The Natural and Unnatural Histories of Covid-19 Contagion

Michael Beenstock

Department of Economics

Hebrew University of Jerusalem

Xieer Dai

Department of Geography

Hebrew University of Jerusalem

April 26, 2020

We study the epidemiology of Covid-19 using an overlapping-generations method in which successive cohorts of infectives are temporarily contagious. We use this method to estimate R0 (natural rate of contagion) and Re (effective rate of contagion) for Covid-19 in various countries and over time using data on confirmed morbidity and estimates of unconfirmed morbidity. We use these estimates to study the effect of mitigation policy on Re. We show that even in the absence of mitigation policy Re tends to decrease by week 3 of the epidemic due to endogenous social distancing. Several methods for estimating the treatment effect of mitigation policy on R suggest that mitigation policy accelerates the decrease in Re. A "Chinese crystal ball" method is proposed for projecting and simulating contagion with an empirical illustration for Israel.

The natural history of contagion is expressed by the basic reproductive number R0, which is the average number of susceptibles that infectives are expected to infect until they cease to be contagious. However, the natural history of contagion is rarely observed because it is concealed by mitigating action undertaken by individuals and governments, such as quarantine, social distancing, travel restrictions and lockdown. The unnatural history of contagion, measured by the effective rate of contagion (Re), is observed instead, which is inevitably smaller than its natural counterpart. This applies to Covid-19 in particular because in contrast to previous epidemics, mitigation policy has been unusually activist.

In this paper we make three contributions. First, we propose an "overlapping generations method" for calculating R. We argue that this method provides more reliable estimates of R than the canonical SIR (susceptible – infective – removed) method widely used by epidemiologists since 1927 (Bailey 1975). It also calculates R in real time. Second, the true number of infectives is typically much larger than the measured number due to delays in diagnosis and because carriers may be completely asymptomatic (Bendavid et al 2020). We propose a method to impute the true number of infectives from official data on infectives. Third, we study the causal effect of mitigation policy on R. We introduce the concept of "corona equilibrium" in which mitigation and R are mutually dependent. To resolve the identification problem for estimating treatment effects in corona equilibrium, we explore a number of methodologies, including triple differences-in-differences, vector autoregressions, and a leading indicator method, which we call the "Chinese crystal ball".

Measuring R in Real Time

Overlapping Generations Method

Let τ denote the infective period measured in days, and suppose that infectives are diagnosed immediately. "Corona time" begins on day 0 when the first infectives are diagnosed. We denote the true number of ever-infectives by C and corona days by t, hence C_0 is the initial number of infectives. We assume that the daily rate of contagion is $R0/\tau$ so that after τ days infectives infect R0 susceptibles, i.e. the temporal distribution of contagiousness is uniform. The number of new infectives on day 1 is C_0R0/τ, so that the number of infectives at the end of day 1 is $C_1 = C_0$

+C_0R0/τ. Until $t = \tau$ the number of infectives is $C_t = (1 + R0/\tau)C_{t-1}$. Repeated backward substitution implies that $C_t = C_0(1 + R0/\tau)^t$ and $C_\tau = C_0(1 + R0/\tau)^\tau$.

On day $\tau + 1$ the original infectives cease to be contagious. Henceforth, we distinguish between the ever-infected (C) and contagious carriers (C*), or infectives. According to the "perpetual inventory method" the number of contagious infectives equals the sum of the changes in the number diagnosed over the previous τ days. Hence, $C_t^* = C_t - C_{t-\tau} = \Delta_\tau C_\tau$. The number of ever-infected by the end of day $\tau + 1$ is $C_{\tau+1} = (1 + R0/\tau)(C_\tau - C_0)$ where the last term is the number of infectives on day τ. By day 2τ all infectives up to day τ cease to be contagious. More generally, $C_{\tau+j} = (1+ R0/\tau)(C_{\tau+j-1} - \Delta C_{j-1})$, which implies that by day 2τ the number of ever-infected is:

$$C_{2\tau} = C_\tau\left(1 + \frac{R0}{\tau}\right)^\tau - \frac{R0}{\tau}\left(1 + \frac{R0}{\tau}\right)^{\tau-1} C_0 \qquad (1)$$

Suppose initially there is only one infective ($C_0 = 1$), infectives are contagious for 14 days ($\tau = 14$), and by the end of day 14 there are 70 infectives ($C_\tau = 70$), the implicit value for R0 during the first "generation" is 4.9636, which is the average rate of infection during the first 14 days of the epidemic. If this rate of infection applied over the next 14 days, equation (1) implies that the number of ever-infected by the end of day 28 of the epidemic will be 4881.7. If the number of ever-infected is less than this, R0 during the second generation must have decreased. For example, if the number is 4000, equation (1) implies that R0 during days 14 – 28 is 4.6968 instead of 4.9636.

$R0/\tau$ refers to the daily rate of contagion (ROD), which varies directly with R0 and inversely with the duration of contagion (τ). The number of ever-infected obviously varies directly with the daily rate of contagion since:

$$\frac{\partial C_t}{\partial ROD} = \ln(1 + ROD)\,[C_t - (1 + ROD)C_{t-\tau-1}] > 0 \qquad (2)$$

Where $t = \tau + j$. Equation (2) is positive because the term in square brackets is positive. Given ROD, the number of ever-infected varies directly, but perhaps less obviously, with the duration of contagion since:

$$\frac{\partial C_t}{\partial \tau} = R0Dln(1 + R0D)C_{t-\tau-1} \; > 0 \qquad (3)$$

Finally, when ROD is not given, the separate effects of R0 and are determined by:

$$dC_t = \ln(1 + R0D)\,(C_t - C_{t-\tau-1})\frac{1}{\tau}(dR0 - R0Dd\tau) \qquad (4)$$

The number of ever-infected varies directly with R0 and inversely with the duration of contagion (τ). The reason for this apparent paradox is that when τ increases, the daily rate of contagion (R0D) decreases. Therefore, what matters is not R0, but the daily rate of contagion and the duration of contagion.

The counterpart of equation (1) for generation g is:

$$C_{g.\tau} = C_{g-1.\tau}\left(1 + \frac{R0}{\tau}\right)^{\tau} - \frac{R0}{\tau}\left(1 + \frac{R0}{\tau}\right)^{\tau-1} C_{(g-2)0} \qquad (5)$$

Where at the end of generation g, gτ = t, g-1.τ = t - τ and g-2.1 = t - 2τ. In summary, this is a τ- order overlapping generations (OLG) model in which infectives are contagious for τ days. These implicit values may vary over time, as indicated, and may be calculated for further generations of infectives as the epidemic continues. In particular, the OLG method may be used to calculate how Re changes after mitigating actions such as travel bans, lockdowns etc. We have ignored the fact that the number of susceptibles decreases over time because in the early stages of epidemics the ever-infected constitute a tiny fraction of the population.

The "doubling time" (the time in which C doubles) is equal to $ln2/\ln(1 + R0D)$. For example, if R0 = 3 and τ = 14 days, the doubling time is 3.23 days. The doubling time varies inversely with R and directly with τ. Doubling time has become a popular concept for measuring contagion during the corona epidemic.

OLG Estimates of R using Morbidity Data

Many health authorities have set isolation periods at two weeks, suggesting that τ is 14 days. However, clinical evidence suggests that viral shedding may have a mean of 20 days (Zhou et al 2020) and that contagiousness my not be uniformly distributed. Since OLG estimates the daily rate of contagion (R0/τ), the estimate of R0 varies proportionately with τ. In the previous numerical example, in which the doubling time is 3.23 days, R0 would have been 4.286 instead of 3 if τ is 20 instead of 14 days. Hence, given the OLG estimate of R0D, R0 is proportionate to τ.

Figures 1 – 3 report rolling OLG estimates of R based on official morbidity data as of April 24 for a variety of countries and Chinese provinces assuming τ is 14 days. In Figures 1 and 2 time is measured in corona days where day 0 occurs when morbidity is 100 in large countries and proportionately smaller in less populated countries. This is why the last data point appears to refer to different days. In Figure 1 the epidemic occurred earliest in Italy and latest in Russia. In Figure 2 the epidemic occurred earliest in Singapore and latest in Sweden. In Figure 3A the horizontal axis is measured in calendar time. The epidemic occurred earliest in Hubei province and latest in Xingjiang province.

Although R in Figures 1 – 3 vary by country, they share a common dynamic profile. R initially increases steeply before decreasing gradually. Recall that if R exceeds 1 the number of infectives increases, and it decreases when R is less than 1. In the latter case, the battle is won. When R is zero the war is won. Several countries are on course for winning the battle, while the war appears to have been won in most provinces of China and in S. Korea. Japan and Singapore are exceptions in that R did not increase steeply, it has fluctuated around 1 in Japan, and it has increased gradually to 1.5 in Singapore.

Figure 1 Rates of Contagion

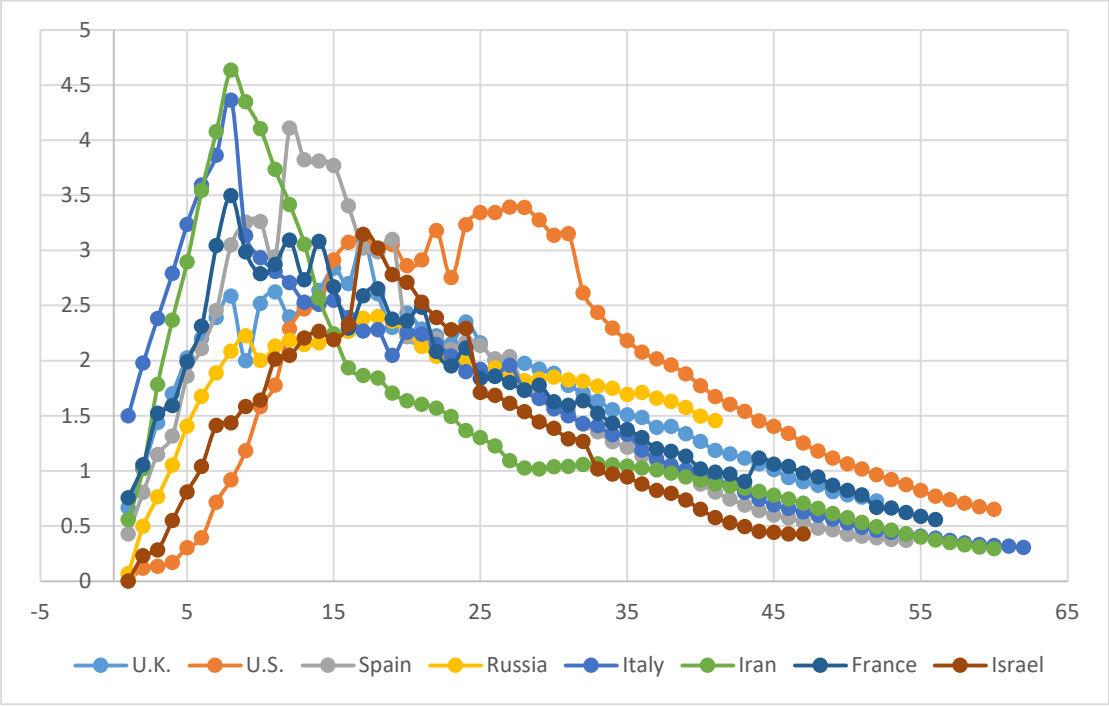
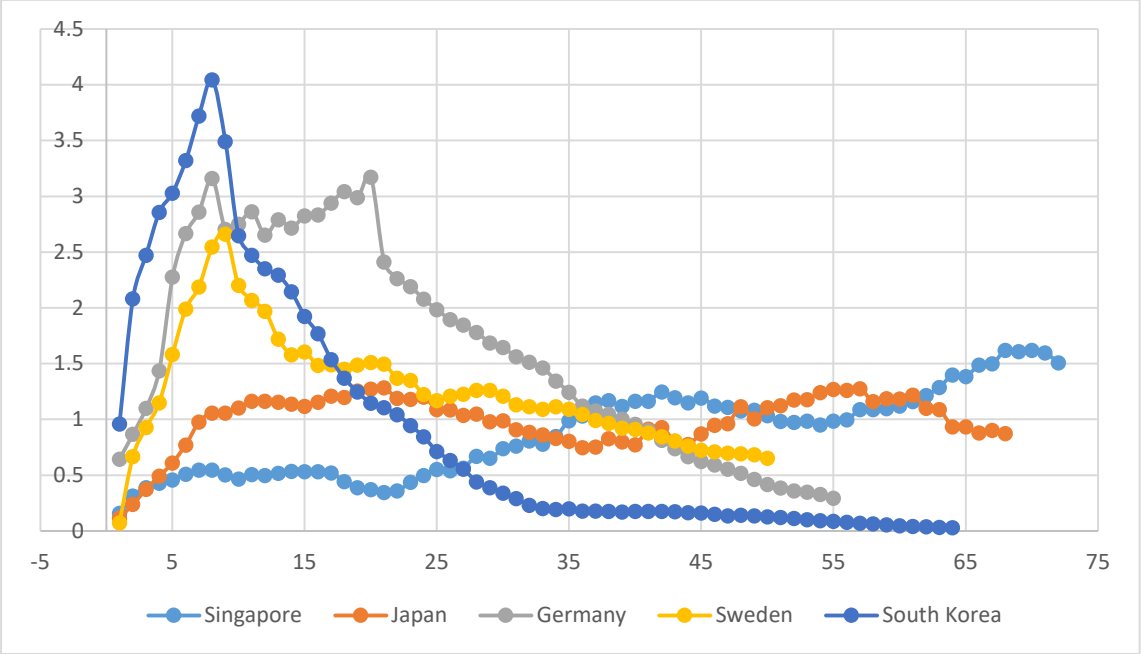


Figure 2 Rates of Contagion in Countries with Minimal Mitigation



Sweden has been a focus of international attention because she has undertaken minimal mitigation (Figure 6). Nevertheless, the dynamic profile for R has followed the international trend, and is currently less than 1. Unfortunately, Sweden is the only example. Hence, it is difficult to judge whether Sweden is an outlier, or whether it may serve as a counterfactual for what might have happened had countries not practiced mitigation, as discussed further below. Also, South Korea has followed a policy of mass testing and isolation, which has enabled it to reduce R to zero without intensive mitigation.

China has had the longest experience with Covid-19. Indeed, we argue below that in many respects, we may look to China to project the future course of the epidemic in other countries. In March, lockdown policy that had been in force since January 24 began to be reversed, and by late April had been reversed completely. Therefore, it is particularly interesting to ask whether the epidemic resurged subsequently. Figure 3B seems to suggest that in some provinces resurgence occurred, especially in Hong Kong and Inner Mongolia. However, almost all of this resurgence was imported because Chinese citizens abroad, who were unable to return to China beforehand, were allowed to return in March. Indeed, there have been few cases of indigenous contagion. In the vast majority of provinces R continues to be almost zero more than a month after the reversal of lockdown.

Figure 3A Rates of Contagion in China

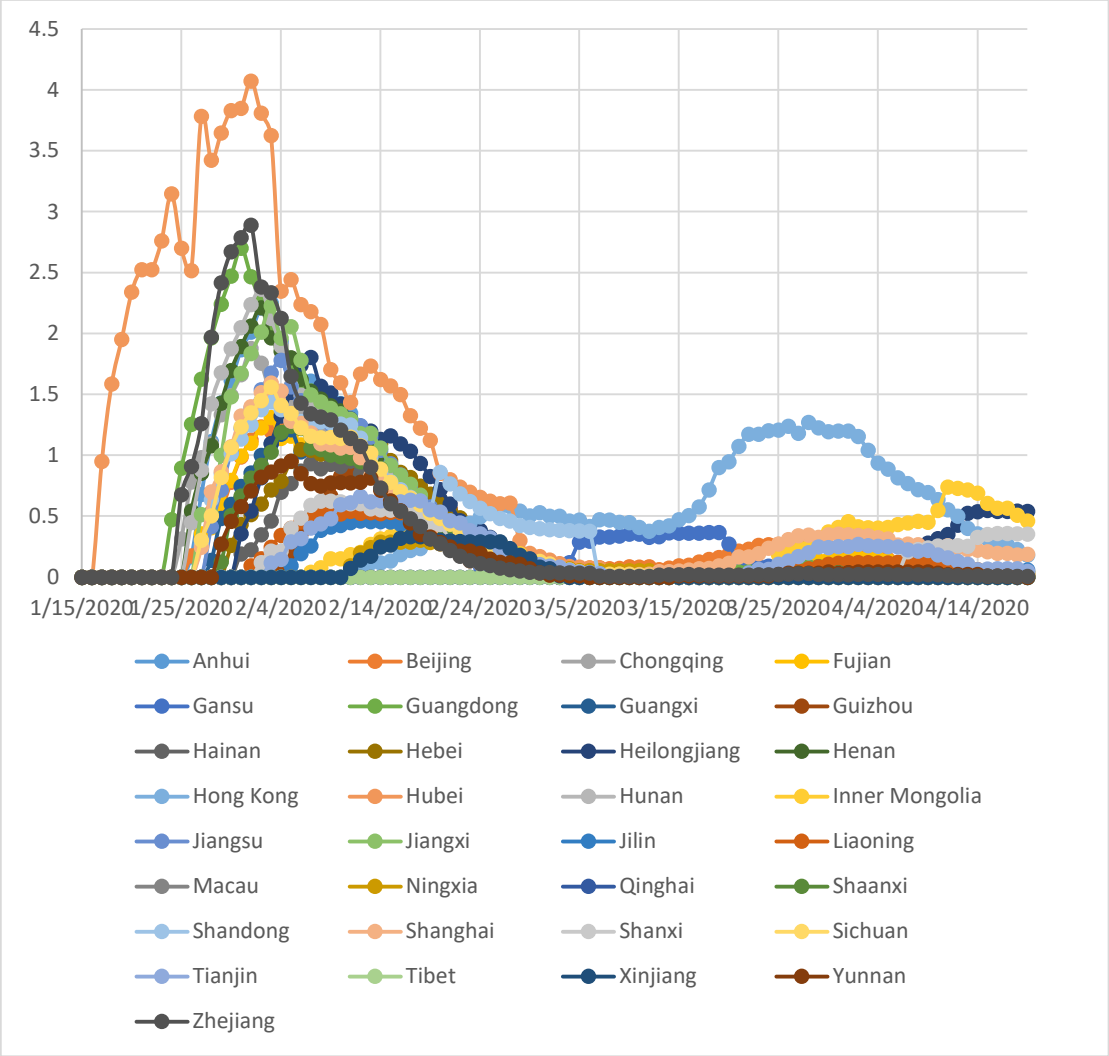
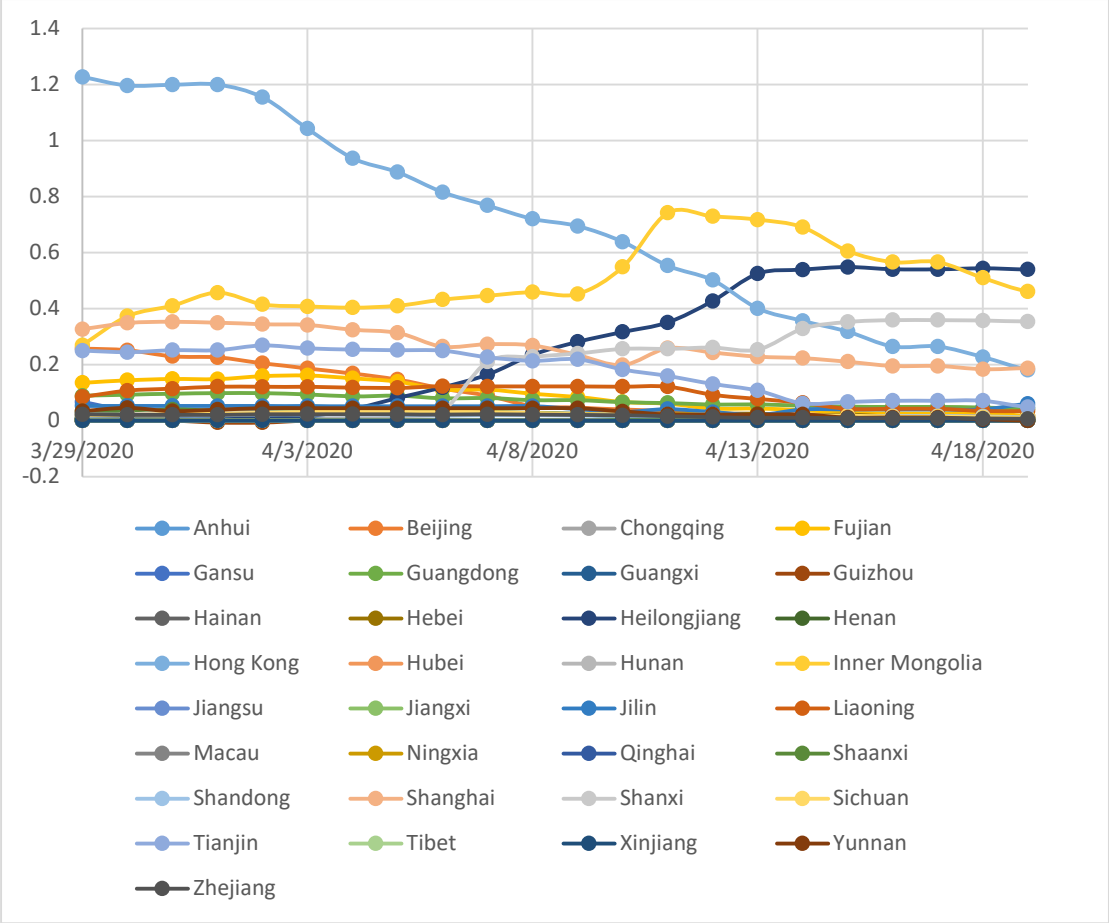


Figure 3B Contagion in China since the Reversal of Lockdown



Diagnostic Delay and Hidden Infectives

Suppose infectives are not diagnosed immediately so that the true number of ever-infected (C) is greater than the number of diagnosed infected denoted by D. We assume that a proportion ϕ of infectives are never diagnosed; they remain hidden as "corona zombies", and that on each day a proportion θ of other infectives become symptomatic and are diagnosed:

$$\Delta D_t = \theta[(1-\phi)C_{t-1} - D_{t-1}] \qquad (6)$$

The terms in square brackets is the backlog of asymptomatic and symptomatic infectives yet to be diagnosed. Note that θ will depend on the testing capacity of the health authorities.

The general solution for the number of diagnosed cases is:

$$D_t = A(1-\theta)^t + \theta(1-\phi)\sum_{j=0}^{t-1}(1-\theta)^j C_{t-1-j} \qquad (7)$$

Where A is an arbitrary constant determined by initial conditions D_0, and C_t is determined as in equation (1). Equation (7) states the number of diagnosed infected lags behind the number of carriers, of whom some will never be diagnosed. If, for expositional simplicity there are C infectives, equation (7) simplifies to:

$$D_t = A(1-\theta)^t + (1-\phi)C \qquad (8)$$

Where $A = D_0 - (1 - \phi)C$. Since $1 - \theta$ is a positive fraction and $\phi - 1$ is a negative fraction, equation (8) states that the number of diagnosed cases converges from below, as expected, to the number of non-hidden infectives. It implies that if x percent of them are diagnosed within d days θ would be:

$$\theta = 1 - \left[\frac{(1-\phi)(x-1)C}{A}\right]^{1/d} \qquad (9)$$

If, according to the Director of CDC, 25 percent of infectives are corona zombies ($\phi = 0.25$), there are a hundred infectives ($C = 100$) of which 5 have been diagnosed ($D_0 = 5$), equation (9) implies that 90 percent ($x = 0.9$) will be diagnosed within 10 days, if $\theta = 0.2$.

By reverse engineering, equation (6) implies that the true number of infectives is:

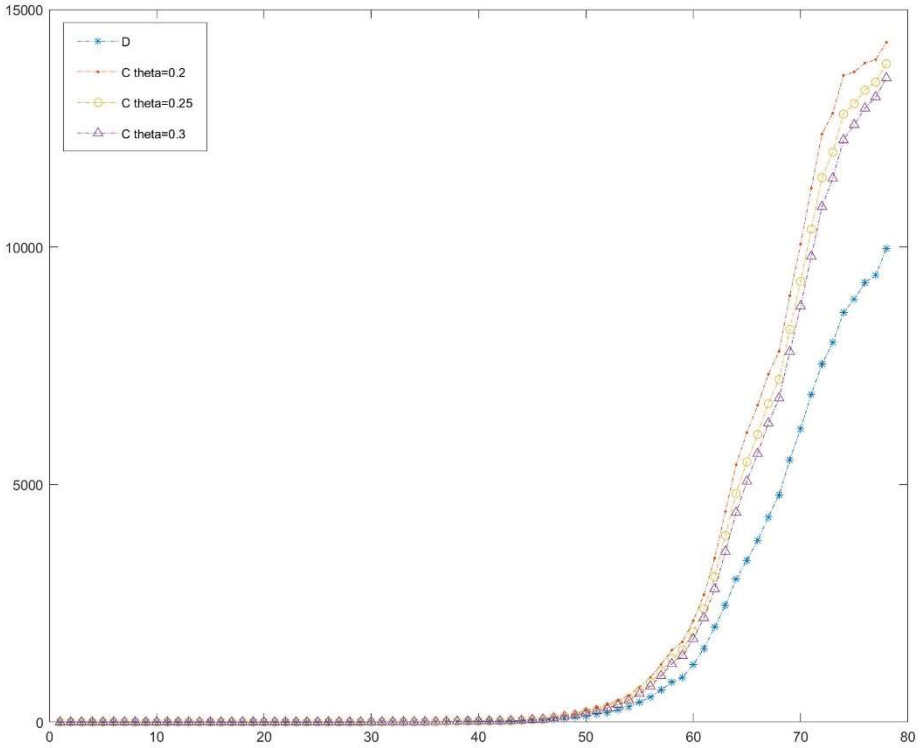
$$C_{t-1} = \frac{1}{1-\phi}\left(\frac{\Delta D_t}{\theta} + D_{t-1}\right) \qquad (10)$$

For example, if $\phi = 0.25$, $\theta = 0.2$, the number of diagnosed cases yesterday was 10,000 (D_{t-1}) and there are 800 newly diagnosed cases today (ΔD_t), the true number of infectives yesterday is 18,666 (C_{t-1}). In principle ϕ may be estimating by randomized testing for covid-19 in the population. Icelandic test data, which are not randomized, suggest that 0.43 percent of the population are asymptomatic carriers of Covid-19. Without supplementary data on the proportion of carriers who eventually became symptomatic, randomized testing provides at best an upper bound for ϕ.

In Figure 4A equation (10) is used to calculate C using Israeli morbidity data for D, setting $\phi = 0.25$ for different values of θ. For these purposes a 4th order moving average of ΔD is specified to ensure that C cannot decrease. The difference between D and C naturally varies inversely with θ. By the end of the period (April 10) the number of diagnosed cases under-estimated the number of undiagnosed cases by

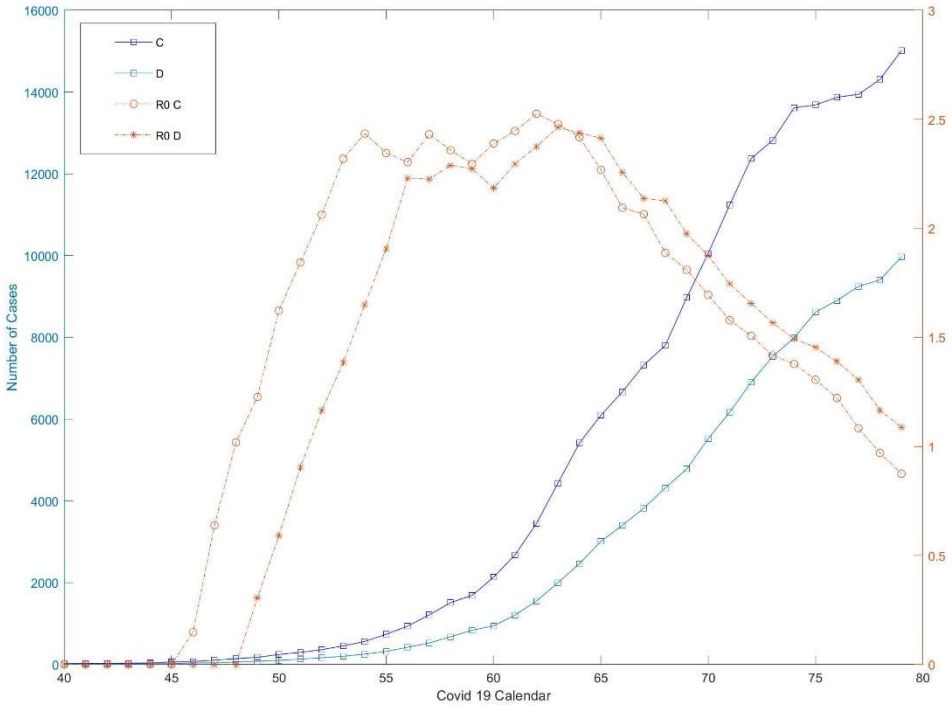
about 29 percent. By construction 25 percent were corona zombies (since $\phi = 0.25$) and 4 percent would have been diagnosed eventually.

Figure 4A Diagnosed and Undiagnosed Covid-19 Carriers in Israel



In Figure 4B we report OLG estimates of R using the imputed data for C from Figure 4A, which are compared with their counterparts using official morbidity data (D). As expected, R is initially larger for C than it is for D. The former precedes the latter by about 5 days. However, the trends are similar, and both measures peak at the same level of R. Subsequently, R based on C is less than R based on D. This means that just as conventional estimates of R understate the rate of contagion when R is increasing, they overstate it when R decreases. It also means that although R based on D lags behind R based on C, the former provides a reliable indicator of trends and turning points in contagion.

Figure 4B OLG Estimates of R including Asymptomatics (Israel)



Quarantine

Let q denote the proportion of infectives in quarantine. Because quarantine shortens the period during which infectives are contagious, the OLG model becomes:

$$C_t = [1 + (1 - q_{t-1})R0D](C_{t-1} - C_{t-\tau-1}) \qquad (11)$$

It is obvious that quarantine mitigates the propagation of coronavirus because it reduces the daily rate of contagion. Since only diagnosed infectives can be quarantined, the proportion of infectives in quarantine at the beginning of day t (the end of day t-1) is assumed to be $q_{t-1} = D_{t-1}/C_{t-1}$. This assumes that once diagnosed, infectives are quarantined.

We modify equation (6) by defining D in terms of the diagnosed who are contagious instead of the ever-diagnosed by subtracting patients who recovered or died:

$$\Delta D_t = \theta[(1 - \phi)C_{t-1} - D_{t-1}] - rD_{t-1} \qquad (12)$$

where the recovery/mortality rate is denoted by r. In the early stages of the epidemic D is zero, because it takes time for infectives to be diagnosed, in which case q is initially zero. Subsequently, q becomes positive, which reduces the propagation of the

epidemic through equation (11). Substituting for q_{t-1} equation (11) may be rewritten as:

$$\Delta C_t = R0D(C_{t-1} - D_{t-1}) - C_{t-\tau-1} - \frac{C_{t-1} - D_{t-1}}{C_{t-1}} C_{t-\tau-1} \qquad (13)$$

Equations (12) and (13) are simultaneous difference equations, which solve for true infectives (C) and diagnosed infectives (D). Because equation (12) is linear, but equation (13) is nonlinear, we are unable to obtain analytical solutions. Nevertheless, they obviously imply that q grows over time from zero to less than one if R0D is positive. This means that OLG estimates of R0 tend to decrease after q becomes positive, as evidenced in Figures 1 – 3.

As suggested by Chudik, Pesaran and Rebucci (2020) q may also increase endogenously. As the public forms expectations about R, it spontaneously engages in social distancing with the result that R begins to decrease. Hence, the profiles in Figures 1 – 3 may occur spontaneously as in Sweden.

Regression Methods and Measurement Error

Suppose regression methods were used to estimate R0D. Denoting 1 + R0D by β, equation (1) suggests that β may be estimated by regression:

$$C_t = \beta(C_{t-1} - C_{t-1-\tau}) + u_t \qquad (14)$$

Where u ~ iiN(0, σ^2) denotes a residual error. Since C is unknown, β is estimated using morbidity data for D, which contains measurement error (m):

$$D_t = (1 - \phi)C_t + m_t \qquad (15a)$$

$$m_t = \rho m_{t-1} + e_t \qquad (15b)$$

Where e is distributed iiN with zero mean. Equations (6) implies that measurement error will be autocorrelated as in equation (15b). Since the unconditional expectation of m is zero, equations (15) imply that morbidity (D) lags behind its true value (C), of which only 1 - ϕ percent of them are eventually diagnosed. Substituting equation (15a) into equation (14) implies that the regression model for D is:

$$D_t = \beta(D_{t-1} - D_{t-1-\tau}) + w_t \qquad (16a)$$

$$w_t = (1 - \phi)u_t + m_t - \beta(m_{t-1} - m_{t-1-\tau}) \qquad (16b)$$

From equation (15b) the sign of the covariance between D_{t-1} and m_{t-1} is ρ - β and the sign of the covariance between $D_{t-1-\tau}$ and $m_{t-1-\tau}$ is negative. Hence, least squares estimates of β will be attenuated (biased downwards) especially if ρ is less than β. Since estimates of β are attenuated, so are estimates of the daily rate of contagion (R0D).

Regression methods may yield unbiased estimates of β provided that the measurement error model is taken into consideration. Substituting equation (6) with iid residual error d into equation (11), implies the following autoregressive model for D:

$$\Delta D_t = -\theta D_{t-1} + \beta\Delta_\tau \Delta D_{t-1} + \beta\theta\Delta_\tau D_{t-2} + \omega_t \qquad (17a)$$

$$\omega_t = \theta(1-\phi)u_{t-1} + d_t - \beta\Delta_\tau d_{t-1} \qquad (17b)$$

Where Δ_τ denotes the "seasonal" difference operator ($\Delta_\tau X_t = X_t - X_{t-\tau}$). Since u and v are iid, so is ω. Therefore, constrained least squares estimates of β and θ are consistent. Note that equation (17a) delivers consistent estimates of β and θ. Hence, it identifies that daily rate of contagion (R0D) and the rate at which Covid-19 carriers are diagnosed (θ). Furthermore, it almost identifies the proportion of Covid-19 zombies (ϕ) since the variance of ω is:

$$\sigma_\omega^2 = \theta^2(1-\phi)^2\sigma_u^2 + (1+2\beta)\sigma_d^2 \qquad (17c)$$

Using estimates of θ and β from equation (17a) and σ_ω^2, combinations of ϕ and the variances of u and d are identified.

Comparison of OLG with SIR

The OLG method is conceptually different to the canonical SIR methodology, which is widely used to estimate R0. Whereas OLG directly calculates R0 from the data, SIR infers R0 by estimating a statistical model in which the entry rate into infection and the exit rate from infection are assumed to be exponential. The SIR estimate of R0 is the ratio of the entry rate to the exit rate. There are two methodological problems with this. If the entry and exit rates are not exponential, the estimate of R0 will be biased. Exponential exit implies that some patients recover almost instantaneously, while others recover extremely slowly. Evidence on recovery from Covid-19 suggests that recovery rates are not exponential (Zhou et al 2020). Second, exit rates are estimated by regressing recoveries on lagged infectives. Since, as discussed above, the latter

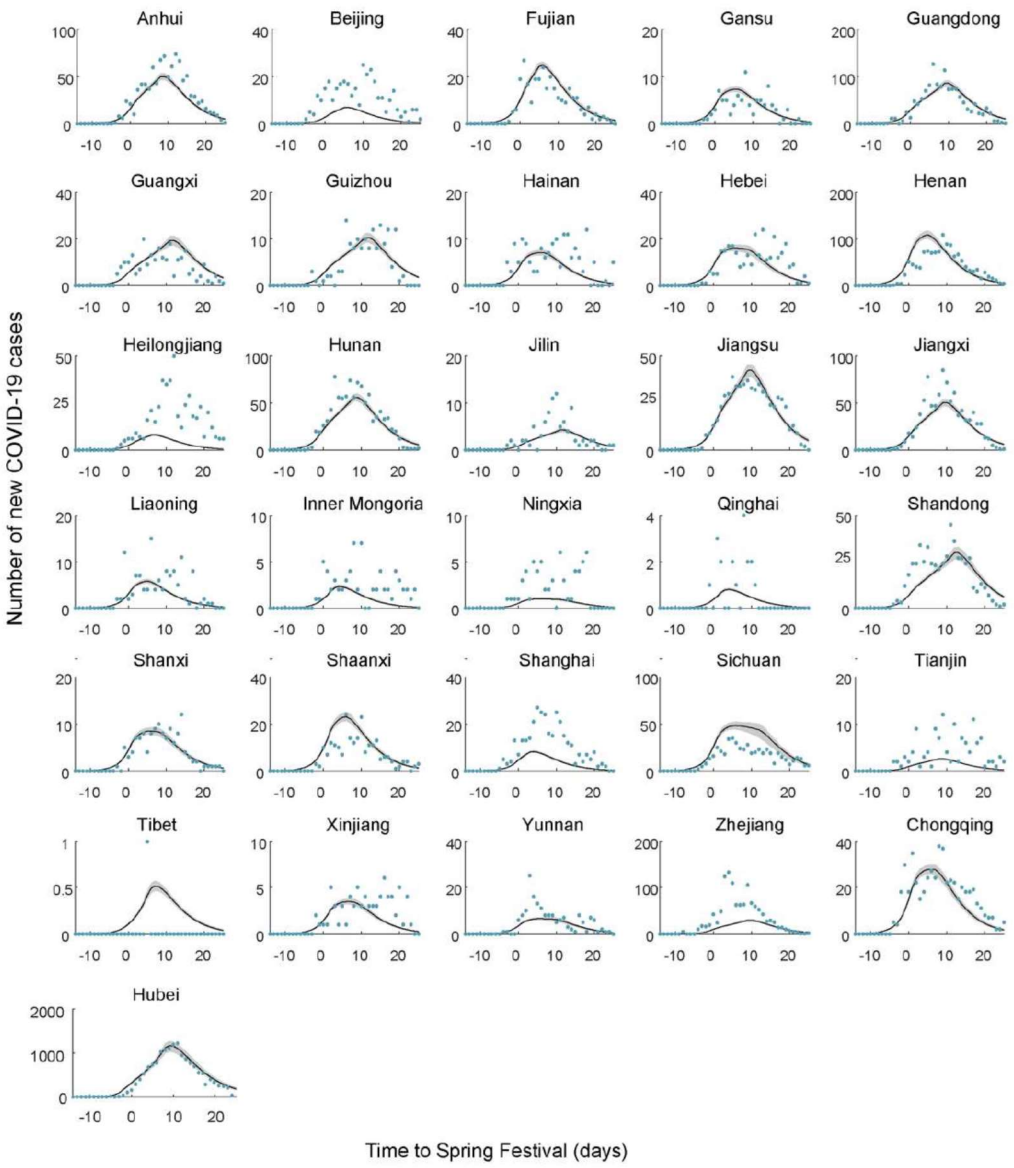
generally contain measurement error, the estimated exit rate is affected by attenuation bias. The SIR estimate of R0 involves both types of bias. A third problem, which arises during the mature stage of epidemics is that susceptibility to Covid-19 is expected to decrease if the least resistant succumb first while the most resistant may not succumb at all. SIR usually ignores the effect of heterogeneity in resistance.

By contrast OLG does not involve statistical methods for inference. Hence, it is less sensitive to measurement error than SIR. On the other hand, OLG makes assumptions about τ. However, as already noted, the nature of the bias is known and equals the ratio of τ to its true value.

A further methodological criticism of SIR concerns the widespread use of the Metropolis – Hastings algorithm used to estimate Bayesian Markov-Chain Monte Carlo (MC^2) models (Berg 2004, Hamra, MacLehose and Richardson 2013). The starting point is that the model is correct, but its parameters are unknown and need to be estimated using data. The method is Bayesian in the sense that priors are chosen for these parameters, such as R0. It is first-order Markov because, for example, recoveries from Covid-19 today are assumed to be strictly proportional to the number of yesterday's infected. Hence the day before yesterday etc does not matter. It is Monte Carlo based because random shocks to exits and entries have to be sampled (using the Gibbs sampler). Parameter estimates for exit and entry rates, and their standard deviations, are represented by their posterior distributions. The latter are generally influenced by their priors and their anterior distributions.

If the priors are chosen badly, so might their MC^2 estimates. Since, recovery from Covid-19 is not exponential, a higher order Markov chain is required. Most importantly, since the model is assumed to be correct, it is never tested empirically. For example, Tian et al (2020) estimate R0 and other parameters in a SIR-type model by MC^2. However, their model poorly fits the data for Covid-19 incidence in almost all provinces in China (Figure 5). The posterior estimates of the entry and exit rates may be statistically significant to Bayesians but not to classical, frequentist statisticians. Wu, Leung and Leung (2020) apply MC^2 to data for Wuhan, and Liu et al (2020) show that estimates of R0 vary widely (from 2 to 6.5) even using the same methodology.

Figure 5 Example of Badness-of-fit of SIR Models Estimated by MC^2



Source: Table SM3 from Tian et al (2020).

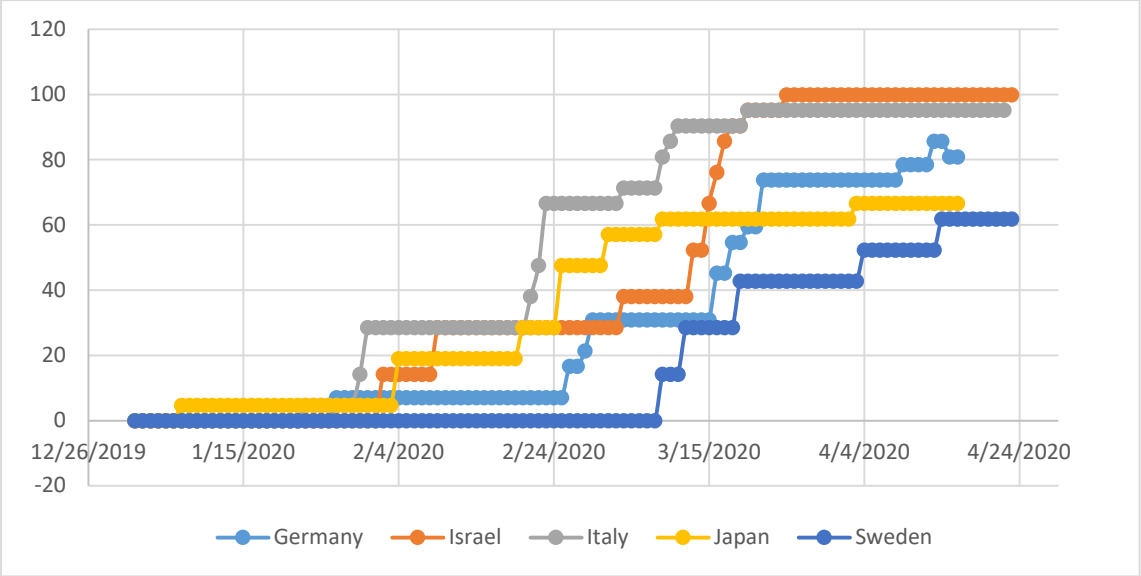
The Effect of Mitigation Policy on Contagion

Ideally, the study of infectious diseases should take place after they have occurred. Although four months have passed since coronavirus broke out in China, most countries have short histories of coronavirus (Figures 1 and 3). These histories get longer by the day, but they are arguably too short to study the treatment effects of mitigation policy on contagion especially because it took time before governments took mitigating action, and because it takes time for mitigation to affect morbidity

through its effect on Re. Many countries ramped-up their mitigation policies during the second half of March (Figures 5 and 6), so it is too early to judge their treatment effects. However, other countries, especially China as well as Spain and Italy undertook such policies earlier.

Consequently, the empirical results we report on the efficacy of mitigation policy are inevitably reported with reservation. On the other hand, mitigation policy in China seems to have stopped coronavirus after two months (Figure 3). Within five weeks Re began to decrease towards 1 in all Chinese provinces, and within ten weeks it continued to decrease towards zero. We therefore look for early signs that the same is happening in other countries. However, whereas in China mitigation policy was very stringent from the start, policy elsewhere was initially more variegated, but eventually became increasingly stringent.

Figure 6 OxCGRT Stringency Scores for Mitigation Policy



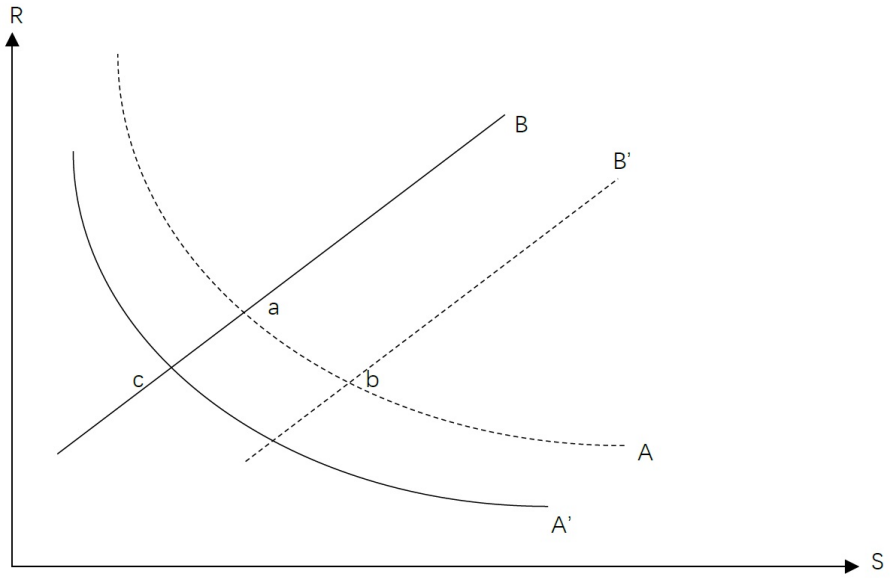
Corona Equilibrium and the Identification of Treatment Effects for Mitigation

We use the stringency index (S) developed by Oxford Covid-19 Government Response Tracker (OxCGRT) to measure mitigation policy in different countries (Petherick, Hale and Phillips 2020). S ranges from 0 to 100 where the latter is maximal stringency. Our purpose is to identify the causal effect of mitigation policy on R when there may be reverse causality from contagion to mitigation policy. Indeed, in countries such as Japan and Singapore where Covid-19 prevalence was

low, mitigation policy has been minimal. By contrast, in China, Italy and Spain where prevalence was large, mitigation policy has been maximal.

Figure 6 illustrates the development of mitigation policy over time in a selection of countries. All countries have tended to ramp-up their mitigation policy over time, but some have done this sooner rather than later. Even Sweden and Japan, which have resisted mitigation have joined the trend. Nevertheless, mitigation in Sweden and Japan continues to be relatively light.

Figure 7 Corona Equilibrium



The identification problem is illustrated in Figure 7 where R is measured on the vertical axis and S on the horizontal. Schedule A plots the hypothesized negative causal effect of mitigation policy on R. The size of the treatment effect varies directly with its slope. Schedule B plots the policy response of mitigation to the rate of contagion. It expresses the idea that if morbidity and mortality are small, governments will mitigate less if at all. Governments more politically prone to mitigation will have flatter B schedules. The data for R and S are jointly determined at point a, where schedules A and B intersect. We refer to this as the "corona equilibrium"

Since the treatment effect of mitigation is reflected in the slope of schedule A, we ideally wish to apply an autonomous increase in S such that schedule B shifts to the right to B' in which the new corona equilibrium would be at point b. For example,

countries with fewer ICUs (Italy?) may resort to more stringent mitigation as in schedule B. The slope of schedule A between points a and b is the local average treatment effect (LATE).

Suppose the population in another country is less susceptible to corona e.g. because of indigenous social distancing as in schedule A', i.e. R is smaller given S. The corona equilibrium will be determined at point c instead of point a, at which both R and S are smaller (Sweden?). The unwitting might think that R is smaller because S is smaller, but this confuses cause and effect. The corona equilibrium in countries with low susceptibility (A') and plentiful ICUs (B) may involve zero mitigation.

To identify the treatment effect of mitigation policy on contagion, we experiment with three different methods. The first involves triple differences (Berck and Villas-Boas 2016). Next, we use time series data to estimate a VAR model for R and S in Israel from which we calculate the impulse response function from shocks to mitigation policy, represented by S, onto contagion measured by R. Third, we estimate a leading indicator model for R in Israel, which is dynamically related to R in Chinese provinces and S in Israel. We refer to the latter as the "Chinese crystal ball", because it exploits the longer history of contagion in China to make inferences regarding the much shorter history of contagion in Israel. Absent is the canonical instrumental variables method, in which mitigation stringency is randomized using instrumental variables for S. We have been unable to find statistically significant instruments for the timing and stringency of mitigation policy.

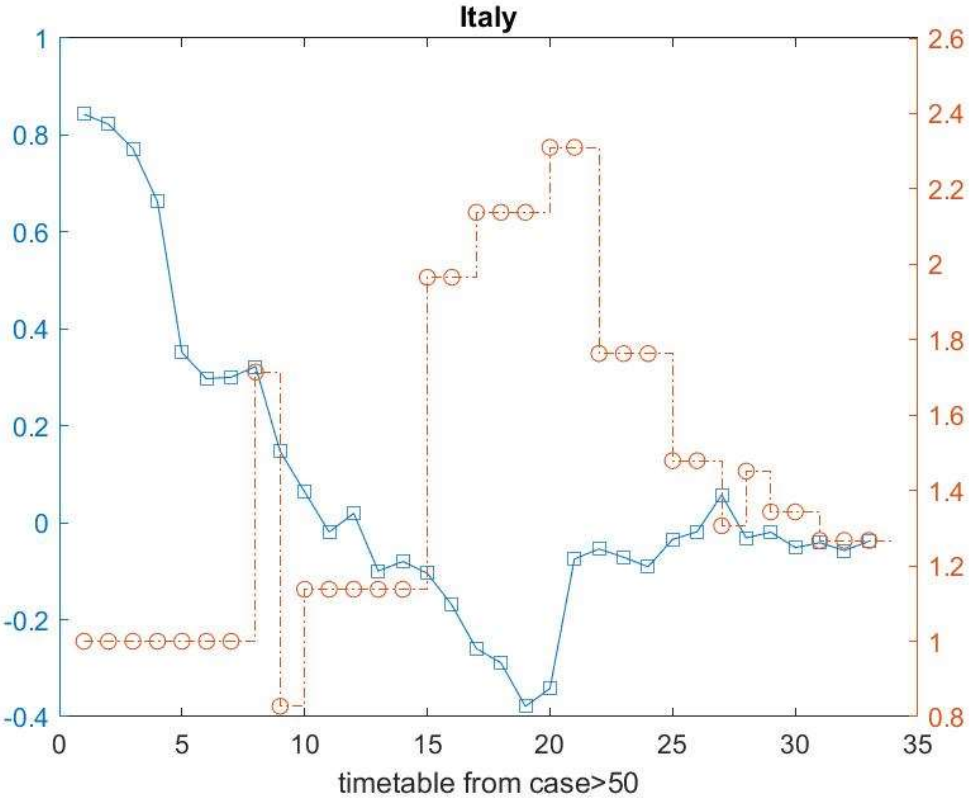
Harti, Wälde and Weber (2020) used structural break tests for nonstationary time series to estimate the treatment effect of mitigation policy in Germany. Since structural breaks may arise for reasons unrelated to mitigation policy, we do not apply their method here. Indeed, Figures 1 – 3 suggest that structural breaks are the norm even in countries that did not mitigate, or mitigated later. Chen and Qiu (2020) extend the SIR model so that the rate of contagion depends on mitigation policies such as school closure and lockdown under the assumption these policies are exogenous.

Triple Differences-in-Differences

We use a triple differences-in-differences (3DID) strategy in which the first two differences refer to the changes in R before and after the change in mitigation policy, and the third difference consists of a comparison with R in a country, which did not

mitigate. For example, Italy adopted a stringent mitigation policy after which R decreased. But so did R decrease in Sweden despite the fact that it refrained from mitigation. If, however, the decrease in R in Italy was greater than in Sweden, the difference may be attributed to the fact that Italy mitigated whereas Sweden did not.

Figure 8 Event Analysis: Italy v Sweden



In Figure 8 the left-hand vertical axis measures the logarithm of the ratio of R in Italy to R in Sweden. The right-hand vertical measures the ratio between stringency in Italy and Sweden. The horizontal axis measures corona time in days as of April 9. Initially R was much higher in Italy than in Sweden (blue). However, by day 19 it was considerably smaller in Italy, and towards the end R in Italy and Sweden were similar. On day 8 Italy began to mitigate, which was intensified on day 15 (brown schedule). On day 22 Sweden began to mitigate but by less than in Italy, hence the decrease in the brown schedule. Did the decrease in relative R result from the increase in relative S, and did the stabilization in relative R result from the subsequent decrease in relative S? Sweden's partial *volte face* over its non-mitigation policy most probably came too late to affect relative R (blue).

Since relative S increased by about 0.9 and the log of relative R decreased by about 0.5, the semi-elasticity of relative R to relative S was about -0.56 for Italy. The semi-elasticity model implies:

$$lnR_A = lnR_B + lnR_A^* - lnR_B^* + \gamma\left[\left(\frac{S_A}{S_A^*}\right) - \left(\frac{S_B}{S_B^*}\right)\right] \qquad (18)$$

where subscripts A and B refer to after and before treatment, * refers here to Sweden, and γ denotes the semi-elasticity. Semi-elasticities relative to Sweden for other countries are reported in Table 2. The absence of countries such as the UK is because mitigation policy came too late to calculate γ. However, in most countries such as Netherlands γ appeared to be zero. The results reported in Table 2 are obviously provisional. With the passage of time it may be possible to affirm them. However, because countries are increasingly adopting similar treatments in terms of more stringent mitigation policies, Sweden included, it naturally becomes more difficult to estimate treatment effects in this way.

Table 2 3DID Estimates of Treatment Effects for Mitigation Policy

Country	Semi-elasticity (γ)
Austria	-1.20
Denmark	-2.00
France	-0.45
Germany	-1.60
Israel	-0.78
Italy	-0.56
Norway	-1.8
Spain	-1.00
Switzerland	-0.43

Data as of April 9.

Equation (18) and the results in Table 2 imply that when Israel ramped-up its mitigation policy in mid-March, R subsequently decreased by 23 percent (approximately 0.64).

Vector Autoregressions

Corona equilibrium theory implies that contagion and policy mitigation are dynamically interdependent. If R increases S tends to increase subsequently, which in turn may reduce R in the future. Vector autoregression (VAR) models capture this dynamic interaction, and their impulse response functions shed light on how autonomous shocks to mitigation policy influence, or Granger-cause R. VAR models also shed light on how autonomous shocks to R Granger-cause mitigation policy. However, it is the former, which concerns us here.

The restricted VAR model for Israel is:

$$R_t = \underset{2.56}{0.410} + \underset{20.44}{0.998}\,R_{t-1} - \underset{-2.34}{0.0047}\,S_{t-2} + \underset{1.47}{0.0127}\,\Delta S_{t-3} + r_t \qquad (19a)$$

Observations: March 11 – April 18 R^2(adj) = 0.932 LM = 0.965 se = 0.196

$$S_t = \underset{4.49}{11.654} + \underset{32.03}{0.883}\,S_{t-1} + \underset{3.14}{0.384}\,\Delta S_{t-2} - \underset{-1.77}{4.659}\,\Delta R_{t-1} + \underset{1.44}{2.032}\,\Delta_3 R_t + s_t \qquad (19b)$$

R^2(adj) = 0.973 LM = 6.29 se = 2.482

where t statistics are reported below their respective parameter estimates, r and s denote innovations, LM denotes the lagrange multiplier statistic for 4th order autocorrelated residuals (not significant), and se denotes the equation standard error. R and S Granger-cause each other. As expected, S Granger-causes R negatively, but so does ΔR Granger-cause S negatively. Also, equation (19a) contains a unit root and equation (19b) contains a near-unit root. After 10 days a temporary impulse equivalent to ΔS = 30 induces a cumulative impulse response in R of -0.15, which continues to intensify through the unit root in R.

Note that Granger-causality is equivalent to causality when S_{t-2} in equation (19a) and R_{t-1} in equation (19b) are weakly exogenous (Greene 2012 p 357), necessary conditions for which are that the VAR innovations, r and s, are serially independent and independent of each other.

Chinese Crystal Ball

Whereas Sweden served as the comparator in the 3DID estimates reported in Table 2, here we use China as a comparator. Whereas Sweden undertook minimal mitigation until recently, the opposite was true in China. Moreover, Chinese exposure to corona is much longer than Sweden's. As noted in Figure 3, in March R converged to zero in all Chinese provinces. If a country adopted Chinese style mitigation, it might be plausible to assume that the temporal diffusion profile of R might be similar to China's. If so, what happened in China to R might serve as a crystal ball, or lead indicator, for countries elsewhere. Since in most other countries mitigation policy was less stringent than in China, the crystal ball needs to take account of mitigation policy.

To motivate the crystal ball model, we propose the following simple hypothesis in which relative contagion varies inversely with relative stringency as in equation (18). Let subscripts i, c, and 0 refer respectively to country i, China and a baseline country, which did not undertake mitigation. Hence, R in China relative to R in the baseline country is assumed to be:

$$\frac{R_c}{R_0} = \alpha_c - \beta_c \frac{S_c}{S_0} \qquad (20a)$$

Normalizing $S_0 = 1$, equation (20a) states that if China had not practiced mitigation, R in China would have been larger than in the baseline country if $\alpha_c - \beta_c$ exceeds 1. Importantly, the decrease in R in China has two components. The first results from maximal stringency. The second results from the decrease in R in the baseline. Similarly, R in country i relative to R in China is assumed to be:

$$\frac{R_i}{R_c} = \alpha_i - \beta_i \frac{S_i}{S_c} \qquad (20b)$$

If $\alpha_i - \beta_i$ exceeds 1, R in country i would exceed R in China if it undertook Chinese mitigation policy. If it did not, S_i/S_c is a fraction and R in country i may be larger or smaller than in China.

Equations (20a) and (20b) imply the following relation between R in country i and the baseline country:

$$R_i = R_0 \left(\alpha_i \alpha_c - \alpha_i \beta_c \frac{S_c}{S_0} - \alpha_c \beta_i \frac{S_i}{S_c} + \beta_i \beta_c \frac{S_i}{S_0} \right) \qquad (20c)$$

According to equation (20c), R in country i varies proportionately with R in the baseline country. If the term in brackets exceeds 1, R in country i is larger than in the baseline country. If mitigation policy in China and country i are proportionate, and α_i - β_i exceeds 1, R in country i varies inversely with stringency in China and country i. Matters are more complicated when mitigation policy is not proportionate because the partial derivatives with respect to stringency in China and country i have ambiguous signs. If mitigation policy is entirely reversed in China and country i, the term in brackets equals $(\alpha_i - \beta_i)(\alpha_c - \beta_c)$. Therefore, if mitigation is entirely reversed, R in country i may be greater or smaller than in the baseline country.

A Crystal Ball for Israel

Representing country i by Israel, the crystal ball hypothesis predicts that R in Israel varies directly in corona time with R in China, and it varies inversely with the stringency of mitigation policy in Israel. In principle, what matters is relative stringency as in the triple difference method, but since stringency in China was maximal (until recently), its effect is absorbed into β_i. Note that the crystal ball hypothesis does not imply that had Israel mitigated maximally as in China, R in Israel would have been the same as R in China. It means, instead, that R in China serves as a lead indicator for R in Israel. Moreover, the relation between R in Israel and R in China is not instantaneous. Instead, it is dynamically related.

We report the following hybrid result for Hubei and Zhejiang provinces, where t refers to corona time in Israel, Hubei and Zhejiang. Note that in terms of calendar time Hubei is 53 days ahead of Israel and 10 days ahead of Zhejiang.

$$\ln R_t = \underbrace{0.338}_{3.02} \ln R_{t-1} + \underbrace{0.626}_{4.85} DlnR_{t-1} + \underbrace{0.443}_{3.69} \ln(1 + Z_{t-1}) - \underbrace{0.089}_{-2.03} Dln(1 + Z_{t-1}) + \underbrace{1.128}_{2.23} \ln(1 + H_{t-1}) - \underbrace{0.528}_{-4.37} \mathrm{Dln}(1 + H_{t-1}) - \underbrace{0.0018}_{-3.30} S_{t-2}H_{t-1} - \underbrace{0.0079}_{-7.10} S_{t-2} \qquad (21)$$

Observations: March 11 – April 18 R(adj) = 0.9821 LM = 5.16 se = 0.073

Where H and Z denote R in Hubei and Zhejiang, and D is a dummy variable, which is 1 until R peaked in Israel and is zero afterwards. Equation (21) implies that the weights on R in Hubei and Zhejiang were larger after R peaked in Israel, and there

was a unit root until R until it peaked. It also implies that mitigation policy in Israel reduces R but this effect varies inversely with R in Hubei. Equation (21) implies, for example, that after R peaked the long run the elasticity of R in Israel with respect to its counterpart in Zhejiang is $0.669\frac{Z}{1+Z}$, which tends to zero as Z tends to zero. The semi-elasticity with respect to stringency is $-(0.0119 + 002H)$, hence the mitigation effect decreases when R in Hubei decreases.

In summary, lead information on R in Hubei and Zhejiang predicts R in Israel, which varies inversely with mitigation policy in Israel relative to mitigation policy in China (which remained constant during the estimation period). The equation implies that with the passage of corona time, R in Israel follows the diffusion profile for R in Hubei, but the level of R in Israel varies inversely with stringency in Israel with a lag. Figure 9A plots in-sample tracking and residuals. Because the time series in equation (21) are nonstationary but the residuals are stationary, the parameter estimates are super-consistent (Greene 2012, p 1005). Consequently, the causal effect of mitigation policy in Israel on R is identified.

Figure 9A Tracking Performance of Crystal Ball Model

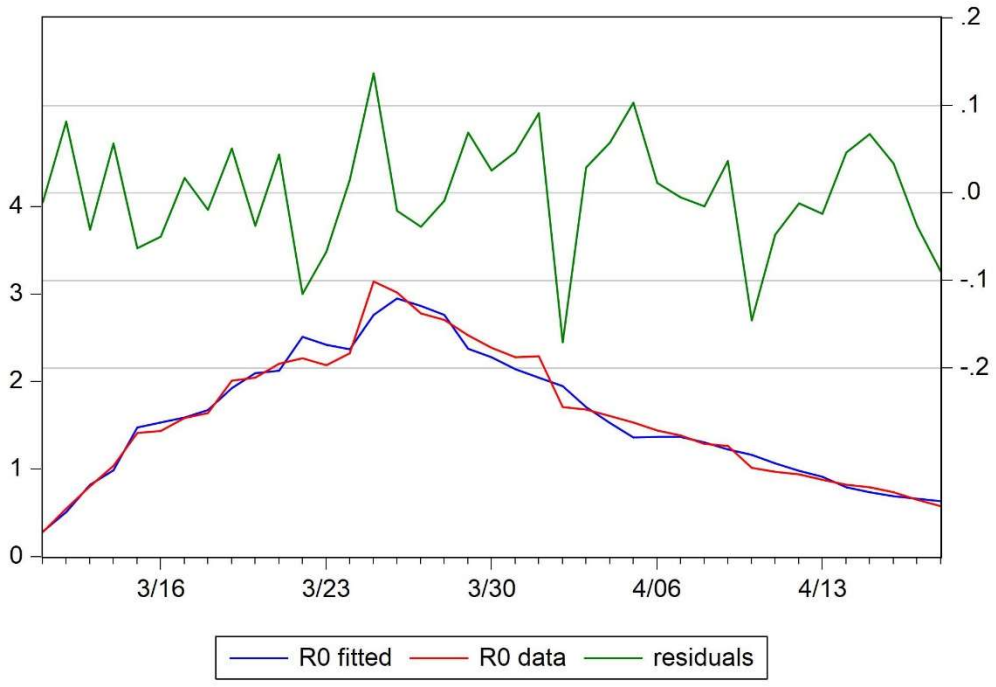
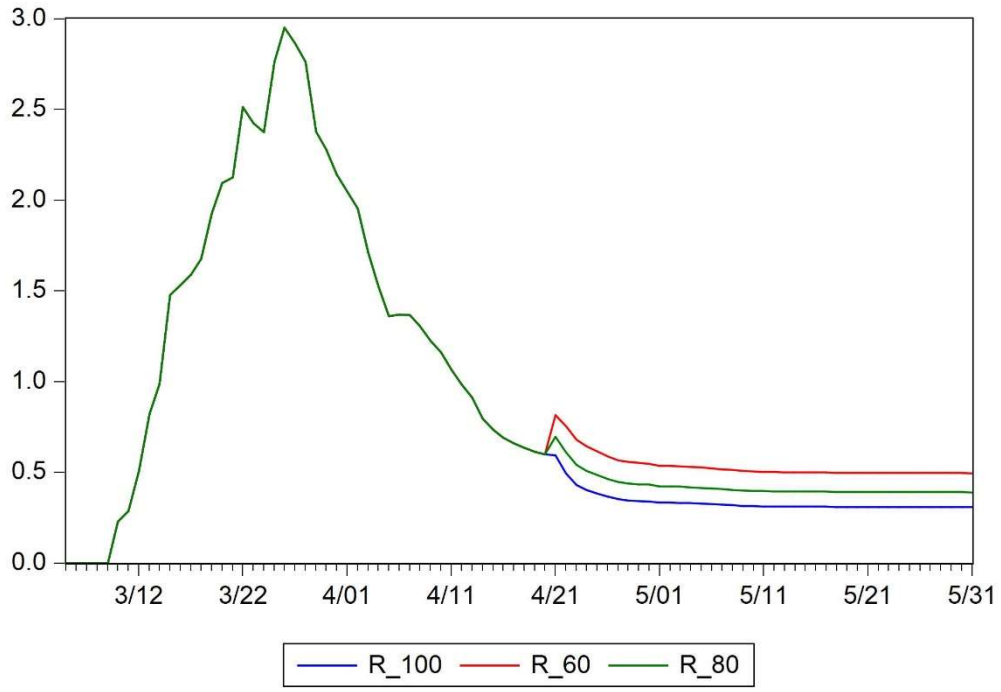


Figure 9B projects R in Israel using equation (21) under various assumptions for mitigation policy. The blue projection is based on the current policy of full mitigation (S = 100), according to which R is projected to decrease further to 0.4. Reversing mitigation policy increases R within the comfort zone of less than 1. The crystal ball model implies that R tends to a small number that exceeds zero.

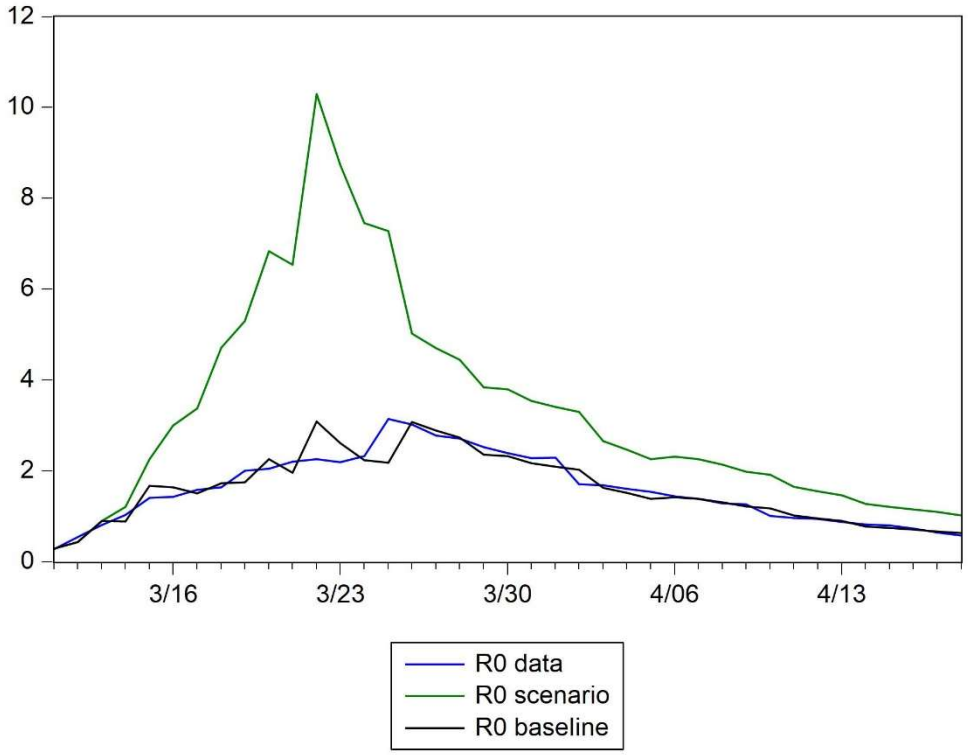
Figure 9B Crystal Ball Projections for R in Israel



The simulations in Figure 9B make the strong assumption that mitigation policy is symmetric; reversing mitigation policy has the same absolute effect on contagion as ramping it up. If mitigation policy increased public awareness of the importance of social distancing, its reversal is unlikely to decrease this awareness. Consequently, mitigation policy might be expected to be characterized by hysteresis. In its extreme form, reversal would have no effect on contagion. More generally, reversal is expected to have a smaller effect on contagion than ramping-up.

Ideally, it would be desirable to study empirically what happens to contagion when mitigation policy is reversed as in China. In the meanwhile, Figure 3B hints that mitigation policy may be asymmetric; it increases R by less when it is reversed than it decreases it when it was ramped-up. If this is true, the adverse effect on R of mitigation reversal in Figure 9B would be overstated.

Figure 9C No Lock Down since mid March



The issue of asymmetric response does not arise in counterfactual simulations, where we ask what would have happened to contagion if the government in Israel had not ramped-up its mitigation policy on March 11. Figure 9C is a counterfactual simulation in which the increase in mitigation stringency from 38 to 100 did not occur. In the absence of lockdown equation (21) implies that R would have increased sharply. This increase is induced by the unit root in equation (21) because the ramping-up occurred before R peaked. Subsequently, however, R decreases such that by the end of the simulation it is less than 2. The simulation shows that mitigation policy "flattened the curve" but it eventually made little difference to R. Lockdown simply brought forward a reduction in R that would have happened sooner or later. However, it avoided a possibly alarming increase in R. Unlike Israel most countries ramped-up mitigation policy after R had already peaked. In this case the crystal ball model would have generated a much milder increase R as in Figure 9B.

The empirical examples for Israel are intended to be illustrative. They may be applied elsewhere. For example, the Chinese crystal ball model may be applied in the United Kingdom or any other country to project contagion and to estimate the effect of mitigation policy and its reversal on the projected course for contagion.

Conclusion

In this paper we have been solely concerned with the measurement of contagion for Covid-19, the treatment effects of mitigation policy on contagion, and the projection and simulation of contagion. We have been concerned with morbidity to the exclusion of mortality. We do not consider here the crucial transitions from morbidity to intensive care and to mortality, which vary widely by country. These transitions have been disastrous for some but not all countries.

Whereas morbidity is essentially epidemiological, mortality is largely organizational and institutional. If there is inadequate provision of intensive care and ventilators, patients who would otherwise be curable die for lack of treatment. The corona epidemic is a drama, which turns into a crisis when there are insufficient or inefficient health services. This fundamental difference is overlooked by epidemiologists. For example, Ferguson et al (2020) treat ICU and ventilator capacity as a hard and fixed constraint rather than an economic phenomenon, which is subject to market forces. Ventilator suppliers have responded to the increase in their price, as a result of which ventilator capacity is expanding. Epidemiological models, such as SIR, are too mechanical and ignore learning-by-doing. They assume that if there is a second wave of Covid-19 we shall have learnt nothing from the first wave to mitigate it. There is little in the way of endogeneity in epidemiological models.

Using Chinese data for January and February, Ferguson et al jumped to the conclusion that eventually 90 percent of the world's population will be infected and hundreds of millions may die. Toda (2020) reached similar conclusions regarding the contagiousness of Covid-19 based on SIR models estimated using cross country panel data for Covid-19 morbidity. It is ironical that by March 16 (when the Ferguson report was published) the rate of contagion had dropped to zero in almost all Chinese provinces, and was heading for zero even in countries that undertook almost no mitigation. It remains to be seen whether their predictions of further even more deadly rounds of Covid-19 are realistic.

References

Bailey N. (1975) *The Mathematical Theory of Infectious Diseases and its Application.* 2nd edition, Griffin, London.

Bendavid E. et al (2020) Covid-19 antibody sero-prevalence in Santa Clara County, California. *mEDRxiv*, https://doi.org/10.101/2020.04.14.20062463.

Berck P. and Villas-Boas S.B. (2016) A note on the triple difference in economic models. *Applied Economic Letters*, 23: 239-242.

Berg B.A. (2004) *Markov Chain Monte Carlo Simulation and their Statistical Analysis*. World Scientific Press, Singapore.

Chen X. and Qiu Z. (2020) Scenario analysis for non-pharmaceutical interventions on global Covid-19 transmissions. *Covid Economics*, 7: 46-67.

Chudik A., Pesaran M.H., and Rebucci A. (2020) Voluntary and mandatory social distancing: evidence on Covid-19 exposure rates for Chinese provinces and selected countries. Federal Reserve Bank of Dallas, https://doi.org/10.24149gwp382.

Ferguson N. et al (2020) Impact of non-pharmaceutical interventions to reduce Covid-19 mortality and health care demand. doi:https://doi.org/10.25561/77482, London (March 16, 2020).

Greene W, (2012) *Econometric Analysis*, 7th edition, Prentice Hall.

Hamra G., MacLehose R. and Richardson D. (2013) Markov chain Monte Carlo: an introduction for epidemiologists. *International Journal of Epidemiology*, 42: 627-634.

Harti T., Wälde K. and Weber E. (2020) Measuring the impact of the German public shutdown on the spread of Covid-19. *Covid Economics*, 1: 25-32.

Liu Y., Gayle A.A., Wilder-Smith A. and Rocklöv J. (2020) The reproductive number of Covid-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*, doi:10.1093/jtm/taa.ao21.

Tian et al (2020) The impact of transmission control measures during the first 50 days of the Covid-19 epidemic in China. *Science*, https://doi.org/10.1126/science.abb6105.

Toda A.A. (2020) Susceptible -infective – recovered (SIR) dynamics of Covid-19 and economic impact. *Covid Economics*, 1: 43- .

Wu J.T., Leung k. and Leung G.M. (2020) Nowcasting and forecasting the potential domestic and international spread of the 2019-nCov outbreak originating in Wuhan, China: a modelling study. *Lancet*, 395: 689-697.

Zhou F. et al (2020) Clinical course and risk factors for mortality of adult inpatients with Covid-19 in Wuhan, China: a retrospective cohort study. *Lancet*, 395: 1054-1062.