MANCHESTER 1824
The University of Manchester

Economics

Discussion Paper

Series

EDP-2016

Severity of the COVID-19 Pandemic in India

Katsushi S. Imai, Nidhi Kaicker & Raghav Gaiha

December 2020

Economics

School of Social Sciences
The University of Manchester
Manchester M13 9PL

Severity of the COVID-19 Pandemic in India

Katsushi S. Imai, (*),

University of Manchester

Nidhi Kaicker,

Ambedkar University,

Raghav Gaiha

University of Manchester & University of Pennsylvania

This Version: 31 December 2020

Abstract

This is one of the first econometric analyses of severity of COVID-19 pandemic in India measured using two related but distinct measures of mortality up to 31 October 2020 based on the Cumulative Severity Ratio (CSR). The CSR measures the additional pressure on our fragile and ill-equipped healthcare system, while its first difference helps monitor the progression of fatalities. These measures are supplemented by a measure of infection cases. Another important contribution of this analysis is the use of rigorous econometric methodologies drawing upon random effects models and Tobit models for the weekly panel of 32 states/union territories. 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 infection cases include income, gender, multi-morbidity, urbanisation, lockdown and unlock phases, weather including temperature and rainfall, and the retail price of wheat. Given the paucity of rigorous econometric analyses, our study yields policy insights of considerable significance.

Key words: COVID-19, Cumulative Severity Ratio, Daily Severity Ratio, Random-Effects Model, India, Maharashtra

JEL Codes:C23; I18; N35; O10

(*) Corresponding Author: Katsushi S. Imai (Dr) Address: Department of Economics, School of Social Sciences, Arthur Lewis Building; University of Manchester, M13 9PL, UK, Email address: Katsushi.Imai@manchester.ac.uk

Acknowledgements: This study is funded by 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 Agarwal for her outstanding research support. The views are personal and not necessarily of IFAD.

Severity of the COVID-19 Pandemic in India

1. Introduction

More than one year has passed since the 1st of December 2019 when the first case of COVID-19 was confirmed in China (Wu et al., 2020) and nearly 11 months have passed since the first positive COVID-19 case was registered in India on 30 January, 2020, in Kerala. As of 20th December 2020, the total coronavirus infection cases in India were 10,170,470 (the second next to USA) with death numbers 147,386 (the third next to USA and Brazil).[1] There are significant geographical variations, however. One state alone (i.e., Maharashtra) has recorded close to one fifth of the total cases, and one third of the total deaths. Despite a rapid progress in medical research on COVID-19, what non-medical factors, in particular socio-economic factors, are associated with the COVID-19 pandemic in India remain largely unknown. Our focus in the present study is on the socio-economic, meteorological and geographical factors associated with the *severity* of the COVID-19 pandemic in India. Despite a surge in the studies about the socio-economic impacts of COVID-19[2], there have not been many studies on the determinants of COVID-19 infections in developing countries, including India. Our study seeks to provide important policy insights for policymakers into the policies about COVID-19 not only in India but in other developing countries.

We carry out regression analyses for the weekly and monthly panel datasets of 32 Indian states/union territories in March-October 2020 to understand the pandemic in these states in the national context to identify the determinants of the COVID-19 pandemic.[3] We also focus on the state of Maharashtra where the pandemic has been severest. Maharashtra, home to around 10 per cent of the total population of India[4], 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.

The research questions we propose to ask are: (i) What are the factors associated with the severity and infection cases of COVID-19 pandemic in India?; and (ii) How has the pandemic of COVID-19 developed in Maharashtra in comparison with other Indian states? Given the nature of the data (i.e. the state-level panel data), it would be difficult to identify the causal relationship.[5] However, even if we cannot identify the causality, it is our view that detailed

[1] Source: https://www.covid19india.org/ (accessed on 26 December 2020).

[2] See Susskind and Vines (2020) for the comprehensive review.

[3] The states are Andaman and Nicobar Islands, Andhra Pradesh, Arunachal Pradesh, Assam, Bihar, Chandigarh, Chhattisgarh, Delhi, Goa, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Karnataka, Kerala, Madhya Pradesh, Maharashtra, Manipur, Meghalaya, Mizoram, Nagaland, Odisha, Puducherry, Punjab, Rajasthan, Sikkim, Tamil Nadu, Telengana, Tripura, Uttar Pradesh, Uttarakhand, West Bengal. The selection is determined by the availability of the data.

[4] Based on Census 2011 estimates

[5] For instance, we interact dummy variables capturing different phases of lockdown policies with a state dummy (e.g. whether Maharashtra or not), but the causality rests on the assumption that the trend of the pandemic development is not different between Maharashtra and non-Maharashtra. However, while this 'parallel trend assumption' holds in a pre-lockdown (15 March -24 March) and Phase 1 (25 March-14 April), it unrealistic to assume that the macro environment is same for Maharashtra and other states and there were no physical interactions to influence the pandemic development.

analyses of the factors correlated with the pandemic development would provide some policy insights into policymakers.

The rest of the paper is organised as follows. The next section reviews the emerging literature on the COVID-19 pandemic with a particular focus on the study on the determinants of the COVID related deaths and illness as well as infections. Section 3 defines the severity ratios we use in the present study to capture the severity of pandemic and offers a statistical description of the data. Section 4 specifies econometric models we employ to assess the severity of the pandemic. Section 5 summarises main econometric results. Section 6 concludes with a few policy implications.

2. Literature Review

Despite a surge of the studies on COVID-19 in economics or social sciences, they are mostly about the impact or the consequences of COVID-19 pandemic. The empirical literature on the socio-economic factors associated with the COVID-19 pandemic or infections are still scarce in India or in other developing countries. This section provides a selective review of the emerging literature on the determinants of the COVID-19 pandemic in India.

Of particular interest is Joe et al. (2020) as it conducts a detailed statistical study of factors associated with the COVID-19 pandemic mortality. They use crowdsourced data (https://www.covid19india.org/) to provide preliminary estimates for age-sex specific COVID-19 case fatality rate (CFR) for India. CFR is estimated as the ratio of confirmed deaths in total confirmed cases.[6] Binomial confidence intervals are given for the CFR estimates. Also, an adjusted-CFR is estimated to capture the potential mortality among the currently active infections. Their main findings are as follows. As of May 20, 2020, males share a higher burden (66%) of COVID-19 infections than females (34%). However, the infection is more or less evenly distributed between males and females in under-five as well as elderly age groups where the CFR among males and females is 2.9% and 3.3%, respectively. The age-specific COVID-19 CFR reflects 'Nike-swoosh' pattern with elevated risks among the elderly. According to the World Health Organization, the CFR for India after standardisation based on the world standard population structure is 3.34%, while its adjusted-CFR is estimated to be 4.8% (Joe et al., 2020, pages 8-9). The authors conclude that (i) males have higher overall burden, but females have a higher relative-risk of COVID-19 mortality in India, and (ii) elderly males and females both display high mortality risk and require special care when infected. As the period which this study covers ends on May 20, 2020, well before the huge surge in COVID-19 cases - inevitably constrained by the timing of the study -, there is a need for covering a more recent period in order to inform policymakers of the results based on the updated data. Our study covers more recent periods and draws upon panel data modelling to allow for lockdown phase and unobservable state effects, such as cultural or institutional factors that are specific to each state and unlikely to change in the short-run.

[6] The results, however, should be interpreted cautiously as the incentives as well as the ability to take COVID-19 tests - which can be constrained by physical abilities or availabilities of the tests - can considerably differ across different sub-populations classified, for example, by age and gender.

In an innovative contribution, Banerjee et al. (2020) conducted a large-scale messaging campaign in West Bengal, India. Twenty-five million individuals were sent an SMS containing a 2.5-minute clip. All messages encouraged reporting symptoms to the local public health worker. Messages were randomized at the PIN code level. As control, three million individuals received a message pointing them to government information. The authors find that the campaign (i) doubled the reporting of health symptoms to the community health workers; (ii) reduced travel beyond one's village in the last two days and, increased estimated handwashing when returning home; (iii) spilled over to behaviours not mentioned in the message – for example, mask-wearing- but increased slightly while distancing and hygiene both increased in the sample where they were not mentioned by similar amounts as where they were mentioned; (iv) spilled over onto non-recipients within the same community, with effects similar to those for individuals who received the messages. While their findings imply that better health systems and information are key to mitigating the spread of the COVID-disease, they may also increase the cases to be reported due to the increased awareness or the lower costs to visit hospitals. This is an important dimension in our study, but to our knowledge, the data to capture the quality and the quantity of health infrastructure and services at state levels (i.e. 'input' variables in the health production function) are unavailable. We use the output variable in the health production function, namely, the ratio of morbidity over 60 years old which tends to be correlated with the health infrastructure and services. We also use lagged retail price of food commodities for wheat to capture the availability and the access to nutrients[7], which are closely associated with the food security, another important ingredient of health. In one specification, we have controlled for the sex ratio, the number of females per 1000 males, at state levels.[8]

As reviewed by Das et al. (2020), recent studies on the determinants of COVID-19 predominantly focus on the meteorological variables (e.g. Ma et al., 2020) and few studies focus on socio-economic determinants. After controlling for temperature and moisture indices, Das et al. have found that the living environment deprivation (in terms of housing conditions, asset possession and water access/ population and household density).was an important determinant of spatial clustering of COVID-19 hotspots in Kolkata megacity. While we cannot include such detailed data for our study at the national level, we control for weekly temperature and rainfall as well as the ratio of urbanisation at state levels.

It is evident that socio-economic factors influence the COVID-19 pandemic and infections, but virtually no studies have taken them into account in India[9], particularly at the national level. An important exception is Olsen et al. (2020) who have estimated a hierarchical and multilevel

[7] We have also tried lentil and rice prices, but have not found any statistically significant associations.

[8] Drawing on the large panel of their own survey data from eight OECD countries, Galasso et al. (2020) found that women are more likely to perceive COVID-19 as a very serious health problem and to comply with public policy measures.

[9] Outside India the rigorous studies on socio-economic determinants of the COVID-19 are scarce. Khalatbari-Soltani et al. (2020) reviewed 29 studies across different countries which reported the characteristics of patients with COVID-19 and their potential risk factors, only one study reports the occupational position of patients with mild or severe disease. The authors conclude that there is a need for the studies on socio-economic factors associated with the COVID-19 pandemic.

model to estimate the determinants of the risk of death due to COVID-19 in 11 states of India[10] taking into account the factors at both individual and district levels. The authors combined the National Family Health Survey for 2015/16, Census data for 2011, and estimates of COVID-19 deaths cumulatively up to June 2020 from How India Lives. Olsen et al. (2020) found that people living in urban areas, belonging to the Scheduled Caste, being smokers, who are males with more exposures to activities outside home, above 65 years have a higher risk for the COVID-19 death. While our study cannot incorporate all the factors, it will cover a few important variables, such as urbanisation, morbidity above 60 years and income per capita.

Acharya and Porwal (2020) have constructed the aggregate vulnerability index at state and district levels based on National Family Health Survey Data in 2015/16 with a focus on several dimensions, such as demographic and socio-economic variables. They found that among eight states that have contributed to over 80% of the confirmed COVID-19 cases in India as of June 17, 2020, five states had a high vulnerability index value and the remaining three had medium vulnerability (e.g. Maharashtra with 33% of the total COVID cases and the vulnerability index 0.829, the seventh from the bottom). Though Acharya and Porwal have not estimated the vulnerability index using the actual COVID-19 data, their analysis implied the importance of socio-economic factors, which is consistent with Olsen et al. (2020).

Our study builds on the existing literature on the determinants of the COVID-19 in India in some important ways. First, our study extends the analysis to 31 October, 2020, and thus helps us capture the surge in the Covid pandemic. We use a measure of COVID-19 severity, namely, the cumulative severity ratio (CSR). 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 first difference of CSR is taken to capture 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 in monitoring the progression of the pandemic-whether it is intensifying, weakening or unchanged. We use panel models that allow use of time-invariant fixed effects.

3. Data and Variables

(1) Definitions of Severity Ratios

A new indicator 'relative severity' proposed by the World Bank illustrates the unequal distribution and progression of covid-19 deaths across states (Schellenkens and Sourrowuille , 2020). 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 provide a state-specific measure of the severity of the pandemic. Given that a majority of COVID-19 deaths occurred in hospitals,

[10] The justification for omitting the Southern and North-eastern regions, where social norms, the social structure and cultural norms are rather distinctive, is not persuasive as unobservable state effects would capture these differences. This comment is also applicable to the omission of the cities of Bangalore and Chennai in South India, as also omission of the states of Kerala and West Bengal, due to the historic differences in state level policy and politics. The present study covers these omitted regions.

CSR is likely to be highly correlated with the excess burden on the health system. In addition to this ratio (which will be denoted as Cumulative Severity Ratio, or CSR), Schellenkens and Sourrowuille have also defined a Daily Severity Ratio (DSR) which tracks the progression of the severity of the pandemic in each region. To calculate the DSR, the number of COVID-19 deaths on a particular day are divided by the expected daily deaths under the assumption of no-pandemic, i.e. annual deaths divided by 365 (in a pre-COVID year). We have modified CSR and DSR to capture excess mortalities. CSR has been re-defined as the ratio of the sum of 'accumulated COVID-19 oriented death numbers and the expected death numbers' to 'the expected deaths from all causes' in a certain period. Likewise, DSR is modified as the ratio of 'the sum of daily COVID-19 death numbers and the expected daily death numbers' to 'the expected daily death numbers.'

Algebraically,

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

where

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

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

The state-wise Covid-19 data are collated from Ministry of Health and Family Welfare, India.[11] The data on past mortality patterns is based on the State-wise Number of Registered Deaths in 2017 from the Ministry of Health and Family Welfare, Government of India.[12] For the purpose of the deriving CSR, the number of reported deaths in 2017 is scaled down from annual estimates to the length of the pandemic in each state, calculated as the number of days since the first death in the state till the data point (t), with the cut-off date 31 June 2020. For the DSR, the denominator used in the ratio is total number of deaths in each region in 2017 / 365.[13] As

[11] The data are available from https://www.mohfw.gov.in/ (accessed on 30/12/2020). COVID-19 death figures were collected by the Integrated Disease Surveillance System, which should effectively cover the deaths which were clearly due to COVID-19. However, the figures may be underestimated because 'suspected' or 'probable' deaths were excluded (Chatterjee, 2020). While it is ideal for us to correct for any measurement errors, we do not have access to any other data sources to make any adjustment and we have decided to use the official death figures.

[12] In India, where death registration is incomplete and where most deaths are not assigned a cause of death by a trained medical professional, the CFR is unreliable. Although it has improved, one in five deaths in 2017 was not registered in any vital statistics database. Deaths where a cause of death was identified were an even smaller fraction of total deaths-less than one out of five. Additional problems are caused by delays in compiling results and missing information. The last year for which causes of death or death registration statistics were compiled was 2017 (Gupta, 2020). Two points are pertinent. First, registered deaths are underestimated as more than a minor fraction are not registered. There is, however, no reason to believe that underestimation was higher in 2017 than in any previous year. Indeed, the underestimation has reduced. Since more recent data on registered deaths is not available, we had no option but to rely on registered deaths in 2017. Second, for the reasons stated, the data by cause of-death are worse in terms of reliability. This of course does not matter as we rely on all cause-deaths.

[13] A question is whether the death numbers in 2017 would 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/1,000 population since

we discuss later, we will use as a dependent variable the first difference of CSR for the weekly panel as its level is non-stationary and both the level and the first difference of CSR for the monthly panel. Descriptive statistics of the variables are presented in Appendix Table 1.

2012 and the year 2017 is not an exceptional year. Second, while India has experienced frequent and widespread droughts, there was no major drought in 2017. The death numbers in 2017 would thus serve as a reasonable counterfactual for the present analysis of Covid-19 (https://www.indexmundi.com/, accessed on 18 July 2020).

Figure 1. Trend of Cumulative Severity Ratio – Selected States (13-03-2020-31-10-2020 (%)

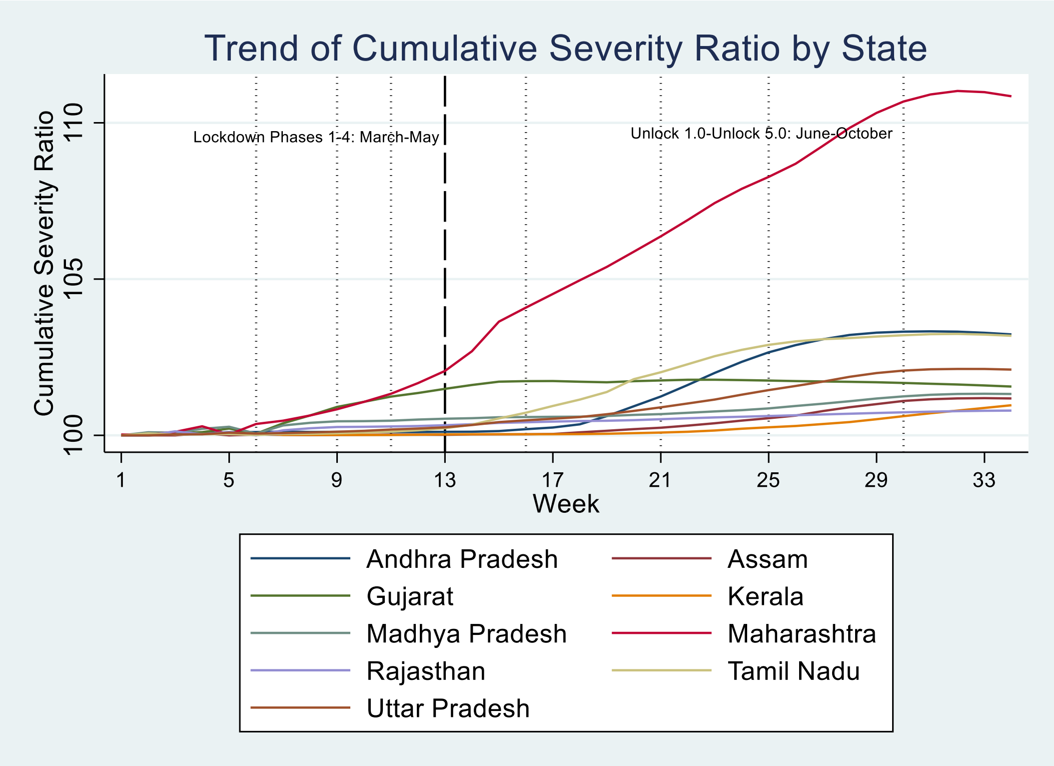
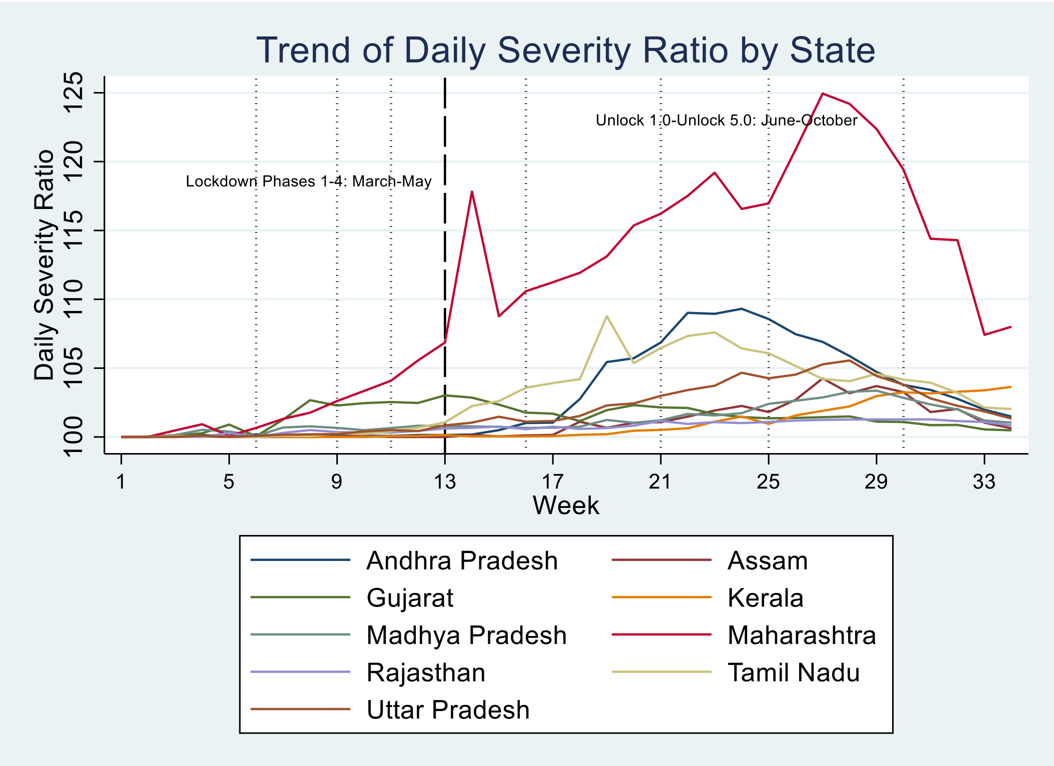


Figure 2. Trend of Daily Severity Ratio – Selected States (13-03-2020-31-10-2020)(%)



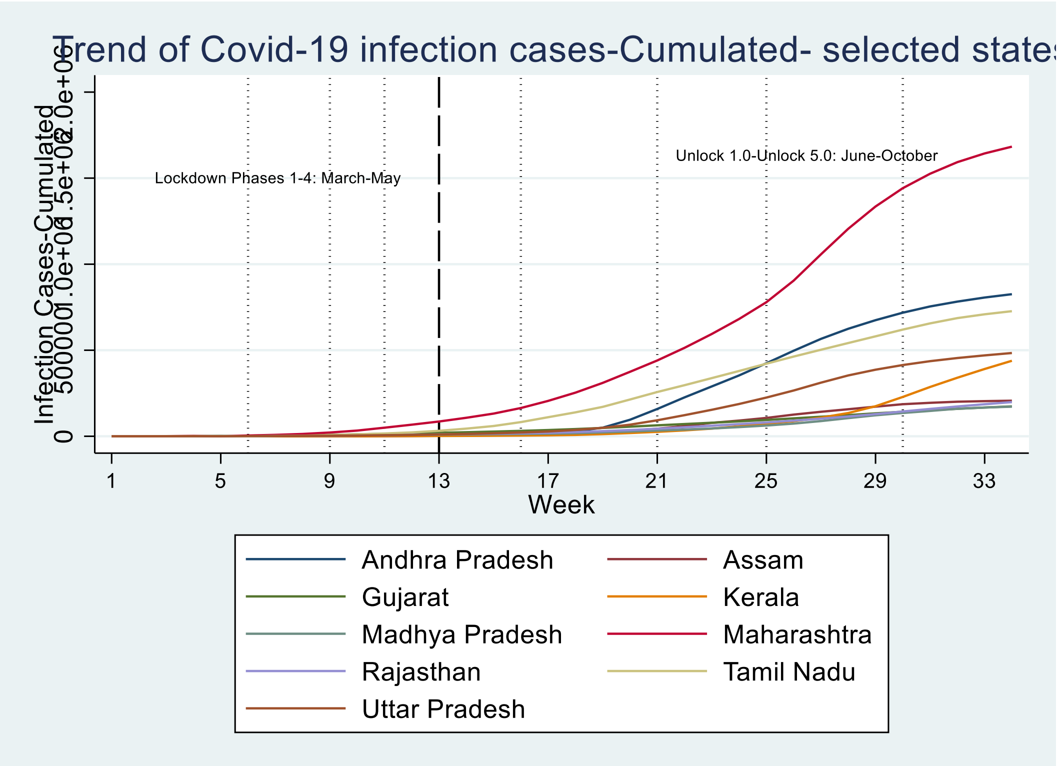
Figures 1 and 2 show the trends of CSR and DSR for relatively large states to avoid cluttering of the graphs. CSR and DSR are aggregated for each week, from Week 1 (starting on 13 March 2020) to Week 34 (on 29 October 2020). It is noted that during this study period the Indian government made serious efforts to prevent the spread of COVID-19 starting from a draconian lock down policy to relatively loose restrictions. The entire period is roughly divided into 'Lockdown' phases from March to May and 'Unlock' phases from June to October, as indicated by a dashed line in the graphs. The former is divided into 4 phases: Phases 1 to 4 and the latter into 5 Phases: Unlock 1.0 to Unlock 5.0 as shown by dotted lines in the graphs. The first lockdown (Phase 1) spanned a period of 21 days from 25 March to 14 April in which nearly all factories and services were suspended, barring "essential services". The second lockdown (Phase 2) started on 15 April and continued until 3 May, with conditional relaxations for regions where the Covid-19 spread had been contained. With additional relaxations, the phase three of the lockdown (Phase 3) was from 4 May to 17 May, and the fourth phase (Phase 4) was from 18 May to 21 June. Unlock 1.0 (1-30 June) was the first phase of the reopening in stages, with an economic focus where shopping malls, hotels and restaurants reopened. In Unlock 2.0 (1-31 July) the lockdown measures were restricted only to the contaminated zones and some inter- and intra-state travels were permitted. Further removing restrictions (e.g. night curfews) occurred Unlock 3.0, while Maharashtra and Tamil Nadu imposed a lockdown (1-31 August). Unlock 4.0 (1-30 September) was characterised by permissions of gathering at marriages/funerals, while wearing face masks became compulsory in public places and Unlock 5.0 (1-31 October) by opening cinemas and a gradual restarting of onsite teaching at schools at the discretion of state governments. How these government policies effectively influence the COVID-19 infection cases or fatalities are debatable and essentially an empirical question. Some authors have constructed the panel data across different countries and have estimated the effects of government policies on the COVID-19 infections. For instance, Chen et al. (2020) estimated the effects of various non-pharmaceutical interventions by governments to prevent the spread of COVID-19 on the country-level effective reproductive rate (Rt) for a panel of 75 economies and have found that, while lockdown measures lead to reductions in Rt, gathering bans are more effective than workplace and school closures. How these policies are effective in India remains uncertain, which would justify our focus on different phases.

We observe in Figures 1 and 2 a gradual increase in both CSR and DSR from the latter half of Phase 1 in Maharashtra and Gujarat. However, Maharashtra has seen a continuous rise in both CSR and DSR until Unlock 4.0-5.0 where CSR exceeded 110%. DSR reached 125% in Unlock 4.0. Evidently, Maharashtra has experienced the severest pandemic. However, the state has seen a gradual decline in DSR from mid-September to October 2020. On the other hand, CSR remained stable at around 102% in Gujarat from June to October. DSR has also remained stable in Gujarat after late July. Tamil Nadu experienced a sharp rise in CSR in July and August (Unlock 2.0-3.0). Its DSR became the second worst next to Maharashtra from mid-June to the end of July with a gradual decline after mid-August.

Andhra Pradesh saw a rise in CSR from early July. Its CSR became the second highest roughly at around 103% next to Maharashtra on 18 September. DSR in Andhra Pradesh was the second highest in late July to early October with its peak nearly 110% in late August. DSR has declined since then. Uttar Pradesh has seen a rise in CSR from July to October. Other states in the graphs, namely, Madhya Pradesh, Rajasthan, Assam, and Kerala have experienced a gradual increase in CSR, but the pandemic measured by CSR or DSR were not as severe as the states mentioned above. We observe a large variation in levels of the severity across different states.

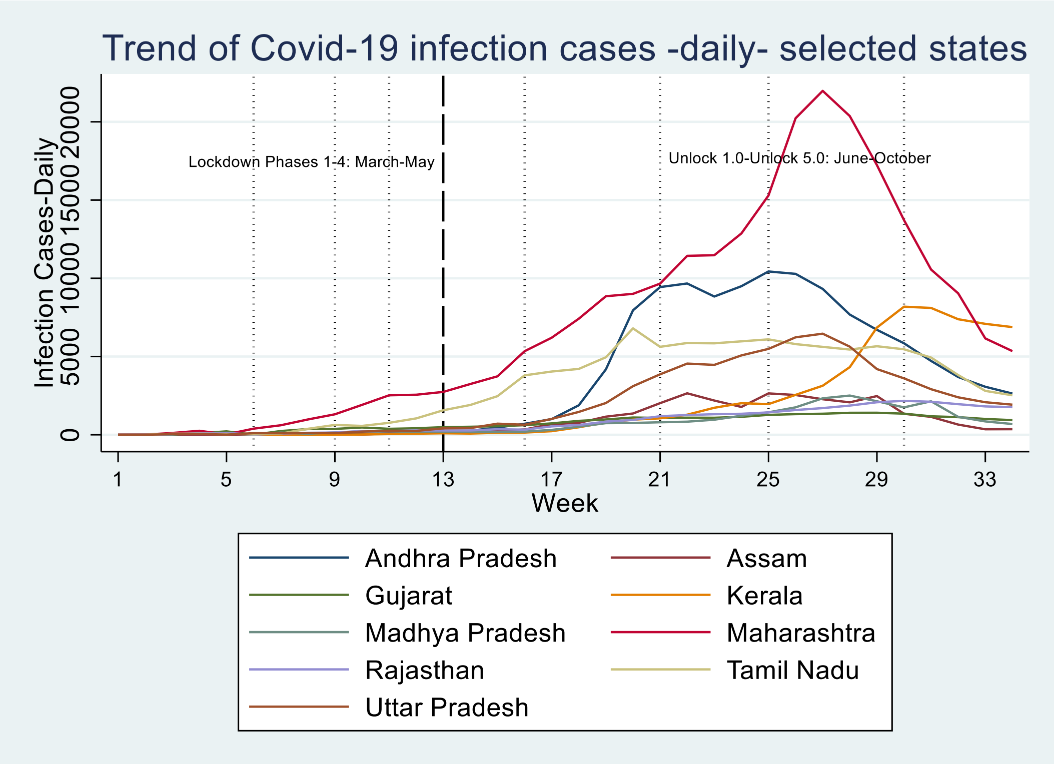
We see large variations in other states/union territories not highlighted in Figures 1-3. For instance, in Goa, CSR increased from 102% in July to 113% in October 2020. Sikkim's CSR remained at 100% (i.e. no extra mortality due to COVID-19) until the end of July, but its CSR has suddenly risen afterwards and climbed to 115% at the end of October. Puducherry has shown a similar trend with a rise in CSR in August and September and CSR surged to 110%. In Uttarakhand CSR gradually increased from 101% in July to 106% in October. Jammu and Kashmir has seen a similar rise in CSR from 101% to 106% in July-October due to a surge in DSR in the same period. On the other hand, CSR and DSR remained very low, such as Odisha and Mizoram.[14]

Figure 3. Trend of Cumulative COVID-19 infection cases (13-03-2020-31-10-2020)



[14] A full set of results will be provided on request.

Figure 4. Trend of Daily COVID-19 infection cases – Selected States (13-03-2020-31-



Figures 3 and 4 report the trends of cumulative and daily infection cases of COVID-19 on the basis of weekly averages in selected states. In terms of infection cases, Maharashtra has experienced by far the severest pandemic, although the number of daily infection cases has started to decline after 18 September. On the other hand, Kerala's daily infections suddenly rose from mid-September to the end of October - leading to a steep increase in cumulated cases. In other states in Figures 3 and 4, daily cases were highest in September and started to decline marginally in October. Most of the other states have shown similar trends of cumulative and daily cases where the latter declined gradually (Appendix Figures 1-3). One notable exception is West Bengal where daily cases continue to rise in September and October. The daily cases have exceeded 4000 and DSR has spiked to 104% in October in West Bengal.[15]

Given the curves in Figures 1 and 2, we have carried out unit-root tests for CSR based on the weekly panel data. To normalise the infection cases we have taken the logarithm of the number of cases and applied the unit-root tests to it. Table 2 gives the results of the panel-unit root tests for CSR, its first difference, the log of cases as well as the log of retail prices series of wheat, one of the explanatory variables. We apply Levin-Lin-Chu (LLC) (Levin et al., 2002) and Im-Pesaran-Shin (IPS) tests (Im et al, 2003). LLC tests the null hypothesis that each time series contains a unit root against the alternative hypothesis that each time series is stationary in which the lag order is permitted to vary across individuals. IPS test is not as restrictive as the LLC test, since it allows for heterogeneous coefficients. The null hypothesis is that all individuals follow a unit root process against the alternative hypothesis allowing some (but not all) of the individuals to have unit roots. We apply the specifications with and without a time trend. We

[15] However, the cases gradually declined in November and December in West Bengal in more recent periods not covered by this study.

determine the number of lags by Akaike Information Criteria (AIC).[16] Three states with missing observations (Kerala, Meghalaya and Punjab) have been dropped to make the panel balanced.

Table 1 shows that CSR is I(1) (non-stationary) as its first difference is stationary. The log of cases and the log of wheat prices are stationary. Given that CSR is not stationary, the OLS or the static panel data model, such as fixed-effects or random-effects models cannot be applied. As all the explanatory variables – including the wheat prices and weather variables - are stationary, they are not co-integrated. So we will use the first difference of CSR or the log of infection cases as a dependent variable for the weekly panel. We have also taken the monthly averages of the data and constructed the monthly panel – where the stationarity is not an issue due to a small T (Pesaran, 2011). For the monthly panel we use the level of CSR, its first difference, or the log of cases as a dependent variable.

[16] We have also applied other alternatives of panel unit root tests and the results are broadly similar.

Table 1. Results of Unit-root Tests for Weekly Panel

		Levin-Lin-Chu (LLC) no trend		Levin-Lin-Chu (LLC) with trend		Im-Pesaran-Shin (IPS) no trend		Im-Pesaran-Shin (IPS) no trend	
Panel structure	N (no of centres)	29		29		29		29	
	T (no of periods)	34		34		34		34	
	Panel means	No		No		No		No	
CSR	Average lags [*1]	0.9		0.86		0.93		1.62	
(level)	adjusted t or W-t-bar[*2]	4.48		-1.43	*	7.02		1.96	
		I(1)		I(1)		I(1)		I(1)	
CSR (first difference)	Average lags [*1]	0.41		0.24		0.45		0.48	
	adjusted t or W-t-bar[*2]	-15.45	***	-16.57	***	-14.63	***	-13.82	
		I(0)		I(0)		I(0)		I(0)	
log Cases	Average lags	3.72		3.93		0.68		0.79	
	t (adjusted)	-5.09	***	-4.64	***	-8.69	***	-9.003	***
		I(0)		I(0)		I(0)		I(0)	
Wheat Price (retail, log)	Average lags [*1]	0.48		0.45		0.69		0.55	
	adjusted t or W-t-bar[*2]	-12.66	***	-13.47	***	-12.72	***	-12.81	***
		I(0)		I(0)		I(0)		I(0)	

Notes: 1. Lags are determined by Akaike Information Criteria (AIC).

*2. adjusted t is reported for LLC and W-t-bar is reported for IPS.

*3. The threshold signficance level is at 5%.

4. Regression analyses on the determinants of the severity of COVID-19 pandemic

(a) Model Specification

Towards an explanation of the regional variation in the severity of COVID-19 pandemic, we use a panel of 32 states/union territories for which the data on various variables are available covering the period from 13 March 2020 to 31 October 2020. As noted earlier, we have organised the data as weekly or monthly panels where all the variables on the daily basis are averaged for each week or month. Because of the missing observations on a few variables, our estimation is based on 1041 observations for the weekly panel (32 states times 32.5 weeks).

We regress a dependent variable, either the Cumulative Severity Ratio (CSR), the Daily Severity Ratio (DSR) or the confirmed COVID-19 infection cases on a number of explanatory variables. CSR captures overall development of the COVID-19 pandemic while DSR denotes how the severity progresses over time. We also have used the number of the infection cases given that a surge in the confirmed cases is closely associated with the COVID-19 pandemic with some lags.

We have selected the explanatory variables, while constrained by the data availability, to reflect the growing empirical literature on the causes of COVID-19 pandemic or infections. The time-variant explanatory variables are weather variables, namely, temperature and rainfall as well as the lagged commodity price (wheat retail price). We have also used a number of time-invariant variables such as the log of per capita income, urbanization, presence of more than one morbidity condition among those above 60 years and the sex ratio (the number of females per 1000 males). The model also includes a few phase dummies. The first phase dummy variable captures the first lockdown - a period of 21 days from 25 March 2020 to 14 April 2020 where nearly all factories and services were suspended, barring "essential services" (which serves as the baseline case in the model). The second 'lockdown dummy' takes 1 for the period from 15 April 2020 to 30 June 2020 and 0 otherwise. In this period, there has been staged relaxations of the lockdown, such as, re-opening of shops, while various restrictions were maintained.

Methodologically, we employ pooled OLS (with both state fixed effects and phase or month dummies), a random-effects model and a random-effects Tobit model so that we can model time-invariant unobservable state or union territory characteristics (e.g. institutional or cultural factors specific to each state or union territory. In a random-effects model, state and phase/month fixed effects are included by applying the mixed-effects model (Bell and Jones, 2015). In the meantime, as some states/union territories had zero cases or deaths in early periods, the random-effects Tobit model is also estimated as a robustness check to take account of left-censoring of the dependent variable in case we estimate either CSR or the log of cases.

We estimate the following equation. We have taken the logarithm of most of the explanatory variables to capture the relative effect, or the elasticity.[17]

$$DCSR_{it} = \beta_0 + \beta_1 \, log \, Per \, Capita \, Income_i + \beta_2 Multimorbidity \, above \, 60_i + \beta_3 Urbanisation_i + \beta_4 Sex \, Ratio_i + \beta_5 Wheat \, Price_{it-1} + \beta_6 Temperature_{it} + \beta_7 Rainfall_{it} + Phase \, (or \, Month) \, Dummies_t \, \beta_8 + \mu_i + e_{it} \, \ldots\ldots\ldots \, (1)$$

In Equation (1) i stands for state (from 1 to 32) and t for week from 13 March to 31 October 2020 (1 to 34) for the weekly panel data and March to October (1 to 8) for the monthly panel data. We have taken the averages of daily data for each month and have constructed the monthly panel data. A dependent variable is $DCSR_{it}$ (the first difference of CSR) or $logCovidCases_{it}$ (the log of the daily infection cases- averaged over a week) - both of which are I(0) - for the weekly panel data and CSR_{it}, $DCSR_{it}$ or $logCovidCases_{it}$ (the log of infection cases averaged over a month) for the monthly panel as in Equation (2).[18] We have also regressed DSR_{it} (the average of daily severity ratio, the flow measure) for the monthly panel to see if the results are similar to those for $DCSR_{it}$ (monthly changes in the cumulative severity ratio, the stock measure).

$$logCovidCases_{it} = \beta_0 + \beta_1 \, log \, Per \, Capita \, Income_i + \beta_2 Multimorbidity \, above \, 60_i + \beta_3 Urbanisation_i + \beta_4 Sex \, Ratio_i + \beta_5 Wheat \, Price_{it-1} + \beta_6 Temperature_{it} + \beta_7 Rainfall_{it} + Phase \, (or \, Month) \, Dummies_t \, \beta_8 + \mu_i + e_{it} \, \ldots\ldots\ldots \, (2)$$

[17] For the descriptive statistics of the variables used, see Appendix Table 1.
[18] The state-wise estimates of daily confirmed COVID-19 Cases are taken from the official website of the Ministry of Health and Family Affairs, Government of India.

The selection of explanatory variables is guided by the emerging empirical literature on the determinants of the COVID-19 pandemic reviewed in Section 2 where not only meteorological factors but also socio-economic and demographic factors are closely associated with the degree of the COVID-19 pandemic or infections. $Per\ Capita\ Income_i$(PCI) denotes income at state level that is measured by per capita net state domestic product (in Rs., divided by 1000).[19] PCI captures not only 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. We include the proportion of elderly people who suffer from more than one non-communicable diseases (NCDs) at state levels ($Multimorbidity\ above\ 60_i$).[20] This is the proportion of population in the age group 60+ reporting more than one NCD (e.g. cardiovascular diseases, diabetes, hypertension, among others). To capture the degree of urbanisation, we also insert $Urbanisation_i$, the share of the population living in urban areas. The idea is that a higher population density and urbanization would increase interactions among people and raise both CSR and DSR. Furthermore, we have inserted $Sex\ Ratio_i$ (the number of females per thousand male), as it is well documented that, while COVID-19 infection rates are broadly similar between men and women, men are more likely to suffer from severe illness or die as a result of COVID infections in China (Jin et al., 2020) and in Europe (Gebhard et al., 2020). However, given the preference of boys over girls in many states of India, more developed States with lower poverty (e.g., Kerala) tend to have a higher sex ratio and these states may have a better health system. So the effect of the sex ratio on the COVID-19 may be ambiguous in India.[21] We also control for the effect of the retail price of wheat to examine whether the lagged food price has any association with the COVID-19 pandemic. An increase in the wheat price may lead to the difficulty of accessing food or a macronutrient where lower calorie intake could impair immunity, but in the meantime, it may induce substitution into inferior cereals, such as, *ragi* or maize, which may result in better nourishment (Gaiha et al., 2014). Our results are consistent with the latter hypothesis.

It is widely debated whether weather influences the COVID-19 infection cases and/or linked deaths. A recent study used the data on daily death numbers from Wuhan, China, in January-February in 2020 and found that death counts are positively associated with temperature and negatively with relative humidity (Ma et al., 2020). We have collected the daily data on temperature, rainfall, and relative humidity from MERRA (Modern-Era Retrospective analysis for Research and Applications – Version 2 web service) and have taken either week or month averages. It delivers time series of temperature (at 2m), Relative humidity (at 2m) and rainfall. The data source is a NASA atmospheric reanalysis of the satellite era using the Goddard Earth Observing System Model (GEOS-5) and focuses on historical climate analyses for a broad

[19] The data on state per capita incomes is obtained from the state economic surveys, and demographic data including population density and urban population are taken from the Census estimates.

[20] The data on this variable are based on authors calculations from *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 were completed in 2004-5; and a second round of IHDS reinterviewed most of these households in 2011-12. IHDS has been jointly organized by researchers from the University of Maryland and the National Council of Applied Economic Research (NCAER), New Delhi.

[21] To avoid multicollinearity one specification includes either 'per capita income and multi-morbidity or 'urbanisation and the sex ratio' as covariates.

range of weather and climate time scales (GMAO, 2015). Due to the high correlation between rainfall and relative humidity, we use the variables, $Temperature_{it}$ and $Rainfall_{it}$.

To capture the time and policy effects, we have included eight dummy variables for Lockdown Phases 2 to 4 and Unlock 1.0 to 5.0 for the weekly panel and seven monthly dummies for the monthly panel. This is aimed to capture the associations of severity of COVID-19 with the lockdown and unlock policies announced by the Government of India. Equation (1) has been estimated by pooled OLS with state and phase/month fixed effects, random-effects model or mixed-effects model (Bell and Jones, 2015) and random-effects Tobit model with phase/month fixed effects.

$$DlogCSR_{it} = \beta_0 + \beta_1\, D\_Maharashtra_i + Phase\ Dummies_t\, \beta_2 + D\_Maharashtra_i * Phase\ Dummies_t\, \beta_3 + e_{it} \ \ldots\ldots\ldots \ (3)$$

Given that the pandemic in Maharashtra has been by far the severest, Equation (2) is estimated to see how the development of the pandemic in Maharashtra differs from the other states. We estimated the case with $D\_Maharashtra_i$,and a vector of phase dummies (or month dummies) (i.e. without the interaction terms in equation (3)) as a reference case in order to interpret the estimated coefficients of β_3.

5. Results

We show the results of our regression analyses in Table 2, Table 3 and Table 4 corresponding to Equations (1), (2) and (3). The main findings are summarised below.

In Table 2, the results based on the weekly panel are shown in Columns 1 and 2 and those based on the monthly panels are in Columns 3-8. We have found that log per capita income state at the state level is positively associated with DCSR, weekly and monthly changes in cumulative severity ratio of COVID-19 (a proxy for the development of the pandemic) (Columns 1, 3, 5 and 7) as well as CSR, monthly CSR, after controlling for state and phase/month fixed effects. For instance, a 1% increase in per capita income is on average associated with 0.33% increase in the change in CSR (Column 1). As CSR is measured in percentages, and not in the logarithm, this increase is substantial and implies that the state with a higher income tends to see a faster change in CSR or a more rapid escalation of the pandemic. The reason is associated with the fact that a higher income level tends to associate with more production, transportion and movement of people and goods even in the lockdown phases. Consistent results are found for the monthly data. A 1% increase in per capita income is associated with 4.4%-5.6% increase in CSR and a nearly 2% increase in the change in CSR on a monthly basis. If we replace CSR with DSR, we find that a 1% income increase is significantly associated with a 5.4% increase in DSR.[22]

[22] The results on DSR are not shown but will be provided on request.

We have also found positive association between CSR or DCSR and multi morbidity for monthly data, but not weekly data. The estimated coefficient varies across different models, but for instance, based on Column 4 on CSR, we observe that a 1% increase in urbanisation is associated with a 0.36% increase in CSR as consistent with Das et al. (2020) and Olsen et al. (2020). As expected, we find that the share of those among the elderly with multi-morbidity conditions is positively associated with DCSR or CSR (e.g. a 1% increase in the share tends to lead to a 0.54%-0.67% increase in CSR, Columns 3 and 5). Similar results are obtained for DSR.

If women's number per 1000 men decreases by 1, this is on average associated with 0.05-0.06% increase in CSR (or 0.02 increase in DCSR). A consistent result has been found for DSR as well. However, the sign is reversed in Table 2. That is, the higher ratio of women is associated with the higher level of infection but the lower level of fatalities. Whether this reflects any gender difference in the risks for infection and fatalities is not clear, but the results indicate that demography is an important determinant of the COVID-19 pandemic.

Lagged retail price of wheat is negatively and correlated with DCSR or CSR and the estimates are statistically significant in all the cases except Columns 3 or 4 (RE model applied to the monthly panel). Columns 1 and 2, based on weekly data, and 7 and 8, based on monthly data, show a similar level of coefficient estimates. A 1% increase in wheat price tends to lead to 0.17-0.19% decrease in changes in CSR, while the estimated coefficient of the Tobit model suggests that a 1% price increase is correlated with a 4.5-4.6% decrease in CSR (in level). The results overall suggest a negative effect of wheat prices on the CSR which could be due to shift to cheaper and more nutritious cereals. However, once we take the second or third lags, the coefficient estimates are negative but not significant.

We have controlled for temperature and rainfall to reflect the empirical literature on the determinants of the COVID-19 pandemic. While the estimated coefficient of temperature is positive and that of rainfall is negative in all the cases, we only find a positive and statistically significant estimate for temperature in Columns 5 and 6 (Tobit for CSR) and a negative and significant estimate for rainfall in Columns 7 and 8 (Random effects model for DCSR). We refrain from inferring any associations between climatic conditions and the pandemic once time and state effects are accounted for. Table 2 also shows coefficient estimates of state dummies for selected states. They do not necessarily match the rankings of CSR in Figure 1 or Appendix Figures 1-2, as estimated coefficients of state dummies have been obtained after conditioning on other covariates, such as per capita income. However, Maharashtra tends to have a higher coefficient estimate when DCSR or CSR statistically significant (e.g. Columns 1, 3 and 5). Phase or month dummies show that not only the level of CSR but also its change tends to increase in later periods, which implies that the pandemic has worsened over time. A decrease in DCSR from Unlock 4.0 to Unlock 5.0 (Columns 1 and 2) indicate that worsening of the pandemic slowed down in October.

Table 2. Determinants of Cumulative Severity Ratio of COVID-19

	(1).	(2).	(3).	(4).	(5).	(6).	(7).	(8).
Data	Weekly	Weekly	Monthly	Monthly	Monthly	Monthly	Monthly	Monthly
Dependent Variable Level/First Difference	FD	FD	Level	Level	Level	Level	FD	FD
	Cumulative Severity Ratio	Cumulative Severity Ratio	Cumulative Severity Ratio	Cumulative Severity Ratio	Cumulative Severity Ratio	Cumulative Severity Ratio	Cumulative Severity Ratio	Cumulative Severity Ratio
Model	Random Effects [1]	Random Effects	Random Effects	Random Effects	Tobit	Tobit	Random Effects	Random Effects
Explanatory Variables	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)
log Per Capita Income	0.329 [2,3]		4.473		5.599 [2,3]		1.955	
	(5.04) ***		(4.62) ***		(3.83) ***		(8.71) ***	
Multi-morbidity * (%)	0.025		0.536		0.669		0.215	
	(4.74) ***		(3.13) ***		(3.59) ***		(6.50) ***	
Rate of Urbanisation (%)		0.01		0.355		14.63		0.129
		(1.08)		(2.02) **		(1.83) *		(3.28) ***
Sex Ratio		-0		-0.05		-0.06		-0.021
		(4.01) ***	(0.00)	(4.67) ***	(0.00)	(2.75) ***		(6.28) ***
log Wheat Prices(-1)	-0.185	-0.18	-1.781	-1.62	-4.603	-4.47	-1.817	-1.758
	(1.95) *	(1.88) *	(0.75)	(0.66)	(2.20) **	(2.08) **	(3.99) ***	(3.81) ***
Temperature	0.005	0.01	0.097	0.097	0.173	0.178	0.026	0.026
	(0.89)	(0.70)	(1.34)	(1.31)	(2.29) **	(2.25) **	(0.83)	(0.77)
Rainfall	-0.001	-0	-0.017	-0.01	-0.005	-0	-0.026	-0.025
[Selective State Dummies][1]	(0.27)	(0.25)	(0.46)	(0.37)	(0.16)	(0.05)	(1.80) *	(1.69) *
D_Maharashtra [4]	0.224	0.15	4.376	-0.88	5.521	-0.25	1.703	-0.166
	(5.58) ***	(1.77) *	(3.39) ***	(0.59)	(3.64) ***	(0.10)	(6.47) ***	(0.41)
D_Andhra Pradesh	0.02	0.25	0.201	2.2	1.168	3.776	0.47	1.463
	(0.58)	(3.01) ***	(0.20)	(1.85) *	(0.99)	(1.86) *	(2.47) **	(5.39) ***
D_Assam	0.19	0.16	3.368	6.476	4.875	12.85	1.601	2.605
	(3.33) ***	(1.44)	(2.50) **	(1.89) *	(2.82) ***	(1.98) **	(6.12) ***	(3.72) ***
D_Gujarat	-0.18	-0.22	-1.187	-4.84	-1.328	-5.44	-0.8	-2.086
	(4.93) ***	(3.91) ***	(2.66) **	(2.89) ***	(1.30)	(2.42) **	(4.15) ***	(5.85) ***
D_Kerala	-0.933	0.44	-18.5	0.572	-21.65	3.593	-7.009	1.313
	(6.53) ***	(1.70) *	(4.05) ***	(0.26)	(4.17) ***	(0.65)	(7.93) ***	(1.41)

D_Madhya Pradesh	**0.115**		**-0.07**		**2.218**		0.442		**3.062**		1.016		**0.888**		0.035	
	(2.90)	***	**(1.99)**	**	**(2.64)**	**	(0.54)		**(2.24)**	**	(0.71)		**(5.84)**	***	(0.17)	
D_Rajasthan	0.065		-0.09		**2.018**		0.752		**2.776**		1.672		**0.685**		0.038	
	(1.52)		(1.47)		**(1.76)**	*	(0.57)		**(1.90)**	*	(0.76)		**(2.89)**	***	(0.11)	
D_Tamil Nadu	**-0.127**		0.16		**-2.068**		-2.67		**-1.558**		-0.97		**-0.534**		-0.338	
	(7.93)	***	(0.97)		**(4.57)**	***	(1.39)		**(1.76)**	*	(0.24)		**(5.32)**	***	(0.52)	
D_Uttar Pradesh	**0.357**		-0.06		**5.826**		**6.314**		**6.341**		6.616		**2.011**		1.596	
	(6.86)	***	(0.27)		**(7.69)**	***	**(1.80)**	*	**(3.47)**	***	(1.07)		**(12.93)**	***	(1.62)	
D_Lockdown Phase 2	0.013		0.02		**-4.049**		**-4.11**		**-4.947**		**-5.02**		**-0.914**		**-0.921**	
(D_April)[8]	(0.57)		(0.67)		**(4.71)**	***	**(4.68)**	***	**(9.22)**	***	**(9.09)**	***	**(2.84)**	***	**(2.78)**	***
D_Lockdown Phase 3	0.045		0.05		**-4.06**		**-4.13**		**-4.791**		**-4.89**		**-0.833**		**-0.843**	
(D_May)[8]	(1.03)		(1.11)		**(4.20)**	***	**(4.24)**	***	**(7.80)**	***	**(7.77)**	***	**(2.41)**	**	**(2.36)**	**
D_Lockdown Phase 4	-0.029		-0.03													
	(0.69)		(0.58)													
D_Unlock 1.0	0.044		0.05		-3.781		-3.88		-4.335		-4.47		-0.653		-0.683	
(D_June)[8]	(0.91)		(0.87)		(4.03)	***	(4.05)	***	(7.18)	***	(7.16)	***	(1.87)	*	(1.89)	*
D_Unlock 2.0	0.082		0.08		-3.386		-3.5		-3.906		-4.05		-0.406		-0.434	
(D_July)[8]	(1.48)		(1.42)		(4.02)	***	(4.05)	***	(6.72)	***	(6.73)	***	(1.17)		(1.21)	
D_Unlock 3.0	0.37		0.38		-2.278		-2.37		-2.618		-2.73		0.354		0.346	
(D_August)[8]	(1.98)	**	(1.97)	**	(3.63)	***	(3.68)	***	(4.50)	***	(4.55)	***	(0.58)		(0.55)	
D_Unlock 4.0	0.205		0.21		-0.961		-1.01		-1.187		-1.25		0.52		0.52	
(D_September)[8]	(3.65)	***	(3.61)	***	(2.31)	**	(2.39)	**	(2.31)	**	(2.36)	**	(0.97)		(0.94)	
D_Unlock 5.0	0.118		0.12													
(October)[8]	(2.61)	**	(2.62)	**												
Constant	-5.021		2.62		25.289		110.9		-2.267		69.03		-25.23		13.8	
	(2.21)	**	(2.48)	**	(1.15)		(4.80)	***	(0.08)		(1.83)	*	(2.38)	**	(1.92)	
State Fixed Effects[1]	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
No of Observations(N)	1041		1008		223		223		216		216		223		223	
(Left censored)									36		36					
No of states(n)	32		31		32		32		32		32		32		32	
No of weeks (T)	32.5		32.5		7		7		7		7		7		7	
Wald chi2	63.03		61.1		281		273.6		298.4		290.7		279.2		276	
(p value)	(0.02)	**	(0.02)	**	(0.00)	***	(0.00)	***	(0.00)	***	(0.00)	***	(0.00)	***	(0.00)	***
R squared within	0.0298		0.03		0.4416		0.446		-		-		0.184		0.186	
R squared betweem	1		1		1		1		-		-		1		1	
R squared overall	0.0571		0.06		0.6025		0.605		**-**		**-**		0.314		0.313	

Breush and Pagan Test	0	0	0	0	-	-	0	0
(p value)	(1.00)	(1.00)	(1.00)	(1.00)	-	-	(1.00)	(1.00)
Hausman Test [7]	0	0	0	0	-	-	0	0
(p value)	(1.00)	(1.00)	(1.00)	(1.00)	-	-	(1.00)	(1.00)

Notes: 1. State dummies or fixed effects for the other states have been included in all the cases. That is, the model has been estimated as by a mixed effects model.

2. *** = Significant at 1% level. ** = Significant at 5% level. * = significant at 105 level.

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

4. D_ stands for a dummy variable (taking 1 or 0).

6. Statistically signicaint cases are highlighted as bold numbers.

7. Hausman tests were carried out between FE and RE models.

8. Monthly dummies have been used instead of phase dummies in the case of monthly panel data.

Table 3 shows the results on infection cases. Here, based on the unit-root test results, log of infection cases is dependent variable in all the cases. Columns 9-12, based on the weekly panel, while Columns 13-16 on the monthly panel where both random-effects models and Tobit models are applied, since there are some states with no cases at the onset of the pandemic. It is notable to find that many of the parameter estimates on CSR or DCSR in Table 2 are reversed in Table 3. For instance, log per capita income is negative and significant in all the cases.[23] That is, if income increases by one percentage point, the number of cases tends to decrease by 2.8-3.0% with no causality implied by these results, after controlling for state fixed effects and phase/month dummies. Interpreting the results in Tables 2 and 3 together, a state with a higher income tends to experience the worse pandemic at relatively low case numbers on average. This is counter-intuitive at first sight if we assume that income leads to more interactions among people leading to more cases, but we may conjecture that a relatively rich state may be able to carry out more tests, but not necessarily have better capacity to cope with fatalities.

On the other hand, states with a higher share of the elderly with morbidity conditions tend to have lower COVID-19 cases (where a 1% increase of the former is associated with a 0.28% decrease in the cases), while the row (unconditional) correlation between the variable is positive. It is conjectured that, while morbidity conditions among the elderly can lead to fatalities once they are infected, they may not influence the probability of being infected at the population level. Urbanisation is not significantly associated with the number of cases (except a negative and significant coefficient based on Tobit, Column 12). Sex ratio is positive and significant, implying that the states with more females per 1000 males tend to have more infection cases (an increase of one woman per 1000 men is associated with 0.03% increase in the cases). As in Table 2, retail prices of wheat are negatively correlated with the log of infection cases where a 1% fall in wheat prices tend to lead to -1.3 to -1.7% increase in infection cases. It is conjectured that higher wheat price induces a shift towards inferior but more nutritious cereals (Gaiha et al., 2014).[24]

On the effect of weather, both temperature and rainfall have a positive association with COVID-19 cases, that is, hot and rainy weather conditions may lead to higher infection rates. One degree increase in temperature is associated with 0.13-0.17% increase in the number of cases on average, other factors held constant. In contrast, a 1 mm increase in rainfall is associated with 0.03-0.06% increase in the cases. Phase or month dummy variables show that the number of cases tends to be larger in later months or phases. State dummy variables show, after controlling for covariates (e.g. income), that Maharashtra, Andhra Pradesh, Gujarat, Kerala, and Tamil Nadu are the states which exhibit a higher number of infection cases than other states.

[23] The raw (unconditional) correlation coefficient between CSR and log of PCI is 0.153, while that between the log of cases and the log of OCI is -0.063.

[24] We have carried out sensitivity tests for taking different lags of wheat prices. If we take the second lag of wheat price, the results are broadly the same. If we take the third lag, the sign is negative and statistically significant for the weekly panel and negative and non-significant for the monthly panel.

Table 3. Determinants of COVID-19 Infection Cases

	(9).	(10).	(11).	(12).	(13).	(14).	(15).	(16).
Data	Weekly	Weekly	Weekly	Weekly	Monthly	Monthly	Monthly	Monthly
Dependent Variable Level/First Difference	Level	Level	Level	Level	Level	Level	Level	Level
	Cases (log)	Cases (log)	Cases (log)	Cases (log)	Cases (log)	Cases (log)	Cases (log)	Cases (log)
Model	Random Effects [1]	Random Effects	Tobit	Tobit	Random Effects [1]	Random Effects	Tobit	Tobit
Explanatory Variables	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)	Est. Coef. (Z value)
log Per Capita Income	-3.039 [2,3]		-3.016		-2.818 [2,3]		-2.815	
	(7.77) ***		(8.33) ***		(8.15) ***		(4.73) ***	
Multi-morbidity *	-0.278		-0.278		-0.28		-0.281	
(%)	(4.48) ***		(7.09) ***		(3.20) ***		(3.88) ***	
Rate of Urbanisation		-0.127		-0.13		-0.154		-0.156
		(1.40)		(2.19) **		(1.21)		(1.58)
Sex Ratio		0.035		0.035		0.031		0.031
		(5.22) ***		(5.97) ***		(4.25) ***		(3.41) ***
log Wheat Prices(-1)	-1.325	-1.301	-1.344	-1.32	-1.716	-1.678	-1.707	-1.668
	(1.91) *	(1.86) *	(4.44) ***	(4.29) ***	(2.06) **	(1.98) **	(2.33) **	(2.22) **
Temperature	0.135	0.133	0.138	0.136	0.168	0.17	0.169	0.171
	(1.97) **	(1.87) *	(8.21) ***	(7.75) ***	(1.76) *	(1.71) *	(5.35) ***	(5.18) ***
Rainfall	0.028	0.029	0.029	0.03	0.06	0.061	0.06	0.061
[Selective State Dummies][1]	(2.83) ***	(2.84) ***	(5.63) ***	(5.65) ***	(2.93) ***	(2.97) ***	(4.41) ***	(4.40) ***
D_Maharashtra [4]	3.473	5.104	3.463	5.142	3.513	5.602	3.501	5.624
	(6.50) ***	(5.22) ***	(10.00) ***	(6.61) ***	(4.58) ***	(4.16) ***	(5.72) ***	(4.54) ***
D_Andhra Pradesh	1.604	-0.343	1.602	-0.314	1.674	0.022	1.668	0.026
	(4.28) ***	(0.72)	(5.81) ***	(0.59)	(3.24) ***	(0.05)	(3.49) ***	(0.03)
D_Assam	-2.839	-3.412	-2.831	-3.452	-2.43	-3.487	-2.434	-3.521
	(7.00) ***	(2.26) **	(7.02) ***	(3.45) ***	(5.31) ***	(1.62)	(3.52) ***	(2.04) **
D_Gujarat	1.754	2.865	1.737	2.884	1.699	3.151	1.692	3.169
	(4.17) ***	(4.00) ***	(6.09) ***	(4.71) ***	(2.94) ***	(3.12) ***	(3.77) ***	(3.10) ***
D_Kerala	10.323	-2.629	10.33	-2.516	10.636	-1.251	10.67	-1.206
	(6.20) ***	(1.18)	(9.08) ***	(1.48)	(4.52) ***	(0.44)	(5.22) ***	(0.46)

D_Madhya Pradesh	-1.333		0.216		-1.324		0.202		-1.165		0.149		-1.167		0.14	
	(5.56)	***	(0.47)		(3.72)	***	(0.58)		(4.13)	***	(0.23)		(2.03)	**	(0.25)	
D_Rajasthan	-1.226		0.059		-1.226		0.035		-1.117		-0.087		-1.122		-0.102	
	(2.92)	***	(0.07)		(3.40)	***	(0.07)		(1.78)	*	(0.08)		(1.85)	*	(0.12)	
D_Tamil Nadu	3.998		2.313		3.994		2.388		4.172		3.261		4.169		3.297	
	(18.82)	***	(1.45)		(16.48)	***	(1.89)	*	(13.66)	***	(1.54)		(10.88)	***	(1.67)	*
D_Uttar Pradesh	-5.905		-3.391		-5.888		-3.484		-5.876		-4.453		-5.874		-4.504	
	(22.66)	***	(1.42)		(11.65)	***	(2.10)	**	(25.98)	***	(1.35)		(7.45)	***	(1.69)	*
D_Lockdown Phase 2	2.428		2.448		2.449		2.471		-7.748		-7.768		-7.754		-7.775	
(D_April)[8]	(7.27)	***	(7.08)	***	(13.18)	***	(12.91)	***	(15.07)	***	(15.02)	***	(34.98)	***	(34.23)	***
D_Lockdown Phase 3	2.92		2.95		2.937		2.969		-6.62		-6.636		-6.625		-6.641	***
(D_May)[8]	(6.38)	***	(6.21)	***	(14.17)	***	(13.90)	***	(9.23)	***	(9.14)	***	(25.97)	***	(25.51)	
D_Lockdown Phase 4	3.571		3.621		3.589		3.64		-4.769		-4.78		-4.776		-4.787	
	(8.33)	***	(8.15)	***	(16.70)	***	(16.44)	***								
D_Unlock 1.0	5.229		5.272		5.247		5.291									
(D_June)[8]	(14.65)	***	(14.12)	***	(30.62)	***	(29.78)	***	(9.49)	***	(9.29)	***	(18.88)	***	(18.38)	
D_Unlock 2.0	6.282		6.291		6.3		6.31		-3.7		-3.747		-3.708		-3.755	
(D_July)[8]	(17.55)	***	(16.86)	***	(37.68)	***	(36.43)	***	(8.08)	***	(7.96)	***	(14.95)	***	(14.68)	
D_Unlock 3.0	7.493		7.51		7.511		7.528		-2.502		-2.541		-2.51		-2.549	**
(D_August)[8]	(19.66)	***	(19.04)	***	(44.89)	***	(43.50)	***	(6.26)	***	(6.22)	***	(10.15)	***	(10.02)	
D_Unlock 4.0	8.666		8.699		8.689		8.723		-1.241		-1.251		-1.246		-1.256	
(D_September)[8]	(23.09)	***	(22.64)	***	(57.99)	***	(56.50)	***	(4.28)	***	(4.23)	***	(5.65)	***	(5.56)	
D_Unlock 5.0	9.67		9.708		9.703		9.741									
(October)[8]	(22.53)	***	(22.37)	***	(66.05)	***	(64.51)	***								
Constant	3.108		-62.21		2.018		-62.61		1.557		-57.91		1.177		-58.13	
	(0.13)		(4.52)		(0.29)		(8.36)		(0.05)		(2.76)		(0.10)		(4.55)	
State Fixed Effects[1]	Yes		Yes		Yes		Yes		Yes		Yes		Yes		Yes	
No of Observations(N)	1041		1008		1041		1008		223		216		223		216	
(Left censored)					18		18						18		18	
No of states(n)	32		31		32		31		32		31		32		31	
No of weeks (T)	32.5		32.5		32.5		33.5		7		7		1		1	
Wald Chi2	11714	***	11100	***	11900	***	11357	***	4740	***	4769	***	4149	***	3969	***
R squared within	0.8886		0.8875		-		-		0.9189		0.4538		-		-	
R squared betweem	1		1		-		-		1		1		-		-	
R squared overall	0.9219		0.9212		-		-		0.9494		0.6283		-		-	

Breush and Pagan Test	0	0			0	0
(p value)	(1.00).	(1.00).			(1.00).	(1.00).
Hausman Test [7]	0	0			0	0
(p value)	(1.00).	(1.00).			(1.00).	(1.00).

Notes: 1. State dummies or fixed effects for the other states have been included in all the cases. That is, the model has been estimated as by a mixed effects model.

2. *** = Significant at 1% level. ** = Significant at 5% level. * = significant at 105 level.

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

4. D_ stands for a dummy variable (taking 1 or 0).

6. Statistically signicaint cases are highlighted as bold numbers.

7. Hausman tests were carried out between FE and RE models.

8. Monthly dummies have been used instead of phase dummies in the case of monthly panel data.

Table 4. Roles of Maharashtra (relative to the rest of India)

	(17).		(18).		(19).		(20).		(21).		(22).		(23).		(24).	
Data	Weekly		Weekly		Weekly		Weekly		Monthly		Monthly		Monthly		Monthly	
Dependent Variable																
Level/First Difference	FD		FD		Level		Level		Level		Level		Level		Level	
	Cumulative Severity Ratio		Cumulative Severity Ratio		Cases (log)		Cases (log)		Cumulative Severity Ratio		Cumulative Severity Ratio		Cases (log)		Cases (log)	
Model	OLS		OLS		OLS		OLS		OLS		OLS		OLS		OLS	
Explanatory Variables	Est. Coef. (Z value)		Est. Coef. (Z value)		Est. Coef. (Z value)		Est. Coef. (Z value)		Est. Coef. (Z value)		Est. Coef. (Z value)		Est. Coef. (Z value)		Est. Coef. (Z value)	
D_Maharashtra	0.211	[1,2]	-0.034		4.032		4.813		3.282	[2,3]	0.003		4.097		4.937	
	(4.88)	***	(0.42)		(23.52)	***	(7.03)	***	(3.53)	***	(0.42)		(13.94)	***	(7.47)	***
D_Lockdown Phase 2	0.016		0.008		3.736		3.74		0.046		0.035		4.039		4.058	
(D_April)	(1.04)		(0.52)		(6.91)	***	(6.68)	***	(0.32)		(2.09)	**	(4.87)	***	(4.67)	***
D_Lockdown Phase 3	0.047		0.04		4.279		4.266		0.149		0.12		5.512		5.505	
(D_May)	(1.05)		(0.88)		(7.08)	***	(6.82)	***	(1.09)		(2.93)	***	(6.50)	***	(6.20)	***
D_Lockdown Phase 4	-0.019		-0.028		5.206		5.197									
	(0.89)		(1.35)		(9.43)	***	(9.09)	***								
D_Unlock 1.0	0.053		0.032		6.852		6.871		0.271		0.196		7.564		7.582	
(D_June)	(3.04)	***	(2.13)	**	(16.77)	***	(16.23)	***	(2.21)	**	(3.45)	***	(10.34)	***	(9.90)	***
D_Unlock 2.0	0.085		0.072		7.9		7.927		0.599		0.462		8.808		8.838	
(D_July)	(3.31)	***	(2.77)	***	(20.06)	***	(19.42)	***	(4.04)	***	(4.45)	***	(12.35)	***	(11.83)	***
D_Unlock 3.0	0.37		0.364		9.084		9.126		1.663		1.5		9.993		10.037	
(D_August)	(2.28)	**	(2.17)	**	(23.33)	***	(22.62)	***	(4.47)	***	(4.07)	***	(14.20)	***	(13.62)	***
D_Unlock 4.0	0.204		0.193		10.078		10.129		2.968		2.769		10.987		11.04	
D_September	(6.24)	***	(5.78)	***	(26.75)	***	(25.95)	***	(4.75)	***	(4.34)	***	(15.75)	***	(15.13)	***
D_Unlock 5.0	0.112		0.11		10.687		10.742		3.852		3.622		11.596		11.653	
D_October	(3.72)	***	(3.63)	***	(28.20)	***	(27.36)	***	(6.02)	***	(5.59)	***	(16.69)	***	(16.03)	***
D_Lockdown Phase 2*D_Maharashtra			0.249				-0.077				0.249				-0.588	
D_April*D_Maharashtra			(2.16)	**			(0.09)				(14.90)	***			(0.68)	
D_Lockdown Phase 3*D_Maharashtra			0.175				0.502				0.823				0.247	
D_May*D_Maharashtra			(1.85)	*			(0.59)				(20.02)	***			(0.28)	
D_Lockdown Phase 4*D_Maharashtra			0.28				0.349				0				0	
			(2.95)	***			(0.43)				(0.00)				(0.00)	
D_Unlock 1.0*D_Maharashtra			0.628				-0.527				2.295				-0.547	

D_June*D_Maharashtra		**(4.24)** ***		(0.73)		**(40.47)** ***		(0.71)
D_Unlock 2.0*D_Maharashtra		**0.393**		-0.81		**4.254**		-0.934
D_July*D_Maharashtra		**(4.65)** ***		(1.13)		**(40.96)** ***		(1.25)
D_Unlock 3.0*D_Maharashtra		**0.174**		**-1.273**		**5.118**		**-1.397**
D_August*D_Maharashtra		(0.93)		**(1.80)** *		**(13.88)** ***		**(1.90)** *
D_Unlock 4.0*D_Maharashtra		**0.314**		**-1.562**		**6.25**		**-1.686**
D_September*D_Maharashtra		**(3.42)** ***		**(2.22)** **		**(9.79)** ***		**(2.31)** **
D_Unlock 5.0*D_Maharashtra		0.023		**-1.687**		7.242		**-1.812**
D_October*D_Maharashtra		(0.20)		**(2.42)** **		(11.19) ***		**(2.49)** **
Constant	-5.021	2.623	25.289	110.9	-5.021	2.623	25.289	110.903
	-2.21	2.48	1.15	4.8	-2.21	2.48	1.15	4.8
No of Observations(N)	1041	1041	1073	1073	255	255	255	255
No of states(n)	32	32	32	27	32	32	27	27
No of weeks (T)	32.5	32.5	33.5	33.4	7.9	7.9	67.7	33.4
F	15.53 **	306.3 ***	1.92E+02 ***	1960.3 ***	16.5 ***	18 ***	6.77E+01 ***	67.77681 ***
R squared	0.0298	0.0318	0.6676	0.6685	0.382	0.4173	0.7051	0.7059

Notes:

1. *** = Significant at 1% level. ** = Significant at 5% level. * = significant at 105 level.
2. The numbers in brackets show z values. They are based on robust standard errors.
3. D_ stands for a dummy variable (taking 1 or 0).
4. Statistically signicaint cases are highlighted as bold numbers.

Table 4 summarises the results based on Equation (3) to compare the trends of DCSR, CSR and the log of cases between Maharashtra and the rest of India. Maharashtra has a 0.21% higher weekly change in CSR (Column 17), a 3.28% higher monthly CSR (21), a 4.03-4.10% larger number of weekly/monthly infection cases (19 and 23). Broadly consistent with Figure 1, the state experienced highest rise in CSR during Unlock 1.0 to be followed by Unlock 2.0 and Unlock 4.0 (Column 18). However, the interacted effects of Maharashtra and phases or months show a different pattern in Columns 20 and 24. Infection cases in Maharashtra in Lockdown Phase 1 or March are 4.81% or 4.94% higher than the rest of India (the top row). This relative gap has not statistically changed from Lockdown Phase 2 (or April) to Unlock 2.0 (or July) as the interactions between a dummy variable for Maharashtra and phase or month dummies remained statistically non-significant. During this period, Maharashtra continued to experience a proportional increase in infection cases in comparison with other states. From Unlock 3.0 (August) to Unlock 5.0 (December), the proportional gap started to decrease as implied by negative and significant parameter estimates of interaction terms in both columns. This reflects that daily cases peaked out in September (Figure 4), while there were other states where daily infection cases increased (e.g. Kerala, West Bengal, see Figures 1 and 2 and Appendix Figure 3) or did not decrease at a faster rate than Maharashtra (e.g. Andhra Pradesh, see Figures 3 and 4). This reflects the effort by the government of Maharashtra (e.g. lockdown policies), though continued policy attention is necessary for Maharashtra given that it still records the highest level of cases.[25]

6. Conclusion

Here we focus on the significance of our analysis. To the best of our knowledge, this is one of the first rigorous econometric analyses of severity of COVID-19 pandemic, measured based on the index of excess mortality called Cumulative Severity Ratio (CSR) and its first difference (DCSR) up to 31 October, 2020. As emphasized earlier, the CSR measures the additional pressure on our fragile and ill-equipped healthcare system, and the DCSR helps monitor the progression of fatalities. Analysis of COVID-19 cases adds a related but distinct dimension of the surge of the COVID -19 dimension. Another important contribution of this analysis is the use of rigorous econometric methodology: a random effects model and a random effects Tobit model, the latter of which takes into account the fact that some states did not record COVID-19 oriented death in early phases. 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 severer pandemic reflected by a larger CSR or DCSR include higher income at the state level, a higher share among the elderly with multi-morbidity conditions, urbanisation, a lower share of females in the population, lower local retail prices of wheat, and lockdown and unlock phases. On the other hand, the factors associated with a higher number of infection cases - which are different from the above factors - include lower income, a lower share of the elderly with multi-morbidity conditions, a higher share of females in the

[25] There could be of course underlying causal relationships between the state-level lockdown policy of Maharashtra and outcome variables, but our results indicate the correlations, not causality, due to aforementioned reasons.

population, a lower wheat prices as well as hotter and/or rainier wheather conditions. Given the paucity of rigorous econometric analyses, our study yields policy insights of considerable significance.

Yet another important result is the positive association between the severity measures and per capita income, implying that higher incomes are associated with higher mortalities. The underlying mechanisms include greater economic activity, more travel and intermixing, and consequently, higher exposure to the infection and higher risk of dying if denied medical assistance. A negative association between infections and per capita income may plausibly reflect greater awareness of potential benefits of social distancing, personal hygiene and sanitation, and of wearing masks. But further investigation is necessary to resolve this puzzle.

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

The negative association of both CSR and DCSR with the sex ratio, the number of women per 1000 men, means that the state with a higher share of women (i.e. a higher sex ratio) tends to have a lower severity ratio or its increase (or a milder pandemic) after controlling for state fixed effects. This is consistent with Joe et al. (2020) who argued that the evidence from various countries suggests that men are at greater risk of both infections and deaths, and that males are at a greater disadvantage than females with the case fatality rate (CFR) of 3.3% and 2.9%, respectively. In a statistical analysis, the authors show that the CFR among males is usually higher than females for most of the age groups. Male and female CFR also have distinct patterns with greater disadvantage for male survival in under-five as well as in older age groups.

A few limitations are briefly noted. One is that the analysis lacks household data. Instead, state is the unit of analysis. As there is considerable variation in COVID-19 fatalities within states, we are unable to capture this variation. Another limitation is lack of data on health capacity and infrastructure for measuring the response to the COVID pandemic. A third limitation is that we are unable to assess the impact of return migration on the villages/small towns to which they belong.

Despite these limitations, we offer a few policy perspectives. First, our results imply a large degree of heterogeneity in severity of the COVID-19 pandemic and its slowdown. The heterogeneity may be due to under-funding of the health sectors in several states. So, the first priority is to increase substantially the funding for the health sector. Including the private sector, the total health expenditure as a percentage of GDP is estimated at 3.9%. Out of the total expenditure, effectively about one-third (30%) is contributed by the public sector. This contribution is low as compared to 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 (e.g. by engaging with the private sector on a larger scale). In a nuanced and coherent proposal, in the context of vulnerability of the old suffering from NCDs to high risk of dying from COVID-19, a case could be made to develop a fully integrated population- based healthcare system that brings together the public and private sectors and the allopathic and indigenous systems, and is well-coordinated at different levels of service delivery platforms-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.

Another major concern is that the response to COVID-19 infection depends on the immune system of the individuals. Specifically, individuals with poor nutritional status are likely to have weak immune system. A significant proportion of women in the age-group 15-49 years, for example, are undernourished and this makes them more vulnerable to COVID-19 morbidity and mortality. As risks of chronic diseases accumulate over a life span, the old (60 years and above) tend to be more vulnerable to diabetes and cardio-vascular diseases, and thus exposes them to higher risks of COVID-19 morbidity and mortality. Amid elevated risks to lives and livelihoods, there is also a surge in hunger and food deprivation in both rural and urban areas. Besides, disruption of healthcare services is inimical to nutritional health (Joe et al. 2020). So food security is a major policy challenge (Reardon, et al. 2020).

A somewhat baffling result is that higher incomes are associated with higher severity of COVID-19. If we juxtapose this finding with the positive association of urbanisation, 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 different thresholds of income that affect COVID-19 severity, which we have not established. Another is the life-style and associated NCD incidence. As obesity tends to be higher in urban areas, mainly because of sedentary life-styles and rich diets (e.g. eating out, fast food) and consequently incidence of NCDs, and severity of COVID-19, effective solutions must be found to address these concerns. As tax policies may have limited impact (e.g., higher taxes on cigarettes, alcohol), greater emphasis from credible sources may induce behavioural changes. A study carried out by one of the authors, points to the important role of mass media and social networks in influencing behavioural responses (Kulkarni et al. 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, there is every possibility of avoiding a second or third wave. Another important observation is that prolonged lockdowns are not the answer to future waves of COVID-19. Neither School closures are sustainable nor could the economy be refrigerated again. What matters most is a mix of combination prevention that includes handwashing, respiratory hygiene, mask-wearing, physical distancing and avoiding mass gatherings, some of which received greater attention during the Unlock phases.

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

References

Acharya, R., & Porwal, A. (2020). "A vulnerability index for the management of and response to the COVID-19 epidemic in India: an ecological study. *The Lancet Global Health,* 8(9), e1142-e1151.

Banerjee, A., M. Alsan, E. Breza, A.G. Chandrasekhar, A. Chowdhury, E. Duflo & P. Goldsmith-Pinkham (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", NBER Working Paper 27496, Cambridge, MA 2138.

Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects modeling of time-series cross-sectional and panel data. *Political Science Research and Methods,* 3(1), 133-153.

Chatterjee, P. (2020). Is India missing COVID-19 deaths?. *The Lancet*, 396(10252), 657.

Chen, L., Raitzer, D., & Hasan, R. (2020). What Works to Control COVID-19? Econometric Analysis of a Cross-Country Panel, ADB Economics Working Paper Series, No.625 https://www.adb.org/publications/control-covid-19-economic-analysis-cross-country-panel (accessed on 26/12/2020).

Das, A., Ghosh, S., Das, K., Basu, T., Dutta, I., & Das, M. (2020). 'Living environment matters: Unravelling the spatial clustering of COVID-19 hotspots in Kolkata megacity, India', *Sustainable Cities and Society,* Forthcoming.

Gaiha, R., Jha, R., and Kulkarni (2014). *Diet, Malnutrition and Disease, The Indian Experience*, New Delhi, Oxford University Press.

Galasso, V., Pons, V., Profeta, P., Becher, M., Brouard, S., & Foucault, M. (2020). Gender differences in COVID-19 attitudes and behavior: Panel evidence from eight countries. *Proceedings of the National Academy of Sciences,* 117(44), 27285-27291.

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 Differences*, 11(1), pp.1-13.

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

Gupta, A. (2020) "Indian health data will miss recording many Covid and lockdown deaths", The Print, 18 April, https://theprint.in/opinion/indian-health-data-will-miss-recording-many-covid-and-lockdown-deaths/403432/ (accessed on 30/12/2020).

Horton, R. (2020)"Offline: The Second wave", *Lancet,* 27th June.

Im, Kyung So, M. Hashem Pesaran, and Yongcheol Shin. 2003. "Testing for unit roots in heterogeneous panels." *Journal of Econometrics* 115: 53–74.

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 Health*, 8, p.152.

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*, 2(1).e17.

Khalatbari-Soltani, S., Cumming, R. G., Delpierre, C., & Kelly-Irving, M. (2020). Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *Journal of Epidemiol Community Health, 74*(8) 620-623.

Kulkarni, V. S. & R. Gaiha (2020) :"Tragic Choices in Health Care during Covid", *Business Standard*, May, 19; https://www.business-standard.com/article/opinion/tragic-choices-in-health-care-during-covid-120051901758_1.html.

Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu. 2002. "Unit root tests in panel data: Asymptotic and finite-sample properties." *Journal of Econometrics* 108: 1–24.

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 Environment*, p.138226.

Olsen, W., Bera, M., Dubey, A., Kim, J., Wiśniowski, A., & Yadav, P. (2020). Hierarchical Modelling of COVID-19 Death Risk in India in the Early Phase of the Pandemic. *The European Journal of Development Research*, 1-28, Published Online. https://doi.org/10.1057/s41287-020-00333-5.

Pesaran, M. H. (2011) 'On the Interpretation of Panel Unit Root Tests', http://www.econ.cam.ac.uk/people-files/emeritus/mhp1/wp11/Interpretation-Panel-Unit-September-2011.pdf

Rao, N. (2018) "Healthcare Financing: Who is paying?", PRS Legislative Research.

Reardon, T., A. Mishra, C.S.R. Nuthalappati, M.F.Bellemare, & D. Zilberman (2020) "COVID-19's Disruption of India's Transformed Food Supply Chains", *Economic and Political Weekly*, 2 May.

Schellenkens, P., & D. Sourrowuille (2020). 'COVID-19 Mortality in Rich and Poor Countries: A Tale of Two Pandemic"s:, Policy Research Working Paper 9260, World Bank, Washington DC.

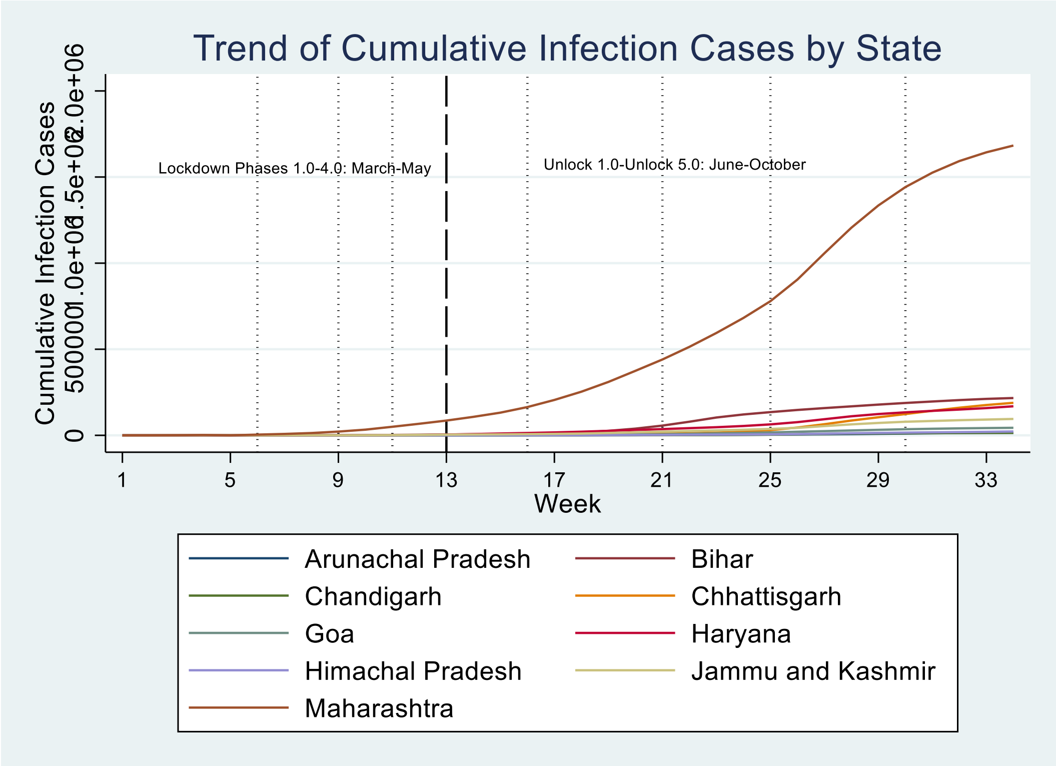
Susskind, D., & D. Vines (2020). "The economics of the COVID-19 pandemic: an assessment". *Oxford Review of Economic Policy*, 36 (S1), S1–S13.

Wu, Y. C., Chen, C. S., and Chan, Y. J. (2020). "The outbreak of COVID-19: An overview". *Journal of the Chinese Medical Association,* 83(3), 217.

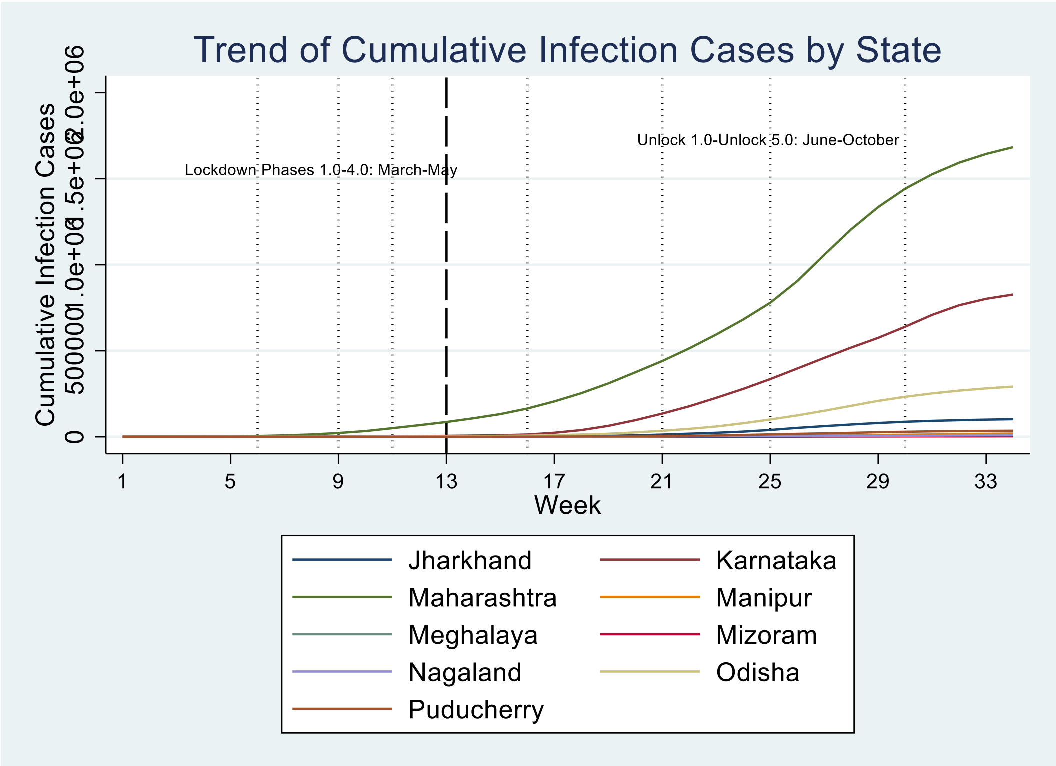
Appendix Table Descriptive Statistics

	Weekly Panel					Monthly Panel				
Variable	Obs.	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max
Cumulative Severity Ratio (CSR) of COVID-19 (%)	1,073	101.4	2.8	100.0	119.8	255	101.2	2.4	100.0	117.0
The first difference of CSR	1,041	0.1	0.7	-1.9	19.1	223	0.6	1.3	-6.0	9.4
log Cumulative Covid infection cases	1,073	7.6	4.3	-6.9	14.3	255	7.0	4.5	-6.9	14.3
log per capita income (Rs.)	1,073	11.5	0.5	10.3	12.9	255	11.5	0.5	10.3	12.9
Rate of Mult-morbidity	1,073	6.6	5.9	1.6	37.6	255	6.7	6.2	1.6	37.6
Rate of Urbanisation (%)	1,039	34.0	18.3	0.8	97.3	247	34.1	18.3	0.8	97.3
Sex Ratio (no. of females per 1000 males)	1,039	947.4	50.5	818.0	1084.0	247	947.8	51.3	818.0	1084.0
log of retail price of wheat	1,073	3.3	0.2	3.0	4.1	255	3.4	0.2	3.0	4.1
temperature	1,073	299.3	5.0	275.2	311.1	255	299.4	5.1	275.9	308.9
rainfall	1,073	7.78	9.53	0	61.15	255	7.4	7.7	0.0	51.6
D_Lockdown Phase 2	1,073	0.08	0.24	0	1					
D_Lockdown Phase 3	1,073	0.06	0.20	0	1					
D_Lockdown Phase 4	1,073	0.06	0.20	0	1					
D_Unlock 1.0	1,073	0.11	0.30	0	1					
D_Unlock 2.0	1,073	0.12	0.32	0	1					
D_Unlock 3.0	1,073	0.12	0.32	0	1					
D_Unlock 4.0	1,073	0.18	0.38	0	1					
D_Unlock 5.0	1,073	0.15	0.36	0	1					
D_April						255	0.13	0.33	0	1
D_May						255	0.13	0.33	0	1
D_June						255	0.13	0.33	0	1
D_July						255	0.13	0.33	0	1
D_August						255	0.13	0.33	0	1
D_September						255	0.13	0.33	0	1
D_October						255	0.13	0.33	0	1

Appendix Figure 1: Trend of Cumulative COVID-19 infection I (13-03-2020-31-10-2020)



Appendix Figure 2: Trend of Cumulative COVID-19 infection II (13-03-2020-31-10-2020)



Appendix Figure 3: Trend of Cumulative COVID-19 infection III (13-03-2020-31-10-2020)