The University of Manchester
Global Development Institute

Global Development Institute

Working Paper Series

2020-047

October 2020

ISBN: 978-1-912607-05-1

Severity of the Covid-19 pandemic in India. The case of three states: Maharashtra, Jharkhand and Meghalaya

Nidhi Kaicker[1]

[1] Ambedkar University, Delhi, India

Email: nidhi@aud.ac.in

Katsushi S. Imai[2]

[2] Department of Economics, University of Manchester, UK

Email: katsushi.imai@manchester.ac.uk

Raghav Gaiha[3]

[3] GDI, University of Manchester, UK; Population Studies Centre, University of Pennsylvania,Philadelphia, USA

Email: rgaiha@sas.upenn.edu

Cite this paper as:

Kaicker, N., Imai, Katsushi S. and Gaiha, R. (2020) Severity of the Covid-19 pandemic in India. The case of three states: Maharashtra, Jharkhand and Meghalaya. GDI Working Paper 2020-047. Manchester: The University of Manchester.

Abstract

This is the first econometric analysis of the severity of the Covid-19 pandemic measured using two related but distinct measures of mortality up to 21 June 2020.One is the Cumulative Severity Ratio(CSR) and the other is the Daily Severity Ratio (DSR). The CSR measures the additional pressure on India's fragile and ill-equipped healthcare system, the DSR helps monitor the progression of fatalities. Another important contribution of this analysis is the use of rigorous econometric methodology: random-effects models and Hausman–Taylor models. Although the rationales vary, they yield a large core of robust results. The specifications are rich and comprehensive, despite heavy data constraints. The factors associated with the CSR and DSR include (lagged) Covid-19 cases, income, age, gender, multi-morbidity, urban population density, lockdown phases within three states, Maharashtra, Jharkhand and Meghalaya, and weather, including temperature and rainfall and their interactions with the two state dummies. Given the paucity of rigorous econometric analyses, our study yields policy insights of considerable significance.

Keywords

Covid-19, Cumulative Severity Ratio, Daily Severity Ratio, random-effects model, Hausman–Taylor model, Jharkhand, Maharashtra, India

JEL Codes

C23, I18, N35, O10

Acknowledgements

This study was funded by the Asia & the Pacific Division, IFAD. We are grateful to the Country Director, Ms Rasha Omar, for her support and advice; we are also indebted to Radhika Aggarwal for her outstanding research support. We also appreciate the advice of two anonymous reviewers and Frances Brown's meticulous editing. The views expressed here are personal and not necessarily those of IFAD.

1 Introduction

On 21 June 2020, the cut-off date for this analysis, the total confirmed Covid-19 cases in India crossed the 400,000 mark and the total deaths stood at more than 13,000. There are significant geographical variations, however. One state alone (Maharashtra) recorded close to a third of the total cases, and 45% of the total deaths. Our focus in the present study is on the *severity* of the Covid-19 pandemic in three Indian states, namely, Maharashtra, Jharkhand and Meghalaya, to contrast the former and the latter two. In addition, we carry out regression analyses for 27 Indian states to understand the pandemic in these states in the national context, in order to identify the determinants of the Covid-19 pandemic both nationally as well as in the three states under scrutiny.[1]

Maharashtra, home to around 10% of India's total population,[2] and classified as one of the richest states–based on per capita income–has recorded the highest number of cases and deaths linked to the Covid-19 virus so far. Jharkhand is home to around 2% of the country's population and is classified as one of the poorest states, based on per capita income. The number of confirmed Covid-19 cases in the state is around 2,000 and the number of linked deaths is negligible. Meghalaya is a small state in northeast India, home to only 0.2% of the total population of the country, with a moderate per capita income. This state had recorded only one death up to 21 June 2020, and fewer than 50 confirmed cases.

While these three states are far from representative of India, we will carry out statistical analyses to contrast the statistics of the severity of the Covid-19 pandemic in these states. Our focus on these states is guided primarily by data availability, but we also attempt to describe how the Covid-19 pandemic developed in Maharashtra by comparing its trend with those of Jharkhand and Meghalaya. This descriptive analysis is extended to the panel data analyses, where Maharashtra is essentially contrasted not only with Jharkhand and Meghalaya but also with 24 other states, to see how its situation has been specific, or generalised. As there have not been many studies on the determinants of Covid-19 infections in developing countries outside China, our study will provide important policy insights for policy makers not only in India but in other developing countries.

The research questions we will ask in this paper are as follows:

(1) How has the Covid-19 pandemic developed differently at state levels, namely, in Maharashtra, Jharkhand and Meghalaya? How was the development of Covid-19 in Maharashtra different from its development in Jharkhand/Meghalaya or other Indian states?

[1]The states we included in our regression analyses are Andhra Pradesh, Arunachal Pradesh, Assam, Bihar, Chhattisgarh, Goa, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Manipur, Meghalaya, Mizoram, Nagaland, Odisha, Punjab, Rajasthan, Tamil Nadu, Tripura, Uttar Pradesh, Uttarakhand and West Bengal. The selection was determined by the availability of the data.
[2]Based on 2011 census estimates.

(2) What are the determinants of the severity of the Covid-19 pandemic in India?

The rest of the paper is organised as follows. The next section reviews the emerging literature on the Covid-19 pandemic. Section 3 offers a detailed statistical description of the pandemic in Maharashtra, Jharkhand and Meghalaya. Section 4 defines the severity ratios we use in the present study to capture the severity of the pandemic. Section 5 specifies the econometric models we employ to assess this severity. Section 6 summarises the main econometric results. Section 7 focuses on the significance of our findings from a broader policy perspective. Finally, Section 8 delineates the policy challenges.

2 Literature review

Although there has been a spate of global and multi-country studies–notably, Brown et al (2020), IMF (2020), Jorda et al (2020), Oldekop et al (2020) and Vadlamannati et al (2020) – and a few on India –notably, Ray and Subramanian (2020), Sen (2020), Banerjee et al (2020) and Joe et al (2020)-our literature review is confined to selected studies focused on India.

Of particular interest is Joe et al (2020), as it presents a detailed statistical study of factors associated with Covid-19 mortality. The authors used crowd-sourced data to provide preliminary estimates of the age- and sex-specific Covid-19 case fatality rate (CFR) for India.[3]The CFR was estimated as the ratio of confirmed deaths to total confirmed cases. Binomial confidence intervals were given for the CFR estimates. Also, an adjusted-CFR was estimated to capture the potential mortality among the currently active infections. The authors' main findings are as follows. As of 20 May 2020, males shared a higher burden (66%) of Covid-19 infections than females (34%). However, the infection was more or less evenly distributed between males and females in the under-five as well as elderly age groups, where the CFR among males and females was 2.9% and 3.3%, respectively. The age-specific Covid-19 CFR reflects a 'Nike-swoosh' pattern with elevated risks among the elderly. The World Health Organization world standard population structure standardised CFR for India was 3.34%. The adjusted CFR was estimated to be 4.8%. Early evidence thus indicates that males have a higher overall burden, but females have a higher relative risk of Covid-19 mortality in India. Elderly males and females both display a high mortality risk and require special care when infected. As the period which this study covers ended on 20 May, well before the huge surge in cases and, to a lesser extent, in fatalities, there is a need for coverage of a more recent period. We have thus carried out our analysis covering the more recent period (up to 21 June), which should be more rewarding from a policy perspective. Panel models with fixed or random effects that allow for the lockdown phase and unobservable state effects, as we have used, are likely to yield more robust results and policy insights.

[3] See https://www.covid19india.org/.

Ray and Subramanian (2020) asked an important question: why is there such a forceful emphasis on a draconian lockdown? Their conjecture is that it is precisely because Covid-19 poses a visible threat that Indian elites all know about, and are vulnerable to it. In catering to these fears, the Indian government is also seeking international recognition as a leader against the pandemic. For the elites, a lockdown often implies a loss of cleaners and maids. But there is a huge price to be paid, in the horrors and privations that the poor and truly vulnerable in the country must suffer as their livelihoods fail, often leading to an 'invisible'death. According to Ray and Subramanian, this invisibility is only in part a result of poverty. In greater part, it is because non-Covid deaths are diffuse; they are classified under various headings, and so lack the capacity to attract focal attention. Our understanding of media reports differs, however, as horrific images of the dead lying next to Covid-19 patients– and abandoned at cremation grounds –following the rapid surge in infections in late May–June, are flashed relentlessly across TV channels, imparting greater visibility. So perhaps the distinction between the visible and invisible is overdrawn.

In an innovative contribution, Banerjee et al (2020) conducted a large-scale messaging campaign in West Bengal. Twenty-five million individuals were sent an SMS containing a 2.5-minute clip. All messages encouraged people to report symptoms to the local public health worker. Messages were randomised at the PIN-code level. As a control, three million individuals received a message directing them to government information. The campaign, first, doubled the reporting of health symptoms to the community health workers; second, reduced travel beyond people's home village in the previous two days and increased estimated hand-washing when returning home; third, spilled over to behaviours not mentioned in the message – for example, mask-wearing increased slightly, while distancing and hygiene increased in the sample where they were not mentioned by similar amounts to where they were mentioned; and fourth, spilled over onto non-recipients within the same community, with effects similar to those for individuals who received the messages.

Our study builds on Joe et al (2020) in some important ways. First, it extends the analysis to 21 June 2020, and thus succeeds in capturing the surge in the pandemic during late May and June. We used two distinct but related measures of Covid-19 severity, namely, the cumulative severity ratio (CSR) and the daily severity ratio (DSR).The CSR takes Covid-related deaths over a period since the occurrence of the first death relative to deaths in a pre-pandemic year over the same duration. This unravels the cumulative pressure on India's fragile and ill-equipped health system. The DSR, on the other hand, is a flow measure of the pandemic based on the new Covid-related deaths in comparison with the deaths in a pre-pandemic year. It helps us monitor the progression of the pandemic –whether it is intensifying, weakening or unchanged. We used panel models that allowed us to use random- effects and the Hausman–Taylor model that allowed the use of time-invariant fixed effects. Their results are largely similar, implying a large core of robust data. Our specifications included state-wise lockdown effects, and other interactions of states with, say, per capita income. Moreover, we took into account the association of Covid severity with temperature and rainfall. Thus, our study yields rich and robust policy insights.

3 Descriptive analysis

The numbers of Covid-19 cases and linked deaths for the three states (see Figures 1, 2 and 3) suggest an unequal distribution of the impact of the pandemic, which raises the following questions. First, how unequally severe has the pandemic been so far across states over time? Second, what are the explanations behind these contrasts? These questions will be addressed by detailed econometric analyses for the three states, as well as in the panel regression analyses for 27 states.

The roman numbers in each of the figures indicate the phases of the various lockdowns imposed by the government during this period. The first lockdown spanned a period of 21 days, from 25 March to 14 April 2020, when nearly all factories and services were suspended, barring 'essential services'. The second started on 15 April 2020 and continued until 3 May, with conditional relaxations for regions where the Covid-19 spread had been contained. With additional relaxations, phase three of the lockdown ran from 4 to 17 May, while the fourth phase ran from 18 May to 21 June. Phase 5 of the lockdown (1–30 June), also known as 'Unlock 1.0' was the first phase of the reopening in stages, with an economic focus.[4]

[4]Our analyses range from the period starting 13 March 2020 (pre-lockdown) until 21 June 2020. Hence, wherever mentioned, phase 5 of the lockdown refers to the data for the first three weeks of June only. As the panel regression analyses use the daily weather data, we use the data up to 21 June 2020.

Figure 1: Covid Cases & Deaths in Jharkhand

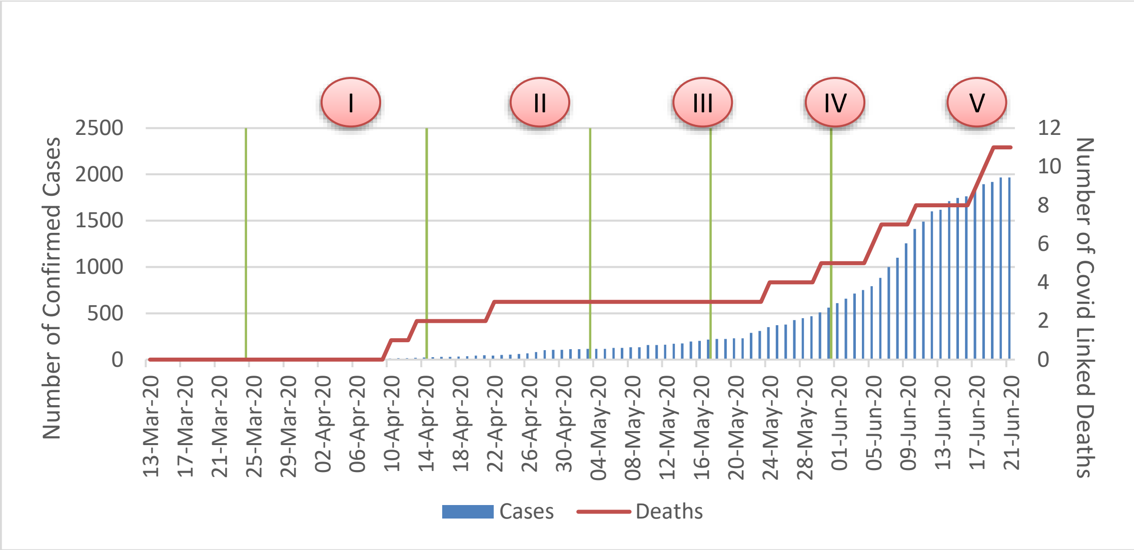


Figure 2: Covid Cases & Deaths in Maharashtra

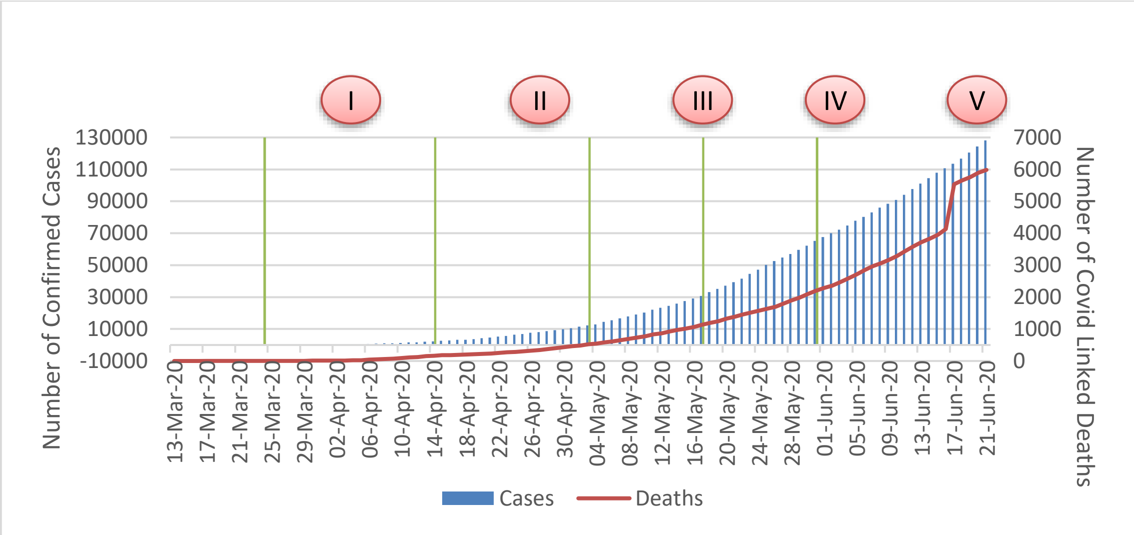
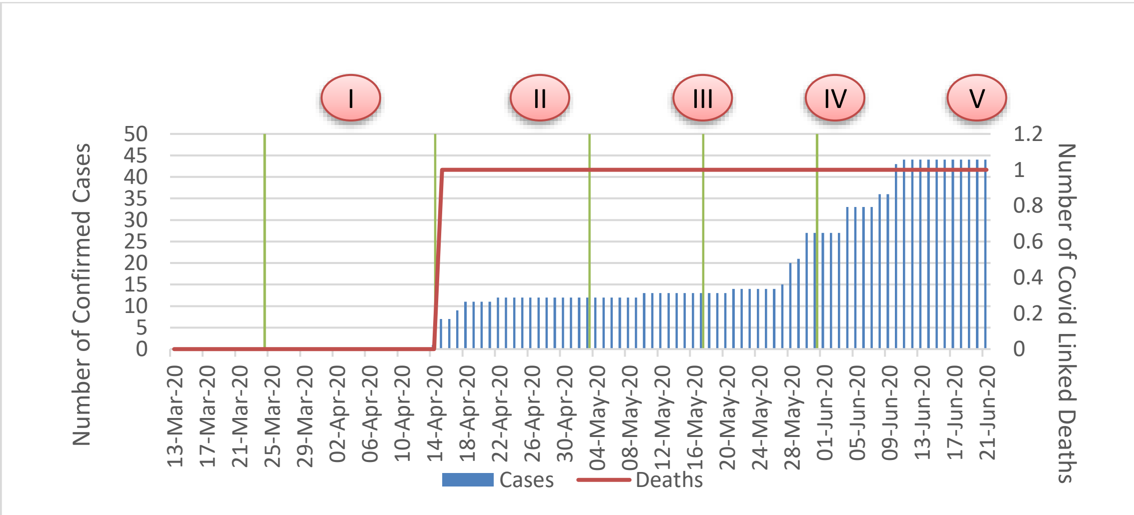


Figure 3: Covid Cases & Deaths in Meghalaya



4 Definitions of severity ratios

A new indicator, 'relative severity', proposed by the World Bank, is being used to illustrate the unequal distribution and progression of Covid-19 deaths across countries.[5] The relative severity ratio is defined as the ratio of the total deaths attributable to Covid-19 over a given period to the expected total deaths from all causes under the counterfactual assumption that the pandemic had not taken place, over a base period of the same length. Comparison with pre-pandemic mortality patterns provides a state-specific measure of the severity of the pandemic, and the excess burden on the health system. In addition to this ratio (which will be denoted CSR), a DSR is also calculated which tracks the progression of the severity of the pandemic in each region. To calculate the daily severity ratios, the number of Covid-19 deaths on a particular day are divided by the expected daily deaths under the assumption of no-pandemic, ie annual deaths divided by 365 (in a pre-Covid year).

Algebraically,

$$Cumulative\ Severity\ Ratio_t = \frac{Cumulative\ Covid\ Deaths_t}{(\frac{No.of\ Deaths\ in\ a\ pre\ pandemic\ year}{365} * Length\ of\ Pandemic_t)},$$

where

$$Length\ of\ Pandemic_t = No.of\ Days\ between\ Date\ of\ First\ Covid\ Linked\ Death\ and\ t\ in\ the\ Region$$

$$Daily\ Severity\ Ratio_t = \frac{New\ (daily)\ Covid\ Deaths_t}{(\frac{No.of\ Deaths\ in\ a\ pre\ pandemic\ year}{365})}$$

The Covid-19 data were collated from the Ministry of Health and Family Welfare. The data on past mortality patterns are based on the 2017 state-wide number of registered deaths from the same ministry. For the purpose of the CSR, the number of reported deaths in 2017 was scaled down from annual estimates to the length of the pandemic in each state, calculated as the number of days since the first death there until the cut-off date for this analysis, ie21 June 2020. For instance, in Maharashtra, the first death was reported on 17 March 2020, implying a 97-daylength for the pandemic. The expected total deaths under the no-pandemic situation were calculated as the total number of deaths in each region in 2017 * 97 days/365.For the DSR, the denominator used in the ratio was total number of deaths in each region in 2017/365.[6]

[5]For details, see Schellenkens and Sourrowuille (2020).

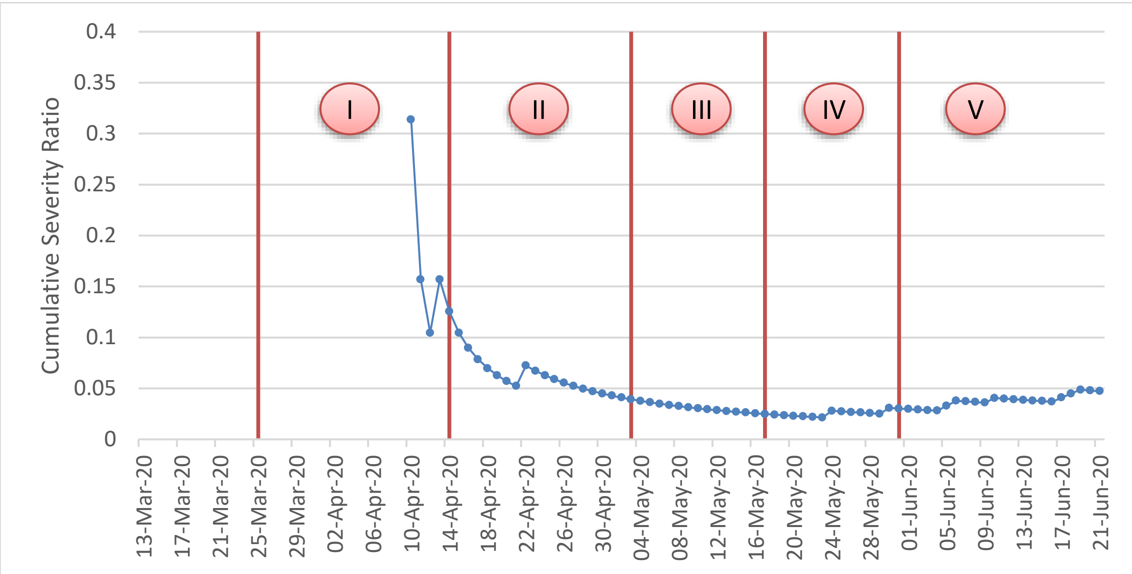
[6]One question is whether the death numbers in 2017 can serve as a valid counterfactual. First, the national-level death rate has been fairly stable and gradually declining from 7.4 to 7.3 deaths per 1,000 population since 2012; 2017 was not an exceptional year. Second, while India has experienced frequent and widespread droughts, there was no major drought in 2017. The death

4.1 CSR–trends

Figure 4 shows the trend in the CSR for Jharkhand. The first Covid-19 linked death was recorded on 10 April 2020. The CSR showed a downward trend in April and May, and a rising one after the completion of the fourth lockdown. The average CSR values for Jharkhand were 0.17% during the first lockdown phase, decreasing to 0.06% during the second lockdown and stable at around 0.03% in the third and fourth lockdown phases, as well as during the first three weeks of the fifth phase (the Unlock 1.0 period).

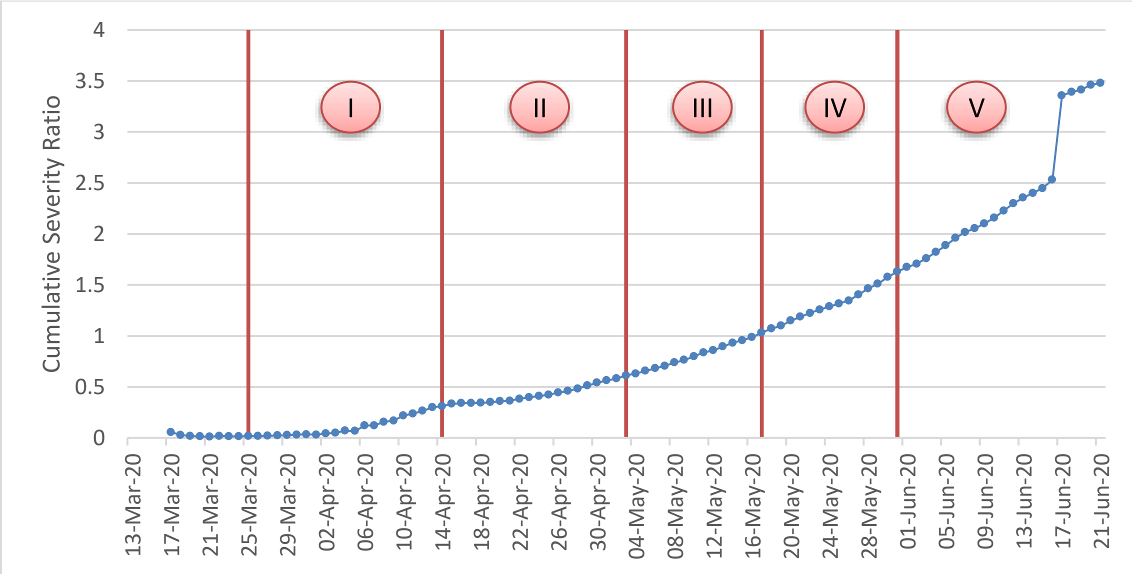
As can be seen from Figure 5, in Maharashtra, the cumulative severity ratio rose gradually during the first two lockdown phases, and more swiftly during the third and fourth lockdowns. There was a sudden spike in the number of deaths on 17 June, as a result of which the cumulative ratio curve spiked upwards. The average CSR values for Maharashtra were 0.11% during the first lockdown phase, 0.43% during the second, 0.82% during the third, 1.33% during the fourth and 2.41% in the first three weeks of the Unlock 1.0 phase.

Figure 4: Cumulative Severity Ratio in Jharkhand



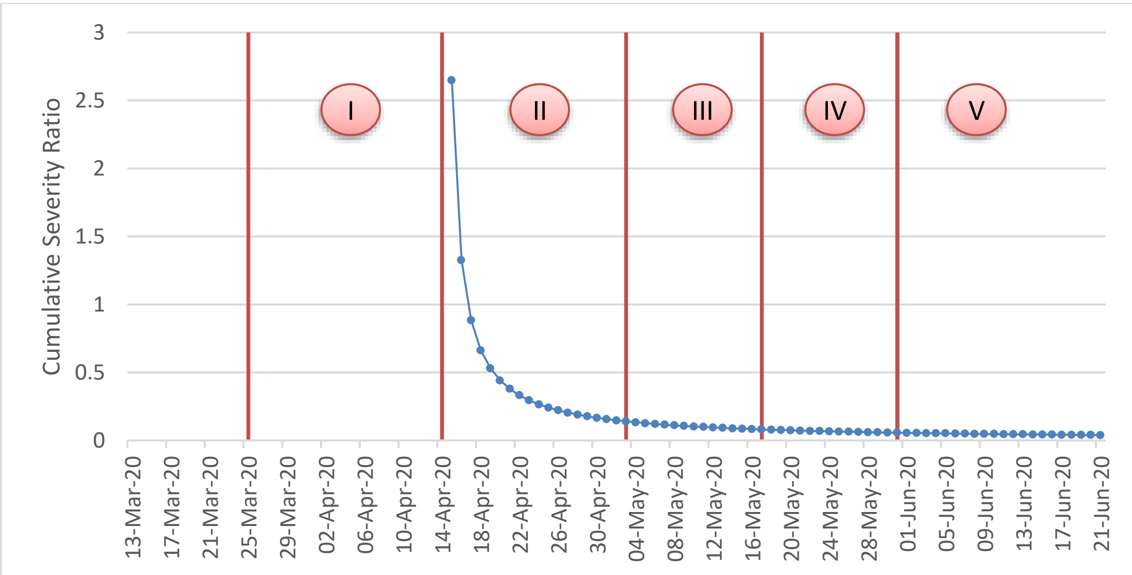
numbers in 2017 thus serve as a reasonable counterfactual for the present analysis of Covid-19.https://www.indexmundi.com/, accessed 18 July 2020.

Figure 5: Cumulative Severity Ratio in Maharashtra



The CSR curve for Meghalaya represents a rectangular hyperbola (see Figure 6), primarily because the first and only death up to 21 June in that state took place on 17 April. Since then, as the length of the pandemic has increased, the CSR has declined because of a constant numerator (number of cumulative Covid-19-linked deaths), and a higher denominator thanks to the longer duration (number of deaths in a pre-pandemic year in a period of the same duration).

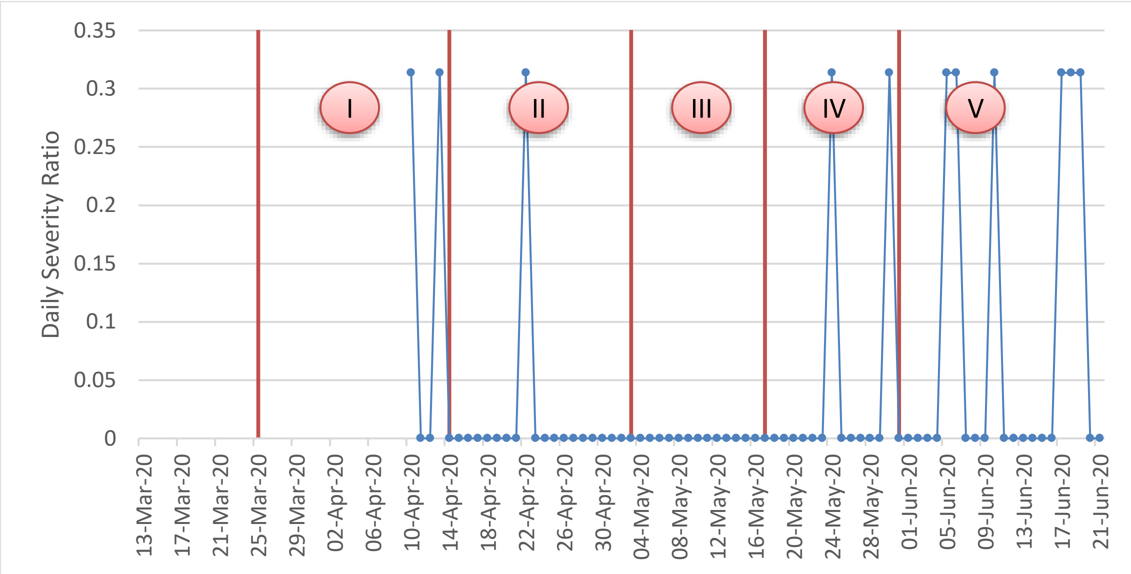
Figure 6: Cumulative Severity Ratio in Meghalaya



4.2 Progression in the DSR

Figures 7, 8 and 9 show the progression in severity ratios for the three states from the first death in each state until the cut-off date for this analysis.

Figure 7: Daily Severity Ratio in Jharkhand



If we look at the DSR graph for Jharkhand, we understand that only 11 deaths have been reported so far, and no more than one death has been reported on any single day.

The DSR graph for Maharashtra reveals the striking progression of the pandemic. In the first lockdown phase, the average DSR was 0.42%, it then crossed the 1% mark and reached a maximum of 1.41% during this phase. In the second to fourth lockdown phases and the fifth Unlock 1.0 phase, there was not a single day when zero deaths were reported. In fact, the new deaths reported on each day rose swiftly. In the second lockdown phase, the average DSR was 1.07% and the maximum was 2.03%. These numbers rose to 2.47% (average) and 2.78% (maximum) in the third lockdown phase, and almost doubled to 4.28% (average) and 6.54% (maximum) in the fourth phase. In the first three weeks of the fifth phase, Unlock 1.0, an explosion in the number of deaths occurred, taking the average to 10.17% and the maximum to 79.47% (truncated in the graph).

Figure 8: Daily Severity Ratio in Maharashtra

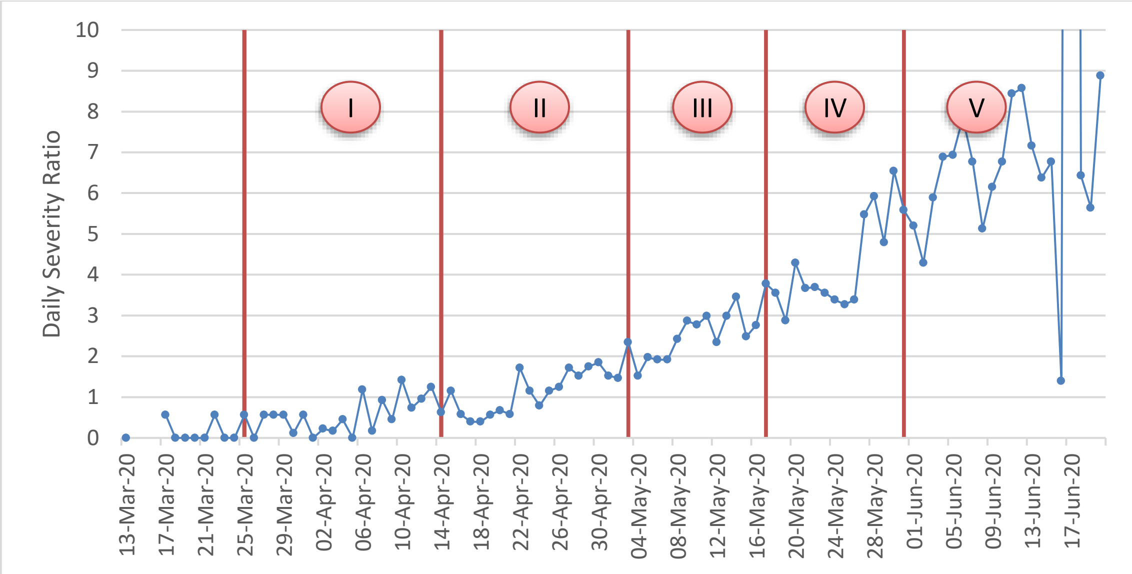
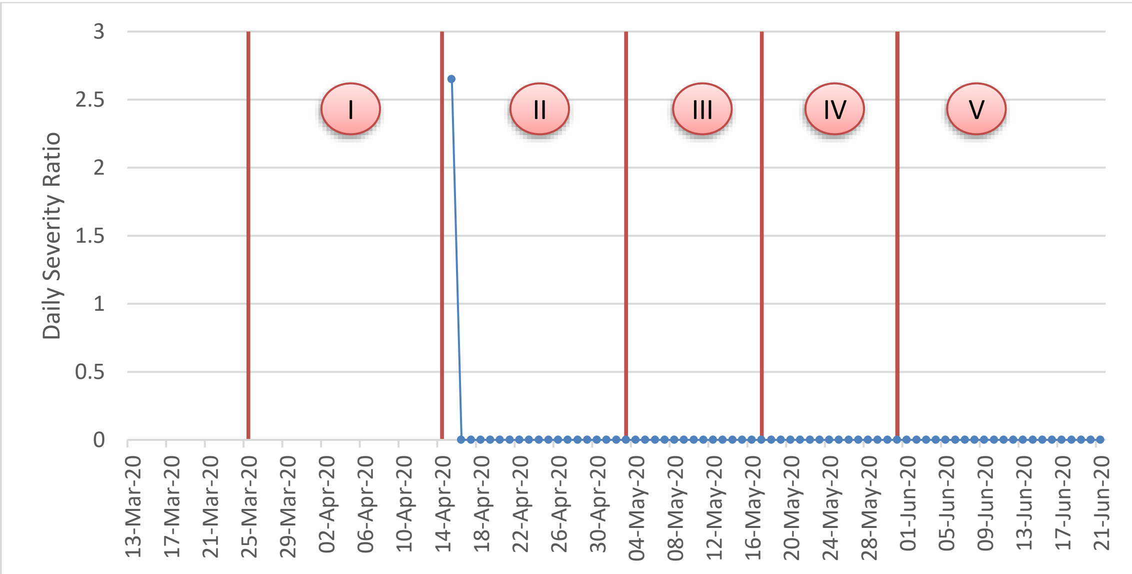


Figure 9: Daily Severity Ratio in Meghalaya



In Meghalaya, only one death has been reported; hence, the daily severity ratio is zero for the entire period barring the day on which the death was reported.

5 Regression analyses on the determinants of the severity of the Covid-19 pandemic

5.1 Model specification

To find an explanation for the regional variation in the severity of the Covid-19 pandemic, we used a panel of 27 states for which the data on various variables are available, covering the period from 13 March to 21 June 2020. The data are organised as weekly panel data, where the daily data are averaged for each week. We regressed a dependent variable, either the CSR or the DSR, on the lagged value of the confirmed Covid-19 cases and a set of lagged time-invariant explanatory factors such as per capita income, urbanisation, population density (a proxy for social distancing and likely contagion effects), and presence of morbidity conditions in the aged (60 years and older/state population). We used the interactions of time dummy variables (with Phases 2, 3, 4 and 5) and the dummy variables for Jharkhand and Maharashtra to compare the severity of the Covid-19 pandemic with other Indian states.[7]The CSR captures the overall development of the Covid-19 pandemic, while the DSR denotes how the severity progresses over time.

Methodologically, we employed random-effects and Hausman–Taylor models (Hausman &Taylor, 1981;Baltagi, 2008) since the fixed-effects model cannot have time-invariant variables as explanatory variables, as it involves the transformation of deriving a deviation from the time-series means of dependent and explanatory variables. A major advantage of using the fixed-effects model is that the unobservable individual effect (in our case, the unobservable effect specific to each state, such as cultural factors or unobservable institutional qualitiy determining the quality of political systems or health services) can be correlated with (time-variant) explanatory variables. A further advantage of a fixed-effects model is that the model allows the correlation of these factors while providing unbiased estimates. However, the fixed-effects model is not suitable for our study, as most of the explanatory variables are time-invariant. As an alternative, the random-effects model assumes that the unobservable individual effect at the state level is *uncorrelated with* explanatory variables. If this assumption does not hold, the estimated coefficients of the random-effects model are biased. A standard procedure is to choose either a fixed-effects model or a random-effects model, based on the Hausman test to compare the estimated coefficients but, if we need to include a number of time-invariant variables in the model, we cannot rely on this test, as the comparison is made for only a subset of explanatory variables of the random-effects model.

In the context of the present study, we have estimated the Hausman–Taylor model. This model allows us to have the time-invariant variables as part of the explanatory variables, while it allows part of the explanatory variables – both time-variant and time-invariant – to be correlated with the error term or made endogenous, as in Equation (2) below. The Hausman–Taylor model draws upon both random- and fixed-effects models and applies an Instrumental Variable (IV) estimation to instrument endogenous variables by instrumental variables based on the deviations from the time-series mean of time-variant exogenous variables, or the mean of time-variant variables.

There are two advantages in applying the Hausman–Taylor model in the present context. First, it is likely that some of the explanatory variables (eg., state-level per capita income) are correlated with the unobservable state-level effect, such as cultural factors; if so, the estimated coefficients of the random-effects model, which assumes no correlation of these variables, may be biased. The Hausman–Taylor model can address this type of endogeneity in the model. Second, because the fixed-effects model cannot be used as

[7]Because observations of a few variables are missing, our estimation is based on 377 observations for the weekly panel (27 states x 13.96 weeks).

a proper baseline model to be compared with the random-effects model in the presence of time-invariant variables, the Hausman–Taylor model can serve as an alternative to the fixed-effects model to decide which model is likely to yield unbiased estimators. We use the Hausman test to compare the random-effects model and the Hausman–Taylor model by hypothesising that the random-effects model is unbiased. We can then decide whether the random-effects model is likely to be unbiased if the vectors of estimated coefficients of the two models are not systematically different. In other words, the Hausman–Taylor model can serve as a robustness check of the random-effects model in the presence of time-invariant explanatory variables.

We estimate the following model by the random- effects model to estimate the logarithm of the CSR as in Equation (1) as well as that of the DSR as in Equation (2). We also take the logarithm of most of the explanatory variables to capture the relative effect, or the elasticity of each determinant.[8]

$$\log CSR_{it} = \beta_0 + \beta_1\, log\ CovidCases_{it-7} + \beta_2 Per\ Capita\ Income_i + \beta_3[Per\ Capita\ Income_i * D_{Maharashtra_t}] + \beta_4[Per\ Capita\ Income_i * D_{Jharkhand_t}] + \beta_5 \log Urban\ Population\ Density_i + \beta_6 \log Rate\ of\ Multi-morbility_i + \beta_7 \log Sex\ Ratio_i + \beta_8 Temperature_{it} + \beta_9 Rainfall_{it} + Phase\ Dummies_t\, \beta_{10} + State\ Dummies_t\, \beta_{11} + [Phase\ Dummies_t * D_{Maharashtra_t}]\, \beta_{12} + [Phase\ Dummies_t * D_{Jharkhand_t}]\, \beta_{13} + [Phase\ Dummies_t * D_{Meghalaya_t}]\, \beta_{14} + \mu_i + e_{it} \ldots\ldots\ldots (1)$$

$$\log DSR_{it} = \beta_0 + \beta_1\, log\ CovidCases_{it-7} + \beta_2 Per\ Capita\ Income_i + \beta_3[Per\ Capita\ Income_i * D_{Maharashtra_t}] + \beta_4[Per\ Capita\ Income_i * D_{Jharkhand_t}] + \beta_5 \log Urban\ Population\ Density_i + \beta_6 \log Rate\ of\ Multi-morbility_i + \beta_7 \log Sex\ Ratio_i + \beta_8 Temperature_{it} + \beta_9 Rainfall_{it} + Phase\ Dummies_t\, \beta_{10} + State\ Dummies_t\, \beta_{11} + [Phase\ Dummies_t * D_{Maharashtra_t}]\, \beta_{12} + [Phase\ Dummies_t * D_{Jharkhand_t}]\, \beta_{13} + [Phase\ Dummies_t * D_{Meghalaya_t}]\, \beta_{14} + \mu_i + e_{it} \ldots\ldots\ldots (2)$$

Here i stands for states (from 1 to 27 – some states were dropped because of lack of availability of a few explanatory variables) – and t for the weeks from 13 March to 21 June 2020 (from the first week to the 14th) for the weekly panel data. The choice of explanatory variables was guided by several studies in the literature. First, socioeconomic factors were identified as important determinants (see, for example, Ehlert, 2020). Second, in addition to the socioeconomic factors, weather conditions are thought likely to influence the severity of the Covid-19 pandemic (Ma et al, 2020; Tosepu et al, 2020). Third, we have considered how past Covid-19 infections dynamically influenced the development of the pandemic.

[8] For the definitions of the variables used, see Appendix Table 1.

In the above equations $logCovidCases_{it-1}$ stands for the log of the first lag of Covid-19 infection. This is lagged by a week in order to consider the time-lag between infections and deaths.[9] The lag between the day when the infection was recorded and that when the death was recorded may be more than a week, but this variable broadly captures how the number of infections results in the severity of the pandemic (i.e. the CSR and DSR). In order to avoid a reduction in the number of observations, we took a lag for the weekly panel. In the Hausman–Taylor model this variable is treated as an endogenous variable to take into account the simultaneity in the occurrence of infection cases and the occurrence or the accumulation of deaths. $Per\ Capita\ Income_i$ (PCI) denotes income at state level that is measured by per capita net state domestic product (in Rs, divided by 1,000).[10] PCI does not only capture overall economic development at state levels. It may also capture health infrastructure or funding at state levels – for which the data are unavailable – in response to the Covid-19 pandemic. Given our study objectives, PCI is

[9]The individual state estimates of daily confirmed Covid-19 Cases are taken from the official website of the Ministry of Health and Family Affairs, Government of India.

[10] The data on state per capita incomes were obtained from the state economic surveys, while demographic data, including population density and urban population were taken from the census estimates.Initially, PCI is not logged to insert the squared term to capture the non-linear relationship. However, we decided to include only $Per\ Capita\ Income_i$ as including both the level and squared term, or only the log of income, makes estimated coefficients difficult to interpret thanksto mild correlations with various covariates (state dummies, interactions, lagged Covid cases, multi-morbidity, urban population density). While there are potentially many more determinants of CSR and DSR, given the availability of the data, adding more variables would make the regression results suffer from a typical symptom of multi-collinearity (eg statistically insignificant or unusually highly significant estimates). The correlation matrix for the weekly panel shows that the highest pairwise correlation coefficient is 0.52 (see Appendix Table 2). The variance inflation factor (VIF) test for the pooled ordinary least squares (OLS) with state dummies shows that the highest VIF is 4.39 (weekly cases), less than 5, implying that the regressions results do not significantly suffer from the multi-collinearity problem.

interacted with a dummy variable for the state of Maharashtra, or Jharkhand, to capture the extent to which the effect of income on the pandemic is different for these states.[11]

To capture the population density in urban areas, we also inserted$\log Urban\ Population\ Density_i$ the log of population density (inhabitants per square kilometer) in urban areas. The idea was that a higher population density and urbanisation would increase the interactions among people and raise both the CSR and DSR.

We also included the logarithm of the rate or proportion of elderly people suffering from more than one non-communicable disease (NCD) at state levels ($\log proportion\ of\ Multi - morbidity$).This is the proportion of population in the age group 60+ reporting more than one NCD (eg cardiovascular diseases, diabetes, hypertension, among others).[12] Furthermore, we have inserted $log\ Sex\ Ratio_i$ (the number of females per thousand males), as it is well documented that, while Covid-19 infection rates are broadly similar between men and women, men have been more likely to suffer from severe illness or die as a result of Covid infections in China (Jin et al, 2020) and Europe (Gebhard et al, 2020). However, given the preference for boys over girls in many states in India, more developed states with a lower poverty rate (eg Kerala) tend to have a higher sex ratio (more girls than boys),as well as a better health system. So the effect of the sex ratio on Covid-19 may be ambiguous in India.

It is widely debated whether the weather influences Covid-19 infections or deaths. A recent study used the data on daily death numbers from Wuhan, China, in January–February 2020 and found that death counts were positively associated with temperature and negatively with relative humidity (Ma et al, 2020).Tosepu et al (2020) found a positive association between temperature and the Covid-19 pandemic in Indonesia. We collected the daily data on temperature, rainfall and relative humidity from MERRA (Modern-Era Retrospective analysis for Research and Applications – Version 2 web service). This delivers time series of temperature (at 2m), relative humidity (at 2m) and rainfall. The data source is a NASA atmospheric re-analysis of satellite data using the Goddard Earth Observing System Model (GEOS-5) and focuses on historical climate analyses for a broad range of weather and climate time scales (GMAO, 2015). Because of the high correlation between rainfall and relative humidity, we use the variables, $Temperature_{it}$ and $Rainfall_{it}$.

[11] An interaction with a dummy variable for Meghalaya was dropped because of collinearity.

[12] The data on this variable is based on the authors' calculations from the *The India Human Development Survey (IHDS*). IHDS is a nationally representative, multi-topic panel survey of 41,554 households in 1503 villages and 971 urban neighbourhoods across India. The first round of interviews was completed in 2004–05; a second round re-interviewed most of these households in 2011–12. IHDS is jointly organised by researchers from the University of Maryland and the National Council of Applied Economic Research (NCAER), New Delhi.

To capture the time and policy effects, the model also has four sets of dummy variables for phase 2, phase 3, phase 4 and phase 5 of the lockdowns announced by the Indian government. A vector of the phase dummy variables is interacted with a vector of dummy variables for Maharashtra, Jharkhand and Meghalaya to capture the effect of phases in these states. μ_i is an unobservable state effect and e_{it} is an independent identically distributed error term. While we estimated both fixed-effects and random-effects models, we present only the results of the random-effects model as the fixed-effects model cannot include time-invariant variables.[13]

We have averaged the daily panel for each week and have constructed the weekly panel data. This is free from daily fluctuations in infection and death cases or in other time-invariant variables, such as temperature or rainfall.

So equations (1), and (2) are estimated by using the weekly panel data for 27 states and 12 weeks. As we discussed earlier, we estimate equations (1) and (2) by using the random-effects model. As a robustness check, the study also employs the Hausman–Taylor model. Equation (1), for example, can be rewritten by grouping the covariates into the four vectors, time-variant and exogenous variables (X_{it}^1) (eg$Temperature_{it}$and $Rainfall_{it}$, phase dummies and their interactions with state dummies), time-variant and endogenous variables (X_{it}^2) (eg($log\ CovidCases_{it-1}$, interactions of $Per\ Capita\ Income_i$ and state dummies), time-invariant and exogenous variables (Z_i^1) (eg state dummies, $\log Sex\ Ratio_{\ i}$), and time-invariant and endogenous variables (Z_i^2) (eg$Per\ Capita\ Income_i$, $\log Urban\ Population\ Density_{\ i}$,$\log Prevalence\ of\ Multi-morbidity_{\ i}$).

Here equation (1) is written as:

$$\log CSR_{it} = \gamma_0 + X_{it}^1\gamma_1 + X_{it}^2\gamma_2 + Z_i^1\gamma_3 + Z_i^2\gamma_4 + \mu_i + e_{it} \ldots\ldots\ldots(3)$$

Here it is assumed that, unlike the random-effects model, the individual effect can be correlated with endogenous variables ($E(\mu_i|X_{it}^2, Z_i^2) \neq$) and it is uncorrelated with exogenous variables ($E(\mu_i|X_{it}^1, Z_i^1) = 0$. Hausman and Taylor (1981) suggest an instrumental variable (IV) estimator which pre-multiplies equation (2) by $\Omega^{-1/2}$, where Ω is the variance covariance term of the error component, $\mu_i + e_{it}$, and performs 2SLS using instruments $[Q, X_{it}^1, Z_i^1]$ in which Q is the transformation matrix (ie based on demeaning transformation) with $\tilde{y} = Qy$ having a typical element $\tilde{y} = y_{it-}\overline{y_i}$, and $\overline{y_i}$ is the individual mean (where y_{it} is $\log CSR_{it}$in our case) (Baltagi et al, 2003, p 363). This is

[13] A statistically insignificant test statistic in the Hausman test implies that there is no significant difference in parameter estimates between random- and fixed-effects models, implying that the assumption for the random-effects model that there is no correlation between the error term (e_{it}) and the state-level individual term (μ_i) is likely to hold, that is, the test favours the random-effects model in all the cases (Table 1 and Appendix Table 3). The Breusch–Pagan Lagrange multiplier test for random effects is insignificant, which suggests that in a choice between OLS and the random-effects model, the latter should win.

equivalent to applying 2SLS to the random-effects model where the vector of time-invariant endogenous regressors, Z_i^2, is instrumented by deviations from the means of time-variant regressors, the mean of exogenous time-variant regressors and exogenous time-invariant regressors $\left[\widetilde{X_{it}^1}, \widetilde{X_{it}^2}, \overline{X_i^1} Z_i^1\right]$. Equation (1) is identified in our case because the number of regressors in X_{it}^1 is much larger than that in Z_i^2 (Baltagi et al, 2003). Our use of weather variables (part of X_{it}^1) is crucial for identifications in this context. This makes sense empirically as (1) fluctuations in weather occur outside the model of determinations of Covid-19 infections and deaths; and (2) in developing countries such as India fluctuations in weather influence macro-level income (where the share of agricultural income is more than moderate).

6 Results

We show the results of our regression analyses in Table 1 (based on the weekly panel data).

Table 1: Determinants of CSR and DSR (weekly panel regression)

Dependent variable	Log CSR		Log DSR	
	Random effects	Hausman–Taylor	Random effects	Hausman–Taylor
	Est. Coef. (z value)	Est. Coef. (z value)	Est. Coef. (z value)	Est. Coef. (z value)
Explanatory variables				
log Covidcases † *[1] L1. *[2, 3]	0.116 (1.76) *	0.116 (1.76) *	-0.043 (0.66)	-0.043 (0.66)
Per capita income/1000 (PCI) †	0.038 (8.12) ***	-0.001 (0.25)	0.049 (11.52) ***	-0.017 (3.57) ***
PCI* D_Maharashtra †	-0.006 (1.93) *	0.018 (5.50) ***	-0.003 (0.86)	0.032 (8.57) ***
PCI* D_Jharkhand †	0.038 (3.77) ***	0.008 (0.62)	0.033 (2.96) ***	-0.01 (0.70)
log Urban Population Density †	5.869 (10.77) ***	0.236 (9.71) ***	7.596 (12.26) ***	0.189 (6.34) ***
log Multi-morbidity †	0.882 (5.40) ***	1.751 (15.05) ***	0.415 (2.08) **	1.58 (10.95) ***
log Sex Ratio †	26.63 (3.64) ***		65.278 (7.71) ***	
Temperature	0.118 (1.39)	0.118 (1.39)	0.19 (2.77) ***	0.19 (2.77) ***
Rainfall	-0.014 (0.94)	-0.014 (0.94)	-0.033 (1.51)	-0.033 (1.51)
D_Phase2 *[4]	0.812 (2.38) **	0.812 (2.38) **	1.129 (2.79) ***	1.129 (2.79) ***
D_Phase3	0.979 (2.41) **	0.979 (2.41) **	1.99 (3.37) ***	1.99 (3.37) ***
D_Phase4	0.959 (2.23) **	0.959 (2.23) **	2.082 (3.40) ***	2.082 (3.40) ***
D_Phase5	1.924	1.924	4.017	4.017

Dependent variable	Log CSR				Log DSR			
	(3.06)	***	(3.06)	***	(5.89)	***	(5.89)	***
D_Jharkhand *5	9.192		3.095		12.772		2.643	
	(5.62)	***	(1.84)	*	(9.99)	***	(1.96)	*
D_Maharashtra *5	6.748		1.876		9.227		2.011	
	(11.39)	***	(10.94)	***	(13.78)	***	(15.37)	***
D_Meghalaya *5	-4.167		-4.629		-4.615		-4.251	
	(5.62)	***	(7.88)	***	(5.96)	***	(6.90)	***
D_Jharkhand* D_Phase2	3.507		3.507		-0.726		-0.726	
	(5.43)	***	(5.43)	***	(1.14)		(1.14)	
D_Jharkhand* D_Phase3	1.724		1.724		-2.788		-2.788	
	(2.74)	***	(2.74)	***	(3.59)	***	(3.59)	***
D_Jharkhand* D_Phase4	1.258		1.258		-2.501		-2.501	
	(1.91)	*	(1.91)	*	(3.29)	***	(3.29)	***
D_Jharkhand* D_Phase5	1.344		1.344		-1.053		-1.053	
	(1.90)	*	(1.90)	*	(1.17)		(1.17)	

Dependent variable	Log CSR				Log DSR			
	Random effects		Hausman–Taylor		Random effects		Hausman–Taylor	
	Est. Coef. (z value)		Est. Coef. (z value)		Est. Coef. (z value)		Est. Coef. (z value)	
Explanatory variables								
D_Maharashtra* D_Phase2	0.87		0.87		2.58		2.58	
	(2.33)	**	(2.33)	**	(7.09)	***	(7.09)	***
D_Maharashtra* D_Phase3	1.075		1.075		2.487		2.487	
	(2.76)	***	(2.76)	***	(4.56)	***	(4.56)	***
D_Maharashtra* D_Phase4	1.457		1.457		2.957		2.957	
	(3.87)	***	(3.87)	***	(5.37)	***	(5.37)	***
D_Maharashtra* D_Phase5	1.52		1.52		2.464		2.464	
	(3.15)	***	(3.15)	***	(3.84)	***	(3.84)	***
D_Meghalaya* D_Phase2	6.187		6.187		-0.82		-0.82	
	(11.77)	***	(11.77)	***	(1.57)		(1.57)	
D_Meghalaya* D_Phase3	3.972		3.972		-2.148		-2.148	
	(5.88)	***	(5.88)	***	(2.53)	**	(2.53)	**
D_Meghalaya* D_Phase4	4.069		4.069		-1.522		-1.522	
	(5.80)	***	(5.80)	***	(1.51)		(1.51)	
D_Meghalaya* D_Phase5	2.346		2.346		-3.791		-3.791	
	(3.31)	***	(3.31)	***	(4.19)	***	(4.19)	***
State dummies*5	Yes		Yes		Yes		Yes	
Constant	-277.85		-45.77		-580.354		-66.669	
	(6.12)		(1.77)	*	(10.95)		(3.20)	
No of observations(N)	377		377		377		377	
No of states(n)	27		27		27		27	
No of days(T)	14		14		14		14	
chi2			77.944	***	.		291.303	***
R squared within	0.5088				0.5298			
R squared between	1				1			
R squared overall	0.8335				0.8254			
Breush and Pagan test	0		-		0		-	

Dependent variable	Log CSR		Log DSR	
(p value)	(1.00).	-	(1.00).	-
Hausman Test *[7]	1	1	1	1
(p value)	(0.00).	(0.00).	(0.00).	(0.00).

Notes: [1] Variables marked by † are treated as endogenous.
*[2] *** = significant at 1% level. ** = significant at 5% level. * = significant at 10.5% level.*
[3] The numbers in brackets show z values. They are based on robust standard errors.
[4] State dummies for all the states have been included in the regressions. However, the results are shown only for Jharkhand and Maharashtra.
[5] D_ stands for a dummy variable (taking 1 or 0).
[6] Statistically significant cases are highlighted in bold.
[7] Hausman tests were carried out between FE and RE (the first and the third columns) and RE and Hausman-Taylor (the second and the fourth).

The key results are summarised as follows. The estimated coefficients of random-effects models are similar to those of the Hausman–Taylor model (as confirmed by the statistically insignificant statistics of the Hausman test). On the other hand, the Hausman test between the random-effects and fixed-effects models favours the former, with the caveat that the comparison is made only for time-variant covariates. Overall, these results suggest that the estimated coefficients of the random-effects model are likely to be unbiased, that is, the assumption of no correlation between explanatory variables and state-level individual effects is not necessarily unrealistic. Therefore, while commenting on the results of the Hausman–Taylor model, we will mainly focus on those of the random-effects model in our summary of the results.[14]

- The coefficient with respect to the confirmed number of Covid infection cases with a weekly lag was positive and statistically significant for the CSR in the weekly case (the first two columns of Table 2) in both the random-effects and Hausman–Taylor models. This is because the Covid-infection cases had an impact on deaths a week later in the CSR, which captures how many deaths have occurred as a result of Covid-19 cumulatively relative to the counterfactuals. In terms of the magnitude, a 1% increase in Covid cases a week before was associated with a 0.11–0.12% increase in the CSR. This is not significant for the DSR, suggesting that there is no clear correlation between Covid cases and the severity ratios.

- The relationship of CSR and DSR with respect to income was positive and significant (in random-effects models).This suggests that states with a higher income tend to be more vulnerable to the risk of people dying than those with low incomes. This was

[14]Another important observation is that the results based on daily panel data are broadly similar to those based on weekly panel data and serve as an additional robustness check of the results presented in this paper. The results based on the daily panel maybe provided upon request.

after controlling for state fixed effects. The Hausman–Taylor Model yields a negative and significant coefficient estimate for DSR.[15]

- The effect of income on the CSR was much higher in Jharkhand and slightly lower in Maharashtra than the overall conditional average effect of income on the CSR (the first column of Table 1).If we rely on the results based on the Hausman–Taylor model, the effect of income on the CSR and DSR was higher in Maharashtra than in the other states.

- A positive and strong association between the CSR (as well as the DSR) and urban population density was observed. This suggests that Covid severity is higher in states with a higher urban population density, possibly because of greater social intermixing. We avoid commenting on the relative size of the effect, as the estimated coefficient is much larger with the random-effects model than with the Hausman–Taylor model. It should be noted that urban population density is instrumented in the latter, which may explain the difference in the size of coefficient estimates.

- The ratio of females to males (sex ratio) is positive and significant for the CSR and DSR in the weekly cases. That is, in the states where there are more women than men, there were relatively more Covid-linked deaths on a cumulative basis. Although our results are not directly comparable with Joe et al (2020), our finding is consistent with their results.[16] This is in sharp contrast to the evidence from outside India (eg China, European countries), where men have been more likely to die as a result of Covid-19, while infection rates are broadly similar between men and women across different age groups (Jin et al, 2020; Gebhardet al, 2020). It is conjectured that the relative vulnerability of women, in particular older women, compared with men is higher in India than elsewhere and thus women in India tend to suffer more than men.

- A higher temperature is associated with a higher severity of Covid-19 for the DSR after controlling for the state fixed effects. This is consistent with recent evidence from China (Ma et al, 2020). It is likely in the Indian context that more deaths tend to occur in hot weather, as this is likely to worsen other illnesses. Rainfall (or relative humidity) is not associated with the CSR or DSR. This is an important result, as little is known about the association between weather and Covid-19 death cases (Ma et al, 2020).

[15] Differences in the results for the random-effects and Hausman–Taylor models arise partly because different sets of states dummies were dropped automatically as a result of collinearity among different covariates. It should also be noted that income is treated as an endogenous variable in the Hausman–Taylor model.
[16]Although an important early contribution, Joe et al (2020) is based on a different dataset that only ran up to 20 May 2020. Besides, it is short on methodological details.

- We also used a different specification, in which we inserted the interaction terms between weather variables (ie temperature and rainfall) and state dummy variables for Maharashtra and Jharkhand, and dropped the interaction terms between income and these state dummies to avoid collinearity. The results are given in Appendix Table 3. Based on the weekly panel data, we found that the effect of temperature on the CSR and DSR in Maharashtra and Jharkhand were significantly higher than in other states, as the interaction terms are positive and statistically significant. The difference is larger for Maharashtra (ie the effect of temperature is larger). In Maharashtra (Jharkhand): a one degree increase in temperature would lead to a 0.66% (0.56%) increase in the CSR on average, other factors being unchanged (based on the random-effects model and the weekly panel). On the other hand, the results for the interaction terms between rainfall and these two state dummies imply that a higher level of rainfall increases both the CSR and DSR in Jharkhand, but only the DSR in Maharashtra. A 1 mm increase in rainfall would lead to a 0.025% (0.016%) increase in the CSR (DSR) in Jharkhand and a 0.012% increase in the DSR in Maharashtra. Given that we controlled for phase effects at state levels, the statistically significant effects of temperature on the Covid-19 pandemic in these states warrants policymakers' immediate attention.

- The presence of more than one underlying condition or NCD among the elderly population (above 60 years) was positively associated with the CSR and DSR. This holds regardless of which econometric model or what type of measure of severity was used. It is well known that patients suffering from diabetes and cardiovascular disease are more vulnerable to dying from Covid-19 (Joe et al, 2020). A related reason is that, with the upsurge of infections and large-scale diversion of healthcare resources to Covid-19 patients, those suffering from these chronic conditions have been neglected and left to die.

- In addition to the above, we used indicator variables for the five phases of the lockdown and opening of the lockdown (the first four phases of the lockdown plus Unlock 1.0). They are highly significant and the estimated coefficients suggest an increase in the CSR as time progressed.

- The results of state dummies imply that, after controlling for all other factors (eg income, urban population density), both the CSR and DSR are much higher in Maharashtra and Jharkhand than in other states, while much lower in Meghalaya. On the other hand, the estimated coefficients of the interaction between the phase dummies and the state dummies for Jharkhand suggest that, while the CSR remained higher than in other states in phases 2 to 5, with the difference from the average effect becoming smaller, the DSR remained relatively low and decreased in phases 3 and 4 in Jharkhand. In Maharashtra, DSR continued to be significantly higher in phases 2–5, which resulted in the CSR becoming relatively higher from phase 2 to phase 4. The DSR in Meghalaya remained relatively low (phases 2–5), while the CSR became relatively high, in particular, in phases 2–4. It should be noted that these results do not exactly match the earlier graphical analyses, as we have taken into account various factors (state-level income, etc) in the regression analyses.

7 Discussion

Here we focus on the significance of our analysis. To the best of our knowledge, this is the first econometric analysis of the severity of the Covid-19 pandemic measured using two related but distinct measures of mortality – the CSR and DSR – up to 21 June 2020. As emphasised earlier, the CSR measures the additional pressure on India's fragile and ill-equipped healthcare system, while the DSR helps monitor the progression of fatalities. Another important contribution of this analysis is the use of rigorous econometric methodology, namely, the random-effects and Hausman–Taylor models. Although the rationales vary, they yield a large core of robust results. The specifications are rich and comprehensive, despite heavy data constraints. The factors associated with the CSR and DSR include (lagged) Covid-19 cases, income, age, gender, multi-morbidity, urban population density, lockdown phases and their interactions with two states, Maharashtra and Jharkhand, and temperature and rainfall and their interactions with the two state dummies. Given the paucity of rigorous econometric analyses, our study yields policy insights of considerable significance.

Although the robust relationship between the severity measures and (lagged) Covid-19 cases is not surprising, it has policy significance as it underlines the urgent need for speedier and more comprehensive testing. While testing has considerably accelerated, large segments of the population –especially in rural and remote areas – remain untested. There is still considerable fear, which has acted as a barrier to more voluntary testing and subsequent quarantining. If Banerjee et al (2020) are right, messaging from sources more credible than official agencies could make a difference.

Another important result is the positive association between the severity measures and per capita income. Although this relationship was observed only in the random-effects specification, it holds for both the CSR and DSR, implying that higher incomes are associated with higher mortality. The underlying mechanisms for this include greater economic activity, more travel and intermixing and, consequently, higher exposure to the infection and a higher risk of dying if denied medical assistance. However, the effect of income on the CSR was much higher in Jharkhand and lower in Maharashtra than the overall conditional average effect of income on the CSR. If there are income thresholds between which the rise in severity varies (not investigated here), these results are plausible.

A not-so-surprising result is the positive association between the CSR (or DSR), and urban population density. Although there has been a large-scale *reverse* migration from urban areas to villages, the indications are that large segments of the population have been forced to return to small towns and cities with only flickering signs of economic revival. In that case, the risks of dying from the pandemic may escalate. Exactly how a balance can be struck between economic revival, expansion of livelihoods and containment of the pandemic is still in the realm of speculation. A related risk of worsening sanitation and hygiene – especially in the slums –is daunting but preventable.

The positive association of both the CSR and DSR with the sex ratio, the number of women per 1000 men, has been controversial. As Joe et al (2020) observe, preliminary evidence from various countries suggests that men are at greater risk of both infections and deaths, but these inferences should not be taken at face value. As of May 2020, the incidence of Covid-19 cases and deaths in India was suggesting that males were at a greater disadvantage than females, with a CFR of 3.3% and 2.9%, respectively. However, it is not evident whether men experience a higher risk of mortality throughout the age spectrum or whether there are sex-differentials in survival risk. It has been argued that pre-existing conditions, behavioural risk factors (smoking) and biological factors all elevate the risk of mortality among males, but these patterns need to be verified with new and emerging information on the Covid-19 outbreak in India. In a statisticalc analysis, Joe et al.(2020) shows that the CFR among males is usually higher than among females in most age groups. Male and female CFRs also have distinct patterns, with lower survival rates in under-five males as well as in older age groups. Females have a higher risk of mortality in the age group 40–49 years; this leads to a marginally higher overall risk of Covid-19 mortality for females. The difference is found to be statistically significant for three age groups (5–19 years, 30–39 years and 40–49 years). Our study corroborates the finding that women are more vulnerable to the risk of dying from Covid-19 through rigorous panel models not used so far. One missing link in both Joe et al (2020) and the present analysis is whether infected women are less likely to receive medical care than infected men. As Amartya Sen (2001) has argued emphatically in several important essays:

Female mortality rates are very significantly higher than what could be expected given the mortality patterns of men (in the respective age groups). This type of gender inequality need not entail any conscious homicide, and it would be a mistake to try to explain this large phenomenon by invoking the occasional cases of female infanticide that are reported from China or India; these are truly dreadful events when they occur, but they are relatively rare. Rather, the mortality disadvantage of women works mainly through a widespread neglect of health, nutrition and other interests of women that influence survival.

Several studies, including Joe et al (2020), have reported greater vulnerability among the old (60 years or more) to the risks of Covid-19 infection and fatality, but few take into account the vulnerabilities of the multi-morbid aged (ie those elderly people suffering from more than one NCD, such as diabetes, cardiovascular disease or hypertension). Our analysis corroborates the finding that the multi-morbid aged are more vulnerable to dying from Covid-19. While it is well-known that those suffering from diabetes and cardiovascular disease are more likely to die from the virus, it is not so well-known that the pandemic caught both policy makers and the health system unprepared. Limited healthcare resources were diverted to the testing and treatment of Covid-19 cases and there was neglect of the elderly with near fatal chronic conditions (Kulkarni & Gaiha, 2020). This further aggravated the healthcare crisis. So, short of a substantial expansion of healthcare resources, their rationing must assign a high priority to elderly people with multi-morbid conditions.

Few studies have established the importance of weather conditions (temperature and rainfall) in explaining the severity of Covid-19 (a notable exception being Ma et al, 2020). Epidemiologists remain divided on this issue. Using alternative specifications, we have found that both high temperature and high rainfall are associated with high severity of Covid-19. Focusing on Jharkhand and Maharashtra, we found that the associations of temperature with the CSR and DSR were significantly higher in those two states than in other states. The difference is larger for Maharashtra. In this state (Jharkhand): a one degree increase in temperature was associated with a 0.66% (0.56%) increase on average, other factors held constant. In contrast, the interaction of rainfall and these two state dummies show that a higher level of rainfall is associated with increases in both the CSR and the DSR in Jharkhand, but only with the DSR in Maharashtra. A 1 mm increase in rainfall was associated with a 0.025% (0.016%) increase in the CSR(DSR) in Jharkhand and a 0.012% increase in the DSR in Maharashtra. These results are particularly important in the present context, as, at the time of writing, the monsoon season is underway and record spikes in Covid-19 cases in recent weeks warrant a quick and substantial step-up in the prevention strategy.

Whether the stringent lockdowns – with huge livelihood losses among migrant labourers and a sharp contraction of the economy from late March to the end of May2020 –were avoidable continues to be debated. Specifically, whether the prevention of deaths through containment of the Covid pandemic justified such a contraction and loss of livelihoods is a dilemma that is far from resolved.[17]Our analysis sheds light on whether the severity of Covid-19 varied in different lockdowns – especially in Jharkhand and Maharashtra. Using indicator variables for the five phases, the estimated coefficients suggest an increase in CSR as time progressed. The results of state dummies imply that, after controlling for all other factors (eg income, urban population density, sex ratio, weather), both the CSR and DSR are much higher in Maharashtra and Jharkhand than in other states, while much lower in Meghalaya. However, the estimated coefficients of the interaction between phase dummies and state dummy for Jharkhand suggest that,

[17]As Nayyar (2020) has eloquently observed: “The lockdown has shut down almost two-thirds of the economy and the collateral damage is enormous. It stranded 25–30 million migrants in cities far away from their homes, deprived of their work and dignity, at the mercy of shelters for food provided by state governments or charities, often hungry and homeless. Manufacturing, mining, construction, trade, hotels and restaurants, and transport, which account for more than 40 per cent of both output and employment, were shut down completely. Thus,150 million people, as much as one-third of our workforce, who are casual labourers on daily wages or workers in informal employment without any social protection, were deprived of their livelihoods.”

while the CSR remained higher than in other states in phases 2 to 5, the difference from the average effect became smaller. The DSR remained relatively low and decreased in phases 3 and 4 in Jharkhand. In Maharashtra, the DSR continued to be significantly higher in phases 2–5 and, associated with it, the CSR rose during phases 2–4. In Meghalaya, a special case, the DSR remained relatively low (phases 2–5), while the CSR became relatively high, particularly during phases 2–4. Therefore serious doubts persist about the efficacy of lockdowns in containing the severity of the Covid pandemic.

A few limitations are briefly noted. One is that the analysis lacks household data. Instead, the state has been the unit of analysis. As there is considerable variation in Covid-19 fatalities within states, we were unable to capture this variation. Another limitation is the lack of data on health capacity and infrastructure for measuring the response to the Covid-19 pandemic. A third limitation is that we have been unable to assess the impact of return migration on the villages and small towns to which people were returning. Nevertheless, while our preferred specification is the random-effects model and there are a few differences between its results and those obtained from the Hausman–Taylor specifications, there is a large core of robust results.

8 **Policy challenges**

A video that went viral depicts corpses wrapped in plastic left in a Covid ward full of patients, because of the lack of space in a hospital morgue. Hospital staff have protested strongly against personal assaults and unsafe conditions, poor equipment and long working hours. In Delhi, some 600 health workers have tested positive for Covid-19, including 329 at the prestigious All-India institute of Medical Sciences. As the virus has spread like wildfire, from cities to rural states such as Jharkhand, where there is just one doctor for every 6,000 people, the quality of healthcare has deteriorated alarmingly. No less chilling is finding places in morgues, cemeteries and crematoria often with abandoned corpses wrapped in white. Delhi has relaxed a ban on traditional funeral pyres made of wood, instituted to reduce pollution, because there are too few gas-fired ovens to meet the spurt in demand.[18]

The first priority must therefore be to increase substantially funding for the health sector. Including the private sector, total health expenditure in India as a percentage of GDP is estimated at 3.9%. Out of this total expenditure, effectively about one-third (30%) is contributed by the public sector. This contribution is low compared with other developing and developed countries (Rao, 2018). But more important than the amount is the quality of health care, for which a prerequisite is drastic reforms in the provision of healthcare services. One option is to engage with the private sector on a larger scale. Views differ, ranging from scepticism and caution to support for the strong complementarities. One view is that there is a need for more caution but recognition of potential opportunities when engaging with the private sector, and for greater clarity on how the government, the private sector and civil society interact. Another important issue is how to manage multi-sectoral governance and cooperation in health, taking advantage of insights from

[18]Adapted from 'Deadly Tide'.*The Economist,* 6 June 2020.

political, behavioural and organisational sciences.[19] Although there is merit in this view, it is not particularly helpful in the context of the Covid pandemic. A contrary but more convincing and elaborately constructed view is that countries that have had a policy-based strategic relationship with the private sector seem to have performed better in controlling the pandemic. Rather than 'arm twisting' the private sector, it is necessary to formulate a stable policy-based strategy to engage with it.[20] Key elements of this engagement are highlighted through different scenarios.

One scenario is a mobile sample-collection and testing facility operated by a private entity in high-density clusters; this could also be used as a fever clinic. This arrangement could operate under the hub–spoke principle. The cost of tests, key performance indicators and a payment system should be worked out in the purchase contract. Hospitalisation of Covid-19 cases cannot be restricted to hospitals in major cities alone. Improving the infrastructure and capacity in tier II and III cities in collaboration with the private sector is critical.

The government could refer patients to empanelled private Covid-19 hospitals, at a fixed package rate. This kind of strategic purchasing or insurance reimbursement (say, under the Pradhan Mantri Jan Arogya Yojana) requires clear policy directions, a robust referral system, agreement on tariffs and a swift reimbursement mechanism. The current government tariffs are far from attractive for the private sector.

The growing scarcity of test kits, ventilators and other biomedical supplies cannot be met by current manufacturers or supply chain sources. Alternative indigenous sources could be explored. A plethora of innovations and prototypes need government laboratories to test them rapidly, approve them and grant a licence for production thatincludes patenting.

In a nuanced and coherent proposal, in the context of the vulnerability of elderly people with NCDs to the high risk of dying from Covid-19, a case could be made for developing a fully integrated population-based healthcare system bringing together the public and private sectors and the allopathic and indigenous systems. The system should be well coordinated at different levels of service delivery platform – primary, secondary and tertiary. It should address acute and chronic healthcare needs, offer accessible, good quality healthcare choices, and be cashless at the point of service delivery. The primary healthcare provider should be a strengthened public care system with a clearly defined role for the private sector, especially in specialised services (Patel et al, 2015).

While the potential benefits of public–private partnerships as delineated here seem high, much will depend not just on the speed of approval of contracts and their effective enforcement but also on effective regulation, with the well-known paradox of who regulates the regulators.

While there is a need for more comprehensive monitoring and surveillance of at-risk individuals (eg those with influenza-like symptoms, people who have had contact with an

[19]For details, see Healthier Societies for Healthy Populations Group (2020).
[20]Yellapa&Venkatraman (2020).

individual testing positive for Covid-19, or those with a travel history to an affected region), this does not provide an accurate estimate of how many persons are infected in the population. The point is that at-risk individuals are not representative of the general population. Establishing this value, as emphasised by Joe et al (2020), is vital to assessing Covid-19-related morbidity and mortality, as India cannot absorb the economic and health impact of the pandemic.

Another major concern is that the response to Covid-19 infection depends on individuals'immune systems. Specifically, individuals with poor nutritional status are likely to have a weak immune system. A significant proportion of women in the age-group 15–49 years are undernourished, for example, and this makes them more vulnerable to Covid-19 morbidity and mortality. As the risks of chronic diseases accumulate over a life span, the old tend to be more vulnerable to diabetes and cardiovascular diseases, which expose them to higher risks of Covid-19 morbidity and mortality. Amid elevated risks to lives and livelihoods, there has also been a surge in hunger and food deprivation in both rural and urban areas. Further, disruption of healthcare services is inimical to nutritional health (Joe et al, 2020). So food security is a major policy challenge.

Taking our scrutiny further, an important point is that Covid-19's most significant effect will be on national food security via its effects on food supply chains (FSCs), as 92% of the food consumed in India is bought from FSCs (Reardon et al,2020). Most of these effects will be on the post-farmgate FSC – the firms and workers in the midstream wholesale, processing and logistics segments, and downstream in retail and food service – and much less on farms and farm workers. The main reasons are: (1) 60% of the FSC and, thus, the formation of food prices and FSC employment come from post-farmgate activities;(2) post-farmgate activities tend to be clustered in peri-urban areas, towns and secondary cities in areas close to farmland;as Covid-19 is transmitted via human contact, greater population densities tend to facilitate its spread; and (3) post-farmgate activities are dominated by large numbers of SMEs that tend to operate in de facto, spontaneous clusters or in dense enclosed areas such as *mandis*, while farms, by comparison, are spread out.

The direct effect of Covid-19 on farms is likely to be limited. Because farms are relatively spread out, the human density aspect of its spread will be lower than in the cities. However, the indirect effect of Covid-19 on farms is likely to be substantial. First, the virus's main effect on farmers will be through deficient effective demand from consumers via the constraints on the midstream and downstream of the FSC and because of a reduction in consumers' real incomes in the crisis. The effect will be strongest on perishable products such as milk, fruit and vegetables, and fish and chicken, which require more handling and are more income-elastic in demand. Second, its effects on the midstream of input supply chains such as fertiliser and seed will hurt farmers.

Contrary to a widely held view of universalising access to Public Distribution System (PDS)-which provides subsidised food- a recent study offers a caveat (Reardon et al, 2020). Of the enormous volume of food consumed, the Indian government directly supplied barely 4% of it. So even if the government doubles its sales or even transfers of food, it would cover 8%, or just one month, of the food market. As Covid-19 disruption of the food economy is likely to persist for several months, a much greater reliance on strengthening food supply chains (FSCs) is likely to have a greater pay-off in terms of food security.

As our analysis confirms robustly, the lockdown and subsequent unlocks have not been effective in curbing Covid-19 morbidity and mortality. There are two key issues. One is revival of the economy – especially of SMEs – through an expansionary fiscal policy, as they are the backbone of the economy and the fate of millions of workers is inextricably tied to them. So an effective fiscal stimulus imperative, something which has been announced more often than implemented.

Behavioural changes evolve over time. A glaring example is sex-selective discrimination in which girls are deprived of food and health care from the womb to adulthood. Thus, not only do they lack immunity against Covid-19, when they succumb to it, they are more often than not denied access to medical care and hence are more likely to die from Covid-19 than men, as confirmed in our analysis. As a vaccine against Covid-19 is months away, apart from strengthening a fragile and collapsing health system, the second key issue is accelerating behavioural changes such as hand-washing, maintaining a safe physical distance from others, eating nourishing food and sharing it equitably within the household. Although official agencies emphasise the imperative of such behavioural changes *ad nauseam*, judging by the continuing spikes in Covid-19 morbidity and mortality, official campaigns have had little effect on or, worse, have failed miserably tocontain the pandemic.

From this perspective, the insights offered in Banerjee et al (2020) merit serious consideration. As noted already, the SMSes to millions in West Bengal produced some remarkable insights. The campaign first doubled the reporting of health symptoms to the community health workers; second, reduced travel beyond one's village in the previous two days and increased estimated hand-washing when returning home; third, spilled over to behaviours not mentioned in the message – for example, mask-wearing –which increased slightly, while distancing and hygiene both increased in the sample where they were not mentioned by similar amounts towhere they were mentioned; and fourth, spilled over onto non-recipients within the same community, with effects similar to those for individuals who received the messages.

All this is fine but replicability across different states remains a challenge.

A somewhat baffling result is that higher incomes are associated with a higher severity of Covid-19. If we juxtapose this finding with the positive association of urban density, a missing link is whether higher income growth is driven by greater urbanisation. As evidence of the growing importance of urbanisation in the growth process has accumulated, there are two conjectures. One is, of course, the different thresholds of income that affect Covid-19 severity, which we have not established. Another is lifestyles and associated NCD incidence. As obesity tends to be higher in urban areas, mainly because of sedentary lifestyles and rich diets (eg eating out, fast food) and consequently the incidence of NCDs and the severity of Covid-19 are also higher, effective solutions must be found to address these concerns. As tax policies (eg higher taxes on cigarettes and alcohol) may have limited impact, placing greater emphasis on recruiting credible sources may induce behavioural changes. A study co-authored by one of the present authors points to the important role of mass media and social networks in influencing behavioural responses (Kulkarni & Gaiha, 2020).

Finally, looking beyond the current pandemic, a perceptive comment by Horton (2020) merits serious consideration. If we are able to diagnose new infections more rapidly, there is hope of exiting lockdown faster and more safely. For example, self-isolation when there are early signs of muscle pain, fatigue, headache, diarrhoea and rashes may lead to every possibility of avoiding a second wave. Another important observation is that prolonged lockdowns are not the answer to future waves of Covid-19. School closures are not sustainable, nor can the economy be refrigerated again. What matters most is a mix of prevention measures that include hand-washing, respiratory hygiene, mask-wearing, physical distancing and avoiding mass gatherings.

In brief, the tidal wave of the corona pandemic calls for extraordinary measures. While some are identified here, their implementation is daunting.

References

Baltagi, B.H. (2008).*Econometric Analysis of Panel Data*. 3rd edn.Chichester: John Wiley.

Baltagi, B.H., Bresson, G. andPirotte, A. (2003). ‘Fixed effects, random effects or Hausman–Taylor? A pretest estimator’.*Economics Letters79*, 361–369.

Banerjee, A., Alsan, M.,Breza, E., Chandrasekhar, A.G., Chowdhury, A.,Duflo, E. and Goldsmith-Pinkham, P. (2020).*Messages on Covid-19 Prevention in India, Increased Symptoms Reporting and Adherence to Preventive Behaviour among 25 Million Recipients with Similar Effects on Non-recipient Members of their Communities*. National Bureau of Economic Research Working Paper 27496. Cambridge MA: NBER.

Brown, C.S., Ravallion, M. and Van de Walle, D. (2020).*Can the World’s Poor protect themselves from the New Corona Virus?*. NBER Working Paper 27200. Cambridge MA: NBER.

Ehlert, A. (2020). ‘The socioeconomic determinants of COVID-19: a spatial analysis of German county level data’. *medRxiv*(pre-peer review print).

Gebhard, C., Regitz-Zagrosek, V., Neuhauser, H.K., Morgan, R. and Klein, S.L. (2020). ‘Impact of sex and gender on COVID-19 outcomes in Europe’. *Biology of Sex Differences11*, 1–13.

Global Modeling and Assimilation Office (GMAO) (2015). MERRA-2 tavg1_2d_slv_Nx: 2d,1-Hourly,Time-Averaged,Single-Level,Assimilation,Single-Level Diagnostics V5.12.4. Greenbelt MD: Goddard Earth Sciences Data and Information Services Center. doi:10.5067/VJAFPLI1CSIV.

Hausman, J.A. and Taylor, W.E. (1981). ‘Panel data and unobservable individual effects’. *Econometrica49*, 1377–1398.

Healthier Societies for Healthy Populations Group (2020).‘Healthier societies for healthy populations.’*Lancet395*, 6 June,1747-1749.

Horton,R. (2020). ‘Offline: the second wave.’*Lancet 395*,27 June,1960.

IMF (2020).‘A crisis like no other’.*World Economic Outlook Update*. Washington DC: IMF.

Jin, J.M., Bai, P., He, W., Wu, F., Liu, X.F., Han, D.M., Liu, S. and Yang, J.K. (2020). ‘Gender differences in patients with COVID-19: focus on severity and mortality’. *Frontiers in Public Health8*, 152, 1-6.

Joe, W., Kumar, A., Rajpal, S., Mishra, U.S. and Subramanian, S.V.(2020). ‘Equal risk, unequal burden? Gender differentials in COVID-19 mortality in India’. *Journal of Global Health Science*.https://doi.org/10.35500/jghs.2020.2.e17.

Jorda, O., Singh, S.R. and Taylor, A.M. (2020).‘The long economic hangover of pandemics’.*Finance and Development,* June, 12-18.

Kulkarni, V.S. andGaiha, R. (2020).‘Tragic choices in health care during Covid’.*Business Standard*, 19 May [available athttps://www.business-standard.com/article/opinion/tragic-choices-in-health-care-during-covid-120051901758_1.html].

Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., Yan, J., Niu, J., Zhou, J. and Luo, B. (2020). ‘Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China’. *Science of The Total Environment724*.https://doi.org/10.1016/j.scitotenv.2020.138226.

Nayyar,D, (2020).‘Let there be light’.*Open*, 25May [available at https://openthemagazine.com/essay/let-there-be-light-2/].

Oldekop, J.A., Horner, R., Hulme, D., Adhikari, R., Agarwal, B., Alford, M., Bakewell, O., Banks, N., Barrientos, S., Bastia, T. and Bebbington, A.J. (2020).‘COVID-19 and the case for global development’. *World Development*.https://doi.org/10.1016/j.worlddev.2020.105044.

Patel, V., Parikh, R., Nandraj, S.,Balasubramaniam, P., Narayan, K., Paul, V.K., Shiva Kumar, A.K., Chatterjee, M. and Srinath Reddy, K. (2015).‘Assuring health coverage for all in India’. *Lancet 386*, 2422–2435.

Rao, N. (2018).‘Healthcare financing: who is paying?’.*The Wire*, April 14.

Ray, D. and Subramanian, S. (2020). ‘India's lockdown: an interim report’. Mimeo.

Reardon, T., Mishra, A.,Nuthalappati, C.S.R., Bellemare, M.F. and Zilberman, D. (2020).‘Covid-19’s disruption of India’s transformed food supply chains’.*Economic and Political Weekly*, 2 May.

Sen, A.K. (2001).‘Many faces of gender inequality’. *Mainstream*, Volume 18 - Issue 22, Oct. 27 - Nov. 09.

Schellenkens, P. andSourrowuille, D. (2020).‘*Covid-19 Mortality in Rich and Poor Countries: A Tale of Two Pandemics*. Policy Research Working Paper 9260. Washington DC: World Bank.

Sen, K. (2020).*Covid-19 and Socio-economic Impact in Asia: The Case of India*.Working Paper 1/2020.Helsinki: WIDER.

Tosepu, R., Gunawan, J., Effendy, D.S., Lestari, H., Bahar, H. andAsfian, P. (2020). ",Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia ". *Science of The Total Environment725*. https://doi.org/10.1016/j.scitotenv.2020.138436.

Vadlamannati, K.C., Cooray, A. and de Soya, I. (2020). *Healthcare Equity and Covid-19*. Working Paper 2020/89. Helsinki: WIDER.

Yellapa, V. and Venkatraman, A.S. (2020) "Paging the private sector in the COVID fight" *The Hindu*, 8 June.

Appendix Table 1: Descriptive statistics

Variable		Mean	Std. dev.	Min	Max	Obs.
log CSR	overall	-5.19	3.47	-9.21	1.24	N =481
	between		2.67	-9.21	-1.11	n =33
	within		2.31	-12.86	1.44	T=14.575
log DSR	overall	-6.77	3.36	-9.21	2.35	N =481
	between		2.58	-9.21	-0.65	n =33
	within		2.19	-14.60	0.03	T=14.575
log Covidcases	overall	3.76	5.00	-9.21	11.73	N =480
	between		3.25	-5.64	8.73	n =32
	within		3.85	-12.18	13.64	T =15
Per capita income(PCI)	overall	120	71.75	30.62	396.94	N =480
	between		72.82	30.62	396.94	n =32
	within		0.00	119.75	119.75	T =15
PCI* D_Maharashtra	overall	4.80	26.77	0.00	153.72	N =480
	between		27.17	0.00	153.72	n =32
	within		0.00	4.80	4.80	T =15
PCI* D_Jharkhand	overall	1.79	9.96	0.00	57.16	N =480
	between		10.10	0.00	57.16	n =32
	within		0.00	1.79	1.79	T =15
log Urban Population Density	overall	7.06	4.09	-9.21	9.15	N =478
	between		4.28	-9.21	9.15	n =32
	within		0.00	7.06	7.06	T = 14.9375
log Multi-morbidity *	overall	1.74	0.59	0.46	3.56	N =420
	between		0.60	0.46	3.56	n =28
	within		0.00	1.74	1.74	T =15
log Sex Ratio †	overall	6.85	0.05	6.71	6.99	N =465
	between		0.06	6.71	6.99	n =31
	within		0.00	6.85	6.85	T =15
log Temperature	overall	300	6.17	275.16	311.13	N =480
	between		5.47	282.78	305.84	n =32
	within		2.99	289.35	309.09	T =15
log Rainfall	overall	4.03	6.46	0.00	45.86	N =480
	between		2.86	0.35	13.23	n =32
	within		5.82	-8.28	44.88	T =15

Appendix Table 2: Correlation matrix

	log Covidcases	Per capita income (PCI)	PCI* D_Maharashtra	PCI* D_Jharkhand	Temperature	Rainfall	Phase2	Phase3	Phase4	Phase5	D_Jharkhand* D_Phase2	D_Jharkhand* D_Phase3	D_Jharkhand* D_Phase4	D_Jharkhand* D_Phase5	D_Maharashtra* D_Phase2	D_Maharashtra* D_Phase3	D_Maharashtra* D_Phase4	D_Maharashtra* D_Phase5	D_Meghalaya* D_Phase2	D_Meghalaya* D_Phase3	D_Meghalaya* D_Phase4	D_Meghalaya* D_Phase5
log Covidcases	1																					
Per capita income(PCI)	-0.08	1																				
PCI* D_Maharashtra	0.18	0.09	1																			
PCI* D_Jharkhand	-0.05	-0.16	-0.03	1																		
Temperature	0.5	-0.02	0.09	0.03	1																	
Rainfall	0.08	0.03	-0.03	-0.01	-0.26	1																
Phase2	0	0	0	0	-0.04	-0.13	1															
Phase3	0.1	0	0	0	0.08	-0.11	-0.1	1														
Phase4	0.17	0	0	0	0.15	0.08	-0.27	-0.03	1													
Phase5	0.41	0	0	0	0.09	0.45	-0.34	-0.3	-0.15	1												
D_Jharkhand* D_Phase2	0	-0.07	-0.02	0.48	0.01	-0.05	0.16	-0.02	-0.04	-0.05	1											
D_Jharkhand* D_Phase3	0.02	-0.07	-0.01	0.43	0.03	-0.05	-0.02	0.16	-0.01	-0.05	0.13	1										
D_Jharkhand* D_Phase4	0.03	-0.07	-0.01	0.43	0.05	-0.04	-0.04	-0.01	0.16	-0.02	-0.01	0.16	1									
D_Jharkhand* D_Phase5	0.06	-0.08	-0.02	0.52	0.01	0.13	-0.05	-0.04	-0.02	0.15	-0.01	-0.01	0.1	1								
D_Maharashtra* D_Phase2	0.08	0.04	0.48	-0.02	0.06	-0.05	0.16	-0.02	-0.04	-0.05	-0.01	-0.01	-0.01	-0.01	1							
D_Maharashtra* D_Phase3	0.09	0.04	0.43	-0.01	0.05	-0.05	-0.02	0.16	-0.01	-0.05	-0.01	-0.01	-0.01	-0.01	0.13	1						
D_Maharashtra* D_Phase4	0.11	0.04	0.43	-0.01	0.05	-0.01	-0.04	-0.01	0.16	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01	0.16	1					
D_Maharashtra* D_Phase5	0.14	0.04	0.52	-0.02	0.02	0.07	-0.05	-0.04	-0.02	0.15	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.1	1				
D_Meghalaya* D_Phase2	-0.08	-0.06	-0.02	-0.02	-0.11	-0.01	0.16	-0.02	-0.04	-0.05	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	1			
D_Meghalaya* D_Phase3	-0.02	-0.06	-0.01	-0.01	-0.08	0.03	-0.02	0.16	-0.01	-0.05	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.13	1		
D_Meghalaya* D_Phase4	-0.01	-0.06	-0.01	-0.01	-0.07	0.22	-0.04	-0.01	0.16	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.16	1	
D_Meghalaya* D_Phase5	0	-0.07	-0.02	-0.02	-0.07	0.25	-0.05	-0.04	-0.02	0.15	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.10	1

Appendix Table 3: Results with the interaction terms between weather and state dummies (weekly panel)

Dependent variable	Log CSR		Log DSR	
	Random effects	Hausman–Taylor	Random effects	Hausman–Taylor
Explanatory variables	Est Coeff (z Value)	Est Coeff (z Value)	Est Coeff (z Value)	Est Coeff (z Value)
log Covidcases † *[1] L1.*[2,3]	0.112 (-1.71) *	0.112 (-1.71) *	-0.046 (-0.69)	-0.046 (-0.69)
Per capita income/1000 (PCI) †	0.038 (-8.12) ***	-0.002 (-0.37)	0.049 (-11.48) ***	-0.018 (-3.66) ***
log Urban Population Density †	5.918 (-10.92) ***	0.239 (-9.79) ***	7.627 (-12.12) ***	0.191 (-6.21) ***
log Multi-morbidity *	0.899 (-5.19) ***	1.776 (-13.71) ***	0.457 (-2.1) **	1.628 (-9.83) ***
log Sex Ratio †	27.525 (-3.76) ***		66.042 (-7.64) ***	
Temperature	0.11 (-1.27)	0.11 (-1.27)	0.179 (-2.61) **	0.179 (-2.61) **
Temperature* D_Maharashtra	0.546 (-7.03) ***	0.546 (-7.03) ***	1.738 (-23.06) ***	1.738 (-23.06) ***
Temperature* D_Jharkhand	0.447 (-5.64) ***	0.447 (-5.64) ***	0.299 (-3.64) ***	0.299 (-3.64) ***
Rainfall	-0.018 (-1.05)	-0.018 (-1.05)	-0.04 (-1.59)	-0.04 (-1.59)
Rainfall* D_Maharashtra	0.009 (-0.55)	0.009 (-0.55)	0.052 (-2.07) **	0.052 (-2.07) **
Rainfall* D_Jharkhand	0.043 (-2.39) **	0.043 (-2.39) **	0.056 (-2.09) **	0.056 (-2.09) **
D_Phase2 *[4]	0.859 (-2.56) **	0.859 (-2.56) **	1.181 (-2.9) ***	1.181 (-2.9) ***
D_Phase3	1.038 (-2.59) **	1.038 (-2.59) **	2.054 (-3.46) ***	2.054 (-3.46) ***
D_Phase4	1.04 (-2.43) **	1.04 (-2.43) **	2.179 (-3.55) ***	2.179 (-3.55) ***
D_Phase5	2.015 (-3.19) ***	2.015 (-3.19) ***	4.131 (-5.99) ***	4.131 (-5.99) ***
D_Jharkhandd *[5]	9.123 (-5.54) ***	2.929 (-1.72) *	12.634 (-9.99) ***	2.429 (-1.8) *
D_Maharashtra *[5]	6.836 (-11.54) ***	1.905 (-10.66) ***	9.318 (-13.68) ***	2.06 (-14.84) ***
D_Meghalaya *[5]	-4.261 (-5.7) ***	-4.705 (-7.95) ***	-4.709 (-5.97) ***	-4.327 (-6.89) ***
D_Jharkhand* D_Phase2	2.392 (-4.26) ***	2.392 (-4.26) ***	-1.496 (-2.8) ***	-1.496 (-2.8) ***
D_Jharkhand* D_Phase3	0.157 (-0.31)	0.157 (-0.31)	-3.854 (-5.58) ***	-3.854 (-5.58) ***
D_Jharkhand* D_Phase4	-1.401 (-2.9) ***	-1.401 (-2.9) ***	-4.251 (-7.03) ***	-4.251 (-7.03) ***
D_Jharkhand* D_Phase5	-0.179 (-0.28)	-0.179 (-0.28)	-2.446 (-3.28) ***	-2.446 (-3.28) ***
D_Maharashtra* D_Phase2	-0.237 (-0.76)	-0.237 (-0.76)	-0.92 (-2.58) **	-0.92 (-2.58) **
D_Maharashtra*D_Phase3	-0.02 (-0.06)	-0.02 (-0.06)	-0.956 (-1.69) *	-0.956 (-1.69) *

D_Maharashtra* D_Phase4	0.228 (-0.57)	0.228 (-0.57)	-0.981 (-1.75) *	-0.981 (-1.75) *
D_Maharashtra* D_Phase5	1.761 (-3.33) ***	1.761 (-3.33) ***	3.077 (-4.9) ***	3.077 (-4.9) ***
D_Meghalaya* D_Phase2	6.19 (-11.72) ***	6.19 (-11.72) ***	-0.819 (-1.56)	-0.819 (-1.56)
D_Meghalaya* D_Phase3	4 (-5.89) ***	4 (-5.89) ***	-2.124 (-2.48) **	-2.124 (-2.48) **
D_Meghalaya* D_Phase4	4.167 (-5.75) ***	4.167 (-5.75) ***	-1.36 (-1.26)	-1.36 (-1.26)
D_Meghalaya* D_Phase5	2.427 (-3.38) ***	2.427 (-3.38) ***	-3.668 (-3.91) ***	-3.668 (-3.91) ***
State dummies *5	Yes	Yes	Yes	Yes
Constant	-281.855 (-6.24)	-43.211 (-1.65)	-582.619 (-10.8)	-63.414 (-3.04)
No of observations(N)	377	377	377	377
No of states(n)	27	27	27	27
No of days(T)	14	14	14	14
chi2		565.214***	.	348.49***
R squared within	0.5125	-	0.5364	
R squared between	1	-	1	
R squared overall	0.8348	-	0.8254	
Breush and Pagan Test (p value)	0 (1.00).	- -	0 (1.00).	- -
Hausman Test *7 (p value)	1 (0.00).	1 (0.00).	1 (0.00).	1 (0.00).

Notes: [1] Variables marked by † are treated as endogenous.

*[2] *** = Significant at 1% level. ** = Significant at 5% level. * = Significant at 10.5% level.*

[3] The numbers in brackets show z values. They are based on robust standard errors.

[4] State dummies for all the states have been included in the regressions.

However, the results are shown only for Jharkhand and Maharashtra.

[5] D_ stands for a dummy variable (taking 1 or 0).

[6] Statistically significant cases are highlighted in bold.

[7] Hausman tests were carried out between FE and RE (the first and the third columns) and RE and Hausman–Taylor (the second and the fourth).