Addressing the Herd Immunity Paradox Using Symmetry, Convexity Adjustments and Bond Prices

Peter Cotton[*]

Apr 26, 2020

Abstract

In constant parameter compartmental models an early onset of herd immunity is at odds with estimates of R values from early stage growth. This paper utilizes a result from the theory of interest rate modeling, namely a bond pricing formula of Vasicek, and an approach inspired by a foundational result in statistics, de Finetti's Theorem, to show how the modeling discrepancy can be explained. Moreover the difference between predictions of classic constant parameter epidemiological models and those with variation and stochastic evolution can be reduced to simple "convexity" formulas. A novel feature of this approach is that we do not attempt to locate a true model but only a model that is equivalent after permutations. Convexity adjustments can also be used for cross sectional comparisons and we derive easy to use rules of thumb for estimating threshold infection level in one region given knowledge of threshold infection in another.

1 An informal overview

This paper introduces fixed income modeling technique and a tiny amount of group theory to epidemiological modeling, and is inspired by a theorem in statistics which has not, to our knowledge, found application in this field. We give an brief overview of the approach, deferring formality to Section 2 onwards.

1.1 A meta-population model

Our ansatz is a collection of non-interacting regions, termed cities for the sake of exposition. Within each city we imagine there are strongly interacting sub-populations called neighborhoods. We assume each city shares the same stochastic generative model. The path taken by each city's epidemic will be different, but we assume that mean quantities for a stochastic model for a city can be viewed as equivalent to an aggregate model for many cities (as if all paths are played out at once).

1.2 Symmetric measurements

Model observables (often called functionals in stochastic modeling) are expectations with respect to the model for one city. For example:

- The overall (mean) early stage growth rate

- The mean across cities of the number of susceptible people at peak infection - where peak infection time is defined city by city.

More generally we might consider only functionals that are symmetric with respect to interchange of neighborhoods. Note that a sum, viewed as a multivariate function, is symmetric with respect to its arguments. We lean heavily on the ramifications of this seemingly banal fact.

1.3 Exchangeable models

We can exploit this symmetry by noting that a model itself (say the one that is the true representation of nature) can also be permuted leaving the functional (such as the expectation of a sum) unchanged. Moreover, taking an average of *all* model permutations also leaves the result unchanged. But the average of all permutations of a model is an *exchangeable model*.

1.4   Equivalent models

We consider two models providing the same mean observable quantities to be equivalent, and to belong to the same equivalence class. An equivalence classes includes an orbit of the symmetric group, the name given to the set of all permutations of a given model. The motivating observation for our approach is expressed in the following equalities, where we use $E[i(t=3)]$, which is to say the number of infected people at $t=3$, as an example of a functional:

$$\overbrace{E_{\phi^*}\left[i(t=3)\right]}^{nature's\ true\ model} = \overbrace{E_{\psi=U\phi^*}\left[i(t=3)\right]}^{nature's\ symmetrized\ model} = \overbrace{E_{\phi\in O(\phi^*)}\left[i(t=3)\right]}^{any\ model\ in\ same\ class}$$

This straightforward observation about reported quantities will be made more formally in the next section to avoid any ambiguity. It suggests that we may be well served by characterizing the orbit generated by a symmetry group which passes through nature's true model - perhaps by finding a representative of each orbit.

1.5   Exchangeable models

Moving to the next stage of the argument, we note that there is precisely one exchangeable model on each orbit. Thus traversing the set of exchangeable models might be a way to quickly examine equivalence classes of models rather than investigating the much larger space of models per se.

1.6   Mixtures of models and de Finetti's Theorem

Moreover, there is a famous decomposition of exchangeable models into independent, identically distributed models. Our idea is that *many* exhangeable models, in some sense, can be represented in terms of mixtures of independent identically distributed models provided that there is strong depedence between neighbourhoods - something we certainly expect for an epidemic.

De Finetti's Theorem is foundational to Bayesian statistics but not to our knowledge utilized in epidemiology. The theorem states that an infinite sequence of *finitely exchangeable* binary random variables can be represented as a mixture of models in which the variables are independent and identically distributed. The exchangeablility criterion means that if we choose any two finite sequences of the variables of equal length then their joint distributions are exactly the same.

In analogy to de Finetti's decomposition we attempt to span the space of exchangeable models (and thereby the equivalence classes of all models) by taking mixtures of

models in which each neighborhood evolves independently. There is no *guarantee* that we can find every relevant exchangeable model this way, since de Finetti's Theorem applies to infinite sequences (and versions that apply to finite sequences usually require signed mixtures).

1.7 Convexity (variation) adjustments

However we are *likely* to quickly understand *many* exchangeable models. The reason is that mixtures of models in which neighbourhoods are independent and identically distributed are very tractable. There are two types of calculation that we present here. One type of mixing varies parameters in deterministic compartmental models. The other utilizes stochastic evolution of infection rate, which can also be interpreted as variation.

As examples of convexity effects arising due to mixing of models with different parameters we make the connection to moment generating functions (relevant to early stage growth) and harmonic means (relevant to peak infection). Both calculations reveal the effect of variation relative to a simpler model in which there is no variation. Calculating the implications of a model with variation as a perturbation of an oft-used constant parameter model (we use SIR as an example) makes it easy to understand the first order impact.

This is also true for variation modeled as paths of a stochastic process. For example a crude but useful rule of thumb to account for stochastic evolution of infection rate presented in Section 4 is:

$$susceptible\ multiplier = 1 + \frac{relative\ variation}{R}$$

This adjustment might help us understand peak infection rate.

1.8 Mean reverting stochastic infection rate

A stochastic variant of the SIR compartmental model in which infections rates follow an Ornstein-Uhlenbeck process (random walk with drift back to a mean reversion level). This choice is well motivated both by empirical epidemiological literature and a physical model for repeat contacts.

1.9 Use of existing results from interest rate modeling

This paper also introduces the use of fixed income modeling techniques in epidemiology. The connection arises because during early stage growth the susceptible population is

close to unity. Growth in the number of infected people is directly analogous to growth of a savings account. For this reason we are spared the difficulty of computing early stage growth adjustments due to stochastic infection rate. Instead, we show that these can be identified with an existing bond pricing formula originally derived by Vasicek in what is considered a landmark paper in interest rate modeling [22]. This connection is formalized in Section 4 where we use negative interest rates to compute mean early stage growth.

1.10 Plausibility of empirical variation

Finally, there is a part of our argument that is mostly implicit - the plausibility of variation. This paper merely assumes variation in infection rate is likely without adding to the case (rather, we indirectly add to the evidence for variation insofar as we show it can explain peak infection rates relative to early stage growth, and we supply formulas for inferring variation).

The empirical literature seems supportive, though a survey is beyond the scope. It is widely believed that infection rate (β as we shall refer to it) exhibits variation across various dimensions because it commingles a variety of factors driving the number of people who will be infected in unit time. Some relate to the physics and biology of transmission, others to the behavior of humans.

For example infection rate as defined in a compartmental model may be influenced by patterns of socializing that differ from place to place [7]. Variation in transmission rates may be driven by culture. A reasonable hypothesis is that household size and mixing of generations therein may be an important dynamic. Economic variation may be a driver also. Some may be better placed to afford protective measures, or voluntarily reduce hazardous work.

At the individual level age [29] dyspnea [26] cardiovascular disease [4], pregnancy [28], [30] and other factors may influence outcomes and aggregate transmission likelihood. Another intriguing possibility is that existing vaccinations may be associated with dramatically lower incidence [13]. If true, those with the vaccinations and those without will constitute two very different groups with two very different turning points.

Environment is important. If airborne transmission [19] is a key factor, or if supply of ultraviolet light [14] is important, these factors will play a bigger role in some parts of a city than others - even in one apartment versus another. Humidity [24] has been studied, and a recent study of SARS CoV-2 in Iran suggests that wind speed may reduce infection - in addition to other climatological factors [2] such as temperature, which is said to drive down infection rates by three percent for every one degree increase [27].

Population density is another prime culprit because of its inherent plausibility and dramatic variation. Additional references are provided in Section 3.4.1.

1.11   Plausibility of variation in mixtures of models

A slightly more subtle remark is that variation in the mixture models might be expected to exceed what seems plausible purely from the standpoint of biology, environment or behaviour. That is because, as noted, the mixture model is not intended to be realistic but rather, *equivalent* to a realistic model. The mixing commingles variation and dependence, and as also noted we expect strong positive dependence between neighbourhoods in nature's model.

1.12   Summary

We take an pragmatic approach to modeling epidemics subject to variation and random evolution. We show that modeled quantities can be computed as corrections to those that would be computed in deterministic homogeneous counterparts. We appeal to de Finetti's Theorem to justify the surprising generality of this strategy, despite the fact that the models we consider are simple mixtures of models - each one of which might not, in isolation, be considered plausible at the fine scale of neighbourhoods. This strategy is now presented with more precision.
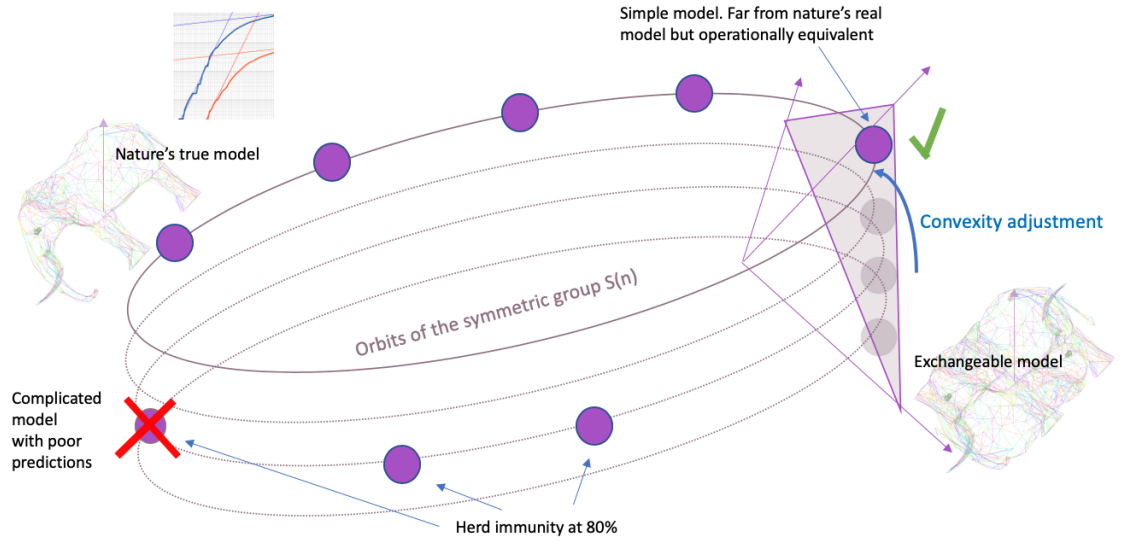
Figure 1: Outline of the strategy adopted in this paper. When only symmetric function-als of a model are required, we are motivated to consider representatives from orbits of the symmetric group only, rather than the full model space. Simple mixtures of models in which neighborhoods are independent might be used to quickly traverse a collection of exchangeable models, one of which might provide the same answers as nature's true model - a model we are unlikely to ever locate. Convexity adjustments relate stochastic model fucntionals to their deterministic counterparts, allowing for a potentially rapid resolution of modeling conundrums precipitated by early onset heard immunity.

2   An approach inspired by de Finetti's Theorem

Our strategy is represented stylistically in Figure 5 and is hopefully clear in broad brush terms from the introduction. Here we go into more detail to avoid ambiguity, though the reader familiar with exchangeability and finite versions of de Finetti's Theorem might safely skim parts of this Section.

2.1   Invariant functionals

Consider the following questions that might be asked of the model

1. When will peak infection occur?

2. What is the distribution of case counts after seven weeks?

3. What is the mean growth rate in the first three weeks?

4. How many people will remain susceptible when infection peaks?

5. How many cases will there be in the limit $t \to \infty$?

Next imagine a city (more generally some population) is divided up in some manner (we use the term neighbourhoods). We imagine some partition $P_1, \ldots, P_n$ and some quantity of interest $X_1, \ldots, X_n$ measured on each. We suppose, however, that only aggregates such as $X = X_1 + \ldots X_n$ matter for decision making purposes. More generally any symmetric function can be allowed, such as the sum of the squares of infections in each neighborhood.

2.2 Symmetric group

We formalize this notion with the symmetric group of permutations. Under consideration is a space of models (joint distributions) on n variables. There is an operation on this collection of models induced by permutations π in the symmetric group $S(n)$ acting on $S = \{1, \ldots, n\}$. Namely if ϕ is any joint distribution over $X = (X_1, \ldots, X_n)$ then $\pi \cdot \phi$ is the function that when applied to $x = (x_1, \ldots, x_n)$ yields

$$
\begin{aligned}
(\pi \cdot \phi)(x_1, \ldots, x_n) &= \phi(\pi(x)) \\
&= \phi\left(x_{\pi(1)}, \ldots x_{\pi(n)}\right)
\end{aligned}
$$

As is well appreciated these actions form a group. If π' is another permutation then applying the definition twice

$$
\begin{aligned}
\pi'((\pi \cdot \phi))(x_1, \ldots, x_n) &= (\pi \cdot \phi)(\pi'(x)) \\
&= \phi(\pi(\pi'(x))) \\
&= \phi((\pi' \circ \pi)(x)) \\
&= ((\pi' \circ \pi)\phi)(x_1, \ldots, x_n)
\end{aligned}
$$

2.3 Exchangeable models as orbit representatives

Let U denote the symmetrizing operator

$$
U := \frac{1}{n} \sum_{\pi \in S(n)} \pi
$$

that averages any function over all possible permutations.[1] This operator remains unchanged when composed with any permutation since

$$\sum_{\pi \in S(n)} \pi = \sum_{\pi \in S(n)} \pi \circ \pi'$$

given that the right hand side is a reordering of the sum on the left. Under the action of U any model ϕ is mapped to an exchangeable model $U(\phi)$ meaning that it is left unchanged by the action of all elements of the symmetric group. Furthermore if ψ_1 and ψ_2 are both exchangeable and $\psi_1 = \pi \psi_2$ for any permutation π then

$$\psi_1 = U\psi_1 = U\pi\psi_2 = \sum_{\pi' \in S(n)} \pi'\pi\psi_2 = U\psi_2 = \psi_2$$

showing that there is one and only one exchangeable model per orbit.[2] These orbits, such as

$$O(\phi) := \{\pi(\phi)\}_{\pi \in S(n)}$$

are for us a useful decomposition of the set of possible models, and the considerations above show that the set of exchangeable models traverse the collection of all orbits.

2.4 Proxy models

Next we formalize the notion that an important planning question for an epidemic might be invariant to permutations. Decision making is said to be driven by functionals of the model. Denoting the space of all models by $\mathcal{M}$, any functional

$$f : \mathcal{M} \to \mathcal{R}$$

on models (joint distributions) can also be acted on by a permutation π.[3] This action is

$$\pi(f)(\phi) := f\left(\pi^{-1}(\phi)\right)$$

---

[1]The notation U for the symmetrizing operator comes from U-statistics.

[2]In the second to last equality we are using the fact that every element of the symmetric group can be considered a permutation of the group itself, in turn because all elements are invertible.

[3]We follow a convention of defining this action in a contragredient manner, but this is choice is not essential to the argument. The functionals need not be scalar valued but in practice we can take them on one at a time.

and we can define a symmetrizing operator, also denoted U, on this space. It may be suggestive to write the application of the functional in bra-ket notation:

$$< \pi f, \phi > = < f, \pi^{-1}\phi >$$

When it is the case that f is a symmetric functional, meaning $U(f) = f$, we have

$$< f, \phi > = < Uf, \phi > = < f, U\phi > \tag{1}$$

where we appeal to the fact that the sum comprising U is left unchanged when we replace every permutation with its inverse, since this amounts to a mere reordering of terms. Equation 1 is the key to our suggested approach, for if we now suppose nature has a true model ϕ^* and we observe $f(\phi^*)$, we also know this value is shard by the same functional applied to a symmetrized version of nature's model.

$$f(\phi^*) = f(U\phi^*)$$

On the surface this is perhaps a controversial way to solve the challenge of how to create an accurate model that is close to reality (i.e. by not really trying). But seemingly what matters is not proximity to truth but proximity to the orbit of the truth. A limitation is that this approach leaves partly unaddressed the challenge of assessing the reasonableness of a *choice of orbit* (beyond its ability to match empirical data) a task not necessarily made obvious by any given representative of that orbit.

2.5   A pseudo-basis for exchangeable models

However, we know that the exchangeable models constitute a traversal set. This particular traversal set may bring an additional advantage provided by, or at least inspired by, de Finetti's Theorem.

In Section 4 we will consider a collection of stochastic infection rate models. However that is not a requirement and the point can be made by considering the more familiar constant parameter SIR model [1]. Suppose we allow β to vary with each realization and leave γ fixed for simplicity.[4]

To avoid confusion between average quantities and varying we shall separate out the

---

[4]If using deterministic evolution seems out of spirit with de Finetti's Theorem, the reader may decide to add some simple perturbation of this model, for instance through the addition of a measurement error distribution on instrumented quantities $i(t)$, $s(t)$ and $r(t)$ - or perhaps a more elaborate scheme involving delays in measurement.

variation as:

$$\beta(\omega) = \bar{\beta}\eta(\omega)$$

where $\bar{\beta}$ is a constant rate of infection and $\eta(\omega)$ is a function defined on the space Ω over which variation takes place. Assume the population density is a measure on the same space and we have both

$$E_\rho[\eta] = \int_\Omega \eta d\rho = \int_\Omega \eta(\omega)\rho(\omega)d\omega = 1$$

and

$$\rho(\Omega) = \int_\Omega \rho(\omega)d\omega = 1$$

where $\rho(\omega)$ is population density.[5] Then $\bar{\beta}$ can be interpreted as a population mean infection rate and $\eta(\omega)$ is the variation of the same.

If this model is applied to a neighborhood P_k and at the same time applied to all other neighbourhoods independently, the city model will be exchangeable. In fact the city model will be iid with respect to neighborhoods. Let us denote such a model as

$$m(\cdot; \omega) = \Pi_{k=1..n}m_k(\cdot; \omega)$$

The mixture model satisfies

$$
\begin{aligned}
\frac{\partial}{\partial t}i(t) &= -\eta(\omega)\bar{\beta}i(t)s(t) \\
\frac{\partial}{\partial t}i(t) &= \eta(\omega)\bar{\beta}i(t)s(t) - \gamma i(t) \\
\frac{\partial}{\partial t}r(t) &= \eta(\omega)\gamma i(t)
\end{aligned}
\tag{2}
$$

once ω is chosen. Assume now that η takes on the law of a known distribution - preferably one with closed form moment generating function or harmonic mean for reasons that will become clear. We allow θ to summarize parameters controlling η. For example η might be a rescaled Beta distribution and $\theta = (\alpha, \beta)$, overloading the use of β just for a moment. Denote a mixing of models induced by η's law as

$$m_\theta = \int m(\cdot, \omega)d\rho \tag{3}$$

This model is a mixture of independent, identically distributed models. The integration

---

[5]The space need not be continuous. For example Ω_1 and Ω_2 may represent disjoint sets populated by 30 percent and 70 percent of the population respectively. There is no loss of generality since $\bar{\beta}$ can be rescaled.

over η will induce dependence between the neighborhoods.

The question arises as to whether there is a θ for which m_θ can serve as a proxy for the truth. As per our discussion, this will be true if nature's model lies in the same orbit and, since m_θ is exchangeable, this also implies that

$$m_\theta = U(\phi^*)$$

for nature's true model ϕ^*.

We provide no guarantee that every exchangeable model can be represented in this fashion (unless ρ is allowed to be a signed measure, which some readers may wish to consider) but as θ varies we may nonetheless traverse a reasonably representative collection of orbits. If the orbit of nature's true model exhibits strong positive correlation between neighborhoods - which in the case of an epidemic it almost certainly does - our chances that it exists on the orbit $O(m_\theta)$ for some choice of θ go up substantially.

3 Convexity adjustments

We now come to the chief advantage of using mixtures of simple models. Predictions of the mixture model can often be written as multiples of those for deterministic models. The multiplier captures the impact of Jensen's Inequality, and can often be expressed in closed form in which case it is particularly easy to invert. Our approach makes it difficult to create very badly mispecified models.

As an aside, Jensen's Inequality has cropped up in numerous biological contexts. The effect of daily fluctuations in temperature on Rhodnius prolixus, a vector of Chagas disease, was studied by Roalandi and Schilman [20]. A Jensen's Inequality effect was noted in regard to prevalence of sigma virus and host Drosophila melanogaster in Georgia peach stands [25]. Closer to our interest, the nature of bias introduced in cohort-based models for disease progression was considered by Elbasha and Chhatwal [8].

3.1 Peak infection

We continue with the same notation as in Section 2. Suppose in a given realization ω peak infection occurs at time t^*. At this moment the proportion of the population yet susceptible is $s(t^*)$. In the classic SIR model we expect

$$s(t^*) = \gamma/\beta$$

but with a mixture model this quantity gets scaled by a quantity we know exceeds unity (by Jensen's Inequality).[6] We have instead

$$s(t^*) \approx E\left[\gamma/\beta\right] = \gamma/\bar{\beta}E\left[\frac{1}{\eta}\right]$$

The ratio of the mixed versus unmixed model with $\beta = \bar{\beta}$ is a multiplicative convexity

---

[6]For the benefit of those not familiar with epidemiological arguments for peak infection: fix two people Irene and Sam, respectively infectious and susceptible. Let β denote the instantaneous probability (per unit time) that Irena both interacts with, and also infects, Sam. Let γ denote the instantaneous probability that Irene ceases to be contagious. In more general compartmental models γ may comprise a sum of hazard rates for recovery, death, secure isolation or other states. In the absence of any other dynamics influencing the number of susceptible people[7] we have the accounting identity

$$\beta s(t^*) = \gamma$$

and an expression for the vulnerable fraction of society

$$s(t^*) = \frac{\gamma}{\beta} \tag{4}$$

as a ratio of instantaneous rates.

adjustment,

$$J = E\left[\frac{1}{\eta}\right]$$

where we would be tempted to use the notation H for herd immunity were it not for the confusion this would cause ($J \approx H^{-1}$ not H, where H is standard notation for the harmonic mean). Furthermore this is only approximate.[8]

To illustrate, suppose that the law of the variation function η is a scaling of a Beta distributed random variable.[9] The mean of $Beta(\alpha, \beta)$ is $\frac{\alpha}{\alpha+\beta}$ and the harmonic mean exists for $\alpha > 1$ and $\beta > 0$ which we assume. The harmonic mean is $\frac{\alpha-1}{\alpha+\beta-1}$ and the ratio of the mean to harmonic mean is

$$J = \frac{\alpha}{\alpha+\beta}\frac{\alpha+\beta-1}{\alpha-1} = \frac{\alpha^2 + \alpha\beta - \alpha}{\alpha^2 + \alpha\beta - \alpha - \beta}$$

which provides the convexity adjustment in closed form.

3.1.1 Unnormalized variation functions

A brief pragmatic comment on calculation of convexity adjustments. Suppose $\eta = c\varphi$ for some constant c and unnormalized variation function φ. Then

$$J = \frac{1}{c}\int_\Omega \frac{1}{\varphi}d\rho = \frac{\int_\Omega \frac{1}{\varphi}d\rho}{\int_\Omega \varphi d\rho}$$

and it is often more convenient to compute a ratio of two unnormalized integrals than worry about constants. In particular we will consider cases where $\varphi \propto \rho^\alpha$ and thus

$$J(\alpha) = \frac{\int_\Omega \rho(\omega)^{1-\alpha}d\rho}{\int_\Omega \rho(\omega)^{1+\alpha}d\rho} = \frac{I(-\alpha)}{I(\alpha)}$$

for some integral I with parameter α. Evidently it is not necessary to carry through all constants in the integrals in order to determine $J(\alpha)$ or a ratio of two $J(\alpha)$.

---

[8]A slightly more careful analysis would consider the case when $\beta(\omega) < \gamma$ and the first order condition is no longer relevant. A better approximation is provided by

$$s(t^*) = \gamma/\beta E\left[\frac{1}{\max(\eta, \eta_{min})}\right]$$

where $\eta_{\min} = \gamma/\beta$. For smaller variation this may be splitting hairs.

[9]We need to scale so that the mean of η is unity. Let $\Omega = (0, \beta/\alpha)$ and $\rho = \frac{1}{\beta/\alpha}$ be the uniform measure and ν be defined by the inverse cumulative distribution of a beta distributed variable that has been scaled by $\beta/\alpha > 0$.

14

3.1.2 Interpreting reported R values based on peak infection

With these thoughts in mind, if a reported number ($\hat{R}$ say) is known to be derived from a peak infection estimate, which is to say

$$\hat{R} = 1/\hat{s}(t^*)$$

for some estimate $\hat{s}(t^*)$ of peak s we should be careful to interpret this directly as an estimate of the population size and not a link to mean infection rate via the SIR identity $R_0 = \gamma/\beta$. Even if we bravely identify $s_{\mathcal{T}}$ with $s(t^*)$ we would be better served by

$$\bar{\beta} = J\gamma\hat{R}$$

instead of $\beta = \gamma R_0$. The mean rate of contagion $\bar{\beta}$ we infer is *higher* than a constant β that might be inferred using a homogeneous model.

3.1.3 Peak infection at different times

In the preceding discussion of peak infection we implicitly assumed critical times $\mathcal{T}^* = \{t_\omega\}_{\omega\in\Omega}$ in the evolution of an ensemble of classic SIR compartmental models. Really, peak infection globally occurs when

$$i'(t_\omega) = 0 = \int_\Omega \bar{\beta}\eta(\omega)i(t_\omega,\omega)s(t_\omega,\omega)\rho(\omega) - \gamma i(t_\omega,\omega)\rho(\omega)d\omega$$

and it is bold to approximate this by assuming balance holds for each ω simultaneously:

$$s() = \frac{\gamma}{\bar{\beta}\eta()}$$

Here $s(\omega)\rho(\omega)d\omega = sd\rho$ is interpreted as the count of susceptible people in a sub-population at this turning point t_ω. This gives rise to the approximation

$$s(t^*) \approx s_{\mathcal{T}} = \int s(t_\omega)d\rho = \frac{\gamma}{\bar{\beta}} \int_\Omega \frac{1}{\eta}d\rho$$

However the first approximation may be poor. In the appendix we consider a further convexity adjustment taking into account the differing times of peak infection.

On the other hand as variation is increased the set $\mathcal{T}$, the times t_ω at which individual peaks are reached, will become more disperse so the estimate $\hat{s}(t^*)$ will be lower than it

otherwise would be with no variation. In fact we will have

$$\hat{s}(t^*) = s_{\mathcal{T}}/F = \frac{1}{F} \int s(t_\omega) d\rho$$

for some ratio $F > 1$ depending on choice of $\bar{\beta}$ and η (also γ and $i(0)$).[10] Thus a more careful inference would be

$$\hat{R} = \frac{1}{\hat{s}(t^*)} = \frac{1}{Fs_{\mathcal{T}}} = \frac{\beta_h}{F\gamma} = \frac{1}{F}\frac{1}{J}\frac{\bar{\beta}}{\gamma}$$

leading to

$$\bar{\beta} = FJ\gamma\hat{R}$$

in place of the usual $\beta = \gamma R_0$ in the SIR model.

3.2   Early stage growth

Convexity adjustments are also available for early stage growth. We consider mean growth starting with the same proportion of initially infected people $i(0)$ where again infection is drawn at random according to $\beta(\omega) = \bar{\beta}\eta(\omega)$ as before.[11] We further assume $s(t) \approx 1$ since we assume the measurement takes place in the early stage of the epidemic. So approximately we have

$$i'(t;\omega) = \bar{\beta}\eta(\omega)i(t) - \gamma i(t)$$

Thus by integration

$$\frac{i(t)}{i(0)} = \int \frac{i(t;\omega)}{i(0)} d\rho(\omega) = e^{-\gamma t} E_\rho \left[e^{t\bar{\beta}\eta} \right]$$

We might refer to the quantity

$$\beta_g := \frac{1}{t} log E_\rho \left[e^{t\bar{\beta}\eta} \right]$$

defined for some time t that is well before herd immunity, as the *growth effective infection rate*. It plays a similar role to β in the constant parameter SIR model, at least as far as growth in aggregate infections is concerned. The growth effective infection rate is a

---

[10]The mnemonic here is that F stands for Fourier. For readability all ratios F, J and G to follow are defined so as to be greater than 1.

[11]We are attempting to determine properties of an exchangeable model that lies on the same orbit as a realistic one. For a discussion of interacting groups and early stage growth see [17] for example.

multiple of $\bar{\beta}$,

$$\beta_g = G(t)\bar{\beta}$$

where we introduce G and note that again by Jensen's Inequality $G(t) > 1$. The growth ratio G will depend on the parameters of η and the time t marking the end of the interval $(0, t)$ where we measure growth.

It will be apparent that G is closely related to the moment generating function of the variation function η. In financial terminology it also plays a role analogous to a yield. We will draw a closer connection to finance in Section 4. Note for now that dividing by $\bar{\beta}$ we have

$$G(t) = \frac{\log E_\rho[e^{t\bar{\beta}\eta}]}{\bar{\beta}t}$$

showing that adjusted growth as a multiple of $\bar{\beta}$, when time is measured in units of $1/\bar{\beta}$, is equal to the extent to which the logarithm of the moment generating function of η grows faster than linear time. In those more sensible time units

$$G'(t') = \frac{\log E_\rho[e^{t'\eta}]}{t'}$$

is slightly more convenient expression of G with $t' = \bar{\beta}t$.

3.2.1 Gaussian approximation

For gaussian η

$$E_\rho\left[e^{t\bar{\beta}\eta}\right] \approx e^{t\bar{\beta}\bar{\eta} + \frac{t^2}{2}\bar{\beta}^2 E[(\eta-1)^2]}$$

is exact and this approximation may be useful elsewhere. The exponent is the variance σ^2 of the variation function η. Thus measured growth to time t will be

$$g(t) = \bar{\beta} - \gamma + \frac{t}{2}\bar{\beta}^2\sigma^2$$

and effective β_g given by

$$\beta_g = \bar{\beta}\overbrace{\left(1 + \frac{t}{2}\sigma^2\right)}^{G(t)}$$

although this demands care in interpretation for non-gaussian variation. The issue is that $\bar{\beta}t$ may not be small so approximations for small times may have limited utility. Early stage growth is often measured after a few multiples of $1/\bar{\beta}$ to minimize small sample effects, censoring and related measurement inaccuracies. It may also be prudent to choose t larger than $1/\gamma$, the typical recovery time, to reduce the impact of short

17

time effects such as an initial, rapid decrease in novel interactions.

If η varies by a large percentage over some regions of the probability space we must be additionally careful. Therefore in what follows we assume t is several multiples of $1/\bar{\beta}$. To the extent that analytic approximations are useful one might steer towards the expansions around $t\bar{\beta} = \infty$ rather than $t = 0$. We shall also compute G for several distributions of η in Section 6.4.

3.2.2 Growth adjustments under stochastic evolution

Thus far we have considered deterministic evolution. However to prepare for Section 4 we briefly remark that this is not a prerequisite and we can allow stochastic dynamics. In the field of quantitative finance, considerable work has gone into the finding of solutions of the form

$$u(t, x_0) = E\left[\exp\left(\int_0^t x_s ds\right) |x(0) = x_0\right]$$

where x_s is a stochastic process and the integral can be viewed as a Feynman-Kac formula solving a diffusion equation. One interpretation arises when x_s is a short rate (instantaneous rate of interest) and the integral is the balance in an account after time t when all interest is reinvested. Assuming a convenient expression or method of calculation, the growth convexity adjustment is facilitated by the fact that

$$\frac{E\left[\exp\left(\int_0^t x_s ds\right) |x(0) = x_0\right]}{\exp\left(E\left[\int_0^t x_s\right] ds\right)} = \frac{E\left[\exp\left(\int_0^t x_s ds\right) |x(0) = x_0\right]}{\exp\left(\int_0^t E[x_s] ds\right)} \tag{5}$$

is tractable (where the equality glosses over some technicalities).

More frequently, at least in the fixed income literature, formulas as provided for the closely related quantity

$$u(t, x_0) = E\left[\exp\left(\int_0^t -x_s ds\right)\right]$$

whose interpretation is the price of a bond with maturity t. The insurance and failure rate literature also provided examples of convenient choices of stochastic process x_s where x_s is interpreted as an instantaneous rate of failure of a machine, or default of a bond issuer and the integral represents a probability of survival (either a real world probability or more likely, in the case of credit default swaps, a pricing measure).

For some solutions there is a trivial connection between the moment generating function and the survival function. We give an example in Section 4 of using $-x$ in

place of x to compute mean growth for an epidemic.[12]

3.2.3 Limitations

The convexity adjustment G ratios assume $s(t) \approx 1$ and will be invalid if the susceptible population rises to an appreciable level for any appreciable mass of Ω.

3.3 Combining convexity adjustments longitudinally

Having considered the impact on both early stage growth and peak infection, we turn to the questions of how observations of the former might inform the latter - and how one might find models which are capable of fitting both types of evidence for some choice of parameters.

3.3.1 Forecasting peak infection

The mean infection rate $\bar{\beta}$ and variation η describe the system and thus must relate the two, but they are not directly observed. Without the use of convexity adjustments one might be tempted to follow the usual prescription:

1. Estimate R_0 by some means

2. Forecast $s(t^*) = 1/R_0$.

where theoretically $R_0 = \gamma/\beta$. The evidence from Sweden seems to suggest that R_0 estimated in step one leads to an underestimate of $s(t^*)$ in step 2. Referring to Britton's use of this methodology for Stockholm [3], we note his estimates of parameters as of April 15th, a full month after interventions, of $\gamma = 0.132$, $\beta = 0.209$ and thus $R = 1.58$. This implies a mean generation time of 7.5 days and a threshold infection level of $1 - s(t^*) = 0.37$. However infection rates appear to have dropped in Stockholm well before reaching this level of penetration. An even bigger puzzle exists if we assume consensus R_0 estimates from earlier studies with $R_0 \approx 2$ or $R_0 \approx 2.5$.

Here is a possible explanation. Assuming γ is known and the true system behaves is as we have described it - as a mixture with varying β - the first step will in fact produce an estimate of the growth effective infection rate β_g due to the relation $R_0 = \gamma/\beta_g$. Here the denominator is related to $\bar{\beta}$ and η by

$$\beta_g = G(t)\bar{\beta}$$

---

[12]Depending on the model in question, reflection under the x-axis may not be possible. However it may be possible to recycle existing analytics in other ways such as reflection in time.

as we have seen.[13]

We have also determined that the susceptible population at peak $s(t^*)$ relates to $\bar{\beta}$ and η via

$$s(t^*) = \beta_h/\gamma = \frac{1}{J}\frac{\bar{\beta}}{\gamma}$$

Thus we might write

$$s(t^*) = \frac{1}{J}\frac{1}{G}\frac{\beta_g}{\gamma} = \frac{1}{J}\frac{1}{G}\frac{1}{R_0} \tag{6}$$

Thus we are presented with two convexity adjustments J and G that may combine to resolve the apparent discrepancy between early stage growth and peak infection. The threshold number of infections will be higher due to two applications of Jensen's Inequality.[14]

One can add a third convexity effect if we also consider the impact of peak infection at different times.

$$s(t^*) = \frac{1}{F}\frac{1}{J}\frac{1}{G}\frac{1}{R_0}$$

In future work we hope to provide simple estimates of F.

3.3.2 Inferring the amount of variation

Assume now some empirical data at different stages of epidemics and a desire to build models that can, for some choice of parameters, explain multiple types of observation coherently. In particular suppose we are presented with data for the two functionals we have paid attention to: early stage growth and peak infection susceptible population. Data for the latter is hard to come by but let us take a ratio of β to γ of approximately $5/4$ as illustrative.

On the other let us also suppose, again just for illustration, that this ratio implies 1.7 times less early stage growth than is implied by reproduction numbers on the lower end of the spectrum.[15]

---

[13]The term $G(t)$ will be a function of the interval $(0, t)$ over which a measurement of early stage growth is made.

[14]Strictly speaking, Jensen's Inequality will not apply if we use a more careful approximation to peak infection, as noted in the Appendix, but J is likely to exceed 1.

[15]In the case of constant γ the time until an infectious person transmits to another is exponentially distributed. This leads to a relationship between reproduction number R, growth r and recovery γ

$$R = 1 + \frac{r}{\gamma}$$

consistent with $R = \beta/\gamma$ and

$$r = \beta - \gamma$$

as we expect. When variation is introduced into γ the reproduction number will be reduced relative

We propose to use the relationships established in Section 3 to quickly identify the parameter θ which controls the variation in η and thus β. In other words, we solve

$$\underbrace{1.7}_{Swedish\ puzzle\ factor} = \overbrace{\frac{mean\ growth}{SIR\ growth}}^{growth\ convexity} \times \overbrace{\frac{arithmetic\ mean\ \beta(t^*)}{harmonic\ mean\ \beta(t^*)}}^{herd\ convexity}$$

for θ or more abstractly, solve for η. We emphasize again that the variation implied by this procedure must be viewed in the light of de Finetti's Theorem. The mixing serves as a proxy for *dependence* between neighborhoods that are otherwise independent, so it need not represent a seemingly realistic level of actual spatial variation.

Referring back to our original Figure 5, it is apparent that inverting convexity adjustments can move us much closer to the correct orbit. This approach stands in stark contrast to other modeling strategies which involve modifications to elaborate simulations that attempt to model reality at a granular level (using geospatial demographic data, transport networks and so forth). These models may in some respect be closer to the truth. However they may easily be further from the equivalence class of which nature's model is a member.

3.3.3   An iterative procedure

Inverting the convexity adjustments is unlikely to present serious difficulty but for concreteness, one simple iterative approach would simply assume - in each loop. - equal residual convexity adjustments. Speaking generally about any two functionals f_1 and f_2 represent measurements, with respective convexity adjustmets J_1 and J_2, once might suppose γ is known and $\bar{\beta}$ and θ are to be determined. We assume a missing factor of λ where, in the case of Sweden's puzzling herd immunity, we might set $\lambda = 1.7$ say. As a first pass we express this model mispecification as

$$f_1(\phi(\bar{\beta}_0, \theta)) \approx J_1(\theta, \bar{\beta}) f_1(\phi(\bar{\beta}_0, \eta \equiv 1)) \approx \sqrt{\lambda} \overbrace{f_1(\phi^*)}^{observed}$$

where ϕ^* is nature's unseen true model but $f_1(\phi^*)$ is observed, and

$$f_2(\phi(\bar{\beta}_0, \theta)) \approx J_2(\theta, \bar{\beta}) f_2(\phi(\bar{\beta}_0, \eta \equiv 1)) \approx \frac{1}{\sqrt{\lambda}} \overbrace{f_2(\phi^*)}^{observed}$$

---

to the reproduction number computed using the mean of γ. However we shall limit our attention to convexity adjustments for the variable β, referring the reader to Wallinga and Lipsitch for further discussion of mean generation times [23].

and $\bar{\beta}_0$ is some first guess of infection rate parameter taking into account the conflicting evidence.

For instance if threshold infection implies $\beta = 1.25\gamma$ and growth implies $\beta = 1.65\gamma$ then we take $\beta^* = 1.45\gamma$. Here ϕ^* is nature's true model and the right sides of these equations are observed quantities.[16] Each of these equations might be solved for θ independently, leading to a next guess at θ (say the average of the two) and a next value of λ. This simple approach is motivated by the observation that J's may be entirely independent of $\bar{\beta}$ altogether, or depend only weakly on the same.

3.4 Cross-sectional use of convexity adjustments

Before leaving the topic of convexity adjustments we point our that longitudinal prediction along the trajectory of a single epidemic is not the only application. Since variation driven adjustments provide corrections for measurable quantities relative to deterministic compartmental models, they can also provide a very convenient way to relate two different epidemics *at the same stage* subject to different conditions.

Moreover, the convexity adjustments per se may not vary strongly with underlying assumptions used for the deterministic models. So it is possible that this sort of comparison may sidestep problems associated with calibration of compartmental models. To illustrate we consider threshold infection levels. The measurable of interest is the number of susceptible people.

This time we assume variation η is driven by population density. Though this is no the only source of variation it is certainly significant and, moreover, the degree of variation also varies considerably. This is manifest in the way population density takes very different *shapes*, referring to the geometry of population density. Figure 2 shows the population density for London, whereas the equivalent for Paris is depicted in Figure 3.[17]

Empirical study of population density and its effect on epidemic growth is a muddied picture with some authors appealing to the likelihood of a large impact (e.g. [21]) whereas in a study of pine beetles a thinning of the population did not noticeably change an epidemic's equilibrium [16]. Population fluctuations appear to drive pathogen spread in Niger [9]. Different rates of microparasite transmission were observed in cities versus countries, providing indirect evidence of the role of density [10]. HIV transmission

---

[16]This may well be the case at time of writing. If f_1 and f_2 represent measurements of early stage growth of an epidemic and threshold infection, as we have been discussing, then evidence from Sweden may be interpreted as an onset of herd immunity with $s(t) \approx 0.8$ corresponding to a ratio of β to γ in the SIR model of 5/4, much lower than R values implied by spread of the disease.

[17]Population graphics provided by Matt Daniels of The Pudding.
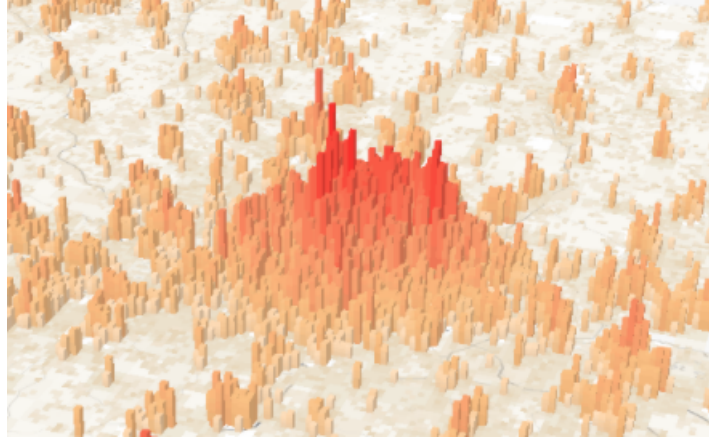
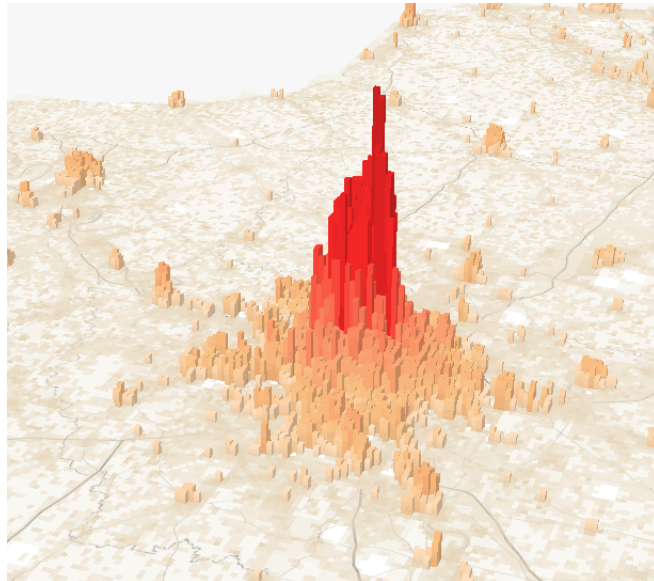Figure 2: Population density for London.

Figure 3: Population density for Paris

reconstructed with phylogenetic methods revealed density dependence [15].

COVID-19 cases in China and the U.S. have seemingly concentrated in areas of high population density - though there are of course confounding factors such as proximity to international airports and movements of people. Without seeking to evaluate the evidence we suppose merely in this section that population density plays a role - possibly a weak one.[18]

Focusing on cross-section comparison, we suppose two towns: Town A and Town B. We assume the epidemic has long since run its course in Town A. We are given the task of estimating herd immunity levels and reproduction numbers for Town B. This town shares many of the same properties as Town A including the same mean contagiousness - which is to say the likelihood of a contagious person infecting another. The only difference between the two towns we can identify is that Town A's population profile is shaped like a bell curve whereas Town B's falls away more like an exponential distribution.

3.4.1 Rules of thumb for peak infection from population shape

Town A		Town B	
Shape	Gaussian	Shape	Exponential
Reproduction	5.0	Reproduction	R_b ?
Herd	$h_a = 0.8$	Herd	h_b ?
Vulnerable	$v_a = 0.2$	Vulnerable	v_b ?

Shape	Coef	Shape	Coef
Flat	0	Skirt $(1/r^4)$	4
Gaussian	2	Gamma $(rexp(-r)$	4
Exponential	4	Gamma $(r^2exp(-r))$	3
Skirt $(1/r^3)$	6		

Based on *ratios of convexity adjustments* to be provided, we suggest the following way to answer this challenge. A calculations such as the following can be used to adjust for population shape:

$$v_b = v_a \left\{ 1 + \underbrace{0.1}_{power\ law} (\overbrace{4}^{shape\ b\ coef} - \overbrace{2}^{shape\ a\ coef}) \right\} = 1.2v_a = 0.24$$

---

[18]In fact the convexity adjustments work when infection is proportional to a power of population density less than 1 as well as greater than 1, so one might even use them under the seemingly unlikely scenario that density reduces infections.

from whence it follows that $h_b = 1 - v_b = 0.76$ and $R_b = 1/0.76 = 4.16$. This heuristic is derived from first order approximations of two convexity adjustments performed back to back. The reader may choose to view them as reproduction number interpretations chained as follows:

$$R_a \to R_{hom} \to R_b$$

where R_{hom} represents a reproduction number that that would exist in the absence of any variation. We probably never get to see R_{hom}, however, because there isn't a country, city or neighbourhood in the world lacking variation. Instead, we first *infer* a true underlying level of contagiousness in Town A by subtracting a convexity effect, and then add back the convexity correction for Town B.

The coefficients provided in Table 3.4.1 are simply the first terms in a Taylor expansion of the product of convexity adjustments. We provide some formalization of these calculus exercises.

3.4.2   Rotation invariant population densities and power law

We identify the probability space with the plane $\mathcal{R}$. The link between density and contagiousness is muddied [21] [15] but even a weak relationship may be important. We shall for simplicity assume that contagiousness β is a function of population density via a power law.

$$\beta(\rho) \propto \rho^\alpha \tag{7}$$

for some $0 < \alpha < 1$ and population density ρ. Further, we shall assume that density is a radial function from the origin and thus so is $\beta(\cdot) = \beta(r)$ is as well. We can write

$$\frac{\beta(r)}{\bar{\beta}} = \eta(r) \propto \rho(r)^\alpha$$

As per our comments in Section 3.1.1 it is marginally cleaner to work with integrals of an unnormalized variation ϕ and determine the ratio of the integral.

3.4.3   Gaussian population density

If a city's population density varies as $\rho(r, \theta) = \frac{1}{2\pi} e^{-r^2/2}$ as we move out a distance r from the center then we take $\phi(r) = e^{-r^2 \alpha/2}$

$$\iint \rho\phi = \frac{1}{2\pi} \int_0^\infty e^{-\frac{1+\alpha}{2} r^2} r \, dr = \frac{1}{1+\alpha}$$

25

Recycling the same integral with $\alpha \mapsto -\alpha$ and taking a ratio:

$$J(\alpha) = \frac{1+\alpha}{1-\alpha}$$

Merely as an example, $J(a = 1/4) = 5/3$ if we use a power law coefficient loosely informed by Measles outbreaks in the United States.

3.4.4 Exponentially distributed populations

Next consider a city with more sprawl. Set $\rho(r,\theta) = \frac{1}{2\pi}e^{-r}$ to create a city whose density falls off more abruptly at first, but with a much longer tail. With $\phi(r) = e^{-\alpha r}$ we have

$$\iint \rho\phi = \int_0^\infty e^{-(1+\alpha)r}r\,dr = \frac{1}{(1+\alpha)^2}$$

yielding

$$J(\alpha) = \left(\frac{1+\alpha}{1-\alpha}\right)^2$$

which is the square of the Jensen's ratio for a Gaussian city. Comparing with the Gaussian case and holding everything else constant, even a ten percent relationship between log contagiousness and log density might imply a twenty percent correction in the number of people who escape the virus when we translate the data from a Gaussian city to an Exponential one.

More examples of shape adjustments are given in Section 6.5

4 Vasicek meta-population mixture models

We have outlined a strategy using mixtures of independent identically distributed (iid) models that are carefully interpreted as representatives of equivalence classes of models as per Figure 5. We now apply the approach using models where each neighborhood's epidemic follows a stochastic compartmental model. Infection rate follows an Ornstein-Uhlenbeck process, which implies early stage growth follows an Ornstein-Uhlenbeck process as well.

This approach also adopts the suggestion in Section 3.2.2, namely that early stage growth calculations can be borrowed from the fixed income literature. We call this a Vasicek compartmental model after the corresponding work on interest rates by Vasicek (in which short rate follow an Ornstein-Uhlenbeck process) and we borrow his formula for the price of a bond that pays no coupons.

The reader may choose to view this as a somewhat arbitrary choice of basis and thus a mere illustration of technique. However we provide in Section 7 some motivating remarks leaning both on empirical work [5] and also a more theoretical justification in terms of a physical model for repeat contact probability [6].

4.1 Stochastic process for infection

In contrast to most of the discussion in Section 3, we allow infection rate $\beta(t)$ to follow a stochastic process rather than a deterministic process. However in a manner similar to the preceding it will have parameters driven by $\omega \in \Omega$ once again, where Ω is an additional probability space having nothing to do with the evolution of the stochastic process driving β. A general example is

$$d\beta(t) = \kappa(\omega) \left(\beta_\infty(\omega) - \beta(t) \right) dt + \sigma_\beta^2(\omega) dW_t$$

with initial condition $\beta(t = 0; \omega) = \beta_0(\omega)$. Here β_∞ is a mean reversion level and the rate $\kappa(\omega)$ controls how quickly reversion occurs. This is inserted into the SIR differential equations as before. Here the initial value of infection rate, variance, mean reversion level and pull κ can all be varied with ω. Notice that if we define

$$r(t) = \beta(t) - \gamma$$

then $dr(t) = d\beta(t)$ and $r(t)$ may be seen to satisfy a similar stochastic differential equation. Add zero to the drift:

$$dr(t) = \kappa \left(\overbrace{\beta_\infty(\omega) - \gamma}^{r_\infty} - \overbrace{(\beta(t) - \gamma)}^{r(t)} \right) dt + \sigma_\beta^2(\omega)dW_t$$

to see this, noting that $r(t)$ also has different initial condition $r_0 = \beta_0 - \gamma$.

4.2 Iterated expectations

We begin with an elementary observation that applies if evolution of $\beta(t)$ is unrelated to choice of ω, or more accurately if it is conditionally independent. By the law of iterated expectations we have

$$E\left[e^{\int_0^t r(s)ds}\right] = E_\rho\left[\overbrace{E_W\left[e^{\int_0^t r(s)ds}\right]}^{inner}\right]$$

and thus

$$\frac{E\left[e^{\int_0^t r(s)ds}\right]}{e^{\int_0^t E[r(s)]ds}} = \frac{E_\rho\left[\overbrace{E_W\left[e^{\int_0^t r(s)ds}\right]}^{inner}\right]}{e^{E_\rho \int_0^t E_W[r(s)]ds}} \tag{8}$$

which will simplify in some cases, as we shall see, to give simple additive corrections for early stage growth arising from variation.

4.3 Mean (inner) early stage growth

Unlike the deterministic mixture model, the mean growth is not obvious even for fixed ω. The mean across different paths generated by W_t requires an expectation of the exponential of a stochastic integral. Fixing ω for now, expected growth in infections is given by

$$g_t := \frac{1}{t} \log E_W\left[e^{\int_0^t r(s)ds}\right] = \frac{\tau(t)r_0(\omega) + B(t;\omega)}{t} \tag{9}$$

where $\tau(t)$ and $B(t)$ are functions not depending on the initial rate $r_0(\omega)$ and we have written the expectation with a W subscript to emphasize that this is with respect to the law of the stochastic process (not the probability space Ω that might do its own mixing independently). Here $r_0(\omega)$ appears only as a linear term. The function τ depends only

28

on κ_ω

$$\tau(t) := \frac{1 - e^{-\kappa t}}{\kappa}$$

and is approximately equal to t for small times. The function $B(t; \kappa, \sigma)$ is given by

$$\frac{B(t; \omega)}{t} := \left(r_\infty(\omega) + \frac{\sigma_\omega^2}{2\kappa_\omega^2} \right) \frac{t - \tau}{t} - \frac{\sigma_\omega^2}{4\kappa_\omega^2} \frac{\tau^2}{t} \tag{10}$$

and revealed in the next section to be a thinly veiled bond price formula. The initial growth rate is $r_0(\omega)$ as we expect and the eventual growth rate is

$$g_\infty(\omega) = r_\infty(\omega) + \frac{\sigma_\omega^2}{2\kappa_\omega^2}$$

although this holds only so long as $s(t) \approx 1$. We have some flexibility to match an infection rate decaying towards a lower level, but actually the σ^2 terms allow us to fit slightly more complicated term structures, as is revealed by rearranging the growth equation.

$$g_t(\omega) = \overbrace{r_\infty(\omega) + \frac{\sigma^2}{2\kappa^2}}^{g_\infty} + \frac{\tau}{t} \left\{ (r_0(\omega) - r_\infty(\omega)) - \frac{\sigma^2}{2\kappa^2} \left(1 + \frac{\tau}{2} \right) \right\} \tag{11}$$

Here we might view the functions 1, τ/t and $(1 + \tau/2)(\tau/t)$ as a basis for shapes taken by growth as a function of time. Or, separating the roles of the parameters, we might introduce ratios

$$\nu_0 = \frac{r_0 - r_\infty}{r_\infty} \tag{12}$$

$$\nu_\sigma = \frac{\sigma}{\sqrt{2\kappa r_\infty}} \tag{13}$$

so that ν_δ represents the ratio of the mean infection rate attenuation to eventual infection rate and ν_σ is the ratio of the ergodic standard deviation of the infection rate process to the long term mean. Then growth is written

$$\frac{g_t}{r_\infty} = 1 + \nu_\delta \frac{\tau}{t} - \nu_\sigma^2 \frac{\tau^2}{\kappa t} \tag{14}$$
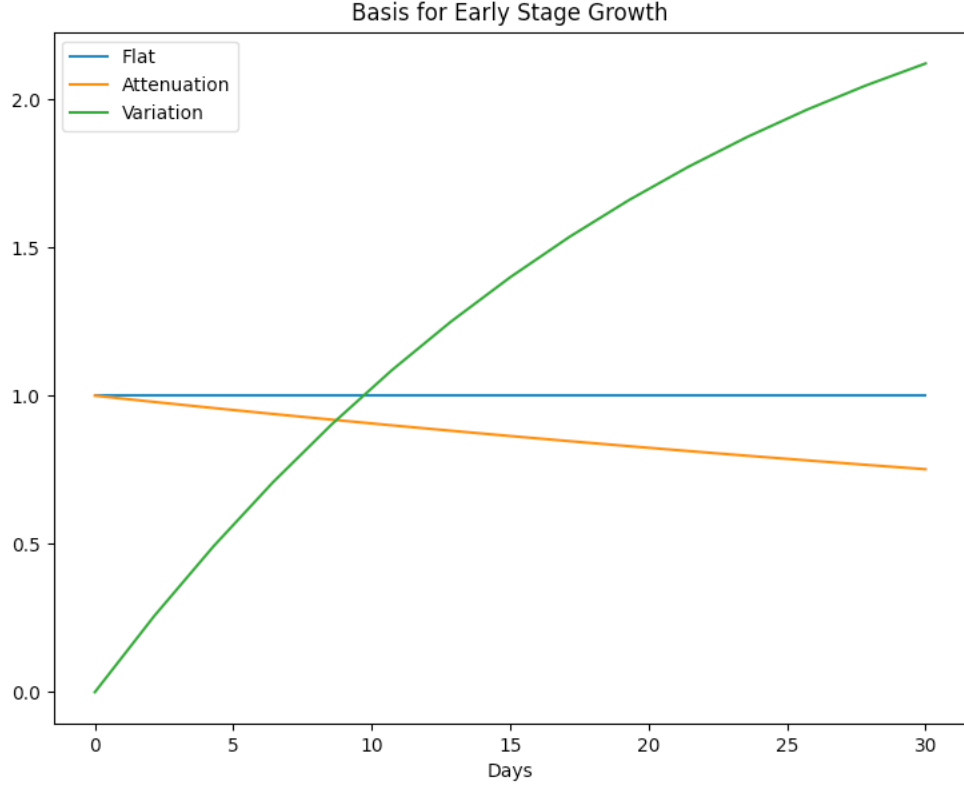
Figure 4: Basis for early stage growth shape using Equation 14. The difference between initial infection rate and long term infection rate is multiplied by the orange function τ/t which can be interpreted as including an attenuation due to repeat contacts. Variation drives a yield component shown in green, which has been scaled down by a factor of 200 relative to the term $\frac{\tau^2}{\kappa t}$ in Equation 14. The green line would be multiplied by ν_σ^2 where $\nu_\sigma = \frac{\sigma}{\sqrt{2\kappa}r_\infty}$ is the ratio of the long run standard deviation of instantaneous growth to the long run mean of the same.

We plot this basis in Figure 4, which might be viewed as a non-orthogonal basis. One might write

$$\frac{g_t}{r_\infty} = \begin{bmatrix} 1, \nu_0, \nu_\sigma^2 \end{bmatrix} \begin{bmatrix} 1 \\ \tau/t \\ -\frac{\tau^2}{2\kappa t} \end{bmatrix}$$

to emphasize a natural choice of parameters r_∞, ν_0 and ν_σ^2, yield basis and resultant ratio to long term growth.

Two of these basis functions can be accommodated in the deterministic setting also, but we can revert to that by setting $\sigma \equiv 0$ rather than treating it separately.

4.4 Relationship to term structure modeling

Our terminology is suggestive and as noted the growth formula in Equation 11 appears elsewhere in another guise. It is the price of a bond paying no coupons in a Vasicek short rate model [22]. This correspondence obviates a full derivation, though that is a straightforward if lengthy exercise.[19] Skipping past that thanks to Vasicek's formula, we need only observe that the price of a zero coupon bond with maturity t, mean reversion level x_∞, initial short rate x_0 and κ and σ as before is

$$B(x_0) = exp\left(E\left[-\int_0^t x_s ds \right] \right) = e^{-x_0\tau + \Lambda(t)}$$

where

$$\Lambda(t; x_0) = \left(x_\infty - \frac{\sigma^2}{2\kappa^2} \right)(\tau(t) - t) - \frac{\sigma^2 \tau(t)^2}{4\kappa}$$

and τ as above. If we let $x = -r$ with $x_0 \to -r_0$ and $x_\infty \to -r_\infty$ we have

$$\begin{aligned}
dr_t = -dx_t &= -\kappa\left(-r_\infty - (-r)\right) - \sigma dW_t \\
&= \kappa\left(r_\infty - r\right) + \sigma d(-W_t)
\end{aligned}$$

and since W_t and $-W_t$ have the same law, r_t follows a "Vasicek process". Thus by using negative interest rates in this manner we have

$$\begin{aligned}
\overbrace{\frac{1}{t}\log E\left[\int_0^t r_s ds \right]}^{mean\ early\ stage\ growth} &= \frac{1}{t}\log E\overbrace{\left[-\int_0^t x_s ds \right]}^{Vasicek\ bond\ price} \\
&= \frac{1}{t}\left(-(-r_0)\tau + \Lambda\left(t; x_\infty = -r_\infty, x_0 = -r_0\right) \right) \\
&= \frac{1}{t}\left(r_0\tau + \left(-r_\infty - \frac{\sigma^2}{2\kappa^2} \right)(\tau - t) - \frac{\sigma^2\tau^2}{4\kappa} \right) \\
&= \frac{1}{t}\left\{ r_0\tau + \underbrace{\left(r_\infty + \frac{\sigma^2}{2\kappa^2} \right)(t - \tau) - \frac{\sigma^2\tau^2}{4\kappa}}_{B(t)} \right\}
\end{aligned}$$

and we have justified Equation 9.[20]

---

[19]The Vasicek bond price formula is a convexity calculation for a normal random variable. Most of the terms arise from the formula for the variance of the Ornstein-Uhlenbeck process, one half of which contributes to the difference between $E[exp(\int_0^t r_s ds)]$ and $exp(E[\int_0^t r_s ds])$.

[20]The exponentially exploding bond price may seen counterintuitive. However a bond becomes extremely expensive when it is an attempt to shield wealth from the confiscatory effect of negative
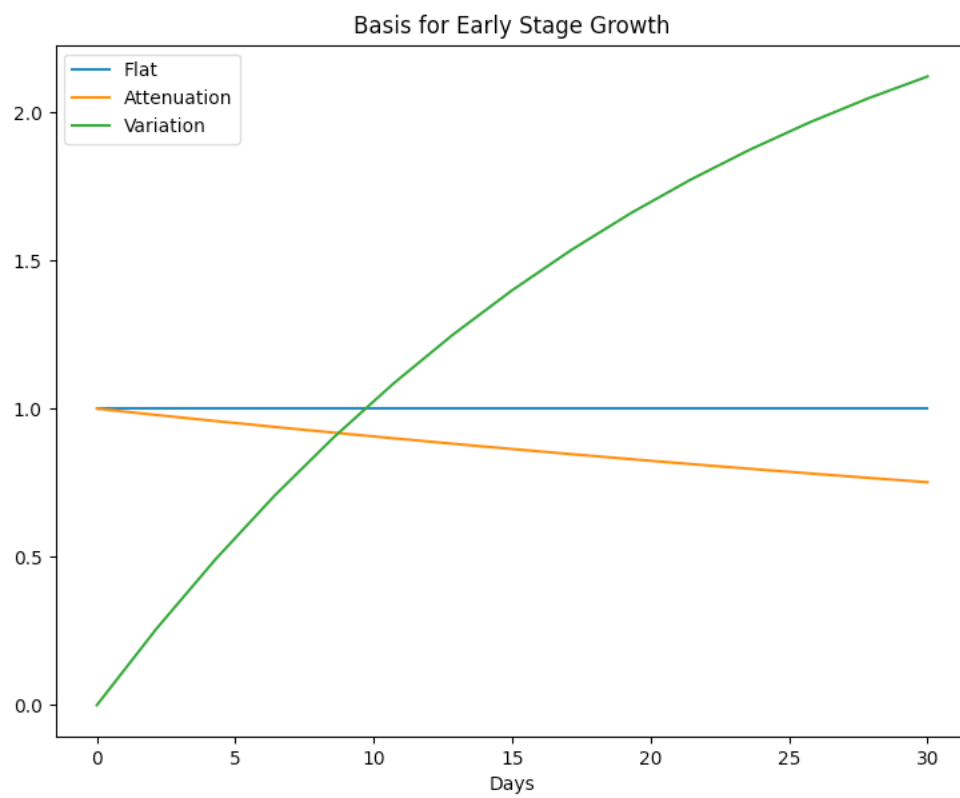
Figure 5: Caption

4.5 Inner convexity

Again let us fix $\omega \in \Omega$. From the "inner" mean growth rate defined in Equation 9 and given in Equation 11, we derive the convexity adjustment defined in Equation 5 for mean early stage growth relative to the deterministic model simply by setting $\sigma(\omega) \equiv 0$. It is apparent from Equation 10 that

$$
\begin{aligned}
G(t) &= \frac{1}{t} \log \left(\frac{E_W \left[\exp \left(\int_0^t r_s ds \right) | r(0) = r_0 \right]}{\exp \left(\int_0^t E[r_s] ds \right)} \right) \\
&= B(t; \sigma) - B(t; \sigma = 0) \tag{15} \\
&= \frac{\sigma_\omega^2}{2\kappa_\omega^2} \frac{t - \tau}{t} - \frac{\sigma_\omega^2}{4\kappa_\omega^2} \frac{\tau^2}{t} \\
&= \frac{\sigma_\infty^2}{\kappa_\omega} \left\{ \frac{t - \tau}{t} - \frac{\tau^2}{t} \right\} \rightarrow \frac{\sigma_\infty^2}{\kappa_\omega} \tag{16}
\end{aligned}
$$

where we recall that

$$
\sigma_\infty^2 = \frac{\sigma^2}{2\kappa}(\omega)
$$

is the long term ergodic variance of the infection rate β. This is the multiplicative adjustment that needs to be made to account for variation induced by different paths taken by infection in different places, as compared with simply assuming infection rate is the same everywhere.

4.6 Overall convexity

Referring to Equation 8 we next consider the overall convexity taking into account variation due to $\omega \in \Omega$, and show that this is easily computed for some choices. In other words we allow η to determining the law of $\beta_0 = \bar{\beta}\eta$ in analogy with the deterministic SIR case considered in Section 2. Using an expression for the mean exponentialed integral analogous to Equation 11, we have

$$
G(t) = \frac{1}{t} \log \frac{E_\rho \left[\exp \left(r_\infty(\omega) + \frac{\sigma^2}{2\kappa^2} + \tau \left\{ (r_0(\omega) - r_\infty(\omega)) - \frac{\sigma^2}{2\kappa^2} \left(1 + \frac{\tau}{2} \right) \right\} \right) \right]}{E_\rho \left[\exp \left(r_\infty(\omega) + \tau (r_0(\omega) - r_\infty(\omega)) \right) \right]} \tag{17}
$$

This will simplify in some cases we now consider.

---

instantaneous interest rates.

4.7 Independent variation

Suppose now that under the measure ρ the variables r_∞, σ, κ and r_0 vary independently. The numerator in Equation 17 then factors into product terms, many of which cancel immediately with terms on the denominator.

$$
\begin{aligned}
G(t) &= \frac{1}{t} \log \frac{E_\rho \left[\exp\left(r_\infty(\omega)t + \tau(r_0(\omega) - r_\infty(\omega))\right) \right] E_\rho \left[\exp\left(t\frac{\sigma^2}{2\kappa^2} + \tau\left\{-\frac{\sigma^2}{2\kappa^2}\left(1 + \frac{\tau}{2}\right)\right\}\right) \right]}{E_\rho \left[\exp\left(r_\infty(\omega)t + \tau(r_0(\omega) - r_\infty(\omega))\right) \right]} \\
&= \frac{1}{t} \log E_\rho \left[\exp\left(t\frac{\sigma^2}{2\kappa^2} + \tau\left\{-\frac{\sigma^2}{2\kappa^2}\left(1 + \frac{\tau}{2}\right)\right\}\right) \right] \\
&= \frac{1}{t} \log E_\rho \left[\exp\left(t\frac{\sigma^2}{2\kappa^2}\right) \right] + \frac{1}{t} \log E_\rho \left[\exp\left(-\tau\frac{\sigma^2}{2\kappa^2}\left(1 + \frac{\tau}{2}\right)\right) \right]
\end{aligned}
\tag{18}
$$

which we recognize as a sum of convexity adjustments that might be independently computed. For example the first term is a convexity adjustment pertaining to the ergodic mean variance. If we write

$$
\sigma_\infty^2(\omega) = \sigma_\infty^2 \eta(\omega)
$$

for some convenient choice of variation function η then the approach of Section 3 applies (assuming independence allows this to be recycled in the second term also).

4.8 Additive convexity adjustments for growth

To focus on this first term in Equation 18 momentarily, which is the only one to survive several mean generation times:

$$
\frac{1}{t} \log E_\rho \left[\exp\left(t\frac{\sigma^2}{2\kappa^2}\right) \right] = \frac{1}{t} \log \exp\left(tE\left[\frac{\sigma^2}{2\kappa^2}\right] G_\Omega\right) = G_\Omega + G_W
$$

where G_Ω is a convexity adjustment arising from variation of σ_∞^2 over Ω. On the other hand G_W might also be called a convexity adjustment although it arises from the *existence* of σ_∞^2 and is proportional to the mean value of the ergodic variance.

$$
G_W = E_\rho \left[\frac{\sigma^2}{2\kappa^2}\right]
$$

In this way a simple meta-population model with two types of variation is easy to relate to its deterministic counterpart at least as far as early stage growth is concerned. This property arises from choice of an infection rate falling into the affine class, as we

remarked earlier, which is evidently a tractable choice. In contrast, if early stage growth is not linear in the initial value of β_0 then we have more work to do. Likewise if a complicated dependency is assumed between the variables r_∞, σ, κ and r_0 as functions of ω then no simple calculation such as this may apply. This is not to suggest that there are not other ways to make progress simplifying Equation 11 for other choices.

4.9   Peak infection

The instantaneous infection rate at the onset of herd immunity will determine whether balance occurs between recovery and new infections (as compared with the mean growth up until that point in time).[21] The mean and variance of $\beta(t)$ are

$$E[\beta(t)] = \beta_\infty + e^{-\kappa t}(\beta_0 - \beta_\infty) \tag{19}$$

$$var(\beta(t)) = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa t}\right) \tag{20}$$

so with

$$\sigma_\infty = \frac{\sigma}{\sqrt{2\kappa}}$$

as shorthand then the ergodic distribution is $\beta(t \to \infty) \sim N(\beta_\infty, \sigma_\infty^2)$. As noted in Equation 6 an approximate convexity adjustment for peak infection might break down into two components: one arising from deviation of the harmonic mean from the arithmetic mean and the other arising from a superposition effect.

Some approximations for peak infection might be couched in terms of an auxiliary function

$$h(a, b, c; \mu, \sigma) := \frac{1}{E\left[\frac{1}{a+b\max(X,c)} | X \sim N(\mu, \sigma^2)\right]} \tag{21}$$

which is a generalization of the harmonic mean. To see this, note that the susceptible population can be approximated as

$$s(\omega, t) := E_W\left[\frac{\gamma}{max(\beta(t), \gamma)}\right] = E_W\left[\frac{1}{max(\beta(t)/\gamma, 1)}\right] = h(1, 0, 1; \mu, \sigma)$$

where

$$\mu = \frac{\beta_\infty + e^{-\kappa t}(\beta_0 - \beta_\infty)}{\gamma}$$

---

[21]There is a tendency to conflate these two quantities, and the corresponding R values, in a constant parameter compartmental model.

and
$$\sigma = \sigma_\infty(1 - e^{-2\kappa t})/\gamma$$

are the mean and variance of the infection rate. This approximation assumes that when $\beta(t) < \gamma$ the epidemic never really gets going and the susceptible population remains close to unity.

4.10   Harmonic mean of gaussian

Using the Taylor expansion

$$\frac{1}{1+x} \approx 1 - x + x^2 - x^3 + x^4 + O(x^5)$$

valid for $x \in (0,1)$ let $\{\epsilon_i\}_{i=1}^n$ be independent standard normal. In the limit $n \to \infty$ we have

$$
\begin{aligned}
h(0, 1, \infty; \mu, \sigma) &= \frac{1}{E\left[\frac{1}{X} | X \sim N(\mu, \sigma^2)\right]} = \frac{1}{E\left[\frac{1}{\mu+\sigma Z} | Z \sim N(0,1)\right]} \\
&\approx \frac{n}{\sum_{i=1}^n \frac{1}{\mu+\sigma\epsilon_i}} \\
&\approx \frac{n}{\frac{1}{\mu}\sum_{i=1}^n \left(1 - \frac{\sigma}{\mu}\epsilon_i + \frac{\sigma^2}{\mu^2}\epsilon_i^2 - \frac{\sigma^3}{\mu^3}\epsilon_i^3 + \dots\right)} \\
&= \frac{\mu}{1 - \frac{\sigma}{\mu}\frac{1}{n}\sum_{i=1}^n \epsilon_i + \frac{1}{n}\sum_{i=1}^n \frac{\sigma^2}{\mu^2}\epsilon_i^2 - \frac{1}{n}\sum_{i=1}^n \frac{\sigma^3}{\mu^3}\epsilon_i^3} \\
&\approx \mu\left(1 + \frac{\sigma}{\mu}\frac{1}{n}\sum_{i=1}^n \epsilon_i - \frac{1}{n}\sum_{i=1}^n \frac{\sigma^2}{\mu^2}\epsilon_i^2 + \frac{1}{n}\sum_{i=1}^n \frac{\sigma^3}{\mu^3}\epsilon_i^3 + \dots\right) \\
&\to \mu\left(1 - \frac{\sigma^2}{\mu^2} - 3\frac{\sigma^4}{\mu^4} - \dots\right)
\end{aligned}
$$

because the odd powers have zero mean and the even powers are the well known moments of the normal distribution. The simplest convexity adjustment for peak infection using

only the first term is

$$
\begin{aligned}
J &= \left(1 - \frac{\sigma^2}{\mu^2}\right)^{-1} \\
&\approx 1 + \frac{\sigma_\infty(1 - e^{-2\kappa t})/\gamma}{\left(\frac{\beta_\infty + e^{-\kappa t}(\beta_0 - \beta_\infty)}{\gamma}\right)^2} \\
&= 1 + \frac{\gamma \sigma_\infty(1 - e^{-2\kappa t})}{(\beta_\infty + e^{-\kappa t}(\beta_0 - \beta_\infty))^2} \\
&\rightarrow 1 + \frac{\gamma}{\beta_\infty}\frac{\sigma_\infty}{\beta_\infty} \qquad\qquad (22)
\end{aligned}
$$

where we recognize the two ratios.

4.11   A simple adjustment to peak infection susceptibles

The adjustment in Equation 22 is an easy mental calculation as $R_\infty := \beta_\infty/\gamma$ is a kind of limiting R_0 value (to mix conventions somewhat) whereas $\sigma_\infty/\beta_\infty$ is the limiting *relative* variation of infection. Thus to translate relative variation into a herd immunity adjustment we can use the very coarse but easy to remember formula:

$$
susceptible\ multiplier = 1 + \frac{relative\ variation}{R}
$$

This, we emphasize, is the number we should multiply the susceptible population by to correct an estimate that is derived by assuming infection follows a deterministic trajectory. Note that even for R values such as $R = 2$ or $R = 4$ this is an appreciable correction. We can also use more accurate approximations using the derivation of Equation 22, of course.

5 Summary

It is possible that some compartmental models for disease spread suffer from a pronounced flaw of averages, and it is notable that in comparing model predictions for early stage growth and peak infection, these errors combine as we have shown in Section 3. However we need not discard the elegance and simplicity of compartmental models, if simple corrections for variation are available.

We have treated the predictions of models exhibiting variation as perturbations of models that do not. This follows a long tradition in applied mathematics and in particular homogenization [18]. It helps us build on the intuition of the simpler models while still capturing the first order effects.

We have shown that mixture models and stochastic models are not only interesting in their own right, and may resolve serious model mispecification issues, but also find interpretation within an approach that is motivated by de Finetti's Theorem. In this setting, mixture models with tractable convexity adjustments can help locate a model that is equivalent to nature's, at least as far as symmetric aggregate quantities are concerned.

In this approach it is not a requirement that the amount of variation implied by convexity adjustments be plausible purely from the point of view of population variation. And it may be possible to provide actionable intelligence without ever coming close to the true model.

From an analytical perspective it is clear from the size of the adjustments that care is required any time we have variation that reflects genuine heterogeneity. This is true of cross sectional comparisons also, if the amount of variation can be expected to vary. Population density may have only a weak relationship to contagion but because it varies so dramatically a convexity adjustment may be warranted.

Finally, we have appealed to the usefulness of growth models of a different kind, taken from the interest rate literature. Though this connection is limited to early stage growth the flexibility provided may be of benefit to epidemiology. Stochastic growth rates and complex term structures are a given in that field. So our hope is that this isn't the last use of Fixed Income modeling techniques for disease modeling.

6 Appendix A: Additional convexity adjustments

We provide further examples of convexity adjustments for peak infection, early stage growth and population profile.

6.1 Peak infection

Section 3.4.1 considers examples where η depends on population density ρ. Here are some example where it does not.

6.1.1 Inverse Gamma

We suppose the law of η is an Inverse Gamma distribution with parameters $\alpha > 1$ and β chosen such that the mean $\frac{\beta}{\alpha-1} = 1$. That is to say $\beta = \alpha - 1$. The mean of the inverse is α/β and thus

$$J(\alpha) = E\left[\frac{1}{\eta}\right] = \frac{\alpha}{\alpha - 1}$$

In passing, we note that this example might be formalized by letting the probability space be $\Omega = (0, 1)$ with uniform ρ equivalent to Lebesgue measure and η the inverse cumulative distribution function for the Inverse Gamma function. In similar fashion other examples can be made to satisfy the definition we provided for the variation function η, although when η does not depend on ρ this may seem a tad ceremonial.

6.1.2 Lognormal

Suppose instead the law of η is log-normally distributed, which is to say that with mild abuse of notation

$$\eta \sim e^{\mu + \sigma Z}$$

for standard normal Z. The mean is

$$E[\eta] = e^{\mu + \frac{1}{2}\sigma^2}$$

whereas the harmonic mean is

$$H(\mu) = e^{\mu - \frac{1}{2}\sigma^2}$$

The ratio J is the ratio of the mean to the harmonic mean and thus

$$J(\sigma) = e^{\sigma^2}$$

We set $\sigma^2 = -2\mu$ so that the mean is 1. So alternatively we can write

$$J(\mu) = e^{-2\mu}$$

bearing in mind $\mu < 0$.

6.1.3 Pareto

Fix $\alpha > 1$ and let $\sigma = 1 - \frac{1}{\alpha}$. Assume the law of η follows a Pareto distribution, which is to say that the probability it exceeds x is $\left(\frac{x}{\sigma}\right)^\alpha$ for $x > \sigma$. We have

$$J = \overbrace{\frac{\alpha\sigma}{\alpha - 1}}^{mean(=1)} \Big/ \overbrace{\sigma\left(1 + 1/\alpha\right)}^{harmonic\ mean} = \frac{1}{1 - \frac{1}{\sigma^2}}$$

6.2 Alternative harmonic mean of gaussian

The following approximation of the harmonic mean of a gaussian not centered near zero is provided by Dimitri Offengenden. One starts with the following approximation

$$\frac{1}{1 + x} \approx 1 - x + \frac{3}{4}x^2 - x^3$$

which Offengenden finds is approximately valid for $x \in (0, 1)$ with error no more than one to two percent, and at the same time quite accurate for larger values of x (fitting considerably better than the Taylor expansion, for example). Let $\{\epsilon_i\}_{i=1}^n$ be independent standard normal. In the limit $n \to \infty$ we have

$$
\begin{aligned}
h(0, 1, \infty; \mu, \sigma) &= \frac{1}{E\left[\frac{1}{X} | X \sim N(\mu, \sigma^2)\right]} = \frac{1}{E\left[\frac{1}{\mu + \sigma Z} | Z \sim N(0, 1)\right]} \\
&\approx \mu\left(1 + \frac{\sigma}{\mu}\frac{1}{n}\sum_{i=1}^n \epsilon_i - \frac{1}{n}\sum_{i=1}^n \frac{3}{4}\frac{\sigma^2}{\mu^2}\epsilon_i^2 + \frac{1}{n}\sum_{i=1}^n \frac{1}{4}\frac{\sigma^3}{\mu^3}\epsilon_i^3\right) \\
&\to \mu\left(1 - \frac{3}{4}\frac{\sigma^2}{\mu^2}\right)
\end{aligned}
$$

where ϵ_i are independent standard normal.

6.3 Peak infection with modified harmonic mean

In Section 4.9 we considered an approximation using a modified harmonic mean h. There are no doubt many ways to approximate this. Here we make two observations only

1. It suffices to consider the case of standard normal

2. Option pricing technology can be used

Recall the auxiliary function that generalizes the harmonic mean:

$$h(a, b, c; \mu, \sigma) := \frac{1}{E\left[\frac{1}{a+b\max(X,c)} \big| X \sim N(\mu, \sigma^2)\right]}$$

Let $Y := \frac{X - \mu_X}{\sigma_X}$ be an affine transformed version of X. As Y is manifestly a standard normal random variable, and since

$$\begin{aligned}
\max(X, 1) &= \max(Y\sigma_X + \mu_X, 1) \\
&= \mu_X + \max(Y\sigma_X, 1 - \mu_X) \\
&= \mu_X + \sigma_X \max\left(Y, \frac{1 - \mu_X}{\sigma_X}\right)
\end{aligned}$$

it is clear that

$$h(1, 0, 1, \mu, \sigma) = E\left[\frac{1}{\mu_X + \sigma_X \max\left(Y, \frac{1-\mu_X}{\sigma_X}\right)}\right] = h\left(\mu_x, \sigma_x, \frac{1 - \mu_x}{\sigma_x}; 0, 1\right)$$

allowing one to specialize to the case of standard normal random variables when approximating h. Furthermore

$$\begin{aligned}
h(a, b, c; 0, 1) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}x^2}}{a + b\max(x, c)} dx & (23) \\
&= \frac{1}{2\pi} \frac{1}{a + bc} \int_{-\infty}^{c} e^{-\frac{1}{2}x^2} dx + \frac{1}{2\pi} \int_{c}^{\infty} \frac{e^{-\frac{1}{2}x^2}}{a + bx} dx & (24)
\end{aligned}$$

where the first term is merely a multiple of the cumulative normal distribution function. In Section ?? we considered the case $c = -\infty$ where only the second term survives. The technique used there to approximate $\frac{1}{a+bx} = \frac{1}{a}\frac{1}{1+\frac{b}{a}x}$ could be applied to convert the second term into a sum of normal options.[22] This might be overkill, admittedly, and the reader likely won't need reminding that this class of approximations assumes independent equilibria are reached between populations - likely an assumption leading to greater error.[23]

---

[22]We again thank Dimitri Offengenden for the suggestion.

[23]In future work we hope to use these convexity adjustments as features for predicting actual peak infection levels in agent simulations, thereby providing some calibration useful for actual epidemics.

6.4 Early stage growth

In order to provide more tractable examples of early stage growth adjustments due to variation, it obviously behoves us to consider variation functions with known moment generating functions. These examples are standard. We merely enforce normalization, which sometimes reveal bounds on the growth convexity adjustments. Recall that we use

$$t' = \bar{\beta}t$$

to measure time in units that are convenient for this purpose.

6.4.1 Gamma

Let $\Omega = (0, 1)$ with uniform measure ρ and define η by the inverse cumulative distribution function of the Gamma distribution. Then the law of η is Gamma distributed and

$$E[e^{t'\eta}] = (1 - t\theta)^{-k} \tag{25}$$

where k and θ are the Gamma distribution parameters. We must set $k = 1/\theta$ to achieve a mean $E[\eta] = 1$. The probability density function is

$$\frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-x/\theta}$$

but will only serve us if θ is small enough that $t' < 1/\theta$, for otherwise Equation 25 is invalid. If it is valid then

$$G(t) = -k\frac{log(1 - \bar{\beta}t/k)}{\bar{\beta}t} \approx 1 + \frac{1}{2}\theta\bar{\beta}t < \frac{3}{2}$$

where the inequality follows from the limitation $t' < 1/\theta$ corresponding to $t < \frac{1}{\bar{\beta}\theta}$.

6.4.2 Poisson

Similarly we can allow η to be Poisson distributed with parameter $\lambda = 1$. Then

$$E[e^{t'\eta}] = e^{e^{t'}-1} \tag{26}$$

and thus

$$G(t) = \frac{e^{\bar{\beta}t} - 1}{\bar{\beta}t} \approx 1 + \frac{1}{2}\bar{\beta}t + \frac{1}{6}\bar{\beta}^2 t^2 + O(t^3)$$

Excess growth is unbounded. However a more careful analysis would curtail growth for some values of η where the approximation $s(t) \approx 1$ is no longer valid.

6.4.3 Negative binomial

Let η be distributed as the the negative binomial distribution with parameters r and p, and with $p = \frac{1}{1+r}$ so that $E[\eta] = 1$. Then

$$E[e^{t'\eta}] = \frac{(1-p)^r}{(1-pe^{t'})^r} \tag{27}$$

and so

$$G(t) = r \frac{\log \frac{1-p}{1-pe^{\bar{\beta}t}}}{\bar{\beta}t}$$

is the excess growth.

6.4.4 Binomial

With η Binomial(n,p) and $p = 1/n$ a similar calculation yields

$$G(t) = \frac{\log(1 - p + pe^{\bar{\beta}t})}{\bar{\beta}t} \approx 1 + \frac{1}{2}t\bar{\beta}(1-p) + \frac{1}{6}\bar{\beta}^2(2p^2 - 3p + 1)t^2 + O(t^3)$$

6.4.5 Linear

Let $\Omega = (0,2)$ with $\eta(\omega) = \omega$ and $\rho(\omega) = \frac{1}{2}$ so that $E[\eta] = 1$ as required. Since

$$E[e^{t\bar{\beta}\eta}] = \frac{e^{2\bar{\beta}t} - 1}{2\bar{\beta}t}$$

we have

$$G(t) = \frac{1}{\bar{\beta}t} \log\left(\frac{e^{2\bar{\beta}t} - 1}{2\bar{\beta}t}\right) \approx 1 + \frac{1}{6}\bar{\beta}t$$

in agreement with the previous gaussian approximation for small t after we note that the variance of η is $\sigma^2 = 1/3$. But when t is several multiples of of the timescale $1/\beta$ the limit $t \gg 1/\bar{\beta}$ becomes more relevant. We have

$$G(t) \to 2$$

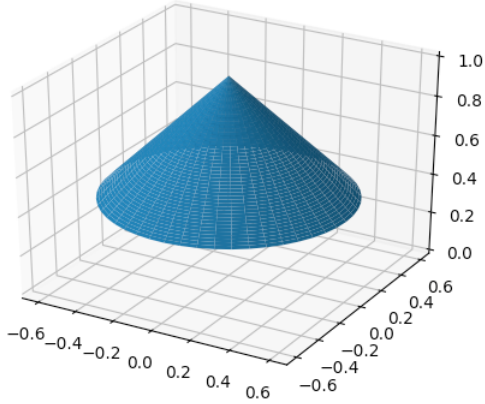Figure 6: An example of a stylized population density profile resembling a circus tent.

6.4.6 Quadratic

In similar fashion we could quadratic η. We set $\Omega = (0, \sqrt{3})$, $\rho = \frac{1}{\sqrt{3}}$ and $\eta(\omega) = \omega^2$. Then $E[\eta] = 1$ and $E[(\eta - 1)^2] = \frac{4}{5}$. We find

$$E[e^{t\bar{\beta}\eta}] \;\; = \;\; \frac{1}{\sqrt{3}} \int_0^{\sqrt{3}} e^{\bar{\beta}t\omega^2} d\omega$$

Using a series expansion at $\bar{\beta}t \to \infty$ we approximately have

$$\frac{\log E[e^{t\bar{\beta}\eta}]}{t} \to \frac{1}{t} \log\left(e^{3\bar{\beta}t} \left\{ \frac{1}{\bar{\beta}t} + \frac{1}{36\bar{\beta}^2 t^2} \right\} \right) \to 3\bar{\beta}$$

showing that if the limit is relevant, the growth inflated by recovery is three times what it would be β were to be held constant.

$$G(t) \to 3$$

6.5 Additional population profile adjustments

Following on from Section 3.4.1 we provide more examples of convexity adjustments when contagion varies as a power law of population.

6.6 Linear drop in population density

The city's population takes the form of a circus tent shown in Figure 6.

Here we assume the city has a radius of b and the tent's roof, pitched at a height 1, falls off at a 45 degree angle. Quite unlike a circus tent it is assumed that the height of this roof represents population density, which drops abruptly to zero as we reach a distance b from the center of town.[24] A calculation analogous to those above yields

$$J(a,b) = \frac{\frac{1}{(a^2-5a+6)} - \frac{(1-b)^{2-a}(1+b(2-a))}{(2-a)(3-a)}}{\frac{1}{a^2+5a+6} - \frac{(1-b)^{(a+2)}((a+2)*b+1)}{(2+a)(3+a)}}$$

The multiplicative correction takes values on the order of 1.5 for $a \approx 1/4$ and $b \approx 0.9$, for example. This function is plotted in Figure 7

6.6.1   Power law population distribution

An example when J is unbounded is provided by a "skirt". Suppose $\rho(r) \propto 1/r^3$ and the population lives outside the unit circle. Then

$$J(\alpha) = \frac{\int_1^\infty \frac{1}{r^3(1-\alpha)} r dr}{\int_1^\infty \frac{1}{r^3(1+\alpha)} r dr} = \frac{1+3a}{1-3a} \to \infty$$

It is of note that the correction is inaccurate beyond small values of α. The size of the immune group cannot exceed unity and this is not taken into account.

6.7   Gamma distributed populations

With density proportional to re^{-r} and $\beta(r)$ the α power of the same,

$$
\begin{aligned}
J_1(\alpha) &= \frac{\int_0^\infty (rexp(-r))^{1+\alpha} r dr}{\int_0^\infty (rexp(-r))^{1-\alpha} r dr} \\
&= \frac{(1+\alpha)^{-\alpha-3}\Gamma(3+\alpha)}{(1-\alpha)^{\alpha-3}\Gamma(3-\alpha)} \\
&\approx (3/2+\gamma\alpha)^2 + O(\alpha^3) \\
&\approx 1 + 4.15\alpha
\end{aligned}
$$

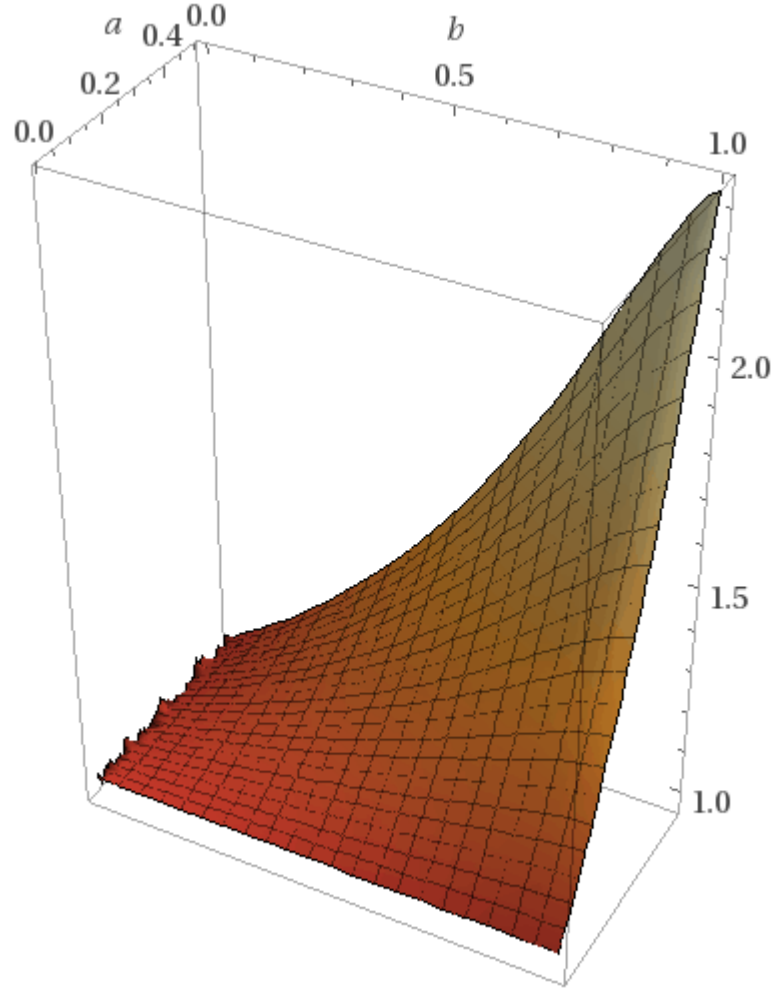[24]That is to say that all citizens reside in a conical version of the Luxor Hotel.

Figure 7: Convexity adjustment to the susceptible population as a function of parameters a and b determining contagiousness in a tent-shaped population profile. Population density falls off linearly as we move away from the origin until we reach a radius $b < 1$ as shown in Figure 6. Contagiousness is related to density by a power law with coefficient $a \ll 1$. Even a weak relationship between population density and contagiousness can imply a significant ratio between the number of susceptible people at peak infection and the number of susceptible people at peak infection when we assume homogeneous density.

where γ is the Euler-Mascheroni constant, $\gamma \approx 0.57721$. If instead density is proportional to $r^2 e^{-r}$ then

$$
\begin{aligned}
J_2(\alpha) &= \frac{\int_0^\infty (r^2 exp(-r))^{1+\alpha} r \, dr}{\int_0^\infty (r^2 exp(-r))^{1-\alpha} r \, dr} \\
&= \frac{(1+\alpha)^{-4-2a}\Gamma(4+2\alpha)}{(1-\alpha)^{2a-4}\Gamma(4-2\alpha)} \\
&\approx (1 - (1/3 + 2\gamma)\alpha)^2 + O(\alpha^3) \\
&\approx 1 + 2.97\alpha
\end{aligned}
$$

Notice that these adjustment are material even if the power law α relating population density to contagiousness is relatively small, say $\alpha \approx 0.1$.

7 Appendix B: Justifying decaying mean-reverting infection rate

We provide further motivation for the choice of Ornstein-Uhlenbeck infection rates

7.1 Form suggested on phenomenological basis

Chowell et al [5] propose modeling infection rates as time varying:

$$\beta(t) = \beta_0 \left\{ (1 - \phi)e^{-qt} + \phi \right\} \tag{28}$$

We refer the reader to the extensive empirical survey by these authors that led them to suggest this model for early stage growth. The reader will recognize this as the mean of an Ornstein-Uhlenbeck process.

7.2 Repeat contact in a continuous physical model

In [6] the probability of novel contact in an infinite agents model is shown to take the form

$$\beta(t; \beta_0, \alpha_0) = \beta_0 \left(\overbrace{\frac{1 - e^{-\alpha_0 t}}{\alpha_0 t}}^{Ein(\alpha_0 t)} \right)$$

This model assumes that agents populate the plane $\mathcal{R}^2$ and make visits to nearby locations with probability proportional to a bivariate gaussian distribution. When probability of repeat contact is taken into account we can approximate the system by a compartmental model with delay differential equation for infection. In this approach the infection rate is multiplied by an average over the infected cohort

$$\bar{P}(t; \alpha) = \frac{\int_{s=0}^{t} Ein(\alpha(t - s)) \overbrace{\left[\frac{\partial i(s)}{\partial s} + \gamma i(s) \right]}^{new\ infections} e^{-\gamma(t-s)}ds}{\int_{s=0}^{t} \left[\frac{\partial i(s)}{\partial s} + \gamma i(s) \right] e^{-\gamma(t-s)}ds}$$

Because this is a vintage effect it may be seen to fall within a broad class of models considered by Kermack and McKendrick [1], [11] [12]. This particular choice leads to infection rates resembling the patterns seen in Figure 8, which can be seen to be reasonably well approximated by our choice. The reader is referred to the paper for further details [6].
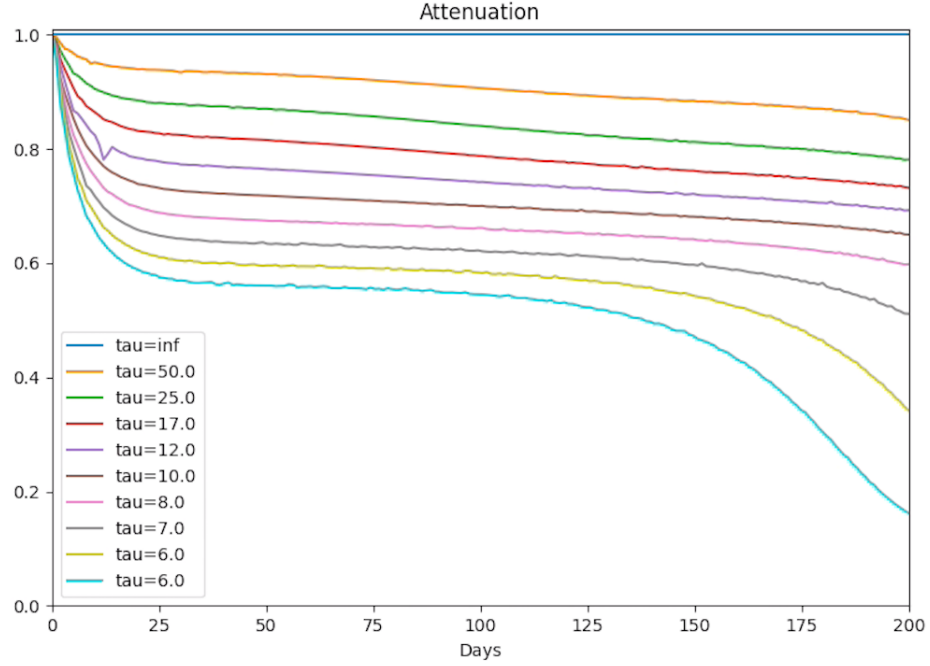
Figure 8: A physical infinite agents model provides further motivation for choice of decaying mean-revering infection rate. In [6] the variation in infection rate $\beta(t)$ revealed by numerical solution of delay differential equations in a modified SIR compartmental model. The form taken accounts for repeat contacts. It can be seen that infection rate is reasonably well approximated by an Ornstein-Uhlenbeck process - at least for early stage growth and up until the time of peak infection - the period relevant to convexity adjustments presented here (in the late stage of the epidemic a shifting in the vintage of the infected cohort may lead to a second drop in infection rate).

References

[1] A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 1927.

[2] Mohsen Ahmadi, Abbas Sharifi, Shadi Dorosti, Saeid Jafarzadeh Ghoushchi, and Negar Ghanbari. Investigation of effective climatology parameters on COVID-19 outbreak in Iran. *Science of the Total Environment*, 2020.

[3] Tom Britton. Basic estimation-prediction techniques for Covid-19, and a prediction for Stockholm. Technical report, 2020.

[4] C. Chen, J. T. Yan, N. Zhou, J. P. Zhao, and D. W. Wang. Analysis of myocardial injury in patients with COVID-19 and association between concomitant cardiovascular diseases and severity of COVID-19. *Zhonghua xin xue guan bing za zhi*, 2020.

[5] Gerardo Chowell, Lisa Sattenspiel, Shweta Bansal, and Cécile Viboud. Mathematical models to characterize early epidemic growth: A review, 2016.

[6] Peter Cotton. Repeat Contact and the Spread of Disease. 2020.

[7] S. Y. Del Valle, J. M. Hyman, H. W. Hethcote, and S. G. Eubank. Mixing patterns between age groups in social networks. *Social Networks*, 2007.

[8] Elamin H. Elbasha and Jagpreet Chhatwal. Characterizing Heterogeneity Bias in Cohort-Based Models. *PharmacoEconomics*, 2015.

[9] Matthew J. Ferrari, Ali Djibo, Rebecca F. Grais, Nita Bharti, Bryan T. Grenfell, and Ottar N. Bjornstad. Rural-urban gradient in seasonal forcing of measles transmission in Niger. In *Proceedings of the Royal Society B: Biological Sciences*, 2010.

[10] B. T. Grenfell and B. M. Bolker. Cities and villages: Infection hierarchies in a measles metapopulation. *Ecology Letters*, 1998.

[11] W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics-II. The problem of endemicity. *Bulletin of Mathematical Biology*, 1991.

[12] W. O. Kermack and A. G. McKendrick. Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity. *Bulletin of Mathematical Biology*, 1991.

[13] Danielle Klinger, Ido Blass, Nadav Rappoport, and Michal Linial. Significantly Improved COVID-19 Outcomes in Countries with Higher BCG Vaccination Coverage: A Multivariable Analysis. *medRxiv*, 2020.

[14] Wladyslaw J Kowalski. 2020 COVID-19 Coronavirus Ultraviolet Susceptibility 2020 COVID-19 Coronavirus Ultraviolet Susceptibility. *purple sun*, 2020.

[15] Gabriel E. Leventhal, Huldrych F. Günthard, Sebastian Bonhoeffer, and Tanja Stadler. Using an epidemiological model for phylogenetic inference reveals density dependence in hiv transmission. *Molecular Biology and Evolution*, 2014.

[16] Chris J.K. MacQuarrie and Barry J. Cooke. Density-dependent population dynamics of mountain pine beetle in thinned and unthinned stands. *Canadian Journal of Forest Research*, 2011.

[17] Hiroshi Nishiura, Gerardo Chowell, Muntaser Safan, and Carlos Castillo-Chavez. Pros and cons of estimating the reproduction number from early epidemic growth rate of influenza A (H1N1) 2009. *Theoretical Biology and Medical Modelling*, 2010.

[18] James Nolen, George Papanicolaou, and Olivier Pironneau. A framework for adaptive multiscale methods for elliptic problems. *Multiscale Modeling and Simulation*, 2008.

[19] Jonathan M. Read, Ken T.D. Eames, and W. John Edmunds. Dynamic social networks and the implications for the spread of infectious disease. *Journal of the Royal Society Interface*, 2008.

[20] Carmen Rolandi and Pablo E. Schilman. The costs of living in a thermal fluctuating environment for the tropical haematophagous bug, Rhodnius prolixus. *Journal of Thermal Biology*, 2018.

[21] Patrick M. Tarwater and Clyde F. Martin. Effects of population density on the spread of disease. *Complexity*, 2001.

[22] Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 1977.

[23] J. Wallinga and M. Lipsitch. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society B: Biological Sciences*, 2007.

[24] Jingyuan Wang, Ke Tang, Kai Feng, and Weifeng Lv. High Temperature and High Humidity Reduce the Transmission of COVID-19. *SSRN Electronic Journal*, 2020.

[25] Marta L. Wayne, Gabriela M. Blohm, Mollie E. Brooks, Kerry L. Regan, Brennin Y. Brown, Michael Barfield, Robert D. Holt, and Benjamin M. Bolker. The prevalence and persistence of sigma virus, a biparentally transmitted parasite of Drosophila melanogaster. *Evolutionary Ecology Research*, 2011.

[26] Chaomin Wu, Xiaoyan Chen, Yanping Cai, Jia'an Xia, Xing Zhou, Sha Xu, Hanping Huang, Li Zhang, Xia Zhou, Chunling Du, Yuye Zhang, Juan Song, Sijiao Wang, Yencheng Chao, Zeyong Yang, Jie Xu, Xin Zhou, Dechang Chen, Weining Xiong, Lei Xu, Feng Zhou, Jinjun Jiang, Chunxue Bai, Junhua Zheng, and Yuanlin Song. Risk Factors Associated with Acute Respiratory Distress Syndrome and Death in Patients with Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Internal Medicine*, 2020.

[27] Yu Wu, Wenzhan Jing, Jue Liu, Qiuyue Ma, Jie Yuan, Yaping Wang, Min Du, and Min Liu. Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Science of the Total Environment*, 2020.

[28] Nan Yu, Wei Li, Qingling Kang, Zhi Xiong, Shaoshuai Wang, Xingguang Lin, Yanyan Liu, Juan Xiao, Haiyi Liu, Dongrui Deng, Suhua Chen, Wanjiang Zeng, Ling Feng, and Jianli Wu. Clinical features and obstetric and neonatal outcomes of pregnant patients with COVID-19 in Wuhan, China: a retrospective, single-centre, descriptive study. *The Lancet Infectious Diseases*, 2020.

[29] Jin jin Zhang, Xiang Dong, Yi yuan Cao, Ya dong Yuan, Yi bin Yang, You qin Yan, Cezmi A. Akdis, and Ya dong Gao. Clinical characteristics of 140 patients infected with SARS-CoV-2 in Wuhan, China. *Allergy: European Journal of Allergy and Clinical Immunology*, 2020.

[30] L. Zhang, Y. Jiang, M. Wei, B. H. Cheng, X. C. Zhou, J. Li, J. H. Tian, L. Dong, and R. H. Hu. Analysis of the pregnancy outcomes in pregnant women with COVID-19 in Hubei Province. *Zhonghua fu chan ke za zhi*, 2020.