COVID19: Erroneous Modelling and Its Policy Implications*

Yinon Bar-On, Weizmann Institute of Science †

Tatiana Baron, Ben Gurion University ‡ **Ofer Cornfeld, BFI**§
Ron Milo, Weizmann Institute of Science†

Eran Yashiv, Tel Aviv University and CEPR ¶

March 11, 2021

Abstract

Research in Economics on COVID19 posits an economy subject to disease dynamics, which are often seriously mis-specified in terms of speed and scale. Using a social planner problem, we show that such misspecifications lead to misguided policy. Erroneously characterizing a relatively slow-moving disease engenders dramatically higher death tolls and excessive output loss relative to the correct benchmark. We delineate the latter, employing epidemiological evidence on the timescales of COVID19 transmission and clinical progression. The resulting sound model is simple, transparent, and novel in Economics.

Key words: optimal policy, public health, GDP loss, COVID19, disease dynamics and scale, misspecification.
JEL No. E61, E65, H12,I12,I18, J17.

We thank Philipp Kircher, Marc Lipsitch, and Ben Moll and for useful exchanges and seminar participants at UCL and LSE for comments. Any errors are our own.

†Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot, Israel.

‡Department of Economics, Ben Gurion University, Beer Sheva, Israel.

§BFI, Tel Aviv, Israel.

¶*Corresponding author.* The Eitan Berglas School of Economics, Tel Aviv University, Israel. E-mail: yashiv@tauex.tau.ac.il; phone:+972-3-640-9233.

1 Introduction

Since March 2020 there has been a rapidly expanding research effort dedicated to COVID19 analysis across disciplines, inter alia, in Economics. A typical analysis posits an economy, which is subject to a model of COVID19 dynamics. One type of economic analysis describes a planner problem that seeks to derive optimal policy. The latter trades off the costs of public health outcomes, such as breach of ICU capacity and death, with the economic costs of suppression policy, including declines in production. Other papers model the decentralized economy and the optimal decisions of agents, emphasizing individual epidemic-related behavior, as well as externalities. In both cases the dynamics of the disease and its features are at the core of the analysis.

This paper makes two contributions: one is to show that there is often serious misspecification of the model of the epidemic in Economics, with errors in the set-up and in the parameterization, at odds with the epidemiological evidence. The underlying cause for the misspecification is the failure to make the distinction between two dynamic aspects of the disease: the epidemiological, in particular, the timing and duration of the infectiousness period; and the clinical, namely durations till symptoms onset, hospitalization, and death. Due to erroneous modelling, wrong values are assigned to key parameters of disease dynamics, while important parameters are omitted. As a consequence there is misleading characterization of a relatively slow-moving disease.

Why should economists care about a model of disease dynamics? This is important for the attempt to derive an optimal intervention plan. Getting the speed of the disease wrong implies getting the timing of interventions wrong, in a situation where timing is crucial. For example, in NYC, imposing lockdown a week too late resulted in a five-fold increase of the epidemic, with over 750,000 people infected on the last week of March 2020. Issuing a stay-at-home order on March 15th, or earlier, could have saved thousands of lives.[1] We place the issue of policy timing at the center of the analysis. We use an optimal social planner framework to analyze costs in terms of death tolls and output loss. These are shown to depend on the timing of interventions and are significantly affected by the erroneous modelling of disease dynamics.

The second contribution is to place the modelling of COVID 19 in Economics on the foundations of an epidemiological analysis of the SARS-CoV-2 properties, particularly, its transmission timescales. The main elements of the ensuing model are two blocks: an infection transmission block, where the number of new cases is determined, and a clinical block, which characterizes the dynam-

[1] Stadlbauer et al (2021), report that on March 29, 2020, the cumulative infection rate was 2.2% and a week later, it was 10.1%. Seroprevalence tests require two weeks to detect the disease. This, then, implies that if a stay-at-home order were issued a week before March 22, the actual time of the policy decision, disease spread could have been significantly suppressed. The death toll would have been reduced by 6,000 lives, based on the assumption of an Infection Fatality Rate (IFR) of 0.8%.

ics within the health system – development of symptoms, hospitalization, ICU admission, and recovery or death. We present models of these two different dynamics, including the relevant parameterizations, based on epidemiological and clinical evidence. The distinction between the two dynamic models is novel in the Economics literature and accurate from the epidemiological perspective. It results in a simple and transparent structure.

We proceed as follows. We start by specifying a model of COVID19, which is sound in terms of epidemiology and contrast its dynamic properties with the specifications prevalent in the Economics literature. We then show that a widely-used specification is unable to reproduce the dynamics of the disease observed in the data, while the correct model can replicate it well. Subsequently, we discuss the dynamics of the disease implied by each model. Finally, we illustrate the implications of erroneous disease modelling for optimal policy. To this end, we embed the correct model, as well as the alternative, prevalent specifications in a simple planner problem and derive optimal lockdown timing. We find dramatic consequences of the specification errors here: they distort the policymaker decisions, under-estimate the scale of the epidemic, and result in higher death tolls and excessive output losses. The magnitudes are big; death tolls deteriorate by hundreds of thousands of people and output losses increase by 20%. We note that the specification errors in question bear no relation to assumptions about individual behavioral responses. They lie in a different dimension, that of understanding fundamental disease properties and distinguishing between its epidemiological and clinical aspects.

The analysis points economic researchers at the correct way to model the dynamics of the disease and will be useful for other epidemics beyond COVID19, as the discussion is pertinent to other infectious diseases. Note, in this context, that the set of epidemics since 1980 is quite large and includes, inter alia, HIV/AIDS, SARS, H5N1, Ebola, H7N9, H1N1, Dengue fever, and Zika. We see the analysis here as complementary to work on the importance of the correct modelling of population heterogeneity, such as Ellison (2020).

The structure of the paper is the following: Section 2 very briefly presents the parts of the epidemiological and economics literatures relevant for the current discussion. Section 3 discusses the epidemiological models, both the preferred model (including the appropriate parameterization), and widely-used models, which are the subject of the current critique. Section 4 examines the empirical fit of these models. Section 5 discusses the epidemic dynamics implied by each model. Section 6 analyzes the economic policy implications of using the different models, and the costs involved when employing erroneous ones. Section 7 concludes.

2 Literature

This paper relates to two literatures.

One is the part of the Epidemiology literature, which has modelled epidemic dynamics using a compartmental approach. The approach was pio-

neered by the seminal work of Kermack and McKendrick (1927). The ensuing family of models identifies epidemiological states and considers the flow rates between compartments containing individuals in each disease state. In this paper we explore three variants of this model. For reviews of this approach and its underlying rationale, see Champredon, Dushoff, and Earn (2018). Within this strand, pandemic or epidemic management policy is a key topic of study. Prominent examples of such studies, pre-COVID 19, include Mills, Robins, and Lipsitch (2004) and Wallinga, van Boven, and Lipsitch (2010).

The other is the Economics literature on COVID19. Avery et al (2020) discuss its connections with the afore-cited Epidemiology literature. Many papers have been making use of epidemiological models and are thus subject to the current analysis. These include models in Macroeconomics, International Economics, Public Economics, and Labor Economics. Within this burgeoning literature, we briefly mention key papers which have examined optimal lockdown policy using the concept of a social planner. They study the health-related losses due to the pandemic and the economic consequences of public health policy. In this framework an objective function is defined, with values taking into account economic losses and the value of statistical life. Thus, tradeoffs are measured and alternative policies can be evaluated. The planner constraints include, inter alia, the disease dynamics typically examined within the SIR epidemiological model. Prominent contributions include Abel and Panageas (2020), Acemoglu, Chernozhukov, Werning, and Whinston (2020), Alvarez, Argente, and Lippi (2020), and Jones, Philippon, and Venkateswaran (2020). Depending on the exact formulation, we show below how erroneous use might lead to work with misspecified models, with substantial consequences for policy. Two key properties of disease dynamics, its scale and speed, are at the center of misspecification.

3 Modelling the Epidemic

This section presents three models of the epidemic. We first present our preferred specification, which relies on up-to-date epidemiological evidence (subsection 3.1). We then present two alternative specifications, prevalent in the afore-cited Economics literature (sub-section 3.2).

3.1 The Model

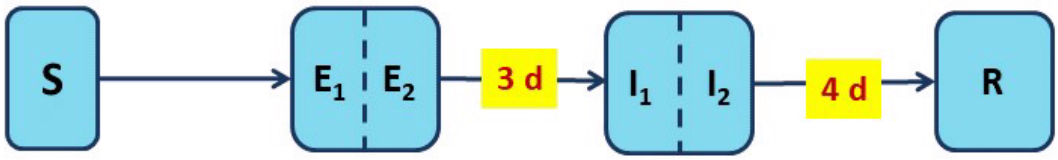
We analyze the dynamics of the epidemic within two complementary blocks – infection transmission and clinical progression. The former block is characterized by the SEIR-Erlang model and reflects the epidemiological properties of COVID19. The clinical block characterizes the development of symptoms, hospitalization, ICU admission, and recovery or death, and is needed to describe the dynamics in the public health system. This structure is novel and proves extremely useful, as in COVID19 the infectiousness period and the clinical progression of the disease follow different timescales, demonstrated below. Separating them makes the model rich, yet simple and transparent, ensuring

that each parameter has a clear empirical counterpart in the data. Moreover, the shortcomings of prevalent specifications, that we present in sub-section 3.2, are rooted in the inability of these models to capture both infection transmission and clinical properties of COVID19 using a single set of parameters.

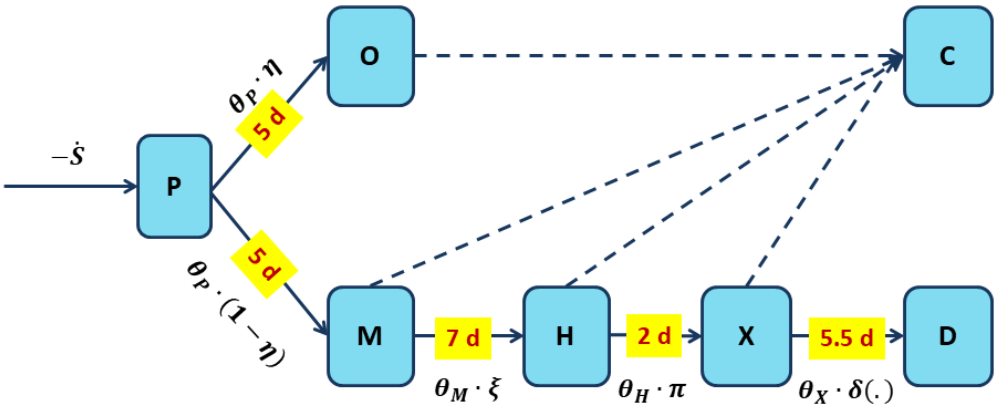
3.1.1 The SEIR-Erlang Block

Before contacting the disease for the first time, a person is Susceptible (S). Once a person gets infected, disease progression is split into distinct compartments – Exposed (E), Infectious (I), and Resolved (R). We denote by $\beta(t)$ the infections transmission rate, σ, the transition rate from E to I, and γ, the transition rate from I to R. An infected individual spends some time in each compartment before moving on to the next one. The person is infectious only when in the I compartment, but not when residing in the preceding E compartment. The time durations spent in the E and I compartments are known as the latent and infectious periods, respectively. Once people move to the Resolved stage, they no longer participate in disease transmission. Graphically, this model is presented in panel a of Figure 1. We provide references below to the durations noted in the figure.

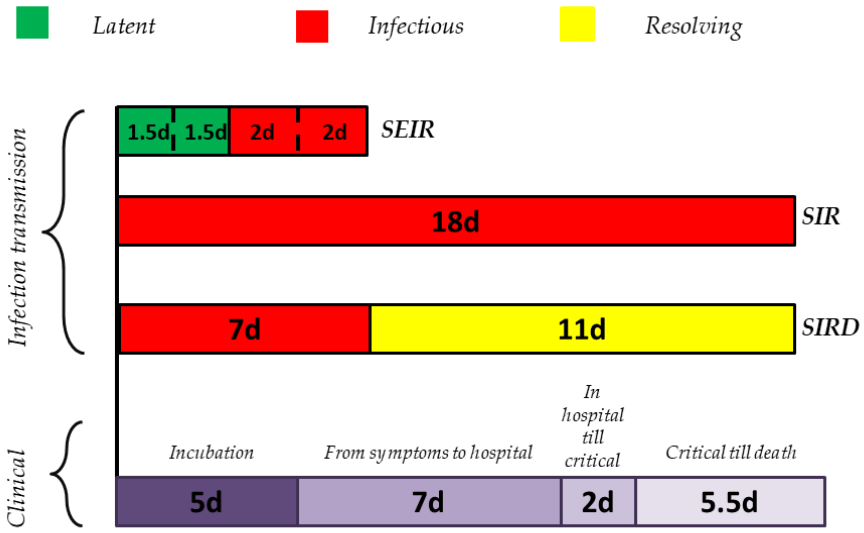
Figure 1: The Model



a. The Infection Transmission Block – $SEIR$



b. The Clinical Block



c. Timescales of the Three Models

Note that with Poisson transition rates between compartments, the residence times in each of them are distributed exponentially, and thus have zero mode. Exponential distributions capture the mean but not the mode of the biologically accurate distributions of residence times, because in reality what most people spend in each stage is close to the mean of the distribution, rather than zero. Therefore, we split the E and I compartments into two sub-compartments and double the rate of transition. Now, the latent and infectious periods are the sum of the time spent in the E_1 and E_2 or I_1 and I_2 sub-compartments, respectively. Their distribution is the sum of exponentially distributed random variables, a special case of the Gamma distribution, known as the Erlang distribution. The means of Erlang distributions remain $1/\sigma$ and $1/\gamma$, but the modes are now near the means, as they should be. In the remainder of the paper we shall refer to this model as the $SEIR$ model, without noting the number of sub-compartments.

The following equations describe this block. Throughout, all stock variables are expressed as a fraction of the population.

$$\begin{aligned}
\dot{S}(t) &= -\beta(t) \cdot (I_1(t) + I_2(t)) \cdot S(t) && (1)\\
\dot{E}_1(t) &= \beta(t) \cdot (I_1(t) + I_2(t)) \cdot S(t) - 2\sigma E_1(t) && (2)\\
\dot{E}_2(t) &= 2\sigma E_1(t) - 2\sigma E_2(t) && (3)\\
\dot{I}_1(t) &= 2\sigma E_2(t) - 2\gamma I_1(t) && (4)\\
\dot{I}_2(t) &= 2\gamma I_1(t) - 2\gamma I_2(t) && (5)\\
\dot{R}(t) &= 2\gamma I_2(t) && (6)
\end{aligned}$$

An important parameter is the reproduction number $\mathcal{R}_t$, which is the average number of people infected by a person, and is given by:

$$\mathcal{R}_t = \frac{\beta(t)}{\gamma} \tag{7}$$

We use $\mathcal{R}_t$ for the reproduction number at date t and denote the basic reproduction number by $\mathcal{R}_0$ at the initial stage, when $S(0) = 1$. We shall also be discussing the effective reproduction number, defined as:

$$\mathcal{R}_e = S(t)\mathcal{R}_t \tag{8}$$

The path assumed for the reproduction parameter merits discussion. We model a time-varying parameter, $\mathcal{R}_t$, reflecting both rational individual behavior and the effects of suppression policy. We take into account that individuals adjust to the new environment and behave differently, both with and without policy interventions. In particular, as the epidemic unfolds, people become increasingly aware of the risks and adjust their behavior. This adjustment is manifested in avoiding or reducing social contact and taking precautions, such as wearing masks. These changes happen in part as a direct result of government Non-Pharmaceutical Interventions (NPIs) and in part as a voluntary response. It is a rational choice to adopt new norms of behavior, even when restrictions

by the government are weakened or removed. As a result, the speed of disease transmission declines relative to its value at the start of the epidemic. Specifically, when we discuss policy interventions below, the rate $\mathcal{R}_t$ will depend on the regime – either lockdown, to be denoted $\mathcal{R}_L$, or out of lockdown, work, to be denoted $\mathcal{R}_W$. In Section 6 we elaborate on the empirically-based parameterization of $\mathcal{R}_0, \mathcal{R}_L$ and $\mathcal{R}_W$.

3.1.2 The Clinical Block

The clinical block describes the progression of the infected through the healthcare system, depending on the development and severity of symptoms.

We postulate the following. Once infected, a person enters an incubation period, a P state, during which there are no symptoms. This period lasts for $1/\theta_P$ on average.[2] Following the incubation period, a person either remains asymptomatic (O) or develops symptoms (M). Denote the share of asymptomatic cases by η. The others $(1-\eta)$ develop symptoms, and with probability ξ are hospitalized (H). A given share π of patients become critically ill, that is develop conditions requiring transition to ICU (denoted X). Once critically ill, a fraction $\delta(\cdot)$ dies. We specify the death probability in this critical state X as:

$$\delta(X(t), \overline{X}) = \delta_1 + \delta_2 \cdot \frac{\mathbf{I}(X(t) > \overline{X}) \cdot \left(X(t) - \overline{X}\right)}{X(t)} \tag{9}$$

where $\overline{X}$ denotes ICU capacity and $\mathbf{I}$ is the indicator function. At any stage, a person may recover (C). The clinical block is represented graphically in panel b of Figure 1.

The analytical description of the symptomatic branch is:

$$\begin{aligned}
\dot{P}(t) &= \beta(t) \cdot (I_1(t) + I_2(t)) \cdot S(t) - \theta_P \cdot P(t) && (10)\\
\dot{M}(t) &= (1-\eta) \cdot \theta_P \cdot P(t) - \theta_M \cdot M(t) && (11)\\
\dot{H}(t) &= \xi \cdot \theta_M \cdot M(t) - \theta_H \cdot H(t) && (12)\\
\dot{X}(t) &= \pi \cdot \theta_H \cdot H(t) - \theta_X \cdot X(t) && (13)\\
\dot{D}(t) &= \delta(X(t), \overline{X}) \cdot \theta_X \cdot X(t) && (14)
\end{aligned}$$

The parameters θ_P, θ_M, θ_H, and θ_X denote the average time that passes between the stages of infection, symptoms onset, hospitalization, ICU admission, and death, respectively.

3.1.3 The Connection to Economic Analysis

We posit that the number of people who can work daily, $N(t)$, is given by:

$$N(t) = N^{SS} \cdot \rho \cdot (1 - D(t) - X(t) - H(t) - \phi M(t)) \tag{15}$$

[2]Note that the incubation period is governed by θ_P and is different from the latent period which is governed by σ. In fact, recent epidemiological evidence indicates that the average incubation period can be twice as long as the average latent period (Bar-On et al. (2020)).

where N^{SS} is steady state employment, $0 < \rho \leq 1$ is the fraction able to work given any policy restrictions, and $0 \leq \phi \leq 1$ is the fraction of people with symptoms who do not work.

3.1.4 Parameterization

The parameterization of this model needs to be both epidemiologically- and clinically-based. In Table 1 we present the relevant values for the two blocks, where we rely on sources in the epidemiological and medical literatures published in Science, Nature, the Lancet, and JAMA, as detailed in the table's notes.

Table 1: Epidemiologically- and Clinically-based Parameterization

	Interpretation	**Range of Estimates**	**Preferred**	**Parameter value used**
a. The Infection Transmission Block (SEIR)				
σ	latent period duration	$3-5$ days	3 days	$1/3$
γ	infectious period duration	$4-5$ days	4 days	$1/4$
b. The Clinical Block				
θ_P	incubation period	$5-6$ days	5 days	$1/5$
θ_M	days from symptoms till hospitalization	7 days	7 days	$1/7$
θ_H	days in hospital till ICU	2 days	2 days	$1/2$
θ_X	days in ICU before death	5.5 days	5.5 days	$1/5.5$
η	Prob. to be asymptomatic	$20\% - 50\%$	50%	0.5
ξ	Prob. to get hospitalized when symptomatic	$\frac{\#Hospitalized}{\#Infected} = [2\% - 4\%]$	4%	$\frac{0.04}{1-0.5} = 0.08$
π	Prob. of ICU admission	$10\% - 40\%$	40%	0.4

Notes:

1. Sources – for panel a: Bar-On et al (2020); He at al (2020); Li et al (2020); Tian et al (2020);

For panel b – Bar-On et al (2020); Huang et al (2020); Richardson et al (2020); Salje et al (2020).

2. $\frac{\#Hospitalized}{\#Infected} = \frac{\#Hospitalized}{\#Symptomatic} \cdot \frac{\#Symptomatic}{\#Infected} = \xi \cdot (1-\eta) \implies \xi = \frac{\frac{\#Hospitalized}{\#Infected}}{1-\eta}$

Note that the implied Infection Fatality Rate (IFR) is 0.8%,[3] consistent with the estimates of the Imperial College COVID19 Response Team (2020). Additionally, based on Bar-On et al (2020), we set $\delta_1 = 0.5$. In the U.S. economy, ICU capacity is $\overline{X} = 1.8 \times 10^{-4}$, based on an estimate of approximately 58,100 ICU beds by the Harvard Global Health Institute.[4] Finally, we set $\delta_2 = 0.5$ to capture the fact that with extreme loads on the public health system, the probability to die increases to 1 for each patient in need of an ICU bed.

3.2 Alternative Specifications

The overwhelming majority of Economics papers on COVID19 model both clinical outcomes and infection dynamics within a single block. We proceed by presenting the SIR model and its calibration, and a modification ($SIRD$), designed to better capture the dynamic path to death. Panel c of Figure 1 provides a graphical illustration.

3.2.1 The Widely-Used SIR Model

Economists modelling the dynamics of COVID19 have been using in many cases versions of the SIR model with the following structure.

$$\dot{S}(t) = -\beta(t) \cdot I(t) \cdot S(t) \tag{16}$$
$$\dot{I}(t) = \beta(t) \cdot I(t) \cdot S(t) - \gamma I(t) \tag{17}$$
$$\dot{R}(t) = \gamma I(t) \tag{18}$$

Whenever numbers of deceased and recovered are needed the following equations are used:

$$\dot{D}(t) = \mu \dot{R}(t) \tag{19}$$
$$\dot{C}(t) = (1-\mu)\, \dot{R}(t) \tag{20}$$

where D is deceased, C is recovered and μ is the infection fatality rate.

A prevalent calibration is given by:

$$1/\gamma = 18 \Longrightarrow \gamma = 1/18$$

This is derived as follows:

a. The duration of the disease till death is taken to be 18 days.[5]

b. The latter is also taken to be the duration of the Infectious stage, and given by $1/\gamma$. The infection transmission rate $\beta(t)$ is then pinned down by a given value of $\mathcal{R}_t$ and the length of the infectious stage, $1/\gamma$.

[3] This rate is given by $IFR = \xi \cdot \pi \cdot \eta \cdot \delta_1$

[4] See https://globalepidemics.org/our-data/hospital-capacity/

[5] Our preferred value, given in Table 1 above, is 19.5 days, which is not very different from the value here.

The key point here is that the SIR model misses the clinical block. It confounds time to death (from the onset of infection) with the length of the Infectious period. We show below that this mis-specification has profound implications for the speed and scale of the disease.

3.2.2 The SIRD Model

Some modelers modify the SIR model, to take into account the fact that the time from infection until a person is no longer infectious is relatively short (7 days on average), though it takes longer till one recovers or dies. They thus replace equation (18) by:

$$\dot{R}(t) = \gamma I(t) - \theta \cdot R(t) \tag{21}$$

where θ defines the duration of the resolving stage R. Replacing equations (19)-(20), one gets:

$$\dot{D}(t) = \mu \cdot \theta \cdot R(t) \tag{22}$$

$$\dot{C}(t) = (1-\mu) \cdot \theta \cdot R(t) \tag{23}$$

This model is denoted $SIRD$ and is usually calibrated with $\gamma = 1/7$.

4 Empirical Fit

Before analyzing the properties of these models, we use data on COVID 19 deaths in NYC to examine their empirical fit, or lack thereof. We chose NYC as the epidemic there reached a very significant scale. Moreover, this is a case where both data, that we use, and empirical studies of the disease, that we cite, are readily available. At the heart of the analysis is the derivation of estimates of two values for the reproduction number $\mathcal{R}_t$ – one during the early outbreak (denoted by $\mathcal{R}_0$) and the other during lockdown (denoted by $\mathcal{R}_L$). We explain the method of derivation, present the ensuing estimates, and discuss model-data fit.

4.1 Derivation of Reproduction Parameter Estimates

At the stage of the initial exponential growth of the disease, it can be shown that there is a relation between the rate of exponential growth λ and the reproduction number $\mathcal{R}_0$.

For the case of the SIR model and $SIRD$ the following relation obtains (see equation 3.1 in Wallinga and Lipsitch (2007)):

$$\mathcal{R}_0 = 1 + \frac{\lambda}{\gamma} \tag{24}$$

For the $SEIR$ model with m, n sub-compartments the relation is given by (see equation 4 in Wearing et al (2005)):

$$\mathcal{R}_0 = \frac{\lambda\left(\frac{\lambda}{\sigma m}+1\right)^m}{\gamma\left(1-\left(\frac{\lambda}{\gamma n}+1\right)^{-n}\right)} \tag{25}$$

After the initial $\mathcal{R}_0$, we use the effective reproduction number $\mathcal{R}_e$, as discussed in sub-section 3.1.1.

For our empirical exercise, we use data on daily deaths in NYC.[6] It covers 236 days, from the first death in mid March 2020 to the end of October 2020, and is the sum of confirmed and probable deaths from COVID19. The death count is smoothed using a 7-day centered moving average.

The following data-fitting exercise is undertaken. First we estimate the disease exponential growth or decline rate, λ, based on the periods of initial eruption and subsequent contraction using the smoothed death series. For both these periods, we run a Poisson (log-linear) regression to estimate λ:

$$\log\left(daily\ death\ count\right) = const + \lambda t \tag{26}$$

Second, based on the estimated λ values we derive the corresponding values of $\mathcal{R}_0$ and $\mathcal{R}_L$ in each model using equations (24)-(25) for the initial outbreak and lockdown periods, respectively.

Third, we try to fit the entire death series by minimizing the squared deviations of the simulated series of daily deaths from the corresponding data series, refining our derived values of $\mathcal{R}_0$ and $\mathcal{R}_L$ from the first step. Within this procedure, we need to derive a value also for $\mathcal{R}_W$, one that captures disease growth after the initial (and harshest) measures have been removed. In particular, we solve the following minimization problem:

$$\min_{T_0,T_1,\mathcal{R}_0,\mathcal{R}_W,\mathcal{R}_L,\tau_0} \int_{t=0}^{t_{end}} \left(D\left(t\right) - D_{actual}\left(t-\tau_0\right)\right)^2 dt \tag{27}$$

where D_{actual} is the data death series and $D(t)$ is the death series predicted by the model; τ_0 is the time needed to adjust the death series to model dynamics, given the duration from infection to death; lockdown is imposed between T_0 and T_1, such that:

$$\mathcal{R} = \begin{cases} \mathcal{R}_0 & t \leq T_0 \\ \mathcal{R}_L & T_0 < t \leq T_1 \\ \mathcal{R}_W & T_1 < t \end{cases}$$

[6] Source is in data page
https://www1.nyc.gov/site/doh/covid/covid-19-data.page
and the data set is to be found at
https://github.com/nychealth/coronavirus-data/blob/master/archive/probable-confirmed-dod.csv

4.2 Data Estimates and Model-Data Fit

Table 2 presents the resulting estimates for the λs (the exponential growth and decline rates) and the corresponding estimated values of $\mathcal{R}_0$ and $\mathcal{R}_L$ for the second step above.

Table 2: Parameter Estimates for NYC Data

dates		**point estimates**	**95% Confidence interval**
March 14 - April 02	Exponential Growth	$\lambda = 0.20$	$[0.18, 0.22]$
	Reproduction number	$\mathcal{R}_0$	
	SEIR	2.76	$[2.53, 3.00]$
	SIRD	2.40	$[2.26, 2.54]$
	SIR	4.60	$[4.24, 4.96]$
April 12 - May 05	Exponential Decline	$\lambda = -0.059$	$[-0.060, -0.058]$
	Reproduction number	$\mathcal{R}_L$	
	SEIR	0.81	$[0.80, 0.81]$
	SIRD	0.69	$[0.68, 0.70]$
	SIR	N/A	N/A

Some of the problems with the SIR model are apparent in Table 2. Under exponential growth, one obtains the implausibly high estimate of $\mathcal{R}_0 = 4.60$, which does not conform most estimates in the literature. In the exponential decline phase, the decline is too steep to be matched by the SIR model; there is no positive $\mathcal{R}_L$ that can account for the observed exponential decline in fatalities during lockdown, so there is no $\mathcal{R}_L$ value that can be derived. In other words, the case of NYC shows that the widely-used SIR specification cannot reproduce the observed decline in fatalities.

For the $SIRD$ model, the estimates in Table 2 are in line with the Fernandez-Villaverde and Jones (2020) estimates,[7] produced from the same dataset, and which are not far from the preferred $SEIR$ model estimates.[8]

The findings of Table 2 can be summed up as follows.

The exponential growth rate λ is estimated between March 14 and April 2 to be 0.20 . Under $SEIR$, these values correspond to $\mathcal{R}_0 = 2.76$, under $SIRD$ to $\mathcal{R}_0 = 2.40$, and, under SIR, to the substantially higher and implausible value of $\mathcal{R}_0 = 4.60$.

[7] Their estimates are somewhat attenuated, as they use the HP filter to further smooth the data.

[8] As external validation of these estimates, $\mathcal{R}_t$ values computed from the site http://metrics.covid19-analysis.org/ for NY county in the state of NY, in the same period, start at around 2.50. This website and the associated $\mathcal{R}_t$ analysis was developed by Xihong Lin's Group in the Department of Biostatistics at the Harvard Chan School of Public Health.

The exponential decline rate λ is estimated between April 12 and May 5 to be -0.059. At the beginning of exponential decline, around April 12th, we postulate that 15% of the population is already infected,[9]and thus $\mathcal{R}_L = \mathcal{R}_e/0.85$. Under $SEIR$, these λ values correspond to $\mathcal{R}_L = 0.81$ and under $SIRD$ to $\mathcal{R}_L = 0.69$. As noted, such a decline is too steep to be matched by the SIR model.

Next, as explained above, looking at the entire series of fatalities in NYC from mid March to the end of October 2020, we refine our parameterization of $\mathcal{R}_0$ and $\mathcal{R}_L$, and obtain a value for $\mathcal{R}_W$.[10] Minimizing the distance between the data and the fatalities series implied by the SEIR model, now we get $\mathcal{R}_0 = 2.70$ at the outbreak, before March 23, and $\mathcal{R}_L = 0.70$ during lockdown, lasting till April 20.[11]

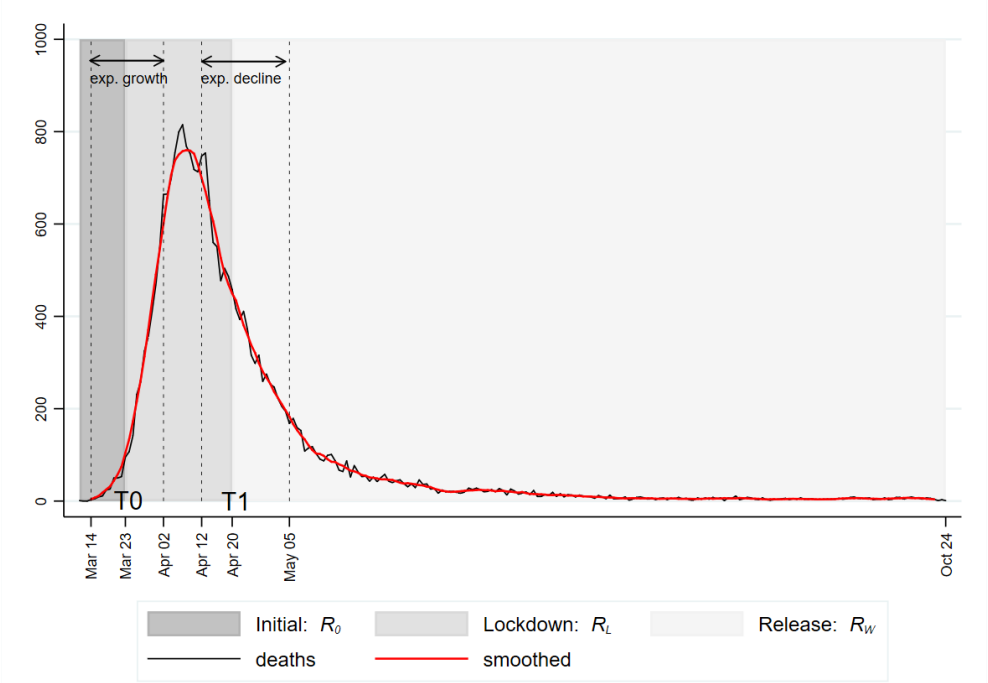
Panel a of Figure 2 plots the daily death flow (both the raw data and the smoothed series).

[9]Based on the seroprevalence tests findings of Stadlbauer et al. (2021) and seroconversion timescales reported in Kai-Wang To et al. (2020), the infection rates in NYC as of April 12, 2020 were at least 15%. Additionally, our postulated assumptions are later confirmed by the simulation results (reported below).

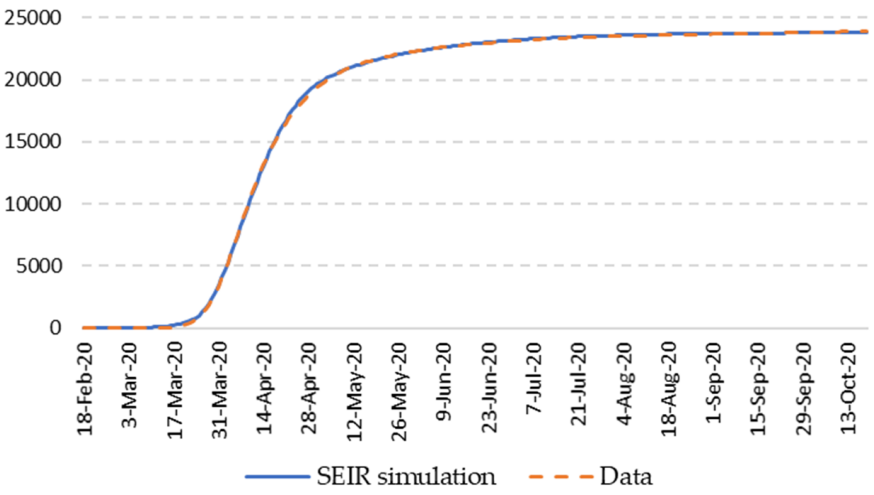
[10]We keep using the ICU capacity constraint for the U.S., as in Table 1. The same source gives a somewhat higher capacity for NYC alone, but we prefer to take the more conservative estimate. In any event, both estimates produce very similar results.

[11]Stay-at-home orders went into effect on March 22, 2020. The policy lasted until May 15, 2020

Figure 2: NYC deaths and SEIR model fit



a. NYC Daily Deaths



b. NYC Cumulative Deaths – Data and Simulated SEIR Model

In the period after lockdown, our procedure identifies a reproduction parameter ($\mathcal{R}_W$) which is lower than the initial value ($\mathcal{R}_0$) and higher than the lockdown value ($\mathcal{R}_L$), i.e., $\mathcal{R}_W = 1.10$.[12] At the release of lockdown, time T_1 (see panel a in Figure 2), around 20% of the population had been infected, so $S(T_1) = 0.80$. This value of $\mathcal{R}_W = 1.10$ generates an effective $\mathcal{R}_e = 0.80 * \mathcal{R}_W = 0.88$ leading to a gradual decline of the daily death series, as seen in the figure.

Panel b of Figure 2 shows an excellent fit of the resulting simulated series to the data series (correlation of 0.999). Thus, the correctly-specified $SEIR$ model fits NYC fatality data extremely well. Under the $SIRD$ specification (not shown), the NYC death data are also well captured. The SIR specification fails to do so in a fundamental way, not producing a simulated dynamic path, as its time-scales are unable to generate a path similar to the data.

The key take away point is that in NYC data, the prevalent SIR specification runs into problems when trying to reproduce observed disease dynamics. It manages to capture the initial fast growth phase only by assuming an implausibly high reproduction number, and it is fundamentally unable to capture the fast subsequent decline in fatalities. The $SEIR$ and $SIRD$ models fit the data dynamics well.

5 Disease Dynamics

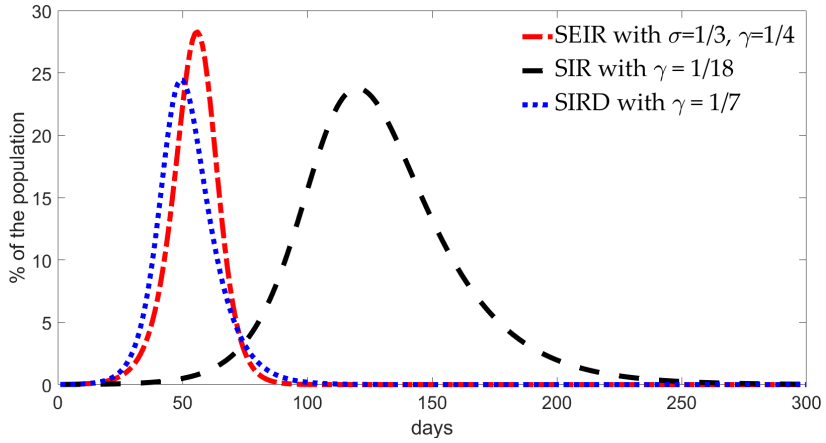
We now turn to discuss the dynamics of the epidemic that are implied by the three models. This discussion does not depend on the specific data set or empirical findings of the preceding section and applies universally. It uses simulation, relying on the epidemiological parameters discussed for each model in Section 3 and the initial value of the reproduction parameter, $\mathcal{R}_0$, at 2.50, which is a widespread estimate[13] Note that this analysis focuses on the basic, unmitigated properties of the disease as implied by the different specifications. In contrast, in real world data, one observes a disease, which is subject to suppression measures.

Figure 3 illustrates the development of the disease, as measured by the stock of infectious and exposed people (panel a) and the critically ill (panel b), under the three models. The $SEIR$ model is shown by the red line (dash-dotted); the SIR model by the black line (dashed); and the $SIRD$ model by the blue line (dotted). Table 3 presents the numerical values of the parameters and indicators which describe these dynamics.

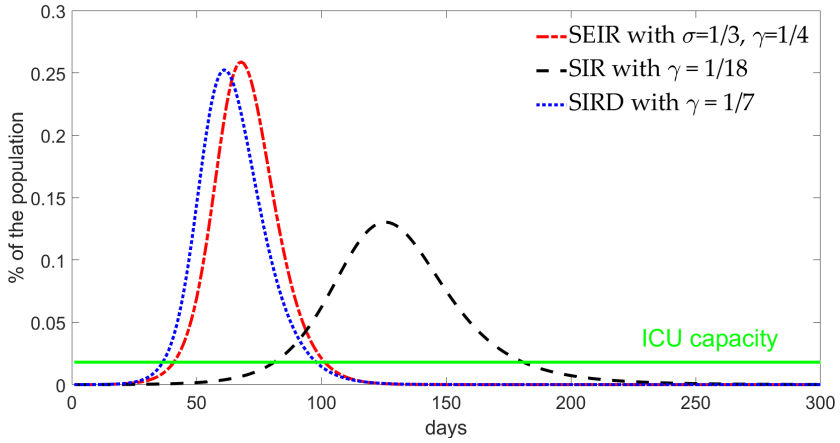
[12] Recall that we have defined, in sub-section 3.1.1, two regimes – lockdown $\mathcal{R}_L$ and out of lockdown, work $\mathcal{R}_W$, reflecing changes in behavior relative to $\mathcal{R}_0$.

[13] The approach of this section is different from the one used in the preceding Section 4, which had derived $\mathcal{R}_t$ from daily death data for NYC. The values used below, though, are close.

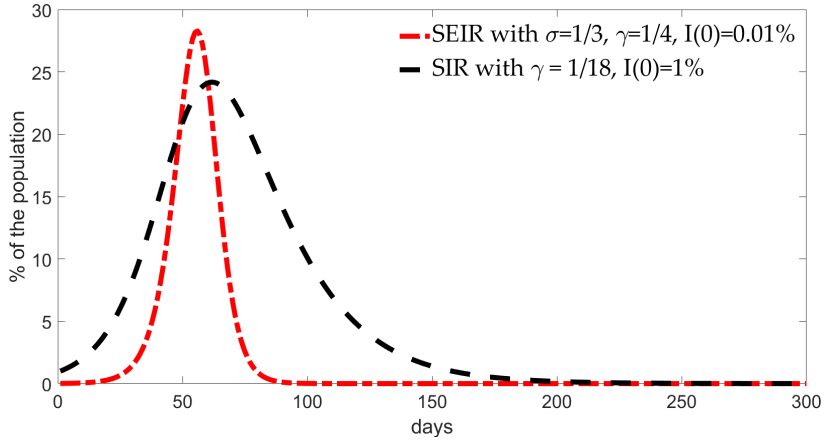
Figure 3: Disease Dynamics



a. Exposed and Infectious ($E + I$)



b. Critically Ill (X)



c. $SEIR$ seed 10^{-4}, SIR seed 10^{-2}

Table 3: Properties of the Three Models

	$SEIR$	SIR	$SIRD$
Parameterization			
σ	$1/3$	$-$	$-$
γ	$1/4$	$1/18$	$1/7$
θ	$-^{(a)}$	$-^{(a)}$	$1/11$
$Scale^{(b)}$	$E_1 + E_2 + I_1 + I_2$	I	I
Implied transmission rate given $\mathcal{R}_0 = 2.50$			
β	$\mathcal{R}_0 \cdot 1/4 = 0.625$	$\mathcal{R}_0 \cdot 1/18 = 0.139$	$\mathcal{R}_0 \cdot 1/7 = 0.357$
Implied growth rate, doubling time and disease scale at peak given $\mathcal{R}_0 = 2.50$			
$\lambda^{(c)}$	0.18	0.08	0.21
$t^{(d)}$	3.91	8.32	3.23
$Scale^{*(e)}$	0.27	0.23	0.23

Notes:

(a) there is no duration for the R state in these models.

(b) scale of the disease – the number of people who are either infectious or exposed (i.e., will become infectious).

(c) exponential growth rate.

(d) doubling time.

(e) scale of the disease at the peak.

The key difference between the models lies in the implied transmission rate $\beta(t)$, as seen in the fourth row of Table 3. Specifications that assume a long infectious period ($\frac{1}{\gamma}$) have to posit a low transmission rate $\beta(t)$ in order to match the particular value of $\mathcal{R}_0$ used, while specifications that assume the epidemiologically-grounded short infectious period, posit a relatively high $\beta(t)$ (see equation (7)).

We draw from Figure 3 and Table 3 the following key lessons.

a. *The disease in the widely-used SIR model is much slower than in the SEIR model.* A specification with a very long infectious period , as in the SIR model with $1/\gamma = 18$, implies a much lower transmission rate $\beta(t)$ and therefore much slower disease progression. It implies a growth rate of 8% and a doubling time of 8.3 days. As a result, the epidemic is spread out in time and it takes almost 330 days for it to die out. By contrast, the correctly-specified $SEIR$ model implies much faster dynamics. The epidemic starts aggressively with growth rates of 18%, and cases rise very fast, doubling every 3.9 days. The epidemic also dies off quickly under $SEIR$; the entire episode ends twice as fast as under the SIR model specification.

b. *The disease in SIRD is somewhat faster than in SEIR.* Ignoring the short latent period (E), as in $SIRD$, has moderate effects on epidemic dynamics. In $SIRD$, relative to the $SEIR$ model, the epidemic develops somewhat faster at the beginning, because there is no delay between the moment a person becomes infected and the moment he or she starts spreading the disease.

c. *Widely-used SIR and SIRD imply a much lower disease scale than SEIR.* At the peak, in both SIR and $SIRD$ specifications, the number of infectious/exposed people reaches 23% of the population, whereas in the $SEIR$ model, a higher level of disease is reached. There is a difference of 3.5 percentage points relative to the other models, or 11.6 million people in the case of the entire U.S. economy. This can be seen in the higher peak of the red line of $E + I$ in Figure 3 and in the numbers presented in Table 3.

d. *Substantially delayed pressure on ICU capacity in SIR.* Panel b of Figure 3 shows that with a slow moving disease, implied by a long infectious period as in the widely-used SIR model, ICU capacity is breached on day 82, and peak demand exceeds capacity by a factor of 7, whereas in the epidemiological-grounded $SEIR$ model it is breached much earlier, on day 41, and peak demand exceeds capacity by a factor of 14. The $SIRD$ specification performs close to $SEIR$ in this case.

e. *Implications for initial conditions.* Under equal initial conditions, it takes much more time for the epidemic to gain pace under the SIR model than under $SEIR$ so that the peak arrives almost two months later relative to the correct model, and is more moderate in scale relative to SEIR . One can try to 'circumvent' this problem by assuming a higher initial seed of the infection. Panel c of Figure 3 compares the $SEIR$ model with initial seed of 10^{-4} and the SIR model with initial seed of 10^{-2}. It shows that assuming a higher initial seed places SIR on a relatively similar timescale as $SEIR$ in terms of the timing of the peak. However, two serious problems remain. First, at peak, the implied number of infectious individuals is still significantly lower under SIR, which distorts the

problem of a policymaker constrained by a number of ICU beds. Second, assuming a seed of 1% of the population implies, in terms of the U.S. economy, that the epidemic has started when over 3.3 million people were infected. This is a highly implausible assumption, given actual data on the path of known cases and on deaths.

The separation of the infection generation block from the clinical block lies at the heart of the differences between the prevalent SIR set up and the benchmark $SEIR$ model. Targeting two separate timescales with one parameter (γ) leads to severe distortions of the implied dynamics of the disease. The $SIRD$ model presented in sub-section 3.2.2 alleviates the problem somewhat by adding a parameter θ thus enabling separate targeting of $\mathcal{R}_0$ based on the duration of infectiousness period and of duration-till-death. This makes the dynamic properties of SIRD relatively close to those of the correct SEIR model. But, as shown above, its scale distortion is substantial and, as we show below, it remains problematic when coming to formulate policy.

6 Erroneous Modelling: Implications for Optimal Policy

A key aim of this paper is to show the implications of the modelling of the disease for policy responses and to highlight how erroneous modelling is costly. To do so, we use an optimizing planner model of the kind used in the papers cited in Section 2. We simulate optimal policy undertaken when the planner uses each of the three models discussed above, but actual disease dynamics are given by the afore-cited $SEIR$ model. We set up a planner model and calibrate it (6.1); we then explain the simulation methodology (6.2) and present the results (6.3).

6.1 The Planner Model and Its Calibration

The planner minimizes the following loss function:

$$\min_{T_0,T_1} V = \int_{t=0}^{T_V} e^{-rt} \left(\frac{Y^{SS}}{N^{SS}} \left(N^{ss} - N\left(t\right) \right) + \chi Y^{SS} \dot{D}\left(t\right) \right) dt + R_D(T_V) + R_Y(T_V) \tag{28}$$

subject to equations (1)-(15).

The loss function is minimized in PDV terms (r is the discount rate), where at finite point T_V (which we set at 540 days) a vaccine is found and the pool of susceptibles drops to zero, so that the disease stops growing.[14] After time T_V, there is a residual death toll $R_D(T_V)$ and residual output loss $R_Y(T_V)$, which accompany the decline of the epidemic.

[14] We make the assumption of $T_V = 540$ given the progress actually made in 2020 and the start of vaccination in December 2020. In terms of the model, T_V refers to the time of sufficient vaccination; with logistics, production times, gradual take-up rates, etc. an expected 540 days seems reasonable at the time of writing (March 2021).

The loss function includes two terms. One is lost output $\frac{Y^{SS}}{N^{SS}}$, due to a decline in employment N relative to steady state N^{ss}; equation (15) provides the connection between employment and the epidemiological states. Essentially, employment declines because of lockdown measures (expressed in the parameter ρ) and because people fall ill or die. The second term is the value of lost life. This is a function of the flow of fatalities $\dot{D}(t)$ translated into lost output terms using the value of statistical life (with the parameter χ). This flow is affected by the breach of ICU modelled in equation (9) above. To work within a realistic but simple set-up, we let the planner decide on when to start (T_0) and stop (T_1) a full lockdown.

To calibrate the model, we use the parameterization of Table 1. For the values of the time-varying $\mathcal{R}_t$ under different planner policies – see the discussion in sub-section 3.1.1 above – we use data estimates for the U.S. as follows.[15]

i) Initial level. We set

$$\mathcal{R}_0 = 2.50 \tag{29}$$

We get the value of 2.50 in equation (29) by using the methodology of Fernandez-Villaverde and Jones (2020), adapted to our model, and U.S. daily death flow data taken from Johns Hopkins University CSSE (2020). This yields estimates of $\mathcal{R}_0$ values of 2.67 on March 17, 2020 and 2.48 on March 18, 2020. This is the reproduction number during the initial phase of the epidemic, before significant lockdowns were imposed in the U.S.

ii) Subsequent values. To reflect the fact that over the course of the initial outbreak and following it, individuals change their modes of behavior and economic activity, including compliance with NPIs, we allow the reproduction number in subsequent periods to be lower than the initial $\mathcal{R}_0$. As noted above, we posit that there is a value of $\mathcal{R}_t$ during times of lockdown, to be denoted $\mathcal{R}_L$, and another value at other times, denoted $\mathcal{R}_W$ ("work"). Both are lower than $\mathcal{R}_0$ to take into account the fact that individuals have adjusted to the new environment and are taking more precautions. When lockdowns are in place, policy and individual responses together engender $\mathcal{R}_L < \mathcal{R}_W$.

For their calibration, we rely on two sets of estimates.

First, Karin et al (2020) review the literature and estimate values for $\mathcal{R}_L$ and $\mathcal{R}_W$.[16] These relate to developed countries with a population density of over 100 people per square km. The estimates indicate a value of 1.50 for $\mathcal{R}_W$ as the upper bound; the estimates for $\mathcal{R}_L$ range from 0.6 to 0.9 with a value of 0.80 as the estimate for NYC, the only U.S. location examined.

Second, we use the U.S. estimates of Fernandez-Villaverde and Jones (2020) for the biggest 15 states in the U.S., covering 65% of the U.S. population[17] We look at the minimal and maximal values of the estimated $\mathcal{R}_t$ series from April

[15]Note that in Section 4 we have derived estimates of the reproduction parameter for NYC only.

[16]The full details of their analysis, including references and the code, are found at: www.github.com/milo-lab/

[17]Due to insuffiicent estimates, we exclude the state of New Jersey. As noted, these authors infer $\mathcal{R}_t$ from daily death flow data taken from Johns Hopkins University CSSE (2020).

1, 2020 till September 30, 2020. According to the Oxford Stringency Index this period covers lockdowns and release in all of the states, at different points in time. These $\mathcal{R}_t$ values are indicative of $\mathcal{R}_L$ and $\mathcal{R}_W$: $\mathcal{R}_L$ cannot be lower than the minimal data value, and $\mathcal{R}_W$ cannot be higher than the maximal data value. The median (average) minimal value across the 15 states is 0.68 (0.61) and the median (average) maximal value across the 15 states is 1.42 (1.49).

Given these two sets of estimates we posit values that are conservative, in the sense that $\mathcal{R}_L$ and $\mathcal{R}_W$ are calibrated at relatively high values:

$$\mathcal{R}_t = \begin{Bmatrix} 1.50 & \mathcal{R}_W, work \\ 0.80 & \mathcal{R}_L, lockdown \end{Bmatrix} \tag{30}$$

(iii) Dynamics of the reproduction parameter. To capture the gradual nature of learning and adjustment of individual behavior, we posit that a certain minimal time should pass under lockdown before the reproduction number declines from its initial value $\mathcal{R}_0$ to $\mathcal{R}_W$. To calibrate the path at this time span, we look at two sources..

a. Using the Fernandez-Villaverde and Jones (2020) methodology applied to our model, we get that it takes 8 days to get from $\mathcal{R}_t = 2.48$ to $\mathcal{R}_t = 1.50$. This decline took place in the third week of March, when lockdowns only started to unfold. Thus, we interpret this decline mainly as the rational adjustment of behavior.

b. We use Imperial College COVID-19 Response Team (2020) estimates $\mathcal{R}_t$ for U.S. states since the start of the epidemic. We focus on the initial decline of $\mathcal{R}_t$ when suppression measures have been undertaken across the US and assume a log-linear decay function[18]

$$\ln \mathcal{R}(T_1^s) = \ln \mathcal{R}(T_0^s) - \widehat{\alpha}_s^{\log -linear} \cdot (T_1^s - T_0^s) \tag{31}$$

$$\widehat{\alpha}^{\log -linear} = \frac{\sum_{1}^{45} \alpha_s^{\log -linear}}{S} \tag{32}$$

We obtain two alternative estimates for the average speed $\widehat{\alpha}^{\log -linear}$ of a $\mathcal{R}_t$ decline, depending on the definition of the decline in period T_0^s to T_1^s:

a. The speed $\widehat{\alpha}^{\log -linear} = 0.027$ per day was obtained when $T_0^s =$ the day of the first $\mathcal{R}_t$ observation in the state, and $T_1^s =$ end of the decline (i.e., the point where $\mathcal{R}_t$ is not statistically different from 1 at 5%). It thus takes $\frac{\ln(2.5)-\ln(1.5)}{0.027} =$ 19.2 days to get from 2.50 to 1.50.

b. The speed $\widehat{\alpha}^{\log -linear} = 0.065$ per day was obtained when $T_0^s =$ the day the highest $\mathcal{R}_t$ observed in the state; T_1^s as in (a). It takes $\frac{\ln(2.5)-\ln(1.5)}{0.065} = 7.9$ days to get from 2.50 to 1.50.

[18]Out of the fifty states and DC, six were not included in this analysis (AK, HI, MT, ND, SD, WY) because their initial values of the reproduction number were already below unity at the start of the epidemic.

The decay time of $\mathcal{R}_t$ is 8 days based on Fernandez-Villaverde and Jones (2020) national death data, or 8 or 19 days based on the state-level data, Imperial model estimates. Again, we adopt a conservative calibration and assume that 14 days must pass before $\mathcal{R}_t$ declines from 2.50 to 1.50.

Referring to the U.S. economy, as we do throughout, we use $\rho = 0.65$ for the fraction of workers able to work in a lockdown, consistent with the values used in Kaplan, Moll, and Violante (2020),[19] and $\chi = 85.7$ for the value of lost life.[20]

6.2 Simulation Methodology

To analyze the costs of basing policy on a mis-specified model, we proceed as follows. We assume that the disease always behaves according to the epidemiologically based $SEIR$ model of sub-section 3.1.1 above. But the planner uses one of the three models discussed above, i.e., the correct one or one of the two erroneous ones, when deriving the optimal intervention timings T_0 and T_1.

In particular, we use the two specifications discussed above: (1) SIR with $\gamma = 1/18$, and (2) $SIRD$ with $\gamma = 1/7$. For each model, we simulate optimal policy while the disease in fact behaves according to $SEIR$, and record ensuing deaths and ICU breaches.

To derive an optimal policy under each epidemiological model, we use a numerical solver in Mathematica (NDSolve, see Abell and Braselton (2016)). We find the values of control variables (T_0, T_1) that minimize the cost function in (28) by conducting an exhaustive search of the control variables space. First, this is done using a hierarchical search on a coarse grid of integer values such that $0 \leq T_0 \leq T_1 \leq 540$ and T_0 and T_1 are multiples of 8. For each combination of T_0 and T_1 we solve the continuous time system of ODE describing the stocks dynamics in each model, according to the calibration described above. Initial values for this system are set from a seed of 0.01% of the population (100 people per million). Infectious (and latent in $SEIR$) compartments (and subcompartments in $SEIR$) are initialized so they are consistent with an exponential growth rate of the disease; susceptibles ($S(t)$) form the rest of the pool. The clinical block is initialized to 0 at $t = 0$.

Using the solution of the ODE system – a set of interpolated functions describing stocks dynamics – we are able to evaluate the planner's objective (28), which is a function of these stocks (through death $D(t)$ and employment $N(t)$). To find the global minimum of the objective function (28), we maintain a set of

[19] This value is reinforced by the findings in Dingel and Neiman (2020) about remote work. We set $\phi = 0$.

[20] We compute the value of life as follows:

$$\begin{aligned} \chi &= \frac{\text{expected years remaining} \cdot \text{value of statistical life}}{\frac{Y}{POP}} \\ &= \frac{14 * 400,000}{65,351} = 85.7 \end{aligned}$$

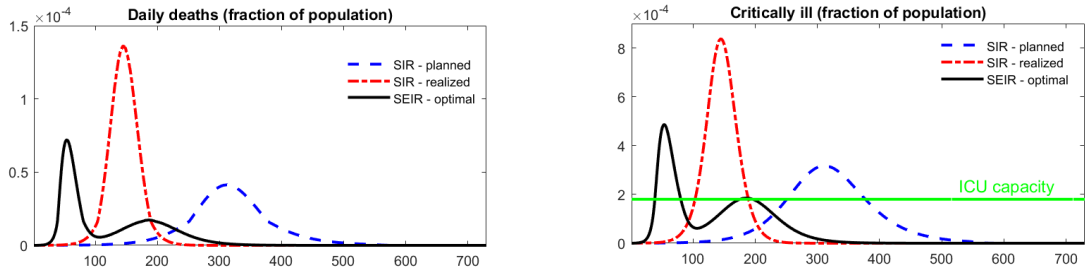
The resulting value conforms the high end of the estimates discussed in Hall, Jones, and Klenow (2020).

the best possible minima, spanning the control variables space. We then recursively refine the grid, until the desired granularity of 1 day is reached for both T_0 and T_1.

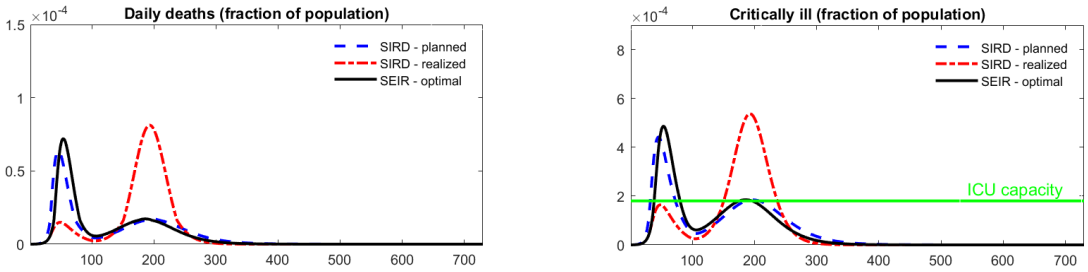
6.3 Results

For each model, we present both the planned outcome and the realized outcome obtained by applying the policy to the actual disease. This exercise illustrates the price of deriving policy based on erroneous assumptions. Figure 4 and Table 4 report the results.

Figure 4: Planner Policy in the Three Models



a. SIR compared to SEIR



b. SIRD compared to SEIR

Table 4: Optimal Policy Timing and Outcomes

	Timing		Planner cost V		Output loss V_Y		D per 10^6	
	T_0	T_1	planned	realized	planned	realized	planned	realized
$SIR, \gamma = 1/18$	0	14	0.484	0.651	0.020	0.025	5,609	7,421
$SIRD, \gamma = 1/7$	30	91	0.426	0.561	0.064	0.066	4,278	5,895
$SEIR, \sigma = 1/3, \gamma = 1/4$	37	84	0.418	0.418	0.051	0.051	4,335	4,335

Notes:

1. The values in the table are defined as follows:

$$\min_{T_0,T_1} V = \int_{t=0}^{540} e^{-rt} \left(\frac{Y^{SS}}{N^{SS}} \left(N^{ss} - N\left(t\right) \right) + \chi Y^{SS} \dot{D}\left(t\right) \right) dt + R_D(540) + R_Y(540)$$

$$V_Y = \int_{0}^{540} e^{-rt} \left(\frac{Y^{SS}}{N^{SS}} (N^{SS} - N(t)) \right) dt + R_Y(540)$$

2. D is the stock of fatalities, given here in terms of persons per 1 million.

We see that the optimal timing using the $SEIR$ model of the actual disease, is to lock on day 37 for 47 days. This epidemiologically-correct timing implies two waves of the epidemic, whereby deaths are minimized. This is so as the planner accurately spreads the burden on ICU so that capacity is breached in the first wave and fully utilized in the second (see the black solid lines in panels a and b of Figure 4).

Optimal timing is very different when the planner assumes a slowly moving disease using the SIR model with $\gamma = 1/18$. Looking at the top row of Table 4, one sees that the planner locks immediately for only 14 days, the minimal time necessary to bring $\mathcal{R}_W$ down to 1.50. The planner builds on a slow-moving disease that will not massively breach ICU capacity and not cause many deaths (see planned dynamics shown in the dashed lines of panel a of Figure 4). However, in reality, the disease is much faster ($SEIR$ is much faster than SIR, as shown above), and it erupts immediately after release, breaching ICU capacity by a factor of 4 and increasing the death toll by 71% relative to the epidemiologically-correct policy. The costs of loss of output are of course low due to the very short lockdown, but total planner costs are 56% higher than under the epidemiologically-relevant strategy, due to a much higher death toll.

Under the $SIRD$ model with $\gamma = 1/7$, optimal lockdown timing is closer to the epidemiologically-grounded one, and is in fact even more conservative, with the lockdown starting one week earlier and ending one week later than under the correct $SEIR$ specification. However, timing is crucial here and seemingly more stringent policies can be as dangerous as more relaxed ones. By starting the lockdown too early relative to the policy based on the epidemiological evidence, the planner suppresses the first wave and under-utilizes ICU capacity at the beginning of the disease. After release, a second wave develops, which is much higher than the first one, with a massive breach of ICU capacity by almost three-fold and a high death toll. This type of mis-specification implies a death toll that is 36% higher relative to the epidemiologically-correct policy.

When planning interventions to manage epidemics, timing is of the essence. Locking too early, when lockdown cannot last too long, means that the second wave will be high and might breach the capacity constraint of the public health system. Locking too late, when the disease is growing exponentially, poses an immediate threat to ICU capacity and results in excess deaths. These are the kinds of outcomes that emerge when the policymaker derives lockdown timing based on erroneous assumptions on the dynamics of the disease. Using the epidemiologically-grounded model to guide policy is crucial to get the timing right and avoid unnecessary loss of life and output.

We have shown that a SIR model is very far from the $SEIR$ specification in terms of both implied disease dynamics and optimal interventions timing. We have also shown that the $SIRD$ model is close to the epidemiologically-grounded model in terms of disease dynamics, but that it implies significantly different outcomes when coming to derive and implement optimal policy timing.

Two additional notes should be made. One is that we deliberately consider a simple planner problem with only two control variables T_0 and T_1 and a binary intervention type, either full lockdown or full release. This is to illuminate the basic intuition of the policy distortions that arise due to the mis-specification of the epidemiological model. With more sophisticated tools, such as a variable lockdown or lockdown of selected population groups, planner outcomes would by definition be better under any epidemiological model. However, the costs of misspecification will persist. Any intervention, even the most sophisticated one, has a start and end date. The planner, facing a trade-off between fatalities and output losses, will get these dates wrong when perceiving the disease as slow-moving and moderate in scale. We have demonstrated that such misperceptions cost human lives and output. We show examples in an analysis of a much richer planner problem in Bar-On, Baron, Cornfeld, Milo, and Yashiv (2021), where the costs of mis-specification remain high.

The second note is that it is useful to put the fatalities numbers here (Table 4) in perspective. One comparison is to the real world. U.S. death numbers are currently (March 2021) around $530,000$ or $1,640$ per million, 38% of the $SEIR$ model scenario here. This difference arises because the model is computed over two years while the real world numbers pertain to a year, and because U.S. policymakers have imposed longer lockdowns than the planner, having access to wider policy choices. Another comparison is to the papers which model a SIR-based planner. These present relatively high numbers of deaths, typically ranging from 1% to 2% of the U.S. population, namely 3.2 to 6.4 million fatalities. These numbers are 7 to 13 times higher than the current real world numbers. At $10,000 - 20,000$ per million, they are 1.35 to 2.7 times higher than the worst case SIR scenario here.

7 Conclusions

The key take away is that erroneous modelling of epidemic dynamics is of crucial importance for economic analysis. Seemingly innocuous modelling choices have direct consequences for the implied scale and speed of the disease, which in turn impact optimal policy planning. The outcomes analyzed in research in Economics on COVID19 usually include death tolls and GDP loss. If the epidemic is deemed slower and less severe than it really is, the consequences are dramatically higher death tolls and excessive output loss relative to the case where the policymaker bases interventions on the correct epidemiological model.

In companion work (Bar-On, Baron, Cornfeld, Milo, and Yashiv (2021)), we use the epidemiologically-grounded model presented here to analyze an optimal planner model. Exploring the stringency of lockdown policies and its timing, the emerging optimal policy is quite different from the one proposed thus far in the Economics literature, and is shown to improve on real-world outcomes.

References

[1] Abel, Andrew B. and Stavros Panageas, 2020. "Optimal Management of a Pandemic in the Short Run and the Long Run," NBER Working Paper No. 27742.

[2] Abell, Martha L. and James P. Braselton, 2016. **Differential Equations with Mathematica**, 4th edition, Academic Press.

[3] Acemoglu, Daron, Victor Chernozhukov, Ivan Werning, and Michael D. Whinston, 2020. "A Multi-Risk SIR Model with Optimally Targeted Lockdown" **AER Insights,** forthcoming.

[4] Alvarez, Fernando, David Argente, and Francesco Lippi, 2020. "A Simple Planning Problem for COVID-19 Lockdown," **AER Insights,** forthcoming.

[5] Avery, Christopher, William Bossert, Adam Clark, Glenn Ellison, and Sara Fisher Ellison, 2020. "An Economist's Guide to Epidemiology Models of Infectious Disease" **Journal of Economic Perspectives** 34, 4, 79–104.

[6] Bar-On, Yinon, Tanya Baron, Ofer Cornfeld, Ron Milo, and Eran Yashiv, 2021. "When to Lock – Not Whom: Managing Epidemics Using Time-Based Restrictions," working paper, available at https://www.tau.ac.il/~yashiv/

[7] Bar-On, Yinon, Ron Sender, Avi Flamholz, Rob Phillips, and Ron Milo, 2020. "A Quantitative Compendium of COVID-19 Epidemiology," arXiv:2006.01283; available at https://arxiv.org/abs/2006.01283

[8] Champredon, David, Jonathan Dushoff, and David J.D. Earn, 2018. "Equivalence of the Erlang-distributed SEIR Epidemic Model and the Renewal Equation," **SIAM Joural of Applied Math** 78, 6, 3258–3278.

[9] Dingel, Jonathan I. and Brent Neiman, 2020. "How Many Jobs Can be Done at Home?" **Journal of Public Economics** 189, 104235

[10] Ellison, Glenn , 2020. "Implications of Heterogeneous SIR Models for Analyses of COVID-19," NBER Working Paper No. 27373

[11] Fernandez-Villaverde, Jesus and Chales I. Jones, 2020. "Estimating and Simulating a SIRD Model of COVID-19 for Many Countries, States, and Cities," NBER Working Paper No. 27128.

[12] Greenstone, Michael and Vishan Nigam, 2020. "Does Social Distancing Matter?" BFI working paper.

[13] Hall, Robert E, Charles I. Jones, and Peter J. Klenow, 2020. "Trading Off Consumption and COVID-19 Deaths,"**Minneapolis Fed Quarterly Review** 42, 1, 2-14.

[14] He, Xi , Eric H. Y. Lau, Peng Wu, Xilong Deng, Jian Wang, Xinxin Hao, Yiu Chung Lau, Jessica Y. Wong, Yujuan Guan, Xinghua Tan, Xiaoneng Mo, Yanqing Chen, Baolin Liao, Weilie Chen, Fengyu Hu, Qing Zhang, Mingqiu Zhong, Yanrong Wu, Lingzhai Zhao, Fuchun Zhang, Benjamin J. Cowling, Fang Li, Gabriel M. Leung, 2020. "Temporal Dynamics in Viral Shedding and Transmissibility of COVID-19," **Nature Medicine** 26, 672–675.

[15] Huang, Chaolin, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, Zhenshun Cheng, Ting Yu, Jiaan Xia, Yuan Wei, Wenjuan Wu, Xuelei Xie, Wen Yin, Hui Li, Min Liu, Yan Xiao, Hong Gao, Li Guo, Jungang Xie, Guangfa Wang, Rongmeng Jiang, Zhancheng Gao, Qi Jin, Jianwei Wang, Bin Cao, 2020. "Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China," **Lancet** 395: 497–506

[16] Imperial College COVID-19 Response Team, 2020. "Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries," available at https://dsprdpub.cc.ic.ac.uk:8443/bitstream/10044/1/77731/10/2020-03-30-COVID19-Report-13.pdf

[17] Johns Hopkins University CSSE, 2020. "2019 Novel Coronavirus COVID-19 (2019-nCoV) Data Repository," Center for Systems Science and Engineering.

[18] Jones, Callum J., Thomas Philippon, Venky Venkateswaran, 2020. "Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home," NBER Working Paper No. 26984.

[19] Kai-Wang To, Kelvin, Owen Tak-Yin Tsang, Wai-Shing Leung, Anthony Raymond Tam, Tak-Chiu Wu, David Christopher Lung, Cyril Chik-Yan Yip, Jian-Piao Cai, Jacky Man-Chun Chan, Thomas Shiu-Hong Chik, Daphne Pui-Ling Lau, Chris Yau-Chung Choi, Lin-Lei Chen, Wan-Mui Chan, Kwok-Hung Chan, Jonathan Daniel Ip, Anthony Chin-Ki Ng, Rosana Wing-Shan Poon, Cui-Ting Luo, Vincent Chi-Chung Cheng, Jasper Fuk-Woo Chan, Ivan Fan-Ngai Hung, Zhiwei Chen, Honglin Chen, Kwok-Yung Yuen, 2020. "Temporal Profiles of Viral Load in Posterior Oropharyngeal Saliva Samples and Serum Antibody Responses During Infection by SARS-CoV-2: an Observational Cohort Study, " **The Lancet Infectious Diseases** 20, 5,565–74.

[20] Kaplan, Greg, Ben Moll, and Gianluca Violante, 2020. "The Great Lockdown and the Big Stimulus: Tracing the Pandemic Possibility Frontier for the U.S." NBER Working Paper No. 27794.

[21] Karin, Omer, Yinon Bar-On, Tomer Milo, Itay Katzir, Avi Mayo, Yael Korem, Avichai Tendler, Boaz Dudovich, Eran Yashiv, Amos J Zehavi, Nadav Davidovitch, Ron Milo and Uri Alon,

2020. "Adaptive cyclic exit strategies from lockdown to suppress COVID-19 and allow economic activity", MedRxiv, available at https://www.medrxiv.org/content/10.1101/2020.04.04.20053579v4.full.pdf

[22] Kermack, William O., and Anderson G. McKendrick, 1927. "A Contribution to the Mathematical Theory of Epidemics," **Proceedings of the Royal Society London**. Ser. A., 115, 700–721.

[23] Li, Ruiyun , Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, and Jeffrey Shaman, 2020. "Substantial Undocumented Infection Facilitates the Rapid Dissemination of Novel Coronavirus (SARS-CoV-2)," **Science** 368(6490), 489-493.

[24] Mills, Christina E., James M. Robins, and Marc Lipsitch, 2004. "Transmissibility of 1918 Pandemic Influenza, " **Nature** 432, 904-906.

[25] Richardson Safiya, Jamie S. Hirsch, Mangala Narasimhan, et al., 2020. "Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area." **JAMA** 323(20):2052–2059.

[26] Salje, Henrik, Cecile Tran Kiem, Noemie Lefrancq, Noemie Courtejoie, Paolo Bosetti, Juliette Paireau, Alessio Andronico, Nathanael Hoze, Jehanne Richet, Claire-Lise Dubost, Yann Le Strat, Justin Lessler, Daniel Levy-Bruhl, Arnaud Fontanet, Lulla Opatowski, Pierre-Yves Boelle, Simon Cauchemez, 2020. "Estimating the Burden of SARS-CoV-2 in France," **Science** 369, 208–211.

[27] Stadlbauer, Daniel, Jessica Tan, Kaijun Jiang, Matthew M. Hernandez, Shelcie Fabre, Fatima Amanat,, Catherine Teo, Guha Asthagiri Arunkumar, Meagan McMahon, Jeffrey Jhang, Michael D. Nowak, Viviana Simon, Emilia Mia Sordillo, Harm van Bakel, and Florian Krammer, 2021. "Repeated Cross-Sectional Sero-Monitoring of SARS-CoV-2 in New York City," **Nature** 590, 146–150. https://doi.org/10.1038/s41586-020-2912-6

[28] Tian, Huaiyu Yonghong Liu, Yidan Li, Chieh-Hsi Wu, Bin Chen, Moritz U. G. Kraemer, Bingying Li, Jun Cai, Bo Xu, Qiqi Yang, Ben Wang, Peng Yang, Yujun Cui, Yimeng Song, Pai Zheng, Quanyi Wang, Ottar N. Bjornstad, Ruifu Yang, Bryan T. Grenfell, Oliver G. Pybus, and Christopher Dye, 2020. "An Investigation of Transmission Control Measures During the First 50 Days of the COVID-19 Epidemic in China," **Science** 368(6491), 638-642.

[29] Wallingaa, Jacco, Michiel van Boven, and Marc Lipsitch, 2010. " Optimizing Infectious Disease Interventions During an Emerging Epidemic" **Proceedings of the National Academy of Sciences** 107, 2, 923–928.

[30] Wallinga, Jacco and Marc Lipsitch, 2007. "How Generation Intervals Shape the Relationship between Growth Rates and Reproductive Numbers," **Proceedings of the Royal Society** 274, 599–604

[31] Wearing, Helen J., Pejman Rohani, and Matt J. Keeling, 2005. " Appropriate Models for the Management of Infectious Diseases," **PLoS Medicine** 2, 7, 621-627.