

# Failure of monotonicity in epidemic models

Maria Chikina and Wesley Pegden

April 2020

## Abstract

We discuss the failure of monotonicity properties for even simple compartmental epidemic models, for the case where transmission rates are non-constant. We also identify a special case in which monotonicity holds.

## 1 Introduction

Consider a simple SIR (Susceptible, Infected, and Recovered/Removed) model governed by the differential equations

$$\frac{dI}{dt} = I\beta S \cdot \frac{1}{m} - \alpha I \tag{1}$$

$$\frac{dS}{dt} = -I\beta S \cdot \frac{1}{m} \tag{2}$$

$$\frac{dR}{dt} = \alpha I. \tag{3}$$

$$\tag{4}$$

When the transmission coefficient  $\beta$  and the recovery coefficient  $\alpha$  are constant (as is the total population  $m$ ), the final recovered population  $R(\infty) := \lim_{t \rightarrow \infty} R(t)$  (which is the total number of people ever infected) can be found to satisfy the transcendental equation

$$R(\infty) = 1 - S(0)e^{-\frac{\beta}{\alpha m}(R(\infty) - R(0))},$$

by using the chain rule to write

$$\frac{dS}{dR} = -\frac{\beta S}{\alpha m},$$

separating the variables, and integrating. From this it follows that  $R(\infty)$  behaves monotonically with respect to the basic reproduction number  $R_0 := \frac{\beta}{\alpha}$ ; that is, the larger  $R_0$ , the larger  $R(\infty)$ .

The purpose of this note is consider monotonicity properties obeyed by the system in a context where the transmission rate is not a constant but also a

function of time (because of effects of seasonality, for example, or the effects of mitigations). We will see that natural-seeming monotonicity properties fail to hold for even this simplest of epidemic models.

For an SIR model where the transmission rate  $\beta : [0, \infty] \rightarrow \mathbb{R}$  is a function of time, consider the following three monotonicity properties one might hope would be satisfied for such a system, if  $\beta_1$  and  $\beta_2$  are two transmission rate functions, and  $R_1(\infty)$  and  $R_2(\infty)$  are the recovered populations resulting from SIR models with  $\beta_1$  and  $\beta_2$ , respectively.

**Pointwise monotonicity:** If  $\beta_1(t) \leq \beta_2(t)$  for all  $t$ , then  $R_1(\infty) \leq R_2(\infty)$ .

**Shift monotonicity:** If  $\beta_1(t) = \beta_2(t) - \varepsilon$  for all  $t$  for some constant  $\varepsilon > 0$ , then  $R_1(\infty) \leq R_2(\infty)$ .

**Scaling monotonicity:** If  $\beta_1(t) = (1 - \varepsilon)\beta_2(t)$  for all  $t$  for some constant  $\varepsilon > 0$ , then  $R_1(\infty) \leq R_2(\infty)$ .

In this note we show that **all three of these forms of monotonicity can fail**, in general, for SIR models where the transmission rate  $\beta(t)$  is not constant.

On the other hand, we show that pointwise monotonicity *does* hold, in the special case that transmission rates never increase:

**Theorem 1.1.** *Consider functions  $\beta_1, \beta_2$  which are both bounded below by some positive constant. If  $\beta_1(t) \leq \beta_2(t)$  for all  $t$  and  $\beta_1(t)$  is monotone nonincreasing, then an SIR model satisfying (1), (2), (3), governed by the transmission rate function  $\beta := \beta_1$  ends with a greater number of susceptible individuals than when governed by  $\beta := \beta_2$  (assuming identical initial conditions).*

We remark that the lack of monotonicity we point out in this note could be more severe in more complex, realistic epidemic models, where, for example, the presence of heterogeneous transmission and mortality rates can lead to substantial benefits to increasing transmission rates for some populations [1].

## 2 Failures of monotonicity

In the following sections we present examples to demonstrate the failure of each type of monotonicity. The examples are generated using code available at the entry for this manuscript at <http://math.cmu.edu/~wes/pub.html>.

### 2.1 Failure of pointwise monotonicity

The failure of pointwise monotonicity for SIR models is well-known; see [2] for a recent reference discussing examples. A standard example can be constructed by taking

$$\beta_1(t) := \begin{cases} 1 & t < 100 \\ 2 & t \geq 100 \end{cases},$$

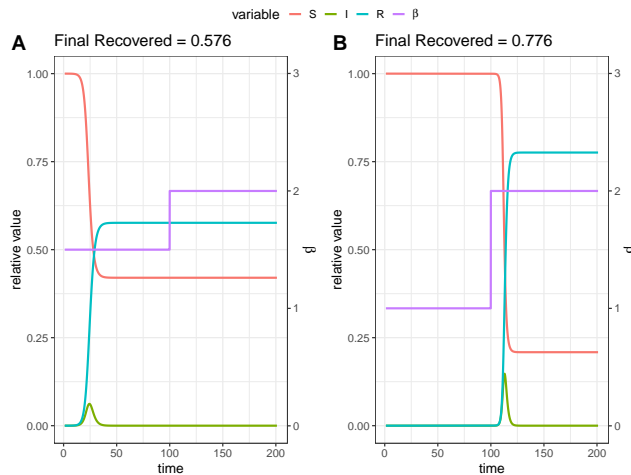


Figure 1: Failure of pointwise monotonicity. The scale for  $S$ ,  $I$ , and  $R$  is on the left, and for  $\beta$ , on the right.

while

$$\beta_2(t) := \begin{cases} 1.5 & t < 100 \\ 2 & t \geq 100 \end{cases}.$$

For convenience we scale time so that  $\alpha = 1$ , here and for the remaining examples, giving the correspondence  $\beta = R_0$ . The result is shown in Figure 1.

Qualitatively, this particular example is typically explained qualitatively by noting the fact that  $\beta_2(t) > \beta_1(t)$  for  $t \in [0, 100]$  results in an initial epidemic in the second scenario, which, while smaller than the epidemic for  $\beta = 2$  seen in the first scenario, is nevertheless large enough to achieve enough population immunity to resist transmission when  $\beta_2(t) = 2$ .

However, it is important to note that there is no general principle asserting that this simple example is the only way in which monotonicity can fail; instead, this simple examples underlines the fact that pointwise monotonicity is not guaranteed by even the simplest forces at play in an epidemic's growth. As we will see, monotonicity can fail in ways which are more surprising than this first example might suggest.

### 3 Failure of shift monotonicity

Perhaps more surprising than the failure of pointwise monotonicity is the fact that even consistent changes to the transmission function over time can be counterproductive for reducing epidemic size.

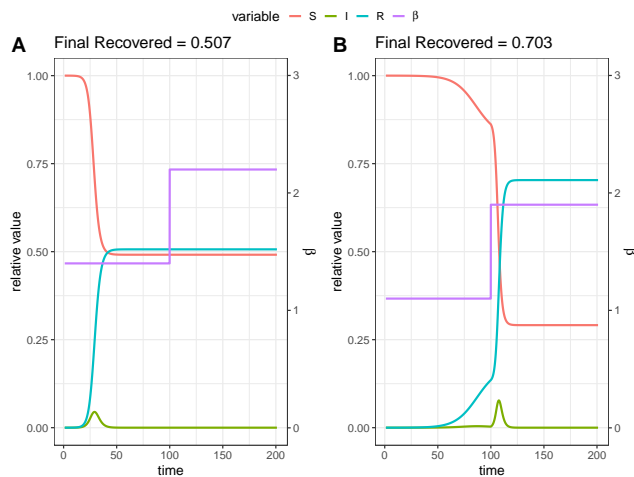


Figure 2: Failure of shift monotonicity. The scale for  $S$ ,  $I$ , and  $R$  is on the left, and for  $\beta$ , on the right.

Consider epidemics governed by the transmission functions

$$\beta_1(t) := \begin{cases} 1.1 & t < 100 \\ 1.9 & t \geq 100 \end{cases},$$

while

$$\beta_2(t) := \begin{cases} 1.4 & t < 100 \\ 2.2 & t \geq 100 \end{cases}.$$

Observe that  $\beta_1(t) = \beta_2(t) - 0.3$  for all  $t$ . (Recall we have set  $\alpha = 1$ .) Nevertheless, more than 50% more total infections occur in an epidemic governed by  $\beta_1$  instead of  $\beta_2$ , as we see in Figure 2. As we will see next, the same surprising result can occur from a geometric scaling, rather than an arithmetic shift.

## 4 Failure of scaling monotonicity

Here we see the same type of example but where the global change in transmission rates is geometric.

We let

$$\beta_2(t) := \begin{cases} 1.5 & t < 100 \\ 2.5 & t \geq 100 \end{cases}.$$

and  $\beta_1 = .7 \cdot \beta_2$ . The result is in Figure 3.

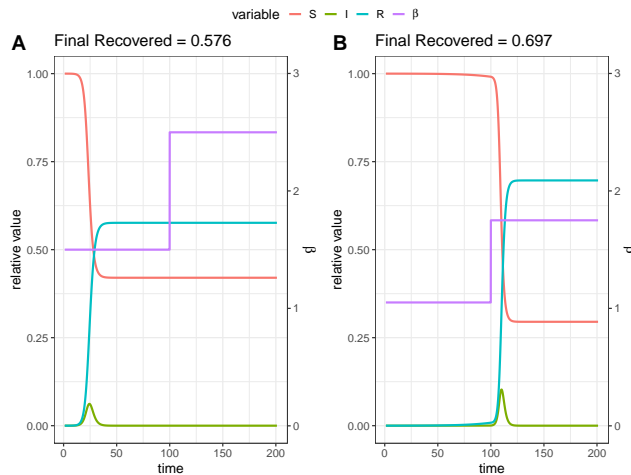


Figure 3: Failure of scale monotonicity. The scale for  $S$ ,  $I$ , and  $R$  is on the left, and for  $\beta$ , on the right.

## 5 Failure of shift for naturalistic transmission variation

To produce examples which are as simple as possible, we have focused on step-functions in the previous sections. This may give the incorrect impression that failure of monotonicity only occurs in the special case that there will be some sudden resumption of high transmission rates in the future. In fact it is easy to construct examples where monotonicity fails even when transmission rates are changing only slowly over time. In this section, we give an example motivated by the possibility of seasonal effects on viral transmission, using a shift of a sine-wave function as the time-varying transmission rate  $\beta$ .

We use

$$\beta_2(t) := \frac{\sin t/20 + 1}{2} + 1.9$$

and

$$\beta_1(t) := \frac{\sin t/20 + 1}{2} + 2.1.$$

Again, we see that a small consistent *decrease* in transmission rates in this scenario can result in more total infections (Figure 4).

## 6 Monotonicity when $\beta$ is decreasing

In this section we prove Theorem 1.1. For this purpose, let the functions  $S_i, I_i, R_i$  of  $t$  for  $i = 1, 2$  be the two SIR models satisfying the systems corresponding to (1), (2), (3) for the given  $\beta_i$ . Note that for simplicity Theorem

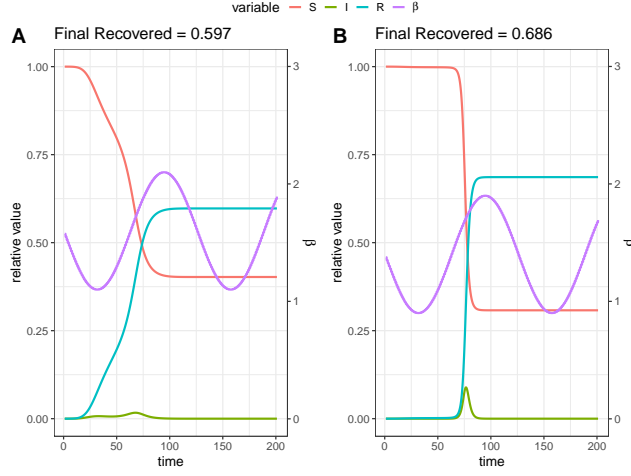


Figure 4: Failure of shift monotonicity for naturalistic variation. The scale for  $S$ ,  $I$ , and  $R$  is on the left, and for  $\beta$ , on the right.

1.1 assumes that the differential equalities hold everywhere, which means that the  $\beta_i$  are necessarily continuous. For simplicity of notation we let  $m = 1$  without loss of generality, so that  $I$ ,  $S$ , and  $R$  are all ratios between 0 and 1.

As is standard, we can view these models under derivatives with respect to  $S$  instead of  $t$ : Beginning from (1), (2), (3), we can apply to the chain rule to get

$$\frac{dI}{dS} = \frac{\alpha}{\beta \cdot S} - 1. \quad (5)$$

We write  $B = 1 - S$  to write

$$\frac{dI}{dB} = 1 - \frac{\alpha}{\beta \cdot (1 - B)}. \quad (6)$$

Considering  $I$  and  $\beta$  as a functions of  $B$ , and assuming for simplicity that the number of initially recovered individuals is 0, the choice of initial value for  $I$  is made by choosing some  $B^{\text{init}} < 1$  such that  $I(B^{\text{init}}) = B^{\text{init}}$ . (In the case where the initial number  $r_0$  of recovered individuals is nonzero, set  $I(B^{\text{init}}) = B^{\text{init}} - r_0 > 0$ .)

Note that we already see from (5) that the SIR model *does* satisfy pointwise monotonicity with respect to  $\beta$  as a function of  $B$ ; the spectacular failure of monotonicity for SIR models occurs precisely because of the interaction between  $\beta$  and the correspondence between  $S$  and  $t$ .

To proceed, we consider the derivative

$$\frac{dt}{dB} = \frac{1}{I \cdot \beta \cdot (1 - B)}. \quad (7)$$

Now we can define  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  by

$$f(I, t, B) = \left( 1 - \frac{\alpha}{\beta_1(t) \cdot (1 - B)}, \frac{1}{I \cdot \beta_1(t) \cdot (1 - B)} \right) \quad (8)$$

so that the SIR model governed by  $\beta_1$  satisfies

$$\frac{d}{dB}(I_1, t_1) = f(I_1, t_1, B).$$

Observe that  $f$  is Lipschitz in  $[B^{\text{init}}, B^{\text{end}}]$  since  $B$  is bounded away from 1, and  $\beta(t)$  and  $I$  are bounded away from 0 on this range.

We write  $f_I$  and  $f_t$  for the first and second coordinates of  $f$ , respectively. Observe that  $f_I$  is decreasing with respect to  $t$ , and  $f_t$  is decreasing with respect to  $I$ . In particular,

$$g(I, t, B) := \left( -f_I(-I, t, B), f_t(-I, t, B) \right)$$

is *quasimonotone* in  $(I, t)$ ; that is,  $g_I$  is increasing in  $t$  and  $g_t$  is increasing in  $I$ .

Moreover, we have that the SIR models governed by  $\beta_1$  and  $\beta_2$ , respectively, satisfy

$$\frac{d}{dB}(-I_1, t_1) = g(-I_1, t_1, B) \quad (9)$$

$$\frac{d}{dB}(-I_2, t_2) \leq g(-I_2, t_2, B), \quad (10)$$

where the inequality follows by our assumption that  $\beta_1(t) \leq \beta_2(t)$ .

As a result, from the relations (9), (10), together with  $I_1(B^{\text{init}}) \leq I_2(B^{\text{init}})$ ,  $t_1(B^{\text{init}}) \geq t_2(B^{\text{init}})$  and the Lipschitz condition for  $f$  (and so  $g$ ) imply that

$$\begin{aligned} I_1 &\leq I_2 \\ t_1 &\geq t_2 \end{aligned}$$

throughout; see [3], page 112, in particular version (b) of the *Comparison Theorem* on the same page.  $\square$

## References

- [1] M. Chikina, W. Pegden. Modeling strict age-targeted mitigation strategies for COVID-19. arXiv preprint arXiv:2004.04144 (2020).
- [2] A. Handel, I.M., Longini Jr, R. Antia. What is the best control strategy for multiple infectious disease outbreaks?, in *Proceedings. Biological sciences* **274** (2007) 833–837. <https://doi.org/10.1098/rspb.2006.0015>
- [3] W. Wolfgang. Ordinary differential equations. Springer-Verlag (1998).