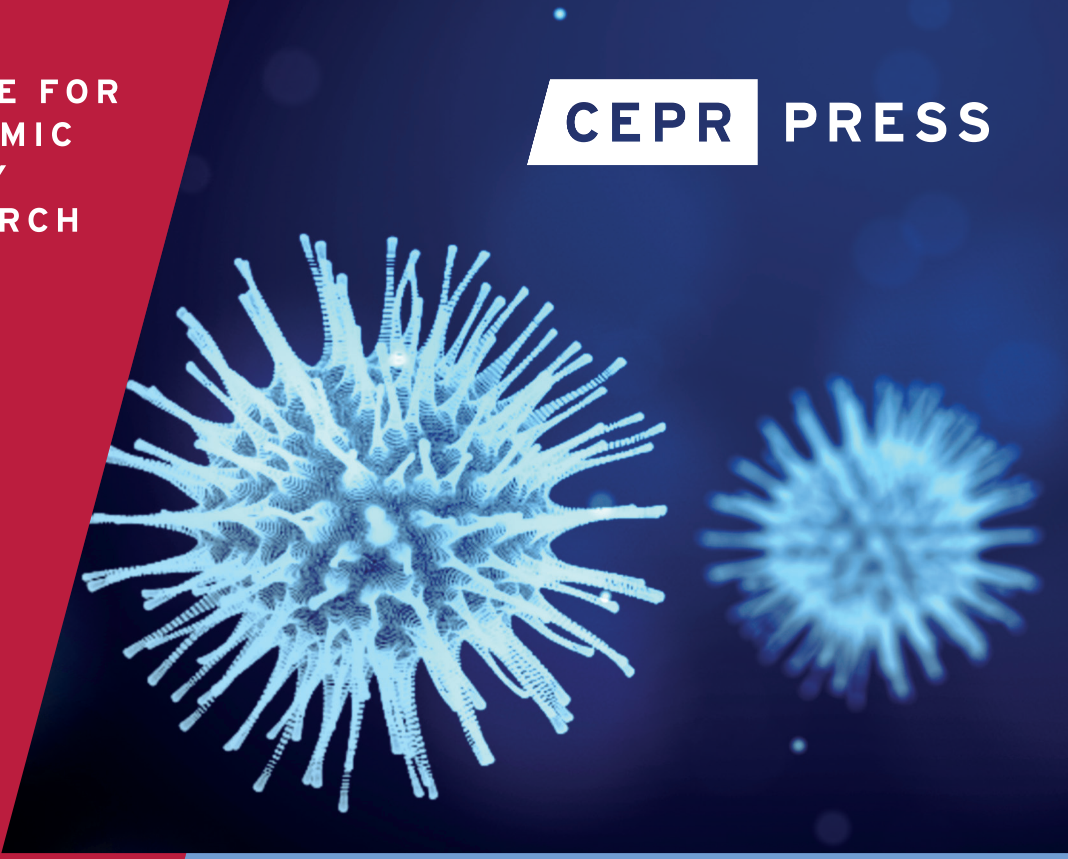
CENTRE FOR
ECONOMIC
POLICY
RESEARCH

CEPR PRESS



COVID ECONOMICS

VETTED AND REAL-TIME PAPERS

ISSUE 37
14 JULY
2020

Covid Economics
Vetted and Real-Time Papers

Covid Economics, Vetted and Real-Time Papers, from CEPR, brings together formal investigations on the economic issues emanating from the Covid outbreak, based on explicit theory and/or empirical evidence, to improve the knowledge base.

Founder: Beatrice Weder di Mauro, President of CEPR
Editor: Charles Wyplosz, Graduate Institute Geneva and CEPR

Contact: Submissions should be made at https://portal.cepr.org/call-papers-covid-economics. Other queries should be sent to covidecon@cepr.org.

The Centre for Economic Policy Research (CEPR)

The Centre for Economic Policy Research (CEPR) is a network of over 1,500 research economists based mostly in European universities. The Centre's goal is twofold: to promote world-class research, and to get the policy-relevant results into the hands of key decision-makers. CEPR's guiding principle is 'Research excellence with policy relevance'. A registered charity since it was founded in 1983, CEPR is independent of all public and private interest groups. It takes no institutional stand on economic policy matters and its core funding comes from its Institutional Members and sales of publications. Because it draws on such a large network of researchers, its output reflects a broad spectrum of individual viewpoints as well as perspectives drawn from civil society. CEPR research may include views on policy, but the Trustees of the Centre do not give prior review to its publications. The opinions expressed in this report are those of the authors and not those of CEPR.

Editorial Board

Ethics

Covid Economics will feature high quality analyses of economic aspects of the health crisis. However, the pandemic also raises a number of complex ethical issues. Economists tend to think about trade-offs, in this case lives vs. costs, patient selection at a time of scarcity, and more. In the spirit of academic freedom, neither the Editors of *Covid Economics* nor CEPR take a stand on these issues and therefore do not bear any responsibility for views expressed in the articles.

Submission to professional journals

The following journals have indicated that they will accept submissions of papers featured in *Covid Economics* because they are working papers. Most expect revised versions. This list will be updated regularly.

American Economic Review

American Economic Review, Applied Economics

American Economic Review, Insights

American Economic Review, Economic Policy

American Economic Review, Macroeconomics

American Economic Review, Microeconomics

American Journal of Health Economics

Canadian Journal of Economics

*Econometrica**

Economic Journal

Economics of Disasters and Climate Change

International Economic Review

Journal of Development Economics

*Journal of Econometrics**

Journal of Economic Growth

Journal of Economic Theory

*Journal of the European Economic Association**

Journal of Finance

Journal of Financial Economics

Journal of International Economics

*Journal of Labor Economics**

Journal of Monetary Economics

Journal of Public Economics

Journal of Political Economy

Journal of Population Economics

*Quarterly Journal of Economics**

Review of Economics and Statistics

*Review of Economic Studies**

Review of Financial Studies

(*) Must be a significantly revised and extended version of the paper featured in *Covid Economics*.

Covid Economics
Vetted and Real-Time Papers

Issue 37, 14 July 2020

Contents

Cities and COVID-19 infections: Population density, transmission speeds and sheltering responses[1]

Michiel Gerritse[2]

Date submitted: 8 July 2020; Date accepted: 9 July 2020

The transmission and incidence of COVID-19 infections differ markedly across areas in the U.S. Using daily infection rates at the county level, this paper explores how population density and the organization of the city correlate to the speed of transmission and shelter-in-place responses (staying home, avoiding travel). Population density is associated with higher transmission speeds in particular at the start of outbreaks. Density is also associated with stronger sheltering responses, but mostly in later phases of the outbreak. There is a considerable additional role of the urban form (i.e., public transport, work-from-home and local incomes), in transmission and sheltering. Over the course of the pandemic, workplace connections are increasingly less likely to predict infection, and phone movement shows that people avoid heavily infected areas. Altogether, this suggests that densely populated places are initially prone to faster viral spread, and later develop stronger sheltering responses. The considerable spatial differences in both the speed of transmission and the mobility responses to local infection could explain differences in the pandemic's toll across cities and counties.

Covid Economics 37, 14 July 2020: 1-26

1 This paper has benefited from comments by Frank van Oort and Daniel Arribas-Bel.
2 Erasmus University Rotterdam, Tinbergen Institute.

Covid Economics 37, 14 July 2020: 1-26

1 Introduction

Four months since the first identified COVID-19 infection, all major cities in the U.S. are infected. However, the rates of infection differ considerably across cities and locations. New York City developed over 2,000 cases per 100,000 inhabitants, while Los Angeles has just over 200: the two largest cities in the U.S. are about a factor 10 apart, in terms of infection rates.[1]

This paper examines whether cities of higher population density experience higher rates of COVID-19 transmissions. It estimates the local rate of virus transmission, and then questions whether density and correlated factors are associated to faster transmission. On average, just over one new infection develops out of an earlier COVID-19 case due to exposure to infectious population. That average, however, masks considerable heterogeneity across areas and over time.

Figure 1 shows a scatter of the log of population density in counties against the estimated rate of transmissions in the county in the month after it developed ten COVID-19 infections. Here, the estimated rate of transmission is, briefly, the number of new cases per capita associated with earlier exposure to infection rates of the county population, conditional on state-day averages.[2] The figure shows that in the first months since developing 10 cases, denser areas generally see quicker spread of the virus.

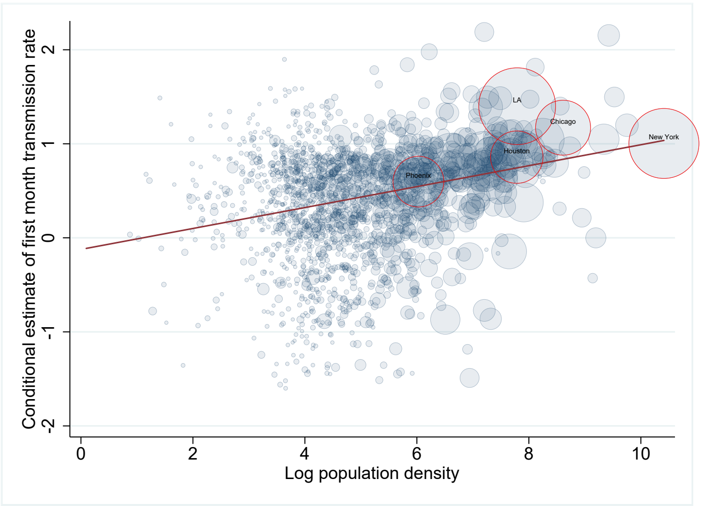


Figure 1: **Estimated virus transmission rates in the month after first case by county.** The transmission rate is estimated as the rate at which 5-day lagged exposures to the virus generate new cases (see text below), conditional on county and state-day fixed effects. Markers are proportional to population size. The counties hosting the five largest cities are highlighted, New York is a combination of five boroughs (New York, Kings, Queens, Bronx and Richmond).

For the main analysis, this paper introduces density and correlated urban variables in an epidemiological regression equation to understand where and when the highest rates of transmissions occured. The estimate for the rate of transmission follows a standard epidemiological approach – not explaining the number of cases, but whether previous infections are associated to subsequent new cases. Using

[1]The mortality rates in excess of the long-run averages show even more marked differences between the state of New York and the state of California.

[2]The infection rate that the population of a county is exposed to, is the weighted average of infection rates in all counties, weighted by the county's commuting flows to those destination counties. For some counties, estimated transmission rates are negative as their infections rises while nearby infection rates fall (or vice versa) conditional on state-day fixed effects.

daily case numbers for roughly 2,000 to 3,000 counties in the U.S., the main model regresses new case development on the 5-day exposure (the number of people multiplied with the average rate of infection that they encounter). The analysis considers different spatial patterns in exposure to infection (in counties of residence, in commuting networks, or in phone-tracked mobility patterns), employing the commuting network for the main analysis. In addition, the paper considers how people's mobility choices changed in the face of exposure to COVID-19 infections.

The results show that density is associated with higher transmission rates of the virus. An additional log point of population density (i.e., twice the density, or roughly the difference between Philadelphia on the one hand and the five central boroughs of New York on the other hand) implies about 10% higher expected transmission rates. This estimate is conditional on state-day fixed effects and it is robust across different definitions of density. However, there is also notable variation over time. The role of density in transmission peaks early on: density is more strongly linked to high transmission rates in March than it is in April or May. Conditional on population density, counties with more public transport use, higher wages and many work-from home jobs see higher transmission rates in the first weeks of March. In April, the role of wages and work-from-home jobs reversed, leading to lower (conditional) transmission on occasion. Public transport is not generally associated with lower transmissions conditional on density.

The results also show that mobility responses to shelter from exposure in areas of infection are stronger in densely populated areas: people are more likley to reduce travel (to work and to other destinations) and to stay home more, when their potential exposure to infectious people rises. Counties with higher population densities and higher wages show stronger sheltering responses. In areas with high numbers of work-from-home jobs, sheltering is weaker than elsewhere in the first month of infections, but stronger after the first month. Consistent with growing avoidance of exposure in mobility patterns, the transmission estimates from workplaces decline over time relative to those from the county of residence: the commuting network becomes less relevant to predict transmission. Patterns from phone tracking data are also consistent with mobility changes that evade exposure: when faced with higher exposure, people travel less in general, and in particular avoid counties with high infection rates as destinations.

This paper provides descriptives that inform several current discussions. The role of population density in the development of COVID-19 cases is hotly debated. Evidence on the role of population density in infectious disease is generally mixed (Li et al., 2018). For COVID-19 in particular, some results show that density and city size aggravates the spread of COVID-19 (Stier et al., 2020; Ribeiro et al., 2020), while others argue that population density plays little role in the development of the disease (Heroy, 2020; Fang and Wahba, 2020). The urban organization correlated to population density may play a considerable role, too. In dense areas, commuters make more extensive use of public transport. The physical proximity and grouping of people in public transport may be a source of contagion (Harris, 2020; Tian et al., 2020). Others, however, argue that car users are more likely to spread the virus as they tend to combine trips to multiple destinations and they are less likely to diminish their mobility than public transport users do (Furth, 2020). Urban population may also live in smaller or more crowded spaces, and in buildings that have system ventilation or plumbing, which are argued to foster transmission (Gormley et al., 2020). Jobs that concentrate in dense areas, such as those in the service industry, may require more interaction, facilitating transmission (Florida, 2020; Almagro and Orane-Hutchinson, 2020). Poorer workers in large cities, or those that have poorer internet infrastructure access, may have few options to shelter in place (Coven and Gupta, 2020; Chiou and Tucker, 2020). This paper provides a first U.S.-wide analysis of counties and how suspected factors correlated to density play a role. Moreover, the current analyses delves into the dynamics, arguing that density may play a different role in different stages of the epidemic.

Covid Economics 37, 14 July 2020: 1-26

The paper also adds sheds light on voluntary distancing and shelter-in-place responses. Several studies show that mobility declines when infection rates rise, but that the decline varies substantially. Income, ethnicity, political preference, job types and means of transport determine the mobility response (e.g., Coven and Gupta, 2020; Engle et al., 2020; Adam Brzezinski, 2020; Crowley et al., 2020). This paper analyzes how people from different areas vary in their shelter-in-place behaviour. Higher population densities, work-from-home job shares and incomes correlate to more sheltering behaviour, but only in later phases of the epidemic. Moreover, while related literature points out that people are less likely to move and more likley to stay home facing exposure, this paper shows that the destinations matters: people cut their movement to heavily infected areas in particular. Hence, the data point to a strong geographical dimension in sheltering responses.

These discussions are also relevant to the policy questions raised by the pandemic. In the short run, lockdown restrictions have presented heavy tolls on citizens. Understanding where transmissions occur fastest could help minimize policy burdens while effectively maximizing virus containment. Potential geographical dimensions to lockdown policies can be informed by an understanding of the role of density and urban context in virus transmission rates. In the longer run, understanding the geography of transmissions may help to predict how cities will develop. Several epidemics, including plagues, cholera, and tuberculosis have left notable imprints on the organisation of cities, giving rise to public parks, spacious street plans and clinical building styles, for instance. The density and connectedness of cities are often argued to have become perils,[3] and evidence on how viral transmissions fare in large cities may show what is in stock for those cities.

This paper details the methodology and data sources before turning to three sets of results: i) on density; ii) on urban factors correlated to density; and iii) on mobility responses.

2 Empirical strategy and data

The transmission rate is estimated from the infection equation of the workhorse Kermack and MacKendrick SIR model. Susceptible persons (S) can be infected by interactions with infectious persons (I). The rate of transmission is β. The number of new cases identified on day t is $i_t = \beta S_{t-\tau} * I_{t-\tau}$, where τ is a time lag between infection and confirmation - for incubation and testing, for instance. The infectious rate is not equal to effective reproduction number ("R"), but it is monotonically related (Wallinga and Lipsitch, 2007).

People can get infected in different locations. I allow an interaction of residents from origin o in destination d, using mobility measures between counties s_{od}. Summing up a county o's interaction across counties d (e.g. Song et al., 2017), the effective exposure is a weighted summation of infection rates:

$$i_{o,t} = \beta S_{o,t-\tau} \sum_{d} s_{o,d,t-\tau} \frac{I_{dt-\tau}}{N_d}, \tag{1}$$

with $s_{odt} = S_{odt-\tau}/S_{ot-\tau}$ is the share of a destination in total interactions of people from origin o in county d. Using $Exposure_{ot} \equiv S_{o,t-\tau} \sum_d s_{o,d,t-\tau} \frac{I_{dt-\tau}}{N_d}$ as the interaction-weighted potential for infection across destinations, the estimating equation becomes $i_{o,t} = \beta Exposure_{ot}$. The transmission rate coefficient β represents the average number of new infections that evolves from the potential for interaction. The coefficient is an overall "force of infection" (Wakefield et al., 2019) , comprising a contact rate at which susceptible and infectious people interact in ways that can transmit the virus, and the biological probability of infection (determined by, e.g., the distance the virus can travel in air and its survival time on surfaces).

[3]https://blogs.lse.ac.uk/covid19/2020/05/13/covid-19-has-turned-cities-main-economic-assets-into-their-worst-enemies/

Covid Economics 37, 14 July 2020: 1-26

Effective contact rates can vary with the characteristics of the location: living in a dense neighborhood may increase the probability of a physically close encounter, or having to travel by subway may imply standing close to multiple people in a confined space. Using x for locational characteristics, such as population density or transport mode use, I estimate the equation as follows: $\frac{dI_{ot}}{dt} = \beta(x)\,Exposure_{ot}$, where conditioning β on x allows the transmission rate to vary according to urban characteristics x.

Estimating equation (1) to describe infectious rates requires a few additional steps. First, I estimate the equation in first differences, to eliminates unobservables at the county-level. Factors such as international travel or overall susceptiblity could cause elevated infection rates over the whole period. The level of infection rates is not what this equation seeks to identify: it identifies whether infections are likely to cause other infections, whether they be in an environment of low or high infection rates.[4] Second, there are considerable spatial patterns in the infection rates, and state-level policy responses affect neighboring counties similarly. To account for time-varying shocks, such as news or federal policies; and in particular for time-state-varying shocks, such as state-level medical or containment responses, a state-day fixed effect α_{st} controls for average daily state-level growth rates in the number of cases. Moreover, as it is plausible that reporting and testing varies by day of the week (e.g., slightly fewer cases are reported on Mondays in the data), daily fixed effects account for such measurement differences.[5] The estimating equation is, then:

$$\Delta i_{ot} = \beta(x)\,\Delta Exposure_{ot} + \alpha_{st} + u_{ot}. \tag{2}$$

The standard errors for the coefficient β are clustered two-way at the county level and at the day level. The regression is weighted by 2018 census population weights per county. An augmented Pesaran test show no signs of residual serial correlation.

2.1 Infectious time and detection delays

Estimating the infectious rate equation (2) requires specifying for how many days people are infectious. In the data, I proxy the number of infectious people on a given day by the number of people that test positive in the following days. Use virological assessments (Siordia Jr, 2020; Pung et al., 2020), I will examine effective infectious period of 3 to 9 days. A patient infected on day t develops symptoms after an estimated 5-day period of incubation. Infectiousness likely starts earlier, as of day 3 after infection. The patient may remain biologically infectious for close to three weeks. However, the most common transmissions are likely to occur within a shorter time frame. Viral loads decline after 7 days for the majority of patients, and severe cases are more likely to be identified in the first or second week, after which the patient liklley reduces infectious interaction. In addition, half of the transmissions are estimated to occur in the first few days, as the patient is infectious but has not developed symptoms yet (He et al., 2020). Hence, the effective average number of days to the end of infectiousness since infection, T, may plausibly be less than a week. The number of infectious people on day t is proxied as the number of people testing positive in the ensuing days, consistent with being infectious at t: $I(t, T) = \sum_{\theta=t-T,\ldots,t} i_\theta$.

To define how reported infections arise out of exposure, a second parameter is needed: the average number of days between infection and testing positive, τ, which represents a transmission detection delay. The virus is detectable about 3 days from infection, and test results may not be instant, putting a plausible lower limit of 3 days on τ. The upper limit is less certain. Patients either see viral loads decline after a week, or tend to develop severe pneumonia, which makes detection after 8 days less likely.

To accomodate plausible values for the two relevant parameters (days of infectiousness T, detection

[4] A county-level fixed effect could also control for county-level time-invariant unobserved variables, but given that the county's earlier infections are measured in the exposure variable, such county-level fixed effects would cause a Nickell bias.

[5] Note that as both the dependent and the independent variables are expressed in numbers confirmed cases, underreporting of cases does not necessarily lead to a bias in the estimate of transmission.

Covid Economics 37, 14 July 2020: 1-26

delay τ) I estimate the baseline model for a parameter space varying either parameter from 3 to 9 days. One objection is worth pointing out: if the infectiousness period is assumed larger than the detection delay, then the measure of exposure for a patient identified on day t contains the infections at day t itself. While this is biologically possible (when a patient infects another patient, the infector could be identified after the infectee) it is undesirable in the estimation of equation 2, as the right-hand side variable for exposure would contain the dependent variable.

2.2 Data

The main source of data is the New York Times Github repository of daily case counts per county. The counts include laboratory-confirmed and probable cause (documented clinical criteria without other diagnosis) following protocal of the Council of State and Territorial Epidemiologists. The New York Times data are compiled of state and county governments and local health departments. Cases are assigned to the location of identification and treatment where possible. The epidemiological data are supplemented with commuting data from the most recent 5-year American Community Surveys (ACS). Selected covariates (public transport use shares among commuters, wages, unemployment health insurance coverags and occupation distributions) are drawn from the ACS in 2018 accessed through IPUMS (Ruggles et al., 2020). Phone mobility data between counties are from PlaceIQ as distributed by Couture et al. (2020). Due to minimum observation requirements for the use of data, around 2,000 of the more than 3,000 counties in the U.S. are covered in the PlaceIQ data. The coverage is visualized on the map in Appendix A. These data record by county of residence what shares of the phone in the sample "ping" (register to the network) in different counties. Changes to trip frequency by purpose are from the Google mobility reports (Google, 2020). Job characteristics on physical proximity[6] are obtained by connecting O*net dictionary of occupation data on job characteristics to the ACS job distributions by county, and the share of work-from-home jobs is calculated as the (weighted) share of workers that have a teleworkable job in the Dingel and Neiman (2020) classification. Not all counties are covered in the ACS, leading to smaller samples whenever variables from the ACS are used.

All code is available from the website of the author.

3 Results

The baseline transmission equation (2) is first estimated for different parametrizations of transmission detection days and infectiousness days, and for different definitions of mobility. Figure 7 summarizes the results visually: it shows the coefficient estimates and Root Mean Squared Errors of the models for regression fit, exploring four definitions of interaction (via the commuting network, within counties only, via the phone-tracked mobility across counties - normalized and absolute).

Restricting attention to below-diagonal estimates (infectiousness time shorter than detection times), the transmission coefficient estimates for commuting based, residence based and normalized ping based mobility are very similar, at around 1.03 for all measures. The coefficient for non-normalized pings is lower, but as the interaction terms s_{od} sum to more than one, the scale of the variable cannot be compared. The fit across the sample with a common time frame (phone data are reported with a longer lag), judged by root mean squared errors, is best for the commuting network estimates of exposure. The baseline results are reported for the parametrization of 6 days of incubation and 5 days of infectiousness, but the results that follow are comparable for other below-diagonal parametrizations. At this parametrization, the coefficient for transmission in the long sample is 1.06 with 95% confidence interval of 1.01 to 1.12, based on day and county clustered standard errors.

[6] Appendix A provides a brief description of the merging procedure.

3.1 Population density

How does population density change the expectation of transmission rates? Table 1 reports regressions that allow transmission speeds to vary with different definitions of density.

Column 1 of Table 1 interacts the exposure with the most common measure of density: the log of the number of people per squared kilometer. The interpretation of the interaction coefficient is that if the density is twice as high (one log point increase), then the estimated infectiousness rate is higher by the magnitude of the coefficient. One log point of density is the difference between Newark and San Francisco, or between Atlanta and Chicago, and the standard deviation of log density across U.S. counties is just under two. That implies that a density twice as high yields about 0.06 points higher transmission rate. By this measure, San Francisco is expected to have about 35% quicker transmission rates than a county of median density.

Columns 2 and 3 of Table 1 employ different definitions of density. Column 2 uses the number of people over the built-up surface area.[7] This increases the measure of population density in particular in counties with large shares of natural land. Using this definition, the estimated transmission rate is weaker and the estimated standard errors are larger. Column 3 uses the density of homes, rather than people. The estimated transmission rates is similar when using people density or home density.

Column 4 splits up the estimated transmission rates by quintiles of density. The top most dense quintile displays the highest estimated transmission rates, and the least dense display the lowest estimated transmission rates.[8]

The process of transmissions may have looked different in the onset of the crisis than it did in the later weeks. Columns 6 to 10 of Table 1 repeat the regressions on the first 30 daily observations ("1st month") after a county has developed 10 cases. The first month of infections shows a larger role for density: the impact of log density on transmission speed is roughly twice as high and significant for all definitions of density. More clearly than before, in column 8, the most dense quintile shows significantly higher transmission rates. The row "Equal coefficient (p)" reports a (joint) F-test for equality of the coefficient for interaction between exposure and the density measure in the full sample and in the sample of first month after infection. All reject at the 5% level, implying that the coefficients are significantly higher in the first-month sample than outside that sample.

[7] The built up land area share is calculated from the National Land Cover Database. It is calculated in QGIS as the share of pixels in each county polygon classified as "Developed Open Space", "Developed Low Density", "Developed Medium Density" 'or "Developed High Density" out of total count of non-open water and non-perennial snow/ice pixels in the county polygon.

[8] This pattern persists when looking at deciles rather than quintiles as reported in Table 5, pointing to significantly higher transmissions in the most extremely populated areas.

Covid Economics 37, 14 July 2020: 1-26

Covid Economics 37, 14 July 2020: 1-26

Table 1: Transmission equation conditioned on different definitions of population density

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		full sample				1st month after first 10 cases		
exposure	0.65***	0.88***	0.70***		0.40	0.80***	0.49*	
	(0.18)	(0.12)	(0.19)		(0.27)	(0.19)	(0.27)	
log density	1.42**				1.09			
	(0.60)				(1.10)			
log density X exposure	0.06**				0.11***			
	(0.02)				(0.04)			
log built density		0.80**				0.68		
		(0.37)				(0.72)		
log built density X exposure		0.03				0.06**		
		(0.02)				(0.02)		
log house density			1.51**				1.17	
			(0.62)				(1.15)	
log house density X exposure			0.05*				0.10***	
			(0.02)				(0.03)	
Q1/5 exposure				0.60***				0.85***
				(0.14)				(0.13)
Q2/5 exposure				0.87***				1.07***
				(0.07)				(0.06)
Q3/5 exposure				0.88***				0.89***
				(0.06)				(0.05)
Q4/5 exposure				0.90***				0.89***
				(0.04)				(0.06)
Q5/5 exposure				1.06***				1.14***
				(0.02)				(0.05)
Equal coefficient (p)					0.04	0.03	0.04	0.04
Observations	407,685	403,681	407,264	407,630	73,014	72,672	72,984	73,014
state-day FE	yes	yes	yes	yes	yes	yes	yes	yes

Exposure based on 5 days infectious time and 6 days detection
*** p<0.01, ** p<0.05, * p<0.1
Weighted by county population. Twoway (county and day) clustered standard errors

As the stage of development of the epidemic may matter for transmissions, I estimate the transmission rate by week w. Using a weekly coefficient β_w, I estimate

$$\Delta i_{ot} = \sum_w \beta_w x \Delta Exposure_{ot} + \alpha_{st} + u_{ot}. \tag{3}$$

and subsequently calculate the weekly transmission estimates by tertile of density (low density, medium density and high density).

Figure 2 plots the weekly development of transmission estimates for counties of different densities. The coefficients reflect the estimate β_w of the regression equation. The whiskers indicate 95% confidence intervals based on twoway county and day-clustered standard errors. Most strikingly, the largest transmission estimates are reported for high-density counties in the week following March 23 onward. By the week following April 6th, the transmission confidence intervals are around 1 for all density groups, if slightly lower in the least populated counties. Less densely populated counties generally develop transmissions later; and conditional on state-year fixed effects, no transmission can be identified before March 2 in counties of medium or low population density.

The elevated estimates of transmission coincide with a sharp rise in daily cases per capita in the

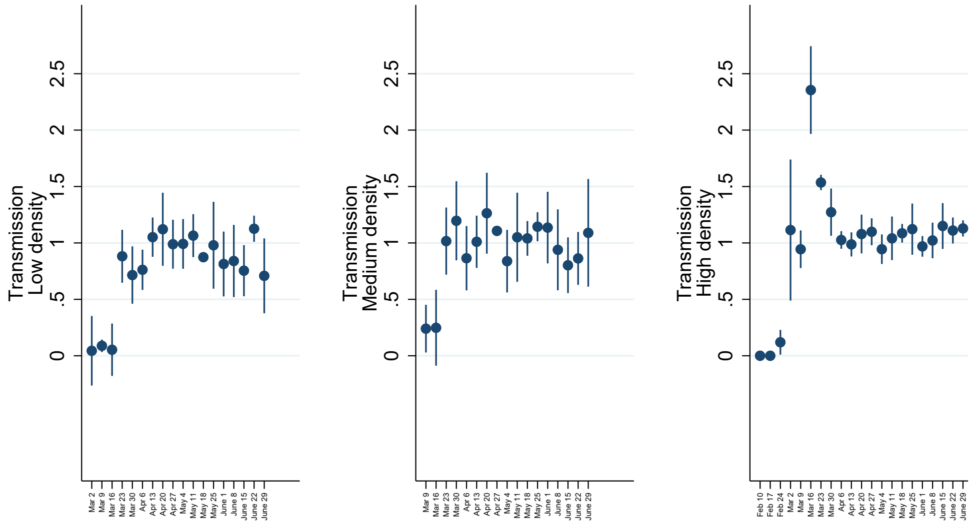


Figure 2: Weekly transmission estimates by tertile of population density

most densly populated counties early on on the outbreak. Appendix C shows a plot of cumulative cases per capita for the three groups, showing an earlier rise in cases in densely populated areas. In the regression, the number of counties with non-zero cases contributing to identification rises sharply in the week of March 16, from 421 to 1,116. Moreover, the number of new cases rise quickly in the vicinity of New York City.

Does population density play a statistically significant role in transmission estimates? To understand this, I allow the interaction between the log of density and exposure to vary over time. Controlling for week-varying impacts of density as well as week-specific exposure impacts unrelated to density, I estimate:

$$\Delta i_{ot} = \sum_{w} \beta_{1w} \Delta Exposure_{ot} * \log density_o + \sum_{w} \beta_{2w} \Delta Exposure_{ot} + \sum_{w} \beta_{3w} \log density_o + \alpha_{st} + u_{ot}. \tag{4}$$

Here, the coefficients β_{1w} estimate, by week, how a unit increase in density affects the rate at which exposure leads to new cases. This is different from overall development of new cases associated with density that are not predicted by previous exposure: the impact of density unrelated to earlier infections is captured in β_{3w}. Denser areas may develop more cases not predicted by local exposure, for instance due to infected patients arriving at the local airport.

Figure 3 plots the weekly development of estimates of the coefficients β_{1w}. The coefficient for the interaction of exposure and density is insignificant up to the week of March 16, and then turns significant with a magnitude of 0.7. This suggests that the transmission rate 0.7 points higher if an area is twice as dense - or that San Francisco would be predicted to have close to 5 times higher transmission speeds as a county of median population density. The timing coincides, again, with a sharp increase in cases in denser places (see Figure 8). In the ensuing weeks up to May 18, the interaction coefficient for log density and exposure indicates that a log point density is associated with about 10% higher infectiousness estimates.

CEPR PRESS

Covid Economics 37, 14 July 2020: 1-26

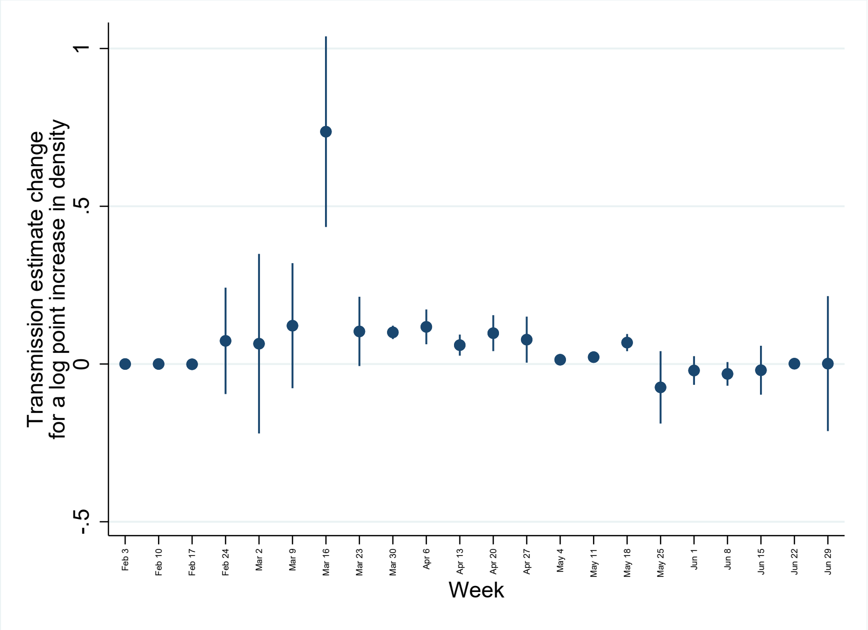


Figure 3: The role of population density in estimated transmission rates by weekWeekly transmission estimates by tertile of population density

3.2 Covariates of density effects

Population density correlates with several plausible explanations for elevated rates of infections. In dense cities, people may share smaller spaces, they make different transport choices or they have different types of jobs, for instance. Several arguments have been proposed for mechanisms through which the spread of the COVID-19 virus can be faster in large or dense cities. Citizens of larger cities rely more on public transport, which may make them vulnerable to transmission: both due to the geographic mobility of infected people and the time spent in close vicinity (Harris, 2020; Tian et al., 2020). Larger cities tend to have workers with different incomes, jobs types, and working facitilies (a.o. Chiou and Tucker, 2020; Florida, 2020; Almagro and Orane-Hutchinson, 2020). The characteristics of service jobs disproportionately found in cities, such as the need for physical proximity (e.g. hairdressers), and the ability to work from home may determine commuting and on-the-job infection rates. On the other hand, cities host relatively many jobs that allow working from home. Higher incomes correlate to job conditions and infrastructure as well as (spacier) living conditions. Health insurance coverage may affect choices for interaction, although the direction is not clear (the cost of infection lead uninsured individuals to avoid interaction, but it might also lead them to avoid hospitalization). Unemployment, on the other hand, could reduce mobility. Finally, age structure could play a role in virology and mobility: children appear to be weaker spreaders of the virus, but they are arguably more mobile, whilst elderly people may be more likely spreaders (Dowd et al., 2020).

To understand which of the proposed mechanisms may explain the impacts of density, I include exposure interactions with a selection of variables that could explain virus transmissions.[9] Table 2

[9]The dataset (available from the author's website) includes some other variables, including different types of health insurance coverage, mandated paid sick leave areas, poverty and jobs requiring exposure to infection. These do not change the conclusions.

Covid Economics 37, 14 July 2020: 1-26

reports regressions of the form:

$$\Delta i_{ot} = \beta_0 \Delta Exposure_{ot} + \beta_2 \Delta Exposure_{ot} * \log density + \beta_3 \Delta Exposure_{ot} * x_o + \alpha_{st} + u_{ot}. \quad (5)$$

In this regression equation, β_3 identifies how the transmission rate varies with some variable x_o, conditional on the log population density of that area. Unemployment is associated with significantly lower transmission (conditional on the effect of wages). On a 10% significance level the coefficient for public transport is significant too. A sample standard deviation increase in public transport is associated with about 2% faster transmission. Whether conditioning on density or not (columns 2 and 3), the coefficients for the other factors are very comparable, suggesting that their impacts may not correlate much with density.

When focusing on areas in the first month of infections, the results change (columns 4-6). The coefficients for density are more pronounced; and the role of public transport changes: no effect is found conditional on density.

Table 2: Transmission equation and covariates related to density

	(1)	(2)	(3)	(4)	(5)	(6)
		Full sample			First month	
exposure	1.12***	7.84**	7.84**	0.42	7.84	12.79**
	(0.19)	(3.50)	(3.31)	(0.34)	(5.23)	(5.95)
exposure interacted with..						
log density	0.03*	0.00		0.11**	0.16***	
	(0.02)	(0.03)		(0.04)	(0.04)	
public transport share		0.05*	0.05***		-0.02	0.08*
		(0.03)	(0.02)		(0.05)	(0.04)
health insurance coverage		0.03	0.03		1.02	0.98
		(0.34)	(0.36)		(0.98)	(0.99)
log average wage		-0.55	-0.55		-0.31	-0.55**
		(0.37)	(0.42)		(0.24)	(0.24)
unemployment rate		-5.14***	-5.14**		-4.72**	-4.93*
		(1.96)	(2.21)		(2.33)	(2.53)
proximity index of jobs		-0.26	-0.27		-1.28	-1.85
		(0.86)	(0.95)		(1.29)	(1.37)
work-from-home		0.31	0.31		-2.04	-1.17
share of jobs		(3.22)	(3.48)		(1.40)	(1.19)
share elderly (>70)		-0.55	-0.55		0.00	0.69
		(0.73)	(0.80)		(2.04)	(2.17)
share children		0.13	0.13		-0.03	1.06
		(1.47)	(1.75)		(2.67)	(2.93)
Observations	45,880	45,880	45,880	12,234	12,234	12,234
R-squared	0.86	0.86	0.86	0.87	0.87	0.87
state-day FE	yes	yes	yes	yes	yes	yes

Exposure based on 5 days infectious time and 6 days detection time
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Weighted by county population. Twoway (county and day) clustered standard errors

The role of these urban factors could vary substantially over time, as the overall role of density varies, too. Figure 4 below plots how the transmission rate varies with the three covariaties most highly correlated to population density in the dataset: public transport, wages, and the share of work-from-

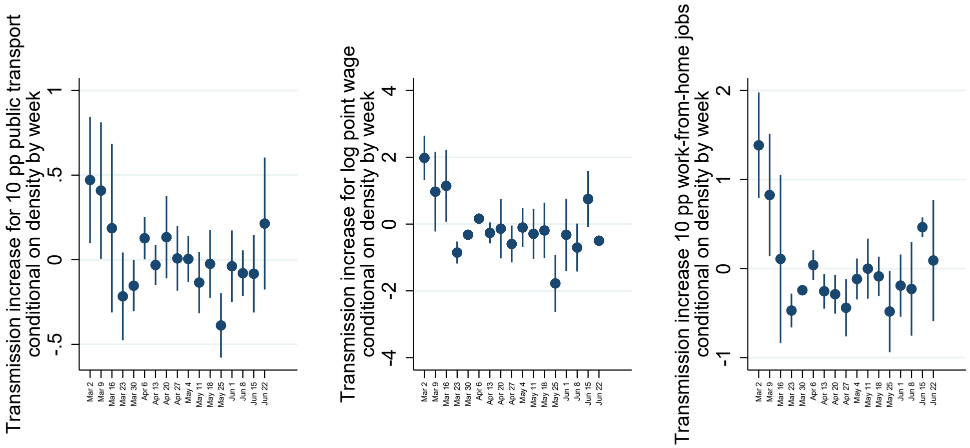


Figure 4: Change in the estimated transmission rates with an increase in public transport use (per 10 percentage point), wage (per log point) and the share of work-from-home jobs (per 10 percentage point), by week.

home jobs in total jobs. The regression equation is:

$$\begin{aligned} \Delta i_{ot} \quad = \quad & \sum_w \beta_{1w} \Delta Exposure_{ot} + \sum_w \beta_{2w} \Delta Exposure_{ot} * \log density_o + \sum_w \beta_{3w} \Delta Exposure_{ot} * x_o \\ & + \sum_w \alpha_{2w} * \log density_o + \sum_w \alpha_{3w} * x_o + \alpha_{st} + u_{ot}, \end{aligned} \tag{6}$$

where β_{3w} measures how a unit increase in the covariate x_o changes the transmission rate in week w, conditional on the weekly impact of density.

Figure 4 graphs the contributions of public transport, wages and work-from-home jobs by week, conditional on the impact of density. The unit of measure for public transport is 10 percentage points in the share of commutes. The point estimate in the first panel hence implies that if a location has 40% instead of 30% of its commuters in public transport, the transmission rate is 0.4 points higher. This interpretation is conditional on the weekly impact of population density. As the outbreak progresses, public transport is not significantly or occasionally negatively associated with transmission. Across all weeks, the weekly impact of public transport use by itself, α_{3w} (i.e., its impact independent of infection rates) is insignificant.

The second and third panel of Figure 4 repeat the analysis for wage levels and work-from-home job shares. Conditional on weekly density effects, a log point higher wage is associated with 2 points increase in the transmission rate in the week of March 2. This impact turns negative or insignificant over the course of the pandemic. Similarly, locations with large share of jobs that permit working from home see high transmission rates at first: if the share of jobs that permit working from home is 10 percentage points higher, the estimated transmission rate is 1.2 points higher. This effect is sizable, but the county cross-sectional standard deviation in work-from-home jobs shares is about 0.04. In the weeks following March 23rd, 12 out of the 14 estimates point to an association of work-from home jobs with lower transmission speeds.

3.3 Sheltering and mobility differences across locations

3.3.1 Trip reductions

Plausibly, the mobility responses of people faced with exposure to the virus differ by location. If an area of high density forces people to interact more, then its inhabitants might stay home and avoid travel. To examine how people adjust their trips to exposure, I employ two sources of phone-tracking data. First, I use Google's mobility reports, which tracks trips by destination by county. I employ the percentage change in measured trip frequency in a county compared the its frequency of the same trips in the period January 3 to February 6, 2020. Second, I use patterns of "pings" based on PlaceIQ (Couture et al., 2020), calculating what share of the pings in a county in the last 14 days were in the same county of residence. Higher shares of pings in the home county indicate that the phone spent more time in the home county.

The regression is as follows:

$$\begin{aligned} \Delta trips(\%)_{ot} &= \beta(x) \sum_{d} s_{od} * \text{infectious rate}_{dt} + \alpha_o + \alpha_{st} + u_{ot} \qquad (7) \\ \text{infectious rate}_{dt} &= \frac{\text{infectious}_{dt}}{pop_d/1000} \end{aligned}$$

The outcome, $\Delta trips(\%)_{ot}$, measure how trip frequency changed relative to the baseline frequency of trips. The explanatory variable of interest, $\sum_d s_{od}*\text{infectious rate}_{dt}$, measures the weighted average of infections per 1,000 people. The weights s_{od} are commuting flow shares, and add up to 1 for all destinations for a given origin o. As a result, the variable measures the infectious rate per 1,000 people that a typical commuter from o faces at time t. As before, the number of infectious people is a 5-day window of cases. Hence, the coefficient β measures the percentage change in number of trips undertaken that is associated with one extra infection per 1,000 people. The estimating equation has no lagged dependent variable, so trip changes are estimated with a county fixed effects in addition to the state-day fixed effect. Note that as the dependent variable is specified in changes relative to a baseline period of trip frequency, this takes out the average frequencies in the mobility.

First, to check whether mobility responds to infections, I regress trip frequency changes on infection rates, conditional on county and state-year fixed effects. The results in Appendix E report the regressions. They show that one infection per 1,000 people is associated with 0.34% fewer workplaces trips, and similarly intuitive results on transit, parks, shopping and recreation. Additionally, they show that staying home is more likely, both in the Google and PlaceIQ data.

The first question of this section is whether density and other covariates influence the sensitivity of trip choice to exposure. Table 3 presents regressions that include an interaction of exposure with density and other covariates. Population density significantly amplifies the mobility responses to exposure: for a county with one log point higher density an additional exposure of 1 cases per 1,000 people is associated with 0.09% lower travel to work and 0.05% higher home tracking percentages (or: when exposure increases by a standard deviation, a standard deviation in log density is associated to a 1.7%p stronger reduction in travel to work and a 0.9%p stronger increase in home stays; or at sample means, an additional log point of population density is associated with 15% stronger declines in travel to work trips). Introducing other covariates in column 3, the share of jobs that can be done from home significantly amplifies the decline in travel to work, when exposure increases. A ten percentage point increase in the share of teleworkable jobs leads to an additional 0.5% reduction in travel to work. For staying at home, in column 8, only wages have a significant impact to increase the rate at which people stay at home when exposure increases.

The responses look considerably different in the first month after a county develops 10 cases, as

Covid Economics 37, 14 July 2020: 1-26

reported in columns 4 and 5, and columns 9 and 10 of Table 3. Density plays an insignificant (work travel, column 4) or substantially smaller (home stay, column 9) in the first month of infections, as compared to the full period. Counties with larger shares of jobs that can be done from home saw significantly weaker reductions in travel to work in the first month as opposed to stronger reductions later on, and significantly weaker increases in homestaying. Possibly, the option to work from home was only effectuated after a month of epidemic. Areas with higher wages, on the other hand, do see a stronger reduction in travel to work and a stronger increase in homestays in the first month.

Table 3: Mobility responses to infection rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mobility	Work	Work	Work	Work	Work	Home	Home	Home	Home	Home
Sample		full sample		1st month			full sample		1st month	
Infectious rate d	-0.60***	0.36**	2.29	-0.02	4.12**	0.32***	-0.13*	-3.05**	-0.05	-1.46*
interacted with	(0.08)	(0.15)	(2.38)	(0.07)	(1.68)	(0.05)	(0.07)	(1.29)	(0.04)	(0.73)
log density		-0.09***	-0.05	-0.01	-0.03*		0.05***	0.02	0.01**	0.02***
		(0.02)	(0.04)	(0.01)	(0.02)		(0.01)	(0.02)	(0.01)	(0.01)
public transport			0.02		0.00			-0.01		-0.00
			(0.03)		(0.01)			(0.01)		(0.00)
log wage			-0.05		-0.55***			0.24*		0.18***
			(0.28)		(0.19)			(0.14)		(0.07)
work-from-home			-0.46*		0.46***			0.15		-0.15***
jobs			(0.27)		(0.13)			(0.11)		(0.04)
Observations	116,121	22,980	22,980	7,541	7,541	60,594	22,945	22,945	7,514	7,514
county FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
state-day FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Exposure based on 5 days infectious time and 6 days detection time
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Twoway (county and day) clustered standard errors

3.3.2 Transmissions on the commuting network

Do mobility choices affect the rate of spread of the virus? Although exposure following the commuting network predicts new cases best, the relevance of workplace infections may change over time. To examine changes over time, I estimate weekly transmission rates, but now I include a measure of workplace-level exposure in addition to the commuting network exposure in a variation of the original equation (2):

$$\Delta i_{ot} = \sum_{w} \beta_w \Delta Exposure_{ot} + \sum_{w} \beta_w^d Exposure_{ot}^d + \alpha_{st} + u. \quad (8)$$

Here the original variable $Exposure_{ot} = S_{o,t-\tau} \sum_d s_{o,d,t-\tau} \frac{I_{dt-\tau}}{N_d}$ is the same as before, and $Exposure_{ot}^d = S_{o,t-\tau} \sum_{d \neq o} s_{o,d,t-\tau} \frac{I_{dt-\tau}}{N_d}$ is the commuting exposure measured in all counties except in the county o where the infection is confirmed. The interactions $s_{o,d,t-\tau}$ are the same between the two variables, but $s_{o,d,t-\tau}$ is not included in $Exposure_{ot}^d$ for $o = d$. Hence, coefficient β_w measures the transmission as before and coefficient β_w^d is an estimate of the additional transmission from workplace counties. If the transmission rates are equal (under the assumption that commuting patterns reflect likelihoods of interaction), $\beta_w^d = 0$. To save on space, I plot the coefficients of the two terms in Figure 5.

Figure 5 shows that exposure in workplace counties initially has no significantly different likelihood of leading to new infections. Over the course of the crisis, and in particular from early March on, counties are less likely to develop more cases if the counties in which its inhabitants work have higher

Covid Economics 37, 14 July 2020: 1-26

Covid Economics 37, 14 July 2020: 1-26

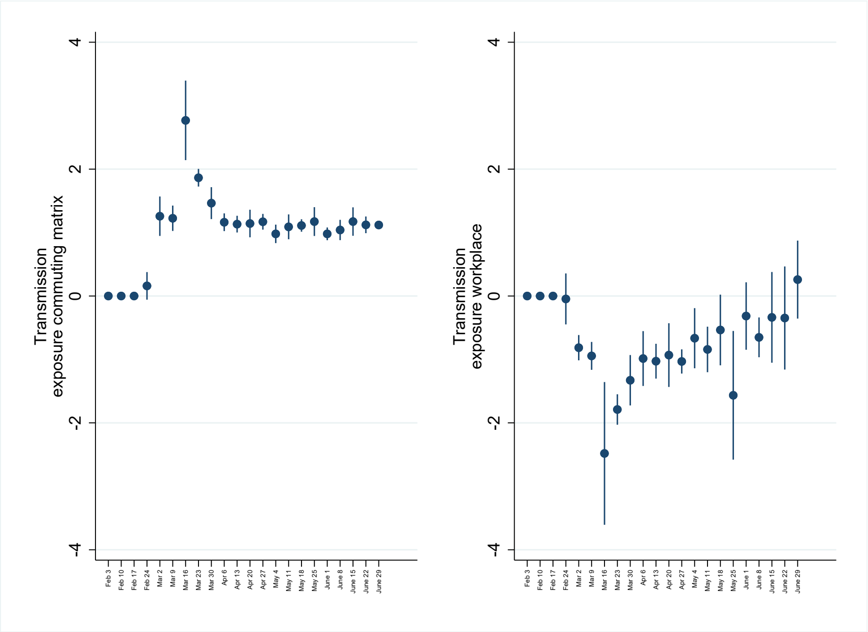


Figure 5: Transmission estimate from the commuting network and conditional transmission from non-home counties

infection rates. Workplaces outside the own county become less important in explaining infections. Yet, from April onward, workplace infections slowly reclaim their role as their transmission coefficient grows, with no significant difference between home and workplace infection likelihood in the most recent weeks.

3.3.3 Mobility destination choices

Do travelers avoid places with high rates of infections? Previous results suggest that overall mobility declined. Phone movement data may reveal how travel behavior changes when faced with exposure in different locations. I use the PlaceIQ data on movement between counties (Couture et al., 2020) that record what share of the phones residing in a county has "pinged" (registered to a cell phone tower) in another county over the previous two weeks. I estimate the regression equation:

$$pingshare_{odt} = \beta \frac{cases}{cap(k)}_{dt} + \alpha_{od} + \alpha_{ot} + u_{odt} \tag{9}$$

The dependent variable pingshare$_{odt}$ is the share of phones residing in origin county o that ping in destination county d over the two weeks of period t. The independent variable $cases/cap(k)_{dt}$ is the mean daily number of new cases per 1,000 inhabitants over the preceding two weeks in destination d. The coefficient β measures the association between infectious rates in d and the share of phones that have been active in d over the period. As the phone-tracking data at any day is based on two weeks of historical data, I take the first and 16th day of every month to avoid overlap in the time periods. The fixed effect α_{od} differences out any time-invariant variables that are specific to the origin-destination pair, such as distance and the road networks connecting them. This set of fixed effects also differences out origin and destination variation that does not vary over time, such as population density.

It is well possible that people become less mobile in general when faced with higher infection rates in their vicinity. The fixed effect α_{ot} controls for the average mobility response in the origin county. Hence, β measures whether people visits destination with higher infection rates less, compared to other destinations.

Table 4 presents estimates for how (phone-tracked) mobility is associated with infection rates. Column 1 reports an estimate only conditional on the time-invariant origin-destination variables, such as distance. The magnitude is substantial: an additional daily case per 10,000 inhabitants in the preceding two weeks in a destination is associated with a reduction of about 11 percentage points in the share of phones that moves to that area. The impact is likely to be overstated. High infection rates in nearby counties correlate with various other reasons not to travel, such as infection rates in the origin county, overall uncertainty, stay-at-home orders, or reliance on the same local health care infrastructure.

Column 2 of Table 4 exploits the bilateral structure of the data to control for the overall willingness to travel at the origin, by including an origin-time fixed effects. The origin-time fixed effect controls for overall declines in the propensity to be mobile, including home county infection rates, updated information on viral progression or time-varying state-level policies. The interpretation of the coefficient in column 2 is that the ping share of citizens from a given origin county declines by 5 percentage points faster to a destination with an extra infection per 10,000 inhabitants as compared to the other destinations. The magnitude of this coefficient is less than half of the coefficient reported in column 1. Hence, the decline in travel to a county with higher infection rates is driven in part by active avoidance of travelers into that county, and in part by overall reductions in willingness to travel by people that are likely to travel to the county.

Column 3 of Table 4 reports a regression that estimates the coefficient using two intervals per month. The coefficient turns negative in April, suggesting avoidance of counties with infection rates developed over the course of the outbreak. The coefficient is estimated to be closer to zero as time progresses. This may point to lass active avoidance by June.

Columns 4 to 7 estimate whether the mobility-deterring effect of infections depends on density and transport choice. In column 4, an additional case per 1,000 people is associated with 28 percentage point lower phone registration share when the county has a log point higher density. Column 5 suggests that the decline in travel to an infected area is significantly stronger when more of commuters between the areas use public transport. The sample of county pairs for which transport choice data are available is limited, leading to a significant reduction in the sample size. When including the interaction with density on column 6, however, no significant effects of public transport are found. Column 7 reiterates the regression with density interacted with exposure in that limited sample, showing substantially larger coefficient magnitude for the density interaction, in the reduced sample compared to the full sample.

Do containment policies distort these estimates? Regressions in Appendix F suggest that issuance of stay-at-home orders substantially reduces the transmission estimate (by about 0.8 points). Other policies (restaurants, schools, foreign travel) show no significant impact. However, these estimates are conditional on the state-day fixed effects, and there may be limited variation within states in the timing of these policies. Density shows no significant moderation of the impact of stay-at-home orders on transmission rates.

Covid Economics 37, 14 July 2020: 1-26

Covid Economics 37, 14 July 2020: 1-26

Table 4: Share of phones by origin that are active in a destination county in the preceding two weeks

Share of phones active in destination county	(1)	(2)	(3)	(4)	(5)	(6)	(7)
cases per 1,000 cap	-1.15***	-0.53***		0.73**	-23.71***	125.49***	127.45***
(daily over 2 weeks)	(0.17)	(0.06)		(0.29)	(4.30)	(19.84)	(19.61)
interacted with							
log density				-0.28***		-18.41***	-18.71***
				(0.29)		(2.45)	(2.42)
public transport					-29.99***	-4.97	
					(6.94)	(5.82)	
(fortnight of..)							
March 15st			822.56**				
			(380.25)				
April 1st			-10.74***				
			(2.60)				
April 15th			-3.36***				
			(0.60)				
May 1st			-1.35***				
			(0.21)				
May 16th			-0.62***				
			(0.10)				
June 1st			-0.76***				
			(0.10)				
June 15th			-0.46***				
			(0.07)				
Observations	21,810,660	21,810,660	21,810,660	17,985,505	42,196	42,196	42,196
origin-destination FE	yes	yes	yes	yes	yes	yes	yes
origin-time FE	no	yes	yes	yes	yes	yes	yes

Destination county-day clustered standard errors
*** p<0.01, ** p<0.05, * p<0.1

4 Conclusions

Population density is associated with higher rates of COVID-19 transmission. The role of density concentrates in particular on the initial phase of the epidemic, and on locations of very high density. As the epidemic progresses, the association between density and viral transmission becomes smaller. Higher income, high work-from-home job shares and public transport use concentrate in dense areas and are initially linked to additional increases in virus transmission.

People in densely populated areas are more likely to shelter when faced with the risk of infection. However, sheltering only starts after the first month of local infections. Phone mobility also shows less movement towards infected areas, even relative to overall movement declines. Avoidance is less strong when commuters use public transport. Consistent with reduced mobility, new cases become less likely to be linked to workplace infections during the outbreak.

The results are consistent with the idea that larger and denser cities were vulnerable to faster spread of the virus mostly at the onset of the epidemic. The differences between dense cities and other areas fade over time. The decline in transmissions in cities relative to other areas is possibly driven in part by their stronger sheltering responses.

At least two qualifications are in order with these results. First, the analysis is descriptive, as it describes the statistical association of density with transmission. It cannot rule out that variables correlated to density but not included in the analysis form underlying drivers of viral transmissions. Additionally, the sample used for identification is effectively changing over time, especially when using fixed effects. The outbreak seems to progress from more dense areas into less dense areas, so examining the development of transmission coefficients over time might conflate changes in the transmission process with changes in the places where they occur, at least condtional on the fixed effects. Second, the county-level infection data are from a journalistic source. Case counts, especially for the most recent few days, are occasionally revised. As the data collection is ongoing, results may change over the course of the epidemic. The code is available from the author's website, and it also updates all results with the most recent data when run.

References

Adam Brzezinski, Guido Deiana, V. K. (2020). The covid-19 pandemic: Government versus community action across the united states. *CEPR Covid Economics: Vetted and Real-Time Papers*, 7:115–156.

Almagro, M. and Orane-Hutchinson, A. (2020). The determinants of the differential exposure to covid-19 in new york city and their evolution over time. *Covid Economics: Vetted and Real-Time Papers*, (13).

Chiou, L. and Tucker, C. (2020). Social distancing, internet access and inequality. Working Paper 26982, National Bureau of Economic Research.

Couture, V., Dingel, J., Green, A., Handbury, J., and Williams, K. (2020). Exposure indices derived from PlaceIQ movement data. https://github.com/COVIDExposureIndices/COVIDExposureIndices.

Coven, J. and Gupta, A. (2020). Disparities in mobility responses to covid-19. Technical report, NYU Stern Working Paper.

Crowley, F., Daly, H., Doran, J., and Ryan, G. (2020). Covid-19, social distancing, remote work and transport choice. *CEPR Covid Economics: Vetted and Real-time papers*, 30.

Dingel, J. I. and Neiman, B. (2020). How many jobs can be done at home? *CEPR Covid Economics: Vetted and Real-time papers*, 1:16–24.

Dowd, J. B., Andriano, L., Brazel, D. M., Rotondi, V., Block, P., Ding, X., Liu, Y., and Mills, M. C. (2020). Demographic science aids in understanding the spread and fatality rates of covid-19. *Proceedings of the National Academy of Sciences*, 117(18):9696–9698.

Engle, S., Stromme, J., and Zhou, A. (2020). Staying at home: Mobility effects of Covid-19. *CEPR Covid Economics: Vetted and Real-time papers*, 4:86–102.

Fang, W. and Wahba, S. (2020). Urban Density Is Not an Enemy in the Coronavirus Fight: Evidence from China. World Bank Sustainable Cities blog, April 20, 2020.

Florida, R. (2020). The geography of the coronavirus. *Bloomberg: Citylab, April 3, 2020*.

Furth, S. (2020). Automobiles Seeded the Massive Coronavirus Epidemic in New York City. Market Urbanism, April 19, 2020.

Google (2020). Google mobility reports retrieved from https://www.google.com/covid19/mobility/.

Gormley, M., Aspray, T. J., and Kelly, D. A. (2020). Covid-19: mitigating transmission via wastewater plumbing systems. *The Lancet Global Health*, 8(5):e643.

Covid Economics 37, 14 July 2020: 1-26

Harris, J. E. (2020). The subways seeded the massive coronavirus epidemic in New York City. *NBER Working Paper*, (w27021).

He, X., Lau, E. H., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y. C., Wong, J. Y., Guan, Y., Tan, X., et al. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature medicine*, 26(5):672–675.

Heroy, S. (2020). Metropolitan-scale covid-19 outbreaks: how similar are they? *arXiv: Populations and Evolution*.

Li, R., Richmond, P., and Roehner, B. M. (2018). Effect of population density on epidemics. *Physica A: Statistical Mechanics and its Applications*, 510(C):713–724.

Pung, R., Chiew, C. J., Young, B. E., Chin, S., Chen, M. I., Clapham, H. E., Cook, A. R., Maurer-Stroh, S., Toh, M. P., Poh, C., et al. (2020). Investigation of three clusters of COVID-19 in Singapore: implications for surveillance and response measures. *The Lancet*.

Ribeiro, H. V., Sunahara, A. S., Sutton, J., Perc, M., and Hanley, Q. S. (2020). City size and the spreading of COVID-19 in Brazil. *arXiv preprint arXiv:2005.13892*.

Ruggles, S., Flood, S., Meyer, R. G. J. G. E., Pacas, J., and Sobek, M. (2020). IPUMS USA: Version 10.0 [dataset].

Siordia Jr, J. A. (2020). Epidemiology and clinical features of COVID-19: A review of current literature. *Journal of Clinical Virology*, page 104357.

Song, L.-P., Zhang, R.-P., Feng, L.-P., and Shi, Q. (2017). Pattern dynamics of a spatial epidemic model with time delay. *Applied Mathematics and Computation*, 292:390–399.

Stier, A. J., Berman, M. G., and Bettencourt, L. M. A. (2020). Covid-19 attack rate increases with city size. *medRxiv 2020.03.22.20041004*.

Tian, H., Liu, Y., Li, Y., Wu, C.-H., Chen, B., Kraemer, M. U., Li, B., Cai, J., Xu, B., Yang, Q., et al. (2020). An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*, 368(6491):638–642.

Wakefield, J., Dong, T. Q., and Minin, V. N. (2019). *Spatio-temporal analysis of surveillance data*, pages 455 – 476. In: Leonhard Held, Niel Hens, Philip D O'Neill, Jacco Wallinga, Handbook of Infectious Disease Data Analysis, CRC Press.

Wallinga, J. and Lipsitch, M. (2007). How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1609):599–604.

Covid Economics 37, 14 July 2020: 1-26

A Data

A.1 Phone mobility coverage

Figure 6 shows the 2,018 counties included in the PlaceIQ dataset. A county is included if has at least 1,000 registered devices that ping on at least 11 days in any consecutive 14 day period since November 1st, 2019. The residence county of a device is assigned as the county in which the device registers the largeste cumulative time at a residential location.

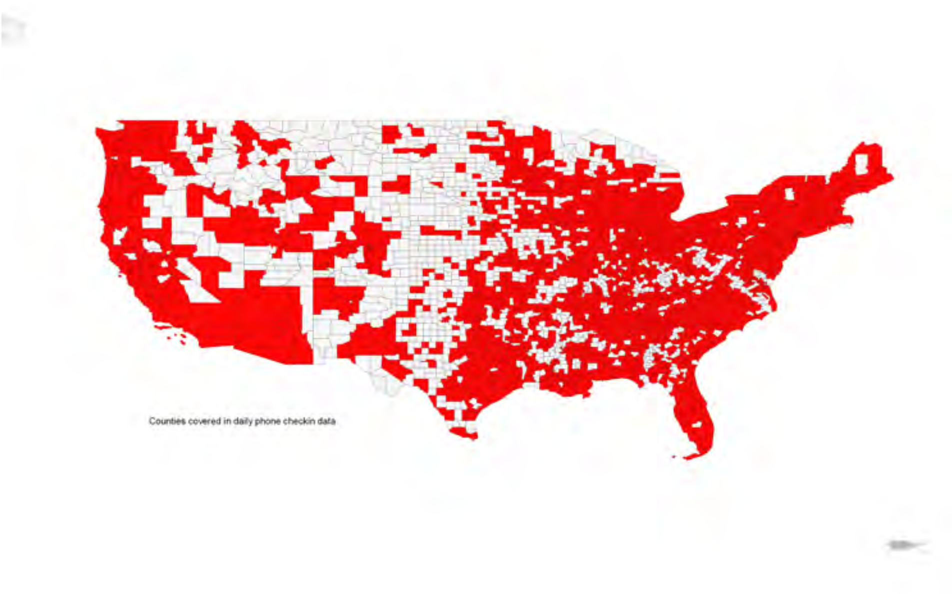


Figure 6: Coverage of counties in PlaceIQ pings.

A.2 Job context matching

In order to derive proximity indexes (and other job characteristics), I connect O*NET job context data on occupations to county-level estimates of the prevalence of those occupations. The county-level job requirements for proximity are approximated as the employment-weighted average of the context scores from O*NET. I use the 2018 crosswalk from the U.S. Census Bureau between the 2018 Census Code and the 2018 2018 SOCcode. The hierarchical procedure is as follows: I merge based on the 6-digit crosswalk, then merge unmatched occupations on a 5-digit crosswalk, and then merge the remaining occupations on a 4-digit crosswalk. Using this, I merge 320 6-digit occupations, followed by 132 original 6-digit occupations on a 5-digit scheme, followed by 43 original 6-digit occupations on 4 digit scheme.

Covid Economics 37, 14 July 2020: 1-26

B Estimates for assumptions on infectious time and incubation

Figure 7 shows the model root mean squared errors (left-hand panels) and the transmission coefficient estimates (right-hand panels). The panel "commutes" defines the interaction s_{od} as the share of commuters from county o that travel to d, according to the 2018 ACS. The panel "residence" defines s_{od} equal to one when the origin equals the destination, and zero otherwise. The panel "normping" uses the share of phones that has pinged in another location, as a proxy for movement of people to that location. The panel "normping" present result from (row-)normalized shares in which the interaction terms for a given origin sum to one when aggregated over destinations. The panel "ping" simply uses the share of phones that has pinged in the destination, which leads to larger numbers s_{od}, because phones can ping at multiple locations. Not normalizing allows for changes over time in the mobility level to play a role, but the interpretation of the estimate of β is changed.

Overall, the best fit, judged as lowest RMSE, occurs for mobility based on commuting patterns, suggesting that infection rates at workplace counties (including the own county) explain new infections. Restricting the assumed exposure to the own county yields somewhat less precise model fits, while using pings (phone-tracked mobility estimates) shows poorer fits. Estimates reported north of the diagonal involve longer infectiousness periods than delay in detection times, such that the number of new cases in a given day is included in that day's exposure measure. Hence, the estimates north of the diagonal, in particular at short time horizons (e.g. 3 days infectiousness, 2 days detection), tend to obtain low RMSEs by construction. Restricting attention to plausible below-diagonal estimates (infectiousness shorter than detection times), the transmission estimates for commuting based, residence based and normalised ping based mobility are very similar, at around 1.03 for all measures.

Covid Economics 37, 14 July 2020: 1-26

Covid Economics 37, 14 July 2020: 1-26

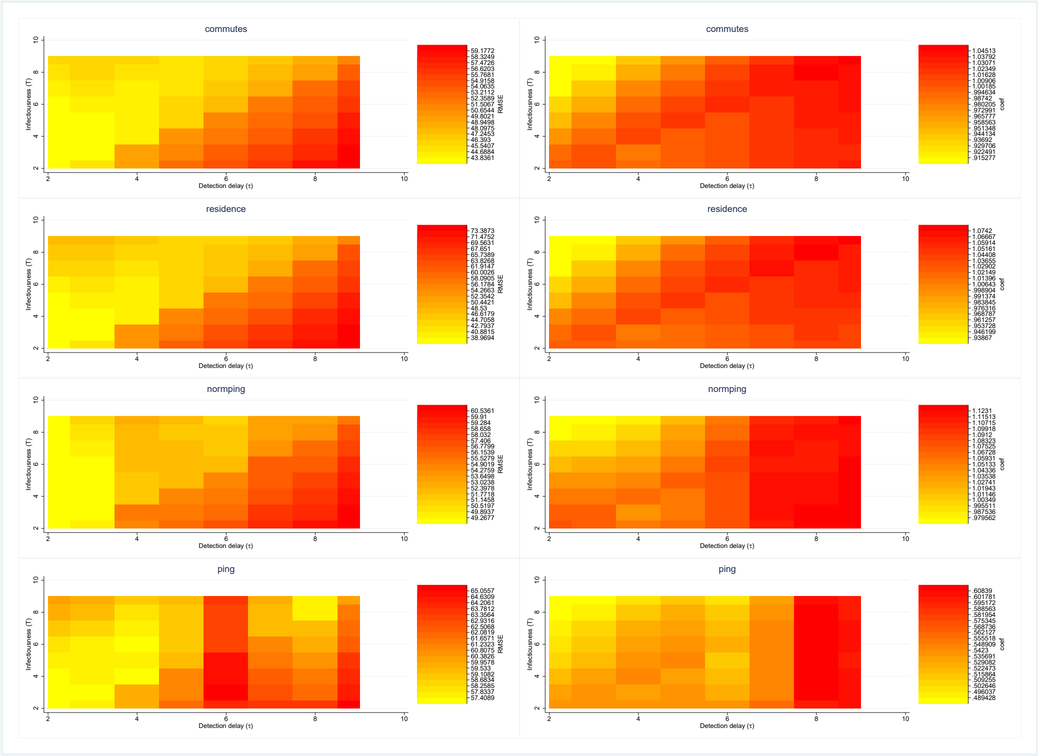


Figure 7: Tranmission coefficient and Root Mean Squared Errors for combinations of exposure (days of infectiousness vertical axis) and lag (days between presumed infection and registration of case).

C Development of infected rates for low, medium and high density counties

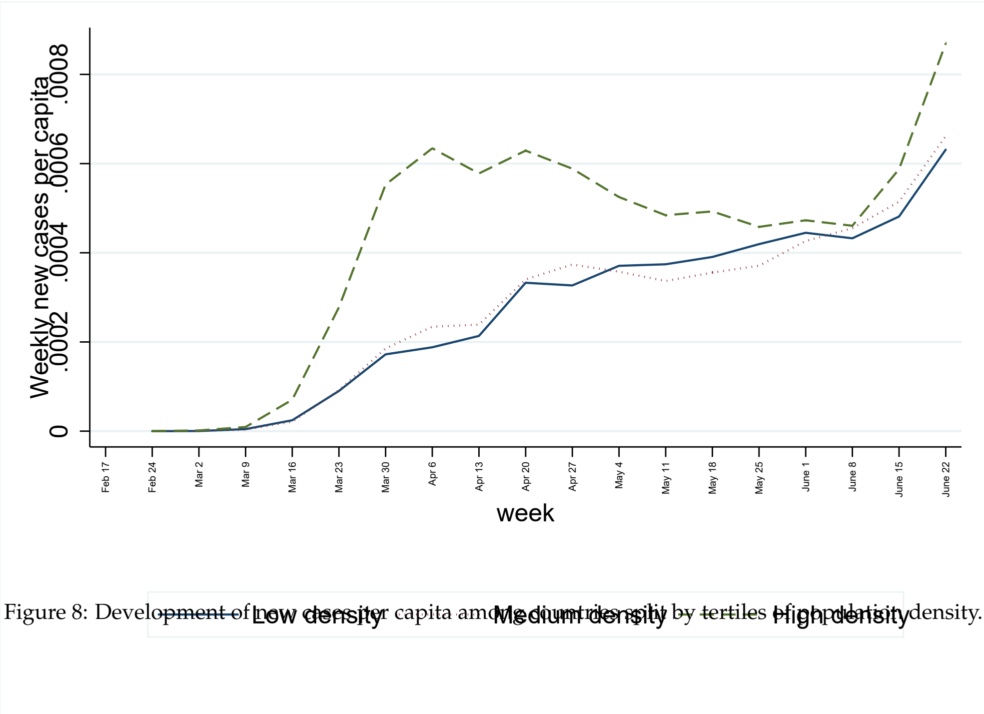


Figure 8: Development of new cases per capita among countries split by tertiles of population density.

Covid Economics 37, 14 July 2020: 1-26

D Estimates by decile of density

Table 5 repeats the regressions for the baseline model, estimating a coefficient by decile of population density.

Table 5: Transmission equation conditioned on deciles population density

	(1)	(2) 1st month after first 10 cases
Q1/10 exposure	0.19	0.86***
	(0.12)	(0.13)
Q2/10 exposure	0.62***	0.85***
	(0.15)	(0.13)
Q3/10 exposure	0.78***	1.12***
	(0.14)	(0.17)
Q4/10 exposure	0.92***	1.05***
	(0.06)	(0.04)
Q5/10 exposure	0.89***	0.84***
	(0.05)	(0.07)
Q6/10 exposure	0.86***	0.92***
	(0.12)	(0.03)
Q7/10 exposure	0.90***	0.94***
	(0.05)	(0.07)
Q8/10 exposure	0.90***	0.86***
	(0.04)	(0.07)
Q9/10 exposure	1.05***	0.95***
	(0.03)	(0.05)
Q10/10 exposure	1.06***	1.14***
	(0.02)	(0.05)
Observations	407,630	73,014
state-day FE	yes	yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Twoway (county and day) clustered standard errors

Covid Economics 37, 14 July 2020: 1-26

CEPR PRESS

E Shelter in place mobility choices

Table 6 reports regressions for different forms of trips. Differently put, one extra infection per 1,000 people leads to 0.34% decline in trips to workplaces. A sample standard deviation increase in the exposure rate is linked to about 3.6% decline in workplace trips. Phone tracking in the residence increases significantly, while tracking in transit stations, parks, groceries, pharmacies, retail and recreation all decline significantly. The share of phone pings in the county of residence also increases significantly, implying that phones leave the resident county less often.

Table 6: Mobility responses to infection rates

	(1) Workplace	(2) Residence	(3) Transit	(4) Parks	(5) Groceries and pharmacies	(6) Retail and recreation	(7) Residential county ping share
share infected	-0.34***	0.21***	-0.39***	-0.43**	-0.37***	-0.36***	0.03***
	(0.05)	(0.04)	(0.10)	(0.21)	(0.08)	(0.07)	(0.00)
Observations	116,114	60,584	48,965	34,928	101,387	105,163	183,748
R-squared	0.94	0.96	0.85	0.77	0.80	0.92	0.96
county FE	yes	yes	yes	yes	yes	yes	yes
state-day FE	yes	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Twoway (county and day) clustered standard errors

Covid Economics 37, 14 July 2020: 1-26

CEPR PRESS

F Containment policies

Table 7: Tranmission rates under different containment policies

	(1)	(2)	(3)	(4)	(5)	(6)
Exposure	0.63	-5.99*	-0.16	-0.20**	-1.20**	-0.49
	(0.39)	(3.50)	(0.21)	(0.10)	(0.58)	(0.63)
Stay-at-home order	6.29***				4.82***	4.26***
	(2.16)				(1.28)	(1.42)
Stay-at-home order X exposure	-0.79*				-0.57*	-0.30
	(0.40)				(0.32)	(0.72)
Public school order		2.32			-3.94	
		(2.13)			(3.67)	
Public school order X exposure		5.83			8.73*	
		(3.52)			(4.41)	
Restaurant order			1.96***		6.09	
			(0.39)		(3.86)	
Restaurant order X exposure			0.01		-5.91	
			(0.23)		(3.64)	
Foreign travel ban				0.87**	2.75	
				(0.41)	(1.66)	
Foreign travel ban X exposure				0.05	-1.24	
				(0.13)	(0.95)	
log density X exposure						0.08
						(0.16)
Stay-at-home order X						-0.00
log density X exposure						(0.16)
Observations	162,716	162,716	162,716	162,716	162,716	162,635
R-squared	0.93	0.93	0.93	0.93	0.93	0.88
county FE	yes	yes	yes	yes	yes	yes
state-day FE	yes	yes	yes	yes	yes	yes

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Twoway (county and day) clustered standard errors

Covid Economics 37, 14 July 2020: 1-26

Propagation of epidemics' economic impacts via production networks: The cases of China and ASEAN during SARS and COVID-19

Ammu George,[1] Changtai Li,[2] Jing Zhi Lim[3] and Taojun Xie[4]

Date submitted: 4 July 2020; Date accepted: 7 July 2020

Two decades after the SARS outbreak, Asia is confronted with COVID-19 which has caused a greater economic impact to the region. In this paper, using China and the ASEAN's experiences of SARS and COVID-19 as a case study, we aim to identify the economic impact of a pandemic that is associated with global production linkages. We construct a novel general equilibrium model of production networks with epidemiological dynamics. Calibrating the model with the OECD inter-country input-output tables for the pre-SARS and pre-COVID-19 periods, and controlling for disease dynamics across years, we find that, in the absence of policy intervention, greater importance of China in the global value chains is associated with greater economic impacts, both within China and in the ASEAN region.

Covid Economics 37, 14 July 2020: 27-56

1 Research Fellow, Asia Competitiveness Institute, Lee Kuan Yew School of Public Policy, National University of Singapore.
2 PhD Candidate, School of Social Sciences, Nanyang Technological University.
3 Research Assistant, Asia Competitiveness Institute, Lee Kuan Yew School of Public Policy, National University of Singapore.
4 Senior Research Fellow, Asia Competitiveness Institute, Lee Kuan Yew School of Public Policy, National University of Singapore.

Covid Economics 37, 14 July 2020: 27-56

1 Introduction

Two decades after the outbreak of the Severe Acute Respiratory Syndrome (SARS), Asia was confronted again with Coronavirus Disease 2019 (COVID-19). As of June 6 2020, based on data from the World Health Organization, 6.8 million people have been infected with COVID-19 with 400,000 deaths worldwide. Similar to the SARS episode, soon after the COVID-19 outbreak in China, countries in the Association of Southeast Asian Nations, commonly known as the ASEAN, was one of the first regions affected due to its close geographical proximity, business travel, tourism and supply chain links to China. Aside from the human costs shown in Fig. 1, economies had also suffered. Comparing between the SARS and COVID-19 episodes, Fig. 2 shows more synchronized decline in GDP growth at the beginning of the COVID-19. Some have ascribed the heavier economic damage to the higher contagiousness of the coronavirus, while others have attributed it to the government mandated lockdowns around the world. There has also been a debate on whether China's increasing importance in the global value chains (GVCs) contributed to the current global economic downturn. Nevertheless, the jury is still out on the reason behind COVID-19's greater impact.

In this paper, we focus on one specific transmission channel. We ask: *What is the role of international production networks in the propagation of an epidemic's economic impact?* We derive the answers from the experiences of China and the ASEAN during the SARS and the COVID-19 periods.

Thus far, existing empirical analyses on the impact of disease outbreaks have faced a few challenges. Firstly, each epidemic, for example, the SARS or the 1918 Spanish Flu, has unique characteristics. They could differ in terms of the degrees of contagiousness, or the spreading media. Without sufficient incidents of such pandemics, it is difficult to control for these characteristics in an empirical analysis. An implication of this challenge is that findings from historical events may not be applicable for emerging new epidemics. Secondly, production and consumption linkages between and within countries have evolved. Taking China as an example, its growing importance in the GVCs was coupled with an expanding services sector domestically (Liao, 2020). The simultaneously evolving international and domestic economic conditions raise challenges in pinning down the transmission channels for economic impact.

This paper provides an analysis via a set of counter-factual simulations, which is particularly suited for singling out a specific transmission channel. For this purpose, all we need is the ability in an analytical framework to control for the variables mentioned previously. We construct a multi-country and multi-sector model with production networks. An SIR-Macro framework proposed by Eichenbaum et al. (2020) is used to capture the epidemiological dynamics. With the SIR-Macro model, we calibrate the population dynamics so that human costs of the disease outbreak are the same across scenarios, thus addressing the challenge of different disease characteristics. We then focus on sectoral dynamics using the framework proposed by Krueger et al. (2020): agents voluntarily substitute consumption in the more contagious sectors with those in the less contagious sectors. The setup of global production networks eventually allows this reallocation effect to propagate among trading partners. While calibrating the model, we choose the pre-SARS and pre-COVID-19 years as the normal-time scenarios. In the case of China and the ASEAN, this two time periods are largely similar in terms of domestic consumption and output patterns, but the latter is characterized by more integrated GVC networks. With these in hand, we have a toolkit for analyzing the transmission of an epidemic's economic impacts via global production networks.

We contribute to the growing discussion of the macroeconomic impact of COVID-19 in the context of production networks. Baqaee and Farhi (2020), Bonadio et al. (2020), and Luo et al. (2020) examine the impact on aggregate output based on the US economy. Perhaps the most closely related studies are Çakmaklı et al. (2020); Luo and Tsang (2020). Çakmaklı et al. (2020) consider the context of a small and open economy with trade and capital flows. Luo and Tsang (2020) study the impact on the world due to a shortage of labor in China. In our model, we do not assume exogenous shocks in macroeconomic variables, but an exogenous disease shock in population dynamics. All dynamics in the macroeconomic variables are outcomes of economic agents' welfare-maximizing behaviors following the disease outbreak.

Our innovation is also in the synthesized study of historical global pandemics. As mentioned, due to heterogeneity in disease characteristics, conclusions from

Covid Economics 37, 14 July 2020: 27-56

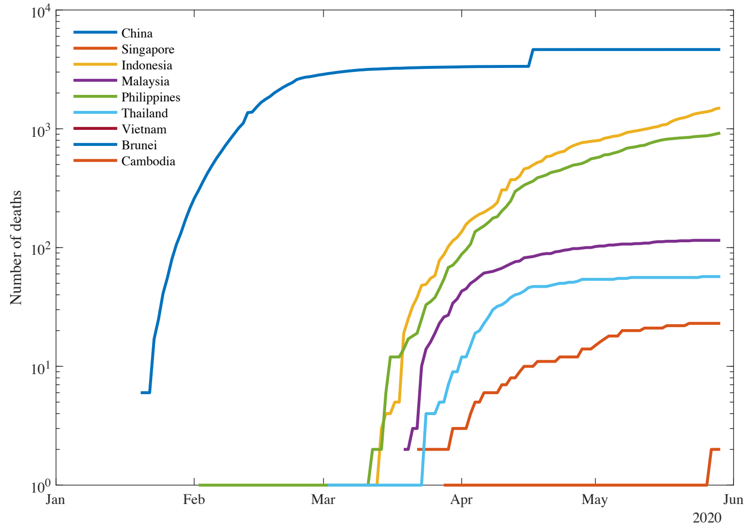


Figure 1: Cumulative death cases of COVID-19 in China and ASEAN.

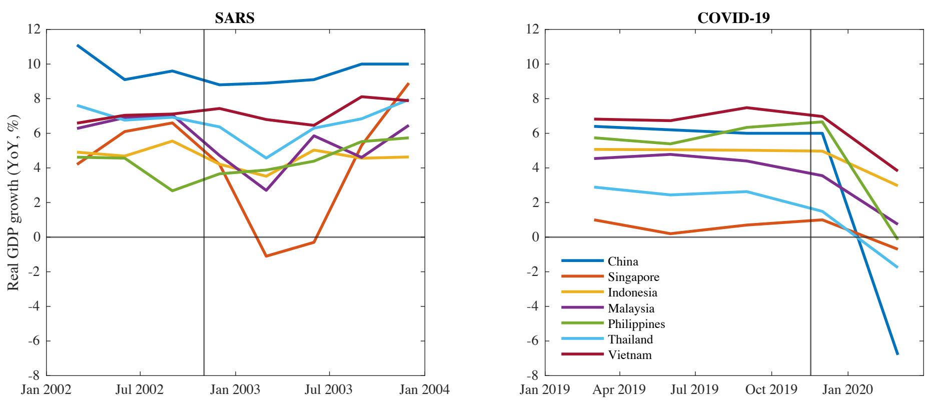


Figure 2: Real GDP of China and the ASEAN countries: Year-on-Year Change. The vertical lines indicate the dates of first cases.

Covid Economics 37, 14 July 2020: 27-56

Covid Economics 37, 14 July 2020: 27-56

historical analyses are often not applicable to new events, unfortunately. Case studies on China and the ASEAN are perhaps among the exceptions due to their common experiences in the SARS and COVID-19, as well as the similarity between the two viruses. It is opportune and instructive to make comparisons between now and then, and across countries on the economic impacts. By simply tweaking the infection parameters, our counter-factual simulations facilitate such comparisons by providing controlled settings across the two events. It is also possible for our model to be calibrated against actual population dynamics in a specific event.

The main findings of the paper show that the increasingly integrated production networks would contribute to the greater economic impacts of COVID-19. This is seen from three aspects. First, the within-country economic impacts on China during the time of COVID-19 is greater than during the SARS period, whereas the ASEAN would experience similar impact across the years. Second, despite the limited sizes, the spillover effects between China and ASEAN are larger during COVID-19 than during SARS. We argue that this is a result of evolving production linkages and trading patterns. Third, containment efforts in China helped reduce economic costs in the ASEAN.

Before proceeding to the remainder of the paper, it is equally important for our readers to bear in mind that the model presented in this paper is only a stylized one. Such simplicity is necessary for us to understand the transmission mechanism via the production networks, which is the key objective of ours. Producing empirically precise results, however, requires models in which the interactions are too complicated to be disentangled. We therefore leave the more complicated models for future research.

The remainder of the paper is organized as follows. Section 2 describes the model. Section 3 discusses the results from our simulations. Section 4 concludes.

2 A Model with International Production Networks

We construct a general equilibrium model with multiple countries and multiple sectors. During normal time, agents consume from all sectors of the economy

Covid Economics 37, 14 July 2020: 27-56

which sell composite items produced by domestic and foreign firms. At the same time, households supply labor to domestic firms of all sectors. The firms produce heterogeneous outputs, using labor from the households and intermediate outputs from all sectors of the economy. The intermediate outputs can be domestically produced or imported. The market is competitive. Nominal variables, including prices, wages, and nominal exchange rates, are assumed to be constant.

In the event of a pandemic, agents become heterogeneous as some of them contract the disease. Agents contract the disease while interacting with infected agents participating in the same consumption activity. The mechanism follows Eichenbaum et al. (2020). In addition, following Krueger et al. (2020), we differentiate the economic sectors by the degree of consumer interaction. A sector is said to be more infectious if consumers in this sector need to extensively interact among themselves, and vice versa. By consuming goods from the high-infection sectors, an agent faces increased risk of contracting the disease which reduces future welfare. All agents are initially susceptible. Upon contracting the disease, they become infected, who subsequently recover or decease. Susceptible, infected, and recovered agents behave differently due to their different utility functions. The aggregate economic outcomes are the resultant interactions of all agents.

2.1 Normal time

A representative agent in country ℓ derives utility from a consumption bundle and disutility from supplying labor. The agent's lifetime utility is given by

$$U_0(\ell) = \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(C_t(\ell), N_t(\ell)\right) \tag{1}$$

$$u\left(C_t(\ell), N_t(\ell)\right) = \log C_t(\ell) - \frac{\theta(\ell)}{2} N_t(\ell)^2 \tag{2}$$

where β is the discount factor, $C_t(\ell)$ is an aggregate of goods from all sectors, and $N_t(\ell)$ is the labor supply to domestic firms. The consumption bundle is a

CES aggregate of goods from all sectors indexed by j:

$$C_t(\ell) = \left(\sum_j \upsilon(j,\ell)^{\frac{1}{\eta}} \cdot c_t(j,\ell)^{1-\frac{1}{\eta}} \right)^{\frac{\eta}{\eta-1}} \tag{3}$$

where $\upsilon(j,\ell)$ is the share of sector-j goods in the consumption basket, $\sum_j \upsilon(j,\ell) = 1$, and $c_t(j,\ell)$ is country ℓ's demand for goods from sector j. η is the elasticity of substitution across varieties of goods. Sector-j goods are packaged by an entrepreneur who aggregates the same goods from around the world

$$c_t(j,\ell) = \left(\sum_m \upsilon^*(j,\ell,m)^{\frac{1}{\zeta}} \cdot c_t^*(j,\ell,m)^{1-\frac{1}{\zeta}} \right)^{\frac{\zeta}{\zeta-1}} \tag{4}$$

where $\upsilon^*(j,\ell,m)$ is the share of sector-j goods imported from country m, $\sum_m \upsilon^*(j,\ell,m) = 1$, and $c_t^*(j,\ell,m)$ is country ℓ's final demand for sector-j goods produced in country m. When $\ell \neq m$, goods are imported from country m to country ℓ. ζ is the elasticity of substitution among goods of different origins. Assume that all prices equal to 1. The budget constraint of the agent in country ℓ is

$$\sum_j c_t(j,\ell) + b_t(\ell) = w_t(\ell) \cdot N_t(\ell) + \pi_t(\ell) \tag{5}$$

where $b_t(\ell)$ is the holding of net foreign assets. $w_t(\ell)$ is the nominal wage. $\pi_t(\ell)$ is the transfer payment from the government. Firms produce output using domestic labor and intermediate outputs from all sectors

$$y_t(j,\ell) = A(j,\ell) \cdot n_t(j,\ell)^{1-\alpha(j,\ell)} \cdot \left(\prod_{k,m} z_t(k,m,j,\ell)^{\gamma(k,m,j,\ell)} \right)^{\alpha(j,\ell)} \tag{6}$$

where $A(j,\ell)$ is the level of technology. $n_t(j,\ell)$ is the labor input for sector j. The intermediate output $z_t(k,m,j,\ell)$ are produced by sector k in country m for use in sector j in country ℓ. Similar to the case of consumption, when $\ell \neq m$, the intermediate goods are imported, and when $k \neq j$, the intermediate output

Covid Economics 37, 14 July 2020: 27-56

is produced by a different sector. $\gamma(k, m, j, \ell)$ is the share in intermediate output required for production. $\alpha(j, \ell)$ is the share of intermediate outputs in all inputs of production. The firms' objective is to maximize the profits

$$\max_{n,z} y_t(j, \ell) - w_t(\ell) \cdot n_t(j, \ell) - \sum_{k,m} z_t\left(k, m, j, \ell\right) \tag{7}$$

Under perfect competition, firms set price at the marginal cost of production:

$$1 = \frac{1}{A(j, \ell)} \left(\frac{w_t(\ell)}{1 - \alpha(j, \ell)}\right)^{1-\alpha(j,\ell)} \prod_{k,m} \left(\frac{1}{\alpha(j, \ell) \cdot \gamma(k, m, j, \ell)}\right)^{\alpha(j,\ell)\cdot\gamma(k,m,j,\ell)} \tag{8}$$

The nominal wage is constant and is a function of $A(j, \ell)$. Maximize Eq. (7) subject to Eq. (6). The first-order conditions posits that for any sector-k intermediate output from country m, the marginal product of labor equals the marginal product of any intermediate output:

$$\frac{\partial y_t(j, \ell)}{\partial n_t(j, \ell)} = \frac{\partial y_t(j, \ell)}{\partial z_t(k, m, j, \ell)} \tag{9}$$

The goods market clears when the supply of sector-j goods meets the demand, consisting of intermediate outputs demanded by domestic and foreign firms of all sectors, and consumption goods demanded by domestic and foreign households, less a lump-sum tax collected by the government:

$$y_t(j, \ell) = \sum_{k,m} z\left(j, \ell, k, m\right) + \sum_{m} c_t^*\left(j, m, \ell\right) - \pi_t \tag{10}$$

The clearing condition in the labor market is

$$\sum_{j} n_t(j, \ell) = N_t(\ell) \tag{11}$$

The international asset market clears when $\sum_{\ell} b_t(\ell) = 0$.

Covid Economics 37, 14 July 2020: 27-56

Covid Economics 37, 14 July 2020: 27-56

2.2 Epidemic

We now describe agents' behavior during an epidemic. There are four types of agents in the society, namely, susceptible, s, infected, i, recovered, r, deceased, d. A susceptible agent may contract the disease while interacting with an infected agent in any market j. The probability of getting infected is given by the aggregate risk from all consumption activities:

$$\tau_t(\ell) = \pi_s(\ell) I_t(\ell) \sum_j \phi(j,\ell) \cdot c_t^s(j,\ell) \cdot c_t^i(j,\ell) \tag{12}$$

where $\pi_s(\ell)$ is general risk of infection in country ℓ. $I_t(\ell)$ is the size of the infected population. $\phi(j,\ell)$ measures the degree of interaction in market j. $c_t^s(j,\ell)$ and $c_t^i(j,\ell)$ are sectoral consumption as defined in Eq. (4), with the superscripts indicating the agent types. As in Krueger et al. (2020), the degree of interaction has a mean value 1:

$$\sum_j \upsilon(j,\ell) \cdot \phi(j,\ell) = 1 \tag{13}$$

The newly infected agents in each period come from the group of susceptible people, and is given by:

$$T_t(\ell) = \tau_t(\ell) S_t(\ell) \tag{14}$$

As a result, the population dynamics evolve according to the following equations:

$$S_t(\ell) = S_{t-1}(\ell) - T_{t-1}(\ell) \tag{15}$$

$$I_t(\ell) = I_{t-1}(\ell) + T_{t-1}(\ell) - (\pi_r + \pi_d)\, I_{t-1}(\ell) \tag{16}$$

$$R_t(\ell) = R_{t-1}(\ell) + \pi_r I_{t-1}(\ell) \tag{17}$$

$$D_t(\ell) = D_{t-1}(\ell) + \pi_d I_{t-1}(\ell) \tag{18}$$

At the beginning of the pandemic, we assume that a very small proportion of the population contract the disease from an unknown source. This is represented by the initial state: $I_0(\ell) = \varepsilon$, $S_0(\ell) = 1 - \varepsilon$, $R_0(\ell) = D_0(\ell) = 0$.

Susceptible Lifetime utility of a susceptible household is expressed in the following Bellman equation:

$$U_t^s(\ell) = u\left(C_t^s(\ell), N_t^s(\ell)\right) + \beta\left[\left(1-\tau_t(\ell)\right)U_{t+1}^s(\ell) + \tau_t(\ell)U_{t+1}^i(\ell)\right] \quad (19)$$

In the next period, with a probability $\tau_t(\ell)$, a susceptible agent contracts the disease and becomes an infected agent, while with a probability $1-\tau_t(\ell)$, the agent remains susceptible. Maximize Eq. (19) subject to Eqs. (5) and (12). The first-order conditions are:

$$\upsilon(j,\ell)^{\frac{1}{\eta}} \cdot \frac{1}{C_t^s(\ell)}\left(\frac{C_t^s(\ell)}{c_t^s(j,\ell)}\right)^{\frac{1}{\eta}} - \theta N_t^s(\ell) = \pi_s(\ell)\lambda_{\tau,t}(\ell)\cdot I_t(\ell)\cdot\phi(j,\ell)\cdot\upsilon(j,\ell)\cdot C_t^i(\ell) \quad (20)$$

$$\beta\left(U_{t+1}^i(\ell) - U_{t+1}^s(\ell)\right) + \lambda_{\tau,t}(\ell) = 0 \quad (21)$$

where $\lambda_{\tau,t}$ is the Lagrangian multiplier for the constraint 12.

Infected and recovered Lifetime utility of an infected household

$$U_t^i(\ell) = u\left(C_t^i(\ell), N_t^i(\ell)\right) + \beta\left[\left(1-\pi_r-\pi_d\right)U_{t+1}^i(\ell) + \pi_r U_{t+1}^r(\ell) + \pi_d \times 0\right] \quad (22)$$

With a probability π_r, the agent recovers. The agent may also die with probability π_d. In the event of death, the agent derives zero utility. First-order condition is:

$$\upsilon(j,\ell)^{\frac{1}{\eta}} \frac{1}{C_t^i(\ell)}\left(\frac{C_t^i(\ell)}{c_t^i(j,\ell)}\right)^{1/\eta} = \theta N_t^i \quad (23)$$

Note that the right-hand side of the equation is independent of j. The solutions to Eq. (23) are

$$c_t^i(j,\ell) = \upsilon(j,\ell)\cdot C_t^i(\ell) \quad (24)$$

$$N_t^i(\ell) = \frac{1}{\theta}\cdot\frac{1}{C_t^i(\ell)} \quad (25)$$

Covid Economics 37, 14 July 2020: 27-56

Covid Economics 37, 14 July 2020: 27-56

Lifetime utility of a recovered agent

$$U_t^r(\ell) = u\left(C_t^r(\ell), N_t^r(\ell)\right) + \beta U_{t+1}^r(\ell) \tag{26}$$

It can be shown that the solutions to a recovered agent's problem are the same as an infected agent's. In what follows, we use superscript i to represent a recovered agent's consumption and labor supply.

Equilibrium The market clearing conditions in Eqs. (10) and (11) are rewritten as follows, taking into account the population dynamics:

$$\begin{aligned} y_t(j,\ell) = & \sum_{k,m} z\left(j,k,\ell,m\right) \\ & + \sum_{m} S_t(m) \cdot c_t^{*s}\left(j,m,\ell\right) + \left[I(m) + R(m)\right] \cdot c_t^{*i}\left(j,m,\ell\right) - \pi_t(j,\ell) \\ \sum_j n_t(j,\ell) = & \, S_t(\ell) N_t^s(\ell) + \left(I_t + R_t\right) N_t^i(j,\ell) \end{aligned}$$

By Walras' law, the net foreign assets sum to 0

$$\sum_{\ell} S_t(\ell) b_t(\ell) + \left[I_t(\ell) + R_t(\ell)\right] \left[w_t N_t^i(\ell) + \pi_t(\ell) - \sum_j c_t^i(j,\ell)\right] = 0. \tag{27}$$

2.3 Data and Parameterization

Our main data source is the inter-country input-output table (ICIO) database from the Organisation for Economic Co-operation and Development (OECD). There are 67 economies in the ICIO, including China and eight countries from the ASEAN, each with 36 sectors. The ASEAN countries included in the ICIO are Brunei Darussalam, Indonesia, Cambodia, Malaysia, Philippines, Singapore, Thailand, and Viet Nam. The table describes the input-output linkages between any pair of sectors from the same or different countries. In this paper, we are particularly interested in the linkages among sectors in China and in the ASEAN economies.

Covid Economics 37, 14 July 2020: 27-56

The parameters pertaining to the steady state of the model are calibrated using the ICIO. Since the region experienced both the SARS and the COVID-19, we compare the economic environments prior to both disease outbreaks. For the impact of the SARS episode, we choose 2002 to be the base year, since the SARS outbreak happened in early 2003. For the COVID-19 outbreak, we choose 2015, which is slightly earlier than the beginning of COVID-19. Two reasons explain our choice of base year for the COVID-19 episode. Firstly, according to Luo and Tsang (2020), the later half of the 2010s saw a series of events that disturbed the global trade patterns, particularly the trade war. Data from a slightly earlier year help capture a normal-time scenario prior to the disease outbreak. Secondly, the Chinese economy has undergone a transformation with increasing concentration of services sectors taking over the manufacturing sectors since early 2010s. We find that in year 2015, the distribution of goods and services sectors within China was largely similar to that in the pre-SARS period. As we are interested in the different impacts due to cross-nation production networks, this similar distribution of domestic sectors helps minimize any effects that could arise from a changing domestic economic landscape. In the remainder of the paper, since we calibrate the disease dynamics to be the same, the terms 'SARS' and 'COVID-19' merely refer to the time when they happened.

We reduce the dimensions in the ICIO in this paper. We keep the elements for China, and aggregate those for the ASEAN countries. All elements of the other countries are grouped as the Rest of the World (ROW). The sectors are grouped into two broad sectors, namely the goods and the services. The detailed groupings are summarized in Appendix B. Since the original ICIO tables are expressed in nominal terms, we rebase the values for the pre-SARS period to 2015 dollar by multiplying them with the gross world consumer price inflation between 2002 and 2015. The inflation data is obtained from the World Bank, and the deflator is found to be 1.62. The reduced input-output tables in 2015 price are shown in Section 2.3. Notably, outputs in both China and the ASEAN have expanded faster than the ROW between pre-SARS to pre-COVID-19 periods. We also see that the ASEAN has used more intermediate inputs from China for production.

Two sets of parameters are calculated based on the input-output tables in Sec-

		China		ASEAN		ROW		Final demand			Total use
		Goods	Services	Goods	Services	Goods	Services	China	ASEAN	ROW	
China	Goods	2.018	.468	.015	.002	.158	.049	1.325	.006	.190	4.231
	Services	.479	.379	.003	.002	.033	.023	.836	.005	.087	1.847
ASEAN	Goods	.028	.001	.547	.127	.146	.039	.011	.445	.134	1.477
	Services	.006	.001	.157	.209	.038	.033	.005	.407	.073	.929
ROW	Goods	.239	.006	.160	.021	14.345	5.332	.091	.062	16.279	36.535
	Services	.074	.009	.062	.047	7.616	14.927	.042	.057	31.416	54.250
Value added		1.387	.983	.534	.521	14.199	33.848				51.471
Total inputs		4.231	1.847	1.477	.929	36.535	54.250	2.310	.982	48.179	150.741

Table 1: Three-country two-sector input-output table, pre-SARS (US$ trillion, 2015p).

		China		ASEAN		ROW		Final demand			Total use
		Goods	Services	Goods	Services	Goods	Services	China	ASEAN	ROW	
China	Goods	9.768	1.145	.111	.017	.711	.255	5.886	.044	.848	18.785
	Services	2.348	2.101	.009	.004	.059	.040	3.737	.005	.095	8.398
ASEAN	Goods	.116	.012	1.211	.277	.223	.072	.036	.998	.229	3.174
	Services	.017	.007	.356	.503	.051	.102	.021	.923	.111	2.091
ROW	Goods	.893	.070	.241	.038	18.222	5.807	.301	.104	19.803	45.479
	Services	.155	.064	.083	.100	7.910	19.455	.199	.103	38.259	66.328
Tax less subsidies		.688	.354	.030	.015	.940	.649	.568	.057	2.352	5.653
Value added		4.798	4.645	1.134	1.138	17.362	39.948				69.024
Total inputs		18.785	8.398	3.174	2.091	45.479	66.328	10.748	2.233	61.696	218.933

Table 2: Three-country two-sector input-output table, pre-COVID-19 (US$ trillion, 2015p).

Covid Economics 37, 14 July 2020: 27-56

Covid Economics 37, 14 July 2020: 27-56

tion 2.3, corresponding to the pre-SARS and the pre-COVID-19 periods respectively. The steady-state parameters are summarized in Section 2.3. On the consumption side, Section 2.3 calculates the shares of consumption by source and by sector. These values correspond to parameters $\upsilon^*(j, \ell, m)$ and $\upsilon(j, \ell)$. The shares of goods in China and the ASEAN's consumption baskets have remained similar between the pre-SARS and pre-COVID-19 periods, at 0.62 and 0.53, respectively. On the production side, the labor share parameters, $1 - \alpha(j, \ell)$, are shown in Section 2.3. China experienced a decline in labor share in the goods sector, and a slight increase in the services sector. Whereas, in the ASEAN and the ROW, there have been slight declines in the labor share in their services sectors. In Section 2.3, we show the shares of intermediate inputs $\gamma(k, m, j, \ell)$. Cells for domestic intermediate inputs are shaded for ease of reading. We see that China has used less intermediate inputs from outside the country, with the decline being larger for intermediate inputs from the ROW. For production in the ASEAN, more intermediate inputs were from China, while less were from the ROW.

There is a measure for forward and backward linkages which captures the relationship between a particular sector j and other industries from which it purchases or supplies input to. The derivations are presented in Appendix A. The linkages of China and ASEAN are given in Section 2.3. China services' backward linkages decreased from pre-SARS period to pre-COVID-19 period (0.98 to 0.87), implying increased independence, and is less reliant on ASEAN and the rest of the world upstream. Furthermore, its R^2 value increased from 1.53 to 1.74, implying more uneven trading with the different economies. This supports the observation that China participates more unevenly in trade, with an increasing importance in ASEAN and less with the world. For the same period, ASEAN goods backward linkage, and services forward and backward linkages, all increased from pre-SARS period to pre-COVID-19 period. It became more integrated with other industries, supporting the claim that ASEAN has become more integrated with China in the GVCs.

To calibrate the country-specific labor preference parameter $\theta(\ell)$, we make use

	CHN goods				CHN services			
	Forward Linkage	R^2	Backward Linkage	R^2	Forward Linkage	R^2	Backward Linkage	R^2
Pre-SARS	1.19	1.87	1.23	1.75	1.00	1.50	.98	1.53
Pre-COVID-19	1.18	1.96	1.30	1.78	1.05	1.49	.87	1.74
	ASEAN goods				ASEAN services			
	Forward Linkage	R^2	Backward Linkage	R^2	Forward Linkage	R^2	Backward Linkage	R^2
Pre-SARS	1.08	1.58	1.12	1.50	.93	1.49	.87	1.59
Pre-COVID-19	1.07	1.58	1.13	1.50	.93	1.48	.89	1.59

Table 3: Table of Linkages for China and ASEAN

		Pre-SARS				Pre-COVID-19			
ℓ	m \ j	Goods		Services		Goods		Services	
		$v^*(j,\ell,m)$	$v(j,\ell)$	$v^*(j,\ell,m)$	$v(j,\ell)$	$v^*(j,\ell,m)$	$v(j,\ell)$	$v^*(j,\ell,m)$	$v(j,\ell)$
	China	.929		.947		.946		.944	
China	ASEAN	.008	.618	.006	.382	.006	.611	.005	.389
	ROW	.064		.048		.048		.050	
	China	.010		.011		.039		.005	
ASEAN	ASEAN	.867	.522	.867	.478	.871	.527	.895	.473
	ROW	.123		.122		.091		.100	
	China	.011		.003		.041		.002	
ROW	ASEAN	.008	.345	.002	.655	.011	.352	.003	.648
	ROW	.980		.995		.948		.995	

Table 4: Shares of consumption by source and sector

of the representative agent's first-order condition with respect to labor supply:

$$\theta(\ell) = \frac{1}{C(\ell) \cdot N(\ell)}. \tag{28}$$

The productivity parameters, $A(j,\ell)$'s, are calibrated according to Eq. (8).

We calibrate the parameters for the degree of consumer interactions using the labor market proximity index, following Çakmaklı et al. (2020), shown in Appendix B. The proximity index measures the distance between workers in 36 industries. We adopt this index for household consumption, as Krueger et al. (2020) have shown that infection via the labor market is isomorphic to that in the con-

Covid Economics 37, 14 July 2020: 27-56

	Pre-SARS		Pre-COVID-19	
	Goods	Services	Goods	Services
China	.328	.532	.265	.577
ASEAN	.362	.561	.361	.548
ROW	.389	.624	.390	.608

Table 5: Labor shares in total inputs, $1 - \alpha(j, \ell)$.

k	j / ℓ / m	China Goods	China Services	ASEAN Goods	ASEAN Services	ROW Goods	ROW Services
Pre-SARS							
China	Goods	.710	.542	.016	.005	.007	.002
	Services	.169	.438	.004	.005	.001	.001
ASEAN	Goods	.010	.001	.580	.313	.007	.002
	Services	.002	.001	.166	.512	.002	.002
ROW	Goods	.084	.007	.170	.051	.642	.261
	Services	.026	.010	.066	.115	.341	.732
Pre-COVID-19							
China	Goods	.735	.337	.055	.018	.026	.010
	Services	.177	.618	.005	.004	.002	.002
ASEAN	Goods	.009	.003	.602	.296	.008	.003
	Services	.001	.002	.177	.536	.002	.004
ROW	Goods	.067	.021	.120	.040	.671	.226
	Services	.012	.019	.041	.106	.291	.756

Table 6: Shares of intermediate inputs, $\gamma(k, m, j, \ell)$.

Covid Economics 37, 14 July 2020: 27-56

ℓ	j	$\upsilon(j,\ell)$	δ_j^n	$\phi(j,\ell)$
		Pre-SARS		
China	Goods	.618	.710	.704
	Services	.382	1.491	1.478
ASEAN	Goods	.522	.710	.656
	Services	.478	1.491	1.376
ROW	Goods	.345	.710	.581
	Services	.655	1.491	1.220
		Pre-COVID-19		
China	Goods	.611	.710	.701
	Services	.389	1.491	1.471
ASEAN	Goods	.527	.710	.658
	Services	.473	1.491	1.381
ROW	Goods	.352	.710	.584
	Services	.648	1.491	1.226

Table 7: Table of $\phi(j,\ell)$ values

sumption market. We transform the proximity index so that the numbers are normally distributed with mean 1. Let δ_j be the original proximity index for sector j. The normalized proximity index is calculated as $\delta_j^n = \frac{\delta_j - \mu_\delta + 1}{\sigma_\delta} \sim N(1,1)$, where μ_δ is the mean value of the original proximity index, and σ_δ is the standard deviation. We then calculate the mean normalized proximity index for groups of goods and services sectors, respectively. The country-invariant normalized proximity index of sector j, δ_j^n is used with country ℓ's sectoral market shares $\upsilon(j,\ell)$, to calculate the country sectoral-specific value of $\phi(j,\ell)$ according to

$$\phi(j,\ell) = \frac{\delta_j^n}{\sum_j \upsilon(j,\ell)\cdot\delta_j^n} \tag{29}$$

and presented in Section 2.3. Eq. (13) is hence satisfied.

The calibration of π_r, and π_d follows Eichenbaum et al. (2020). The initial size of the infected population is $\varepsilon = 0.001$. The values of $\pi_s(\ell)$ are such that each country's infection curve is peaked at 1% of the population. Such a specification

Covid Economics 37, 14 July 2020: 27-56

controls the spread of the disease in each economy. Specifically, the values of $\pi_s(\ell)$ are calibrated to be 6.4×10^{-8} and 3.77×10^{-7} for the pre-SARS period, and 9.95×10^{-9} and 2.33×10^{-7} for the pre-COVID-19 period, for China and the ASEAN respectively. The elasticity of substitution parameter, η, is set to be 10. The model is numerically solved in Dynare 4.6.1 using the perfect foresight solver. In what follows, we examine the economic impacts of different scenarios of disease outbreaks.

3 Numerical Results

3.1 Epidemic in a Single Country

An epidemic occurs in the country when a small fraction (ε) of the susceptible population is infected by a disease from an unknown source. The population dynamics following the initial disease outbreak are detailed from Eq. (16) to Eq. (18). Fig. 3 shows the hypothetical population dynamics during a single country epidemic in China and ASEAN at the time of COVID-19. The blue line in Fig. 3 depicts the equilibrium population dynamics in China following an epidemic in China. The green line in Fig. 3 shows the equilibrium population dynamics in ASEAN following an epidemic in ASEAN. We control the population dynamics, so that the infected population peaks at 1% . This is done by adjusting the $\pi_s(\ell)$ values for individual countries.

The slight differences in the population dynamics are due to the different sectoral distributions in China and in the ASEAN seen in Section 2.3. In China, the infected population peaks at around 1% of the population in 47 weeks and thereafter falls with fewer susceptible population. Infected population peaks earlier for ASEAN at 36 weeks. A larger services sector in the ASEAN causes the infected numbers to increase faster, and to decay slower. Percentage of susceptible population declines at a faster rate in the ASEAN when compared to China. For all single country epidemics, the recovery rate edges to around 30% of the population in the long run. For a disease outbreak during the time of SARS, the population dynamics are set to be the same as in Fig. 3.

Covid Economics 37, 14 July 2020: 27-56

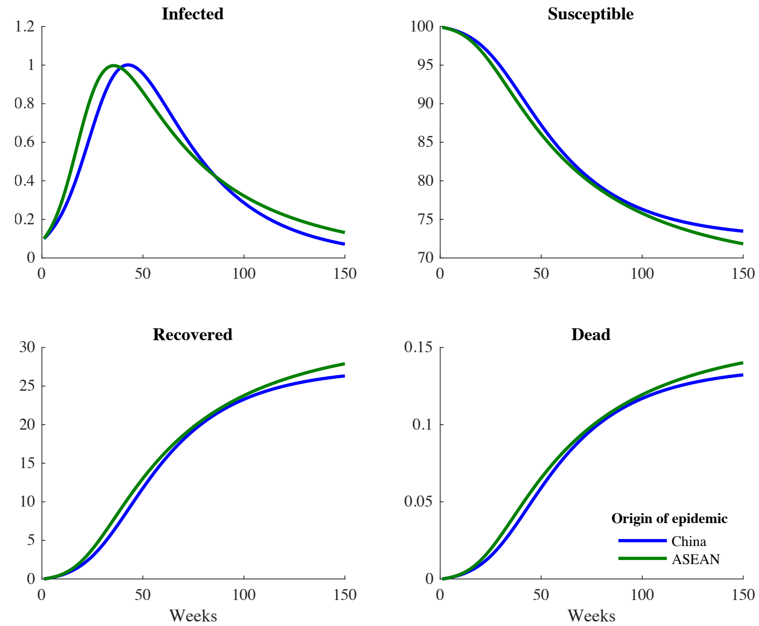


Figure 3: SIRD model population dynamics in China and ASEAN following a single country epidemic.

3.1.1 Within-country Impacts

Our first simulation exercise examines the within-country impact of an epidemic. In Fig. 4, we show China's responses to an epidemic in China (blue lines) and the ASEAN's responses to an epidemic in the ASEAN (green lines). We conduct the simulations for the SARS period (dashed lines) and the COVID-19 period (bold lines).

The immediate effect of the epidemic is seen from the changes in consumption of goods and services. Due to different degrees of contagiousness, consumers choose to consume in the sector that is less infectious in the event of an epidemic. As a result, consumption of goods increased while that of services declined. Because we assume constant prices and nominal exchange rate, the imports for goods and services consumption declined by exactly the same magnitudes as domestic goods and services consumption respectively. The overall effect on consumption is a decline. In general, the ASEAN experienced larger impacts than China at the

Covid Economics 37, 14 July 2020: 27-56

sectoral level, but smaller impacts in aggregate. This result implies that consumers in the ASEAN substitute services consumption with goods consumption more extensively, mitigating the adverse effect at the aggregate level. Comparing between the SARS and the COVID-19 periods, one finds the impacts were greater during the COVID-19 period, and were more distinct in the case of China as seen from the wider gaps between the bold and dashed lines in blue.

Accordingly, a larger domestic demand for the goods sector increases the imports of goods for final consumption. Imports of inputs (both goods and services) to goods sector output also increased. The rise in imports was felt more in ASEAN in both periods when compared to China.

The changes in consumption induced dynamics in sectoral industrial outputs. As expected, the ‘highly infectious’ services sector output declines and the ‘less infectious’ goods sector output rises when faced with an epidemic in both China and ASEAN. The relative rise (decline) in goods (services) sector output is larger in ASEAN than China. It is interesting to note that there is not much difference to ASEAN response to ‘own’ country epidemic (both services and goods sector output) in both periods. However, we see that China’s response to an own country epidemic is larger in the COVID-19 period. In particular, response of China’s goods sector revolved around the steady state in the SARS period. In principle, these equilibrium dynamics are jointly explained by domestic and foreign consumption demand. From a comparison between China’s domestic demand for goods and services, we see negligible change across the two periods. Therefore, foreign demands for consumption and production are likely to account for the different impact of the epidemic in within China.

Similar to the case of consumption, the imports of inputs for goods sector output response is largely muted for China in the SARS period. On the other hand, imports of services for final consumption and imports of inputs for services production decline with the largest dip felt by ASEAN in the COVID-19 period.

The rise in goods sector output of China causes the total exports for final consumption to increase in the COVID-19 period. However, exports for final consumption in ASEAN witnesses a slight dip in COVID-19 as service sector exports decline. Although exports for final consumption rise for China in COVID-19, the

Covid Economics 37, 14 July 2020: 27-56

decline in aggregate consumption leads to the furthest decline in the aggregate output of China in COVID-19.

Two observations can be summarised from this set of results. Firstly, the substitution effects between goods and services are stronger in ASEAN than in China, due to the larger services sector in the ASEAN. Secondly, across time, an epidemic has lead to greater impact on China during COVID-19 than during SARS. In the next set of results, we examine the spillover effect across countries.

3.1.2 Cross-country Impacts

Our second set of results examines the spillover effect of one country's epidemic shock to another. The responses of the macroeconomic variables are shown in Fig. 5. This entails China's response to an epidemic in ASEAN (blue lines) and ASEAN's response to an epidemic in China (green lines). We conduct the simulations for the SARS period (dashed lines) and the COVID-19 period (bold lines).

On account of the epidemic, a lower demand for the 'highly infectious' services and a higher demand for the 'less infectious' goods arise from the 'other' country. Hence, the immediate impacts we observe are a rise in goods sector output and a decline in services sector output.

The epidemics in China have led to some interesting dynamics in the ASEAN. Overall consumption in the ASEAN increased in both the SARS period and COVID-19 period. Two factors have contributed to the rise in consumption; labor hours and net foreign assets. First, we see from Fig. 5 that in both periods, ASEAN's labor hours fell as the decline in service sector output outweigh the rise in goods sector output. Second, we also see that ASEAN experienced a fall in the net foreign assets in both periods. Net foreign assets in our model is equivalent to net exports (both inputs and final goods). Hence, the contraction in ASEAN's net foreign assets can be attributed to the lower services demand from China. The dynamics of labor hours and net foreign assets affect total consumption through the budget constraint in Eq. (5). With wage normalized as one and no change in transfers from the government, the relatively bigger decline in net foreign assets in comparison to labor hours caused total consumption to fall in ASEAN.

Comparing between the SARS and COVID-19 periods, we notice that the im-

Covid Economics 37, 14 July 2020: 27-56

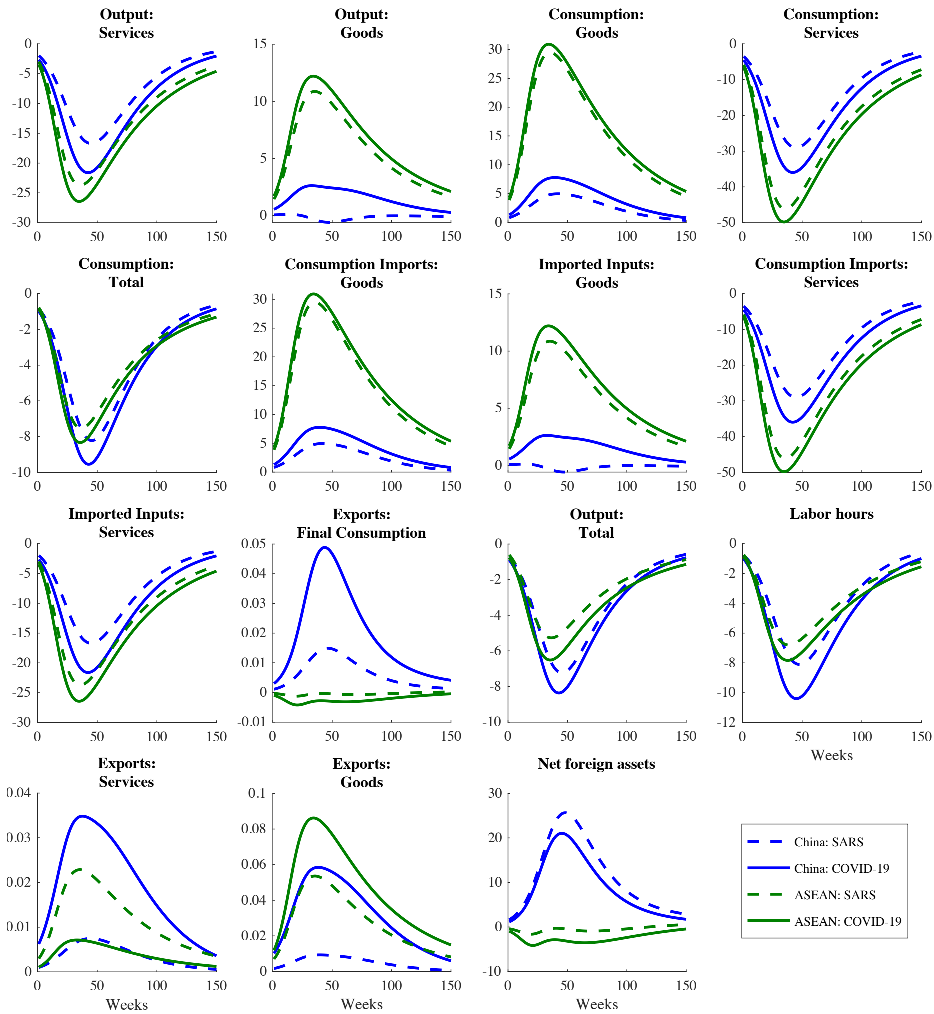


Figure 4: Within-country response to an epidemic.

Note: Green lines indicate ASEAN's response to an epidemic in ASEAN. Blue lines indicate China's response to an epidemic in China. Bold lines refer to COVID-19 period and dashed lines refer to SARS period. X axis represents weeks and Y axis represents percentage deviation from the initial state.

Covid Economics 37, 14 July 2020: 27-56

pacts are in general larger for ASEAN during the later period. Two reasons have led to this outcome. First, China consumed more goods during the COVID-19 period than the SARS period as shown in the third panel of Fig. 4. As China's goods sector accounted for more than 60% of overall consumption, this increase in goods consumption translated into huge increase in the demand for the ASEAN's goods output. Second, the ASEAN's goods sector relies more on China's intermediate inputs during the COVID-19 period than the SARS period. This is seen from Section 2.3 in which the intermediate inputs from China for the ASEAN's goods production increased from 1.6% to 5.5%. As China's goods output increased, so did the supply of intermediate inputs for the ASEAN's production. The increased use of intermediate inputs from China further increased the impacts. The combined effect of demand- and supply- side factors together explain the substantial increase in the ASEAN's goods output. The larger decline in the ASEAN's services sector output, on the other hand, was mainly pulled by China's decreased demand for services. Although the services sector in the ASEAN also used more intermediate inputs from China's goods sector, this supply-side effect is not large enough to offset the drop in demand during the COVID-19 period.

China experienced smaller impacts from disease outbreaks in the ASEAN. The main difference lies in the fact that the increased goods demand from the ASEAN causes China to accumulate net foreign assets. Larger demand for goods sector output also increased the labor hours in China. The relatively larger rise in net foreign assets than labor hours caused aggregate consumption in China to decline through Eq. (5). Across the two periods, China imported less from ASEAN for final consumption, while the ASEAN had imported more goods from China. The increased exports from China to the ASEAN meant that during an epidemic, as the ASEAN increased its consumption of goods, China benefited from increased goods production and total output.

3.2 Response to a pandemic

A pandemic arises when ASEAN and China face an epidemic simultaneously. The path of the population dynamics are similar to Fig. 3. Hence, there are no observable spillover effects of the population dynamics across countries when a

Covid Economics 37, 14 July 2020: 27-56

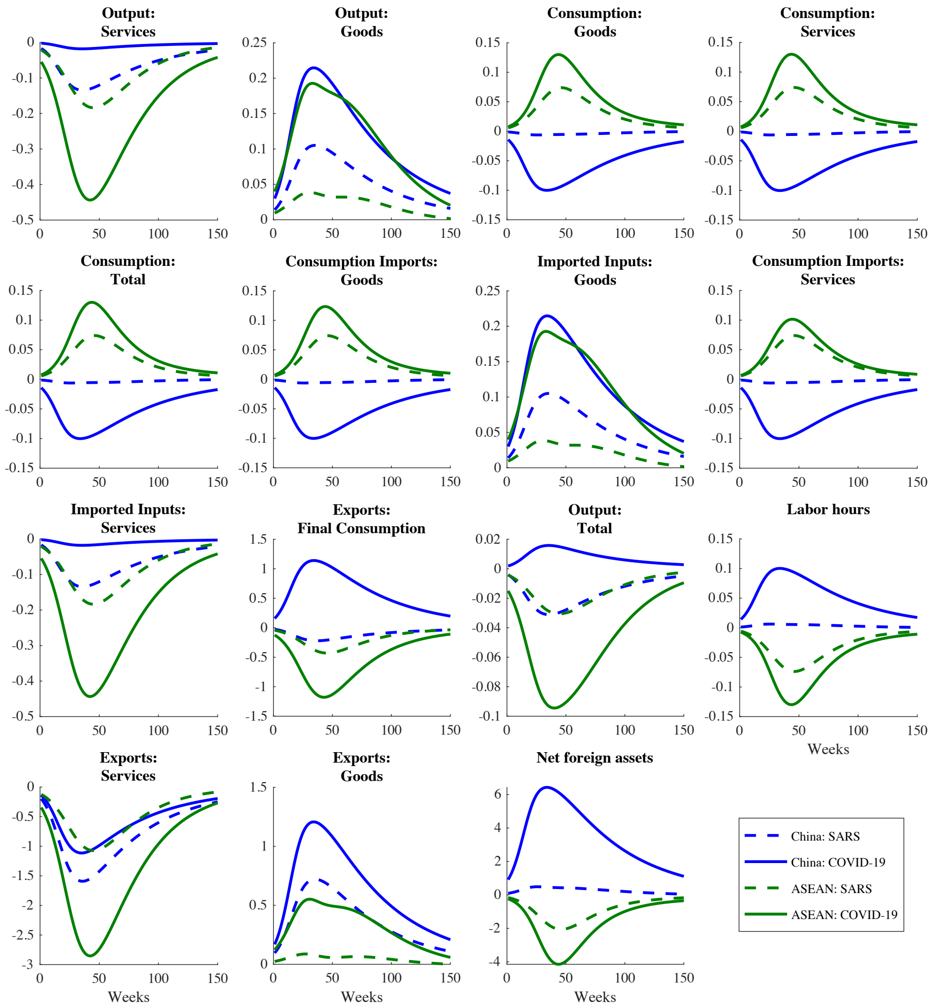


Figure 5: Cross-country response to an epidemic.

Note: Green lines indicate ASEAN's response to an epidemic in China. Blue lines indicate China's response to an epidemic in ASEAN. Bold lines refer to COVID-19 period and dashed lines refer to SARS period. X axis represents weeks and Y axis represents percentage deviation from the initial state.

pandemic occurs. In Fig. 6, we see the response of China (blue lines) and ASEAN (green lines) to a pandemic. The dynamics of Fig. 6 closely follows the dynamics of Fig. 4 for all variables except for exports for final consumption, services exports, goods exports and net foreign assets. This indicates the dominance of own country effects (own country epidemic) over cross country effects (other country epidemic). Exports are primarily driven by the demand from the ”other country”. Hence, we see that cross country effects has a larger impact on exports and net foreign assets. Dynamics of exports for final consumption, services exports and goods exports in Fig. 6 resemble Fig. 5.

In Fig. 6, net foreign assets response of China is very similar in both periods. The amount of rise in China’s net foreign assets at the peak of the epidemic is very large when compared to ASEAN who witnesses a subdued decline in net foreign assets. This signifies the larger role of China in trade linkages in the context of a pandemic.

4 Conclusion

In this paper, we discuss the different economic impacts of a pandemic shock that are associated with evolving economic landscapes and production linkages. We use China and ASEAN in this case study due to the fact that both countries / regions have been hit by both the SARS and the COVID-19. We introduce an SIR framework to a general equilibrium model of production networks. In order to identify the effect of economic landscapes and production linkages, we control the population dynamics so that they are similar in both China and the ASEAN.

Our results support that the greater economic impact of a pandemic over the years is associated with China’s evolving role in the global value chains over the years. Our results for the within-country impact show that China would experience greater impact on aggregate consumption and production in COVID-19 period than in SARS period, while the impact on the ASEAN would be similar. Moreover, our results for cross-country spillover impact show that the ASEAN would experience much greater decline in aggregate output in COVID-19 period than in SARS period, should a pandemic arise in China.

Covid Economics 37, 14 July 2020: 27-56

Covid Economics 37, 14 July 2020: 27-56

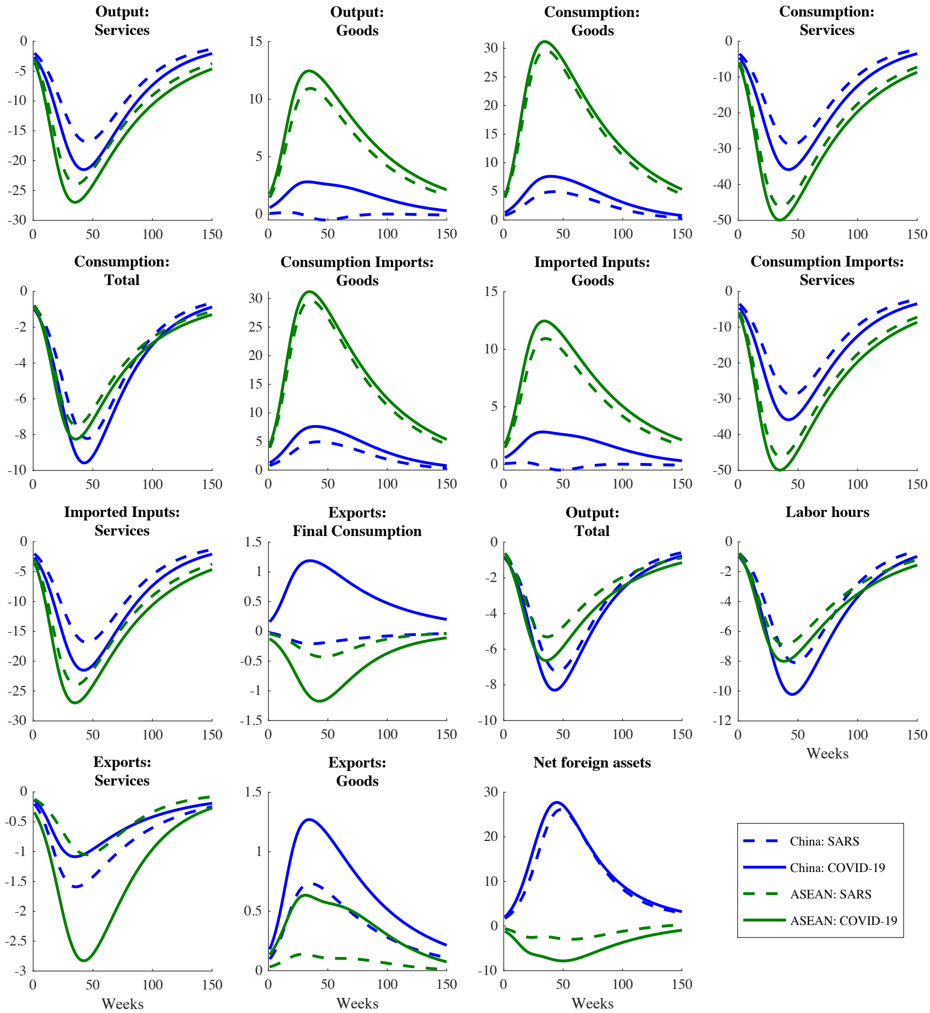


Figure 6: Response to a pandemic.

Note: Green lines indicate ASEAN's response to a pandemic. Blue lines indicate China's response to a pandemic. Bold lines refer to COVID-19 period and dashed lines refer to SARS period. X axis represents weeks and Y axis represents percentage deviation from the initial state. $\pi_s(\ell)$ values remain the same as in the single country epidemic.

It is worth noting a few key limitations of our study which pave way for future research. Firstly, we have ignored frictions in the labor market in this paper. An implication is that workers are free to find jobs in the less infectious sector should they leave the high infectious sector. Our frictionless labor market leads to an under-estimation of the level of economic losses. Nevertheless, we are still able to derive relative impacts of the pandemic between the scenarios discussed. Secondly, we have conducted the analyses without a formal assumption on the policy interventions. Our intention was only to study the propagation channels of pandemic shock within and across countries. However, our sensitivity analysis show that once the infection rate in China has lowered, either due to a change in the virus' characteristics, or a containment effort by the government, the spillover effect to the ASEAN can be reduced accordingly. We therefore encourage future research to focus on the impact on the labor market, as well as the interaction between policy interventions and economic outcome.

Covid Economics 37, 14 July 2020: 27-56

References

Baqaee, D. R. and E. Farhi (2020, May). Nonlinear production networks with an application to the covid-19 crisis. NBER Working Paper No. 27281.

Bonadio, B., Z. Huo, A. A. Levchenko, and N. Pandalai-Nayar (2020, May). Global supply chains in the pandemic. NBER Working Paper No. 27224.

Eichenbaum, M., S. Rebelo, and M. Trabandt (2020, March). The macroeonomics of epidemics. Draft, Northwestern University.

Krueger, D., H. Uhlig, and T. Xie (2020, May). Macroeconomic dynamics and reallocation in an epidemic. *Covid Economics 5*, 21–55.

Liao, J. (2020, February). The rise of the service sector in China. *China Economic Review 59*, 101385.

Luo, S. and K. P. Tsang (2020). China and World Output Impact of the Hubei Lockdown During the Coronavirus Outbreak. *Contemporary Economic Policy*. https://onlinelibrary.wiley.com/doi/pdf/10.1111/coep.12482.

Luo, S., K. P. Tsang, and Z. Yang (2020, April). The impact of stay-at-home orders on usÅ output: A network perspective. Working paper.

Çakmaklı, C., S. Demiralp, □. Kalemli-Özcan, S. Yesiltas, and M. A. Yildirim (2020, May). Covid-19 and emerging markets: An epidemiological multi-sector model for a small open economy with an application to turkey. Working Paper 27191, National Bureau of Economic Research.

Covid Economics 37, 14 July 2020: 27-56

A Forward and Backward Linkages

The forward and backward linkages which capture the relationship between a particular sector j and other industries from which it purchases or supplies input to. In an economy with n industries, the vector of sector linkages, $\mathbf{L}$ can be given as:

$$\mathbf{L} = \frac{n}{|\mathbf{1}^{\intercal}\mathbf{\Lambda}\mathbf{1}|}\mathbf{\Lambda}^{\intercal}\mathbf{1} \tag{A.1}$$

where $\mathbf{\Lambda}$ is the $n \times n$ matrix of coefficients representing the change in output in response to an increased dollar of final demand (backward linkage)/available input for production (forward linkage). $\mathbf{1}$ is a $n \times 1$ column vector of 1s.

The coefficient of variations, R^2 can also be calculated and used to compare how evenly a sector interacts with all other sectors. A larger value of R^2 implies that the sector mainly trades with a few industries, while a smaller R^2 implies that the sector interacts evenly with many industries. The vector of R^2 values, $\mathbf{CV}$ can be calculated as such

$$\mathbf{CV} = \frac{n}{\sqrt{n-1}}(\mathbf{\Lambda}^{\intercal}\mathbf{1})^{-1}(\mathbf{I} \circ (\mathbf{\Lambda}^{\intercal}\mathbf{K}\mathbf{\Lambda}))^{\frac{1}{2}} \tag{A.2}$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{K} = \mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^{\intercal}$ and $\circ$ represents the Hadamard product.

Covid Economics 37, 14 July 2020: 27-56

CEPR PRESS

B Sector Groupings

Table B.1: Sector groupings

ICIO Code	Industry	Proximity Index	Broad sector
D01T03	Agriculture, forestry and fishing	0.86	Goods
D05T06	Mining and extraction of energy producing products	1.08	
D07T08	Mining and quarrying of non-energy producing products	1.06	
D09	Mining support service activities	1.21	
D10T12	Food products, beverages and tobacco	1.12	
D13T15	Textiles, wearing apparel, leather and related products	1.09	
D16	Wood and products of wood and cork	1.03	
D17T18	Paper products and printing	1.08	
D19	Coke and refined petroleum products	1.11	
D20T21	Chemicals and pharmaceutical products	1.06	
D22	Rubber and plastic products	1.10	
D23	Other non-metallic mineral products	1.08	
D24	Basic metals	1.09	
D25	Fabricated metal products	1.08	
D26	Computer, electronic and optical products	1.03	
D27	Electrical equipment	1.07	
D28	Machinery and equipment, nec	1.06	
D29	Motor vehicles, trailers and semi-trailers	1.09	
D30	Other transport equipment	1.06	
D31T33	Other manufacturing; repair and installation of machinery and equipment	1.07	
D35T39	Electricity, gas, water supply, sewerage, waste and remediation services	1.08	Services
D41T43	Construction	1.21	
D45T47	Wholesale and retail trade; repair of motor vehicles	1.13	
D49T53	Transportation and storage	1.18	
D55T56	Accommodation and food services	1.26	
D58T60	Publishing, audiovisual and broadcasting activities	1.11	
D61	Telecommunications	1.07	
D62T63	IT and other information services	1.01	
D64T66	Financial and insurance activities	1.02	
D68	Real estate activities	1.10	
D69T82	Other business sector services	1.09	
D84	Public admin. and defence; compulsory social security	1.16	
D85	Education	1.22	
D86T88	Human health and social work	1.28	
D90T96	Arts, entertainment, recreation and other service activities	1.18	
D97T98	Private households with employed persons	—	

Covid Economics 37, 14 July 2020: 27-56

CEPR PRESS

Covid Economics Issue 37, 14 July 2020

Change points in the spread of COVID-19 question the effectiveness of nonpharmaceutical interventions in Germany[1]

Thomas Wieland[2]

Date submitted: 6 July 2020; Date accepted: 8 July 2020

Covid Economics 37, 14 July 2020: 57-72

Nonpharmaceutical interventions against the spread of SARS-CoV-2 in Germany included the cancellation of mass events (from March 8), closures of schools and child day care facilities (from March 16) as well as a "lockdown" (from March 23). This study attempts to assess the effectiveness of these interventions in terms of revealing their impact on infections over time. Dates of infections were estimated from official German case data by incorporating the incubation period and an empirical reporting delay. Exponential growth models for infections and reproduction numbers were estimated and investigated with respect to change points in the time series. A significant decline of daily and cumulative infections as well as reproduction numbers is found at March 8, March 10 and March 3, respectively. Further declines and stabilizations are found in the end of March. There is also a change point in new infections at April 19, but daily infections still show a negative growth. From March 19, the reproduction numbers fluctuate on a level below one. The decline of infections in early March 2020 can be attributed to relatively small interventions and voluntary behavioural changes. Additional effects of later interventions cannot be detected clearly. Liberalizations of measures did not induce a re-increase of infections. Thus, the effectiveness of most German interventions remains questionable. Moreover, assessing of interventions is impeded by the estimation of true infection dates and the influence of test volume.

1 The author would like to thank Sarah Landmann for moral support and helpful discussions as well as Jefferson Burrowes for proofreading.

2 Karlsruhe Institute of Technology, Institute of Geography and Geoecology.

1 Background

Assessing the effectiveness of nonpharmaceutical interventions (NPIs) in the SARS-CoV-2/COVID-19 context is a topic of growing relevance. Nevertheless, findings documenting the impact of these measures have not been homogeneous within the literature; whether with respect to single countries [1-11], or in terms of international comparisons [12-17]. The question of whether “lockdowns” – including contact bans, curfews or closures of schools and child day care facilities – succeed or fail in reducing infections is a key concern for policymakers, as such measures are accompanied by consequences in terms of economic, social and psychological effects on societies. All European countries introduced NPIs to reduce infections, ranging from appeals to voluntary behaviour changes and the cancellation of mass events (Sweden) to strict curfews (e.g. France, Italy). Being one of the most affected countries (in terms of confirmed prevalence), Germany introduced a strict strategy incorporating three bundles of measures (1. cancellation of mass events after March 8, 2. closure of schools and child day care facilities between March 16 and 18, and 3. a contact ban, bans of gatherings and closures of “nonessential” services from March 23).

There have been some approaches to assessing the interventions in Germany: Dehning et al. [1] utilized epidemiological models (the SIR [susceptible-infected-recovered] model and its extensions) combined with Bayesian inference to find change points in infections over time with respect to the aforementioned measures. They identified impacts of all three bundles of interventions and on this basis have explicitly outlined the importance and necessity of the contact ban for reducing new infections. In a series of studies [2-5], German economists investigated structural breaks in time series of cumulated infections and growth rates. Their inferred change points have been interpreted in a similar way, i.e., in support of the measures. An additional modelling approach using a modified SIR model [4] also outlines the impact of NPIs on infections.

The common denominator in the approaches mentioned above [1-5] is the application of disease case data from the Johns Hopkins University (JHU). This data differs from the official German case data provided by the Robert Koch Institute (RKI) in terms of both precision and detail, with importantly, the latter dataset including information about the date of onset of symptoms for most cases [18,19]. This information is essential because it helps to estimate the true infection dates. In the aforementioned studies [1-5], information of this type was not available, which has therefore required assumptions to be formulated regarding the time between infection and reporting. The SIR modeling study [1] has already been criticized in terms of its underestimation of this delay and the related results [20]. Moreover, studies utilizing epidemiological models [1,4,6] require assumptions on the transmission process of the disease (e.g., spreading rate, contacts per capita) or other unknown epidemiological parameters. Both aspects raise the question whether the previous assessments of NPIs in Germany are reliable.

2 Aims

The aim of this study is to assess the effectiveness of NPIs towards the SARS-CoV-2 spread in Germany (from March 8, 16 and 23, respectively), while overcoming the data-related problems mentioned above. The measures are analysed in terms of revealing their impact on infections over time. By using official case data [19], true dates of infection are estimated. Inspired by the methodical approach in previous studies [2-4,11], change points in time series of three indicators (daily and cumulative infections as well as reproduction numbers, all of

Covid Economics 37, 14 July 2020: 57-72

which were calculated based on the estimated infection dates) were detected. The data covers infections from February 15 to May 31, 2020, which means that also possible effects of the easing of measures (from April 20) and the introduction of face masks (from April 27) can be assessed.

3 Estimating the dates of infection

To assess the effectiveness of NPIs, it is the dates of infections of the reported cases which must be regarded, rather than the date of report. However, the real time of infection is unknown, thus, it must be estimated using the reported cases. In simple terms, the time between infection and reporting consists of two time periods: a) the time between infection and onset of symptoms (incubation period), and b) the time between onset of symptoms and the date of report (reporting delay). Thus, to estimate the date of infection, both periods must be subtracted from the date of report [1-5,8,11].

There are several estimations of the SARS-CoV-2/COVID-19 incubation period, ranging from median values of 5.0 to 6.4 days [21,22]. Incorporating the reporting delay, however, is much more difficult. Previous studies investigating the effectiveness of interventions in Germany [1-5] have employed data from the Johns Hopkins University (JHU) which only includes daily infection and death cases. The reporting delay is either assumed to be equal to 2-3 days [2-4] or estimated in the model parametrization [1,5]. In contrast, the data on German cases from the Robert Koch Institute (RKI) includes the reporting date and, for the majority of cases, case-specific dates of onset of symptoms, socio-demographic information (age group, gender), and the corresponding county [18,19]. The data used here is the RKI dataset from June 28, 2020 [19]. In this dataset, there were 193,467 reported infections, for which, the date of onset of symptoms is known in 135,967 cases (70.28 %). The arithmetic mean of the time between onset of symptoms and report (reporting delay) is equal to 6.71 days (SD = 6.19) and the corresponding median equals 5 days. 95 % of the reporting delays lie between 0 (2.5 % percentile) and 21 (97.5 percentile) days. On this basis, we clearly see that assuming this value to be equal to 2-3 days [2-4] is an obvious underestimation. Moreover, exploring the dataset reveals that the reporting delay varies between the age and gender groups of the reported cases and over time, as well as between German counties. These differences indicate that it is difficult to assume or estimate average values for the reporting delay [1,5].

Thus, the estimation of the true infection dates of reported cases was conducted using the information from the RKI case data. In line with previous studies [1-5], the incubation period is assumed to equal 5 days, which is the minimum value reported in the literature [21,22]. Given this time period for the records in the case dataset with known date of onset of symptoms, the date of infection of case i, DI_i, is calculated as the date of onset of symptoms (DO_i) subtracted by the incubation period (IP):

$$DI_i = DO_i - IP$$

Based on the cases with full information, a dummy variable regression model was estimated for the interpolation of the reporting delay for the remaining 57,500 cases. As the reporting delay differs across case-specific attributes, the reporting delay for case i ($RD_{i,agcwt}$) was estimated by including dummy variables for age group a (a = 1, …, A), gender group g (g = 1, …, G), county c (c = 1, …, C) and weekday w (w = 1, …, W) as well as the time trend t:

$$RD_{i,agcwt} = \alpha + \sum_{a}^{A-1} \beta_a D_{agegroup_a} + \sum_{g}^{G-1} \gamma_g D_{gender_g} + \sum_{c}^{C-1} \delta_c D_{county_c} + \sum_{w}^{W-1} \zeta_w D_{weekday_w} + \varphi\, t + \varepsilon_{i,agcwt}$$

where α is the estimated constant, β_a, γ_g, δ_c and ζ_w represent sets of empirically estimated parameters for the *A-1* age groups, *G-1* gender groups, *C-1* counties and *W-1* weekdays, φ is the empirically estimated parameter for the time trend and $\varepsilon_{i,agcwt}$ is the stochastic disturbance term. The model parametrization was conducted via Ordinary Least Squares (OLS) estimation.

In those cases lacking the information on onset of symptoms, the date of infection was calculated as the date of report (DR_i) subtracted by the estimated reporting delay and the incubation period:

$$DI_i = DR_i - RD_{i,agcwt} - IP$$

4 Infection indicators and detection of change points over time

Previous studies with respect to the assessment of interventions have focused on only one indicator such as daily infections [1], cumulative infections [2-5,8,11,15], or reproduction numbers [7,17]. To arrive at a more holistic picture, three indicators are used: a) the daily new infections, b) cumulative infections and c) the daily reproduction numbers. The estimated infections dates (DI_i) were summarized over days which results in the daily new infections at time *t* (I^D_t) and the corresponding cumulative infections at time *t* (I^C_t). The reproduction number for time *t* (R_t) was computed according to the calculation provided by the Robert Koch Institute [18] as the quotient of infections in two succeeding 4-day intervals (implying a generation period of 4 days):

$$\mathrm{R_t} = \frac{\sum_{t}^{t-3} I_t^D}{\sum_{t-4}^{t-7} I_t^D}$$

The period under study includes the infections from February 15 (first proven “super spreading event” in Germany, the “Kappensitzung” in Gangelt, North Rhine Westphalia) to May 31, resulting in N = 107 daily observations. The final date is estimated by the last available date of report (June 27) subtracted by the 97.5 % percentile of the incubation period (5.6 days) and the 97.5 % percentile of the reporting delay (21 days).

For the analysis of infections over time, phenomenological models have the advantage that they only incorporate time series of infections and do not require further assumptions concerning the transmission process of the disease under study [23,24]. Thus, the time series of all three indicators were analysed using exponential growth models in their semilog form, which means that the dependent variables (I^D_t, I^C_t and R_t) were transformed via natural logarithm. The model parametrization was conducted via Ordinary Least Squares (OLS) estimation. The corresponding slope parameter of the independent variable (time), here denoted as λ, represent the average growth rate per time unit (days) and λ*100 equals the percentage change per day:

$$\ln(I_t^D) = \alpha^D + \lambda^D\, t + \mu_t^D$$

$$\ln(I_t^C) = \alpha^C + \lambda^C\, t + \mu_t^C$$

$$\ln(R_t) = \alpha^R + \lambda^R\, t + \mu_t^R$$

where α^D, α^C, α^R, λ^D, λ^C and λ^R are the parameters to be estimated and μ^D_t, μ^C_t and μ^R_t represent the stochastic disturbance term in each model.

The detection and dating of change points was conducted using a fluctuation test (recursive estimation test) and *F* statistics, which incorporates comparing the regression coefficients of a time series with *M* breakpoints (and, thus, *M*+1 segments) to the full sample estimates (no segmentation). Within these tests, structural breaks in the time series can be identified. The optimal number of breakpoints and their attribution to the specific observation at which point they occur (which means a dating of the breakpoint, including the computation of confidence intervals) was conducted using the Bai-Perron algorithm. The statistically optimal number of *M* breakpoints is inferred by comparing model variants with zero to five breakpoints (corresponding to one to six segments). The variant which minimizes the residual sum of squares (RSS) and the Bayesian information criterion (BIC) is considered to be the optimal solution [25,26]. Thus, the exponential growth functions shown above are divided into *M*+1 segments, in which the regression coefficients in each *m* segment (*m* = 1, …, *M*+1) are constant. The analysis was conducted in *R* [27] using the package *strucchange* [26].

5 Results

Fig. 1 shows the daily reported cases in the RKI dataset, the corresponding daily onsets of symptoms (incorporating the reporting delay) and the daily infections (incorporating the reporting delay and the incubation period) from February 15 to May 31, 2020. Fig. 2 presents the estimated infections and reported cases on the level of calendar weeks along with additional information about the number of conducted SARS-CoV-2 tests [28]. Obviously, the time series are not simply shifted by the average delay between infection and report. The differences between the temporal development of infection and report curves can be attributed to temporal, case-specific, and regional differences in the reporting delay. Furthermore, all results emerging from time series of infections shown below have to be interpreted whilst taking into consideration the changing number of tests conducted weekly. Specifically, we can see an increase in the number of tests by a factor of 2.73 from calendar week 11 (127,457 tests) to 12 (348,619 tests), followed by smaller fluctuations in the succeeding weeks.

Fig. 3, 4 and 5 show the results of the time series analysis of daily infections, cumulative infections, and reproduction numbers, respectively. The top-left plot shows the optimal structural breaks in time series and the corresponding slopes (exponential growth rates) for each model segment. The top-right plot displays the explained variance (R^2) and the point estimate confidence intervals for each model segment. The bottom-left plot presents the corresponding model diagnostics (BIC and RSS) for the model variants with one to five segments, and the adjacent plot shows the model fit on condition that no structural breaks occur. With respect to daily infections (fig. 3), the best model fit minimizing BIC and RSS incorporates three breakpoints and four model segments, respectively. Obviously, a model without breakpoints does not fit the time series appropriately. The three significant structural breaks are on March 8 (95 % confidence intervals: March 7 to March 9) and March 24 (CI [23, 25]) as well as April 19 (CI [18, 20]). The first break on March 8 reduces the growth rate from 0.229 (CI [0.217, 0.240]), which represents an average daily increase of 22.9 % (February 15 to March 8), to -0.013 (CI [-0.025, -0.001]), which means a daily decrease equal to 1.3 % (March 9 to March 24). From March 25, the daily infections decrease by 5.4 % per day (-0.054, CI [-0.057, -0.050]) until April 19. From April 20, the decline of new infections slows down but the daily growth rate is still negative with -3.0 % (-0.030, CI [-0.033, -0.028]).

Covid Economics 37, 14 July 2020: 57-72

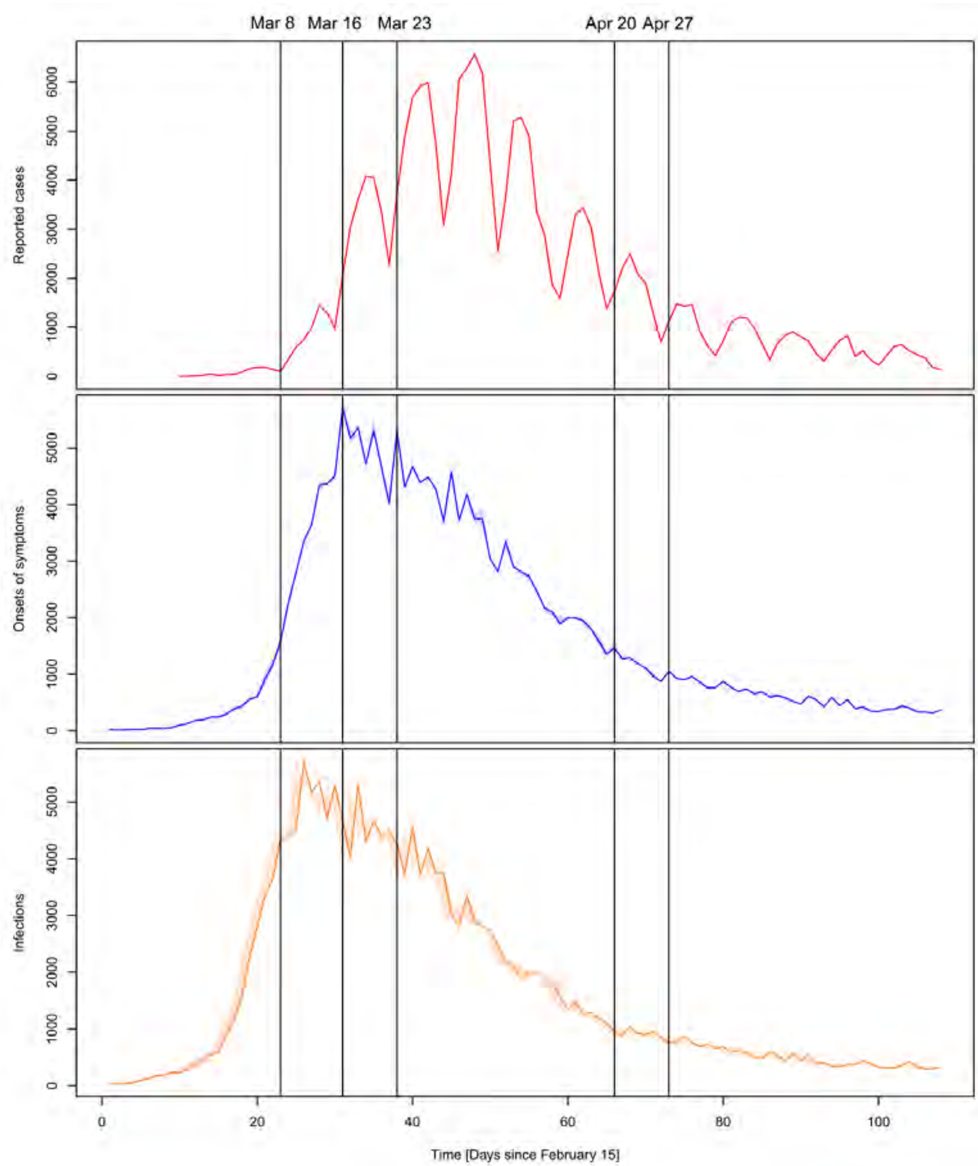


Figure 1: Daily values of reported cases, onsets of symptoms and infections from 15 February to 31 May 2020.

Covid Economics 37, 14 July 2020: 57-72

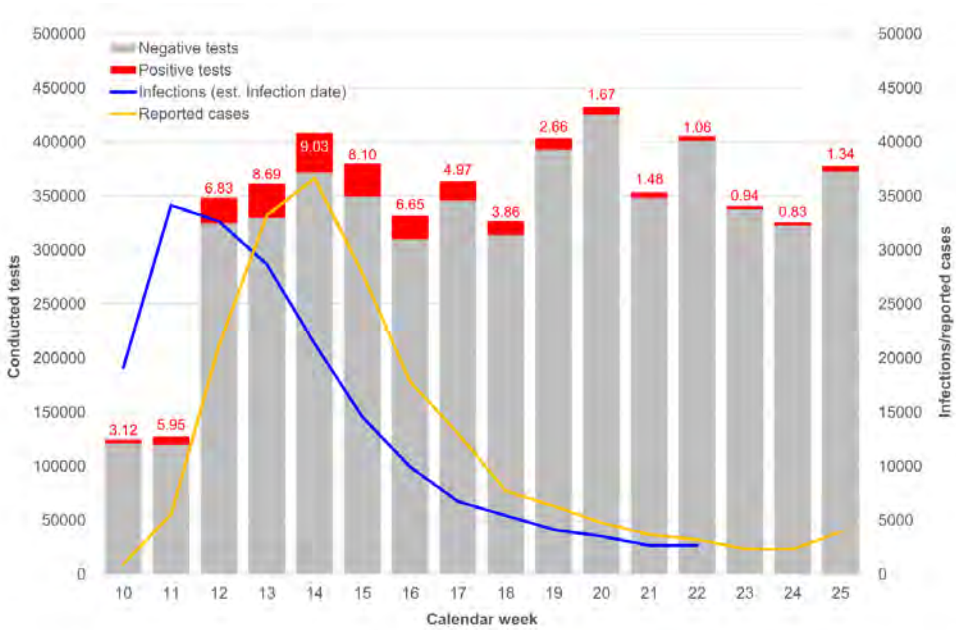


Figure 2: Weekly values of reported cases, infections and conducted SARS-CoV-2 tests from calendar week 10 to 25.

The best model solution for the cumulative infections over time also incorporates three breakpoints. The first break occurs on March 10 (CI [9, 11]), at which point the daily growth rate was reduced from 22.8 % (0.228, CI [0.224, 0.232]) to 6.6 % (0.066, CI [0.059, 0.073]). The second break on March 26 (CI [25, 27]) documents a further decrease in daily growth from 6.8 % to 1.9 % (0.019, CI [0.017, 0.020]). The last structural change is detected on April 13 (CI [12, 14]), at which time the daily growth rate shifted from 1.9 % to 0.4 % (0.004, CI [0.003, 0.004]).

With respect to the reproduction number (R), three structural breaks can also be identified. After the first break on March 3 (CI [2, 4]), R starts to decrease by 9.7 % per day (-0.097, CI [-0.107, -0.087]) until March 19 (CI [18, 20]). The break around March 19 initiates a stabilization of the R value with a decrease equal to 0.7 % per day (-0.007, CI [-0,008, -0,005]). From the last change point which occurs at April 23 (CI [22, 27]), the reproduction number still fluctuates on a low level with a daily increase of 0.3 % (0.003, CI [0.000, 0.005]). With few exceptions, from March 19, the daily reproduction number remains below one (ln R < 0).

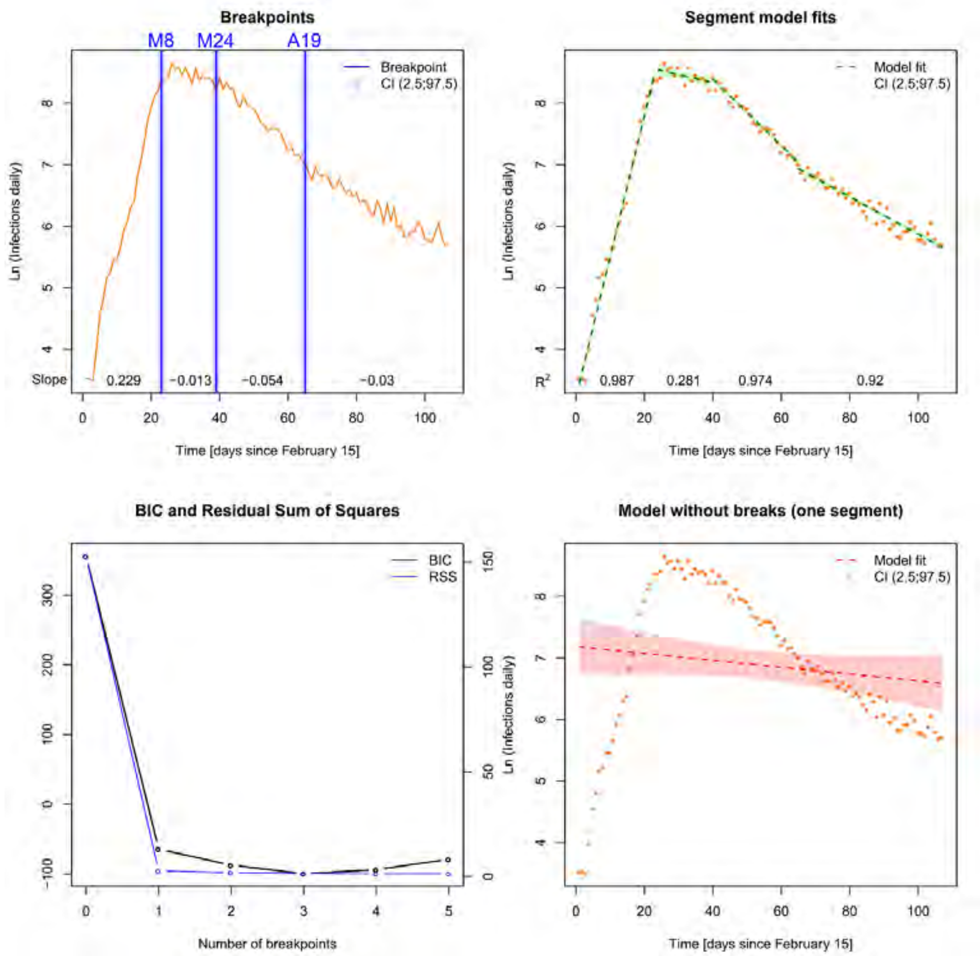


Figure 3: Time series and corresponding break points as well as model diagnostics for daily infections from 15 February to 31 May 2020.

Covid Economics 37, 14 July 2020: 57-72

Covid Economics 37, 14 July 2020: 57-72

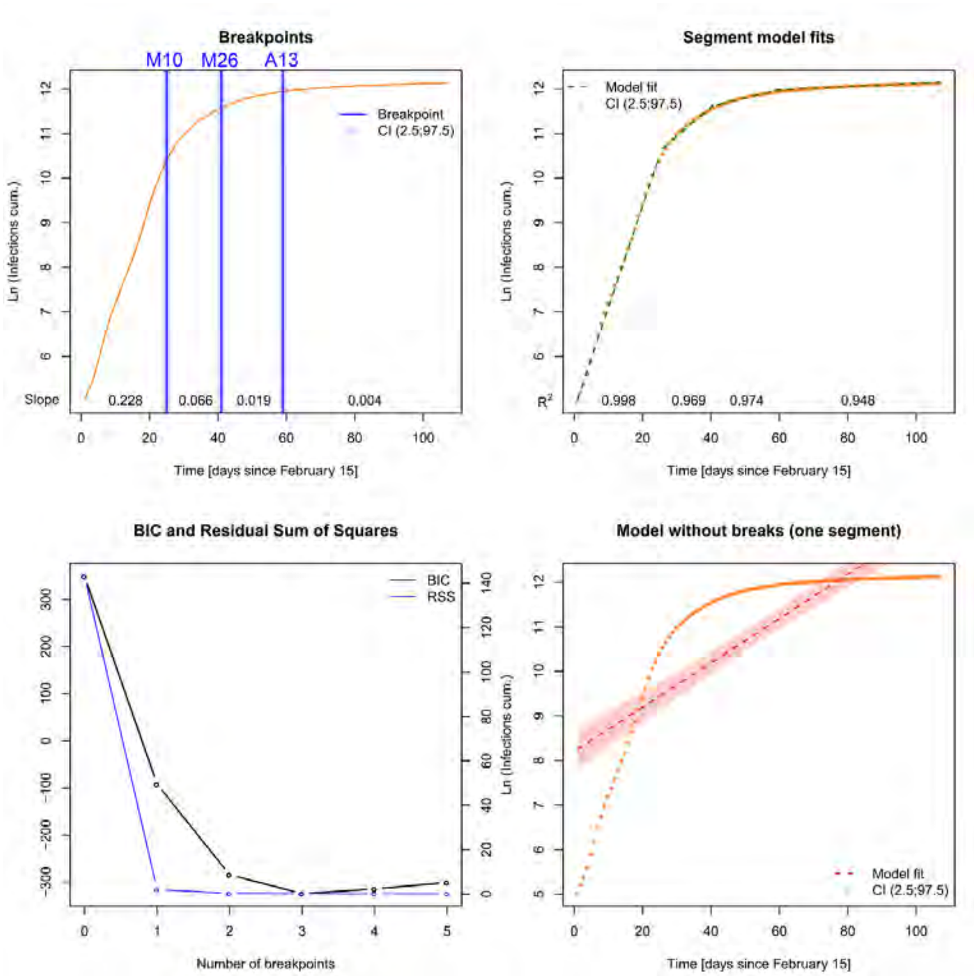


Figure 4: Time series and corresponding break points as well as model diagnostics for cumulative infections from 15 February to 31 May 2020.

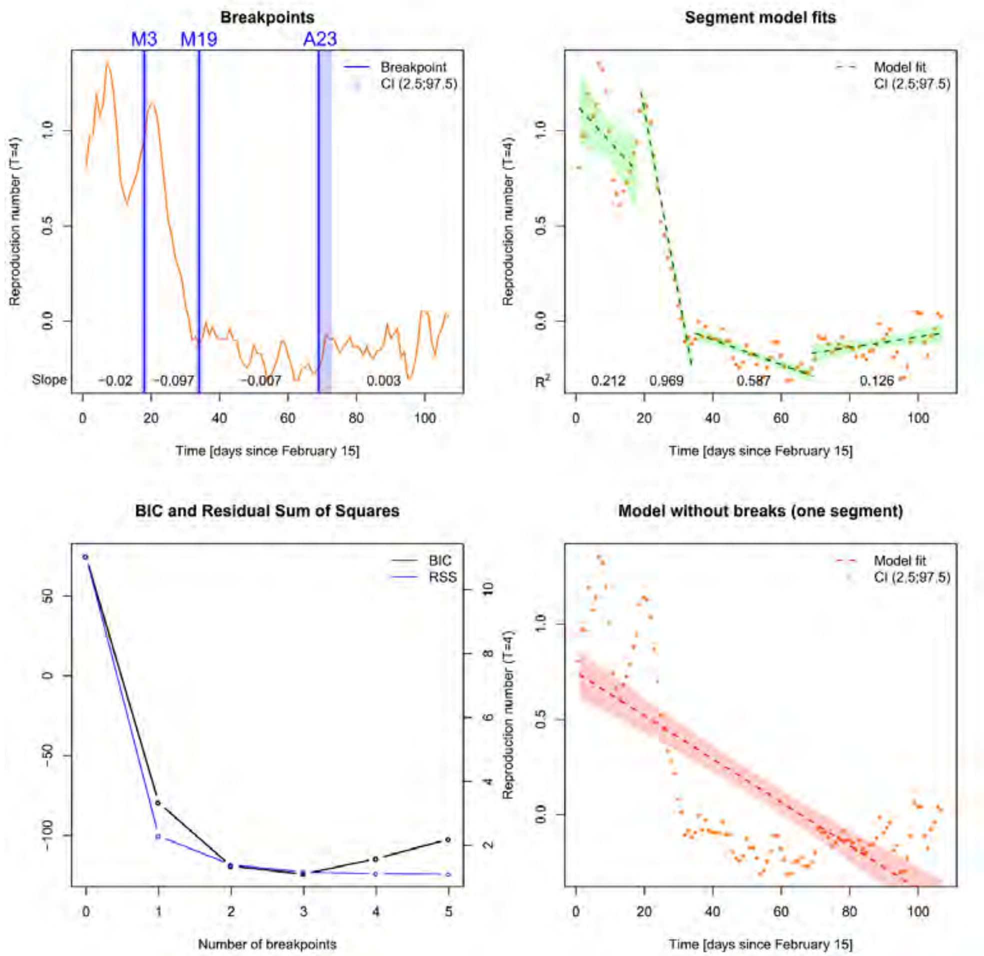


Figure 5: Time series and corresponding break points as well as model diagnostics for reproduction numbers (R) from 15 February to 31 May 2020.

All in all, we find concordant structural breaks for all three indicators in the first third of March 2020. Around March 8, the daily new infections turn from exponential growth to decay and the growth rate of cumulative infections has its highest decrease. This decline occurs although the test volume increased strongly in the succeeding weeks (see fig. 2). Unfortunately, conducted tests cannot be linked to reported cases as both information stem from different data sources. However, the massive increase of testing must have had an influence on the detection of SARS-CoV-2 infections occurred before. It is therefore plausible to assume that if test volume had remained constant over time, fewer infections would have been detected and the decrease of (confirmed) infections would have been stronger. The other breakpoints are not coincident: Whilst structural changes in the daily and cumulative infections occur in the last third of March, there is no corresponding break with respect to reproduction numbers. In the last third of April, we find structural changes with respect to daily infections and reproduction numbers, but the growth rate of infections still remains negative. From March 19, the reproduction numbers, with few exceptions, fluctuate on a level below one (ln R < 0).

6 Discussion

Regarding all three indicators, we find consistent results with respect to a significant decline of infections in the first third of March – about one week before the closing of schools and child day care and two weeks before the full "lockdown" (including the contact ban) came into force. The effect coincides with the cancellation of mass events recommended by the German minister of health, Spahn, on March 8. However, the increased awareness in the general population could have also had a significant impact in terms of voluntary changes in daily behaviour (e.g., physical distancing to strangers, careful coughing and sneezing, thorough and frequent hand washing). Surveys demonstrate an increased awareness towards the Corona threat already in the middle of February [29]. Additionally, voluntary cautious behaviour in the Corona context could also explain the abrupt and unusual decline of *other* infectious respiratory diseases in Germany starting in early March [30].

Previous studies have also found a first slowing of infections in the first third of march [1-5], but the results of the present analysis contradict their findings as the change point in the 10th calendar week is a) the clearest structural break given that it is present for all three indicators, b) the break which initiated a trend change in terms of a decline of daily new infections and c) the most influential break with respect to cumulative infections. Dehning et al. [1] state: "Our results indicate that the full extent of interventions was necessary to stop exponential growth […] Only with the third intervention, the contact ban, we found that the epidemic changed from growth to decay". These statements are based on a negative growth rate (-3 %) having not become apparent before the contact ban came into force. In contrast, given the estimated infection dates in the present study, we see that the growth rates of new infections and reproduction numbers already turn negative on March 3 and 8, respectively. At the same time, the growth rate of the cumulative infections has its biggest decrease across all four segments of the time series. Thus, a decline in infections occurred before school closures and the contact ban came into force.

In the time series studies on the German case [2-5], the closing of infrastructures (schools etc.) in mid-March was found to be the most influential break with respect to cumulative infections. This conclusion cannot be confirmed in the present study, as we cannot find any referring breakpoint with respect to the daily and cumulative infections. If the closures of schools and child day care facilities would have had an impact on infections, there would have been a significant decline of new infections from March 16 to 18 on. The structural break in the reproduction numbers on March 19 initiates a stabilization of the reproduction numbers but not a further decline. Therefore, an impact of school and child day care closures cannot be detected. The influence of the third intervention ("lockdown" including contact ban), which was found to be the most influent factor in the SIR modelling study [1], and an important factor in the previous time series analyses [2-5], remains unclear in the present study as well. There is no structural break in the reproduction numbers which coincides with the contact ban. Significant breaks in daily and cumulative infections occur after the social ban came into force, but not immediately. The mismatches between the present and previous results are obviously related to different data sources, a point underscored by the fact that the modelling approach is similar to some of the previous studies [2-4].

Covid Economics 37, 14 July 2020: 57-72

The impact of first liberalizations of measures from April 20 (e.g., reopening of some “nonessential” retail shops) is plausibly reflected in the temporal development of new infections and reproduction numbers. However, there is no re-increase of new infections as the corresponding growth rate remains negative and the reproduction numbers remain, with few exceptions, below the critical value of one. Moreover, no effect of the implementation of compulsory face masks in retail shops and public transport (starting from April 27) can be detected, as there is no further significant structural break. However, this intervention was implemented at a time where infections were already on a low level. Thus, the effectiveness of this measure cannot be definitely assessed. Further liberalizations starting in the first half of May (e.g., reopening of schools for some age groups, extending emergency childcare) do not show any impact as well.

The current findings support results for Germany inferred from logistic growth models which show a trend change before the contact ban came into force [8]. In addition, a Spanish time series study revealed breakpoints in cumulative infections, with the first occurring about two weeks before the nationwide “lockdown” [11]. Furthermore, the present results tend to support other studies of international comparisons which have found a decline of infections with or without strict interventions [12-16].

Covid Economics 37, 14 July 2020: 57-72

7 Strengths and limitations

One strength of the present study is the relative simplicity of the analysis. The current approach allows for a time-related analysis of NPIs based on a rather simple model which does not require further assumptions concerning the disease under study. Thus, the methodology can be easily transferred to other pandemics, countries, or regions as only time series of infections are necessary. In the future, the research design should be applied to international comparisons, incorporating both Scandinavian and South-European countries. Another strength is the utilization of realistic infection dates, which was not incorporated in previous studies. Moreover, regarding three different indicators allows for a more differentiated picture of infections over time.

The temporal development of the three indicators was also contrasted with conducted tests over time. However, in the absence of daily test data, the impact of changing test volumes was not assessed directly. Another limitation results from the phenomenological nature of the regression models utilized for time series analysis. As the only explanatory variable is time, we can question the impacts of the regarded interventions but cannot explain the factors causing the temporal development of infections directly.

8 Conclusions

This study finds clear evidence of a decline of SARS-CoV-2 infections in Germany at the beginning of March 2020, which can be attributed to relatively small nonpharmaceutical interventions (cancellation of mass events) and voluntary behavioral changes. A trend change of infections from exponential growth to decay was not induced by the “lockdown” measures but occurred earlier. Additional impacts of later NPIs cannot be clearly detected: Firstly, there is no significant effect with respect to infections that could be attributed to school and day-care closures. Secondly, effects which could be related to the contact ban a) do not appear with respect to all three indicators, b) differ in strength and tend towards lower impacts, and c) do not match the time the measure came into force. Thus, the necessity of the second (March 16-18) and the third bundle of interventions (March 23) is questionable because a) the related effects on infections (if any) cannot be unequivocally validated, b) a trend change had already occurred long before they came into force, and c) liberalizations of these measures did not induce a re-increase of infections. We cannot deduce conclusions towards the necessity of compulsory face masks, as this intervention was introduced late. Furthermore, the time series of (confirmed) infections is substantially influenced by temporal changes in the test volume, which leads to a high degree of uncertainty with respect to the data source. Therefore, a future evaluation of NPIs towards SARS-CoV-2/COVID-19 in Germany should consider these questionable effects and uncertainties.

The study reveals three methodological issues for assessing the impact of NPIs which may influence the results enormously. Firstly, the key challenge is the estimation of realistic infections dates from official statistics (which typically do not include this information). This information is essential for the assessment of measures which aim at the reduction of new infections. It is particularly important to include a realistic and differentiated reporting delay. An underestimation of the time between infection and reporting leads to the estimation of infections to a later date than actually occurred in reality. As a consequence, trend changes will also be dated too late, and thus, are attributed erroneously to specific interventions. Secondly, it is important to incorporate several indicators for the pandemic spread. Daily and cumulative infections as well as reproduction numbers, though based on the same initial data, have different meanings. As the results of this study show, significant change points may be found for some indicators but not for others. Thus, assessment of effectiveness of nonpharmaceutical interventions depends on the indicator used which leads to the conclusion that the temporal development of the indicators chosen should be carefully compared. And lastly, quantitative investigations based on empirical case data implicitly assume constant test volumes, which is obviously not true. In the German case, the number of conducted tests for SARS-CoV-2 is not constant over time. An increase (or decrease) of tests may result in an artificial increase (or decrease) of reported infections. Thus, increasing test capacity – which is a key parameter in fighting a pandemic – may result in a statistical source of error when analyzing pandemics over time. All these issues exist regardless of the chosen modeling approach, which suggests a need to shift study design toward prioritizing the handling of data sources rather than refining models.

Covid Economics 37, 14 July 2020: 57-72

References

1. Dehning J, Zierenberg J, Spitzner FP, Wibral M, Pinheiro Neto J, Wilczek M and Priesemann V. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* 2020; eabb9789.
2. Hartl T, Wälde K and Weber E. Measuring the impact of the German public shutdown on the spread of COVID-19. *Covid economics* 2020; 1: 25-32.
3. Hartl T, Wälde K and Weber E. Measuring the impact of the German public shutdown on the spread of COVID-19. *VoxEU*, April 14 2020.
4. Donsimoni J, Glawion R, Hartl T, Plachter B, Timmer J, Wälde K, Weber E and Weiser C. Covid-19 in Deutschland – Erklärung, Prognose und Einfluss gesundheitspolitischer Maßnahmen. *Perspekt. der Wirtschaftspolitik* 2020; ahead of print. DOI: 10.1515/pwp-2020-0019.
5. Weber, E. Which measures flattened the curve in Germany? *Covid economics* 2020; 24: 205-217.
6. Ferguson NM, Laydon D, Nedjati-Gilani G, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri A, Cucunubá Z, Cuomo-Dannenburg G, Dighe A, Dorigatti I, Fu H, Gaythorpe K, Green W, Hamlet A, Hinsley W, Okell LC, Van Elsland S, Thompson H, Verity R, Volz E, Wang H, Wang Y, Walker PGT, Walters C, Winskill P, Whittaker C, Donnelly CA, Riley S and Ghani AC (On behalf of the Imperial College COVID-19 Response Team). Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Report, Imperial College London, UK, March 2020.
7. Stedman M, Davies M, Lunt M, Verma A, Anderson SG and Heald AH. A phased approach to unlocking during the COVID-19 pandemic – Lessons from trend analysis. *Int J Clin Pract*. 2020; 00:e13528.
8. Wieland T. Flatten the Curve! Modeling SARS-CoV-2/COVID-19 Growth in Germany on the County Level. 2020. Pre-print available at: https://www.medrxiv.org/content/10.1101/2020.05.14.20101667v2.
9. Bourdin S, Ludovic J, Nadou F and Noiret G. Does lockdown work? A spatial analysis of the spread and concentration of Covid-19 in Italy. 2020. Pre-print available at: https://ersa.org/wp-content/uploads/2020/05/1-article-covid19vfok.pdf.
10. Orea L and Álvarez IC. How effective has been the Spanish lockdown to battle COVID-19? A spatial analysis of the coronavirus propagation across provinces. Efficiency Series Paper 2/2020, 13 April 2020. Oviedo: Universidad de Oviedo.
11. Santamaria L and Hortal J. Chasing the ghost of infection past: identifying thresholds of change during the COVID-19 infection in Spain. 2020. Pre-print available at: https://www.medrxiv.org/content/10.1101/2020.04.09.20059345v3.
12. Ben-Israel I. The end of exponential growth: The decline in the spread of coronavirus. *The Times of Israel*, 19 April 2020, https://www.timesofisrael.com/the-end-of-exponential-growth-the-decline-in-the-spread-of-coronavirus/ (2020, accessed 21 April 2020).
13. Dimdore-Miles O and Miles D. Assessing the spread of the novel coronavirus in the absence of mass testing. *Covid economics* 2020; 16: 161-176.
14. Born B, Dietrich AM and Müller GJ. Do lockdowns work? A counterfactual for Sweden. *Covid economics* 2020; 16: 1-22.

15. Homburg S. Effectiveness of Corona Lockdowns: Evidence for a Number of Countries. *The Economists' Voice* 2020; ahead of print. DOI: 10.1515/ev-2020-0010.
16. Wood SN. Did COVID-19 infections decline before UK lockdown? 2020. Pre-print available at: https://arxiv.org/abs/2005.02090.
17. Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, Whittaker C, Zhu H, Berah T, Eaton JW, Monod M, Imperial College COVID-19 Response Team, Ghani AC, A. Donnelly CA, Riley SM, Vollmer MAC, Ferguson NM, Okell LC and Bhatt S. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 2020; ahead of print. DOI: 10.1038/s41586-020-2405-7.
18. AD Heiden M and Hamouda O. Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland – Nowcasting. *Epid Bull* 2020; 17: 10-16.
19. Robert Koch Institut. Tabelle mit den aktuellen Covid-19 Infektionen pro Tag (Zeitreihe), dlde/by-2-0, https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/dd4580c810204019a7b8eb3e0b329dd6_0/data (2020, accessed 28 June, 2020).
20. Kuhbandner C, Homburg S, Walach H and Hockertz S. Was Germany's Corona Lockdown Necessary? 2020. Pre-print available at: https://advance.sagepub.com/articles/Comment_on_Dehning_et_al_Science_15_May_2020_eabb9789_Inferring_change_points_in_the_spread_of_COVID-19_reveals_the_effectiveness_of_interventions_/12362645.
21. Backer JA, Klinkenberg D and Wallinga J. Incubation Period of 2019 Novel Coronavirus (2019-nCoV) Infections Among Travellers From Wuhan, China, 20-28 January 2020. *Euro Surveill*. 2020; 25: 2000062.
22. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung S, Yuan B, Kinoshita R and Nishiura H. Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. *J Clin Med*. 2020; 9: 538.
23. Chowell G, Simonsen L, Viboud C and Kuang Y. Is West Africa Approaching a Catastrophic Phase or is the 2014 Ebola Epidemic Slowing Down? Different Models Yield Different Answers for Liberia. *PLoS Curr*. 2014; 6: ecurrents.outbreaks.b4690859d91684da963dc40e00f3da81.
24. Ma J. Estimating epidemic exponential growth rate and basic reproduction number. *Infect. Dis. Model*. 2020; 5: 129-141.
25. Bai J and Perron P. Computation and Analysis of Multiple Structural Change Models. *J. Appl. Econom*. 2003; 18: 1-22.
26. Zeileis A, Kleiber C, Krämer W and Hornik K. Testing and Dating of Structural Changes in Practice. *Comput. Stat. Data Anal*. 2003; 44: 109-123.
27. R Core Team. R: A language and environment for statistical computing, https://www.R-project.org/ (2019, accessed 18 December 2019). Vienna: R Foundation for Statistical Computing.
28. Robert Koch Institut. Täglicher Lagebericht des RKI zur Coronavirus-Krankheit-2019 (COVID-19), 24.06.2020, https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/2020-06-24-de.pdf?__blob=publicationFile (2020, accessed 25 June 2020).
29. Ipsos. Internationale Umfrage: Mehrheit empfindet Coronavirus als große Bedrohung. Press release (12 February 2020), https://www.ipsos.com/sites/default/files/ct/

news/documents/2020-02/ipsos-pi_coronavirus_feb2020.pdf (2020, accessed 24 May 2020).

30. Buchholz U, Buda S and Prahm K. Abrupter Rückgang der Raten an Atemwegserkrankungen in der deutschen Bevölkerung. *Epid Bull* 2020; 16: 7-9.

Covid Economics 37, 14 July 2020: 57-72

CEPR PRESS

Covid Economics **Issue 37, 14 July 2020**

Lockdown, pandemics and quarantine: Assessing the Indian evidence[1]

Saibal Ghosh[2]

Date submitted: 8 July 2020; Date accepted: 9 July 2020

The COVID-19 pandemic has put the global economy under a scanner. India has also been impacted by the pandemic and as a result, policymakers have undertaken significant set of measures to address the challenge. In this context, using daily state-level data, we utilize the staggered timing of the implementation of lockdown to ascertain its impact on the number of Covid19 cases. Our analysis appears to suggest that notwithstanding the lockdown, the number of Covid19 cases increased by 80% and furthermore, there was a differential impact across states, depending on their extent of health preparedness. Robustness tests support these findings.

Covid Economics 37, 14 July 2020: 73-99

1 The views expressed and the approach pursued in the paper reflects the personal opinion of the author.
2 Qatar Central Bank (previously at the Reserve Bank of India and the Centre for Advanced Financial Research and Learning, Mumbai).

Lockdown, pandemics and quarantine: Assessing the Indian evidence

I. Introduction

Once every hundred years, the human race gets infected with a disease that remains etched in memory and in the pages of history, for years to come. The plague in 1720, the cholera outbreak in 1820, the Spanish flu in 1920, all seems to point in this direction. The costs of such diseases are large, especially in terms of human lives. Thus, during the plague of 1720, a reported 125,000 people died. By the time of the subsiding of the Spanish flu in 1920, a reported 50 million died worldwide, making it one of the deadliest in human history.

A more recent case is the coronavirus disease, termed as Covid19 by the World Health Organization (WHO). Believed to have originated in the Wuhan province of China in late 2019, it quickly spread all over the world, engulfing countries and continents in varying degrees. As of May 2020, the virus had spread to over 200-odd countries, affecting nearly 45 million people, of which roughly 16 million have recovered, entailing a recovery rate of around 36%. According to the World Health Organisation, it is believed that the novel coronavirus is 10 times more deadly than the swine flu (WHO, 2020).

The coronavirus has spread at a rapid pace. Initially, it remained confined to Asia and especially the countries surrounding China as free movement of people led to the virus getting transmitted outside the country's borders. As on early January 2020, the first Covid-19 case, recorded outside China, was in Thailand. Subsequently, Europe became a hotspot with the major Alpine countries (Italy, France and Germany) being among the worst hit. In Asia, Iran became one of the worst affected countries. Footprints of the virus also extended far beyond with major Latin American countries as well as Africa not being spared of its impact.

On a global scale, as at end-May 2020, over 6 million Covid19 pandemic cases have been reported, with the United States of America accounting for nearly 30% of the total and the top 15 countries accounting for over 75% of the total. This includes not only emerging economies such as India, Turkey, China, but also developed economies such as UK, France and Germany, including in Latin America. Likewise during the same period, the crude fatality rate – defined as the number of deaths per 100 cases – was in double

Covid Economics 37, 14 July 2020: 73-99

Covid Economics 37, 14 July 2020: 73-99

digits in five of these 15 economies, as compared to a global average of 6.1, being the highest in France (19) and the lowest in Russia (1.2).[1]

The Indian economy has not been immune to these developments. Starting with the identification of the first coronavirus case in the Southern state of Kerala on January 30, 2020, the numbers have exploded at an exponential rate, touching 182,143 at end-May 2020. The crude fatality rate has also increased from 2.6 at end-March 2020 to 2.8 at end-May 2020. Measured in per capita terms, India consistently ranked at the bottom 25 percentile of the scale during March-May 2020, in spite of an increase in the total cases per million persons from 1 in March to 132 in May.

This increase at the all-India level masks the wide divergence across states. To be more specific, after an initial hiatus in February, the numbers suddenly started increasing in March. Out of the total number of active cases at end-March 2020, the top three states accounted for 42% of the total and the top 5 accounted for 55% of the total. By end-May 2020, the share of top three states was 47% and the share of top 5 states was 60% of the total. Maharashtra alone accounted for 18% of the total cases at end-March and 30% at end-May 2020. This is reflected the region-wise figures which shows that the share of Western region, which accounted for nearly 24% of the cases at end-March 2020 increased to 37% at end-May 2020. Concomitantly, the share of Southern region which accounted for 37% of total cases at end-March 2020 declined to 14% at end-May 2020.

To address the challenges, the policymakers undertook a significant number of measures. At the state-level, besides economic stimulus and communication, the concerned governments responded with various declarations of emergency, closure of institutions and public meeting places, and other restrictions intended to contain the spread of the virus. One significant measure in this regard was lockdown. Hailed as 'tough and timely' measure by the World Health Organisation, it has been contended that by promoting social distancing, the lockdown slowed the growth of the pandemic, lowering the doubling rate (i.e., the number of days taken for Covid19 cases to double) from 3.4 in the pre-lockdown period to 13.4 days at end-May 2020 (Government of India, 2020). Whether and how far the lockdown was effective in containing the pandemic across states remains a moot question.

[1] The global data is sourced from the online publication *Our World in Data*. Based at the University of Oxford, the mission of this publication is to present research and data to make progress against the major largest problems such as poverty, disease and hunger. For purposes of comparability, the aggregate information for India is sourced from this publication.

Covid Economics 37, 14 July 2020: 73-99

To inform this debate, this paper studies the effects of the lockdown on the number of active Covid19 cases. The research strategy exploits the exogenous variation in lockdown arising from its staggered implementation across states and adopts an appropriate research design to isolate the impact. We find that despite the lockdown, the number of Covid cases increased on average by 80%. These results differ across states with varying degrees of healthcare preparedness and are consistent with the European evidence which find lockdowns to be ineffective in containing the pandemic. We also take on board other relevant factors and find that notwithstanding their better preparedness in terms of health infrastructure, there was a substantial increase in Covid cases for states with higher net labor inflows, on average. This occurred in spite of the stringency of lockdown implemented in these states. We also explore the relevance of quarantine stringency, an aspect not adequately addressed in prior research. The findings suggest that the quarantine measures exerted a discernible and salutary impact only in states with health infrastructure in the second top quartile, although the evidence in respect of other health-category states is less compelling. Relatedly, we also examine the impact on other outcomes such as fatality rates and continue to find evidence of an increase, notwithstanding the lockdown.

India provides a compelling laboratory among emerging markets to examine this issue in some detail. With over a billion plus people and about 18% of the global population, it is the world's largest democracy. It is also the world’s sixth largest economy in terms of nominal GDP, according to the International Monetary Fund (2019). As a result, the health and prosperity of its people has significant knock-on effects on consumption, trade and investment flows in the global economy. Wolf et. al (2011) for example, show that, going forward, the country’s favourable demographics will enable it to augment its savings capacity and engender a consequent expansion of manufacturing and in the process, outpace the stellar growth achieved by neighbouring China. Second and from a global standpoint, India has been one of the countries with the highest number of Covid cases. To illustrate, as at end-May 2020, the country had over 182,000 cases, accounting for 3% of the global number as compared with just one solitary case at end-January 2020. This has occurred notwithstanding a lockdown being imposed in the country. Third, akin to the US and other developed and emerging economies such as Germany, Australia, Switzerland, Brazil, Russia, South Africa and Mexico, India is a federal polity comprising of states with their own democratically-elected government. The states exhibit marked divergence in terms of geographical heterogeneity and access to healthcare facilities. It therefore remains to be explored as to how the impact of lockdown of Covid cases has differed across states. The findings so obtained could be useful in designing appropriate policy responses in other emerging markets with similar federal structures.

An analysis at the state-level is important for several important reasons. First, given the country's federal structure, the polity comprises of democratically-elected government at the state-. Being a part of the State list, decisions on key public policy issues such as healthcare facilities are made and implemented at the state level.[2] Second, amidst the lockdown, states have responded with different (and multiple) measures to tackle the pandemic. This is also echoed in academic research which observes that the preparedness and response to Covid19 have differed at the state level (The Lancet, 2020).[3] Towards this end, we compute a stringency index which defines the strictness of the lockdown measures across various categories of public interface. This index ranges from zero (minimum) to one (maximum) and provides us with useful insights as to how far these measures were effective in containing the pandemic. And finally, a recent report by National Institute for Transforming India (NITI) *Aayog* finds that states differ widely in terms of their health infrastructure (Government of India, 2019). Towards this end, the Report develops a health index by aggregating several health-related indicators across three domains into a single number.[4] Using this index value, we categorize states into four quartiles and analyse the interaction between health preparedness and lockdown in addressing the pandemic. This provides us with useful insights as to how states with varying degrees of health preparedness were able to address the challenges of the pandemic.

The rest of the analysis proceeds as follows. Section 2 reviews of nascent literature and contextualizes the position of the paper. Section 3 discusses the data and empirical

[2] The Federal and the State government legislate across multiple areas which can be categorized under three heads: Union list (upon which only the Federal government has exclusive powers), State list (which comprise areas under which the legislature of the state make laws) and Concurrent list (upon which both the Federal and State governments can legislate). Illustrative areas under the Union list include defence, foreign affairs, banking, etc., those under the State list include public health and sanitation, police and relief for disabled and unemployed and finally, those under Concurrent list include areas such as education, population control and wildlife protection.

[3] Kerala leveraged on its 2018 experience of the Nipah virus outbreak to contain the current pandemic. Likewise, Odisha's prior experience in dealing with natural disasters meant several precautions were already in place and in case of Maharashtra, drone technology was used to monitor physical distancing during lockdown (The Lancet, 2020).

[4] These domains include health outcomes (focusing on measures such as mortality rate, sex ratio and immunization), governance and information (focusing on the status of the governance structures and information systems within states) and key inputs and processes (focusing on areas such as healthcare quality and availability, staff shortages and birth registration level).

Covid Economics 37, 14 July 2020: 73-99

strategy, followed by the results and robustness (Sections 4 and 5) and the concluding remarks.

II. Literature

On January 30, 2020, the first case of Covid19 was reported in Kerala when a student returning from Wuhan was initially suspected and subsequently confirmed to have contracted this virus. Within the next couple of days, the number increased to three, all cases relating to students returning from Wuhan. After no significant increase in transmission during the month of February, the numbers starting trending upwards since March. Specifically, on March 4, a total of 22 cases were reported across several states, such as Haryana, Delhi, Rajasthan and Uttar Pradesh. Subsequently, the pandemic spread elsewhere, with several states in the Western region becoming the hotspots. To illustrate, the first Covid19 case in Maharashtra was reported on March 9 and likewise, the first such reported case in Gujarat was reported during the third week of March. However, over time, both these states have witnessed a dramatic rise in the number of cases so much so that, at end-May, they accounted for close to 40% of the total Covid19 cases in the country. Along with Delhi and Tamil Nadu, the total number of new cases for these four states has averaged around 15% during Jan-May 2020.

Given its growing importance, research on various aspects of Covid19 is evolving (See, Dixit, 2020). Without loss of generality, research appears to have spawned in two major directions. The first stream of evidence speaks to the literature that examines the policy responses to the pandemic (See, for example the Covid Economics webpage for real-time vetted papers). Theoretically, Caballero and Simsek (2020) find that non-financial supply shocks such as Covid19 amplify into financial shocks that engender contractions in asset valuation and dampen aggregate demand. As the pandemic unfolded, the policymakers responded decisively and aggressively to limit the adverse consequences of the sudden breakdown in economic activity (Arslan et al., 2020). On the fiscal side, governments launched massive stimulus accompanied by funding and credit guarantees, on top of central banks' emergency support measures (IMF, 2020). This has been complemented by prudential policies which enabled countries to sustain credit growth and arrest bank deleveraging (BIS, 2020a). Using cross-country data on previous episodes of pandemic, studies report a contraction in economic activity by 6% (Barro et al., 2020) and a reduction in the natural rate of interest (Jorda et al., 2020).

Covid Economics 37, 14 July 2020: 73-99

A second line of analysis has examined the potential impact of the pandemic on the banking and corporate sectors, including financial markets. The evidence suggests that although bank equity prices witnessed a sell-off during the initial days of the pandemic, subsequent stabilization measures have favoured profitable and well-capitalised banks (Aldasoro et al., 2020). Other studies find that firms with prior experience in dealing with pandemic were better equipped to handle the current challenges (Hassan et al., 2020). Acharya and Steffen (2020) find that US firms with access to liquidity lines were rewarded by stock markets in terms of higher premium as compared to firms for whom such support was weak or non-existent, impelling the latter to make a 'dash for cash'.

Several studies have also examined the response of financial markets. In an early exercise, Bank for International Settlements (2020b) reported that the massive sell-off across asset classes and regions was manifest particularly for countries with close geographical and economic ties to China. Using daily stock price data on publicly held US firms, Ramelli and Wagner (2020) show that the cumulative returns of firms with greater exposure to China gradually retreated, after witnessing an initial decline.

Our study belongs to the literature that explores the impact of lockdown on the real sector and in particular, the daily number of Covid cases. Research arrives at opposite conclusions, with studies for the US emphasizing the efficacy of such lockdowns (Tellis et al., 2020), whereas others, especially for European countries (Born et al., 2020; Meunier, 2020; Stone, 2020), reporting the opposite. Thus, Tellis et al. (2020) report that US states that locked down late suffered a 15-25% higher penetration of the disease, while Stone (2020) find that the social and economic case for the efficacy of lockdown is less compelling.

Our analysis complements the extant literature in a few significant ways. First, this is one of the earliest exercises to provide quantitative evidence regarding the impact of the lockdown on the pandemic for a leading emerging economy. Most related research for India has been descriptive, at best. Persaud (2020) provides a commentary of how to sustainably exit the lockdown, borrowing from the Caribbean experience. Using an event study analysis, Mehrotra (2020) find that the adverse effects of pandemic-related conflicts are higher in low-income districts with inadequate health infrastructure. Utilising granular data for the Indian state of Bihar, Poblete-Cazeneve (2020) report a 44% decline in criminal activity during the lockdown period. Unlike these studies, we focus on the evolution of corona cases after imposition of the lockdown at the state level and find a significant increase, after controlling for other relevant factors.

Covid Economics 37, 14 July 2020: 73-99

Covid Economics 37, 14 July 2020: 73-99

Second, we develop an index of lockdown stringency. Towards this end, we combine the staggered implementation of various measures under lockdown into a single number and correlate it with the extent of health preparedness across states. The findings indicate that the stringency of the lockdown did not exert any perceptible impact on the number of coronavirus cases on average, although there is a differential impact across states with varying degrees of health preparedness.

Relatedly, we also develop an index of quarantine stringency by combining the three major facets of quarantine such as health appraisal, quarantine requirements and passenger obligations into a single number. To the best of our understanding, careful empirical analysis of the impact of quarantine on Covid19 cases has not yet been attempted in the literature. Our findings suggest that although there was no impact in the aggregate, there was a statistically significant impact on Covid19 cases in states which belong to the inter-quartile range of health preparedness.

Lastly, we explore whether the nature of the government matters for coronavirus cases. In an early study, Roubini and Sachs (1989) had observed that the political parties within a coalition government exhibit significant heterogeneity in their objective function, restraining them from undertaking sudden and significant expenditures increases. Recent evidence proffered by Tellis et al. (2020) indicates that US states with Democrat affiliation experience higher impact of the pandemic as compared with those having Republican affiliation.

III. Data and empirical strategy

We utilize data covering all states and union territories, except Ladakh, for which separate data on key macroeconomic and other relevant variables is not reported.[5] As a result, we have data on 29 states and four union territories. We combine two sets of state-level data: (i) daily data and (ii) yearly data.

III.1 Covid data

The key variable is the information on the daily number of new Covid19 cases, extracted from the website of the Ministry of Health and Family Welfare, Government of India and covering the period January 30-May 31, 2020. The beginning date coincides

[5] In August 2019, vide the *Jammu and Kashmir Reorganization Act 2019*, the state of Jammu and Kashmir was reorganized into two Union territories: Jammu and Kashmir and Ladakh, effective from October 31, 2019. Lakshadweep was not included in the study due to paucity of data.

with the reporting of the first Covid19 case in a state and the end date coincides with ending of the fourth phase of the lockdown.[6] The data has been subject to several revisions during this period, with cases being reassigned across states. After these adjustments, we have information on 29 states and a maximum of 4052 state years.

III.2 Annual data

The other state level variables employed include a political variable (i.e., a dummy whether the state government is a majority, else zero) and dummies for the health preparedness of a state, categorized into four quartiles.[7]

III.3 Lockdown and related data

Our crucial variable of interest is the lockdown measures. Although the nation-wide lockdown was imposed on 25 March, states had already begun imposing lockdowns prior to that using their powers under the *Epidemic Diseases Act, 1897*.[8] Indeed on March 11, the Union government urged all states and union territories to invoke the provisions of Section 2 of the *Epidemic Diseases Act, 1897*. Using manifold sources, including Wikipedia, PRS Legislative Research, Confederation of Indian Industry and the coronavirus daily information tracker, we pin down the date of enforcement of lockdown across states.[9] We provide several examples, one from each of the six regions of the country.[10]

> In Rajasthan, the *Epidemic Disease Covid-19 Regulations* was issued on March 17 and on the next day, government imposed state-wide curfew. Thereafter, on March 22, a complete lockdown was announced in the state. It was informed that all government and private offices, malls, shops, factories and public transport will remain closed, while its borders with other states will be sealed. The Union Minister for Culture and Tourism decided to shut all Archaeological Survey of India protected monuments and central museums, effective March 16.

[6] Although a fifth phase of lockdown was announced covering the period June 1-30, 2020, it was much less restrictive in scale and scope than the previous ones.
[7] Thus, states are classified as having low, medium, upper medium and high health preparedness, depending on whether the values of health index belongs to the first (upto 25%), second (above 25 and upto 50%), third (above 50 and upto 75%) and fourth (above 75%) quartile, respectively.
[8] The *Epidemic Diseases Act, 1897* (issued on February 4, 1897) consists of four sections, dealing with empowerment of state governments and prescriptions of regulations therein (Section 2), power to impose penalty on persons who violate the prescribed regulations (Section 3) and protection to persons acting under the Act (Section 4).
[9] Based in New Delhi, PRS Legislative Research is an Indian non-profit organisation that was established in September 2005 as an independent research institute to make the Indian legislative process better informed, more transparent and participatory in nature.
[10] The six regions are Northern, Southern, Eastern, Western, Central and North-Eastern. The discussion in the text follows this sequence.

Covid Economics 37, 14 July 2020: 73-99

Covid Economics 37, 14 July 2020: 73-99

In exercise of its powers under the *Tamil Nadu Public Health Act 1939,* the state government declared coronavirus as a notified disease on March 13. On March 15, in a series of instructions, the government restricted the movement of people in the state. These included, shutting down of establishments, such as, educational institutions (up to Class 5), theatres, malls etc, and banning of inter-state travel for 15 days. Later on March 16, the government announced additional restrictions such as closure of *anganwadis* (i.e., rural child care centres) and making alternate provision of dry ration for children at their homes, swimming pools, amusement parks, gyms, zoos, museums, bars, clubs etc, and all educational institutions (except the conduct of practical exams for class 10 and 12, and various entrance exams). A couple of days later, state borders were sealed off for road traffic effective March 20, except for movement of essential commodities. Public transportation services, such as metro rail and inter-state private buses, were also suspended till March 31. A complete lockdown was announced effective March 24.

On March 13, the state government in Odisha ordered for the closure of cinema halls, swimming pools, gyms and educational institutions (except for holding examinations). The *Odisha Covid-19 Regulations* were notified on March 18 and came into force with immediate effect. Effective March 20, the government also imposed restrictions on the entry of devotees to all places of worship in the state. Subsequently, on March 23, the government issued an order suspending intra-state bus services from March 24 and city bus services in all urban local bodies from midnight of March 23. Prior to that, on March 21, the government announced lockdown in five revenue districts and eight towns of the state until March 29. The lockdown involved (i) suspension of public transport services (ii) closure of all commercial establishments, offices, and factories (iii) banning the congregation of more than seven people at any public place. On March 24, the government extended the lockdown to the entire state. The establishments engaged in the supply of essential goods and services were excluded from this lockdown. Later on March 30, the government sealed its borders with neighbouring states.

In Maharashtra, the *Epidemic Diseases Act 1897* was notified on March 13 and on March 15, the state government ordered the closure of cinema halls, swimming pools, gyms, theatres, and museums until March 31. Subsequently on March 16, all educational institutions and hostels in the state were closed and all exams were also deferred until March 31. The state also sealed its borders with neighbouring states. Vide its order dated March 23, the government enforced a state-wide lockdown with immediate effect till March 31, 2020.

In Uttar Pradesh, the *Uttar Pradesh Epidemic Disease Covid-19 Regulations 2020* was notified on March 14 and came into force with immediate effect. Vide its notification dated March 22, the government ordered the imposition of lockdown in 16 districts starting March 23. People in these districts were advised to avoid going to religious places or assemble for any purpose, including weddings. The lockdown was

> subsequently extended to the entire state on March 24. During March 13-17, the government ordered the closure of educational institutions, cinema halls, museums, and tourist spots in order to prevent public gatherings. On March 20, this was extended to include malls, and all religious, social, and cultural activities.
>
> In Assam, the *Assam Covid-19 Containment Regulations* were issued on March 21. The state government issued an order on March 19 for closure of all museums, libraries, coaching centres, beauty parlours and barber shops. To further restrict the movement of individuals, in order to contain the spread of the disease, the state government enforced a state-wide lockdown from March 24. This involved: (i) sealing the state borders, (ii) suspension of public transport services, (iii) closure of all commercial establishments, offices, and factories, and (iv) banning the congregation of more than five people at public places. The establishments providing essential goods and services were excluded from these lockdown restrictions.

We utilise this information on the staggered implementation of various restrictions to develop a lockdown stringency index. As outlined above, we examine the implementation of restrictions imposed by states across eight areas which include the closure of educational institutions, closure of public places (e.g., malls, cinema halls), closure of public and private transport, closure of borders, banning of tourist (domestic and/or foreign) and closure of religious places. Depending on the date from which these measures were implemented in a state, we code it as one, else zero. As a result, the maximum value across all categories equals eight (all the measures were in place) and a minimum of zero (none of the eight measures were in place). We divide the value for each state-date by eight and arrive at a number which lies in the unit interval. Even within a state, since these restrictions were imposed at different points in time, the average number of days it takes for the quarantine stringency index to assume its maximum value equals 21; the values at the 25th and 75th percentile equal 7 and 35, respectively. Contextually, Elgin et al. (2020) utilise a similar methodology to develop an economic stimulus index and analyse it linkage with several macroeconomic variables in their cross-country study.

Using the lockdown data, we also construct a learning variable such that states affected with the disease learn from those afflicted earlier. In line with recent research (Tellis et al., 2020), this is defined as the number of days between the first case in any given state and the first case across all states.

In Table 1, we report the definition, source and the summary statistics of the relevant variables. The average daily number of cases during this period is 45, although its variability is high. Regarding the independent variables, the lockdown was in place for

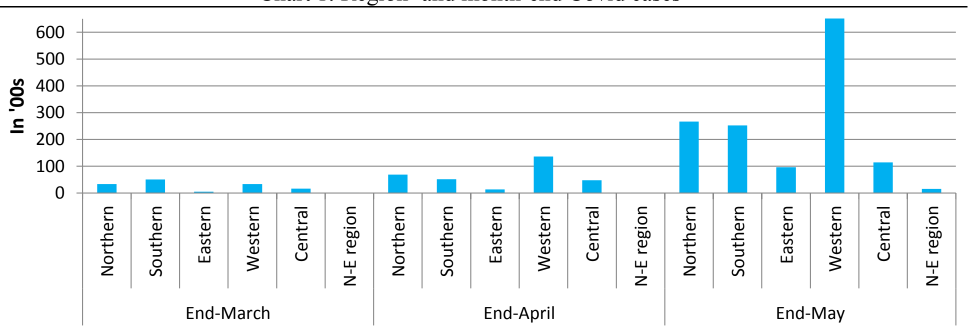
Covid Economics 37, 14 July 2020: 73-99

32% of the period, on average. The levels of lockdown stringency appear modest; its high variability would suggest that not all states have been able to enforce all lockdown-related measures. Among others, the average value of the health index is 54 and in 43% of the states, the government is a single party majority. Across regions, the total numbers started trending upwards since April, with the Western region being the worst-affected by end-May 2020 (Chart 1).

Table 1: Variable definitions and summary statistics

Variable	Definition	Data source	N.Obs	Mean (SD)
Dependent				
Covid	Number of new coronavirus cases in a state (the regression use the natural logarithm of one plus the number of cases)	Wikipedia, PRS, The Hindu	4055	45.56 (199.87)
Fatality	Number of coronavirus-related deaths in a state (the regression use the natural logarithm of one plus the number of cases)	As above	4059	1.28 (6.54)
Independent				
POST	Binary variable equal to one beginning from the day a state has imposed lockdown, else zero	As above	4059	0.32 (0.47)
Learning	Number of days between the first case in any given state and the first case across all states	Author calculations	4059	48.44 (43.71)
Lockdown	Index of lockdown stringency as explained in the text.	Author calculations	4059	0.36 (0.31)
Quarantine	Index of quarantine stringency as explained in the text	Author calculations	4009	0.04 (0.13)
Trend	Number of days since the first reported case in the state	Author calculations	4059	24.09 (27.31)
Time	Number of days since the imposition of the quarantine	Author calculations	2009	0.68 (2.72)
Health	Index of health	Niti Aayog	4059	53.57 (10.80)
Government	Dummy=1 if the government is single party majority, else zero	Election Commission of India	3690	0.43 (0.49)
Migration	Dummy=1 for states which have experienced highest levels of net in-migration, else zero	Census of India	4059	0.18 (0.39)
Civil	Number of civil police/ 100,000 persons (the regressions use the natural logarithm of the variable)	Bureau of Police Research	4059	184.06 (148.41)

Chart 1: Region- and month-end Covid cases



Note: the numbers upto End-March have been multiplied by 10 to ensure its visibility in the scale

Table 2 records the correlation matrix of the relevant variables. The table indicates that the lockdown variable is strongly positively with the dependent variable. However,

Covid Economics 37, 14 July 2020: 73-99

the correlation among them and most of the other relevant correlations are modest, of the order of 20-25%. This is reflected in Charts 2A and 2B which depict the plot between health index and respectively, with the two indices of quarantine. The findings indicate a negative but weak association between health index and lockdown stringency and on the other hand, a positive association between health index and quarantine stringency. The raw correlations do not control for relevant factors and as a result, it becomes important to examine the issue within an econometric framework.

Table 2. Correlation matrix of key variables

	Covid	POST	Learning	Health	Government	Lockdown	Quarantine
Covid							
POST	0.212						
	(0.000)						
Learning	0.282	0.189					
	(0.000)	(0.000)					
Health	0.075	0.085	0.046				
	(0.000)	(0.000)	(0.003)				
Government	0.038	0.165	0.078	0.005			
	(0.021)	(0.000)	(0.000)	(0.742)			
Lockdown	0.123	0.278	0.656	-0.049	0.019		
	(0.000)	(0.000)	(0.000)	(0.002)	(0.230)		
Quarantine	0.232	0.029	0.427	0.061	0.030	0.107	
	(0.000)	(0.066)	(0.000)	(0.000)	(0.068)	(0.000)	

p-Values in brackets

Chart 2A: Lockdown stringency and health index

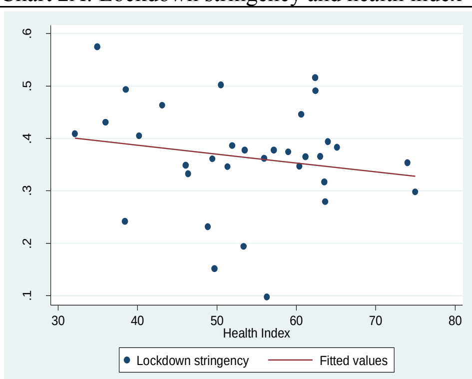
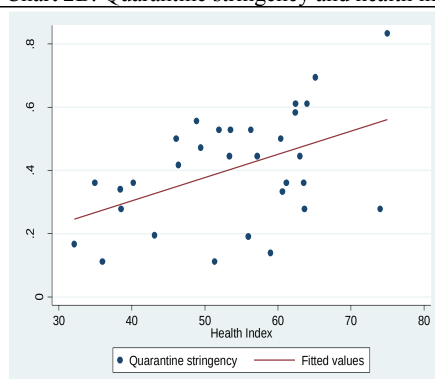


Chart 2B: Quarantine stringency and health index



IV. Lockdown and pandemic

We begin our discussions by investigating the relation between lockdown and coronavirus cases. In this regard, we employ the natural experiment of the lockdown that exploits inter-temporal variation in its implementation across states.

Covid Economics 37, 14 July 2020: 73-99

To study the effects of lockdown on the number of Covid19 cases, we employ the following baseline regression for state s on day d:

$$y_{sd} = \eta_d + \nu_s + \delta\, POST_{sd} + trend_s + Sq.(trend_s) + \varepsilon_{sd} \qquad \textbf{(1)}$$

In Eq.(1), y is the outcome variable, defined as the natural logarithm of one *plus* the number of new Covid cases in a state. *POST* is the measure of lockdown, which equals one for the days after imposition of lockdown, else zero. The coefficient of interest is δ, which depicts the effect of the lockdown on the number of Covid19 cases. To the extent that the lockdown exerts a beneficial impact, one would expect δ to be negative. Since the dependent variable is in logarithm and the key explanatory variable is binary, the point estimates can be interpreted as percentage change due to the imposition of lockdown.

In Eq. (1), we control for state- and day-fixed effects to account for unobservable state- and day-specific characteristics. We measure the number of days since the first reported Covid19 case in a state (*trend*) as a proxy for the spread of the disease. The possible non-linearities are accounted for by including its squared term and ε is the error term

The other measures include a proxy for learning and a political variable (i.e., a dummy whether the state government is a single-party majority, else zero) and dummies for the health preparedness of a state, divided into four quartiles. Throughout, standard errors are clustered by state, to account for possible correlated shocks to state-level Covid cases over time (Bertrand et al., 2004).

IV.1 Baseline results

Regression results are presented in Table 3. In Col.(1), the findings indicate that that notwithstanding the lockdown, Covid19 cases increased by 82%. The point estimate is statistically significant at the 1% level. The point estimate on the coefficient remains broadly unaltered, even when we include additional variables (Cols. 2 and 3) or exclude union territories (Col. 4).

The coefficient on *trend* and its squared term are both positive (Col.1). Both of them are statistically significant at conventional levels. The inflection point in the relationship is reached at 120 days.[11] In other words, once the number of days exceed this threshold, the number of cases increase at an increasing rate. When we add all controls

[11] The inflection point is calculated as the derivative of Covid case with respect to trend. The other reported inflection points are calculated in a similar manner.

(Col. 3), the squared term is observed to be negative, suggesting a concave relationship with an inflection point of 145 days, so that it takes about 5 months on average for the Covid cases to decline. Across columns, we find that it takes around 111 days for cases to decline in high health-index states (Col. 5) and conversely, around two months for cases to exhibit an increase in states with low levels of health preparedness (Col. 8). The average impact in Column 3 earlier therefore appears to reflect the countervailing effect of these opposing magnitudes.

Table 3. Regression results: Impact of lockdown on number of active cases

	All states and Union Territories			Excluding Union Territories	High health index	Upper medium health index	Medium health index	Low health index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
POST	**0.818*****	**0.818*****	**0.800*****	**0.839*****	**-0.298**	**1.590*****	**0.554*****	**1.497*****
	(0.065)	**(0.065)**	**(0.068)**	**(0.069)**	**(0.265)**	**(0.204)**	**(0.121)**	**(0.102)**
Trend	0.048***	0.054***	0.058***	0.059***	0.089***	0.018	-0.013	0.028***
	(0.003)	(0.004)	(0.004)	(0.004)	(0.011)	(0.011)	(0.010)	(0.007)
Sq.(Trend)	0.0002***	-	-0.0002**	-0.00002	-0.0004***	0.0005***	0.0009***	0.00024***
	(0.00003)	0.00008**	(0.00003)	(0.00003)	(0.00006)	(0.0001)	(0.0001)	(0.00007)
		(0.00003)						
Learning		-0.070***	-0.078***	-0.073***	-0.071	0.034	0.020	-0.071***
		(0.015)	(0.016)	(0.016)	(0.054)	(0.033)	(0.037)	(0.027)
Government			-0.817***	-0.812***	-0.562***	1.297***	1.165***	-1.256***
			(0.181)	(0.181)	(0.219)	(0.137)	(0.141)	(0.138)
State FE	YES	YES	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	4052	4052	3684	3438	860	738	981	1105
Adj. R^2	0.7216	0.7224	0.7313	0.7322	0.6579	0.8117	0.6991	0.7877

Standard errors (clustered by state) are in parentheses
*** p<0.01; ** p<0.05; * p<0.10

In Col.(2), we include learning as an additional explanatory variable. The average learning period for the sample states equals 48 days. To the extent that learning provides states with useful insights about possible ways to contain the pandemic, the coefficient is likely to be negative. We find this to be indeed the case. In other words, learning lowers the number of pandemic cases by 7-8%, on average. These findings run contrary to Telly et al. (2020) who report a positive coefficient on this variable in their findings for US states.

Col. (3) includes the nature of government as an explanatory variable, over and above those incorporated earlier. The point estimate on this variable is negative and statistically significant and would suggest that the number of pandemic cases has on average, been lower in states where the government is a single party majority.

The results remain broadly unaltered when we exclude the union territories from the analysis (Col.4).

In Cols. (5)-(8), we repeat our earlier analysis but instead, disaggregate the results by states with varying degrees of health preparedness. The quartile of such preparedness is indicated at the top of each column. Three important findings emerge. First, the results for the entire sample are echoed for all except the high health index states. Second, the coefficient on the lockdown variable is the highest for states in the second top and lowest quartile in terms of preparedness. In terms of magnitude, it is roughly double the coefficient for the entire sample. This outsized response for these states would suggest that the pandemic caught up early in these states, overwhelming their health infrastructure and in the process, compromising on the efficacy of the lockdown. As compared to this, for states with medium health index, the coefficient is much smaller, around two-thirds of the coefficient value for the full sample, presumably reflecting their better preparedness in terms of health infrastructure. In addition, the learning effect is manifest only in case of states with low values of the health index; its relevance in case of the other state categories is less persuasive.

The findings suggest that in spite of the lockdown, there was an increase in Covid cases across states. However, it is possible that factors that drive lockdown could also drive pandemic cases. This concern however, is likely to be compelling, since the exogenous variation in lockdown driven by the staggered implementation across states provides a natural experiment that is less likely to be correlated with lockdown.

IV.2 Lockdown and pandemic: Cross-sectional heterogeneity

It is intuitive that the impact of the lockdown would be higher for states that experience net in-migration. These migrants are typically daily wage laborers and employed in low-skilled jobs, often with no healthcare and related benefits. In that case, we would expect the impact of the pandemic to be relatively higher for such states, notwithstanding the lockdown.

To test this prediction, we estimate specification (2) as under:

$$y_{sd} = \eta_d + \nu_s + \delta\, POST_{sd} + \lambda Migration_s + \theta\,(POST_{sd} * Migration_s) + trend_s + Sq.(trend_s) + \varepsilon_{sd} \quad \textbf{(2)}$$

In Eq.(2), *Migration* is a dummy for states with highest levels of net in-migration and our interest is the point estimate on θ, the remaining variables are as defined earlier. The summary statistics show that roughly a fifth of the states have experienced high levels of in-migration. The effect of net in-migration on Covid cases in the post-lockdown phase can be computed as:

Covid Economics 37, 14 July 2020: 73-99

$$\frac{d\,y_{sd}}{dMigration_s} = \lambda + \theta\ POST_{sd} \quad \textbf{(3)}$$

Table 4. Regression results: Impact of lockdown on number of active cases in net in-migration states

	All states and Union Territories	High health index	Upper medium health index	Medium health index	Low health index
	(1)	(2)	(3)	(4)	(5)
POST * Migration	**1.425*****	**2.680*****	**0.665*****	**1.571*****	**…**
	(0.110)	**(0.190)**	**(0.161)**	**(0.263)**	
POST	0.503***	-1.106***	1.266***	0.381***	1.497***
	(0.068)	(0.241)	(0.223)	(0.121)	(0.102)
Migration	-0.286**	0.448***	0.089	0.429***	…
	(0.126)	(0.161)	(0.108)	(0.079)	
Trend	0.053***	0.091***	0.014	-0.032***	0.028***
	(0.004)	(0.010)	(0.011)	(0.010)	(0.007)
Sq.(Trend)	0.00004	-0.0004***	0.0006***	0.001***	0.0004***
	(0.00004)	(0.00006)	(0.0001)	(0.0001)	(0.00008)
Learning	-0.061***	-0.144***	0.045	0.094***	-0.071***
	(0.016)	(0.046)	(0.031)	(0.035)	(0.027)
Government	-0.812***	-0.157	0.828***	0.319**	-1.256***
	(0.136)	(0.161)	(0.133)	(0.143)	(0.138)
State FE	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES
Observations	3684	860	738	981	1105
Adj. R^2	0.7478	0.7385	0.8168	0.7213	0.7877

Standard errors (clustered by state) are in parentheses
*** p<0.01; ** p<0.05; * p<0.10

In Table 4, we find that across all states, the coefficient on migration is negative and statistically significant with a point estimate equal to -0.29. Taken along with the coefficient on the interaction term, the net impact in the post lockdown phase (i.e., POST=1) is 1.14 (= -0.29+1.43*1). Thus, there is nearly 115% increase in Covid cases for states with higher net labor inflows, on average. These findings lend support to the on-ground evidence which suggests that states with high levels of net worker inflows have borne a major brunt of the pandemic, notwithstanding their higher income and health indices.

This finding is echoed across columns which show that the net effect is nearly triple the average for high health index states, suggesting that despite their preparedness on various facets of health, the high levels of transient workers in the state has increased the Covid case. As compared to this, for states with upper-medium health preparedness, the coefficient is around half the overall value, indicating that worker in-migration in these states has perhaps been much lower.

IV.3 Lockdown and pandemics: Time-series variation

Next, we exploit the time-series variation in the stringency of the lockdown and its impact on Covid19 cases across states. Towards this end, we estimate specification (4) according as:

$$y_{sd} = \eta_d + \nu_s + \delta\, POST_{sd} + \mu\, Stringency_{sd} + \rho\,(POST_{sd} * Stringency_{sd}) + trend_s + Sq.(trend_s) + \varepsilon_{sd} \quad \textbf{(4)}$$

where the variables are as elucidated earlier and the coefficient of interest is ρ. For a state that has imposed a lockdown, this coefficient explores whether the lockdown stringency exerts any perceptible impact on Covid cases. Provided greater stringency exerts a dampening effect, $\rho<0$.

Table 5. Regression results: Impact of lockdown stringency

	All states and Union Territories	High health index	Upper medium health index	Medium health index	Low health index
	(1)	(2)	(3)	(4)	(5)
POST * Lockdown	**0.271**	**5.726*** **	**-4.795*** **	**1.065**	**-4.556*** **
	(0.316)	**(1.171)**	**(1.804)**	**(0.698)**	**(0.569)**
POST	0.674***	-3.227***	4.461***	0.446	4.282***
	(0.194)	(0.655)	(1.112)	(0.332)	(0.364)
Lockdown	1.247***	-1.319***	0.570	2.039***	1.009***
	(0.175)	(0.428)	(0.396)	(0.432)	(0.232)
Trend	0.061***	0.081***	0.030**	0.021**	0.044**
	(0.004)	(0.011)	(0.012)	(0.011)	(0.008)
Sq.(Trend)	0.0006	-0.0004***	0.0004***	0.001***	0.0002***
	(0.0007)	(0.00005)	(0.0001)	(0.0001)	(0.00008)
Learning	-0.062***	-0.081	0.026	0.129***	-0.079***
	(0.015)	(0.052)	(0.032)	(0.037)	(0.026)
Government	-0.751***	-0.482**	1.175***	0.938***	-1.542***
	(0.179)	(0.223)	(0.133)	(0.127)	(0.157)
State FE	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES
Observations	3684	860	738	981	1105
Adj. R^2	0.7367	0.6724	0.8165	0.7192	0.8061

Standard errors (clustered by state) are within parentheses
*** p<0.01; ** p<0.05; * p<0.10

The results in Table 5 indicate several incisive findings. First of all, the coefficient on the interaction term is insignificant for the entire sample. Secondly, at the disaggregated level, the coefficient is positive and statistically significant for high health index states. On the other hand, the coefficient is negative for states with upper-medium and low health indices. As observed earlier, notwithstanding the stringency, the high migrant workforce overwhelmed the health infrastructure in high health index states notwithstanding the beneficial impact of stringency (the negative and statistically significant coefficient on stringency in Col. 2), negating their efficacy. As compared to this, although the impact of stringency was negative for low health index states in post lockdown period, the inefficacy of the stringency measures (the positive and statistically significant coefficient on lockdown in Col.4) exerted a dampening effect overall. For states belonging to the upper medium quartile of health index, although the stringency effect *per se* was ineffective, the overall impact in the post-lockdown period was negative, lowering the Covid cases. However, given the low absolute number of cases in these states, the net effect was an overall spike, in spite of the lockdown.

IV.4 Lockdown and pandemics: Impact of quarantine measures

Next, we examine the impact of quarantine measures on lockdown. To elaborate, on May 24, 2020, the Ministry of Health and Family Welfare issued an Office Memorandum delineating the guidelines for domestic travel (air/train/inter-state bus).[12] The guidelines also explicitly mentioned that states can develop their own protocol as regards to quarantine and isolation as per their assessment.

Utilizing these guidelines, states issued their respective guidelines: while certain states suitably tweaked these guidelines, others retained the same guidelines as issued by the Ministry. Exploiting the variation in state-level guidelines, we develop a state-wise index of quarantine stringency as follows.

We include three facets of quarantine: health appraisal, quarantine requirements and passenger obligations and include two aspects under each of these facets. To measure stringency, we employ a coding process wherein we assign values between zero (if the state has not prescribed any guideline under this aspect) and one (if the state has prescribed the most stringent guidelines under this aspect). As a result, the value of quarantine stringency ranges from zero to six. We scale the actual value for each state by six (the maximum achievable) to arrive at a number between zero (no stringency) and one (maximum stringency).

Under health, we consider two aspects: whether the passenger has to undergo thermal screening and second, whether Covid19 testing is mandatory, random or not prescribed. Thus, if the passenger has to undergo thermal screening, we insert a value of one, else zero. Likewise, if Covid testing is mandatory, we assign a value of one, 0.5 if it is random and zero, otherwise.

Under quarantine requirements, we assign value 1 if institutional quarantine of 14-days is mandatory, 0.5 if institutional quarantine of 7 days is mandatory, else zero.

[12] Office Memorandum No. F.No.Z.28015/19-2020-EMR (Pt.)

Covid Economics 37, 14 July 2020: 73-99

Likewise, we assign value 1 if home quarantine of 21 days is mandatory and values of 0.66 and 0.33 respectively, if home quarantine is for 14 and 7 days, else zero.[13]

As part of passenger obligations, we assign a value of one if using *Aarogya Setu* is compulsory, else zero.[14] In addition, we assign value one if enrolment on relevant government website is compulsory, 0.5 if the passenger has to provide self-declaration on the health status, else zero.

We plot the state-wise lockdown and quarantine stringency measures, in increasing order of the latter (Chart 3). Without loss of generality, we find that states in the Eastern region of the country (e.g., Bihar, Odisha) and Telengana have the least stringent quarantine measures, on average. Using this process, we arrive at an aggregate number of quarantine stringency for each state. An advantage of this index is that is exhibits significant state-wise variation. On the flip side, the index does not display any time-series variation. To address this shortcoming and taking on board the fact that the effective implementation of the guidelines differed across states, we estimate the empirical specification:

$$y_{sd} = \eta_d + \nu_s + \alpha\,(POST_{sd} * Quarantine_s * time_s) + \rho_1\,(POST_{sd} * Quarantine_s) + \rho_2\,(POST_{sd} * time_s) + \rho_3\,(Quarantine_s * time_s) + \delta_1\,POST_{sd} + \delta_2\,time_s + \delta_3\,Quarantine_s + trend_s + Sq.(trend_s) + \varepsilon_{sd} \qquad \textbf{(5)}$$

where *time* is the number of days since the imposition of quarantine stringency in the state; the remaining variables are as defined earlier. The coefficient of interest is α: this three-way interaction term captures the dynamics of quarantine stringency since its imposition on the outcome variable in the post-lockdown phase (See, for example, Berger et al., 2005). The specification includes all possible two-way interactions and single variables, including controls and trend. Table 6 presents the results.

[13] We assign a value of 1 if 21-day home quarantine is mandatory (e.g., Mizoram); accordingly, the assigned values for 14-day and 7-day quarantine are 0.667 (=14/21) and 0.334 (=7/21), respectively.

[14] Developed under the guidance of the Indian Ministry of Electronics and Information Technology and released on April 2, 2020, *Aarogya Setu* is a mobile application to enable contact tracing. The frontline health workers rely on information collected by the application to identify clusters, develop responses and take adequate precautions in their line of work.

Chart 3: State-wise indices of lockdown and quarantine stringency (average value)

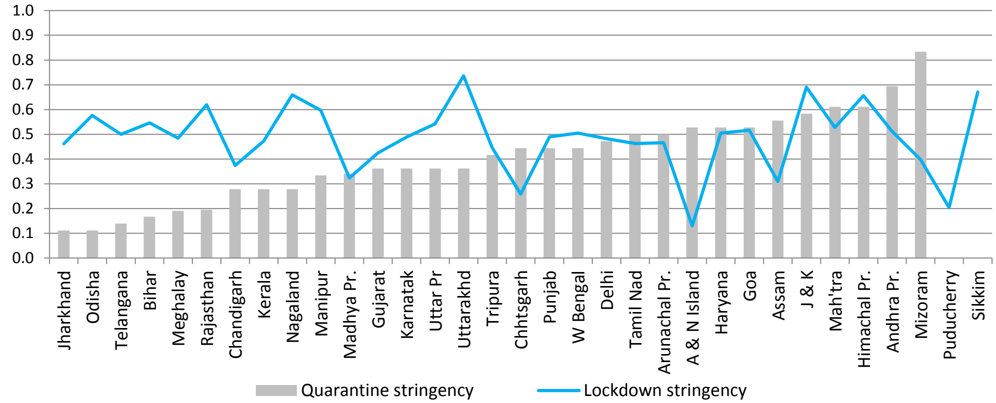


Table 6. Regression results: Impact of quarantine stringency on number of active cases

	All states and Union Territories	High health index	Upper medium health index	Medium health index	Low health index
	(1)	(2)	(3)	(4)	(5)
POST * Quarantine*Time	**-0.024**	**-0.568**	**-1.236****	**4.299*****	**-0.119**
	(0.144)	**(0.359)**	**(0.586)**	**(0.822)**	**(0.221)**
POST*Quarantine	-0.737	-3.437***	0.241	3.696***	-0.128
	(0.635)	(1.088)	(1.510)	(1.332)	(0.948)
POST*Time	0.147**	0.647***	0.274	-2.114***	-0.053
	(0.069)	(0.168)	(0.194)	(0.376)	(0.058)
Quarantine*Time	-0.424***	-0.521*	1.240***	-0.758***	-0.338***
	(0.084)	(0.293)	(0.477)	(0.221)	(0.065)
POST	0.708***	-0.008	1.862***	0.181	1.474***
	(0.075)	(0.328)	(0.254)	(0.128)	(0.506)
Quarantine	1.453***	3.858***	3.497**	-0.069	-1.836***
	(0.439)	(0.963)	(1.551)	(0.718)	(0.518)
Time	0.193***	0.287***	-0.267**	0.357***	0.046
	(0.030)	(0.095)	(0.136)	(0.112)	(0.033)
Trend	0.065***	0.076***	0.019	-0.007	0.039***
	(0.004)	(0.011)	(0.013)	(0.010)	(0.007)
Sq.(Trend)	-0.00001	-0.0002***	0.0003**	0.0009***	0.0002**
	(0.00003)	(0.00006)	(0.0001)	(0.0001)	(0.00007)
Learning	-0.078***	-0.037	0.020	0.021	-0.094***
	(0.016)	(0.049)	(0.034)	(0.036)	(0.027)
Government	-0.934***	-0.658***	0.946***	1.037***	-1.259***
	(0.175)	(0.215)	(0.140)	(0.138)	(0.137)
State FE	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES
Observations	3642	860	738	939	1105
Adj. R^2	0.7407	0.7003	0.8318	0.7473	0.8086

Standard errors (clustered by state) are within parentheses
*** p<0.01; ** p<0.05; * p<0.10

The findings reveal that at the aggregate level, the triple interaction term is not statistically significant. Across columns, the coefficient is negative for upper medium health index states and positive for medium health index states. Thus, the dynamic impact of quarantine was a decline in Covid19 cases in the post-lockdown period for the former and the reverse in case of the latter. Thus, the quarantine measures exerted a discernible and salutary impact only in states with health infrastructure in the upper-medium category, although it was overwhelmed by an adverse impact in states with medium health infrastructure.

In sum, the balance of evidence would suggest that any meaningful analysis of the impact of the pandemic on Covid cases would need to take on board the stringency of the imposed measures as well as the preparedness of the health infrastructure in the state.

V. Robustness checks

In this section, we undertake certain additional tests of the baseline findings. First, we examine the impact of lockdown on fatality rates. Second, we utilize an Instrumental Variable (IV) approach to ascertain the robustness of the findings.

V.1 Impact of lockdown on fatality rates

We run regressions similar to specification (1) earlier, except for the fact that the outcome variable is the natural logarithm of one plus the number of coronavirus-related deaths. The findings reported in Table 7 show that, notwithstanding the lockdown, the number of Covid-related deaths increased by roughly 27% on average, and is higher (31%) when union territories are excluded. Across states with varying degrees of health preparedness, there is no perceptible impact on states with high levels of health preparedness; it is the states in the second top quartile that appear to have been the most affected. Since all regressions take on board the other relevant controls including state- and day-fixed effects, it appears likely that our findings reflect the impact of the lockdown on fatality rates across states.

Covid Economics 37, 14 July 2020: 73-99

Table 7. Regression results: Impact of lockdown on fatality rates

	All states and Union Territories			Excluding Union Territories	High health index	Upper medium health index	Medium health index	Low health index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
POST	**0.272*****	**0.273*****	**0.278*****	**0.312*****	**0.177**	**0.781*****	**0.068****	**0.466*****
	(0.035)	**(0.035)**	**(0.038)**	**(0.038)**	**(0.173)**	**(0.081)**	**(0.033)**	**(0.052)**
Trend	0.006***	0.008***	0.011***	0.012***	0.026***	-0.021***	-0.023***	-0.007**
	(0.001)	(0.001)	(0.002)	(0.002)	(0.005)	(0.005)	(0.005)	(0.003)
Sq.(Trend)	0.00007***	0.00005***	0.00002	-0.00003	-0.0002***	0.0002***	0.0005***	0.0002***
	(0.00001)	(0.00001)	(0.00001)	(0.00007)	(0.00003)	(0.00005)	(0.00007)	(0.00004)
Learning		-0.030***	-0.037***	-0.039***	-0.037	0.018	0.030**	-0.035***
		(0.007)	(0.007)	(0.008)	(0.035)	(0.014)	(0.014)	(0.012)
Government			-1.161***	-1.158***	-0.319***	0.245***	0.408***	-0.342***
			(0.128)	(0.129)	(0.139)	(0.054)	(0.057)	(0.058)
State FE	YES	YES	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	4059	4059	3690	3444	861	738	984	1107
Adj. R^2	0.4787	0.4798	0.4925	0.4976	0.5123	0.4886	0.4617	0.5059

Standard errors (clustered by state) are in parentheses
*** p<0.01; ** p<0.05; * p<0.10

V.2 Instrumental variable estimation

The results thus far suggest that notwithstanding the lockdown, there was an increase in the number of Covid cases across states. However, there are possible endogeneity issues since policy responses are not orthogonal to the incidence of Covid cases. To address this shortcoming, we ascertain the robustness of the results by using an Instrumental Variable (IV) approach. To be more specific, we employ the civil police per 100,000 persons as an instrument. This measure reflects the commitment of the police force towards effectively implementing the lockdown across states.

The findings are highlighted in Table 8. On average, we find that the magnitude of the coefficient on POST is uniformly higher than in the baseline. This is also manifest across states and shows that when statistically significant, the coefficient is uniformly higher as compared with the baseline. Perhaps the most interesting is with regard to states belonging to the lowest quartile of health preparedness, wherein the coefficient is double of that in the baseline, thereby indicating that it is these states that have been the most adversely impacted.

The diagnostics suggest that the instrument chosen is appropriate with a moderate first stage R-squared and the F-test rejects the null hypothesis of weak instrument. The Wooldridge score also rejects the null hypothesis of exogeneity at conventional levels. The over-identification test supports the fact that the chosen instrument is valid. All in all, our results support the findings that the imposition of lockdown did not exert any noticeable impact, on either the number of new Covid cases or even Covid-related fatalities.

Table 8. Regression results: Impact of lockdown on the number of active cases – IV results

	All states and UTs	Excluding UTs	High health index	Upper medium health index	Medium health index	Low health index
	(1)	(2)	(3)	(4)	(5)	(6)
POST (Instrumented)	**1.845*** **	**1.964*** **	**-0.400**	**1.988*** **	**-0.153**	**3.053*** **
	(0.105)	**(0.105)**	**(0.518)**	**(0.251)**	**(0.179)**	**(0.131)**
Trend	0.032***	0.033***	0.057***	-0.006	0.004	-0.015*
	(0.004)	(0.004)	(0.007)	(0.011)	(0.012)	(0.008)
Sq.(Trend)	0.00007	-0.00002	-0.0003***	0.0005***	0.0007***	0.0006***
	(0.00004)	(0.00004)	(0.00006)	(0.0001)	(0.00001)	(0.00009)
Learning	-0.118***	-0.109***	-0.096**	0.012	-0.004	-0.142***
	(0.018)	(0.018)	(0.045)	(0.030)	(0.037)	(0.033)
Government	-1.161***	0.33***	0.358***	0.644***	0.561***	0.239***
	(0.128)	(0.049)	(0.137)	(0.087)	(0.095)	(0.065)
Day FE	YES	YES	YES	YES	YES	YES
Observations	3690	3438	860	738	981	1105
Adj. R^2	0.5877	0.5894	0.4411	0.8295	0.6350	0.7362
Test of endogeneity						
Robust Chi-sq.(1)	95.67	111.62	14.13	3.44	9.72	3.69
	(0.00)	(0.00)	(0.00)	(0.06)	(0.00)	(0.07)
First stage statistics						
Partial R-sq.	0.379	0.395	0.169	0.270	0.429	0.484
F (29, 3528)	83.62	93.99	22.43	38.82	95.33	201.16
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Over-identification test						
Chi-sq. (28)	0.86	0.73	0.27	0.77	0.20	0.11
	(0.61)	(0.65)	(0.73)	(0.51)	(0.69)	(0.78)

Robust standard errors in parentheses
*** p<0.01; ** p<0.05; * p<0.10

VI. Concluding remarks

The role and relevance of lockdowns has come into sharp focus, especially in the wake of the Covid19 pandemic. With 'social distancing' emerging as the new buzzword, countries across the world – both developed and emerging – have resorted to this measure to contain the spread of the disease. The balance of evidence is mixed, with certain studies highlighting its usefulness whereas several others have questioned the efficacy of this measure.

Drawing upon the experiences of other countries, India also implemented a multi-phased lockdown during March-May 2020. As the world's largest democracy implementing such a measure, observers have been skeptical as to how far its positives outweigh the negatives (Livemint, 2020; Wright, 2020). However, the bulk of the extant evidence has been anecdotal in nature, with limited emphasis on careful empirical exploration.

Towards this end, we exploit daily data on coronavirus cases at the state level and examine its association with lockdown, after controlling for other relevant factors. The evidence suggests in spite of the lockdown, there was an increase in the number of Covid cases. This impact varied across states with different degrees of preparedness in terms of their health infrastructure. There was also an increase in Covid-related fatality rates in spite of the lockdown, although these magnitudes are lower as compared to the Covid cases. Robustness checks lend credence to these findings and furthermore indicate significant heterogeneity in response, both across states and over time, depending on several considerations that take on board the realities of the Indian scenario.

We conclude our discussion by highlighting several limitations of the analysis. For one, we consider data till end-May, although the situation is dynamic and still evolving. A more accurate measure of the pandemic would be the number of cases per 100 persons tested. However, owing to paucity of state-level data on this variable during the period considered, it is not possible to consider this in our analysis. The static nature of the population data also does not permit its use as a suitable deflator. In addition, we could not include state-level controls, although our daily data spans two consecutive financial years, simply because data for 2020-21 is not available yet. As well, the number of daily cases reported by each state is influenced by its strategy and resources devoted towards testing individuals, both symptomatic and asymptomatic. The availability of a richer dataset will facilitate a better exploration of these ideas. Third, owing to lack of adequate information, we have also not been able to suitably incorporate information on other measures that states provided and how far it played a role in containing the pandemic. To provide some examples, Assam announced a one-time financial assistance of US $2,000 for its residents stranded in foreign countries. The Government of Manipur has provided Rs. 2000 (≈ US $30) to its citizens (mostly students) stranded in other locations. The Uttar Pradesh Government also announced a financial aid of Rs.1000 (≈ US $15) each for labourer. Several other states also announced support in the form of encouragement grant for healthcare workers (Rajasthan), subsidized food to intended beneficiaries (Maharashtra, West Bengal, Uttar Pradesh), one-time cash support to specified worker categories (Madhya Pradesh, Tamil Nadu). The lack of detailed information on such expenditure including the number of persons supported impedes any meaningful analysis. Besides, the delay in obtaining reliable testing results can further influence the outcomes. These issues can be addressed in future research as more granular data becomes available, in turn, enabling to adequately inform the policy debate.

Covid Economics 37, 14 July 2020: 73-99

References

Acharya, V. and Steffen, S. (2020). The risk of being a fallen angel and the corporate dash for cash in the midst of COVID. *COVID Economics: Vetted and Real-Time Papers*, Issue 10, 44-61.

Aldasoro, I., and Ehlers, T. (2018). The geography of dollar funding for non US banks. *BIS Quarterly Review* (December), 15-26.

Arslan, Y., Drehmann M., and Hofmann, B. (2020). Central bank bond purchases in emerging market economies. *BIS Bulletin* No. 20. (June), BIS: Basel.

Bank for International Settlements (2020a). *Annual Report 2020* (June). BIS: Basel.

Bank for International Settlements (2020b). International banking and financial market developments. *BIS Quarterly Review* (March), 1-14.

Barro, R.J., Ursúa, J.F., and Weng, J. (2020). The coronavirus and the great influenza pandemic: Lessons from the "Spanish Flu" for the coronavirus's Potential effects on mortality and economic activity. NBER Working Paper No. 26866. Cambridge: MA.

Berger, A.N., Clarke, G., Cull, R., Klapper, L., and Udell, G.F. (2005). Corporate governance and bank performance: A joint analysis of the static, selection, and dynamic effects of domestic, foreign, and state ownership. *Journal of Banking and Finance* 29, 2179-2221.

Bertrand, M., Duflo, E. and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119, 249-275.

Born, B., Dietrich, A.M., and Muller, G.J. (2020). Do lockdowns work? A counterfactual for Sweden. *COVID Economics: Vetted and Real-Time Paper*, Issue 16, 1-22.

Caballero, R. J. and Simsek, A. (2020). A model of asset price spirals and aggregate demand amplification of a "Covid-19" shock. CEPR Discussion Paper No. 14627. CEPR: UK.

Dixit, A. (2020). R_0 for Covid19 research: An early estimate and policy implications. Available at <www.princeton.edu/~dixitak>

Election Commission of India (2020). Statistical Reports of General Election to State Legislative Assembly (VidhanSabha). Available at <eci.gov.in/statistical-report>

Elgin, C., Basbug, G. and Yalaman, A. (2020). Economic policy responses to a pandemic: Developing the COVID-19 economic stimulus index. *COVID Economics: Vetted and Real-Time Papers*, Issue 3, 40-53.

Government of India (2019). *Healthy States, Progressive India*. Niti Aayog, Government of India: New Delhi.

Government of India (2020). Daily bulletin on Covid19 (May 31). Available at <PIB%E2%80%99s%20Daily%20Bulletin%20on%20COVID-19-May-31-5-2020.pdf>.

Hassan, T.A, Hollander, S. van Lent, L., and Tahoun, A. (2020). Firm-level exposure to epidemic diseases: Covid-19, SARS, and H1N1. NBER Working Paper No. 26971.Cambridge: MA.

IMF (2019). *World Economic Outlook* (October). IMF: Washington DC.

IMF (2020). *World Economic Outlook* (April). IMF: Washington DC.

Jorda, O., Singh, S., and Taylor, A. (2020). Longer-run economic consequences of pandemic. NBER Working Paper No. 26934. Cambridge: MA.

Livemint (2020). Lockdown impact, as per Google. April 27. www.livemint.com/india/what-google-tells-us-about-lockdown-impact-on-indians

Mehrotra, R. (2020). Contagion and conflict: Evidence from India. *COVID Economics: Vetted and Real-Time Papers*, Issue 25, 122-140.

Meunier, T. (2020). Full lockdown policies in Western Europe countries have no evidence impact on the Covid19 pandemic. *MedRxiv* (DOI: 10.1101/2020.04.24.20078717)

Ministry of Health and Family Welfare website.

Ministry of Statistics and Programme Implementation (2020). GSVA/NSVA by economic activities. Available at <mospi.nic.in/data>

Persaud, A. (2020). Exiting the lockdown sustainably. *Economic and Political Weekly* 55
Poblete-Cazeneve, R. (2020). The great lockdown and criminal activity: Evidence from Bihar. *COVID Economics: Vetted and Real-Time Papers*, Issue 29, 141-163.
Ramelli, S., and Wagner, A.F. (2020). Feverish stock price reactions to Covid19. *Review of Corporate Finance Studies* [DOI:10.1093/rcfs/cfaa012].
Roubini, N., and Sachs, J.D. (1989). Government spending and budget deficits in the industrial economies. NBER Working Paper No. 2919. Cambridge: MA.
Stone, L. (2020). Lockdowns don't work. *Public Discourse* (April). Available at <www.thepublicdiscourse.com/2020/04/62572/>
Tellis, G.J., Sood, N. and Sood, A. (2020). Price of delay in Covid-19 lockdowns: Delays spike total cases, natural experiments reveal. SSRN paper no. 3592912.
The Lancet (2020). India under Covid19 lockdown. Vol.395, 1315.
Wolf, C., Dalal, S., DaVanzo, J., Larson, E., Akhmedjonov, E., Dago, H., Huang, M., and Montaya, S. (2011). China and India 2025: A Comparative Assessment. RAND Corporation, Santa Monica, CA. Available at: [http://www.rand.org/content/dam/rand/pubs/monographs/2011/RAND_MG1009.pdf].
World Health Organisation (2020). World Health Organisation (2020). Opening remarks at the media briefing on COVID-19 (April 13). WHO: Geneva.
Wright, R. (2020). The world's largest coronavirus lockdown is having a dramatic impact on pollution in India. CNN (April 1). Available at edition.cnn.com/2020/03/31/asia/coronavirus-lockdown-impact-pollution-india-intl-hnk/

Covid Economics 37, 14 July 2020: 73-99

Covid Economics Issue 37, 14 July 2020

Information and behavioral responses during a pandemic: Evidence from delays in Covid-19 death reports[1]

Emilio Gutierrez,[2] Adrian Rubli[3] and Tiago Tavares[4]

Date submitted: 7 July 2020; Date accepted: 13 July 2020

Information is an important policy tool for managing epidemics, but issues with data collection may hinder its effectiveness. Focusing on Covid-19 in Mexico, we ask whether delays in reporting deaths affect individuals' beliefs and behavior. Leveraging an online survey, we randomly provide information to respondents either accounting or not for delays in death reports. We find that not accounting for delays leads to a lower perceived risk of contagion and intention to comply with social distancing. An equilibrium model incorporating the endogenous behavioral response documented by our intervention illustrates the effect of reporting delays on the evolution of the epidemic.

Covid Economics 37, 14 July 2020: 100-140

1 The authors acknowledge support from the Asociación Mexicana de Cultura and the ITAM-COVID center. We thank Jose Maria Barrero, Andrew Foster, Miguel Messmacher, and participants at the ITAM Brown Bag seminar for their helpful comments. Gerardo Sánchez-Izquierdo provided outstanding research assistance. This project received approval from the ITAM Institutional Review Board. All errors are our own.

2 Instituto Tecnológico Autónomo de México (ITAM), Department of Economics.
3 ITAM, Department of Business Administration.
4 ITAM, Department of Economics and CIE.

CEPR PRESS

Covid Economics 37, 14 July 2020: 100-140

1 Introduction

Managing the current Covid-19 pandemic is particularly challenging since the usual pharmaceutical instruments, such as vaccines and antivirals, are not yet available on a large scale (Ferguson et al., 2020).[1] Given the reliance on non-pharmaceutical interventions, such as lockdowns and social distancing, it is important to understand which factors mediate their effectiveness.[2] While the literature has analyzed, among others, the roles of sociodemographic characteristics (Papageorge et al., 2020; Knittel and Ozaltun, 2020), political beliefs (Allcott et al., 2020; Baccini and Brodeur, 2020; Barrios and Hochberg, 2020), social capital (Bargain and Aminjonov, 2020; Brodeur et al., 2020; Ding et al., 2020; Durante et al., 2020), and the media (Simonov et al., 2020; Bursztyn et al., 2020), little attention has been given to the role of information.

Information provision is a powerful policy tool that allows individuals to improve their compliance with mitigating actions before, during, and after a pandemic (WHO, 2013, 2005).[3] However, issues with how information is collected may undermine its effectiveness. Specifically, delays in reporting deaths, a widely-used measure of the extent of the epidemic, may affect individuals' beliefs about the state of the pandemic and the likelihood of compliance with protective behaviors.[4] This issue may be particularly challenging in low and middle-income countries, where diminished state capacity may impede the collection of reliable and accurate real-time information.

This paper seeks to fill this gap by asking what is the effect of reporting delays on beliefs about the epidemic and behavior. To answer this question, we first fielded an online survey in Mexico where we randomized information about the epidemic. We compare respondents' beliefs regarding the severity of the epidemic and their reported intentions of complying with the government's shelter in place recommendations between groups that received different information. We then

[1]Although vaccines and medications are being developed and tested, there is no widespread available treatment or inoculation as of this writing. See, for example, https://www.nytimes.com/interactive/2020/science/coronavirus-vaccine-tracker.html?searchResultPosition=1 and https://www.nytimes.com/2020/06/24/health/coronavirus-dexamethasone.html?searchResultPosition=6, last accessed July 1, 2020.

[2]Various papers in different settings have generally found that social distancing measures have a positive impact on containing the epidemic. See, for example, Hsiang et al. (2020); Dave et al. (2020); Alexander and Karger (2020); Juranek and Zoutman (2020); Jinjarak et al. (2020); and Ferguson et al. (2020).

[3]In the context of Covid-19, it has been shown that information-focused policies in the US had the largest impact on limiting mobility (Gupta et al., 2020), and that in Italy information and communication are key tools for managing behavior and expectations (Briscese et al., 2020).

[4]According to Brodeur et al. (2020), the key information during an epidemic is the number of tests, cases, and deaths. Other factors that may lead to imperfect measures of these indicators include differential testing rates, differential times for processing tests, and undercounting of undiagnosed deaths, among others.

Covid Economics 37, 14 July 2020: 100-140

develop a simple equilibrium model that incorporates the behavior response from the information we provided in the survey to illustrate the implications for the evolution of the pandemic.

We focus on Mexico since this middle-income country provides an ideal setting for analyzing this issue for at least three reasons. First, the delays in Covid-19 death reports – defined as the time difference between when a death occurs and when it is reported in the centralized system – are particularly substantial.[5] Gutierrez et al. (2020) documents that these delays are relatively large, heterogeneous across space, and correlated with local measures of the capacity of the public healthcare system. Second, Mexican officials routinely present information on confirmed Covid-19 deaths over time, counting both by the actual date of death and the date on which the death was reported in the centralized system.[6] Lastly, the Mexican government chose a relatively lenient strategy that consisted of mostly optional lockdowns and stay-at-home recommendations, implying that determinants of individual behavior matter a lot for compliance in this setting.[7]

In light of these reporting delays, we conducted an online survey where we (i) elicited baseline priors and behaviors, (ii) provided information about the evolution of the outbreak, and (iii) recovered self-reported measures of perceptions about the risk of contagion and intended compliance with social distancing. Our randomized information treatments presented the cumulative death count either by date reported or by actual date of death from the onset of the epidemic up to 12 days before fielding the survey. Given the reporting delays, the former understated total deaths by 41% on average or up to 2,055 cumulative deaths in a given day relative to the latter.[8] Hence, while both treatments presented truthful information, one of them substantially underestimated the timing of the evolution of the epidemic given the lags in reporting.[9]

Our object of interest is the average difference between respondents that received information based on reported deaths (which we call "lagged") relative to those that saw the cumulative number

[5] Delays in reporting deaths are a well-known problem, documented across a variety of settings (AbouZahr et al., 2015; Bird, 2015).

[6] Information on deaths is presented by date occurred on this government website: https://coronavirus.gob.mx/datos/. During the nightly press conference, information is presented in both formats. See, for example, the first slide of the press conference presentation: https://presidente.gob.mx/conferencias-de-prensa-informe-diario-sobre-coronavirus-covid-19-ssa/.

[7] See, for example, https://www.informador.mx/mexico/No-habra-represion-para-detener-propagacion-del-COVID-19-reafirma-Lopez-Obrador-20200428-0039.html and https://piedepagina.mx/no-tenemos-camas-de-hospital-en-los-parques/, last accessed June 30, 2020.

[8] On average, we find that deaths occurring on a given date are reported with a delay of about six days.

[9] Given the differential reporting delays by date, our treatments also show (slightly) different shapes for the cumulative death curves. Gutierrez et al. (2020) explicitly shows how in this context the epidemic curves, as predicted by a classic epidemiological model, differ when considering either method of counting total deaths.

Covid Economics 37, 14 July 2020: 100-140

of deaths by date occurred ("unlagged" information). Our results indicate that, when unaccounted for, delays in reporting lead to a perception of lower risk of contagion, and lower self-reported intended compliance with social distancing measures.[10]

Informed by these findings and to get a better grasp on their implications for the evolution of the epidemic, we develop a model of equilibrium behavior that showcases the differential responses of agents with lagged and unlagged beliefs. We calibrate the model using specifics of our setting and show the corresponding results. We find that in a world where individuals form beliefs without lags – that is, beliefs based on data without reporting delays – they will adopt mitigating behaviors sooner than a scenario where agents' beliefs are based on lagged data, which in turn leads to a smoother epidemic curve.

The model emphasizes the importance of the type of information presented to individuals during an epidemic. Inaccurate real-time information due to reporting delays leads to individuals being slower to adopt protective behaviors and to more severe epidemic outcomes in terms of cases and deaths. Moreover, the faster speed of the epidemic induced by slower reactions will tend to generate excessive responses later on, which may exacerbate the negative economic impacts of the pandemic.

We contribute to three strands of the growing literature on the economics of Covid-19.[11] First, our paper relates to those that have explored how messages and information affect various outcomes. Akesson et al. (2020) provides different information about the infectiousness of Covid-19, finding that individuals who received the larger estimate of contagion risk were actually less likely to report complying with mitigating behaviors. Binder (2020) randomizes information about the Fed cutting interest rates, increasing consumers' optimism regarding unemployment and inflation. Coibion et al. (2020) randomizes information about different US government policies, finding a null impact on beliefs and spending plans of consumers, likely due to households' priors about the effectiveness of macroeconomic policies. While these studies focus on the effect of receiving information, our paper emphasizes the role of the accuracy of information received.

Second, we contribute to the recent literature that attempts to incorporate changes in behavior over the course of the pandemic into dynamic models that are aimed at predicting the evolution

[10] Our results from this exercise show that despite the fact that government authorities in Mexico publicly present both total deaths by date reported and by date of death, individuals form different beliefs and report different intended behaviors when presented with only one of the two. This suggests that individuals do not fully understand the implications of delays, and do not just incorporate this information when forming beliefs.

[11] See Brodeur et al. (2020) for an overview of the Covid-19 literature in economics.

Covid Economics 37, 14 July 2020: 100-140

of the epidemic over time (Fernández-Villaverde and Jones, 2020; Brotherhood et al., 2020). By explicitly incorporating the endogenous behavioral response resulting from lags in information, we illustrate how this specific channel may determine the evolution of the outbreak.

Lastly, we add to the set of papers focusing on identifying the additional restrictions and challenges that low and middle-income countries face in managing the pandemic and subsequent economic recovery. Various studies have focused on features such as the capacity of the healthcare system, poverty, inequality, and corruption (Gallego et al., 2020; Gottlieb et al., 2020; Loayza, 2020; Monroy-Gómez-Franco, 2020; Ribeiro and Leist, 2020; Walker et al., 2020). We contribute to this line of work by focusing on the potential consequences of issues in collecting reliable real-time information during the pandemic. Given the relationship between reporting delays and state capacity (Gutierrez et al., 2020), this is likely to be an issue for many other low and middle-income countries.

The rest of the paper is organized as follows. Section 2 describes the survey. Section 3 presents the effects from the treatments. Section 4 outlines the equilibrium model and discusses the results. Section 5 concludes.

2 Survey and Descriptive Statistics

In order to explore whether lags in information affect individuals' perceptions about the evolution of the pandemic and, consequently, their behavior, we conducted an online survey with a randomized informational treatment that presented the evolution of total deaths either by date reported or by actual date of death. The full survey consisted of 48 closed-response multiple choice questions and ran from May 28 to June 8, 2020.[12] We recruited participants via email and social media (namely, Twitter), and respondents were not compensated for participating.[13] Our final sample consists of 1,022 completed surveys.

The first set of questions were aimed at recovering socioeconomic characteristics of respondents, as well as their pre-intervention perceptions about the state of the Covid-19 pandemic in Mexico. We asked questions related to age, gender, state of residence, household composition, income,

[12]See online appendix B for a full translation of the survey questions and response options.

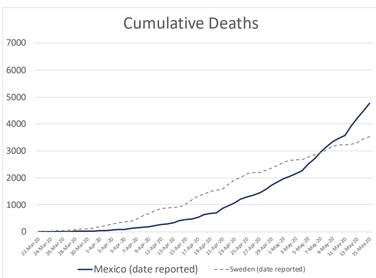
[13]Our mailing list was obtained from ITAM, and consisted of all faculty, administrative staff, and students.

Covid Economics 37, 14 July 2020: 100-140

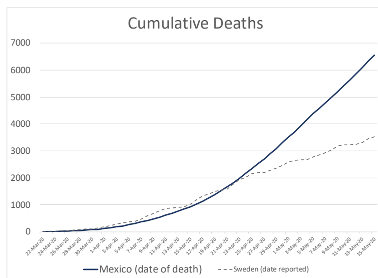
employment status, and a self-reported estimation of the number of Covid-19 cases and deaths by May 20.

After these initial questions, respondents were taken to a new screen showing (randomly) one of the two graphs depicted in Figure 1. A total of 508 participants were shown Figure 1a, which plots cumulative deaths in Mexico by date on which they were reported. The remaining 514 participants were shown Figure 1b, which instead plots cumulative deaths by actual date of occurrence. Both figures show data from March 22 to May 15, using data up to May 27.[14] Additionally, we include the cumulative number of deaths by date reported in Sweden as a reference.[15]

Figure 1:
Information Treatments in the Survey



(a) Cumulative deaths by date reported



(b) Cumulative deaths by date of occurrence

Notes: These graphs show the information treatments that we randomized in the survey. Respondents were shown these exact figures with captions translated into Spanish. Each plot shows data from March 22 to May 15, using information reported up to May 27, 2020. We include the cumulative number of deaths by date reported in Sweden as a reference. The plot on the left shows the cumulative deaths in Mexico based on the date they were reported. The plot on the right shows cumulative deaths by date on which they actually occurred.

Both figures contain truthful information as presented by government authorities themselves. Note that total deaths by date reported in Figure 1a understate total deaths by date occurred (Figure 1b) by 41% on average, with a difference of up to 2,055 deaths on May 11. While not accounting for delays implies that the evolution of Covid-19 deaths appears to be slower, it is not necessarily the case that this would induce a perception of lower risk, especially since neither

[14]This means that we allow for deaths to be reported with a lag of at most 12 days.

[15]The data for Sweden were obtained from https://ourworldindata.org/coronavirus. Sweden followed a similar strategy to Mexico, imposing relatively light restrictions (Juranek and Zoutman, 2020). The trajectory of the epidemic in Mexico had been compared to that in Sweden by Mexican authorities a few days before the survey was implemented. See, for example, https://twitter.com/HLGatell/status/1257694745322819586?s=20 and https://www.milenio.com/politica/ya-aplanamos-la-curva-lopez-gatell, last accessed June 29, 2020.

Covid Economics 37, 14 July 2020: 100-140

plot shows a clear change in the growth rate of deaths. If sophisticated agents were aware of this problem, then they could assess the true risk even when receiving information with delays.

After presenting the corresponding figure, participants answered additional questions regarding beliefs about the risk of contagion associated with attending social gatherings, the expected number of total Covid-19 cases and deaths over the whole epidemic outbreak, and the number of times they expected to leave their home the week after they participated in the survey, as well as four weeks later.

Table 1 tests for balance in observable characteristics between respondents in each of these treatment groups. Columns 1 and 2 present means for the sample that was shown the graph with cumulative deaths by date reported and by date occurred, respectively. Column 3 shows the corresponding difference in means. It is worth highlighting that, due to the nature of the survey conducted, the characteristics of participants suggest they belong to a relatively young, educated, and high-income group in Mexico. More than 78 percent of them live in Mexico City, and more than half of them live in a house with a yard. Evidently, this implies that none of our results can be used to infer the distribution of beliefs and behavior in the general Mexican population.[16] However, given the very small differences in observables between our two treatment groups, we can confidently interpret the results below as the impact of the information provided on the different outcomes.

3 Empirical Strategy and Results

3.1 Empirical Strategy

We explore the impact of the information provided in the survey on four different measures of perceptions about the risk of contagion and the evolution of the epidemic. Specifically, we focus on participants' responses to questions regarding the perceived risk of attending a gathering of 100 people in the week following the survey and in four weeks, as well as the predicted number of total Covid-19 cases and deaths over the course of the current outbreak. In terms of behavior, we focus on questions regarding the number of times respondents expect to leave their home in the

[16]Note also that our sample does not have enough variation to weight it so that it is representative of the entire population in Mexico.

CEPR PRESS

Table 1:
Balance Table for Survey Covariates

	Informational treatments		
	Deaths by date reported	Deaths by date occurred	Difference in means
Female	0.496	0.490	-0.006
	(0.500)	(0.500)	(0.031)
Ages 18-22	0.321	0.383	0.062**
	(0.467)	(0.487)	(0.030)
Ages 23-29	0.274	0.253	-0.021
	(0.446)	(0.435)	(0.028)
Ages 30-49	0.230	0.216	-0.014
	(0.421)	(0.412)	(0.026)
Ages 50+	0.175	0.148	-0.027
	(0.381)	(0.355)	(0.023)
Works	0.409	0.329	-0.081***
	(0.492)	(0.470)	(0.030)
Attends school	0.368	0.416	0.048
	(0.483)	(0.493)	(0.031)
Works and attends school	0.157	0.158	0.000
	(0.365)	(0.365)	(0.023)
Other occupation/employment status	0.065	0.097	0.032*
	(0.247)	(0.297)	(0.017)
Lives in Mexico City	0.776	0.753	-0.023
	(0.418)	(0.432)	(0.027)
Lives in apartment	0.343	0.385	0.043
	(0.475)	(0.487)	(0.030)
Lives in house, no yard	0.124	0.117	-0.007
	(0.330)	(0.321)	(0.020)
Lives in house with yard	0.533	0.498	-0.035
	(0.499)	(0.500)	(0.031)
Household size: 1-2	0.232	0.251	0.019
	(0.423)	(0.434)	(0.027)
Household size: 3	0.207	0.245	0.038
	(0.405)	(0.431)	(0.026)
Household size: 4	0.252	0.226	-0.026
	(0.435)	(0.418)	(0.027)
Household size: 5+	0.561	0.504	-0.057*
	(0.497)	(0.500)	(0.031)
Has HH members over 70 years old	0.159	0.080	-0.080***
	(0.366)	(0.271)	(0.020)
Has HH members 60-70 years old	0.215	0.202	-0.012
	(0.411)	(0.402)	(0.025)
Has HH members 50-60 years old	0.461	0.471	0.010
	(0.499)	(0.500)	(0.031)
Does not seek healthcare when sick	0.140	0.154	0.014
	(0.347)	(0.361)	(0.022)
Self-medicates when sick	0.386	0.381	-0.005
	(0.487)	(0.486)	(0.030)
Observations	508	514	1,022

Notes: This table shows means and standard deviations for a series of covariates asked in the survey before the informational treatment was presented. We show statistics separately for each informational treatment, as well as the difference in the means. Stars denote significance from a difference in means test.
*** p<0.01, ** p<0.05, * p<0.1

Covid Economics 37, 14 July 2020: 100-140

CEPR PRESS

week following the survey and four weeks later. Figures A2 and A3 in the online appendix show histograms of the responses to these six questions.

We estimate the following equation to measure differences in perceptions and behavior between our treatment groups:

$$y_i = \beta_0 + \beta_1 \times [\text{Info By Date Occurred}]_i + \Psi \mathbf{X_i} + \varepsilon_i \tag{1}$$

where y_i is the outcome variable for respondent i, β_0 is a constant, $[\text{Info By Date Occurred}]_i$ is a zero-one indicator for having received the informational treatment that displayed cumulative deaths by actual date of death, $\mathbf{X_i}$ is a vector of observable characteristics as listed in Table 1, and ε_i is the error term. Our estimate of interest corresponds to β_1, which measures the average difference in the outcome variable for survey respondents that were shown the cumulative death toll by date of occurrence with respect to those who were shown the information by date reported.

For simplicity, we construct binary measures of our risk perception variables, where a value of one denotes a high risk perception. Specifically, we assign a value of one if respondents considered the risk of contagion at a social event to be high or extremely high. We also assigned a value of one if participants responded that they expected the total number of Covid-19 cases and deaths to exceed 500,000 and 50,000, respectively. For the social distancing outcomes, our binary variables take a value of one if respondents expected to leave their house three or more times.[17]

3.2 Results

The direction in which the treatment may affect beliefs about the evolution of the epidemic depends on respondents' priors. Before the treatment, survey participants were asked to report their knowledge about the total number of Covid-19 cases recorded in Mexico by May 20, a full week before the launch of the survey. We use the responses to this question to stratify the sample into a low and high prior group.[18] The low prior subsample consists of those that reported that the total

[17]For completeness, we show similar results in online appendix Figures A4 - A6 using indicators for each of the possible responses.

[18]Tables A1 and A2 in the online appendix show balance tables separately for the low and high priors subsamples.

Covid Economics 37, 14 July 2020: 100-140

Covid Economics 37, 14 July 2020: 100-140

number of Covid-19 cases was lower than 50,000 (47.7 percent of the full sample), while the high prior group are those that reported over 50,000 cases.[19]

We present all our results for the full sample and separately for these two subgroups. If our informational treatment is shifting beliefs about the epidemiological curve, then we would expect to see stronger and larger effects among the low prior group, as they are the ones that would update their priors upward.

Table 2 shows our main results. Panel A corresponds to the perceived risk of contagion (next week in columns 1-3, and in four weeks in columns 4-6). Panel B corresponds to the expected number of total cases (columns 1-3) and total deaths (columns 4-6). And Panel C corresponds to our measures of compliance with social distancing (next week in columns 1-3, and in four weeks in columns 4-6). Throughout Table 2, columns 1 and 4 present results for the full sample, columns 2 and 5 restrict to the low prior subsample, and columns 3 and 6 focus on the high prior subsample.

For every risk measure, presenting cumulative deaths by actual date of occurrence seems to shift beliefs towards a higher risk level. For example, for individuals that were shown cumulative deaths by date of death, the fraction of respondents considering that the risk of contagion at a social event is high or extremely high is 0.033 points higher, both for assessments next week and in four weeks. However, only the former is statistically significant. We find similar patterns for respondents predicting a high number of Covid-19 cases and deaths. These differences are larger, as expected, in the low prior sample.

The results in Panel C are consistent with the information on cumulative deaths presented by actual date of death having an effect on expectations to comply with social distancing measures. For expected behavior four weeks after the survey, having been shown the graph by date of occurrence is associated with a significant decrease in the number of times people expect to leave their homes. Once again, the effect is concentrated in the low prior sample.

Notwithstanding the limited statistical power due to the small sample size of the survey and the relatively small differences in the information provided to each group, we interpret the results presented in Table 2 as evidence that the delays with which deaths are reported are very likely to

[19]The true number reported in the nightly press conference on May 20 was 56,594 cumulative cases in the country (see https://twitter.com/HLGatell/status/1263264663283908609?s=20, last accessed June 29, 2020). A histogram with the distribution of the responses to this question is presented in online appendix Figure A1. Stratifying the sample based on individuals' prior regarding total reported deaths by May 20 yields similar results (available upon request).

CEPR PRESS

Covid Economics 37, 14 July 2020: 100-140

Table 2:
Estimates of Informational Treatments on Risk Perceptions and Expected Behavior

	(1) Full sample	(2) Low prior	(3) High prior	(4) Full sample	(5) Low prior	(6) High prior
Panel A: Risk of contagion	High perceived risk of contagion at gathering with 100 people					
	Next week			In 4 weeks		
Information by date occurred	0.0334*	0.0698**	-0.0043	0.0331	0.0737*	0.0047
	(0.019)	(0.027)	(0.026)	(0.026)	(0.039)	(0.035)
Observations	1,022	488	534	1,022	488	534
R-squared	0.055	0.074	0.081	0.028	0.051	0.044
Mean dependent variable	0.90	0.90	0.90	0.79	0.77	0.81
Panel B: Expected toll	High subjective prediction of the full toll of the epidemic					
	>500,000 cases			>50,000 deaths		
Information by date occurred	0.0549*	0.0728	0.0515	0.0383	0.0752*	0.0046
	(0.031)	(0.044)	(0.044)	(0.031)	(0.046)	(0.044)
Observations	1,022	488	534	1,022	488	534
R-squared	0.016	0.054	0.032	0.015	0.032	0.020
Mean dependent variable	0.36	0.34	0.39	0.41	0.41	0.42
Panel C: Social distancing	High number of times expected to leave the house (3+)					
	Next week			In 4 weeks		
Information by date occurred	0.0004	-0.0326*	0.0298*	-0.0553**	-0.0893**	-0.0243
	(0.012)	(0.017)	(0.017)	(0.025)	(0.036)	(0.035)
Observations	1,022	488	534	1,022	488	534
R-squared	0.660	0.658	0.680	0.346	0.354	0.356
Mean dependent variable	0.12	0.11	0.12	0.35	0.33	0.37

Notes: This table presents the results from estimating equation 1. Each panel corresponds to two different questions in the survey converted to a binary measure (see text for details). Columns 1 and 4 show results for the full sample. Columns 2, 5, 3 and 6 stratify the sample by respondents' prior on their knowledge of the number of Covid-19 cases in Mexico up to May 20 into low and high reported cases, respectively. The estimates are the average difference between the responses in the treatment group that received information based on the actual date of death relative to information based on date of reports. Robust standard errors are reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Covid Economics 37, 14 July 2020: 100-140

affect perceptions about the state of the pandemic in Mexico and, consequently, compliance with social distancing measures. These findings also suggest that the individuals surveyed, despite being a selected sample of higher-income respondents and despite the government providing information on both deaths by date reported and by date occurred, do not fully incorporate reporting delays when forming beliefs about the epidemic, and are thus unlikely to be fully aware of these delays. We proceed by taking these insights and incorporating them into an equilibrium model.

4 Model of Equilibrium Behavior during an Epidemic

The previous section documented that reporting delays in the information presented to individuals may affect their beliefs and behavior during an epidemic. In this section, we present an equilibrium model to illustrate how the insights from our online survey may affect the evolution of the epidemic through the endogenous behavioral response of agents. The model's simplicity highlights the potential importance of lagged information due to reporting delays for agents' behavior, but standard extensions could be included.[20]

Following the seminal work of Kermack and McKendrick (1927), agents in the model are compartmentalized into different health states corresponding to susceptible, infected, and recovered. In the specific context of Covid-19, Fernández-Villaverde and Jones (2020) highlights the importance of including an additional state, where recovering agents cannot infect others, and of assuming a time dependent exogenous contact rate to allow for a better fit of the data. Our model also includes a similar state variable, but endogenizes the contact rate through behavioral reactions to the epidemic by introducing an equilibrium concept similar to Brotherhood et al. (2020).

4.1 General Setup

We set up the model in discrete time, where each period corresponds to one day. The economy is populated with a continuum of ex-ante identical agents. Agents can spend time outside or at home. Given an outbreak of Covid-19, let j be an agent's health status. Define $j = s$ (**susceptible**) as the state in which the agent has never been infected. Only susceptible agents spending time outside

[20] Extensions may include features that allow the study of macroeconomic implications (Eichenbaum et al., 2020a), saving decisions (Kaplan et al., 2020), non-pharmacy initiatives (Brotherhood et al., 2020), testing, quarantine, the introduction of vaccines (Eichenbaum et al., 2020b), optimal lockdown policies (Alvarez et al., 2020), and age or asset heterogeneity (Glover et al., 2020; Acemoglu et al., 2020), among others.

Covid Economics 37, 14 July 2020: 100-140

are at risk of becoming infected, in which case the state changes to $j = i$ (**infected**). During this state, agents can infect others with a uniform mixing contact rate. This state subsequently changes to a non-infectious status $j = r$ (**resolving**) with probability γ. The resolving state ends with probability θ, where a share δ of agents die, while the remaining share fully recovers. We assume a recovered agent becomes immune with status $j = c$ (**recovered**). To summarize, $j = \{s, i, r, c\}$ and the future is discounted at a rate β.

Each agent is also endowed with a single unit of time every period, which is divided into n hours outside the home and d at home, such that $1 = n + d$. We assume that the flow utility of being dead is normalized to zero, while being alive generates flow utility from spending hours outside and at home:

$$u(n) = \log(n) + \lambda_d \log(1 - n) + b$$

where $\lambda_d > 0$ is a preference weight for hours spent at home, and b is a positive constant that captures the premium of being alive, so that the flow utility from being alive is larger than zero for reasonable values of n.[21]

Infections and Beliefs. The probability of becoming infected for susceptible agents is assumed to be proportional to the time spent outside the home n and an aggregate transmission risk Π_t:

$$\pi(n, \Pi_t) = n\Pi_t$$

We allow for misperceived beliefs about transmission risk by defining perceived risk as $\widetilde{\Pi}_t$, which may differ from the true Π_t. In the application below, we focus on agents that form beliefs based on reports containing lagged information due to reporting delays. Hence, for simplicity, we assume:

$$\widetilde{\Pi}_t = \Pi_{t-k}, \quad k \geq 0 \tag{2}$$

[21] In practice, we are implicitly assuming that $\underline{n} < n < \bar{n}$, so that b must be larger than $-(\log(n) + \lambda_d \log(1-n))$. Given the calibration described later in Section 4.3, simulation results never hit these bounds.

Covid Economics 37, 14 July 2020: 100-140

meaning that agents form subjective beliefs of contagion risk at time t based on information about the epidemic from k days before. In other words, if $k > 0$, beliefs are formed with a k-day lag, while if $k = 0$, beliefs form from contemporaneous (unlagged) information.[22]

Value Functions. Given the structure described above, we now specify value functions for agents in each of the different states during the course of the pandemic.

For susceptible agents, the value function $V(s,t)$ at time t is given by:

$$V(s,t) = \max_{n\in(0,1)} \left\{u(n) + \beta\left[1-\pi\left(n,\widetilde{\Pi}_t\right)\right]V(s,t+1) + \beta\pi\left(n,\widetilde{\Pi}_t\right)V(i)\right\} \qquad \text{(value susceptible)}$$

where the value of being infected $V(i)$ evolves according to:

$$V(i) = \max_{n\in(0,1)} \left\{u(n) + \beta\left[\gamma V(r) + (1-\gamma)V(i)\right]\right\} \qquad \text{(value infected)}$$

For individuals that start resolving the disease, we assume that they cannot work due to the illness and flow utility is therefore zero. Thus their total value $V(r)$ is given by:

$$V(r) = \beta(1-\theta)V(r) + \beta\theta(1-\delta)V(c) \qquad \text{(value resolving)}$$

Finally, the value for individuals that fully recover is:

$$V(c) = \max_{n\in(0,1)} \left\{u(n) + \beta V(c)\right\} \qquad \text{(value recoverd)}$$

Laws of Motion. Let $n(j,t)$ be the optimal choice of hours spent outside the home for states $j = s,i,c$. Then the following laws of motion characterize the evolution of the population mass in

[22] When $k > 0$, this setup is similar to the notion of adaptive expectations as introduced by Cagan (1956) or Friedman (1957), or a model in which agents herd on epidemic information provided by the government, which may be perceived as better quality (Banerjee, 1992). If $k = 0$, the model is consistent with the concept of rational expectations as presented in Muth (1961).

Covid Economics 37, 14 July 2020: 100-140

the different states of the epidemic:

$$M_{t+1}(s) = M_t(s)(1 - \pi(n(s,t), \Pi_t)) \qquad \text{(mass susceptible)}$$

$$M_{t+1}(i) = M_t(s)\pi(n(s,t), \Pi_t) + M_t(i)(1-\gamma) \qquad \text{(mass infected)}$$

$$M_{t+1}(r) = M_t(i)\gamma + M_t(r)(1-\theta) \qquad \text{(mass resolving)}$$

$$M_{t+1}(c) = M_t(r)\theta(1-\delta) + M_t(c) \qquad \text{(mass recovered)}$$

We can also define additional accounting variables, such as the measure of total Covid-19 deaths:

$$M_{t+1}^{deaths} = M_t^{deaths} + \theta\delta M_t(r)$$

Aggregate Probability of Infection. We assume that the instantaneous rate of infection within a period is given by:

$$\widehat{\Pi}_t = \Pi_0 M_t(i)\, n(i,t)$$

That is, within a period a susceptible agent can have multiple encounters with infected agents given by rate $\widehat{\Pi}_t$, resulting in contagion. Since it only takes one infection to change the status, the probability of an infectious contact within a period becomes:

$$\Pi_t = 1 - \exp\left(-\widehat{\Pi}_t\right) \tag{3}$$

4.2 Definition of the Equilibrium

A *belief-induced equilibrium* in this economy with an initial mass of agents $M_0(j)$ for each $j \in \{s,i,r,c\}$ consists of a sequence of infection rates $\{\Pi_t\}_{t=0}^{\infty}$ and hour allocations $\{n(j,t)\}_{t=0}^{\infty}$, such that:

1. given $\{\Pi_t\}_{t=0}^{\infty}$, induced expectations $\widetilde{\Pi}_t$ are formed from equation 2 with $n(j,t)$ solving the values in equations (value susceptible) to (value recoverd);

2. given $\{n(j,t)\}_{t=0}^{\infty}$ and initial $M_0(j)$ for $j \in \{s,i,r,c\}$, the resulting laws of motion from equations (mass susceptible) to (mass recovered) are consistent with $\{\Pi_t\}_{t=0}^{\infty}$ given the aggregate probability of infection in equation 3.

CEPR PRESS

Covid Economics 37, 14 July 2020: 100-140

4.3 Model Calibration and Results

Note that the stylized structure of the model allows for a simple evaluation of the behavioral impact of lagged information due to reporting delays if agents form beliefs as described in equation 2.[23] Given our survey and recognizing that deaths in Mexico are reported with a 6-day delay on average, we calibrate the model by assuming that delays in reporting induce a belief that lags by $k = 6$ days from the correct one.

The remaining model parameters are summarized in Table 3. The discount factor $\beta = 0.98^{1/365}$ is set to capture a 2% annual interest rate. Parameters associated with infectiousness, the probability of resolving, and the death rate for Covid-19 are calibrated to target standard findings from the medical literature as documented in Bar-On et al. (2020). As for the remaining parameters, we capture certain features of the Mexican economy. We assume that the initial population is 120 million and the number of infected individuals at time zero are 1,200 or 0.001% of the population. We use Mexican time use surveys to calibrate the parameter λ_d by targeting an expenditure of 36% of available hours in activities outside the home prior to the epidemic. The utility function parameter b captures a drop in activity outside the home during the epidemic of 45% as suggested by evidence from Google Mobility data. Lastly, the baseline contagion rate parameter Π_0 is set to generate a basic reproduction number of 1.84 as documented by Marioli et al. (2020).[24]

Table 3:
Baseline Calibration of the Behavioral Contagion Model

Parameter in the model		Value	Target
Discount factor	β	$0.98^{1/365}$	Standard 2% yearly interest rate
Probability of infection	γ	0.166	6 days while infectious (Bar-On et al., 2020)
Resolving probability	θ	0.1	16 days to clear Covid-19 (Bar-On et al., 2020)
Death rate	δ	0.008	From medical literature (Bar-On et al., 2020)
Initial mass of infected	$M_0(i)$	0.001%	1,200 individuals in Mexico
Preference for staying home	λ_p	1.77	36% of hours spent outside home (ENUT)
Preference for staying alive	b	7.4	45% drop in outside home activity (Google Mobility Data)
Baseline contagion rate	Π_0	2.353	Basic reproduction number $R_0 = 1.84$ (Marioli et al., 2020)

Notes: This table shows the values for the parameters used to calibrate the model. ENUT refers to the Mexican Time Use Survey for 2014.

[23] Online appendix C provides a brief discussion of how belief formation about the pandemic affects agents' behavior within the model.

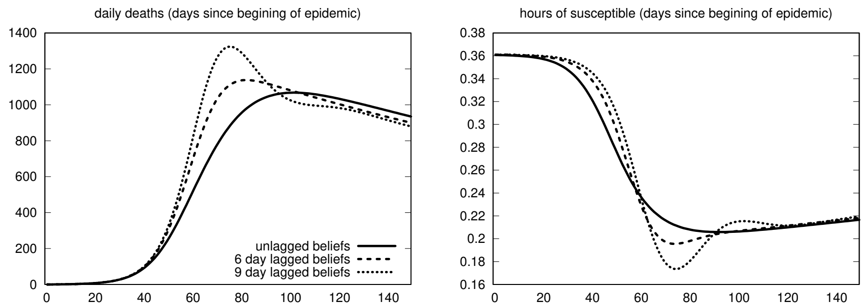
[24] Online appendix C provides additional details on the calibration used in the model simulation.

Covid Economics 37, 14 July 2020: 100-140

We then simulate the model by computing the solution to the equilibrium defined in Section 4.2. We compare the baseline results where $k = 6$ to a case with no delays $k = 0$ (equivalent to unlagged beliefs) and also to a case with more extreme delays with $k = 9$. The results of this exercise are summarized in Figure 2.[25]

The left panel in Figure 2 shows the results for the evolution of daily Covid-19 deaths under each scenario. With perfect information ($k = 0$), it would take 102 days to reach the maximum number of 1,069 daily deaths. If instead agents face a 6-day lag in information about the true contagion rate ($k = 6$), the maximum number of daily deaths would increase to 1,138 and would occur 19 days earlier on day 83 of the epidemic. Lastly, with an even longer delay of 9 days, the maximum number of daily deaths would be 1,324 and would occur on day 76 of the epidemic.

Figure 2:
Simulation Results of Behavioral Model



Notes: These graphs show the simulation results from the model by computing the equilibrium as defined in Section 4.2. We show results considering a 6-day delay in the formation of beliefs consistent with reporting delays, zero delay corresponding to contemporaneous beliefs, and a more extreme case of a 9-day lag. The plot on the left shows the number of daily deaths from Covid-19 over time from the onset of the epidemic. The plot on the right shows the percentage of hours in a day that susceptible individuals (those who have never been infected) spend outside the home since the beginning of the epidemic. The baseline percentage of hours away from home is 36%.

The force behind these different dynamics can be understood by looking at how time outside the home evolves in each scenario. This is shown in the right panel of Figure 2. We highlight three important results in this graph. First, delays slow down the endogenous adjustment in hours spent outside the home as a reaction to the risk of being infected. For instance, on day 45 of the epidemic, hours outside would fall by 15.8% (5.7 percentage points, pp) for $k = 0$ relative to only 11.1% (4 pp) for $k = 6$ and 7.2% (2.6 pp) for $k = 9$.

[25] For additional results and robustness of the model with respect to parameters, see online appendix C.

CEPR PRESS

Second, since agents are slower to adjust in the presence of delays, the overall probability of being infected is larger at the peak of the epidemic. We find a 0.66% probability for $k = 6$ relative to 0.58% for $k = 0$. Under the extreme $k = 9$ scenario, the maximum daily probability of infection reaches 0.81%.

Lastly, the change in behavior over time is also considerably less smooth in the presence of delays. Total hours outside the home for susceptible agents fall from the 36% baseline to a minimum of 21.0% in the model with $k = 0$ in comparison to 19.8% for $k = 6$. In the more extreme case of $k = 9$, hours reach a minimum of 17.8%.

This exercise emphasizes the importance of clear and swift communication during an epidemic. Governments aiming to control the epidemic may care about behavioral responses, which could be affected if agents do not have a correct understanding of the risk of infection due to issues with the type of information available. In particular, reactions may be too slow, thus inducing harsher epidemic outcomes in terms of the daily number of infected individuals and deaths. This may be especially important in the presence of hospital capacity constraints.[26] Moreover, the faster progression of the epidemic when agents are slow to react will tend to generate excessive responses later in the pandemic, thus adding to the economic downturn that would likely be large even in a scenario with fully accurate real-time information.

5 Conclusion

Given the reliance on non-pharmaceutical interventions like social distancing, effecting change in individual behavior is paramount for managing the Covid-19 pandemic. Providing information on the state of the epidemic is an important policy tool, but its effectiveness may be hampered due to data collection issues. In particular, this paper analyzes how individual beliefs and behavior are affected by differing information due to lags in the cumulative death count from reporting delays.

Our randomized informational treatments in an online survey in Mexico show that participants that were shown total deaths over time by date reported – that is, a measure that understates the true death toll because of large reporting delays – were more likely to perceive a lower risk of contagion and to report lower intentions of complying with stay-at-home recommendations.

[26]For example, Gutierrez and Rubli (2020) show a strong relationship between hospital capacity and increases in in-hospital mortality during the 2009 H1N1 epidemic in Mexico.

Covid Economics 37, 14 July 2020: 100-140

CEPR PRESS

We then develop an equilibrium behavioral model that shows that, if individuals receive lagged information because of reporting delays, they are slower to modify their risky behavior, which in turn leads to more severe epidemic outcomes.

Delays in death reports are a common feature across settings, but are likely to be exacerbated by the low state capacity in low and middle-income countries. Hence, our results suggest that data collection issues in these contexts may magnify the extent of the epidemic, adding to the particular challenges facing these countries. Furthermore, other issues with information linked to differential counting of tests, cases, and deaths across and within countries may also affect individual behavior via their effect on perceptions and beliefs, which in turn may limit effective management of the Covid-19 pandemic.

From a policy perspective, our results highlight the importance of collecting and disseminating reliable real-time information on the state of the epidemic, or at least being upfront and clear about the drawbacks of the available data. Evidently, improving data collection in real-time is costly, and scarce resources may be better spent on other mitigation strategies. However, low-cost measures, such as clearly explaining delays and developing correction factors to generate an estimate of the true death toll, could alleviate these shortcomings.

Covid Economics 37, 14 July 2020: 100-140

Covid Economics 37, 14 July 2020: 100-140

References

AbouZahr, C., D. De Savigny, L. Mikkelsen, P. W. Setel, R. Lozano, and A. D. Lopez (2015). Towards universal civil registration and vital statistics systems: the time is now. *The Lancet 386*(10001), 1407–1418.

Acemoglu, D., V. Chernozhukov, I. Werning, and M. D. Whinston (2020). A multi-risk SIR model with optimally targeted lockdown. Technical report, National Bureau of Economic Research.

Akesson, J., S. Ashworth-Hayes, R. Hahn, R. D. Metcalfe, and I. Rasooly (2020). Fatalism, beliefs, and behaviors during the COVID-19 pandemic. Technical report, National Bureau of Economic Research.

Alexander, D. and E. Karger (2020). Do stay-at-home orders cause people to stay at home? Effects of stay-at-home orders on consumer behavior.

Allcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Y. Yang (2020). Polarization and public health: Partisan differences in social distancing during the Coronavirus pandemic. *NBER Working Paper* (w26946).

Alvarez, F. E., D. Argente, and F. Lippi (2020). A simple planning problem for covid-19 lockdown. *Covid Economics 1*(14), 1–32.

Baccini, L. and A. Brodeur (2020). Explaining governors' response to the COVID-19 pandemic in the United States.

Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics 107*(3), 797–817.

Bar-On, Y. M., A. Flamholz, R. Phillips, and R. Milo (2020). Science forum: SARS-CoV-2 (COVID-19) by the numbers. *Elife 9*, e57309.

Bargain, O. and U. Aminjonov (2020). Trust and compliance to public health policies in times of COVID-19.

Barrios, J. M. and Y. Hochberg (2020). Risk perception through the lens of politics in the time of the covid-19 pandemic. Technical report, National Bureau of Economic Research.

Binder, C. (2020). Coronavirus fears and macroeconomic expectations. *Review of Economics and Statistics*, 1–27.

Bird, S. M. (2015). End late registration of fact-of-death in England and Wales. *The Lancet 385*(9980), 1830–1831.

Briscese, G., N. Lacetera, M. Macis, and M. Tonin (2020). Compliance with covid-19 social-distancing measures in Italy: the role of expectations and duration. Technical report, National Bureau of Economic Research.

Brodeur, A., D. M. Gray, A. Islam, S. J. Bhuiyan, et al. (2020). A literature review of the economics of COVID-19. Technical report, Institute of Labor Economics (IZA).

Brodeur, A., I. Grigoryeva, and L. Kattan (2020). Stay-at-home orders, social distancing and trust.

Brotherhood, L., P. Kircher, C. Santos, and M. Tertilt (2020). An economic model of the Covid-19 epidemic: The importance of testing and age-specific policies.

Bursztyn, L., A. Rao, C. Roth, and D. Yanagizawa-Drott (2020). Misinformation during a pandemic. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2020-44).

Cagan, P. (1956). The monetary dynamics of inflation. *Studies in the Quantity. Theory of Money, M. Friedman (ed.), Chicago: University of Chicago Press*.

Coibion, O., Y. Gorodnichenko, and M. Weber (2020). Does policy communication during COVID work? *Covid Economics 1*(29), 1–49.

Dave, D. M., A. I. Friedson, K. Matsuzawa, J. J. Sabia, and S. Safford (2020). Were urban cowboys enough to control COVID-19? Local shelter-in-place orders and coronavirus case growth. Technical report, National Bureau of Economic Research.

Ding, W., R. Levine, C. Lin, and W. Xie (2020). Social distancing and social capital: Why US counties respond differently to COVID-19. *Available at SSRN 3624495*.

Durante, R., G. Gulino, et al. (2020). Asocial capital: Civic culture and social distancing during COVID-19.

Covid Economics 37, 14 July 2020: 100-140

Eichenbaum, M. S., S. Rebelo, and M. Trabandt (2020a). The macroeconomics of epidemics. Technical report, National Bureau of Economic Research.

Eichenbaum, M. S., S. Rebelo, and M. Trabandt (2020b). The macroeconomics of testing and quarantining. Technical report, National Bureau of Economic Research.

Ferguson, N., D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg, et al. (2020). Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID 19 mortality and healthcare demand.

Fernández-Villaverde, J. and C. I. Jones (2020). Estimating and simulating a SIRD model of COVID-19 for many countries, states, and cities. Technical report, National Bureau of Economic Research.

Friedman, M. (1957). Introduction to "A theory of the consumption function". In *A theory of the consumption function*, pp. 1–6. princeton university press.

Gallego, J. A., M. Prem, and J. F. Vargas (2020). Corruption in the times of pandemia. *Available at SSRN 3600572*.

Glover, A., J. Heathcote, D. Krueger, and J.-V. Ríos-Rull (2020). Health versus wealth: On the distributional effects of controlling a pandemic. Technical report, National Bureau of Economic Research.

Gottlieb, C., J. Grobovšek, and M. Poschke (2020). Working from home across countries. *Covid Economics 1*(8), 71–91.

Gupta, S., T. D. Nguyen, F. L. Rojas, S. Raman, B. Lee, A. Bento, K. I. Simon, and C. Wing (2020). Tracking public and private response to the COVID-19 epidemic: Evidence from state and local government actions. Technical report, National Bureau of Economic Research.

Gutierrez, E. and A. Rubli (2020). Shocks to hospital occupancy and mortality: Evidence from the 2009 H1N1 pandemic. Technical report, Working paper.

Covid Economics 37, 14 July 2020: 100-140

Gutierrez, E., A. Rubli, and T. Tavares (2020). Delays in death reports and their implications for tracking the evolution of COVID-19. *Covid Economics 1*(34), 116–144.

Hsiang, S., D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, A. Hultgren, L. Y. Huang, E. Krasovich, et al. (2020). The effect of large-scale anti-contagion policies on the coronavirus (covid-19) pandemic. *MedRxiv*.

Jinjarak, Y., R. Ahmed, S. Nair-Desai, W. Xin, and J. Aizenman (2020). Accounting for global COVID-19 diffusion patterns, January-April 2020. Technical report, National Bureau of Economic Research.

Juranek, S. and F. Zoutman (2020). The effect of social distancing measures on the demand for intensive care: Evidence on covid-19 in Scandinavia.

Kaplan, G., B. Moll, and G. Violante (2020). Pandemics according to HANK. *Powerpoint presentation, LSE 31*.

Kermack, W. O. and A. G. McKendrick (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character 115*(772), 700–721.

Knittel, C. R. and B. Ozaltun (2020). What does and does not correlate with COVID-19 death rates. *medRxiv*.

Loayza, N. V. (2020). Costs and trade-offs in the fight against the Covid-19 pandemic: A developing country perspective.

Marioli, F. A., F. Bullano, S. Kučinskas, and C. Rondón-Moreno (2020). Tracking R of COVID-19: A new real-time estimation using the Kalman filter. *medRxiv*.

Monroy-Gómez-Franco, L. (2020). ¿Quién puede trabajar desde casa? Evidencia desde México.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: Journal of the Econometric Society*, 315–335.

Covid Economics 37, 14 July 2020: 100-140

Papageorge, N. W., M. V. Zahn, M. Belot, E. van den Broek-Altenburg, S. Choi, J. C. Jamison, E. Tripodi, et al. (2020). Socio-demographic factors associated with self-protecting behavior during the COVID-19 pandemic. Technical report, Institute of Labor Economics (IZA).

Ribeiro, F. and A. Leist (2020). Who is going to pay the price of Covid-19? Reflections about an unequal Brazil. *International Journal for Equity in Health 19*, 1–3.

Simonov, A., S. K. Sacher, J.-P. H. Dubé, and S. Biswas (2020). The persuasive effect of fox news: non-compliance with social distancing during the covid-19 pandemic. Technical report, National Bureau of Economic Research.

Walker, P. G., C. Whittaker, O. J. Watson, M. Baguelin, P. Winskill, A. Hamlet, B. A. Djafaara, Z. Cucunubá, D. O. Mesa, W. Green, et al. (2020). The impact of Covid-19 and strategies for mitigation and suppression in low-and middle-income countries. *Science*.

WHO (2005). WHO outbreak communication guidelines. Technical report, World Health Organization.

WHO (2013). Pandemic influenza risk management: WHO interim guidance.

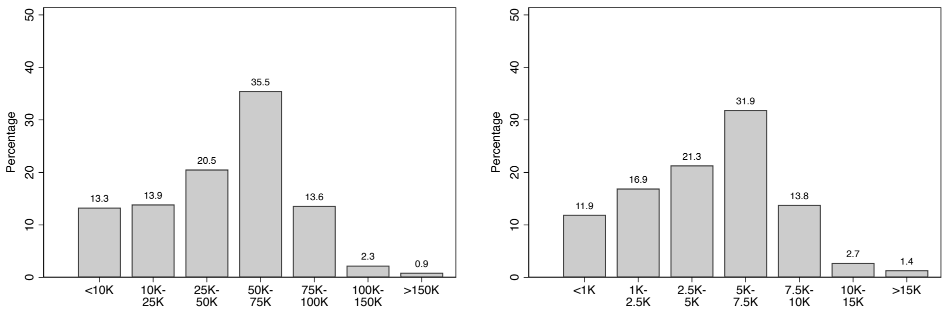
Covid Economics 37, 14 July 2020: 100-140

Covid Economics 37, 14 July 2020: 100-140

Appendix for Online Publication

A Supplementary Figures and Tables for the Survey

Figure A1:
Histograms of Prior Beliefs on Total Cases and Deaths

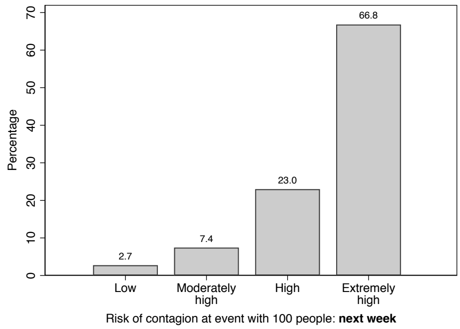


(a) Beliefs on total cases (b) Beliefs on total deaths

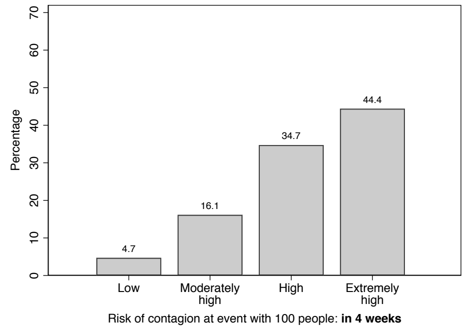
Notes: These graphs show histograms for the questions eliciting beliefs about total cases and total deaths up to May 20 (one week prior to when the survey was launched) for our full sample of participants. Each plot shows the percentage of total respondents that chose each of the answers. The actual number of cumulative cases reported by the government on May 20 was 56,594, and the cumulative deaths reported were 6,090 (see https://twitter.com/HLGatell/status/1263264663283908609?s=20, last accessed June 29, 2020).

Covid Economics 37, 14 July 2020: 100-140

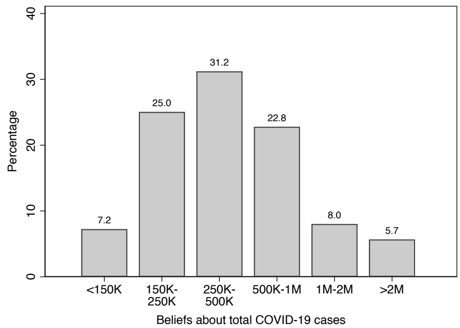
Figure A2:
Histograms of Risk Perceptions and Behavior



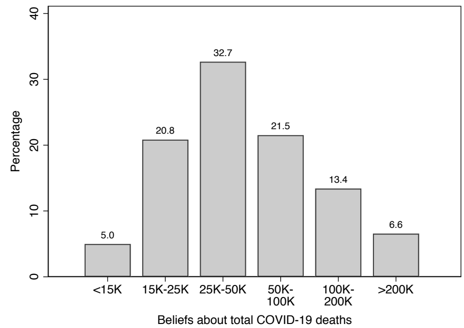
(a) Risk of contagion next week



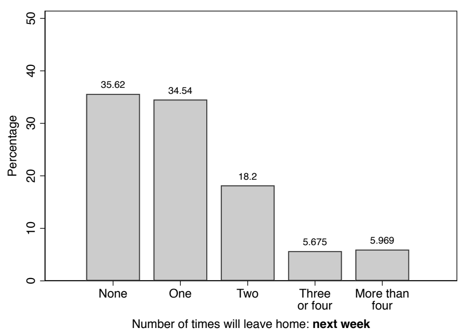
(b) Risk of contagion in 4 weeks



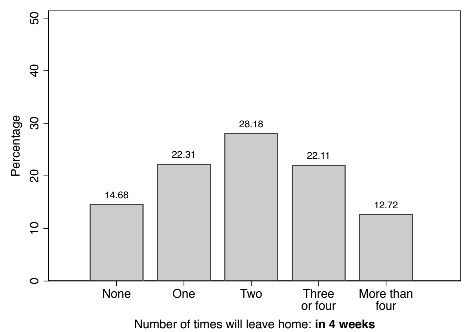
(c) Expected total cases



(d) Expected total deaths



(e) Times leaving the house next week

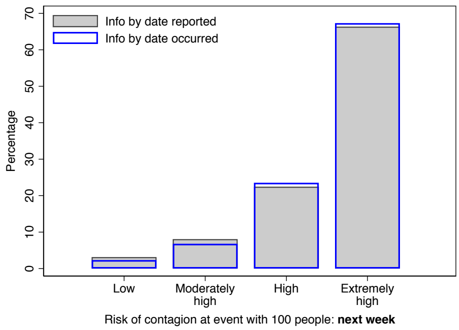


(f) Times leaving the house in 4 weeks

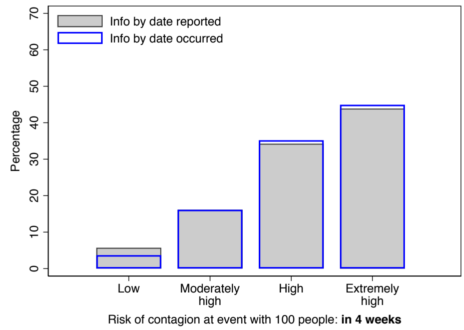
Notes: These graphs show histograms for the six outcome variables related to perceptions and expected behavior elicited in the survey for our full sample of participants. Each plot shows the percentage of total respondents that chose each of the answers.

Covid Economics 37, 14 July 2020: 100-140

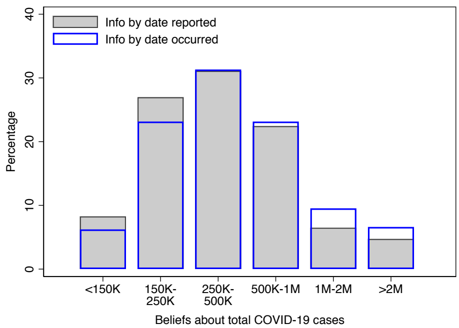
Figure A3:
Histograms of Risk Perceptions and Behavior by Informational Treatments



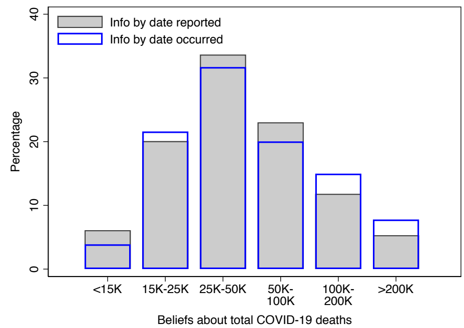
(a) Risk of contagion next week



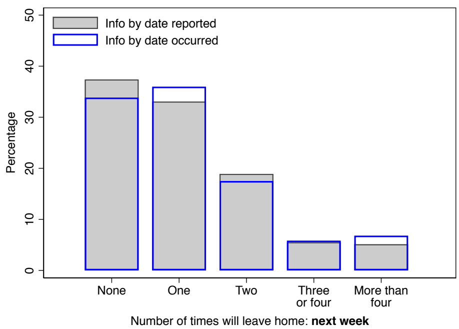
(b) Risk of contagion in 4 weeks



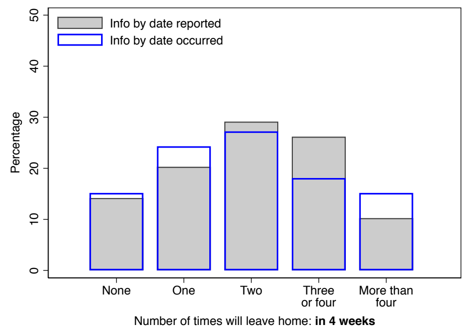
(c) Expected total cases



(d) Expected total deaths



(e) Times leaving the house next week

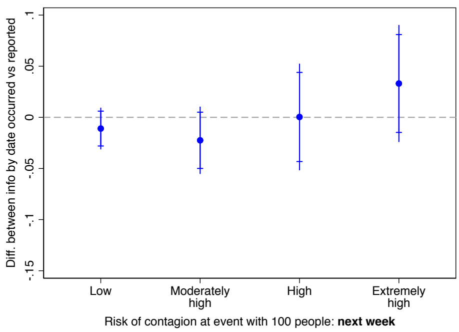


(f) Times leaving the house in 4 weeks

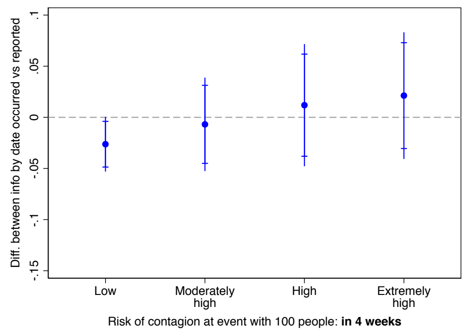
Notes: These graphs show histograms for the six outcome variables related to perceptions and expected behavior elicited in the survey for our full sample of participants. We distinguish between the two informational treatments. Each plot shows the percentage of total respondents that chose each of the answers.

Covid Economics 37, 14 July 2020: 100-140

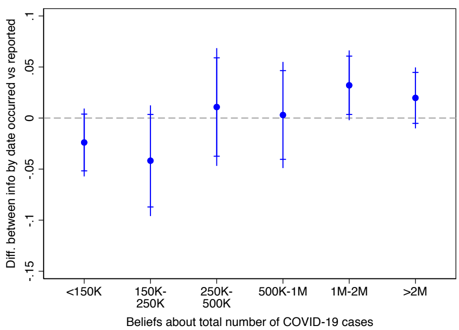
Figure A4:
Estimates of Informational Treatments for Full Set of Responses



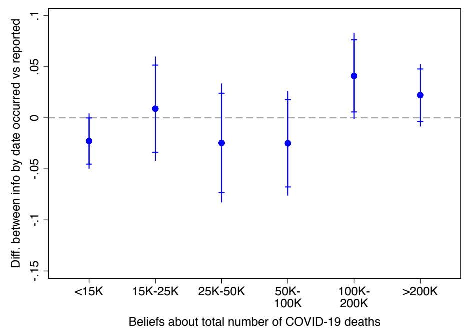
(a) Risk of contagion next week



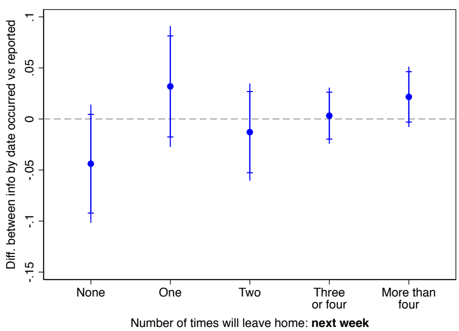
(b) Risk of contagion in 4 weeks



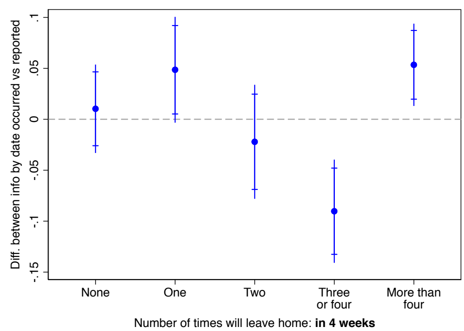
(c) Expected total cases



(d) Expected total deaths



(e) Times leaving the house next week

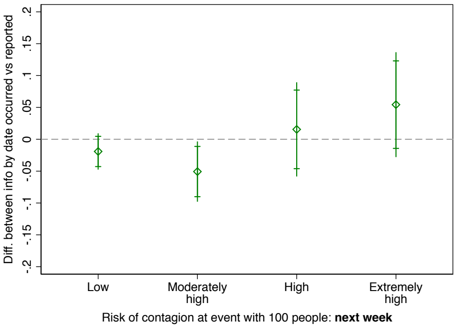


(f) Times leaving the house in 4 weeks

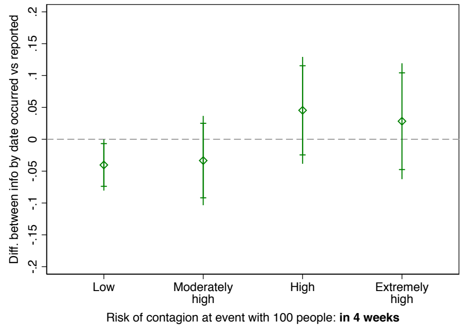
Notes: These graphs show estimates of the difference by informational treatment on the six outcome variables related to perceptions and expected behavior elicited in the survey for our full sample of participants. Each plot shows coefficients from multiple regressions with indicators for each response as the outcome variable. Coefficients correspond to the average difference between respondents that received information based on the actual date of death relative to those that received information based on date of reports. Vertical bars show 95 and 90% confidence intervals.

Covid Economics 37, 14 July 2020: 100-140

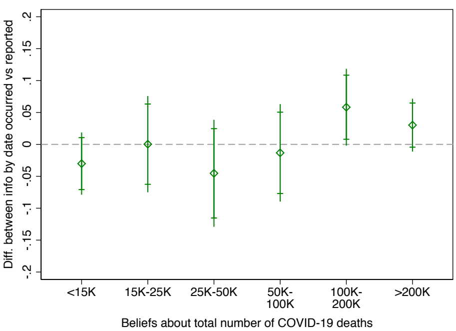
Figure A5:
Estimates of Informational Treatments for Full Set of Responses: Low Prior Sample



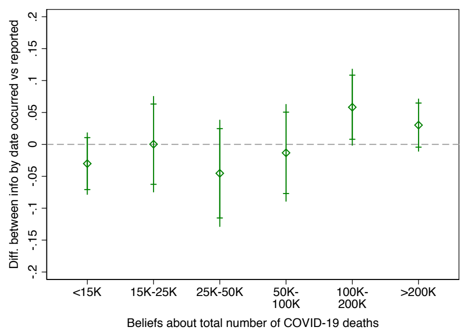
(a) Risk of contagion next week



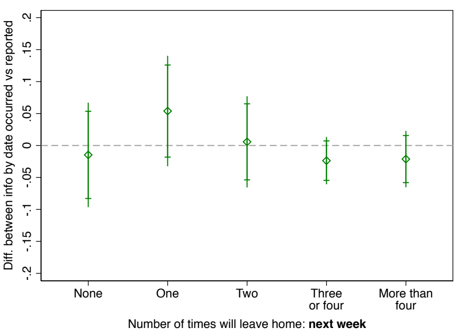
(b) Risk of contagion in 4 weeks



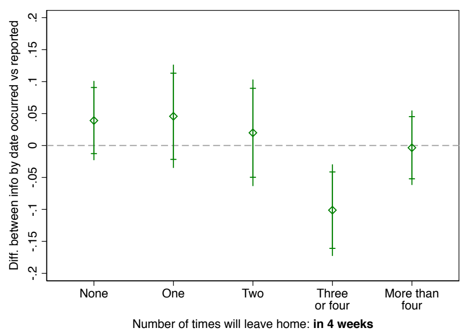
(c) Expected total cases



(d) Expected total deaths



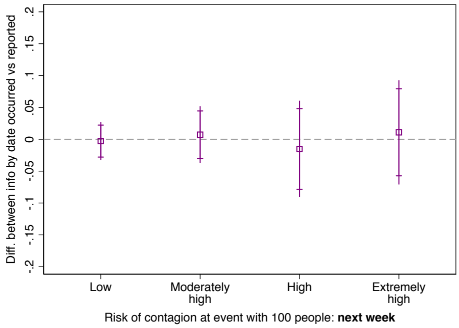
(e) Times leaving the house next week



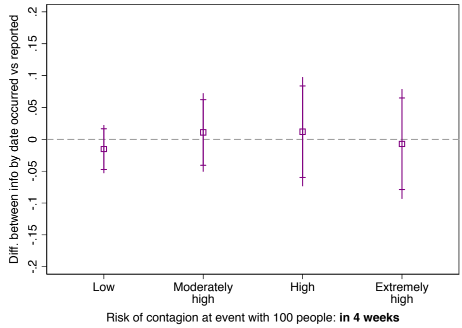
(f) Times leaving the house in 4 weeks

Notes: These graphs show estimates of the difference by informational treatment on the six outcome variables related to perceptions and expected behavior elicited in the survey for the sample of participants with a low prior of total Covid-19 cases as of May 20. Each plot shows coefficients from multiple regressions with indicators for each response as the outcome variable. Coefficients correspond to the average difference between respondents that received information based on the actual date of death relative to those that received information based on date of reports. Vertical bars show 95 and 90% confidence intervals.

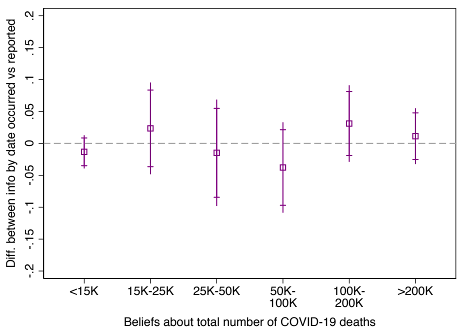
Figure A6:
Estimates of Informational Treatments for Full Set of Responses: High Prior Sample



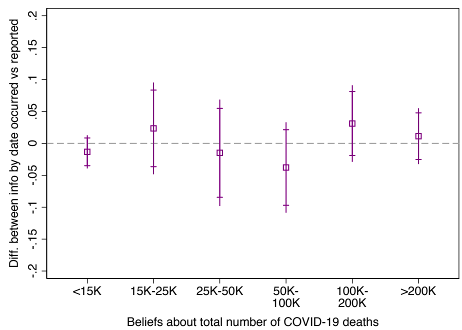
(a) Risk of contagion next week



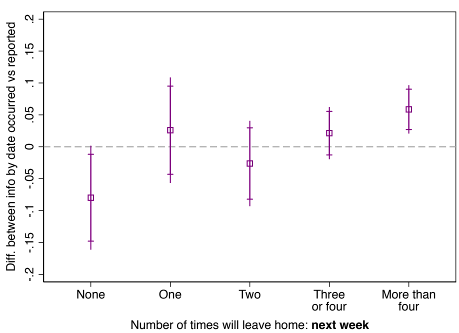
(b) Risk of contagion in 4 weeks



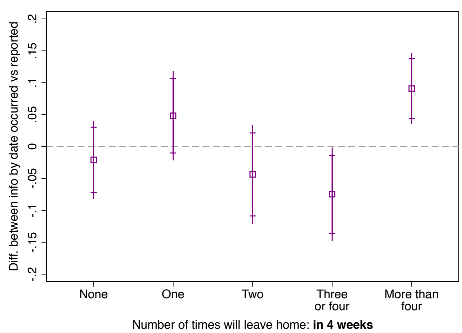
(c) Expected total cases



(d) Expected total deaths



(e) Times leaving the house next week



(f) Times leaving the house in 4 weeks

Notes: These graphs show estimates of the difference by informational treatment on the six outcome variables related to perceptions and expected behavior elicited in the survey for the sample of participants with a high prior of total Covid-19 cases as of May 20. Each plot shows coefficients from multiple regressions with indicators for each response as the outcome variable. Coefficients correspond to the average difference between respondents that received information based on the actual date of death relative to those that received information based on date of reports. Vertical bars show 95 and 90% confidence intervals.

Covid Economics 37, 14 July 2020: 100-140

CEPR PRESS

Table A1:
Balance Table for Survey Covariates: Low Prior Sample

	Informational treatments		
	Deaths by date reported	Deaths by date occurred	Difference in means
Female	0.475	0.500	0.025
	(0.500)	(0.501)	(0.045)
Ages 18-22	0.324	0.412	0.088**
	(0.469)	(0.493)	(0.044)
Ages 23-29	0.252	0.232	-0.020
	(0.435)	(0.423)	(0.039)
Ages 30-49	0.248	0.192	-0.056
	(0.433)	(0.395)	(0.037)
Ages 50+	0.176	0.164	-0.012
	(0.382)	(0.371)	(0.034)
Works	0.420	0.344	-0.076*
	(0.495)	(0.476)	(0.044)
Attends school	0.353	0.432	0.079*
	(0.479)	(0.496)	(0.044)
Works and attends school	0.155	0.152	-0.003
	(0.363)	(0.360)	(0.033)
Other occupation/employment status	0.071	0.072	0.001
	(0.258)	(0.259)	(0.023)
Lives in Mexico City	0.773	0.764	-0.009
	(0.420)	(0.425)	(0.038)
Lives in apartment	0.340	0.396	0.056
	(0.475)	(0.490)	(0.044)
Lives in house, no yard	0.118	0.108	-0.010
	(0.323)	(0.311)	(0.029)
Lives in house with yard	0.542	0.496	-0.046
	(0.499)	(0.501)	(0.045)
Household size: 1-2	0.223	0.240	0.017
	(0.417)	(0.428)	(0.038)
Household size: 3	0.231	0.236	0.005
	(0.422)	(0.425)	(0.038)
Household size: 4	0.214	0.224	0.010
	(0.411)	(0.418)	(0.038)
Household size: 5+	0.546	0.524	-0.022
	(0.499)	(0.500)	(0.045)
Has HH members over 70 years old	0.181	0.076	-0.105***
	(0.386)	(0.266)	(0.030)
Has HH members 60-70 years old	0.206	0.228	0.022
	(0.405)	(0.420)	(0.037)
Has HH members 50-60 years old	0.496	0.432	-0.064
	(0.501)	(0.496)	(0.045)
Does not seek healthcare when sick	0.130	0.108	-0.022
	(0.337)	(0.311)	(0.029)
Self-medicates when sick	0.357	0.396	0.039
	(0.480)	(0.490)	(0.044)
Observations	238	250	488

Notes: This table shows means and standard deviations for a series of covariates asked in the survey before the informational treatment was presented for the sample of participants with a low prior of total Covid-19 cases as of May 20. We show statistics separately for each informational treatment, as well as the difference in the means. Stars denote significance from a difference in means test.
*** p<0.01, ** p<0.05, * p<0.1

Covid Economics 37, 14 July 2020: 100-140

Covid Economics 37, 14 July 2020: 100-140

Table A2:
Balance Table for Survey Covariates: High Prior Sample

	Informational treatments		
	Deaths by date reported	Deaths by date occurred	Difference in means
Female	0.515	0.481	-0.034
	(0.501)	(0.501)	(0.043)
Ages 18-22	0.319	0.356	0.038
	(0.467)	(0.480)	(0.041)
Ages 23-29	0.293	0.273	-0.020
	(0.456)	(0.446)	(0.039)
Ages 30-49	0.215	0.239	0.024
	(0.411)	(0.427)	(0.036)
Ages 50+	0.174	0.133	-0.041
	(0.380)	(0.340)	(0.031)
Works	0.400	0.314	-0.086**
	(0.491)	(0.465)	(0.041)
Attends school	0.381	0.402	0.020
	(0.487)	(0.491)	(0.042)
Works and attends school	0.159	0.163	0.004
	(0.367)	(0.370)	(0.032)
Other occupation/employment status	0.059	0.121	0.062**
	(0.237)	(0.327)	(0.025)
Lives in Mexico City	0.778	0.742	-0.035
	(0.417)	(0.438)	(0.037)
Lives in apartment	0.344	0.375	0.031
	(0.476)	(0.485)	(0.042)
Lives in house, no yard	0.130	0.125	-0.005
	(0.337)	(0.331)	(0.029)
Lives in house with yard	0.526	0.500	-0.026
	(0.500)	(0.501)	(0.043)
Household size: 1-2	0.241	0.261	0.021
	(0.428)	(0.440)	(0.038)
Household size: 3	0.185	0.254	0.069*
	(0.389)	(0.436)	(0.036)
Household size: 4	0.285	0.227	-0.058
	(0.452)	(0.420)	(0.038)
Household size: 5+	0.574	0.485	-0.089**
	(0.495)	(0.501)	(0.043)
Has HH members over 70 years old	0.141	0.083	-0.057**
	(0.348)	(0.277)	(0.027)
Has HH members 60-70 years old	0.222	0.178	-0.044
	(0.417)	(0.383)	(0.035)
Has HH members 50-60 years old	0.430	0.508	0.078*
	(0.496)	(0.501)	(0.043)
Does not seek healthcare when sick	0.148	0.197	0.049
	(0.356)	(0.398)	(0.033)
Self-medicates when sick	0.411	0.367	-0.044
	(0.493)	(0.483)	(0.042)
Observations	270	264	534

Notes: This table shows means and standard deviations for a series of covariates asked in the survey before the informational treatment was presented for the sample of participants with a high prior of total Covid-19 cases as of May 20. We show statistics separately for each informational treatment, as well as the difference in the means. Stars denote significance from a difference in means test.
*** p<0.01, ** p<0.05, * p<0.1

CEPR PRESS

B Survey Text in English

This is an anonymous online survey that is being conducted by Profs. Emilio Gutierrez, Adrian Rubli and Tiago Tavares for an academic project aimed at better understanding the public's perceptions about the evolution of the Covid-19 pandemic in Mexico. Responding to the survey takes approximately 10 minutes. We ask you to please answer to all the questions if you choose to participate. Despite the fact that you received an invitation to participate in this survey via email or social media, the dataset where the information you provide will be stored does not collect any type of personal information (such as your name, phone number or IP address). We take all the relevant measures to safeguard your identity. Clicking on the "accept" button below you certify that you are over 18 years of age, and that you agree to respond to all the questions asked. The information you provide will only be used for academic purposes and statistical analyses, never revealing individual-level responses.

Sociodemographic Questions

Sex: Male / Female / Other or Prefer not to say

What is your age?: 18-22 / 23-29 / 30-39 / 40-49 / 50-59 / 60-69 / 70-79 / 80 or older

The highest schooling degree you have obtained is: Elementary school / Secondary school / Highschool / Undergraduate degree / Graduate degree

Occupation: Works / Attends school / Works and attends school / Unemployed / House work / Retired

Where do you live?: CDMX or its suburbs / Aguascalientes / Baja California / Baja California Sur / Campeche / Coahuila / Colima / Chiapas / Chihuahua / Durango / Guanajuato / Guerrero / Hidalgo / Jalisco / EdoMex outside CDMX metro area / Michoacan / Morelos / Nayarit / Nuevo Leon / Oaxaca / Puebla / Queretaro / Quintana Roo / San Luis Potosí / Sinaloa / Sonora / Tabasco / Tamaulipas / Tlaxcala / Veracruz / Yucatan / Zacatecas

How would you describe the house you live in: Apartment / House with yard / House without yard

Do you have internet access at home (Wi-Fi)?: Yes / No

Do you have access to a computer at home?: Yes, but I share it with others / Yes, and I am the only user / No

Apart from you, how many people live in your home?: 1 / 2 / 3 / 4 / 5 or more

Is anyone in your household aged more than 70?: Yes / No

Is anyone in your household aged between 60 and 70?: Yes / No

Is anyone in your household aged between 50 and 60?: Yes / No

What is your household's approximate monthly income?: 0-4,000 pesos / 4000-10,000 pesos / 10,000-20,000 pesos / 20,000-30,000 pesos / 30,000-40,000 pesos / 40,000-50,000 pesos / 50,000-75,000 pesos / 75,000-100,000 pesos / more than 100,000 pesos

Covid Economics 37, 14 July 2020: 100-140

Do you have access to private health insurance?: Yes / No

Do you have access to health services from IMSS, ISSSTE, PEMEX, SEDENA or SEMAR?: Yes / No

Do you have access to health services from INSABI or Seguro Popular?: Yes / No

When you fall sick, what do you usually do?: Nothing / Take OTCs / Go to a pharmacy-adjacent doctor's office / Go to a doctor's appointment in the private sector / Go to a doctor's appointment in the public sector / Use the medical services at my office or university

Who did you vote for in the last presidential election?: Andres Manuel Lopez Obrador / Ricardo Anaya / Jose Antonio Meade / Other candidate / No vote

What is your opinion about Andres Manuel Lopez Obrador's government's performance?: Completely approve / Approve / Disapprove / Completely disapprove

Covid-19 related questions

How often do you watch the press conference that Dr. Hugo Lopez-Gatell holds daily at 7pm?: Every day / Several times a week / Once a week / Sporadically / Never

How trustworthy do you think is the information about the evolution of Covid-19 shared by Mexican authorities during the daily 7pm press conference?: Very trustworthy / Somewhat trustworthy / Somewhat untrustworthy / Very untrustworthy

Have you received information regarding the evolution of Covid-19 through Facebook?: Yes / No

How trustworthy do you think is the information about the evolution of Covid-19 shared through Facebook?: Very trustworthy / Somewhat trustworthy / Somewhat untrustworthy / Very untrustworthy

Have you received information regarding the evolution of Covid-19 through Twitter?: Yes / No

How trustworthy do you think is the information about the evolution of Covid-19 shared through Twitter?: Very trustworthy / Somewhat trustworthy / Somewhat untrustworthy / Very untrustworthy

Have you received information regarding the evolution of Covid-19 through Whatsapp?: Yes / No

How trustworthy do you think is the information about the evolution of Covid-19 shared through Whatsapp?: Very trustworthy / Somewhat trustworthy / Somewhat untrustworthy / Very untrustworthy

Think of may 20th. According to you, approximately how many Covid-19 cases had be reported by that date?: Less than 10,000 / Between 10,000 and 25,000 / Between 25,000 and 50,000 / Between 50,000 and 75,000 / Between 75,000 and 100,000 / Between 100,000 and 150,000 / More than 150,000

Think of may 20th. According to you, approximately how many Covid-19 deaths had be reported by that date?: Less than 1,000 / Between 1,000 and 2,500 / Between 2,500 and 5,000 / Between 5,000 and 7,500 / Between 7,500 and 10,000 / Between 10,000 and 15,000 / More than 15,000

Covid Economics 37, 14 July 2020: 100-140

CEPR PRESS

What is your opinion about the president's actions in face of the Covid-19 pandemic?: Completely approve / Approve / Disapprove / Completely disapprove

How many times did you leave home last week?: You did not leave home / Once / Twice / Three or four times / More than four times

Information treatments

Cumulative deaths by date reported: The following graph compares the evolution of total Covid-19 related deaths in Mexico and Sweden, from march 22nd to may 15th. The information is presented according to the date on which deaths were reported.

Cumulative deaths by date occurred: The following graph compares the evolution of total Covid-19 related deaths in Mexico and Sweden, from march 22nd to may 15th. For Sweden, rhe information is presented according to the date on which deaths were reported. For Mexico, according to the date on which deaths occurred.

Post-treatment questions

Dr. Hugo Lopez-Gatell has said that the evolution of the pandemic in Mexico is similar to the one experienced by Sweden. In your opinion, the Covid-19 pandemic in Mexico is evolving: Much faster than in Sweden / Faster than in Sweden / Similar to Sweden / Slower than in Sweden / Much slower than in Sweden

What is your opinion about Dr. Hugo Lopez-Gatell and other Mexican health authorities' strategy in face of Covid-19?: Completely approve / Approve / Disapprove / Completely disapprove

When do you expect that Mexico will reach 150,000 total confirmed Covid-19 cases?: Early June / Mid June / Late June / Early July / Mid July / Late July (or later) / There will be less than 150,000 total cases

When do you expect we will reach the maximum number of daily Covid-19 cases in Mexico?: Early June / Mid June / Late June / Early July / Mid July / Late July (or later)

How many cases of Covid-19 do you think will have been confirmed in Mexico by the end of this epidemic outbreak?: Less than 100,000 cases / Between 100,000 and 150,000 cases / Between 150,000 and 250,000 cases / Between 250,000 and 500,000 cases / Between 500,000 and one million cases / Between one and two million cases / More than two million cases

When do you expect that Mexico will reach 15,000 total confirmed Covid-19 deaths?: Early June / Mid June / Late June / Early July / Mid July / Late July (or later) / There will be less than 15,000 deaths

When do you expect we will reach the maximum number of daily Covid-19 deaths in Mexico?: Early June / Mid June / Late June / Early July / Mid July / Late July (or later)

Covid Economics 37, 14 July 2020: 100-140

How many deaths due to Covid-19 do you think will have been confirmed in Mexico by the end of this epidemic outbreak?: Less than 10,000 deaths / Between 10,000 and 15,000 deaths / Between 15,000 and 25,000 deaths / Between 25,000 and 50,000 deaths / Between 50,000 and 100,000 deaths / Between 100,000 and 200,000 deaths / More than 200,000 deaths

Imagine an extremely optimistic scenario (which would only happen with a probability lower than 10 percent). In such scenario, the total number of Covid-19 deaths in Mexico would be : Less than 3,000 deaths / Between 3,000 and 6,000 deaths / Between 6,000 and 9,000 deaths / Between 9,000 and 12,000 deaths / Between 12,000 and 15,000 deaths / Between 15,000 and 18,000 deaths / Between 18,000 and 21,000 deaths / Between 21,000 and 30,000 deaths / Between 30,000 and 50,000 deaths / Between 50,000 and 80,000 deaths / Between 80,000 and 120,000 deaths / More than 120,000 deaths

Imagine an extremely pessimistic scenario (which would only happen with a probability lower than 10 percent). In such scenario, the total number of Covid-19 deaths in Mexico would be : Less than 12,000 deaths / Between 12,000 and 15,000 deaths / Between 15,000 and 18,000 deaths / Between 18,000 and 21,000 deaths / Between 21,000 and 30,000 deaths / Between 30,000 and 50,000 deaths / Between 50,000 and 80,000 deaths / Between 80,000 and 120,000 deaths / Between 120,000 and 180,000 deaths / Between 180,000 and 250,000 deaths / Between 250,000 and 500,000 deaths / More than 500,000 deaths

When do you think that Mexico City will stop being under the maximum alter level due to Covid-19?: Early June / Mid June / Late June / Early July / Mid July / Late July / Early August / Mid August / Late August / September or later

Next week, how many times do you expect to leave home?: Will not leave home / Once / Twice / Three or four times / More than four times

If next week you had to attend a social gathering with 100 people, how high do you think the risk of being infected with the virus would be?: Very high risk / High risk / Moderately high risk / Moderately low risk / Low risk / Very low risk

In four weeks, how many times do you expect to leave home?: Will not leave home / Once / Twice / Three or four times / More than four times

If in four weeks you had to attend a social gathering with 100 people, how high do you think the risk of being infected with the virus would be?: Very high risk / High risk / Moderately high risk / Moderately low risk / Low risk / Very low risk

Do you think that most private universities in Mexico will be back on campus in August?: Yes, everything will go back to normal / Yes, but some courses will still be online / No, all courses will be online next semester

If the 2018 presidential election were today (with the same candidates), who would you vote for?: Andres Manuel Lopez Obrador / Ricardo Anaya / Jose Antonio Meade / Other candidate / Would not vote

Covid Economics 37, 14 July 2020: 100-140

C Additional Details and Results on the Model

C.1 Model Analysis

The optimization of static hours spent outside the home is given by:

$$\bar{n} = \arg\max_{n\in(0,1)} \left(u\left(n\right)\right) = \frac{1}{1+\lambda_p}$$

Note that these are also the hours spent by infected and recovered individuals:

$$n\left(i,t\right) = \overline{n}$$
$$n\left(c,t\right) = \overline{n}$$

This then implies closed-form solutions for:

$$\begin{aligned} V\left(c\right) &= \frac{u\left(\bar{n}\right)}{1-\beta} \\ V\left(r\right) &= \frac{\beta\theta\left(1-\delta\right)V\left(c\right)}{1-\beta\left(1-\theta\right)} \\ V\left(i\right) &= \frac{u\left(\bar{n}\right)+\beta\gamma V\left(r\right)}{1-\beta\left(1-\gamma\right)} \end{aligned}$$

As for the problem of the healthy agents, the first order conditions imply:

$$\begin{aligned} u_n\left(n\right) &= \beta\pi_n\left(n,\widetilde{\Pi}_t\right)\left(V\left(s,t+1\right)-V\left(i\right)\right) \\ \Rightarrow\frac{1}{n}-\frac{\lambda_p}{1-n} &= \beta\widetilde{\Pi}_t\left(V\left(s,t+1\right)-V\left(i\right)\right) \end{aligned}$$

Note that the simple implication of the model says that as long as the value of being healthy is larger that being infected, that is, if $V\left(s,t+1\right) > V\left(i\right)$ and $\widetilde{\Pi}_t > 0$, then $n\left(s,t\right) < \bar{n}$. This means that susceptible agents reduce the number of hours outside the house to prevent becoming infected between period t and $t+1$. Moreover, the larger the perceived infection rate $\widetilde{\Pi}_t$, the larger the response of susceptible agents in terms of how much they decrease their hours spent in the market place (outside the house).

Covid Economics 37, 14 July 2020: 100-140

C.2 Algorithm for the *Belief-Induced Equilibrium* Solution

In order to solve the above model, we use the following algorithm:

1. Choose a sequence for a large T and some sequence $\left\{\Pi_t^{(0)}\right\}_{t=0}^{T}$, making sure that $\Pi_T^{(0)} = 0$.

2. Solve for the values using backward induction and get policies on $n(j,t)$.

3. Compute the path of $\mathcal{M}_t$.

4. Update probabilities $\Pi_t^{(1)}$.

5. Iterate until $\left|\Pi^{(1)} - \Pi^{(0)}\right| < \varepsilon$ for small ε, otherwise set $\Pi^{(0)} = \Pi^{(1)}$ and go back to step 2.

C.3 Details of Model Calibration for Mexico

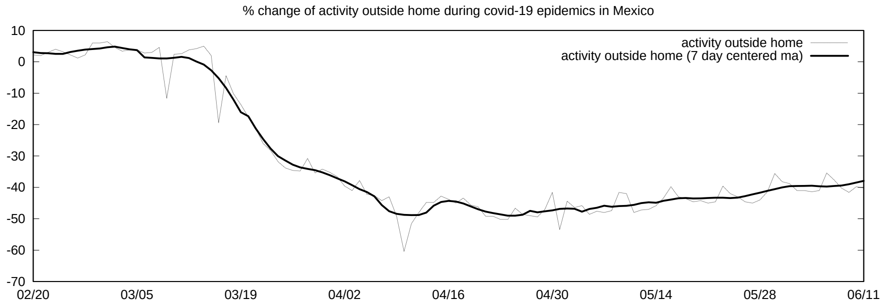
The calibration regarding the parameter λ_p that captures hours spent at vs outside the house in Mexico uses information from the 2014 household time use survey (*Encuesta Nacional sobre Uso del Tiempo 2014*) from the national office of statistics INEGI.[1] From the survey, we consider time spent outside the home as the sum of aggregate hours in market activities and consumption goods, entertainment and social activities, and studying (education) and related activities. As for time spent at home, we aggregate all hours in non-remunerated work at home, and personal activities (including sleeping, eating, and personal hygiene). We conclude that on average a Mexican household spends 36% of total time in activities outside the house, which corresponds to a parameter of $\lambda_p = 1.77$.

As for the parameter that regulates the preferences for staying alive b, we use data from Google Community Mobility Reports for Mexico to determine the reduction in non-home activities during the Covid-19 epidemic.[2] We average all non-home activity (retail and recreation, grocery and pharmacy visits, visit to parks, activity spent in transit, and workplace activity), and measure a 7-day centered moving average. We show the time series for these data in Figure A7. This analysis reveals that activity outside the home decreased by about 45% at the trough of the epidemic, and we use this decline to calibrate the parameter b in the model simulations.

[1] The time use survey data can be accessed at https://en.www.inegi.org.mx/programas/enut/2014/.
[2] Google mobility data can be accessed at https://www.google.com/covid19/mobility/.

Covid Economics 37, 14 July 2020: 100-140

Figure A7:
Mobility as a Response to Covid-19 in Mexico



Notes: This graph shows the percentage change in activity outside the home using data from Google Community Mobility Reports for Mexico (available at https://www.google.com/covid19/mobility/). We show the actual daily data as well as a 7-day centered moving average.

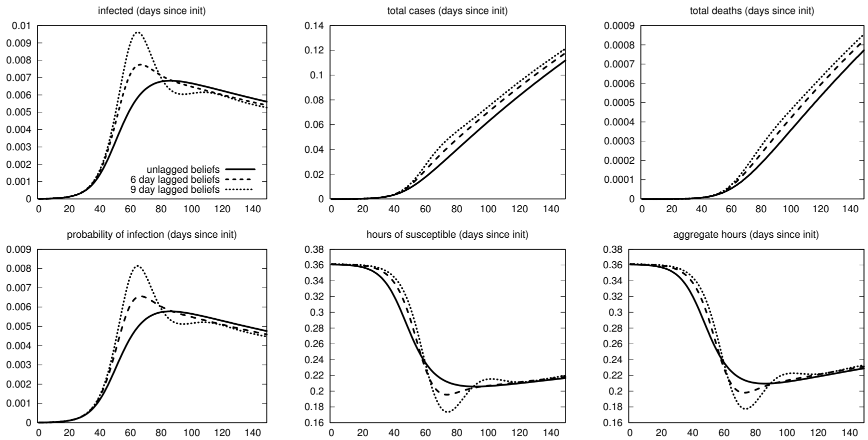
Finally, to calibrate the baseline risk of transmission Π_0, we use estimates of the basic reproduction number R_0 from Marioli et al. (2020) that correspond to a Covid-19 $R_0 = 1.84$ for Mexico. Note that the model counterpart implies that $R_0 = n(s,0)\,\bar{n}\Pi_0/\gamma \approx (\bar{n})^2\,\Pi_0/\gamma$. It follows that for $\bar{n} = 0.36$, $\gamma = 0.166$, and $R_0 = 1.84$, we have $\Pi_0 = 2.35$.

C.4 Additional Results and Robustness of the Model

Figure A8 shows additional results of the simulation of the model without delays $k = 0$ and with delays $k = 6$ and $k = 9$. Additionally, Table A3 shows how the model results change when we either increase or decrease important parameters, while keeping all other constant to the baseline model.

Covid Economics 37, 14 July 2020: 100-140

Figure A8:
Additional Simulation Results of Behavioral Model



Notes: These graphs show additional results of the model simulation. We show results considering a 6-day delay in the formation of beliefs consistent with reporting delays, zero delay corresponding to contemporaneous (unlagged) beliefs, and a more extreme case of a 9-day lag. The top panel shows figures for the percentage of the population that becomes infected on a daily basis since the onset of the epidemic, the total cumulative cases as a percentage of the total population over time, and the total cumulative deaths as a percentage of the population over time. The bottom panel shows the probability of becoming infected since the beginning of the epidemic, the percentage of hours in a day that susceptible individuals (who have never been infected) spend outside the home (as in Figure 2), and the aggregate hours spent outside as the combined hours of susceptible, infected, and recovered agents.

Covid Economics 37, 14 July 2020: 100-140

CEPR PRESS

Covid Economics 37, 14 July 2020: 100-140

Table A3:
Robustness Checks on the Model

	Delay	Peak infections (% of pop)	Days to peak infections	Maximum daily deaths (% of pop)	Total deaths on 120th day (% of pop)	Hrs. susceptible to infection at trough
Baseline	$k = 0$	0.68199	87	0.00089	0.05257	20.58527
	$k = 6$	0.77462	68	0.00095	0.05866	19.55523
	$k = 9$	0.96137	66	0.00110	0.06218	17.35130
Higher death rate	$k = 0$	0.36469	89	0.00096	0.06045	20.19884
$\delta = 0.016$	$k = 6$	0.40438	63	0.00100	0.06736	19.33626
	$k = 9$	0.50476	61	0.00116	0.07144	17.08112
Lower death rate	$k = 0$	1.19222	86	0.00077	0.04278	21.15268
$\delta = 0.004$	$k = 6$	1.38842	71	0.00084	0.04762	19.93799
	$k = 9$	1.70642	70	0.00098	0.05040	17.82880
Higher infection rate	$k = 0$	2.81456	73	0.00217	0.13813	13.06772
$1/\gamma = 10$	$k = 6$	3.54543	57	0.00249	0.15530	11.60159
	$k = 9$	4.92972	56	0.00321	0.16859	8.84075
Lower infection rate	$k = 0$	0.36816	102	0.00058	0.02987	24.40736
$1/\gamma = 5$	$k = 6$	0.39735	82	0.00061	0.03360	23.85656
	$k = 9$	0.45332	78	0.00066	0.03568	22.65139
Higher resolving probability	$k = 0$	0.70261	87	0.00093	0.05844	20.60921
$\theta = 0.2$	$k = 6$	0.79919	68	0.00104	0.06459	19.57020
	$k = 9$	0.99175	66	0.00126	0.06806	17.36984
Lower resolving probability	$k = 0$	0.64418	87	0.00080	0.04218	20.54111
$\theta = 0.05$	$k = 6$	0.73004	67	0.00081	0.04777	19.52695
	$k = 9$	0.90704	65	0.00086	0.05117	17.32289

Notes: This table shows results from changing parameters of the model. We consider a higher and lower death rate, infection rate and resolving probability. For each case, we show estimates from a 6-day delay in the formation of beliefs consistent with reporting delays, zero delay corresponding to unlagged beliefs, and a more extreme case of a 9-day lag. We present the estimates for the peak number of infections (expressed as a percentage of the total population), the number of days it takes from the onset of the epidemic to reach this peak, the maximum number of daily deaths as a percentage of the population, the total number of deaths accrued up to the 120th day as a percentage of the population, and the percentage of hours in a day susceptible to infection at the trough of the curve.

Covid Economics Issue 37, 14 July 2020

Socioeconomic conditions, government interventions and health outcomes during COVID-19[1]

Badar Nadeem Ashraf[2]

Date submitted: 2 July 2020; Date accepted: 3 July 2020

The COVID-19, after hitting hard the developed regions of Europe and the United States, is now fast spreading in relatively less developed regions including the Latin America, South Asia and the African continent. In this paper, we examine the impact of socioeconomic conditions on the health outcomes by COVID-19 and the moderating role of government emergency measures on the relationship between socioeconomic conditions and the health outcomes by COVID-19. Using a panel dataset consisting of 9529 daily observations from 80 countries over the period from January 22 to May 20, 2020, we find that socioeconomic circumstances have strong negative association with COVID-19 confirmed cases and deaths per million people. Quantitatively, a one standard deviation improvement in socioeconomic conditions lowers COVID-19 confirmed cases and deaths per million people by one half. Next, with the help of interaction terms between socioeconomic conditions and government emergency policies, we find that stringent social distancing measures and generous income support programs help to lower the cases and deaths particularly in countries with poor socioeconomic conditions. These findings have important implications to design the right set of government policies to lower the lives losses in countries and regions with poor socioeconomic conditions.

1 I would like to thank an anonymous reviewer for useful comments. The usual disclaimer applies.

2 Associate Professor, School of Finance, Jiangxi University of Finance and Economics.

Covid Economics 37, 14 July 2020: 141-162

Covid Economics 37, 14 July 2020: 141-162

1. Introduction

The outbreak of COVID-19 pandemic, since its start in China in December 2019, is wreaking havoc across the globe. The governments have scrambled with emergency measures to contain the disease on the one hand while to protect the economies on the other hand. The health outcomes in terms of the numbers of total infections and deaths vary by a large extent across countries during the pandemic (Fanelli & Piazza 2020; Okell *et al.* 2020). It has raised the concern regarding which government emergency measures are effective and which are not in controlling the disease. Besides, which other factors, other than the government response, are contributing to such heterogeneous numbers of health outcomes across countries. In this paper, we examine the impact of cross-country differences in pre-pandemic socioeconomic conditions on the health outcomes during the pandemic. We also examine whether government policy measures were effective in eliminating the adverse impact of poor socioeconomic conditions during the pandemic. The main purpose of this exercise is to analyze the impact of socioeconomic conditions on the social distribution of the disease across countries and to suggest policy implications or evaluate the effectiveness of policies targeted to diminish the effect of socioeconomic inequalities on local COVID-19 outbreaks.

The awareness is increasing that social determinants of health, the factors primarily influenced by social policy, have important implications for health inequalities in addition to the access to healthcare (Braveman *et al.* 2011; Braveman & Gottlieb 2014). The World Health Organization defines social determinants of health (SDH) as “the conditions in which people are born, grow, live, work and age” and “the fundamental drivers of these conditions.” Social determinants normally point to health related features of neighbourhoods, such as parks, recreational areas, and access to healthy food, which can influence health-related behaviours. However, Braveman and Gottlieb (2014) argue that among social determinants, socioeconomic factors which include income, wealth, employment status and/or education might be the fundamental drivers of health outcomes and stress the need to further investigate the impact of these factors.

Socioeconomic factors may have both rapid as well as the long-term impact on health outcomes. Rapid health impact channels through the factors such as higher lead ingestion and pollution in substandard housing and neighbourhoods (Brown 1995; Lanphear *et al.* 2001; Lidsky & Schneider 2003), higher social acceptability of risky health behaviours and exposure to violence and alcohol consumption (Bingenheimer *et al.* 2005; Pollack *et al.* 2005), insufficient sleep or affected sleep patterns (Hale 2005; Marco *et al.* 2011), and pressure to go to work (Cook *et al.* 2009). Long-term health impact may channel through the lower availability of fresh produce (Cummins & Macintyre 2006; Gordon-Larsen *et al.* 2006), poverty driven chronic childhood stress (Evans & Schamberg 2009), biological wear and tear from exposure to chronic social and environmental stressors (McEwen & Gianaros 2010) and by regulation of genes controlling immune functioning (Tung *et al.* 2012).

Above factors might have influenced the numbers of infections and deaths from the COVID-19. For example, disease may spread easily among poor families living in small overcrowded substandard housing. Acceptability of risky health behaviours may encourage individuals to go outside and not respect the social distancing measures. Pressure to go for work can also increase the risk of infection. Factors such as insufficient sleep or affected sleep patterns, lower availability of fresh produce, biological wear and tear and weak regulation of genes may result in lower immunity against the novel coronavirus and higher chances of symptomatic infections and deaths. Based on above arguments, our first hypothesis is that poor socioeconomic conditions might result in adverse health outcomes during the COVID-19.

Government policy may moderate the impact of socioeconomic circumstances on health outcomes. In this regard, Braveman and Gottlieb (2014) suggest that poor socioeconomic conditions might have lower impact on health in settings with social support for basic needs. More specifically, Anand and Ravallion (1993) argue that low income may have less impact on health outcomes in places where basic needs, such as food, housing, education and medical care, are provided by the state.

We postulate government interventions can weaken the rapid impact of socioeconomic circumstances on COVID-19 health outcomes. Government interventions may primarily work through the channel of social distancing. So far, without a vaccine, social distancing has proved to be the most important defence against highly contagious COVID-19 (Anderson *et al.* 2020; Okell *et al.* 2020). However, people living under poor socioeconomic conditions are less likely to observe social distancing. For example, Lou *et al.* (2020) and Wright *et al.* (2020) find that compliance with stay-at-home orders during the COVID-19 pandemic varies significantly with income, where lower-income groups are less likely to follow the orders due to work related trips and are more likely to get exposed to the virus. In such settings, implementation of stringent social distancing measures and the provision of more generous income support to low income groups by the government may increase the

Covid Economics 37, 14 July 2020: 141-162

chances of compliance with social distancing thereby reducing the adverse impact of socioeconomic conditions on COVID-19 health outcomes. Based on this discussion, our second hypothesis is that government policies regarding social distancing and income support can moderate the relationship between socioeconomic conditions and health outcomes during the COVID-19 pandemic.

For empirical analysis, we use daily data of COVID-19 confirmed cases and deaths, and government responses from January 22 to May 20, 2020 from 80 countries. We use two alternative proxies including confirmed cases per million people and deaths per million people to measure the national-level health outcomes during the COVID-19 pandemic. Country-level socioeconomic conditions are measured with the socioeconomic conditions index from the International Country Risk Guide dataset. Our empirical strategy is to examine how national-level COVID-19 outbreaks have developed in different socioeconomic settings. We find strong evidence that better socioeconomic conditions lead to superior health outcomes during the pandemic. Specifically, confirmed cases and deaths per million people from COVID-19 are lower in countries with better socioeconomic conditions. These findings are robust when we use education as an alternative proxy of socioeconomic conditions and growth in daily confirmed cases and deaths as alternative measures of health outcomes.

Next, we use interaction terms between socioeconomic conditions and government emergency policies and find that stringent social distancing measures and generous income support programs help to lower the confirmed cases and deaths primarily in countries with poor socioeconomic conditions.

We contribute to the recently emerging literature regarding COVID-19 in two important ways: First, susceptibility studies are being conducted to understand the factors which lead to different responses to the virus. In this regard, Sominsky *et al.* (2020) find that children are less likely to be severely affected by COVID-19. Zhao *et al.* (2020) show that individuals with blood group A are at higher risk to contract COVID-19 compared with individuals with non-A blood groups, whereas individuals with blood group O are at lower risk to get infected compared with non-O blood groups individuals. We add to this literature by showing that individuals living under poor socioeconomic factors are more likely to get infected and die from the virus.

Second, our study also complements recently emerging literature which examines the cross-country determinants of differences in COVID-19 health outcomes. In this regard, Ashraf (2020) examines the impact of political institutions on cross-country differences in deaths by COVID-19 and concludes that it’s too early to blame democracy for the higher

Covid Economics 37, 14 July 2020: 141-162

number of deaths in some countries. Frey *et al.* (2020) find that collectivist cultures result in better health outcomes during the pandemic because individuals from such cultures are more likely to follow social distancing measures. We extend this debate by examining the impact of socioeconomic conditions on the health outcomes during the pandemic.

The rest of the paper proceeds as follows: Section 2 outlines our sample construction procedures. Section 3 presents the empirical methodology briefly. Section 4 reports the results of empirical analyses. Final section concludes the study.

2. Sample construction

We started sample construction by collecting daily accumulated data of COVID-19 confirmed cases and deaths from the website of John Hopkins University, Coronavirus Resource Centre (JHU, CRC). This data was available for more than 200 countries/regions of the world over the period from January 22 to May 20, 2020. Next, we downloaded daily data of government response indexes from the Oxford COVID-19 Government Response Tracker (OxCGRT) database (Hale *et al.* 2020). Lastly, we collected country-level data of main socioeconomic conditions and other control variables from different sources including International Country Risk Guide (ICRG) of The PRS Group (ICRG 2020), World Development Indicators (WDI) database of World Bank and The Polity Project (Marshall & Gurr 2020).

We appended three datasets together to construct the main study sample. Then we refined the main sample first by dropping the countries with missing important data and second by dropping the specific observations with missing values. Our final sample consists of a panel dataset consisting of 9529 daily observations from 80 countries over the period from January 22 to May 20, 2020. Table 1 presents basic information about the sample.

3. Methodology

We specify the following pooled panel ordinary least square regression model to examine the impact of socioeconomic conditions on the health outcomes by COVID-19. Pooled panel model can better capture the impact of country-level factors on micro-level variables (Ashraf 2017; Ashraf & Shen 2019). Since we want to estimate the impact of country-level socioeconomic conditions on individual-level health outcomes, a pooled panel model is a preferred estimation technique for our analysis.

Covid Economics 37, 14 July 2020: 141-162

Covid Economics 37, 14 July 2020: 141-162

Table 1: Sample distribution and basic COVID-19 statistics for sample countries

This table reports the sample countries, the numbers of COVID-19 confirmed cases and deaths in each country on May 20, 2020, confirmed cases and deaths per million people for each country on May 20, 2020, and values of socioeconomic conditions index from the year 2018. It also reports the number of daily observations from each country over the period from January 22 to May 20, 2020.

Sr. No.	Country	Confirmed cases	Deaths	Confirmed cases per million people	Deaths per million people	Socioeconomic conditions index	Observations
1	Albania	964	31	332	11	5.5	120
2	Algeria	7542	568	184	14	5.63	120
3	Argentina	9283	403	211	9	6.5	120
4	Australia	7081	100	283	4	9.25	107
5	Austria	16353	633	1858	72	9.25	120
6	Azerbaijan	3631	43	367	4	6.63	120
7	Bahrain	7888	12	5259	8	6.5	120
8	Bangladesh	20995	314	131	2	2.75	116
9	Belarus	32426	179	3413	19	4.5	120
10	Belgium	55983	9150	5089	832	9.21	120
11	Brazil	271885	17983	1295	86	5.88	119
12	Bulgaria	2292	116	323	16	6.08	120
13	Burkina Faso	796	51	42	3	3.5	117
14	Cameroon	3733	146	149	6	2.5	120
15	Canada	80476	6027	2175	163	9.71	119
16	Chile	43781	450	2432	25	7.33	117
17	China	84063	4638	60	3	7.75	120
18	Colombia	17687	630	361	13	4.79	120
19	Costa Rica	897	10	183	2	6.96	120
20	Cote d'Ivoire	2231	29	93	1	3.67	120
21	Croatia	2234	96	545	23	6.17	104
22	Cuba	1900	79	173	7	6.21	120
23	Cyprus	910	17	758	14	8.08	115
24	Czech Republic	8455	296	769	27	8.96	116
25	Denmark	11117	554	1917	96	9.33	120
26	Dominican Republ	11739	424	1067	39	5.33	115
27	Egypt	14229	680	148	7	4.5	120
28	El Salvador	1413	30	221	5	5.5	118
29	Estonia	1794	64	1380	49	8	120
30	Finland	6443	304	1171	55	9.21	120
31	France	179069	28084	2673	419	8.92	120
32	Germany	178473	8144	2150	98	10	120
33	Greece	2836	165	258	15	5.33	118
34	Honduras	2955	147	314	16	3	120
35	Hungary	3598	470	367	48	8.5	120
36	India	112028	3434	86	3	5.08	120
37	Indonesia	19189	1242	74	5	6.42	120
38	Ireland	24315	1571	5066	327	9.13	120
39	Israel	16667	279	1916	32	9.5	119
40	Italy	227364	32330	3727	530	8.04	119
41	Jordan	672	9	69	1	4.33	120
42	Kazakhstan	6969	35	387	2	7.13	120
43	Kenya	1029	50	21	1	2.13	120
44	Korea, South	11122	264	218	5	9.29	120
45	Kuwait	17568	124	4285	30	9.25	120
46	Lithuania	1577	60	563	21	7.29	120

47	Luxembourg	3971	109	6659	183	9.25	120
48	Madagascar	371	2	14	0	4.46	120
49	Malaysia	7009	114	226	4	10.25	120
50	Mexico	56594	6090	472	51	7.04	120
51	Moldova	6553	228	1872	65	4.5	120
52	Mongolia	140	0	45	0	5.17	120
53	Morocco	7133	194	198	5	6.5	118
54	Namibia	16	0	7	0	6	120
55	Netherlands	44447	5748	2615	338	9.71	120
56	New Zealand	1499	21	312	4	10.17	117
57	Niger	909	55	41	3	3	118
58	Norway	8281	234	1562	44	9.92	120
59	Oman	6043	30	1286	6	6.29	120
60	Peru	94933	2789	3062	90	5	118
61	Philippines	13221	842	120	8	4.79	120
62	Poland	19739	962	519	25	7.71	120
63	Portugal	29660	1263	2966	126	8.63	120
64	Qatar	37097	16	13740	6	8	120
65	Romania	17387	1147	869	57	5.58	120
66	Russia	308705	2972	2205	21	6.38	120
67	Saudi Arabia	62545	339	1895	10	6.04	120
68	Senegal	2714	30	181	2	4.5	120
69	Serbia	10833	235	1548	34	5.21	120
70	Singapore	29364	22	5244	4	9	120
71	Slovenia	1468	105	699	50	7.21	120
72	South Africa	18003	339	316	6	4.04	120
73	Spain	232555	27888	4948	593	7.04	120
74	Sri Lanka	1028	9	49	0	4.5	120
75	Sweden	31523	3831	3152	383	9.33	120
76	Switzerland	30618	1891	3602	222	10.13	119
77	Tunisia	1045	47	95	4	5.67	120
78	United Kingdom	248293	35704	3762	541	8.96	120
79	United States	1551853	93439	4703	283	10.04	120
80	Uruguay	746	20	219	6	7.38	120
	Total	55274.35	3839.75	1547	79	6.824	9529

$$Y_{c,t} = \alpha_c + \beta_1(Socioeconomic\ conditions_c) + \sum_{j=1}^{j} \beta_j X_{c,t}^{j} + \sum_{k=1}^{k} \beta_k X_c^k + \sum_{t=1}^{T-1} \epsilon_t D_t + \varepsilon_{c,t} \_\_\_\_\_ Eq.(1)$$

Here, c and t subscripts represent country and day, respectively. α_c is a constant term. Dependent variable, Y, represents health outcomes by COVID-19 in country c on day t. Socioeconomic conditions is the main variable of interest and represents cross-country differences in socioeconomic circumstances. $X_{c,t}^{j}$ is a set of country-level control variables measured at daily frequency.   is a set of country-level control variables measured as fixed over the sample period. D_t is a set of daily fixed-effects to control for the effect of international factors which might have influenced health outcomes in all sample countries,

Covid Economics 37, 14 July 2020: 141-162

such as the announcements of WHO guidelines, approvals of new medicines and findings of new research, among others. $\mathcal{E}_c$ is an error term. Heteroskedastic-robust standard errors are used to estimate p-values in regressions.

Specifically, we measure country-level health outcomes during the COVID-19 pandemic with two alternative proxies: confirmed cases per million people and deaths per million people for each country on a daily frequency. Higher values of these variables represent adverse health outcomes and vice versa. We do not use case fatality ratio, or alternatively deaths to confirmed cases ratio, to measure health outcomes because this proxy is influenced by a country's COVID-19 testing policy and thus is more problematic. The ratio is lower for countries which performed aggressive testing of suspected patients, even for the individuals with mild symptoms (Morris & Reuben 2020). Therefore, we prefer confirmed cases and deaths per million people as dependent variables in our study over the case fatality ratio.

Socioeconomic circumstances of countries are measured with the socioeconomic conditions index from the ICRG database. Socioeconomic conditions index is measured with three subcomponents including unemployment, poverty and consumer confidence. Index ranges from 0 to 12 where higher values represent better socioeconomic circumstances in terms of lower unemployment and poverty and higher consumer confidence, and vice versa.

The results estimated with pooled panel ordinary least squares regression model might be biased due to the problem of endogeneity. Endogeneity may arise due to reverse causality or omitted variables (Ashraf *et al.* 2016). To eliminate the concern of reverse causality, we ensure our proxy of socioeconomic conditions measures socioeconomic circumstances of countries in the pre-pandemic period. For doing so, we use values of socioeconomic conditions index from the year 2018. Finally, we saturate our regression model by including several related country-level control variables to reduce the omitted variables bias.

$X^j_{c,t}$, which is a set of country-level control variables measured at daily frequency, includes the outbreak stage, stringency index and economic support index. Since the countries which were ahead in local COVID-19 outbreaks than others might have observed higher total numbers of cases and deaths, therefore we include the outbreak stage variable in the model. This variable counts the number of days from the first confirmed case to the May 20th (the day until which we downloaded data) for each country. Stringency index represents government measures, such as closure of schools, workplaces and public places and restrictions on travel, to ensure social distancing among people. Economic support index

Covid Economics 37, 14 July 2020: 141-162

measures the extent of government income support and debt/contract relief to households. Both stringency and economic support indexes range from 0 to 100.

, which is a set of country-level fixed control variables, includes log(total population), log(GDP), urban population to total population ratio, international tourism (arrivals per year), percentage population aged between 15 to 64 years, government general health expenditure as a percentage of GDP, life expectancy at birth, gender ratio and polity. In this regard, log (total population) and log (GDP) control for country size and the level of economic development, respectively. Urban population to total population ratio is added to control for the level of countries' urbanization because contagious diseases like COVID-19 are easy to widespread at places with people in close proximity (Alirol *et al.* 2011). International tourism (annual arrivals) controls for the extent of international travel of countries. Countries with more international travel might have received large numbers of COVID-19 cases in the early stage of the pandemic and resultantly have faced more severe outbreaks later on. Further, since initial evidence suggests a younger population is less at risk due to COVID-19 (JHU-CRC 2020; Shi *et al.* 2020; Worldometers 2020), the percentage population aged between 15 to 64 years is added to control for the differences in age demographics of the countries. Moreover, government general health expenditure as a percentage of GDP controls for the effect that developed healthcare systems might have resulted in better health outcomes during the pandemic. Likewise, life expectancy at birth variable controls for the chances of individuals to live longer in some countries than others. Gender ratio, which equals the male to female ratio, controls for the factor that male members have a higher likelihood to die from the disease than females (Worldometers 2020). Finally, polity controls for the cross-country differences in political institutions.

Next, we modify Eq. (1) as follows to examine the impact of government interventions on the relationship between socioeconomic conditions and health outcomes during the COVID-19.

$$Y_{c,t} = \alpha_c + \beta_1(Socioeconomic\ conditions_c) + \beta_2(Socioeconomic\ conditions_c \times Government\ interventions_{c,t}) + \sum_{j=1}^{j} \beta_j X_{c,t}^j + \sum_{k=1}^{k} \beta_k X_c^k + \sum_{t=1}^{T-1} \epsilon_t D_t + \varepsilon_{c,t} \_\_\_\_\_ Eq.(2)$$

Covid Economics 37, 14 July 2020: 141-162

Here, the interaction term, $Socioeconomic\ conditions_c \times Government\ interventions_{c,t}$, is the main variable of interest where estimated values of β_2 show the impact of government interventions on the relationship between socioeconomic conditions and COVID-19 health outcomes. Specifically, we consider government interventions regarding social distancing measures and income support packages, measured with stringency index and economic support index, respectively, and interact them with socioeconomic conditions index one-by-one. Other variables are the same as in Eq. (1).

4. Empirical analysis

In this section, we present the results of our empirical analysis. First, we report summary statistics and then main regression results.

4.1 Basic sample information and summary statistics of main variables

Table 1 reports the list of sample countries as well as the numbers of confirmed cases and deaths by COVID-19 in each country on May 20, 2020, the last date of our sample period. It also reports the confirmed cases and deaths per million of population for each country on May 20, 2020. The numbers of confirmed cases per million people are higher in countries with relatively small size such as Qatar (13740 cases), Luxemburg (6659), Bahrain (5259), Singapore (5244) and Belgium (5089). On the other hand, the numbers of deaths per million people are higher in relatively developed countries such as Belgium (832 deaths), Spain (593), United Kingdom (541) and Italy (530). Socioeconomic conditions vary by the large extent in sample countries. For instance, Kenya, Cameroon and Bangladesh have the lowest values of 2.13, 2.5 and 2.75, respectively, for the socioeconomic conditions index and thus are countries with the poorest socioeconomic conditions in the sample. On the other hand, Malaysia, New Zealand and Switzerland have the values of 10.25, 10.17 and 10.13, respectively, and are the countries with the best socioeconomic conditions. We exploit this sample diversity to investigate the impact of socioeconomic circumstances on health outcomes during the COVID-19 pandemic.

Table 2 reports the full sample summary statistics of main variables. Confirmed cases and deaths per million people have mean values equal to 407 and 21.99, respectively. Socioeconomic conditions index has a mean value of 6.8 and a standard deviation of 2.15 across mean value. Likewise, control variables also exhibit significant variation.

Covid Economics 37, 14 July 2020: 141-162

Table 2: Summary statistics

This table reports the full sample summary statistics of main variables.

Variables	Mean	Standard deviation	Minimum	Maximum
Confirmed cases per million people	407.371	1009.853	0	13739.629
Deaths per million people	21.997	79.822	0	831.818
Socioeconomic conditions index	6.823	2.152	2.130	10.250
Education	55.119	27.031	3.733	136.603
Outbreak stage	60.292	34.569	1	120
Stringency index	47.104	36.354	0	100
Economic support index	26.491	33.611	0	100
Log (population)	16.617	1.498	13.299	21.060
Log (GDP)	25.937	1.703	22.818	30.601
Urban population to total population ratio	69.406	19.180	16.350	100
International tourism (annual arrivals)	1.30e+07	1.81e+07	1.43e+05	8.68e+07
Percentage population aged between 15 to 64 years	65.801	5.556	47.316	85.257
Govt. general health expenditure as a percentage of GDP	4.424	2.301	0.380	10.475
Life expectancy at birth	76.036	6.086	57.017	83.551
Gender ratio	1.054	0.017	1.011	1.130
Polity	5.546	6.223	-10	10

4.2 Socioeconomic conditions and health outcomes during COVID-19

Next, we estimate Eq. (1) with panel pooled ordinary least squares regressions to examine the impact of socioeconomic circumstances on the numbers of confirmed cases and deaths by the COVID-19. As shown in Table 3, socioeconomic conditions index enters negative, with strong statistical significance at 1% level, with both confirmed cases and deaths per million people. These results suggest that poor socioeconomic conditions result in higher infections and deaths by COVID-19 disease and confirm our first hypothesis.

Quantitatively, the results in Model (1) suggest that in response to one standard deviation increase in socioeconomic conditions index (2.15), the confirmed cases pre million people drop by 204 (-94.59 × 2.15) from its mean value of 407. Likewise, in Model (2), deaths per million people drop by 10 (-4.70 × 2.15) from the mean value of 22. These figures imply that a one standard deviation improvement in socioeconomic conditions index lowers COVID-19 confirmed cases and deaths by one half.

Covid Economics 37, 14 July 2020: 141-162

CEPR PRESS

Covid Economics 37, 14 July 2020: 141-162

Table 3: Socioeconomic conditions and health outcomes during COVID-19: main specifications

This table presents results regarding the impact of socioeconomic conditions on health outcomes during the COVID-19 pandemic. Dependent variable is confirmed cases per million people in Model 1 and deaths per million people in Model 2, and represents health outcomes. Socioeconomic conditions index is the main independent variable of interest and measures cross-country differences in socioeconomic circumstances. Others are control variables. Both models include country-level control variables as well as time fixed-effects dummy variables. Results are estimated with panel pooled ordinary least squares regression method using heteroskedasticity robust standard errors. P-values are given in parenthesis. ***, **,* represent statistical significance at 1%, 5%, and 10% levels, respectively.

Variables	Confirmed cases per million people Model (1)	Deaths per million people Model (2)
Socioeconomic conditions index	-54.950***	-4.696***
	(0.000)	(0.000)
Control variables		
Outbreak stage	12.075***	0.621***
	(0.000)	(0.000)
Economic support index	8.535***	0.602***
	(0.000)	(0.000)
Stringency index	-4.352***	-0.326***
	(0.000)	(0.000)
Log (population)	-433.386***	-11.774***
	(0.000)	(0.000)
Urban population to total population ratio	-3.423***	0.073
	(0.000)	(0.150)
Log (GDP)	363.846***	10.323***
	(0.000)	(0.000)
Percentage population aged between 15 to 64 years	-7.164***	-2.475***
	(0.006)	(0.000)
International tourism (annual arrivals)	0.000***	0.000***
	(0.000)	(0.000)
Govt. general health expenditure as a percentage of GDP	-58.055***	0.099
	(0.000)	(0.804)
Gender ratio	-2710.814***	-98.544***
	(0.000)	(0.001)
Life expectancy at birth	1.781	1.618***
	(0.419)	(0.000)
Polity	-12.323***	-0.303***
	(0.000)	(0.001)
Time fixed-effects dummies	*Yes*	*Yes*
Constant	1757.385***	83.031**
	(0.000)	(0.015)
Observations	9,529	9,529
R-squared	0.430	0.314

Results of control variables are also consistent with the expectation. For instance, the outbreak stage enters positively and significantly showing that countries which were ahead in the early phase of the global outbreak have observed higher numbers of confirmed cases and deaths. Stringency index enters negative suggesting that government response in terms of stringent lockdowns and social distancing measures has proved effective in controlling the spread of highly contagious virus. The share of population aged between 15 to 64 years variable enters negative and significant showing that countries with relatively young populations have fewer infections and deaths. On the contrary, countries with higher numbers of international arrivals have seen the worst outbreaks as the international tourism (arrivals per year) variable enters positive and significant. Life expectancy at birth enters positively. One expected reason behind this finding is that countries with longer life expectancy have a higher share of the old population which is at higher risk due to the specific nature of COVID-19 to attack elderly individuals.

We perform two types of robustness tests to further confirm the above main results. First, we use education as an alternative proxy of socioeconomic conditions. For doing so, we measure cross-country differences in the level of education with the percentage enrolment at tertiary level educational institutions. Higher values of this variable indicate that countries have relatively higher numbers of people with tertiary level education. Like socioeconomic conditions index, education is also measured with its values from the year 2018 to avoid reverse causality problem. We prefer the education variable, measured as the percentage enrolment at tertiary level educational institutions, over the literacy rate to measure the cross-country differences in the level of education. Literacy rate counts the percentage of individuals who can read and write. In contrast, the higher values of percentage enrolment at tertiary level educational institutions indicate that relatively more individuals can afford better education. Further the individuals with a strong educational background are in a better position to acknowledge the risks of the disease. As shown in Table 4[1], the education variable also enters negatively and significantly showing that the countries with higher education levels experience lower numbers of cases and deaths. We also add socioeconomic conditions index together with the education in Table 4 and observe that both variables enter negative and significant further confirming the main results.

[1] For brevity, we only report coefficients of main variables of interest.

Covid Economics 37, 14 July 2020: 141-162

Table 4: Socioeconomic conditions and health outcomes during COVID-19: robustness tests

This table presents robustness results regarding the impact of socioeconomic conditions on health outcomes during the COVID-19 pandemic. Dependent variable is confirmed cases per million people in Models 1-2 and deaths per million people in Models 3-4, and represents health outcomes. Education variable is the main independent variable of interest and is used as an alternative proxy of socioeconomic conditions. Socioeconomic conditions index also measures cross-country differences in socioeconomic circumstances. All models include country-level control variables as well as time fixed-effects dummy variables. Results are estimated with panel pooled ordinary least squares regression method using heteroskedasticity robust standard errors. P-values are given in parenthesis. ***, **,* represent statistical significance at 1%, 5%, and 10% levels, respectively.

Variables	Deaths per million people		Confirmed cases per million people	
	Model (1)	Model (2)	Model (3)	Model (4)
Education	-0.435***	-0.528***	-8.691***	-10.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Socioeconomic conditions index		-7.019***		-98.951***
		(0.000)		(0.000)
Control variables	*Yes*	*Yes*	*Yes*	*Yes*
Time fixed-effects dummies	*Yes*	*Yes*	*Yes*	*Yes*
Constant	56.859	-152.610***	247.314	-2705.847***
	(0.112)	(0.000)	(0.570)	(0.000)
Observations	9,529	9,529	9,529	9,529
R-squared	0.320	0.326	0.449	0.457

Second, we use daily growth in confirmed cases and deaths as alternative proxies to measure health outcomes during COVID-19. Higher growth in daily confirmed cases and deaths imply more severe outbreaks. We use both growth in confirmed cases and deaths as dependent variables one-by-one and re-estimate Eq. (1). As shown in Table 5, socioeconomic conditions index enters negative and statistically significant with both variables. These results again confirm our hypothesis 1 that poor socioeconomic conditions lead to adverse health outcomes during the pandemic.

Covid Economics 37, 14 July 2020: 141-162

Table 5: Socioeconomic conditions and health outcomes during COVID-19: robustness tests

This table presents robustness results regarding the impact of socioeconomic conditions on health outcomes during the COVID-19 pandemic. Dependent variable is daily growth in confirmed cases in Model 1 and daily growth in deaths in Model 2, and represents health outcomes. Socioeconomic conditions index is the main independent variable of interest and measures cross-country differences in socioeconomic circumstances. All models include country-level control variables as well as time fixed-effects dummy variables. Results are estimated with panel pooled ordinary least squares regression method using heteroskedasticity robust standard errors. P-values are given in parenthesis. ***, **,* represent statistical significance at 1%, 5%, and 10% levels, respectively.

Variables	Growth in confirmed cases Model (1)	Growth in deaths Model (2)
Socioeconomic conditions index	-0.013***	-0.005**
	(0.010)	(0.029)
Control variables	*Yes*	*Yes*
Time fixed-effects dummies	*Yes*	*Yes*
Constant	0.785*	0.914***
	(0.072)	(0.003)
Observations	6,819	4,992
R-squared	0.136	0.180

4.3 Government interventions and the impact of socioeconomic conditions on health outcomes during COVID-19

Next, we estimate Eq. (2) to examine the impact of government interventions on the relationship between socioeconomic conditions and health outcomes during the COVID-19. As shown in Table 6, all four interaction terms between socioeconomic conditions index and both government response indexes enter positive and significant with both dependent variables. These results show that government interventions weaken the negative association between socioeconomic circumstances and health outcomes during the COVID-19 pandemic.

Covid Economics 37, 14 July 2020: 141-162

Covid Economics 37, 14 July 2020: 141-162

Table 6: Socioeconomic conditions and health outcomes during COVID-19: moderating effect of government interventions

This table presents results regarding the interaction effect of socioeconomic conditions and government interventions on health outcomes during the COVID-19 pandemic. Dependent variable is confirmed cases per million people in Models 1-2 and deaths per million people in Models 3-4, and represents health outcomes. Socioeconomic conditions index × Stringency index and Socioeconomic conditions index × Economic support index are the main variables of interest and represent the joint effect of socioeconomic conditions and government interventions on health outcomes. Socioeconomic conditions index measures the cross-country differences in socioeconomic circumstances. All models include country-level control variables as well as time fixed-effects dummy variables. Results are estimated with panel pooled ordinary least squares regression method using heteroskedasticity robust standard errors. P-values are given in parenthesis. ***, **,* represent statistical significance at 1%, 5%, and 10% levels, respectively.

Variables	Confirmed cases per million people		Deaths per million people	
	Model (1)	Model (2)	Model (3)	Model (4)
Socioeconomic conditions index	-268.726***	-222.145***	-15.884***	-14.829***
	(0.000)	(0.000)	(0.000)	(0.000)
Socioeconomic conditions index × Stringency index	3.387***		0.177***	
	(0.000)		(0.000)	
Socioeconomic conditions index × Economic support index		4.555***		0.289***
		(0.000)		(0.000)
Control variables				
Outbreak stage	12.395***	10.948***	0.638***	0.541***
	(0.000)	(0.000)	(0.000)	(0.000)
Economic support index	3.453***	-27.046***	0.336***	-1.647***
	(0.000)	(0.000)	(0.000)	(0.000)
Stringency index	-23.752***	-0.500	-1.340***	-0.073*
	(0.000)	(0.311)	(0.000)	(0.065)
Log (population)	-573.154***	-520.630***	-19.130***	-16.244***
	(0.000)	(0.000)	(0.000)	(0.000)
Urban population to total population ratio	0.349	2.779***	0.272***	0.425***
	(0.545)	(0.000)	(0.000)	(0.000)
Log (GDP)	516.534***	448.839***	18.359***	14.551***
	(0.000)	(0.000)	(0.000)	(0.000)
Percentage population aged between 15 to 64 years	-11.285***	-2.414	-2.692***	-2.129***
	(0.000)	(0.297)	(0.000)	(0.000)
International tourism (annual arrivals)	0.000***	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Govt. general health expenditure as a percentage of GDP	-44.298***	-50.081***	0.821**	0.541
	(0.000)	(0.000)	(0.040)	(0.169)
Gender ratio	361.404	1760.356***	63.575**	153.843***
	(0.426)	(0.000)	(0.037)	(0.000)
Life expectancy at birth	19.640***	15.560***	2.558***	2.356***
	(0.000)	(0.000)	(0.000)	(0.000)
Polity	5.636***	6.257***	0.641***	0.767***
	(0.000)	(0.000)	(0.000)	(0.000)
Education	-11.077***	-9.706***	-0.584***	-0.509***
	(0.000)	(0.000)	(0.000)	(0.000)
Time fixed-effect dummies	*Yes*	*Yes*	*Yes*	*Yes*
Constant	-2552.676***	-3977.194***	-144.612***	-233.216***
	(0.000)	(0.000)	(0.000)	(0.000)
Observations	9,529	9,529	9,529	9,529
R-squared	0.513	0.537	0.351	0.378

Following Ashraf *et al.* (2020), we use graphical approach to explain the results of interaction terms. For doing so, we graph the relationship between socioeconomic conditions index and the confirmed cases/deaths per million people at mean and $\pm$ one standard deviation of mean value of two government intervention indexes, one-by-one. Graphs 1, 2, 3 and 4 in Figure (1) are drawn from Models 1 to 4, respectively, of Table 6.

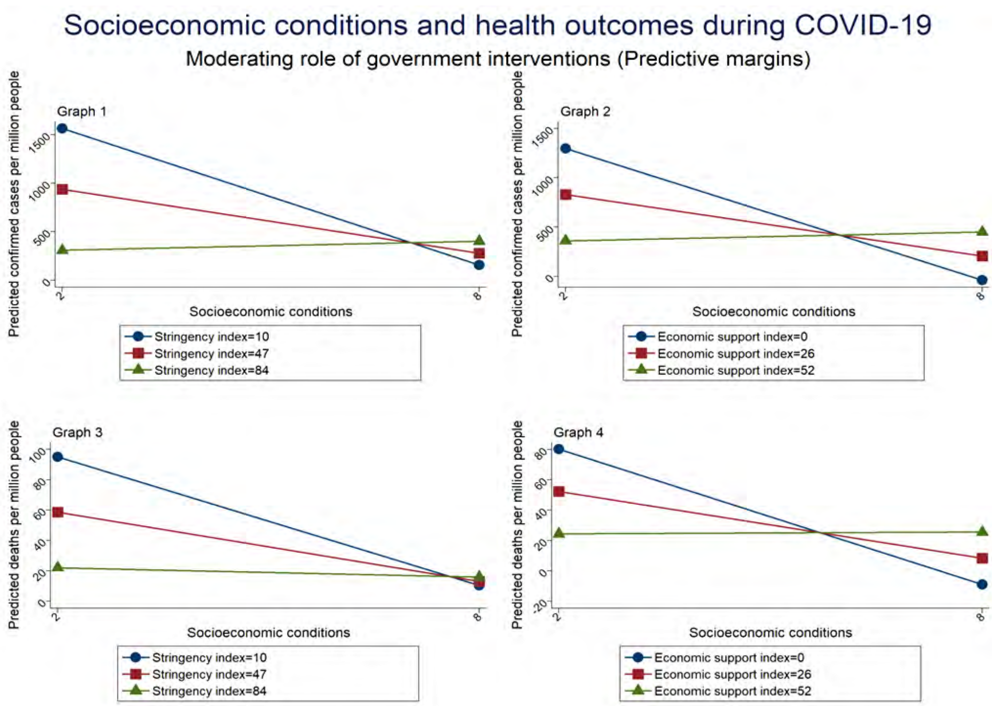


Figure 1 Impact of socioeconomic conditions on health outcomes during COVID-19: moderating effect of government interventions

When government interventions are less stringent, the upper downward sloped lines (with embedded circles) show that as socioeconomic conditions improve from left to right, the confirmed cases and deaths decrease. In contrast, when government interventions are stringent, slopes of lower lines (with embedded triangles) become flat indicating that strict government social distancing measures and generous economic support lower the negative impact of socioeconomic conditions on health outcomes. Specifically, graphs show that government interventions primarily reduce confirmed cases and deaths in countries with poor socioeconomic conditions. For instance, in Graph 1, the predicted confirmed cases are

Covid Economics 37, 14 July 2020: 141-162

CEPR PRESS

slightly higher than 1500 per million people when socioeconomic conditions index equals 2 and stringency index equals 10. However, confirmed cases drop to around 400 per million people if the stringency index increases to 84 at the same level of socioeconomic conditions index. Likewise, in Graph 3, deaths drop from around 100 to 20 per million people when stringency index increases from 10 to 84 at the same level of socioeconomic conditions index. Together, these results suggest that government interventions are effective in reducing the adverse impact of poor socioeconomic conditions on health outcomes during the pandemic.

5. Conclusion

In this paper, we examine the impact of socioeconomic conditions on the health outcomes by COVD-19 in a cross-country setting. We also analyze how government emergency policies moderate the impact of socioeconomic conditions on COVID-19 outbreaks. For empirical analysis, we collect a panel dataset consisting of 9529 daily observations from 80 countries over the period from January 22 to May 20, 2020. COVID-19 health outcomes are measured with two alternative proxies including the confirmed cases per million people and deaths per million people. Country-level socioeconomic circumstances are measured with socioeconomic conditions index from ICRG database. We find that socioeconomic circumstances have strong negative association with COIVD-19 health outcomes. That is, the numbers of confirmed cases and deaths are higher in countries with poor socioeconomic conditions. In quantitative terms, a one standard deviation improvement in socioeconomic conditions lowers COVID-19 confirmed cases and deaths per million people by one half. We also observe that the countries with relatively young populations experienced lower numbers of cases and deaths, while the countries with higher tourist arrivals experienced higher numbers. We observe our findings are robust when we use education as an alternative proxy of socioeconomic circumstances and growth in daily confirmed cases and deaths as alternative measures of health outcomes.

Next, with the help of interaction terms between socioeconomic conditions and government emergency policies, we find that stringent social distancing measures and generous income support programs help to lower the confirmed cases and deaths in countries with poor socioeconomic conditions.

These findings provide useful policy implications at the current stage when COVID-19 cases are increasing in poor socioeconomic settings such as Latin America, South Asia and African continent. Use of an appropriate set of social distancing measures together with

Covid Economics 37, 14 July 2020: 141-162

income support to households might help to curb the severity of the outbreaks. These findings also imply that governments can design specific policies for the areas with poor socioeconomic conditions with potential to become COVID-19 hotspots.

Our findings are likely to be of great interest to the scientific community, governments and general public amid the COVID-19 crisis.

Covid Economics 37, 14 July 2020: 141-162

References

Alirol, E., Getaz, L., Stoll, B., Chappuis, F., Loutan, L., 2011. Urbanisation and infectious diseases in a globalised world. The Lancet Infectious Diseases 11, 131-141

Anand, S., Ravallion, M., 1993. Human development in poor countries: on the role of private incomes and public services. Journal of economic perspectives 7, 133-150

Anderson, R.M., Heesterbeek, H., Klinkenberg, D., Hollingsworth, T.D., 2020. How will country-based mitigation measures influence the course of the COVID-19 epidemic? The Lancet 395, 931-934

Ashraf, B.N., 2017. Political Institutions and Bank Risk-Taking Behavior. Journal of Financial Stability 29, 13-35

Ashraf, B.N., 2020. Devastation Caused by COVID-19: Is Democracy to Blame? Available at SSRN 3596009

Ashraf, B.N., Shen, Y., 2019. Economic policy uncertainty and banks' loan pricing. Journal of Financial Stability 44, 100695

Ashraf, B.N., Zheng, C., Arshad, S., 2016. Effects of national culture on bank risk-taking behavior. Research in International Business and Finance 37, 309-326

Ashraf, B.N., Zheng, C., Jiang, C., Qian, N., 2020. Capital regulation, deposit insurance and bank risk: International evidence from normal and crisis periods. Research in International Business and Finance 52

Bingenheimer, J.B., Brennan, R.T., Earls, F.J., 2005. Firearm violence exposure and serious violent behavior. Science 308, 1323-6

Braveman, P., Egerter, S., Williams, D.R., 2011. The social determinants of health: coming of age. Annual review of public health 32, 381-398

Braveman, P., Gottlieb, L., 2014. The social determinants of health: it's time to consider the causes of the causes. Public health reports 129, 19-31

Brown, P., 1995. Race, class, and environmental health: a review and systematization of the literature. Environ Res 69, 15-30

Cook, W.K., Heller, J., Bhatia, R., Farhang, L., 2009. A health impact assessment of the Healthy Families Act of 2009. Oakland (CA): Human Impact Partners and San Francisco Department of Public Health

Cummins, S., Macintyre, S., 2006. Food environments and obesity--neighbourhood or nation? Int J Epidemiol 35, 100-4

Evans, G.W., Schamberg, M.A., 2009. Childhood poverty, chronic stress, and adult working memory. Proc Natl Acad Sci U S A 106, 6545-9

Fanelli, D., Piazza, F., 2020. Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos, Solitons & Fractals 134, 109761

Covid Economics 37, 14 July 2020: 141-162

Frey, C.B., Chen, C., Presidente, G., 2020. Democracy, Culture, and Contagion: Political Regimes and Countries Responsiveness to Covid-19. Covid Economics 18, 222-238

Gordon-Larsen, P., Nelson, M.C., Page, P., Popkin, B.M., 2006. Inequality in the built environment underlies key health disparities in physical activity and obesity. Pediatrics 117, 417-424

Hale, L., 2005. Who has time to sleep? J Public Health (Oxf) 27, 205-11

Hale, T., Webster, S., Petherick, A., Phillips, T., Kira, B., 2020. Oxford covid-19 government response tracker. Blavatnik School of Government 25

ICRG, 2020. The International Country Risk Guide (ICRG). The PRS Group (available at https://www.prsgroup.com/explore-our-products/international-country-risk-guide/)

JHU-CRC, 2020. How does mortality differ across countries? https://coronavirus.jhu.edu/data/mortality

Lanphear, B.P., Kahn, R.S., Berger, O., Auinger, P., Bortnick, S.M., Nahhas, R.W., 2001. Contribution of residential exposures to asthma in us children and adolescents. Pediatrics 107, E98

Lidsky, T.I., Schneider, J.S., 2003. Lead neurotoxicity in children: basic mechanisms and clinical correlates. Brain 126, 5-19

Lou, J., Shen, X., Niemeier, D., 2020. Are stay-at-home orders more difficult to follow for low-income groups? Working Paper

Marco, C.A., Wolfson, A.R., Sparling, M., Azuaje, A., 2011. Family socioeconomic status and sleep patterns of young adolescents. Behav Sleep Med 10, 70-80

Marshall, M.G., Gurr, T.R., 2020. Polity V Project: Political regime characteristics and transitions, 1800–2018. Dataset User's Manual Center for Systemic Peace

McEwen, B.S., Gianaros, P.J., 2010. Central role of the brain in stress and adaptation: links to socioeconomic status, health, and disease. Ann N Y Acad Sci 1186, 190-222

Morris, C., Reuben, A., 2020. Coronavirus: Why are international comparisons difficult? In: BBC News (available at https://www.bbc.com/news/52311014)

Okell, L.C., Verity, R., Watson, O.J., Mishra, S., Walker, P., Whittaker, C., Katzourakis, A., Donnelly, C.A., Riley, S., Ghani, A.C., 2020. Have deaths from COVID-19 in Europe plateaued due to herd immunity? Lancet (London, England)

Pollack, C.E., Cubbin, C., Ahn, D., Winkleby, M., 2005. Neighbourhood deprivation and alcohol consumption: does the availability of alcohol play a role? Int J Epidemiol 34, 772-80

Shi, Y., Yu, X., Zhao, H., Wang, H., Zhao, R., Sheng, J., 2020. Host susceptibility to severe COVID-19 and establishment of a host risk score: findings of 487 cases outside Wuhan. Critical Care 24, 108

Covid Economics 37, 14 July 2020: 141-162

Sominsky, L., Walker, D.W., Spencer, S.J., 2020. One size does not fit all–Patterns of vulnerability and resilience in the COVID-19 pandemic and why heterogeneity of disease matters. Brain, behavior, and immunity

Tung, J., Barreiro, L.B., Johnson, Z.P., Hansen, K.D., Michopoulos, V., Toufexis, D., Michelini, K., Wilson, M.E., Gilad, Y., 2012. Social environment is associated with gene regulatory variation in the rhesus macaque immune system. Proc Natl Acad Sci U S A 109, 6490-5

Worldometers, 2020. Age, Sex, Existing Conditions of COVID-19 Cases and Deaths. https://www.worldometers.info/coronavirus/coronavirus-age-sex-demographics/

Wright, A.L., Sonin, K., Driscoll, J., Wilson, J., 2020. Poverty and economic dislocation reduce compliance with covid-19 shelter-in-place protocols. University of Chicago, Becker Friedman Institute for Economics Working Paper

Zhao, J., Yang, Y., Huang, H., Li, D., Gu, D., Lu, X., Zhang, Z., Liu, L., Liu, T., Liu, Y., He, Y., Sun, B., Wei, M., Yang, G., Wang, X., Zhang, L., Zhou, X., Xing, M., Wang, P.G., 2020. Relationship between the ABO Blood Group and the COVID-19 Susceptibility. medRxiv, 2020.03.11.20031096

Covid Economics 37, 14 July 2020: 141-162