# CESifo WORKING PAPERS

# Tractable Epidemiological Models for Economic Analysis

*Martín Gonzalez-Eiras, Dirk Niepelt*

CES**ifo**

# Tractable Epidemiological Models
# for Economic Analysis

## Abstract

We contrast the canonical epidemiological SIR model due to Kermack and McKendrick (1927) with more tractable alternatives that offer similar degrees of "realism" and exibility. We provide results connecting the different models which can be exploited for calibration purposes. We use the expected spread of COVID-19 in the United States to exemplify our results.

*Martín Gonzalez-Eiras*
*University of Copenhagen*
*Oster Farimagsgade 5*
*Denmark – 1353 Copenhagen K*
*mge@alum.mit.edu*

*Dirk Niepelt*
*Study Center Gerzensee*
*P.O. Box 21*
*Switzerland – 3115 Gerzensee*
*dirk.niepelt@szgerzensee.ch*
*www.niepelt.ch*

# 1  Introduction

Within weeks, the COVID-19 shock and its broad societal repercussions have not only gripped the personal and professional lives of billions of people but also transformed the economics profession: Epidemiology has become of general interest for economists. As a consequence, the profession has started to intensively analyze the interactions between epidemiological dynamics and economic choices and outcomes. By far the most prominent epidemiological framework underlying this recent wave of work has been the "SIR" framework in the tradition of the classical article by Kermack and McKendrick (1927).[1]

As we point out in this paper the Kermack and McKendrick (1927) framework offers advantages and disadvantages when compared with an alternative framework due to Bailey (1975). For researchers interested in the intersection of epidemiology and economics the choice between the two specifications thus is far from obvious. Against this background, we offer some alternatives.

First, we propose a hybrid model that combines key advantages of the Kermack and McKendrick (1927) and Bailey (1975) specifications. Second, we propose a streamlined model with fewer state variables that captures many of the essential aspects of SIR models but offers more tractability and similar degrees of "realism" and flexibility. Finally, we provide some propositions that help to connect the different models and to exploit the connections for calibration purposes.

The remainder of the paper is structured as follows. In sections 2 and 3 we present the SIR models due to Kermack and McKendrick (1927) and Bailey (1975), respectively. In sections 4 and 5 we propose the hybrid model and the more tractable framework, respectively. Section 6 concludes. Throughout, we use the expected spread of COVID-19 in the United States to exemplify our results.

# 2  Canonical SIR Model

The canonical SIR model due to Kermack and McKendrick (1927) specifies laws of motion in continuous time for the population shares of three groups that differ with respect to their health status. The three groups are the "susceptible," the "infected" or "infectives," and the "removed," and their respective population shares at time $t \geq 0$ are denoted by $x(t)$, $y(t)$, and $z(t)$, respectively, where $x(t) + y(t) + z(t) = 1$.[2] We normalize the mass of the total population at time $t = 0$ to unity. Accordingly, the population shares $x(t)$, $y(t)$, and $z(t)$ correspond to the mass of susceptible, infected, and removed persons.

At time $t = 0$ the population consists of $x(0)$ susceptible persons and a few infected persons, $y(0)$. There are no removed persons at this time, $z(0) = 0$. In each instant after time $t = 0$, infected persons transmit the disease to members of the susceptible group and

---

[1]Early papers in the recent literature that merges epidemiology and economics include Atkeson (2020) and Eichenbaum et al. (2020) but within days, the number of papers has grown substantially, with many contributions being published in CEPR's *Covid Economics: Vetted and Real-Time Papers* series.

[2]We follow the notation introduced by Kermack and McKendrick (1927).

a share of the infected either dies or recovers and develops immunity. Formally,

$$\dot{x}(t) = -b(t)x(t)y(t), \tag{1}$$
$$\dot{y}(t) = -\dot{x}(t) - (c^d + c^r)y(t), \tag{2}$$
$$\dot{z}(t) = (c^d + c^r)y(t). \tag{3}$$

Here, $b(t)$ denotes a possibly time-varying infection rate. It reflects epidemiological factors and, in economic analyses, typically also household, firm, or government choices. The extent to which susceptible persons are infected depends on their number, $x(t)$; the infection rate, $b(t)$; and the population share of infected persons. The number of infected persons increases one-to-one with the susceptible persons that get infected, while a share $c \equiv c^d + c^r$ of the infected population dies or recovers; the coefficients $c^d$ and $c^r$ parameterize the flow into death and recovery, respectively.

Consider the case where $b(t)$ is constant at value $b$. Inspection of equations (1) and (2) reveals that for $bx(0) > c$ the share of infected persons increases until it reaches a maximum when $x(t) = c/b$; thereafter, the share declines. Intuitively, when $x(0)$ falls short of $c/b$ (the "herd immunity level") then there are fewer new infections of susceptible persons than outflows from the infected pool due to recoveries and death.

We state a well-known epidemiological result (e.g., Theorem 2.1 in Hethcote, 2000):

**Proposition 1.** In the canonical SIR model with $b(t) = b > c/x(0)$,

$$x(t) + y(t) = x(0) + y(0) + \frac{c}{b} \ln \left( \frac{x(t)}{x(0)} \right).$$

Accordingly, the maximum value of $y(t)$ equals

$$x(0) + y(0) - \frac{c}{b} + \frac{c}{b} \ln \left( \frac{c/b}{x(0)} \right),$$

and the long-run share of the susceptible population, $x(\infty)$, solves the equation

$$x(\infty) = x(0) + y(0) + \frac{c}{b} \ln \left( \frac{x(\infty)}{x(0)} \right).$$

*Proof.* Dividing equation (2) by equation (1) yields

$$\frac{\dot{y}(t)}{\dot{x}(t)} = -1 + \frac{c}{b}\frac{1}{x(t)} \quad \Rightarrow \quad \dot{y}(t) = -\dot{x}(t) + \frac{c}{b}\frac{\dot{x}(t)}{x(t)}.$$

Integrating yields $y(t) = -x(t) + \ln(x(t))c/b + \text{constant}$. Since $x(0)$ and $y(0)$ are given the constant equals $x(0) + y(0) - \ln(x(0))c/b$ and the first result follows. The second result follows from the fact that $y(t)$ reaches a maximum when $x(t) = c/b$ (see above). The last result follows since $y(\infty) = 0$. □

Note than $x(\infty)$ falls short of the herd immunity level, $c/b$, unless $x(0) = c/b = x(\infty)$ and $y(0) = 0$.[3]

---

[3]Note also, from equation (2), that at the beginning of an epidemic with $x(t) \approx 1$ and $z(t) \approx 0$, $b$ approximately equals the growth rate of the number of persons who are or were infected, $\frac{\dot{y}(t)+\dot{z}(t)}{y(t)+z(t)} = b\frac{x(t)y(t)}{y(t)+z(t)} \approx b$.

**Calibration and Simulation**   We measure time in days and use the spread of COVID-19 in the United States as an example for our analysis. We associate time $t = 0$ with mid March 2020, the date around which public health authorities considered to impose restrictions. We assume that at this time, $z(0)$ equalled practically nil.

Following Atkeson (2020) and the sources cited therein we assume that the flow rate from the infected to the removed population equals $c = 1/18$, corresponding to an exponentially distributed infection duration that averages 18 days.[4]

From Russell et al. (2020), Greenstone and Nigam (2020), and the sources cited therein we infer that the inverse of the infection fatality rate, $c/c^d$, lies in the range $[100, 200]$.

To calibrate $y(0)$ we use data on COVID-19 deaths in mid March 2020 as well as information about $c^d$ and $c$. The number of deaths on March 16 equalled 23.[5] Based on equation (3) we infer that the initial share of the infected population in mid March, $y(0)$, equalled $1.8933 \cdot 10^{-4}$.[6] This compares to a reported case count of 4507, corresponding to $1.3745 \cdot 10^{-5}$ percent of the US population.[7]

Finally, to calibrate $b$ we rely on information in Ferguson et al. (2020), Greenstone and Nigam (2020), and Scherbina (2020). These authors argue, or infer from estimates, that the "basic reproduction number" $\mathcal{R}_0 = b/c$ for COVID-19 equals approximately 2.4 which implies $b = 0.1333$; that in the absence of any intervention the number of infected persons would have peaked in early June or in July; and that the share of the susceptible population would have fallen by roughly 80 percent until October 2020. Given our calibrated values for $y(0)$ and $c$ we compute the paths of $x(t)$ and $y(t)$ for $b$ values around 0.1333 and check when the predicted $y(t)$ series peaks and to what level $x(t)$ falls by early October. Table 1 summarizes the results:

| $b$ | $y(t)$ maximum | $x(1$ October$)$ |
|--------|----------------|------------------|
| 0.1000 | 13 Sept 2020 | 0.2675 |
| 0.1333 | 07 July 2020 | 0.1215 |
| 0.1500 | 20 June 2020 | 0.0844 |
| 0.2000 | 21 May 2020 | 0.0305 |

Table 1: Implications of different values for $b$ in the canonical SIR model.

---

[4]Note that $\int_0^\infty c e^{-ct} t \, dt = 1/c$.

[5]See `https://github.com/nytimes/covid-19-data/blob/master/us.csv`. Regressing the full set of March data on an exponential trend generates a similar point estimate for March 16.

[6]We have $y(0) \cdot$ (US population) = (new deaths)$/c^d$ = (new deaths)$/c \cdot c/c^d$. We use US population = 328 million, new deaths = 23, and $c/c^d = 150$.

[7]See `https://github.com/nytimes/covid-19-data/blob/master/us.csv`. The reported number corresponds to the cumulative case count but there are very few removed cases at the time. Common estimates of the extent of underreporting suggest a factor of ten, in line with our results; see, e.g., `https://www.medrxiv.org/content/10.1101/2020.03.14.20036178v2.full.pdf+html` or `https://fondazionecerm.it/wp-content/uploads/2020/03/Using-a-delay-adjusted-case-fatality-ratio-to-estimate-under-reporting-_-CMMID-Repository.pdf`.
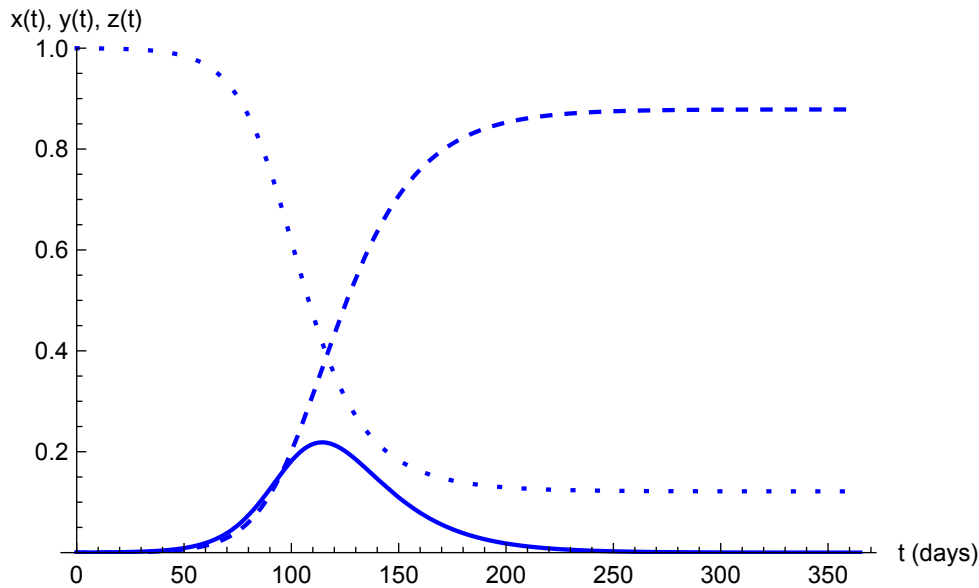
4

Figure 1: Dynamics in the canonical SIR model: $x(t)$ (dotted), $y(t)$ (solid), and $z(t)$ (dashed).

The table shows that, not surprisingly, larger values of $b$ imply a faster transition with an earlier peak of infections. In light of the different forecasts by specialists (quoted above) we choose the baseline value $b = 0.1333$. Figure 1 illustrates the dynamics given these parameter assumptions. Note that in line with the theoretical results discussed before, $y(t)$ reaches its maximum value of $\approx 0.2185$ when $x(t) = c/b \approx 0.4168$.

# 3   Modified SIR Model

One noteworthy feature of the canonical SIR model concerns scale effects in the posited process for infections: Equation (1) implies that a doubling of the population shares $x(t)$ and $y(t)$ leads to a quadrupling of new infections, $-\dot{x}(t)$, although each susceptible person faces just twice as many infected persons as before. This contradicts the fact that infection rates are largely independent of population size:

> "Naively, it might seem plausible that the population density and hence the contact rate would increase with population size, but the daily contact patterns of people are often similar in large and small communities, cities, and regions. For human diseases the contact rate seems to be only very weakly dependent on the population size. ... This result is consistent with the concept that people are infected through their daily encounters and the patterns of daily encounters are largely independent of community size within a given country" (Hethcote, 2000, p. 602).

By implication, analyses based on the canonical SIR model could over estimate the effects of policy interventions that are assumed to affect both $x(t)$ and $y(t)$, as we discussed

in Gonzalez-Eiras and Niepelt (2020a,b).[8]

A second potential problem with the canonical SIR model relates to its limited tractability.[9] While numerical solutions of the model are easy to compute they lack the transparency of closed-form solutions which are more helpful to build intuition, in particular at this stage of the literature on epidemiology and economics and even more so when economic choices are embedded in the epidemiological framework.

A modified SIR model due to Bailey (1975) resolves these issues. It only differs from the canonical SIR with respect to equation (1). In particular, Bailey's (1975)specification of the infection process scales the product of $x(t)$ and $y(t)$ in equation (1) by the total number of infected or susceptible persons; that is, equation (1) is modified to read

$$\dot{x}(t) = -b(t)x(t)\frac{y(t)}{x(t) + y(t)}, \tag{4}$$

where the expression in the denominator replaces $x(t) + y(t) + z(t)$, which equals unity. With this specification, the extent to which susceptible persons are infected depends on their number, $x(t)$; the infection rate, $b(t)$; and the *share* of the infected in the susceptible or infected population. Accordingly, a doubling of the population shares $x(t)$ and $y(t)$ leads to a doubling of new infections, not a quadrupling. We refer to equations (2)–(4) as the modified SIR model.

The system (2)–(4) can be solved as follows (Bohner et al., 2019): Let $\xi(t) \equiv x(t)/y(t)$ for $y(t) \neq 0$. We have

$$\dot{\xi}(t) = \frac{\dot{x}(t)y(t) - x(t)\dot{y}(t)}{y^2(t)} = (c - b(t))\xi(t),$$

such that

$$\xi(t) = \xi(0)e^{\int_0^t (c-b(s))ds} \quad \Leftrightarrow \quad y(t) = x(t)\kappa e^{\int_0^t (b(s)-c)ds}$$

where $\kappa \equiv y(0)/x(0)$. Substituting into equation (4) yields

$$\dot{x}(t) = -b(t)x(t)\frac{\kappa e^{\int_0^t (b(s)-c)ds}}{1 + \kappa e^{\int_0^t (b(s)-c)ds}},$$

which has the solution

$$x(t) = x(0)e^{\int_0^t \frac{-\kappa b(u)}{\kappa + e^{\int_0^u (c-b(s))ds}}du}. \tag{5}$$

Accordingly, we can solve equation (2) for

$$y(t) = y(0)e^{\int_0^t \frac{b(u)}{1+\kappa e^{\int_0^u (b(s)-c)ds}} - c \ du} \tag{6}$$

---

[8]Alvarez et al. (2020) adopt the specification (1)—in their setup a lockdown that reduces the number of susceptible and infected persons by a factor of $\alpha$ each reduces transmission risk per susceptible person by a factor of $\alpha^2$—but they discuss the assumption and indicate some doubt (p. 8). In more recent work, Acemoglu et al. (2020) analyze the implications of scale effects in a model with group specific risks.

[9]See Harko et al. (2014) for a strategy to solve the canonical model.

and equation (3) for

$$z(t) = 1 - x(t) \left(1 + \kappa e^{\int_0^t (b(s)-c)ds}\right) \text{ s.t. (5)}, \tag{7}$$

where we use the fact that the population size equals unity.

The solution simplifies when $b(t)$ is constant at value $b$. In this case equations (5)–(7) reduce to

$$x(t) = x(0)(1+\kappa)^{\frac{b}{b-c}} \left(1 + \kappa e^{(b-c)t}\right)^{-\frac{b}{b-c}}, \tag{8}$$

$$y(t) = y(0)(1+\kappa)^{\frac{b}{b-c}} \left(1 + \kappa e^{(b-c)t}\right)^{-\frac{b}{b-c}} e^{(b-c)t}, \tag{9}$$

$$z(t) = 1 - x(0)(1+\kappa)^{\frac{b}{b-c}} \left(1 + \kappa e^{(b-c)t}\right)^{-\frac{c}{b-c}}. \tag{10}$$

Moreover, we have the following result:

**Proposition 2.** Consider the modified SIR model and let $b(t) = b > c$. Then, $y(t)$ attains the maximum value

$$y^{\max} = y(t^{\max}) = \frac{b-c}{c} \left(\frac{b}{c} \frac{x(0)}{x(0) + y(0)}\right)^{-\frac{b}{b-c}} x(0)$$

at the time

$$t^{\max} = \ln \left(\frac{b-c}{c} \frac{x(0)}{y(0)}\right) / (b-c).$$

*Proof.* From equations (2) and (4), $\dot{y}(t) = 0$ implies $(x(t) + y(t))/x(t) = b/c$. Using equations (8) and (9) this implies

$$1 + \frac{y(0)}{x(0)} e^{(b-c)t} = \frac{b}{c}.$$

The result for $t^{\max}$ then follows. Substituting into equation (9) yields the result for $y^{\max}$. $\qquad\square$

**Calibration and Simulation**   The modified SIR model does not require a new calibration. Figure 2 illustrates the dynamics in the modified SIR model (in red) under the same assumptions about parameter values as before. The blue schedules representing the dynamics in the canonical SIR model are identical to the blue schedules in figure (1). We note that during the early phase of the epidemic the predicted dynamics in the two SIR models are very similar. In fact, given the uncertainty surrounding the epidemiological parameters, the two predictions effectively are indistinguishable. When the difference $b-c$ is increased the similarities become even stronger.

Similarly, infections (represented by the variable $y(t)$), which determine the stress in the health care system and thus constitute a key variable from a policy maker's perspective, peak at nearly the same time in both models, although at different levels. The two models therefore make the same prediction as to when hospital capacities are in highest demand.
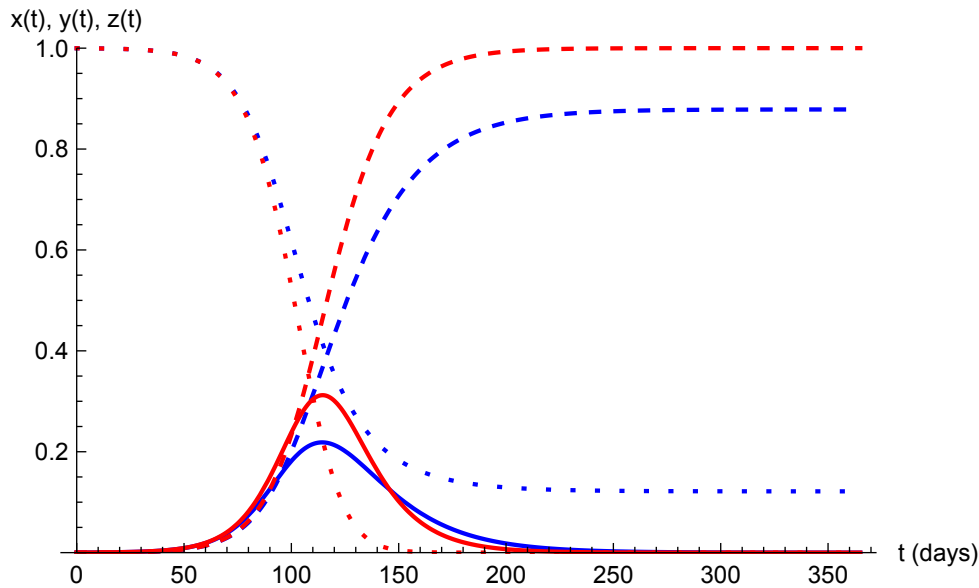
Figure 2: Dynamics in the canonical (blue) and modified (red) SIR models: $x(t)$ (dotted), $y(t)$ (solid), and $z(t)$ (dashed).

But there is also a significant difference between the two model predictions, which concerns the long-run behavior of $x(t)$ (and thus, also $z(t)$). In the canonical model, $x(\infty)$ is strictly positive (albeit small given our calibration), as discussed above, while in the modified SIR model $x(t)$ converges to zero. The two models therefore have different implications for optimal policy when the government's objective (strongly) depends on the population shares $x(t)$ and $z(t)$ in the long run, as seems plausible to expect.

Using proposition 2 we can easily determine the exact date at which infections peak as well as the peak infection level: We find that $t^{\max}$ corresponds to 7 July and $y^{\max} \approx 0.3121$. Alternatively, one may combine information about the likely peak of an infection with proposition 2 to calibrate the model.

## 4 Hybrid SIR Model

We have seen that the canonical and the modified SIR model yield similar predictions for the time profile of infections. At the same time they differ with respect to the implications for the long-run share of susceptible persons, the returns to scale in infections, and their tractability.

Using a simple device, we can form a hybrid model that combines the main advantages of the two frameworks. To this end we rely on the formal structure of the modified SIR model but assume that in addition to the susceptible, infected, and removed groups, there exists a group of "lucky" persons with population share $\bar{x}$. There are no flows into or out of the lucky group and the dynamics of $x(t)$, $y(t)$, and $z(t)$ follow the same laws of motion as in the modified SIR model. At the same time, the assumption that $\bar{x} > 0$ implies that in the long run, not everyone will eventually have been infected at some point, in line
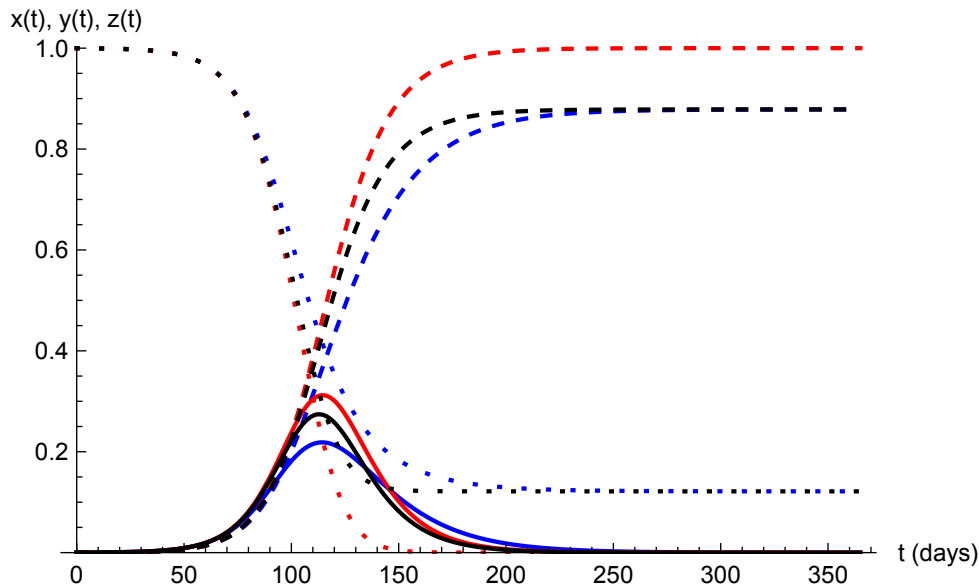
8

Figure 3: Dynamics in the canonical (blue), modified (red), and hybrid (black) SIR models: $x(t)$ (dotted, including $\bar{x}$), $y(t)$ (solid), and $z(t)$ (dashed).

with the canonical model.

The group of susceptible or lucky persons in the hybrid model thus should be interpreted as the counterpart of the susceptible group in the canonical model. The only remaining drawback of the hybrid model is that, unlike in the canonical model where $x(\infty)$ is endogenous, the long-run population share of the group of susceptible or lucky persons in the hybrid model is exogenous (equal to $\bar{x}$).

Summarizing, the hybrid model contains four groups: $\bar{x}$, $x(t)$, $y(t)$, and $z(t)$. The laws of motion for $x(t)$, $y(t)$, and $z(t)$ are identical to those in the modified SIR model (equations (2)–(4)) and $\bar{x}$ is constant over time. Moreover, conditional on $y(0)$, the remaining initial condition is given by $x(0) = 1 - \bar{x} - y(0)$.

**Calibration and Simulation** The hybrid SIR does not require a new calibration. We let $\bar{x} = 0.1215$, corresponding to $x(\infty)$ in the canonical SIR model (see proposition 1). Figure 3 illustrates the dynamics in the hybrid SIR model (in black) under the same assumptions about parameter values as before. The blue and red schedules representing the dynamics in the canonical and modified SIR models, respectively, are identical to the blue and red schedules in figures (1) and (2).

During the early phase of the epidemic the predicted dynamics in the three models are very similar. Using proposition 2 we find that peak infections in the hybrid model occur at $t^{\max}$ corresponding to 5 July, two days earlier than in the modified SIR model, and $y^{\max} \approx 0.2742$. In the long run the dynamics of the hybrid model resemble those in the canonical model, as desired.

9

# 5    Logistic Model

Theoretical research projects typically embed optimizing behavior on the part of economic agents or the government into frameworks that build on the SIR model. This further reduces tractability, even in the modified or hybrid SIR model. As a consequence, (dynamic) optimal policy problems embedded in the dynamic system (2)–(4) generally do not give rise to closed form solutions and researchers have to resort to less transparent, numerical approximations of the model solutions.

Against this background, we propose a simplified epidemiological framework—the logistic model—that represents health dynamics in terms of a single rather than two state variables.[10] This can significantly improve tractability while at the same time imposing only moderate cost in terms of reduced "realism" or flexibility (see Gonzalez-Eiras and Niepelt (2020a,b) for an application). An added benefit is that the modified setup constitutes a special case of the SIR models discussed previously.

To obtain the logistic model we simplify the SIR models along two dimensions. First, we neglect deaths and let $c^d = 0$. Importantly, this does not mean that we disregard the societal cost of deaths, to the contrary. Representing this cost does not require that we explicitly account for the deceased population; it suffices to account for infections and to associate a cost with these infections that reflects the societal cost of the number of deaths that follow from the infections. That is, rather than modeling a time series for the number of deceased as in the SIR models, the simplified setup can capture the costs due to death by a function of the flows into infection, $\dot{x}(t)$ or $\dot{y}(t)$.[11]

Second, we blur the distinction between infected and recovered persons by letting $c^r = 0$ such that infection is an absorbing state and $z(t) = 0$. That is, we assume for example that infected persons are as productive as recovered ones.[12] Again, this does not imply that we disregard the societal cost that infected persons impose on the health care system. Rather than relating this cost to a time series for the number of currently infected (soon to die or to be removed) persons, as in the SIR models, the simplified setup relates this cost to flows from the $x(t)$ to the $y(t)$ pool.

In sum, we view $x(t)$ as the "not yet infected" population and $y(t) = 1 - x(t)$ as the "infected but still productive." By letting $c = 0$ the simplified setup with constant coefficient $b$ reduces the canonical or modified SIR model to a framework where the time paths of $x(t)$ and $y(t)$ follow logistic curves. To see this, note from equations (8) or (9)

---

[10]The SIR models discussed so far feature two state variables, $x(t)$ and $y(t)$; $z(t)$ is implied by the former two.

[11]It is also easily possible to represent a situation where only the total number of deceased persons is of interest: In the SIR model the cost of the total number of deaths is some function of the number of deaths, which appears as a variable in the model and is proportional to $1 - x(\text{inf})$; in the logistic model the cost is a function of the number of deaths which does not appear as a variable in the model but is related to the total number of ever infected persons, $\bar{y}$, which does appear as a variable in the model (see below).

[12]This assumption is particularly realistic when many infected persons do not know of their health status, as appears to be the case with COVID-19.

that $c = 0$ (such that $x(t) + y(t) = 1$) and $b(t) = b$ implies

$$y(t) = y(0)(1 + \kappa)\frac{e^{bt}}{1 + \kappa e^{bt}} = \frac{y(0)}{x(0)}\frac{1}{e^{-bt} + y(0)/x(0)} = \frac{1}{1 + e^{-bt}x(0)/y(0)}.$$

In parallel to the approach taken for the hybrid model we can easily impose that the long-run share of susceptible persons is bounded away from zero, $\lim_{t\to\infty} x(t) \equiv \bar{x} > 0$, or equivalently that $\lim_{t\to\infty} y(t) \equiv \bar{y} < 1$. Specifically, writing the law of motion for $y(t)$ as

$$\dot{y}(t) = by(t)(\bar{y} - y(t))$$

such that $y(t)$ converges to $\bar{y}$ implies the solution

$$y(t) = \frac{\bar{y}}{1 + e^{-b\bar{y}t}(\bar{y}/y(0) - 1)}.$$

Accordingly, we have the following standard result:

**Proposition 3.** Consider the logistic model with $b(t) = b$ and $y(0) < \bar{y} \leq 1$. Then, $\dot{y}(t)$ reaches a maximum at

$$t = \ln\left(\frac{\bar{y} - y(0)}{y(0)}\right)/(b\bar{y}).$$

*Proof.* Solving $\ddot{y}(t) = 0$ (or $y(t) = \bar{y} - y(t)$) for $t$ yields the result. □

**Calibration and Simulation** We can use propositions 2 and 3 to calibrate the logistic model in order to match peak infection rates (or peak health care burdens) as implied by SIR models. To see this, recall that the peak infection rate in the modified or hybrid SIR model occurs at time

$$t^{\max} = \ln\left(\frac{b - c}{c}\frac{x(0)}{y(0)}\right)/(b - c).$$

It follows that the $b$ value in the logistic model, $b^{\log}$ say, that replicates peak infections in the modified or hybrid SIR models with parameter $b$ is given by

$$t = t^{\max} \quad \Leftrightarrow \quad b^{\log} = \frac{(b - c)\ln\left(\frac{\bar{y}}{y(0)} - 1\right)}{\bar{y}\ln\left(\frac{b-c}{c}\frac{x(0)}{y(0)}\right)},$$

where the $x(0)$ value refers to the hybrid SIR model (excluding "lucky" persons). Using our common parameter values $c = 1/18$, $y(0) = 1.8933 \cdot 10^{-4}$, and $\bar{x} = 0.1215 = 1 - \bar{y}$, this yields $b^{\log} = 0.0851$ (for $b = 0.1333$).

Figure 4 illustrates the dynamics in the logistic model (in green). The blue and black schedules representing the dynamics in the canonical and hybrid SIR models, respectively, are identical to the schedules in figures (1) and (3). (To improve legibility the figure does not show the paths in the modified SIR model.) Evidently, the time paths for $x(t)$, $y(t)$, and $\dot{y}(t)$ (scaled) match the corresponding paths in the canonical and hybrid SIR models very closely.

We interpret this as evidence for the usefulness of the approach suggested in this paper.
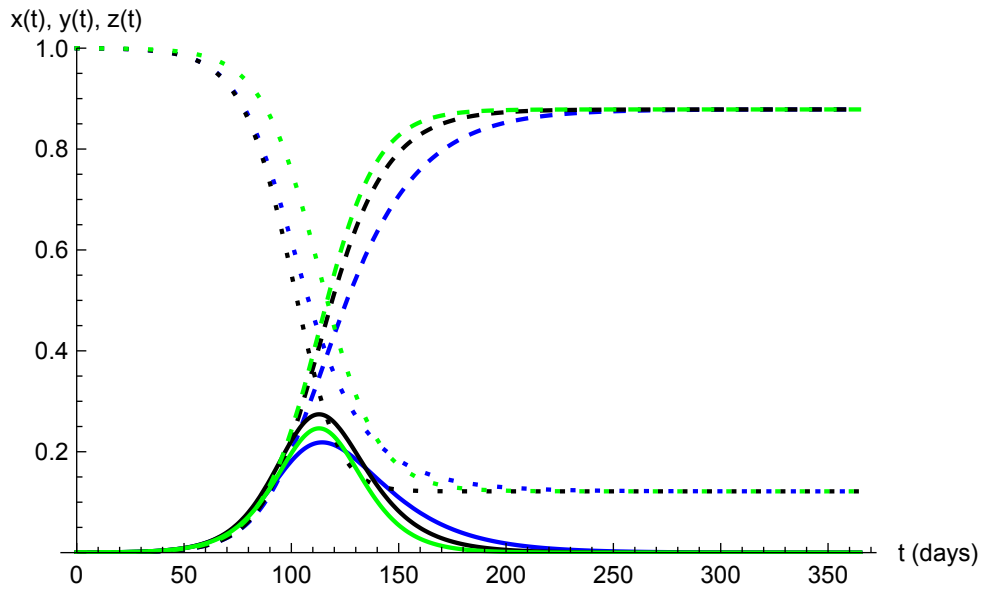
Figure 4: Dynamics in the canonical (blue) and hybrid (black) SIR models and in the logistic model (green). SIR models: $x(t)$ (dotted, including $\bar{x}$), $y(t)$ (solid), and $z(t)$ (dashed). Logistic model: $x(t)$ (dotted, including $\bar{x}$), $\dot{y}(t)$ (solid, scaled), and $y(t)$ (dashed).

# 6 Conclusion

Researchers interested in the intersection of epidemiology and economics might usefully build on variants of the classical SIR models in the tradition of Kermack and McKendrick (1927) and Bailey (1975). In this paper, we discuss two such variants, a hybrid SIR model and the logistic model. We provide some propositions that help to connect the different models and to exploit the connections for calibration purposes, and we exemplify their usefulness against the background of the ongoing COVID-19 epidemic in the United States. Without doubt, future research will continue to build on these models.

# References

Acemoglu, D., Chernozhukov, V., Werning, I. and Whinston, M. D. (2020). A multi-risk SIR model with optimally targeted lockdown, *Working Paper 27102*, NBER, Cambridge, Massachusetts.

Alvarez, F., Argente, D. and Lippi, F. (2020). A simple planning problem for COVID-19 lockdown, *Working Paper 26981*, NBER, Cambridge, Massachusetts.

Atkeson, A. (2020). What will be the economic impact of COVID-19 in the US? rough estimates of disease scenarios, *Working Paper 26867*, NBER, Cambridge, Massachusetts.

Bailey, N. T. J. (1975). *The Mathematical Theory of Infectious Diseases and its Applications*, 2 edn, Hafner Press, New York.

Bohner, M., Streipert, S. and Torres, D. F. M. (2019). Exact solution to a dynamic SIR model, *Nonlinear Analysis: Hybrid Systems* **32**: 228–238.

Eichenbaum, M. S., Rebelo, S. and Trabandt, M. (2020). The macroeconomics of epidemics, *Working Paper 26882*, NBER, Cambridge, Massachusetts.

Ferguson, N. M., Laydon, D., Nedjati-Gilani, G. and collaborators (2020). Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand, *Report 9*, Imperial College, London.

Gonzalez-Eiras, M. and Niepelt, D. (2020a). On the optimal 'lockdown' during an epidemic, *Covid Economics* **7**: 68–87.

Gonzalez-Eiras, M. and Niepelt, D. (2020b). On the optimal 'lockdown' during an epidemic, *Discussion Paper 14612*, CEPR.

Greenstone, M. and Nigam, V. (2020). Does social distancing matter?, *Covid Economics* **7**: 1–23.

Harko, T., Lobo, F. S. and Mak, M. (2014). Exact analytical solutions of the susceptible-infected-recovered (SIR) epidemic model and of the SIR model with equal death and birth rates, *Applied Mathematics and Computation* **236**: 184–194.

Hethcote, H. W. (2000). The mathematics of infectious diseases, *SIAM Review* **42**(4): 599–653.

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

Russell, T. W., Hellewell, J., Jarvis, C. I., van Zandvoort, K., Abbott, S., Ratnayake, R., CMMID COVID-19 working group, Flasche, S., Eggo, R. M. and Kucharski, A. J. (2020). Estimating the infection and case fatality ratio for COVID-19 using age-adjusted data from the outbreak on the Diamond Princess cruise ship, *Eurosurveillance* **25**(12).

Scherbina, A. (2020). Determining the optimal duration of the COVID-19 suppression policy: A cost-benefit analysis, *Economics Working Paper 2020-03*, American Enterprise Institute, Washington.