SCARE:
A statistical – simulation approach to Covid-19

April 23, 2020

André de Palma, ENS Paris-Saclay and CREST, FRANCE

Nathalie Picard, CY Cergy Paris Université, FRANCE

Stef Proost, KUL, BELGIUM

Abstract

We develop an epidemic model to explain and predict the dynamics of the COVID-19 virus. The standard 3-variable model, SIR (Susceptible, Infected and Removed), is extended in the SCARE model where 5 variables are considered: Susceptible, Contaminated, Affected (i.e. sick), Removed and Eliminated (i.e. dead). The variables are related with conversion factors, which are piecewise constant over time, and estimated on empirical (WHO) data. The model is estimated on four countries (Belgium, France, Germany and USA), and provide preliminary predictions of the number of cases and the number of deaths. Parameters are estimated based on the minimization of the distance between observed and simulated data. The objective function combines instantaneous and cumulated cases and deaths.

KEY WORDS: Epidemic model, SIR, Covid-19, differential equations, confinement, SCARE

Acknowledgements

We would like to thank Daron Acemoglu, who emphasized the importance of extending SIR model during his webinar organized by Royal Economic Society on April, 15. We would like also to thank Claude Lefèvre for discussions on various extensions of the SIR model, as well as Philippe Vercruysse for the Belgian data and insights he provided to us.

Introduction

We develop an epidemiological model to explain and predict the dynamics of the COVID-19 virus. The standard 3-variable model, SIR (Susceptible, Infected and Removed) of Kermack & Kendrick (1927) is extended into a 5 variable model (SCARE) that allows to match better the observations. We now distinguish between the Susceptible, Contaminated, Affected (i.e. sick), Removed and Eliminated (i.e. dead). We use the WHO data, but we have to make the assumption that the number of cases (contaminated persons) reported and the number of deaths differ from the number of real cases. The real variables are related with conversation factors, which are piecewise constant over time, to the empirical (WHO) data. The model has 5 parameters and 5 ancillary parameters: the conversion rates (2 with 2 time periods), the number of cases at the starting date of the simulation (1). The parameters are estimated by minimizing the square relative distance between reported and predicted data. The model is estimated for four countries (Belgium, France, Germany and USA), and gives promising results. This SCARE model will be the core of an extended model that also includes economic effects and will allow to test different policies: partial and local confinement, use of different types of tests, use of masks, inter alia.

1 The canonical SIR model

Ross, Kermack and McKendrick have developed a canonical model to describe the transmission of a disease. They consider a population with three groups: Susceptible, Infected, Recovered. Each group is homogenous in the sense that all individuals in the group are statistically identical. Before the disease starts, each individual is Susceptible in the sense that s/he is not infected, but may get the disease. Once s/he gets the disease, s/he becomes Infected. Once Infected, the individual will after a certain time become Recovered. The Susceptible individuals get Infected by a contact with an Infected individual. The Recovered individuals do not infect the Susceptible. All transitions are irreversible in the sense that a Susceptible can only become Infected and an Infected can only become Recovered. Recovered is an absorbing state in the sense that there is no escape from this state. Once the disease starts, there are a few Infected individuals in the population (and no Recovered). Once, by contact, this Infected population contaminates some Susceptible, the disease starts to spread. Then, the number of Susceptible monotically decreases over time and the number of monotically increases over time. By contrast, the number of Infected individuals first increases and then decreases over time. The process reaches a stationary solution where there are no more Infected individuals and some fraction (between 0 and 1) of Susceptible and Recovered.

This dynamics can be described by a system of differential equations. We briefly sketch the SIR model below. It will be extended in Section 1.2. Let denote by t the time. At time t, we denote by $S(t)$ the number of Susceptible individuals, by $I(t)$ the number of Infected individuals and by $R(t)$ the number of Removed individuals. If N denotes the total population, we have at any time $S(t)+I(t)+R(t)=N$. We denote the corresponding fractions of individuals by $s(t)= S(t)/\mathrm{N}$, $i(t)= I(t)/\mathrm{N}$ and $r(t)= R(t)/\mathrm{N}$, with $s(t)+i(t)+r(t)=1$.

The corresponding standard differential equations of the SIR model for the fractions of individuals are the following:

$$\begin{cases} \dfrac{ds(t)}{dt} = -\beta s(t) i(t) \\ \dfrac{di(t)}{dt} = \beta s(t) i(t) - \gamma i(t). \\ \dfrac{dr(t)}{dt} = \gamma i(t) \end{cases}$$

The key parameter is β/γ, where β measures the strength of interactions between Susceptible and Infected and γ measures the recovery rate. On average, each infected individual generates $\beta s(t)$ new infected individuals per unit of time. The β parameter depends on the confinement policy, as will be discussed extensively later on. The recovery rate depends on the health conditions as well as on the intensity of medical care. This suggests the coupling between the epidemic model and the economic model, which will be exploited later on.

In the next subsection, we extend the SIR model into a framework more amenable to explain the experimental data collected for the Covid-19.

2 The SCAR*E* proposed approach

In the SIR model, once an individual is infected, after a while, he transits to a single case, Removed.

Here we consider that an additional category, Carrier: an individual may carry the virus (and be contagious) before showing symptoms. A carrier may transit to 2 different states: he can be affected (show symptoms), or he can recover spontaneously. We also consider another category, Eliminated, corresponding to death. An Affected individual may transit to 2 different states: he can be recover, or he can die.

We thus consider that each individual may be in 1 out of 5 states:

S: Susceptible; C: Carrier; A: Affected; R: Recovered; E: Eliminated.

The fraction of individuals in the corresponding states are denoted by $s(t)$, $c(t)$, $a(t)$, $r(t)$, $e(t)$.

The transition matrix of the SCARE model is presented in Table 1.

To From	**S**	**C**	**A**	**R**	**E**
S	1- $\beta \cdot [c(t)+ a(t)]$	$\beta \cdot [c(t)+ a(t)]$	0	0	0
C	0	1- α - μ	α	μ	0
A	0	0	1- γ - λ	γ	λ
R	0	0	0	1	0
E	0	0	0	0	1

Table 1: Transition probabilities in the SCARE model

The corresponding dynamics is given by:

$$\begin{cases} \dfrac{ds(t)}{dt} = -\beta s(t)\cdot\left[c(t)+a(t)\right] \\ \dfrac{dc(t)}{dt} = \beta s(t)\cdot\left[c(t)+a(t)\right] - (\alpha+\mu)c(t) \\ \dfrac{da(t)}{dt} = \alpha c(t) - (\gamma+\lambda)a(t) \\ \dfrac{dr(t)}{dt} = \mu c(t) + \gamma a(t) \\ \dfrac{de(t)}{dt} = \lambda a(t) \end{cases} .$$

It can easily be checked that $s(t)+ c(t)+ a(t)+ r(t)+ e(t)=1$.

We have assumed here that when an individual is removed, he will never be infected anymore. This is an optimistic hypothesis that we make because there is no clear evidence that a removed individual can be again infected. This hypothesis will be removed if there is enough scientific evidence that it is violated.

3 From SCARE model to WHO daily situation reports

For the ease of comparison with statistics, we rewrite the differential equation in terms of total population in each class:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\tilde{\beta} S(t) \cdot \left[C(t) + A(t) \right] \\ \dfrac{dC(t)}{dt} = \tilde{\beta} S(t) \cdot \left[C(t) + A(t) \right] - (\alpha + \mu) C(t) \\ \dfrac{dA(t)}{dt} = \alpha C(t) - (\gamma + \lambda) A(t) \\ \dfrac{dR(t)}{dt} = \mu C(t) + \gamma A(t) \\ \dfrac{dE(t)}{dt} = \lambda A(t) \end{cases},$$

where $\tilde{\beta} = \dfrac{\beta}{N(t)}$.

WHO reports (WHO (2020) daily the number of new "cases", cumulated number of cases, new deaths and cumulated death in all countries affected by the virus. We thus consider a time interval Δt=1 day. Countries are indexed by *i*, *i*=1,…,*I*.

A new "Case" during day *t*, i.e. during time period [*t*,*t*+1) is an individual who transits from Carrier to Affected (sick).

The actual number of New Cases in country *i* at day *t* is $\omega_i(t) = \alpha_i C_i(t)$. The cumulated number of cases in country *i* at day *t* is $\Omega_i(t) = \sum_{\tau=1}^{\tau} \omega_i(\tau) = \alpha_i \sum_{\tau=1}^{\tau} C_i(\tau)$.

The actual number of New Deaths (Eliminated) in country *i* at day *t* is $\phi_i(t) = \lambda_i A_i(t)$. The cumulated number of Deaths in country *i* at day *t* is $\Phi_i(t) = \sum_{\tau=1}^{\tau} \phi_i(\tau) = \lambda_i \sum_{\tau=1}^{\tau} A_i(\tau)$.

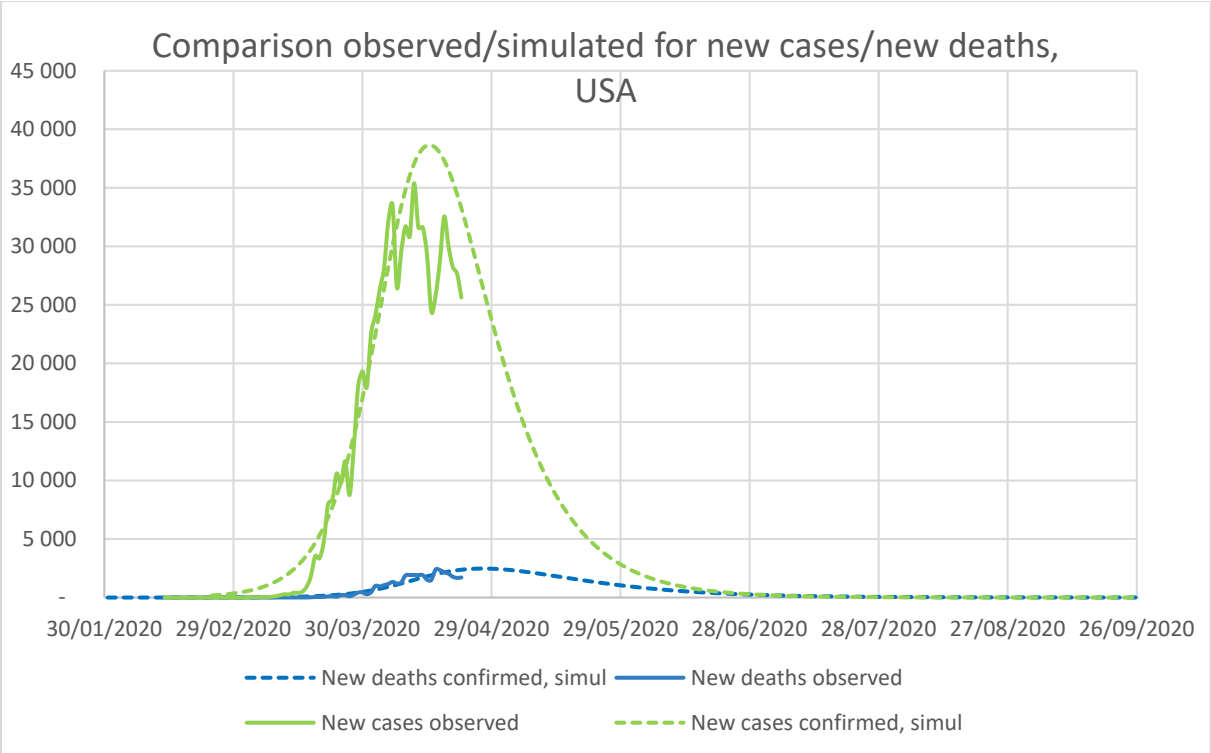
The WHO statistics under-estimate the actual numbers of cases in a way, which varies both between countries and across time. Let $\varsigma_i(t)$ denote the ratio between the observed number of cases, denoted by $\omega_i^o(t)$ and the actual number of cases $\omega_i(t)$. Therefore, $\varsigma_i(t) = \dfrac{\omega_i^o(t)}{\omega_i(t)}$, or $\omega_i^o(t) = \varsigma_i(t)\omega_i(t)$ and the observed cumulated number of cases is given by $\Omega_i^o(t) = \sum_{\tau=1}^{t} \varsigma_i(\tau)\omega_i(\tau)$.

Similarly, the WHO statistics under-estimate (or over-estimate in the case of Belgium) the actual numbers of deaths in a way which varies both between countries and across time. Let $\xi_i(t)$ denote the ratio between the observed number of cases, denoted by $\phi_i^o(t)$ and the actual number of cases $\phi_i(t)$. Therefore, $\xi_i(t) = \frac{\phi_i^o(t)}{\phi_i(t)}$, or $\phi_i^o(t) = \xi_i(t)\phi_i(t)$ and the observed cumulated number of cases is given by $\Phi_i^o(t) = \sum_{\tau=1}^{t} \xi_i(\tau)\phi_i(\tau)$.

We deduce the dynamics of observed cases and deaths from the system above.

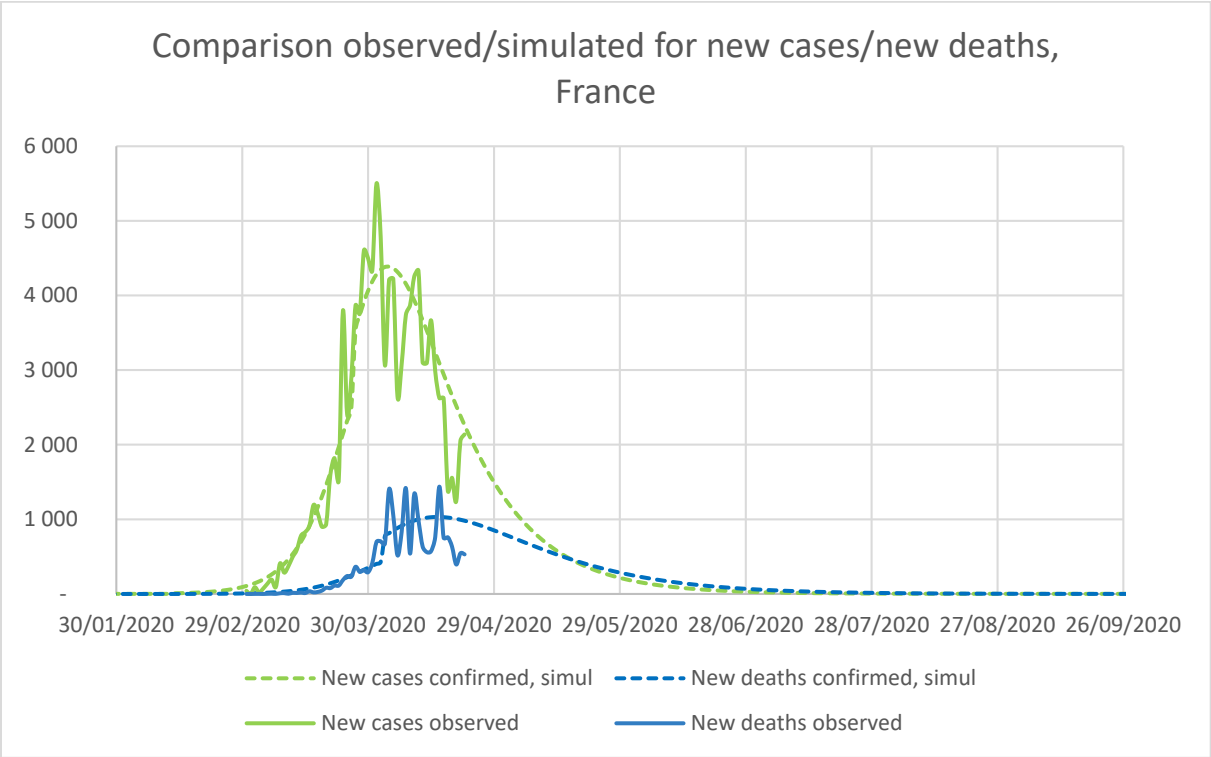
$$\begin{cases} \omega_i^o(t) = \varsigma_i(t)\alpha_i C_i(t) \\ \phi_i^o(t) = \xi_i(t)\lambda_i A_i(t) \\ \Omega_i^o(t) = \sum_{\tau=1}^{t} \omega_i^o(t) \\ \Phi_i^o(t) = \sum_{\tau=1}^{t} \phi_i^o(t) \end{cases}$$

3.1 Calibration of the model in USA



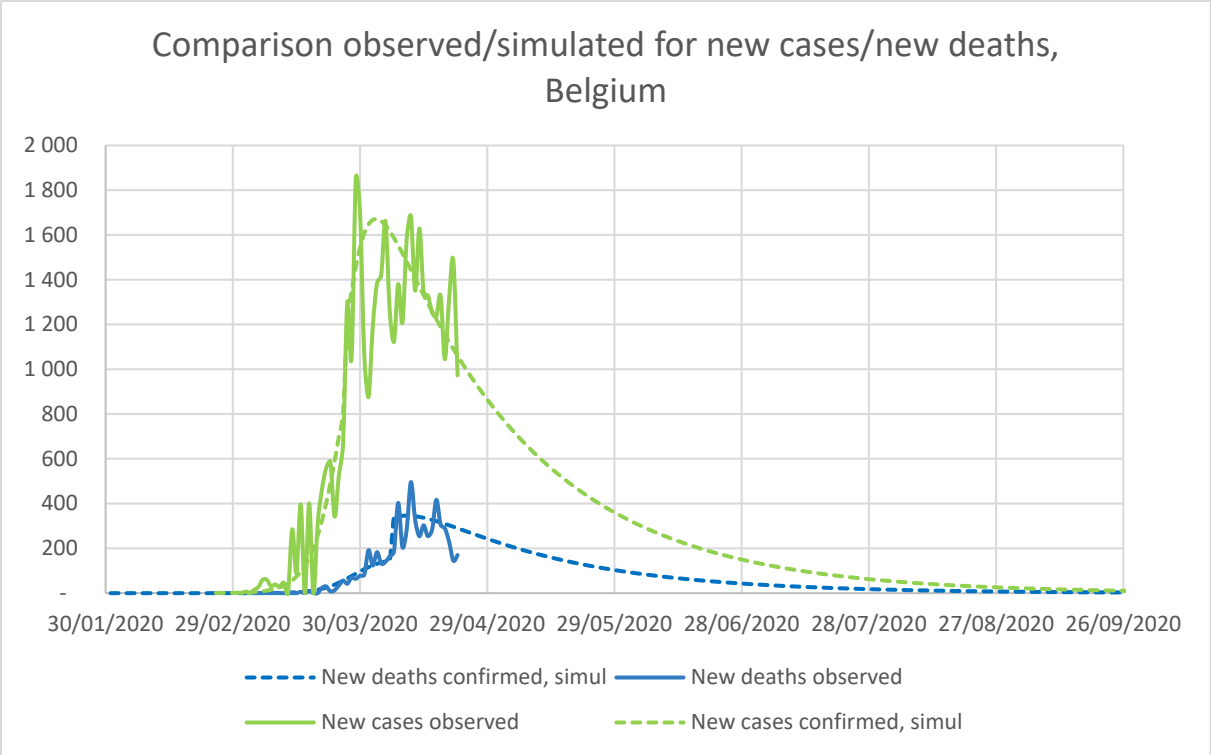
Ancillary parameters: 50% cases observed, 50% deaths observed, 2 carriers imported per 100 000 inhabitants on February, 12 (starting day of simulation).

3.2 Calibration of the model in France



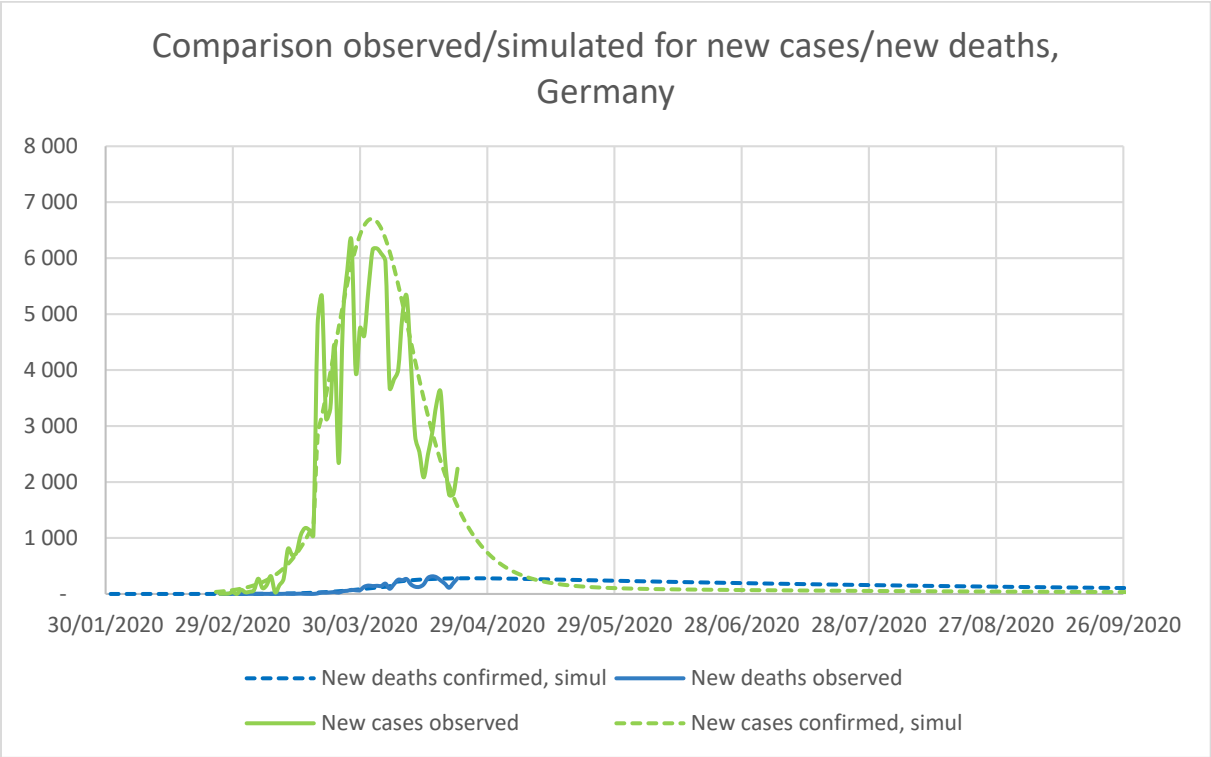
Ancillary parameters: 15% cases observed until March, 26, 20% after, 35% deaths observed until April, 3, 60% after, .8 carrier imported per 100 000 inhabitants on February, 12 (starting day of simulation).

3.3 Calibration of the model in Belgium



Ancillary parameters: 15% cases observed until March, 26, 20% after; 50% deaths observed until April, 7, 115% after (some deaths in nursing homes wrongly assumed caused by covid), 1 carrier imported per 100 000 inhabitants on February, 12 (starting day of simulation).

3.4 Calibration of the model in Germany



Ancillary parameters: 30% cases observed until March, 19, 60% after; 60% deaths observed until April, 7, 90% after, 2 carriers imported per 100 000 inhabitants on February, 12 (starting day of simulation).

3.5 Comparison of calibrated parameters and aggregate statistics across countries

Table 2 reports provisional estimates. Since the data available cover a relatively small period of time (February, 12 to April, 22), the value of estimates remain approximate, although the quality of the fit can be evaluated from the country-specific figures above.

	# daily transmissions /carrier	Transition probability per day from Carrier to Affected	Transition probability per day from Affected to Removed	Transition probability per day from Affected to Eliminated	Transition probability per day from Carrier to Removed
Country	beta	Alpha	Gamma	lambda	mu
France	0.242	0.001	0.05	0.00627	0.0745
Belgium	0.339	0.001	0.139	0.00595	0.0278
Germany	0.387	0.001	0.00602	0.00126	0.2091
USA	0.252	0.001	0.05	0.00501	0.1027

Table 2: Estimated parameters in 4 countries

Country	Cumulated # days sick / inhabitant	max % inhabitants simultaneously sick	final % population immuned	Final death rate
France	0.226	0.421%	95.836%	0.142%
Belgium	0.240	0.440%	99.857%	0.143%
Germany	0.541	0.296%	82.715%	0.068%
USA	0.157	0.298%	89.54%	0.079%

Table 3: Aggregate results in 4 countries

4 Conclusions

The SCARE model, introduced in this paper provides a generalization of the SIR model. It allows a better statistical fit than the SIR model to the available data on Covid-19. In addition, SCARE allows a more careful simulation of the effects of policies on the number of infections and deaths than the SIR model. The first empirical tests for USA, France, Belgium and Germany are promising.

As for now, the epidemiological parameters are assumed to be constant over time. In an extended version, we will consider parameters which vary across time.

In particular, the value of the β contagion parameter will depend on the policy implemented in each country. The SCARE model will be the core of an extended model that not only allows to study the effect of social distance policies as Ferguson et al (2020) but also integrates the economic effects of different confinement policies.

References

N. M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. C. Okell, S. van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. G. T. Walker, C. Walters, P. Winskill, C. Whittaker, C. A. Donnelly, S. Riley, A. C. Ghani. (2020) Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. (Report 9) *Imperial College Response Team*, http://hdl.handle.net/10044/1/77482.

W. Kermack, A. McKendrick, A contribution to the mathematical theory of epidemics. *Proc. R. Soc. A* **115**, 700-721 (1927)

C. Lefèvre, M. Simon (2016). SIR epidemics with stages of infection. *Advances in Applied Probability*, 48(3), 768-791.

WHO (2020) daily reports are available here:

https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/