46] In addition, health sub-centers, which are the first point of contact with the health system, are often reported to be located in the outskirts of habitation to minimize the land cost. Studies have also found health care centers in these regions to be lacking in resources and equipment.[23] In this scenario, with sudden initiation of the lockdown, the shortage

of this equipment in hilly regions could have been systematically higher. This could have led to higher growth in mortality despite lower growth in reported cases during the lockdown period. The findings call for prioritizing health care facilities in hilly regions along with provisions for emergency medical interventions while addressing the spatial inequality for an expedited policy response in the future.

Importantly, in the larger context, this also calls for immediate strengthening and revitalizing the existing Hill Area Development Program (HADP), which has been in the Designated Hill Areas operation since the middle of the decade of 1970 after the initiation of the Fifth Five-Year Plan. The basic objective of the program has been socioeconomic development for residents of hilly areas and living in harmony with ecological development. Accordingly, the plans initiated under the HADP have been aimed at promoting livelihoods through sustainable use of natural resources in the area. However, plans for improving health care in these regions under the aegis of the program have mostly registered poor performance. Reports have also documented that health departments did not have significant participation in the
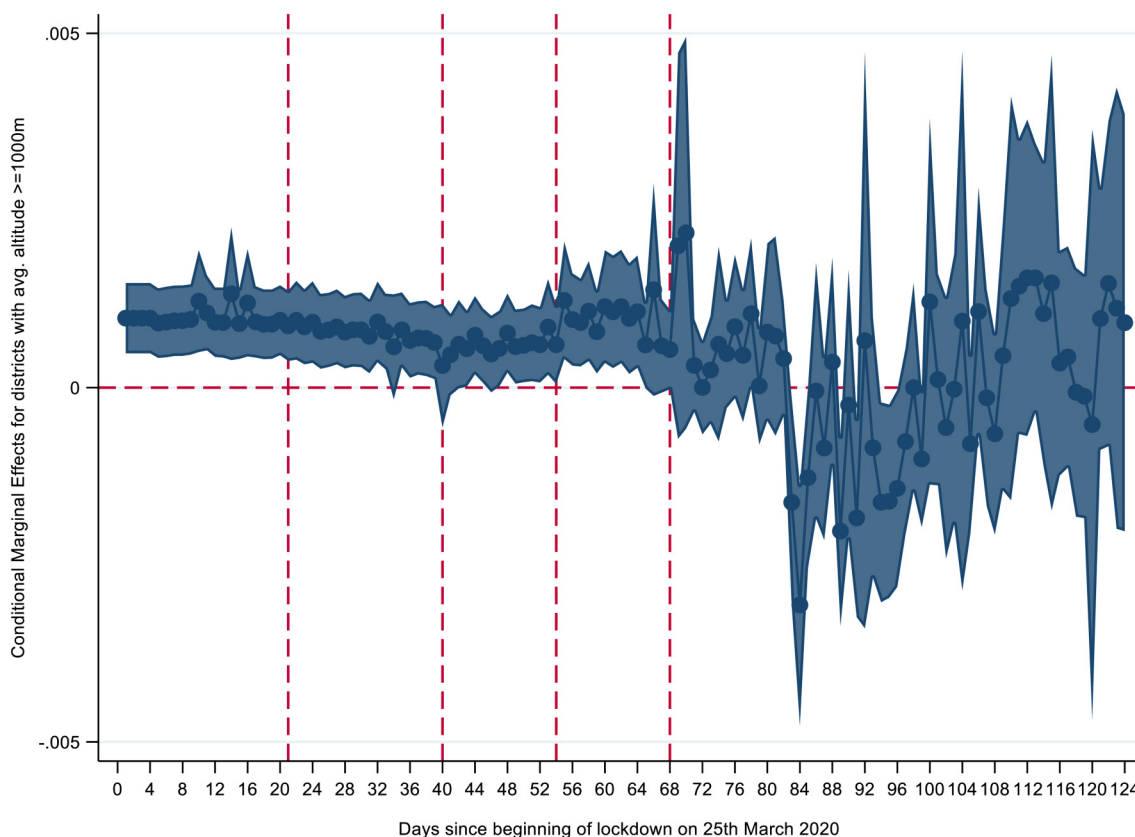
Figure 7. Coefficients of variables of interest from OLS regression for COVID-19 deaths per 10 000 population for each day since initiation of lockdown for higher-altitude districts with average altitude over 1000 meters (until end of July 26, 2020). *Note: 95% Confidence Intervals calculated by clustering standard errors at the district level is given along with the marginal effects. The dependent variable is number of reported deaths per 10 000 population. Vertical lines signify the end of each of the four phases of lockdown.*

Table 2. Accounting for Potential OVB for Higher Altitude Districts (Average Altitude $\geq$ 1000 Meters ($=$ 1)).

Period	δ For Cases Per 10 000 Population	δ For Deaths Per 10 000 Population
1 (March 25–April 14, 2020)	20.39	−1.15
2 (April 15–May 3, 2020)	30.03	−1.51
3 (May 4-17, 2020)	14.49	−1.72
4 (May 18-31, 2020)	12.89	−1.36
5 (June 1-14, 2020)	0.67	−1.78
6 (June 15-28, 2020)	1.40	−2.26
7 (June 29–July 12, 2020)	−0.63	−1.63
8 (July 13-26, 2020)	0.55	−17.54

The command *psacalc* in STATA 14 is used to generate these results.

preparation of the action plan for HADP.[47] Our study underscores the need for greater involvement of health professionals in HADP and renewed emphasis on distributional aspects of health care, including policies that reduce the cost of access to emergency services.

Conclusion

Assessing the spread of COVID-19 infections and deaths in hilly regions is complex. Apart from epidemiological aspects, such as higher exposure to chronic hypobaric hypoxia that can reduce the susceptibility to the virus, other geographic and sociodemographic factors, including uneven terrain, remoteness, and lack of health infrastructure, are pertinent factors connecting the two. In this context, using high-frequency, district-level data on reported cases and fatalities, we study these outcomes of COVID-19 in hilly regions of India. The key findings include lower growth in number of cases per 10 000 population on average during the lockdown periods in hilly regions, the gains of which wither away with the unlocking. Further, we also find systematically higher fatalities during the lockdown in these regions, which reduces only after unlocking.

For policy prescriptions, hilly regions need to better equipped to "respond swiftly" to public health crises.[48] Our findings emphasize the need for medical interventions and maintenance of adequate stocks of life-support devices that includes ventilators and PPEs, among other equipment, in these areas. This becomes particularly important in case of subsequent waves of infection in India. We also argue for

immediate strengthening and revitalizing of the existing HADP. In particular, our findings underscore the need for effective health care provisioning to deal with future pandemics and ensure better health support systems. Because hilly regions are vulnerable to these events, our study provides a basis for further research that can use epidemiological models to forecast COVID-19 spread and fatality differentials in these areas. This would appear especially relevant in the context of COVID-19's ongoing viral evolution to the emergence of new COVID-19 variants.

In addition, this article has multiple implications for further research. First, due to paucity of data, we are unable to pin down the exact mechanisms that can explain our findings. It is possible that better air quality, lower mobility, and remoteness, among other factors that characterize high-altitude areas, may have been influential in reducing the spread of COVID-19 infection. If infection is contracted, impediments in access to health infrastructure and poor connectivity in hilly regions that raises the cost of delivery of public services could have increased the fatality rate in these areas. With availability of disaggregated data and primary surveys, exploring these pathways can be one vital area of future research. Second, further research in other contexts is required to establish external validity of our findings. Existing research from Latin American countries such as Bolivia, Peru, Ecuador, and Brazil indicates that people in high-altitude regions are also less susceptible to COVID-19 infection, but find no evidence of differential fatality rates.[14,15] Future research directed toward the Global South in Southeast Asia and Africa is necessary to understand the implications of health infrastructure and connectivity on spread and fatalities of pandemics in their hilly regions.

Declaration of Conflicting Interests

Funding

ORCID iDs

Upasak Das [iD] https://orcid.org/0000-0002-4371-0139
Prasenjit Sarkhel [iD] https://orcid.org/0000-0001-5486-6959

Supplemental Material

Supplemental material for this article is available online.

References

1. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. *J Epidemiol Commun H*. 2020;74(11):964–968.
2. Martins-Filho PR, De Souza Araújo AA, Quintans-Júnior LJ, Santos VS. COVID-19 fatality rates related to social inequality in northeast Brazil: a neighbourhood-level analysis. *J Travel Med*. 2020;27(7):1–3.
3. El-Khatib Z, Jacobs GB, Ikomey GM, Neogi U. The disproportionate effect of COVID-19 mortality on ethnic minorities: genetics or health inequalities? *E Clinical Medicine*. 2020;23(9):1–2.
4. Cummins S, Curtis S, Diez-Roux AV, Macintyre S. Understanding and representing 'place' in health research: a relational approach. *Soc Sci Med*. 2007;65(9):1825–1838. doi:10.1016/j.socscimed.2007.05.036
5. Xu X, You H, Gu H, et al. Inequalities in health-related quality of life and the contribution from socioeconomic status: evidence from Tibet, China. *BMC Public Health*. 2020;20(1):1–10.
6. Bouchard L, Albertini M, Batista R, De Montigny J. Research on health inequalities: a bibliometric analysis (1966-2014). *Soc Sci Med*. 2015;141:100–108.
7. Dwivedi R, Pradhan J. Does equity in healthcare spending exist among Indian states? Explaining regional variations from national sample survey data. *Int J Equity Health*. 2017;16(1):15.
8. Iyer A, Sen G, George A. The dynamics of gender and class in access to health care: evidence from rural Karnataka, India. *Int J Health Serv*. 2007;37(3):537–554.
9. Awasthi A, Pandey CM, Chauhan RK, Singh U. Disparity in maternal, newborn and child health services in high focus states in India: a district-level cross-sectional analysis. *BMJ Open*. 2016;6(8):e009885.
10. Kulkarni VS, Kulkarni VS, Gaiha R. "Double burden of malnutrition" reexamining the coexistence of undernutrition and overweight among women in India. *Int J Health Serv*. 2017;47(1):108–133.
11. Kar S. Inclusive growth in hilly regions: priorities for the Uttarakhand economy. E/281/2007: IEG Working Paper, New-Delhi. 2007.
12. Mishra PS, Pautunthang N, Marbaniang SP. Geographical divide led inequality in accessing maternal healthcare services between hills and valley regions of manipur state, India. *Clin Epidemiol Glob Health*. 2021;11:100744.
13. Dasgupta P. Cost-disabilities of hill states in India, study report submitted to the fourteenth finance commission of India, Institute of economic growth. 2014. https://fincomindia.nic.in/writereaddata/html_en_files/oldcommission_html/fincom14/others/29.pdf
14. Arias-Reyes C, Zubieta-De Urioste N, Poma-Machicao L, et al. Does the pathogenesis of SARS-CoV-2 virus decrease at high-altitude? *Respir Physiol Neurobiol*. 2020;277:103443. doi:10.1016/j.resp.2020.103443
15. Segovia-Juarez J, Castagnetto JM, Gonzales GF. High altitude reduces infection rate of COVID-19 but not case-fatality rate. *Respir Physiol Neurobiol*. 2020;281:103494. doi:10.1016/j.resp.2020.103494
16. Xi A, Zhuo M, Dai J, et al. Epidemiological and clinical characteristics of discharged patients infected with SARS-CoV-2 on the qinghai plateau. *J Med Virol*. 2020; 92(11): 2528–2535. doi:10.1002/jmv.26032
17. Stephens KE, Chernyavskiy P, Bruns DR. Impact of altitude on COVID-19 infection and death in the United States: a modeling

and observational study. *PloS One*. 2021;16(1):e0245055. doi:10.1371/journal.pone.0245055

18. Wehby GL, Castilla EE, Lopez-Camelo J. The impact of altitude on infant health in South America. *Econ Hum Biol*. 2010;8(2):197–211. doi:10.1016/j.ehb.2010.04.002

19. Kesby JP, Eyles DW, Bume THJ, McGrath JJ. The effects of vitamin D on brain development and adult brain function. *Mol Cell Endocrinol*. 2011;347(1-2):121–127.

20. Cano-Pérez E, Torres-Pacheco J, Fragozo-Ramos MC, García-Díaz G, Montalvo-Varela E, Pozo-Palacios JC. Negative correlation between altitude and COVID-19 pandemic in Colombia: a preliminary report. *Am J Trop Med Hyg*. 2020;103(6):2347–2349.

21. Pun M, Turner R, Strapazzon G, Brugger H, Swenson ER. Lower incidence of COVID-19 at high altitude: facts and confounders. *High Alt Med Biol*. 2020;21(3):217–222. doi:10.1089/ham.2020.0114

22. Pandey J, Singh M. Donning the mask: effects of emotional labour strategies on burnout and job satisfaction in community healthcare. *Health Policy Plan*. 2016;31(5):551–562.

23. Chokshi M, Patil B, Khanna R, et al. Health systems in India. *J Perinatol*. 2016;36(S3):S9–S12. doi:10.1038/jp.2016.184

24. Barah BC. Hill agriculture: problems and prospects for mountain agriculture. *Indian J Agric Econ*. 2010;65(3):584–601.

25. Thayer WM, Hasan MZ, Sankhla P, Gupta S. An interrupted time series analysis of the lockdown policies in India: a national-level analysis of COVID-19 incidence. *Health Policy Plan*. 2021;36(5): 620–629. doi:10.1093/heapol/czab027

26. Hsiao CY, Jha P, Gambhir S, Fu SH. Geospatial relationship between altitude and mortality in India. *Int J Geoinformatics*. 2012;8(3).

27. Daoust JF. Elderly people and responses to COVID-19 in 27 countries. *PloS one*. 2020;15(7): e0235590. doi:10.1371/journal.pone.0235590

28. Mishra S, Gayen A, Haque MS. COVID-19 and urban vulnerability in India. *Habitat Int*. 2020;103. doi:10.1016/j.habitatint.2020.102230

29. Black D. Smoking and the black report. *Lancet*. 1981;317(8214):280.

30. Gray AM. Inequalities in health. The black report: a summary and comment. *Int J Health Serv*. 1982;12(3):349–380. doi:10.2190/XXMM-JMQU-2A7Y-HX1E

31. Jalan J, Ravallion M. Does piped water reduce diarrhoea for children in rural India? *J Econom*. 2003;112(1):153–173.

32. Liu Z, Li J, Rommel J, Feng S. Health impacts of cooking fuel choice in rural China. *Energy Econ*. 2020;89:104811. doi:10.1016/j.eneco.2020.104811

33. Tran BX, Nguyen LH, Nong VM, Nguyen CT. Health status and health service utilization in remote and mountainous areas in Vietnam. *Health Qual Life Outcomes*. 2016;14(1):1–9.

34. Asher S, Lunt T, Ryu M, Paul N. COVID-19 data resources. 2020.

35. Rathore U, Das U, Sarkhel P. Birds of a feather flock together? Diversity and spread of COVID-19 cases in India. Working Paper Series 2021-053, Global Development Institute, University of Manchester.

36. MaCurdy TE, Pencavel JH. Testing between competing models of wage and employment determination in unionized markets. *J Polit Econ*. 1986;94(3):S3–S39.

37. Munshi K, Rosenzweig M. Why is mobility in India so low? Social insurance, inequality, and growth. National Bureau of Economic Research. Working Paper No. w14850. 2009.

38. Roychowdhury P. Peer effects in consumption in India: an instrumental variables approach using negative idiosyncratic shocks. *World Dev*. 2019;114:122–137.

39. Altonji JG, Elder TE, Taber CR. Selection on observed and unobserved variables: assessing the effectiveness of catholic schools. *J Polit Econ*. 2005;113(1): 151–184.

40. Oster E. Unobservable selection and coefficient stability: theory and evidence. *J Bus Econ Stat*. 2019;37(2):187–204. doi:10.1080/07350015.2016.1227711

41. Rathore U, Das U. Health consequences of patriarchal kinship system for the elderly: evidence from India. *J Dev Stud*. 2021;145–163. doi:10.1080/00220388.2021.1939863

42. Jha P, Deshmukh Y, Tumbe C, et al. COVID Mortality in India: national survey data and health facility deaths. *Science*. 2022: eabm5154. doi:10.1126/science.abm5154

43. Planning Commission, Government of India. Report of the task group on problems of hilly habitations in areas covered by the hill areas development programme (HADP/Western Ghats Development Programme (WGDP). 2008. Accessed June 29, 2021. http://www.indiaenvironmentportal.org.in/files/tg11_hillarea.pdf

44. Breevoort A, Carosso GA, Mostajo-Radji MA. High-altitude populations need special considerations for COVID-19. *Nat Commun*. 2020;11(1):1–3. doi:10.1038/s41467-020-17131-6

45. Jayanta S, Bhattacharyya R. Half empty or half filled? Notes on universal health coverage in northeast India. *The Clarion*. 2015;4(1):154–184. https://ssrn.com/abstract=2577328

46. Mustafa A, Shekhar C. Contrast in utilization of maternal and child health services between himalayan region and rest of India: evidence from national family health survey (2015-16). *BMC Pregnancy Childbirth*. 2021;21(1):606. doi:10.1186/s12884-021-04081-0

47. Programme Evaluation Organisation, Planning Commission, Government of India. Report of the evaluation study on hill area development programme in Assam and West Bengal program evaluation organization. 2010. Accessed July 15, 2021. Report of the Evaluation Study on Hill Area Development Programme in Assam and West Bengal.pdf (dmeo.gov.in)

48. Van den Broucke S. Why health promotion matters to the COVID-19 pandemic, and vice versa. *Health Promot Int*. 2020;35(2):181–186. doi:10.1093/heapro/daaa042

Author Biographies

Sushmita Chakraborty has a PhD in human geography and is currently an independent researcher. Her research is primarily focused on environmental issues in social science, health geography, and application of Remote Sensing and Geographical Information System models in social science. (ORCID: 0000-0002-0731-4374)

Upasak Das is a Presidential Fellow of Economics of Poverty Reduction at the Global Development Institute at the University of Manchester. He is also an affiliate of the Center for Social Norms and Behavioral Dynamics at the University of Pennsylvania. His primary research interests include development economics, health, education, social norms, and social protection programs. (ORCID: 0000-0002-4371-0139)

Udayan Rathore is an economist by training. He is currently a consultant with OPM India's Research and Evidence team (formerly Statistics, Evidence, and Accountability Programme [SEAP]). His areas of interest include research on livelihood, health, education, and social safety net programs. (ORCID: 0000-0003-1941-1005)

Prasenjit Sarkhel is an Associate Professor, Department of Economics, University of Kalyani. His current areas of research include public policy for health care and education, impact of environmental regulation, and non-market valuation. (ORCID: 0000-0001-5486-6959)