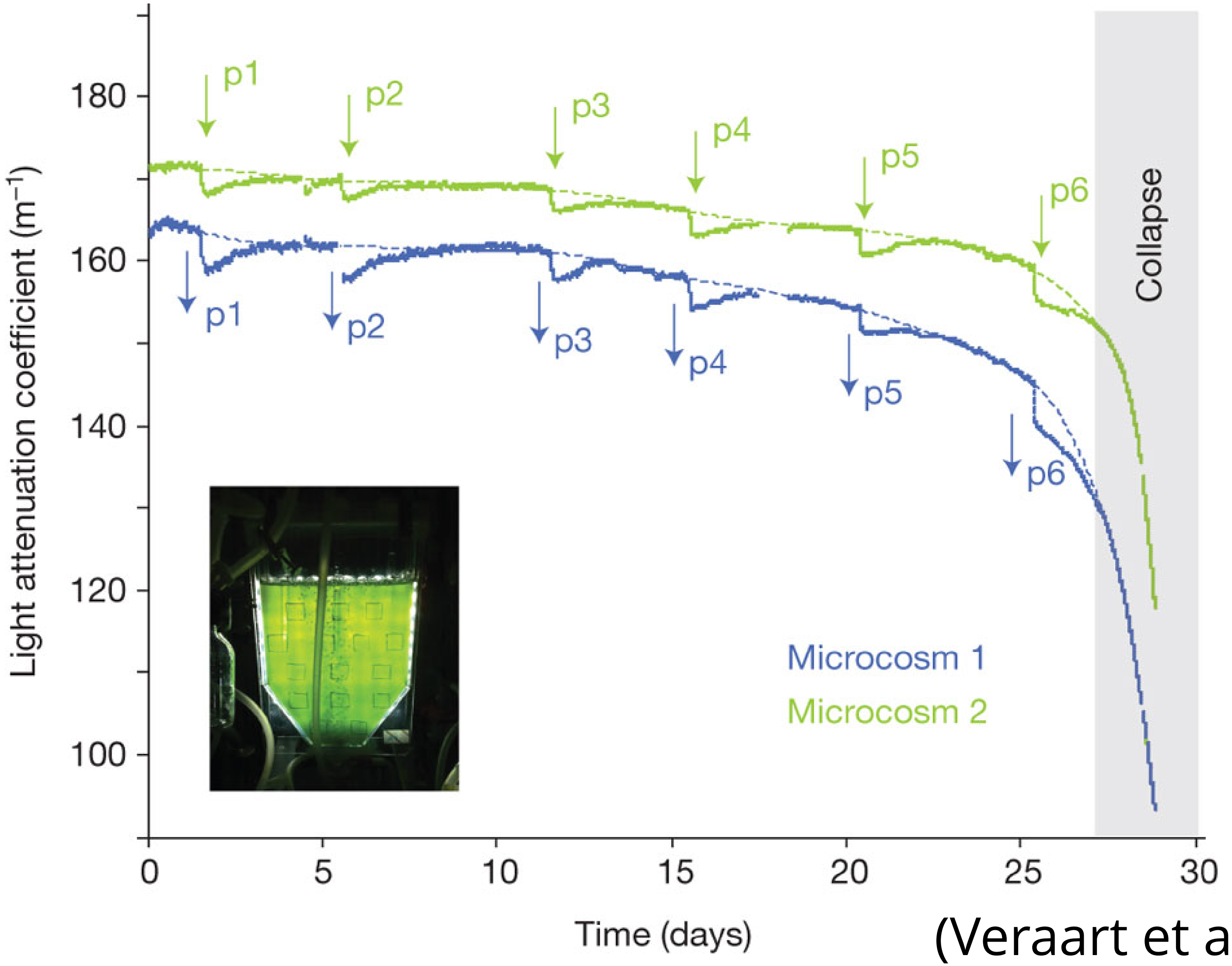
Disentangling reporting and disease transmission using second order statistics

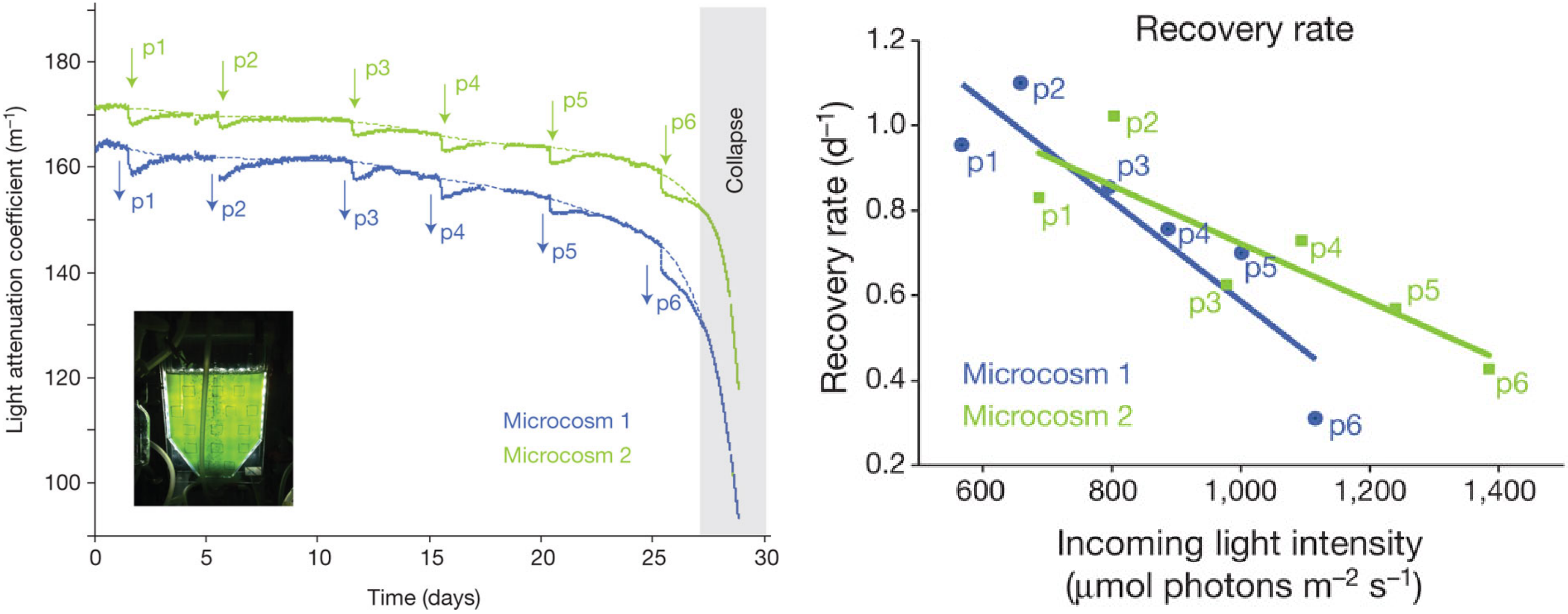
Eamon O'Dea and John Drake

Slower decay rates can indicate upcoming collapse



(Veraart et al. 2012)

Slower decay rates can indicate upcoming collapse



(Veraart et al. 2012)

Slower decay rates can indicate upcoming collapse

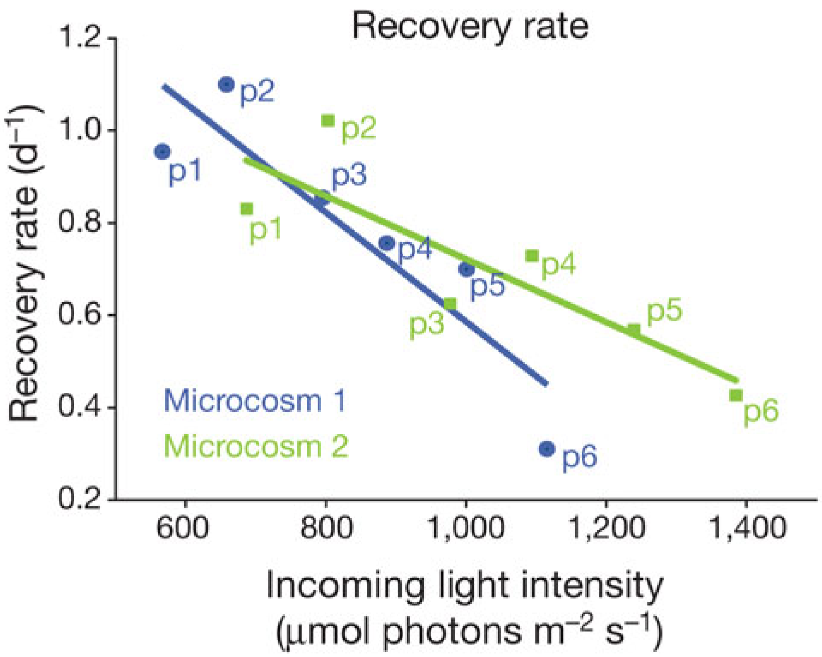
$$\dot{x} = r - x^2$$

$$x^* = \pm\sqrt{r} \text{ if } r > 0$$

$$\frac{\mathrm{d}\dot{x}}{\mathrm{d}x}\Big|_{x=\sqrt{r}} = -2\sqrt{r}$$

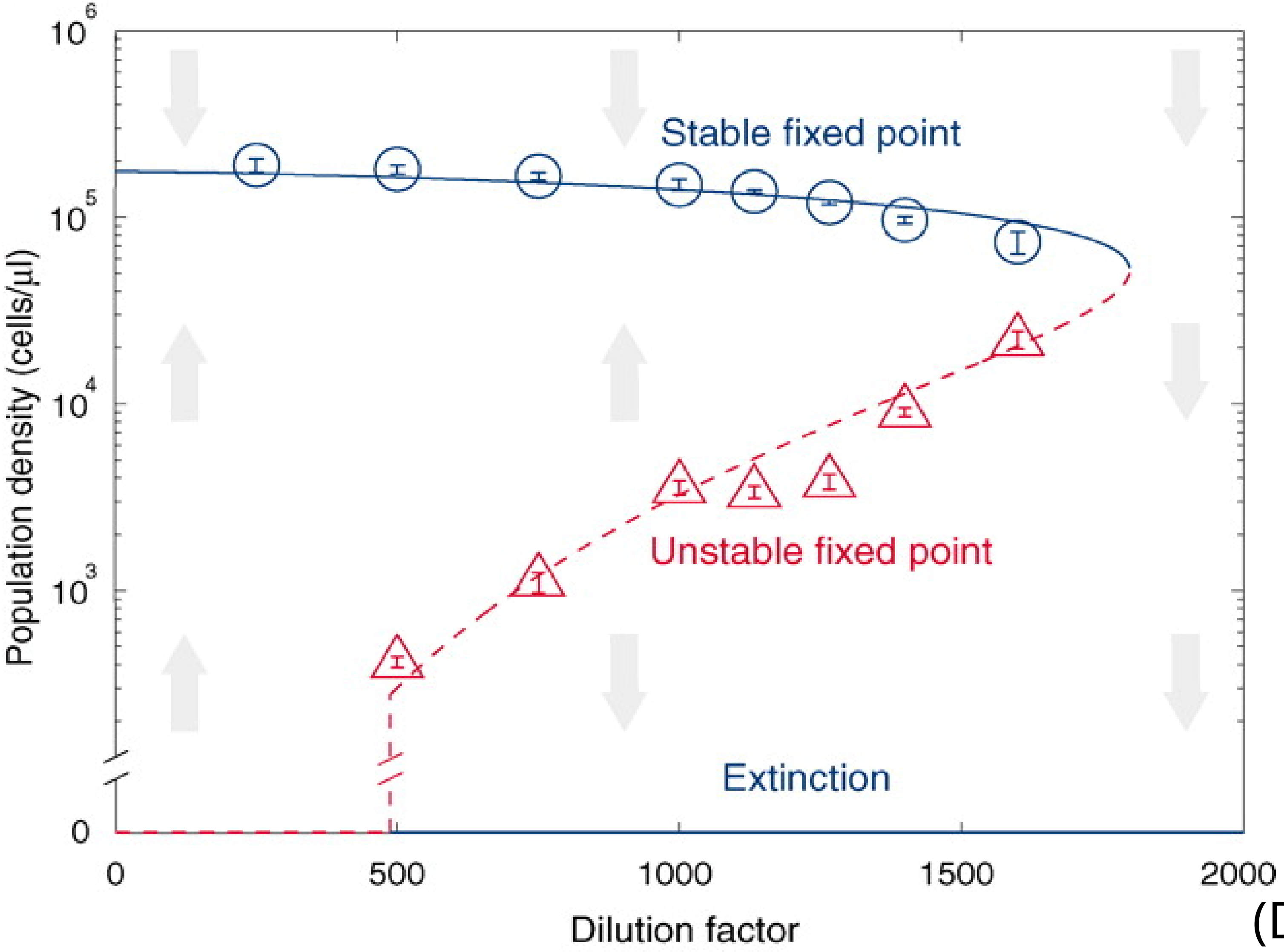
$$p := \text{perturbation}$$

$$\dot{p} = -2\sqrt{r}p$$



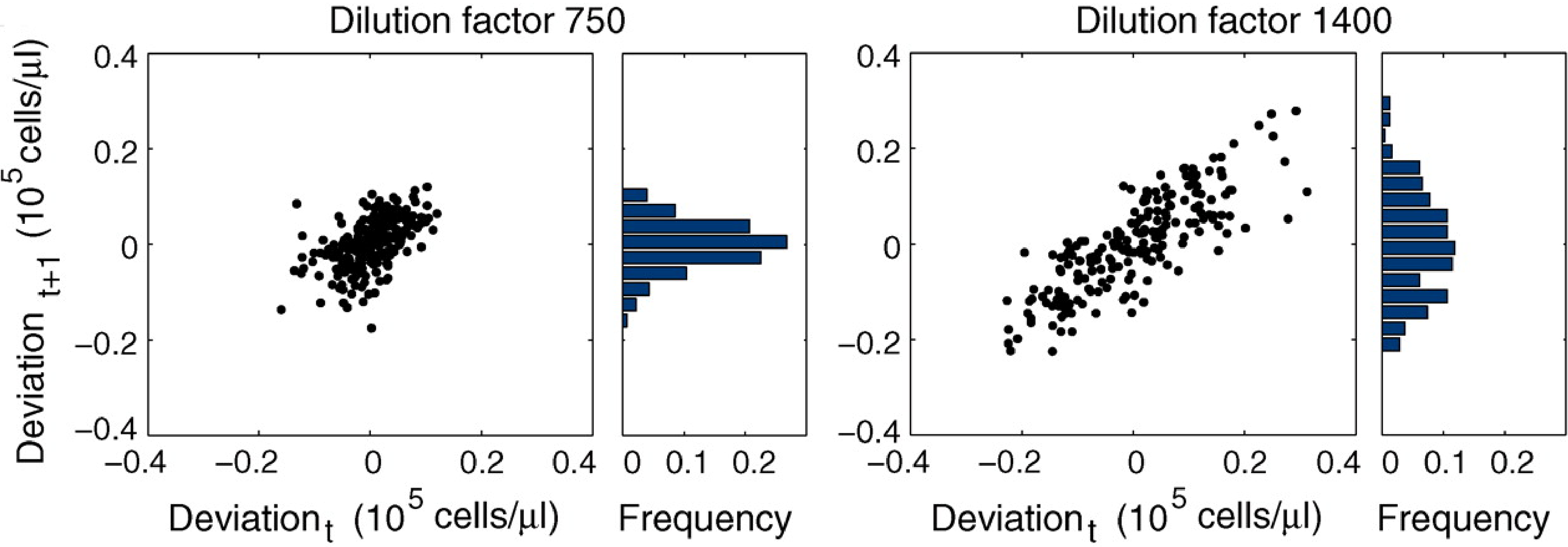
(Veraart et al. 2012)

Yeast system exemplifying a model of collapse



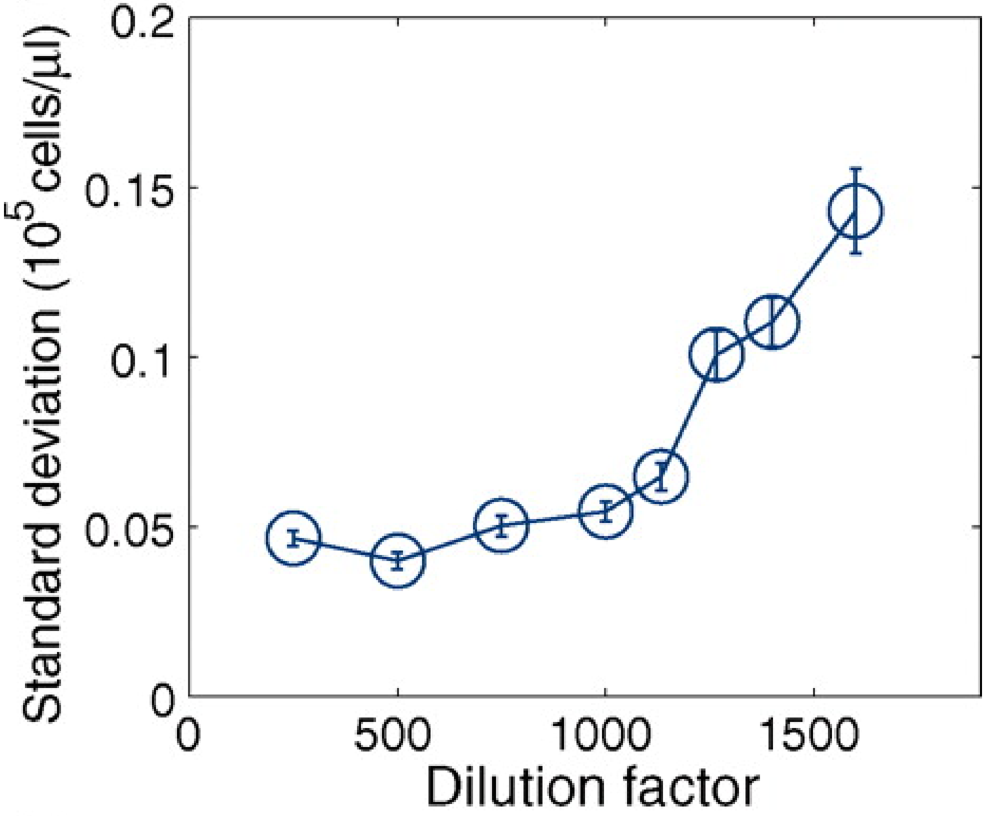
(Dai et al. 2012)

Spread of deviations depended on distance to bifurcation point



(Dai et al. 2012)

Spread of deviations depended on distance to bifurcation point



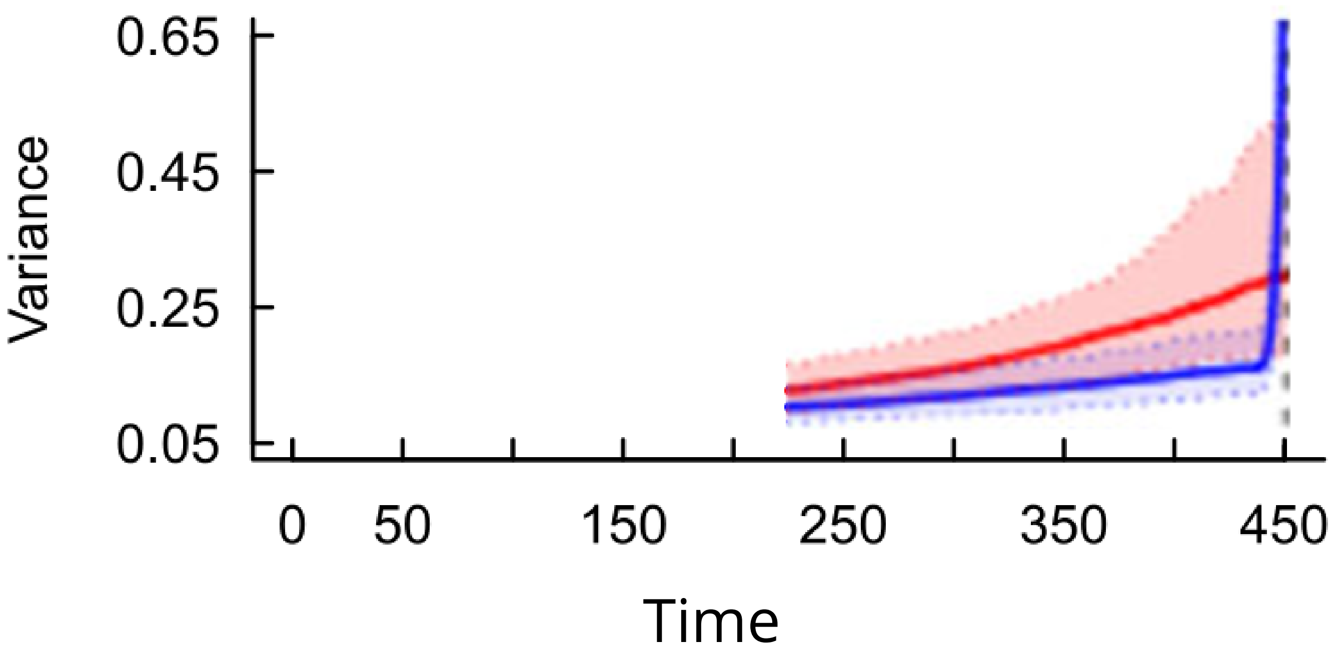
Second order

$$\langle p(t)^2 \rangle = \sigma^2 / (4\sqrt{r})$$

$$\mathrm{d}p = -2\sqrt{r}p\mathrm{d}t + N(t,\sigma)\sqrt{\mathrm{d}t}$$

(Dai et al. 2012)

Variance also increased in a model of disease emergence

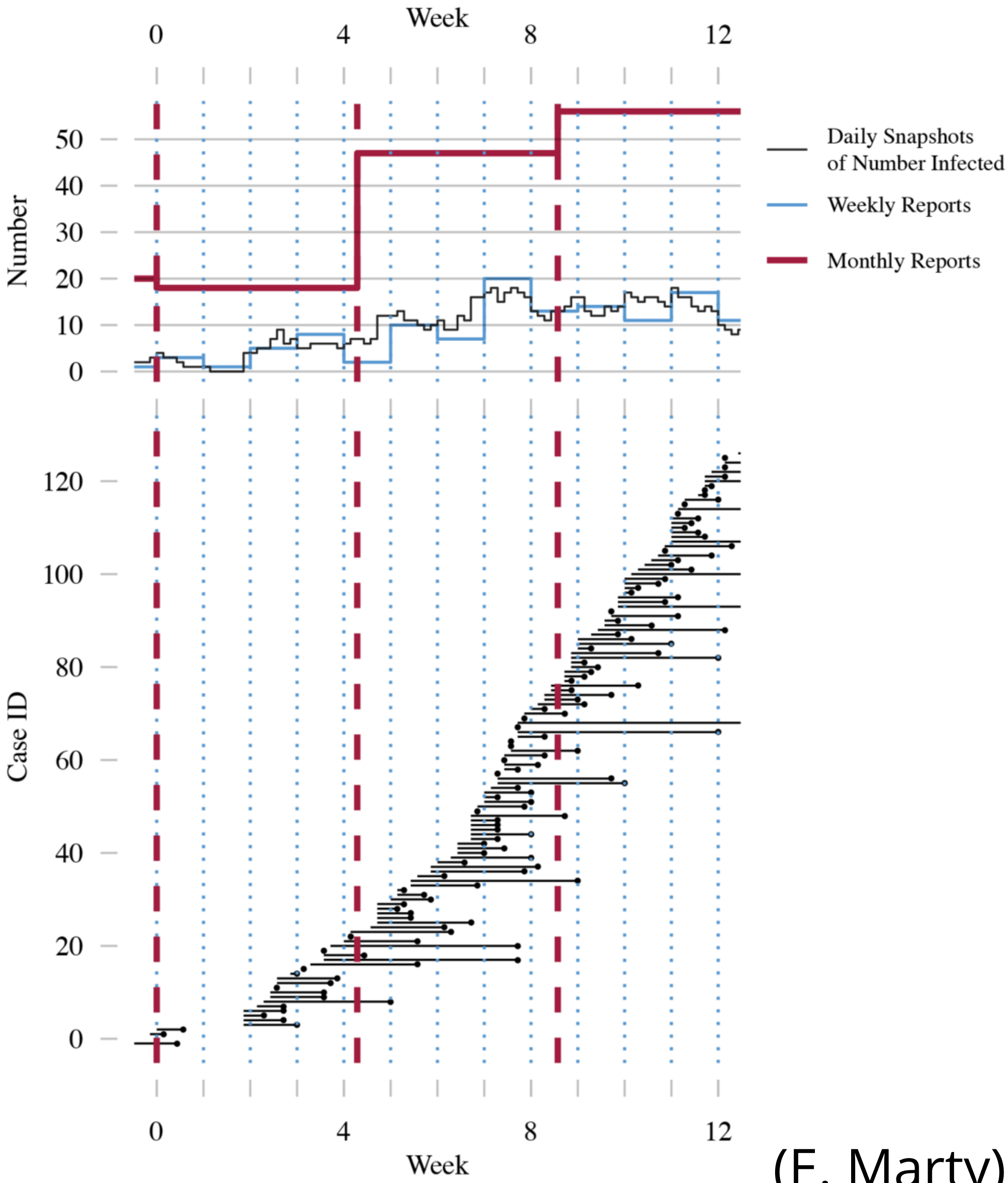


$$\mathrm{d}p = -rp\mathrm{d}t + N(t,\sigma)\sqrt{\mathrm{d}t}$$
$$\langle p(t)^2 \rangle = \sigma^2/(2r)$$

(O'Regan and Drake 2013)

Questions

1.Will indicators based on case data behave similarly?

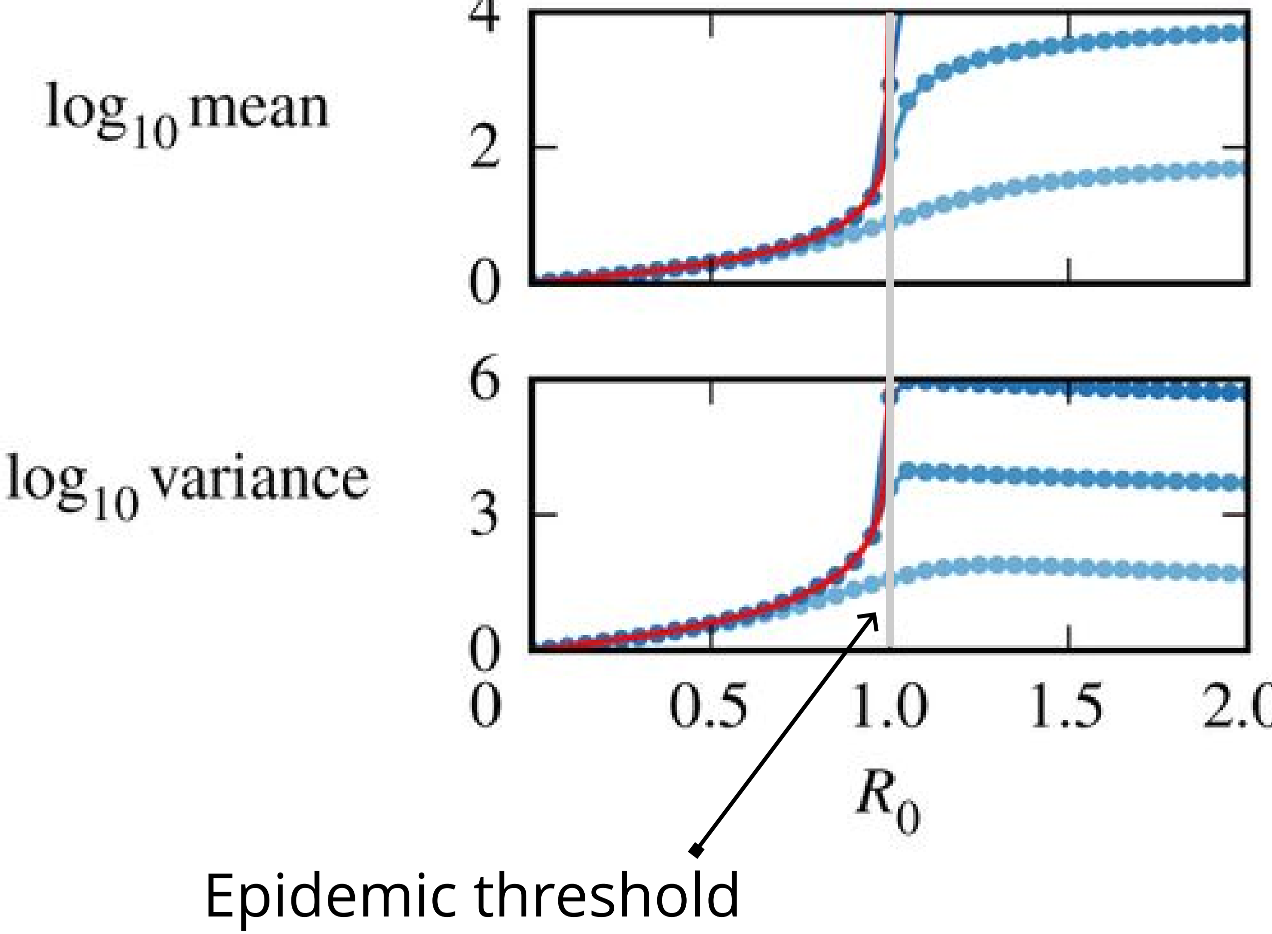


(E. Marty)

Questions

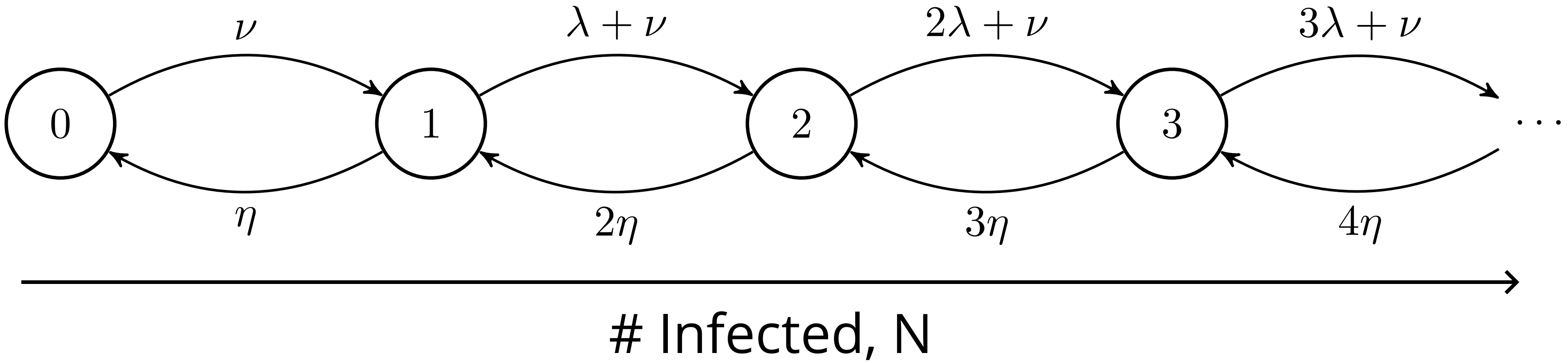
1. Will indicators based on case data behave similarly?

2. What advantages do second-order indicators have over a rolling mean?



(Brett et al. 2017)

We modeled disease spread with a BDI process

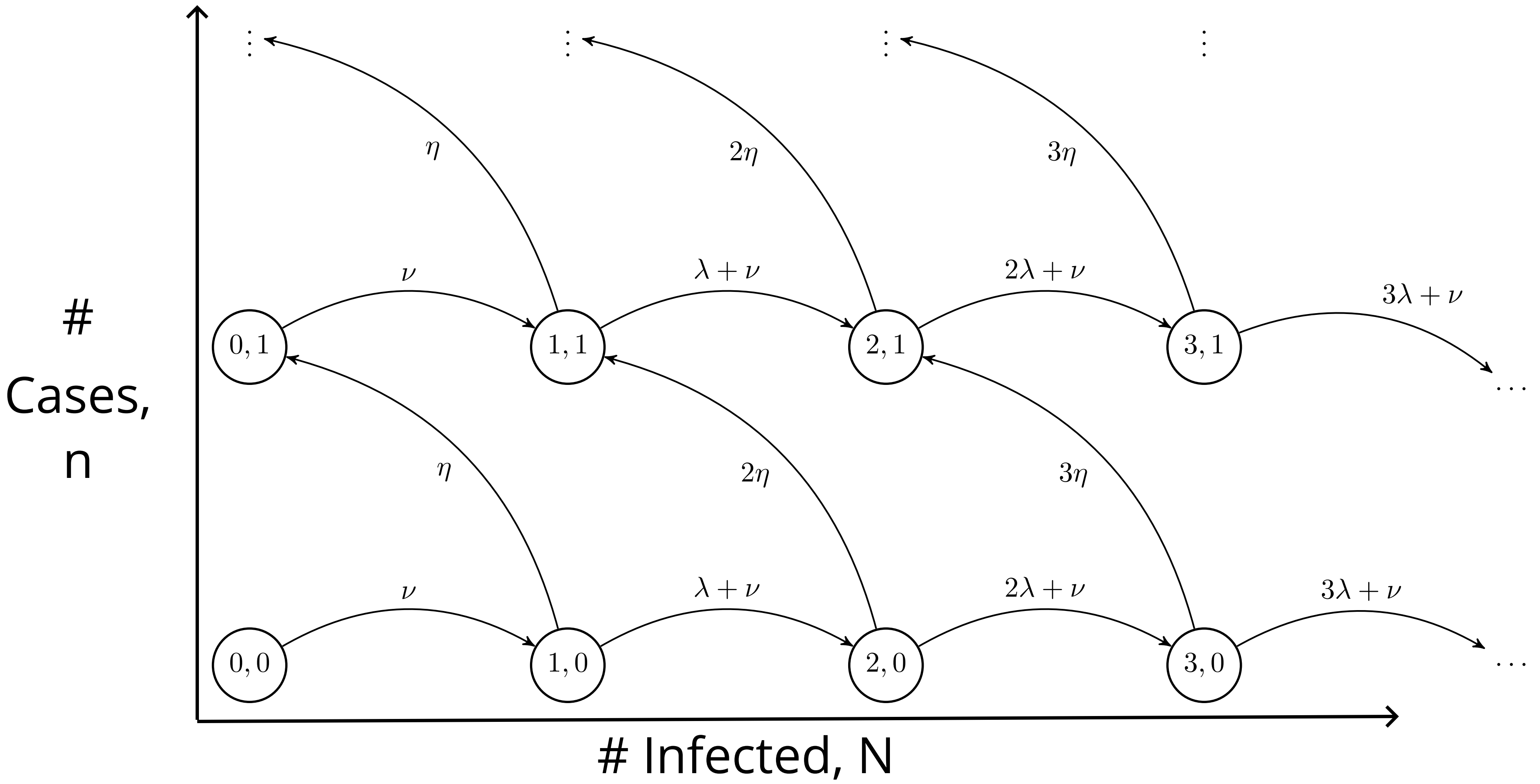


$$\lambda := \text{transmission (Birth) rate}$$

$$\eta := \text{removal (Death) rate}$$

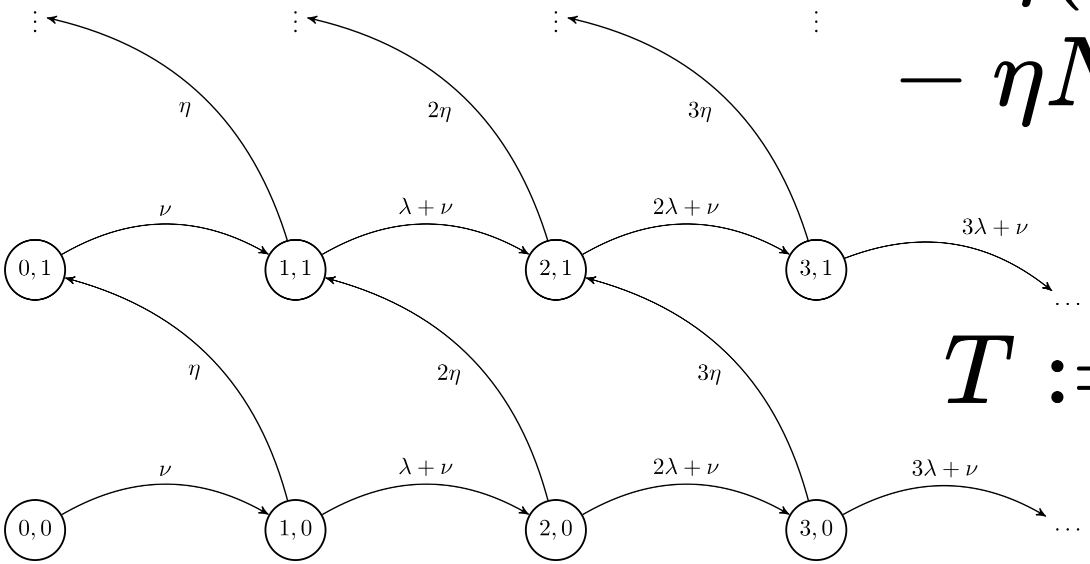
$$\nu := \text{importation (Immigration) rate}$$

Cases were modeled as the count of removals



Cases were modeled as the count of removals

$$\begin{aligned}\frac{\mathrm{d}P_{N,n}}{\mathrm{d}T} = &[\lambda(N-1)+\nu]P_{N-1,n} \\ &-[\lambda N P_{N,n}+\nu]P_{N,n} \\ &+\eta(N+1)P_{N+1,n-1} \\ &-\eta N P_{N,n}\end{aligned}$$



$$T := \text{reporting period}$$

Moments of cases *n* were found with generating functions

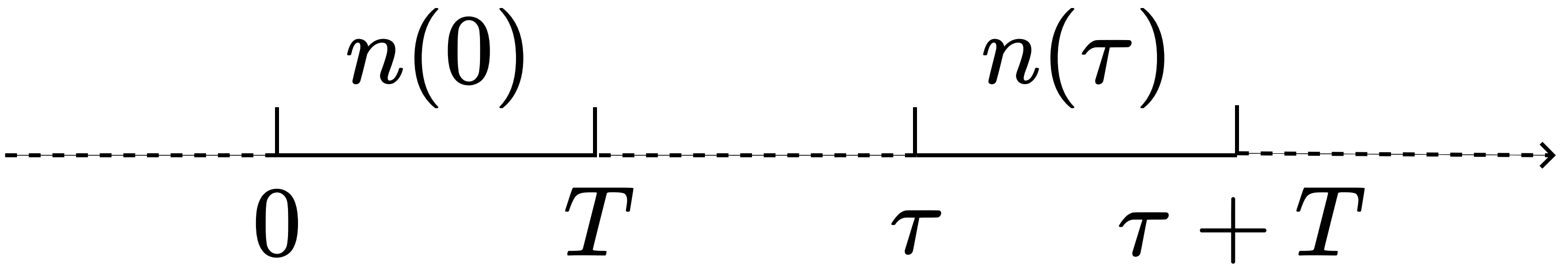
$$q(s,z;T) := \sum_{N,n}(1-s)^N(1-z)^n P_{N,n}(T)$$

$$\begin{aligned}\langle n\rangle &= \eta T\nu/(\eta-\lambda)\\ n^{[2]} &:= \langle n(n-1)\rangle/\langle n\rangle^2\\ &= 1+\tfrac{2\lambda}{\nu(\eta-\lambda)T}\left(1-\tfrac{1-\exp((\lambda-\eta)T)}{(\eta-\lambda)T}\right)\end{aligned}$$

Decreasing functions of distance to threshold: $\eta-\lambda$

(Hopcraft et al. 2014)

The autocorrelation of *n* was also found via a bilinear moment



$$g(\tau;T) := \frac{\langle n(0)n(\tau)\rangle}{\langle n\rangle^2} \qquad \gamma := T(\eta-\lambda)/2$$

$$g(\tau;T) = 1 + \frac{\lambda}{\gamma^2\nu}\sinh^2(\gamma)\exp(-(\eta-\lambda)\tau)$$

$$\begin{aligned}\rho(\tau,T) &:= [\langle n(0)n(\tau)\rangle - \langle n\rangle^2]/\operatorname{var} n \\ &= \frac{g(\tau;T)-1}{n^{[2]}-1+\langle n\rangle^{-1}}.\end{aligned}$$

Decreasing functions of distance to threshold: $\eta - \lambda$

Questions

1. Will indicators based on case data behave similarly?

2. What advantages do second-order indicators have over a rolling mean?

Answers

1. Many should also increase as threshold approached:
   - mean
   - factorial moment
   - decay time of autocorrelation

2. None. But what if not all cases are reported?

Some moments of reports are unaffected by reporting prob.

$$m := \text{number of reports}$$
$$\xi := \text{reporting probability}$$

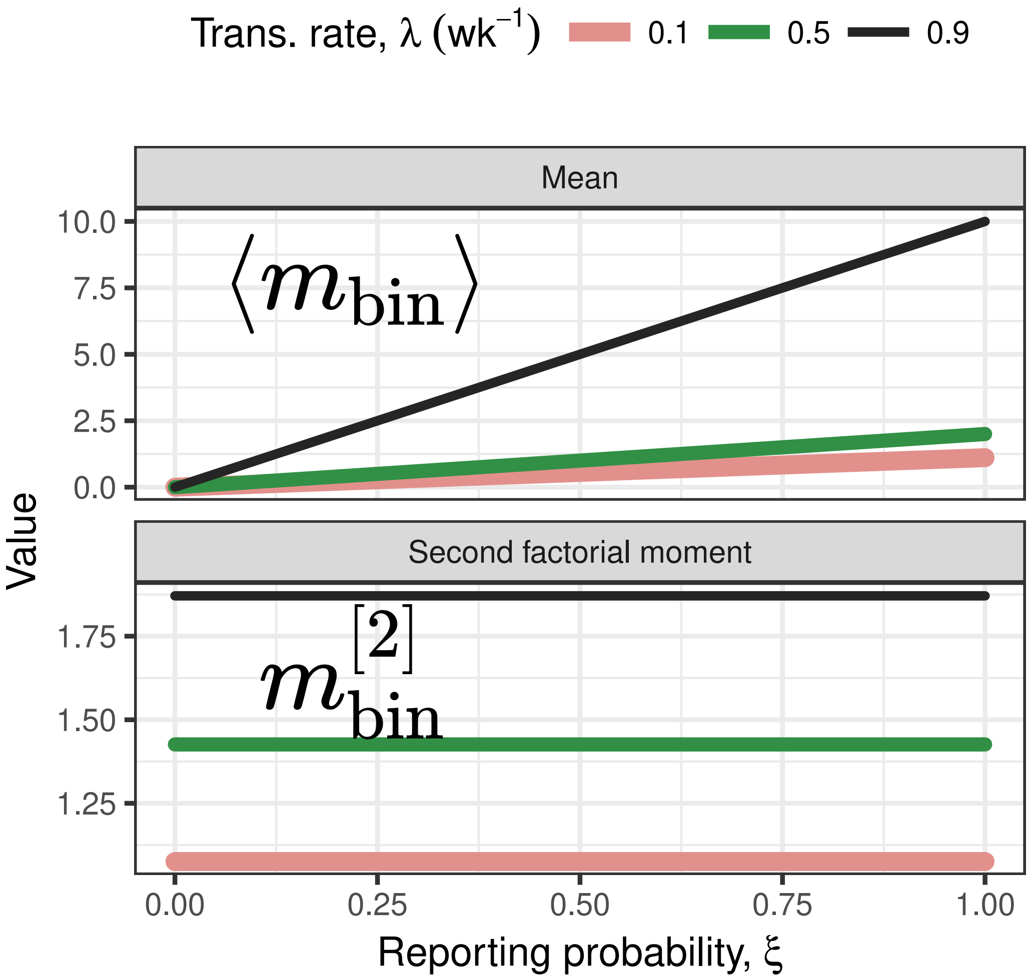
$$q_{\text{bin}}(s, z; T) = q(s, z\xi; T)$$
$$\langle m_{\text{bin}} \rangle = \xi \langle n \rangle$$
$$m_{\text{bin}}^{[2]} = n^{[2]}$$

(Hopcraft et al. 2014)

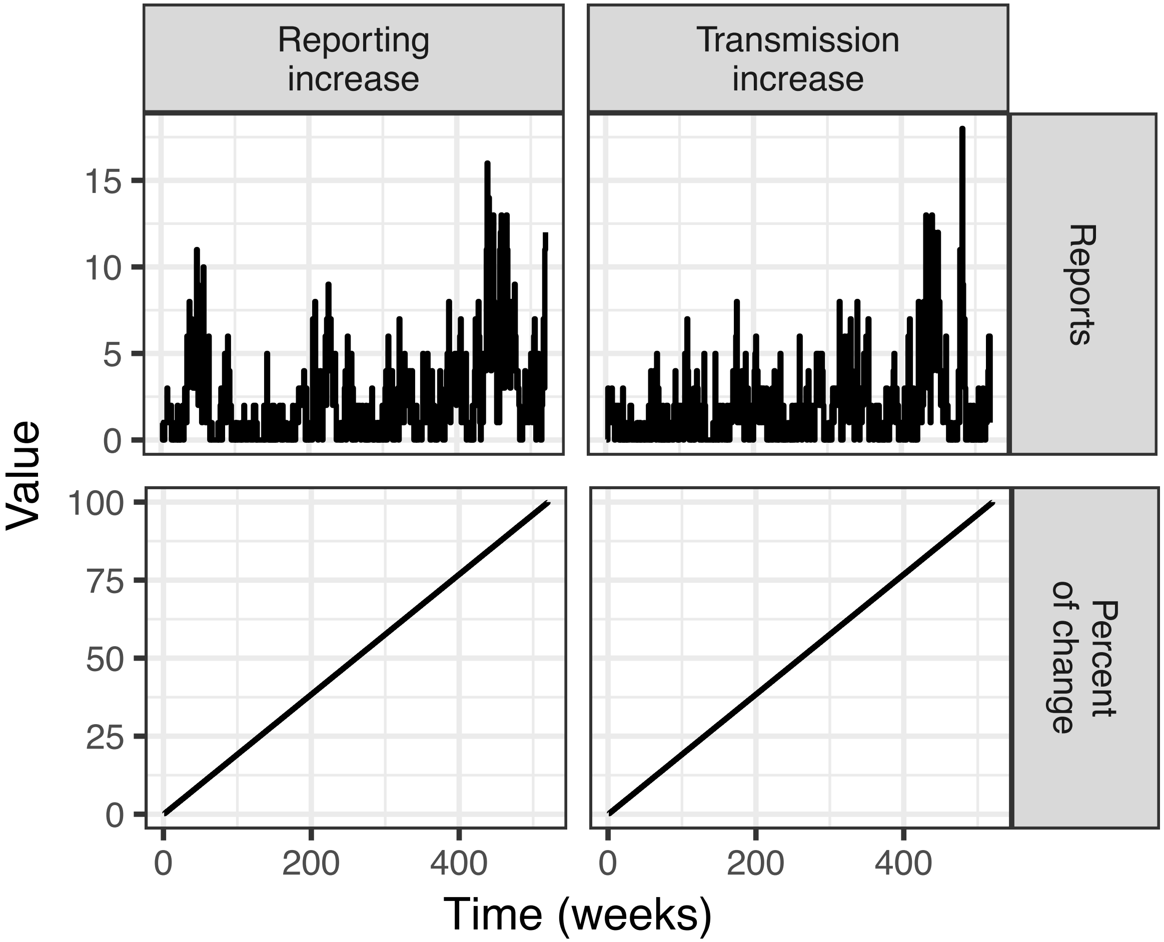
Same result holds when allowing for overdispersion

$$\begin{aligned} m_{\mathrm{nb}} &:= \xi n + e \\ e &:= \text{error term} \\ \phi &:= \text{dispersion parameter} \\ \langle e^2 | n \rangle &= \xi n + (\xi n)^2/\phi \\ \langle m_{\mathrm{nb}} \rangle &= \xi \langle n \rangle \\ m_{\mathrm{nb}}^{[2]} &= (1 + 1/\phi)(n^{[2]} + 1/\langle n \rangle) \end{aligned}$$

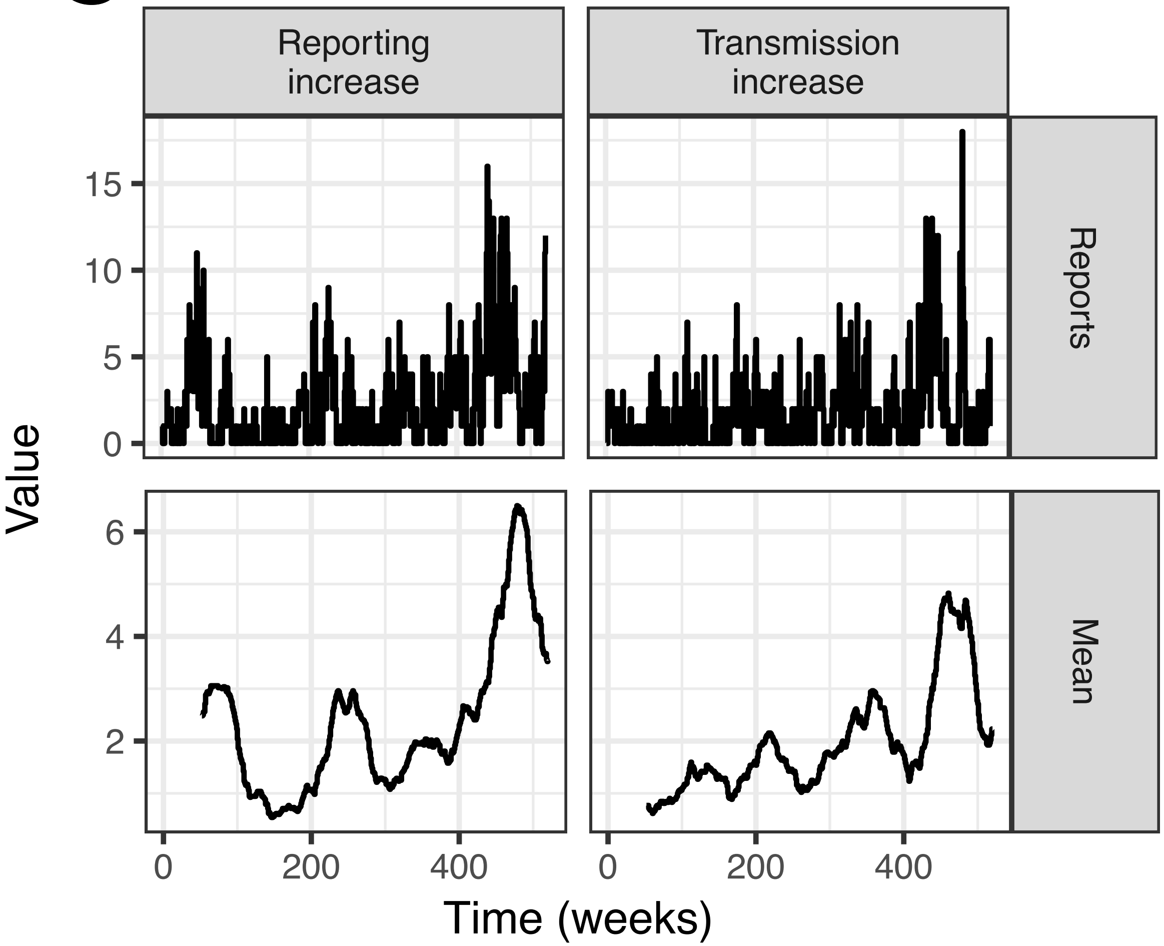
Second-order indicators can be more specific



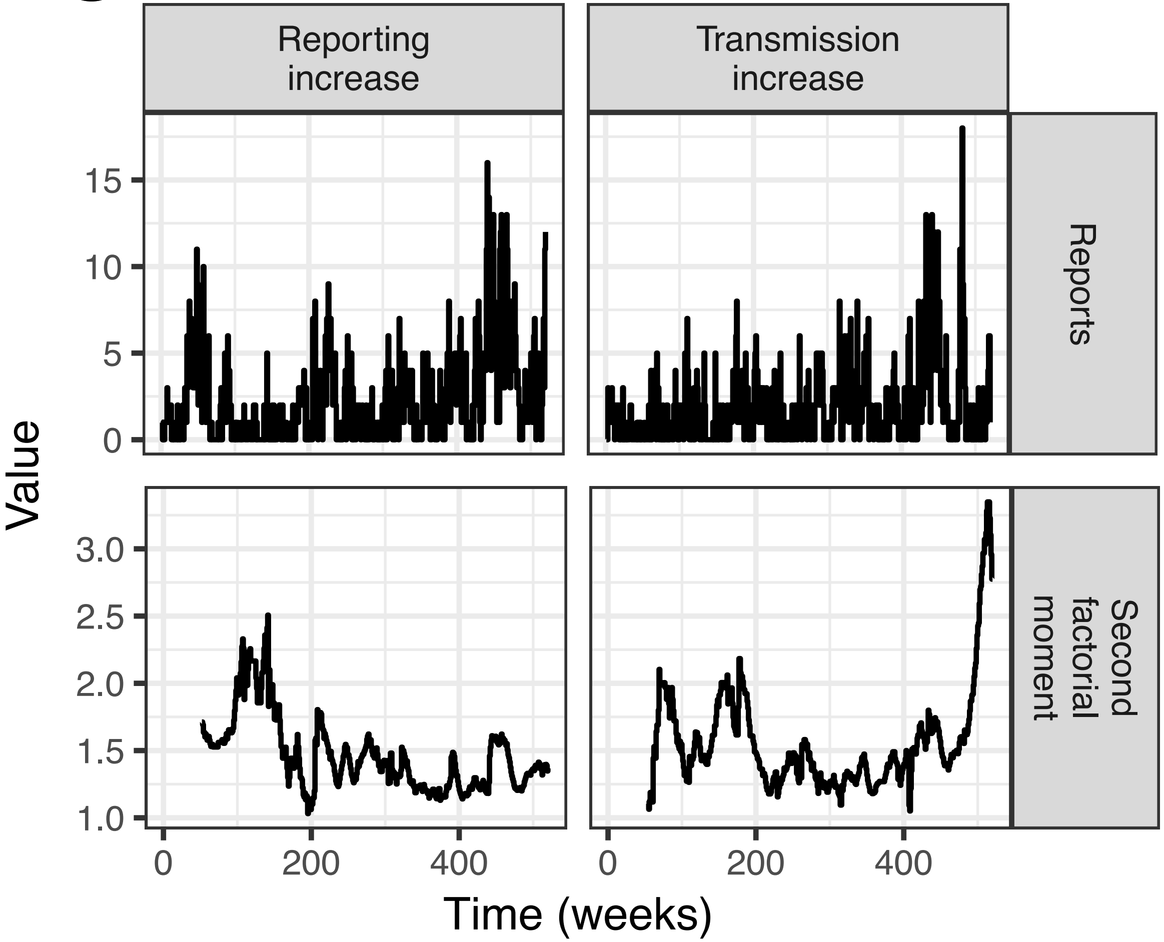
We simulated increasing reports under two scenarios



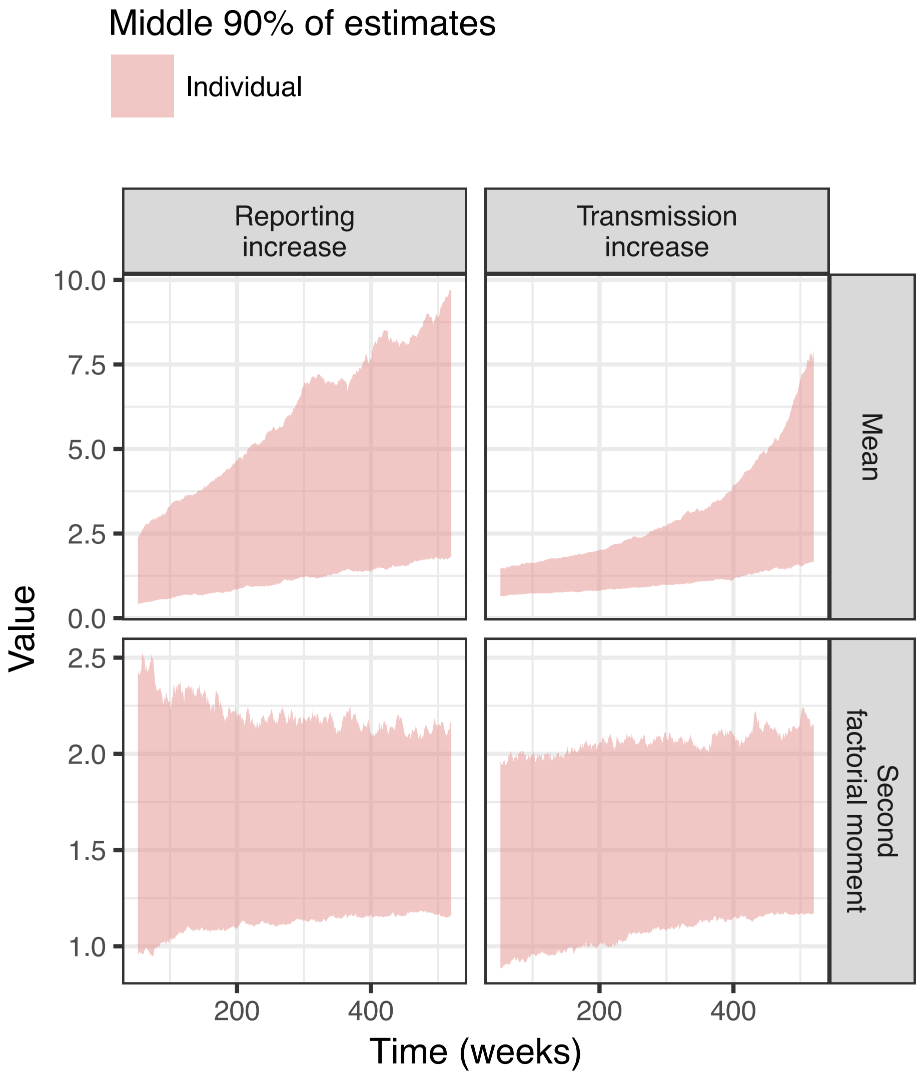
We estimated moments in moving windows for each series



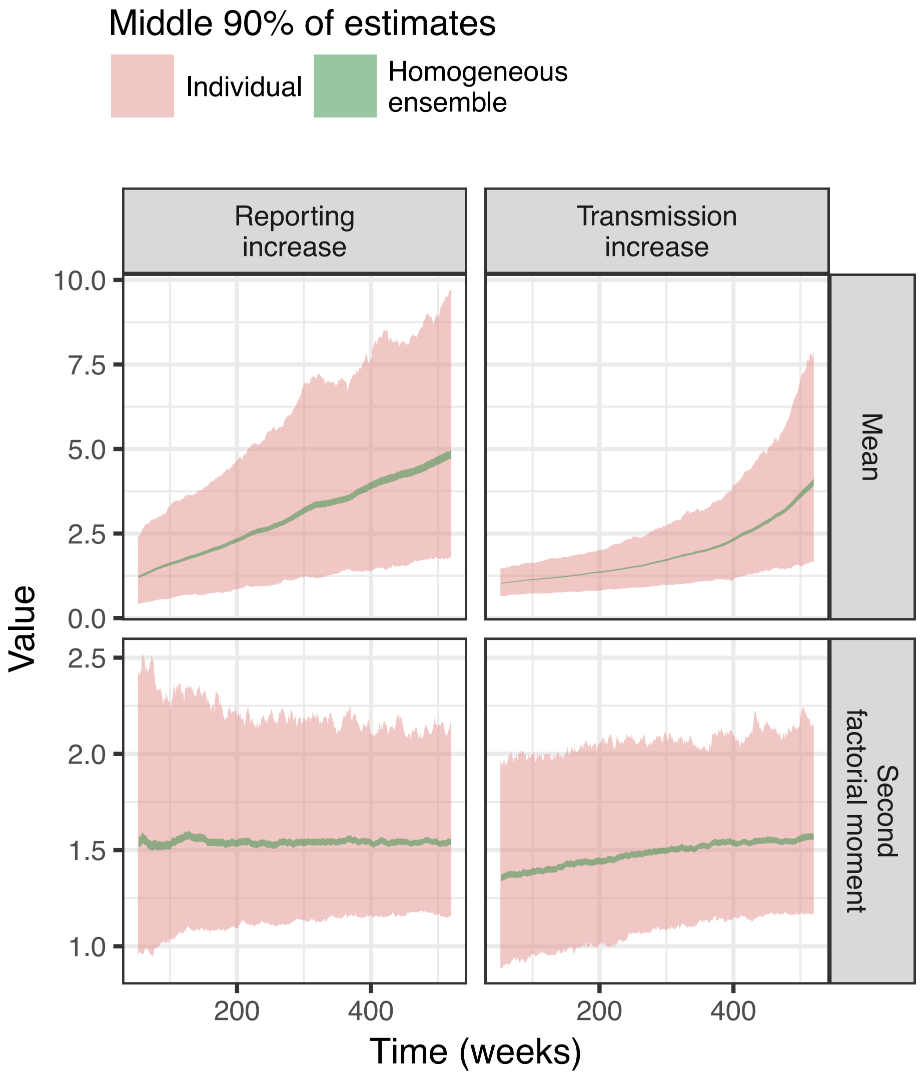
We estimated moments in moving windows for each series



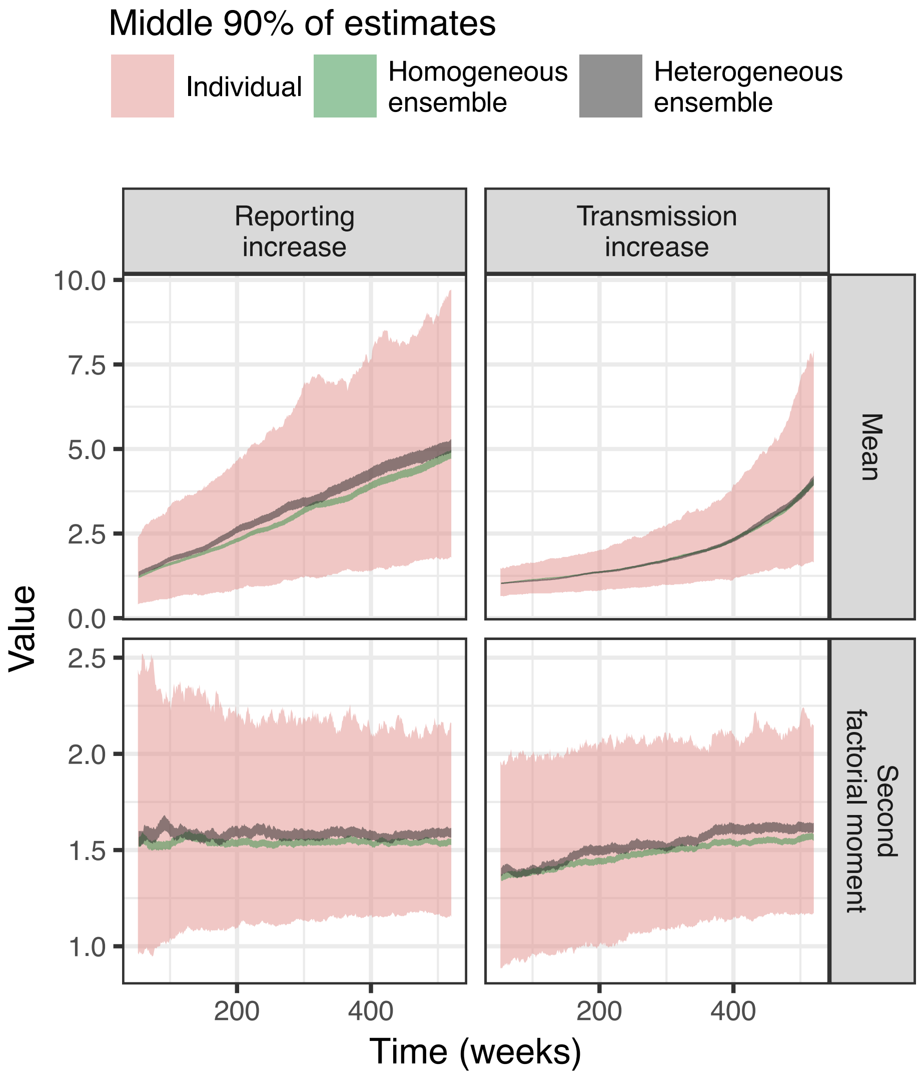
Expected trends were seen, but the noise was high



The noise can be reduced if multiple time series available



Some variation across time series is not a problem

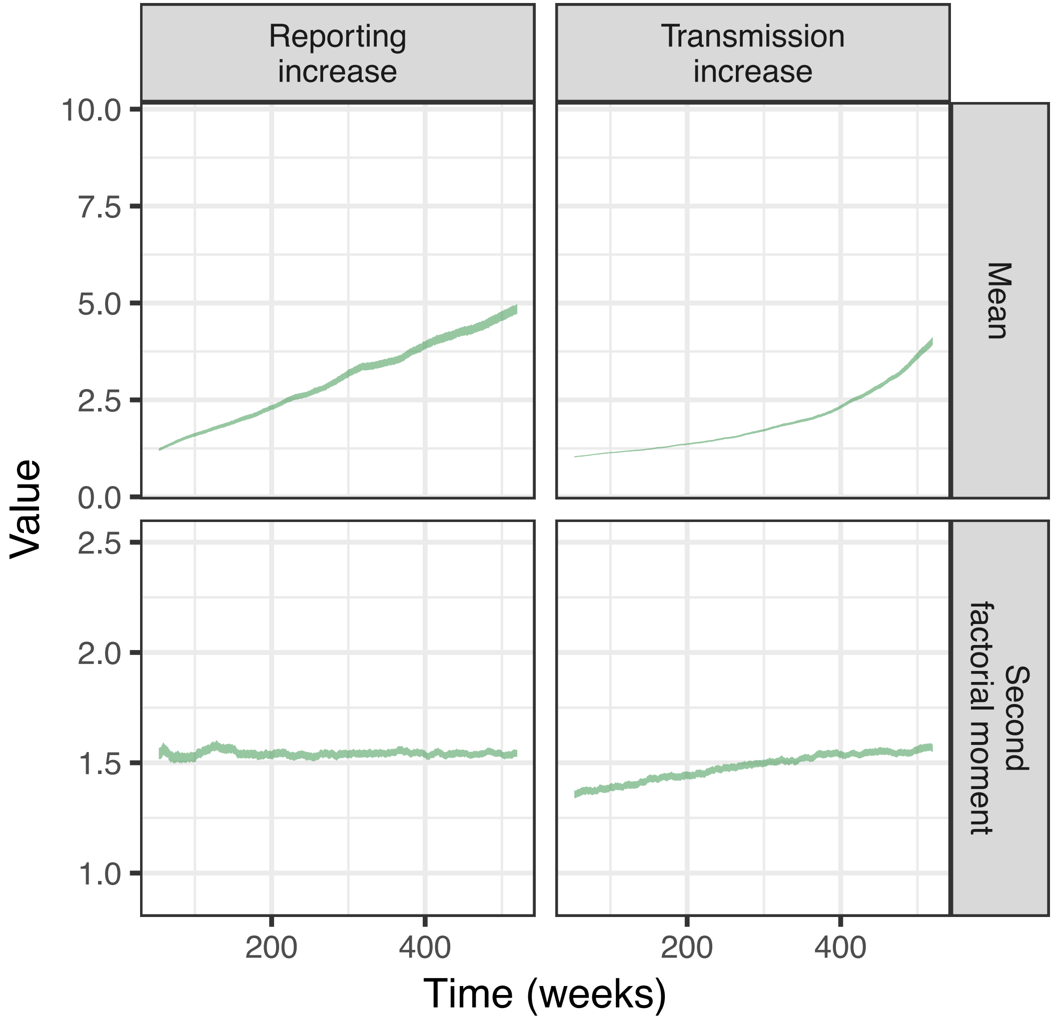


Questions

1. Will indicators based on case data behave similarly?
2. What advantages do second-order indicators have over a rolling mean?

Answers

1. Many should also increase as threshold approached.
2. Some of them are insensitive to changes in reporting probabilities



We are looking for applications, but there are several potential problems

- Stationarity assumptions not valid
- Real ensembles too small and heterogeneous
- Single-type BDI model inadequate
- Observation model inadequate

Thanks!

- NIH Award Number U01GM110744
- Comments from Tobias Brett and Pejman Rohani