When ideas go viral: Early warning signals in theoretical and real-world social contagion systems

Contact: alexis_lerner@brown.edu

Alexis H. Lerner[1,2] and John M. Drake[1]

[1]Odum School of Ecology at the University of Georgia, [2]Brown University

Introduction

The idiom "going viral" is not far from accurate: consumer fads are a type of social contagion in which infected individuals manifest their symptoms by purchasing a transiently popular product or buying into a transiently popular idea. Fads share some important epidemiological features with biological infectious agents; both involve bistable homogenous well-connected complex systems[1] with seemingly spontaneous outbreaks. Here we analyze consumer fad outbreaks as we would an infectious disease epidemic. Can explosive popularity be predicted by early warning signals (EWS) such as critical slowing down (CSD)?

The major phases of a faddish cycle:[2]

① Potentially extensive incubation: few utilize the innovation;
② Take-off: popularity rises explosively;
③ Ascendancy: high levels of usage;
④ Rapid decline leading to a low equilibrium level of usage.

We suggest these phases be considered in analogy to the trajectory of an SIR epidemic since the cycle seems well-aligned with how pathogen outbreaks may occur.[3] Moreover, we hypothesize that studying the dynamic transition from phase 1 to 2 may reveal CSD as a predictor of fad outbreak.

Objective

Extend rumor-spreading theory and epidemiological compartmental models to consumer behavior.

- How can a rumor-spreading model take into account heightened likelihood of S-I state-switching with repeated exposure?
- How does such a model behave as it is pushed through a critical threshold?
- Which early warning signals are detectable in that model, and which perform most robustly?

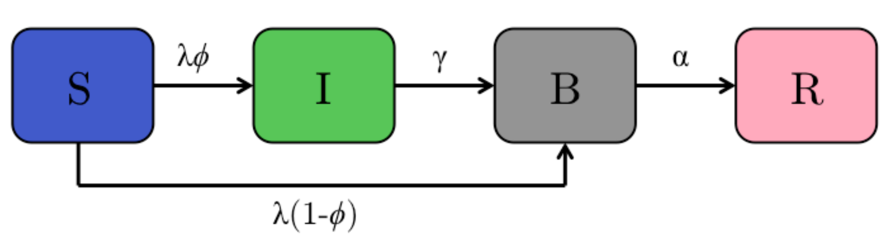
Analyze real-world consumer fad data using early warning signals.

- Does a social contagion outbreak exhibit CSD?
- Which early warning signals and analysis parameters yield most interpretable results?

SIBR Model and Accumulating Buzz

Observational learning theory posits that convergent - and sometimes explosive - behavior can be explained by individuals having access to the past decisions of others.[2,4] This is especially true of consumer fads in that decisions are often informed more by societal conformity than product usefulness. Moreover, the internet has created much faster and larger information cascades, leading to phenomena such as viral videos gaining millions of viewers in days. To capture the "passive transmissibility" of social contagion, we propose a simple compartmental model for social contagion divided into four classes: **Susceptible, Infected, Buzz, and Recovered (Fig. 1).** .

Figure 1: SIBR flowchart



This SIBR model assumes transmissive capability of both active adopters/advocates (analogous to rumor-spreading) and "buzzers": those who were recently active participants in its direct advocacy, or who are passive participants (onlookers, listeners, inquirers) observed by susceptibles.

The SIBR system of ordinary differential equations:

$$\frac{dS}{dt} = -S(\beta_I I + \delta\beta_B B + \epsilon) \quad (1)$$

$$\frac{dI}{dt} = S(\beta_I I + \delta\beta_B B + \epsilon)\phi - \gamma I \quad (2)$$

$$\frac{dB}{dt} = S(\beta_I I + \delta\beta_B B + \epsilon)(1-\phi) + \gamma I - \alpha B \quad (3)$$

$$\frac{dR}{dt} = \alpha B \quad (4)$$

$$\frac{d\delta}{dt} = I + B \quad (5)$$

The SIBR model integrates contagion ubiquity over time with the cumulative factor δ. A social contagion becomes more passively transmittable (β_B) as the number of infecteds and buzzers increases over time. The cumulative factor δ is directly multiplied to β_B as a means of increasing the force of infection with increasing popularity. Unlike conventional hinge models of R_0, the SIBR model has a built-in positive feedback loop that is capable of pushing a non-critical system to its critical threshold.

$$\lambda = \beta_I I + \delta\beta_B B + \epsilon \quad (6)$$

$$R_{CE} = \frac{\beta_I \phi}{\gamma} + \frac{\delta\beta_B}{\alpha} \quad (7)$$

$$\delta_{crit} = \frac{\alpha}{\beta_B}\left(1 - \frac{\beta_I\phi}{\gamma}\right) \quad (8)$$

The force of infection (Eq. 6) results from direct endorsement, ubiquity, and spontaneous exposure. The cumulative effect-dependent reproduction number (Eq. 7) was derived using the next generation matrix method.[5] The critical threshold occurs when R_{CE}=1 and can be solved for δ_{crit} with a given set of parameters (Eq. 8).

Table 1: SIBR definitions

S	Susceptibles
I	Infecteds; adopters/high-level transmitters
B	Buzzers; low-level transmitters
R	Recovereds
δ	Cumulative factor; contagion ubiquity
λ	Force of infection
β_I	Transmissibility due to infecteds
β_B	Transmissibility due to buzzers
ε	Spontaneous exposure rate
ϕ	S→I conversion factor
1-ϕ	S→B conversion factor
γ	I→B recovery rate
α	B→R recovery rate

Consumer Fad Case Study

To characterize CSD in a real-world consumer system, a data set for a faddish product line was analyzed using the R package spaero. The product in question was a handmade crocheted hat modeled off of the hair of Cabbage Patch Kids dolls, sold by the shop The Lillie Pad on the online marketplace Etsy. It was offered on the site for approximately two years (June 2011-July 2013) before exploding in popularity, landing product profiles by the Huffington Post, Today, ABC News, Perez Hilton, Daily Mail, and others.

Figure 7: Cabbage patch hat outbreak EWS analysis

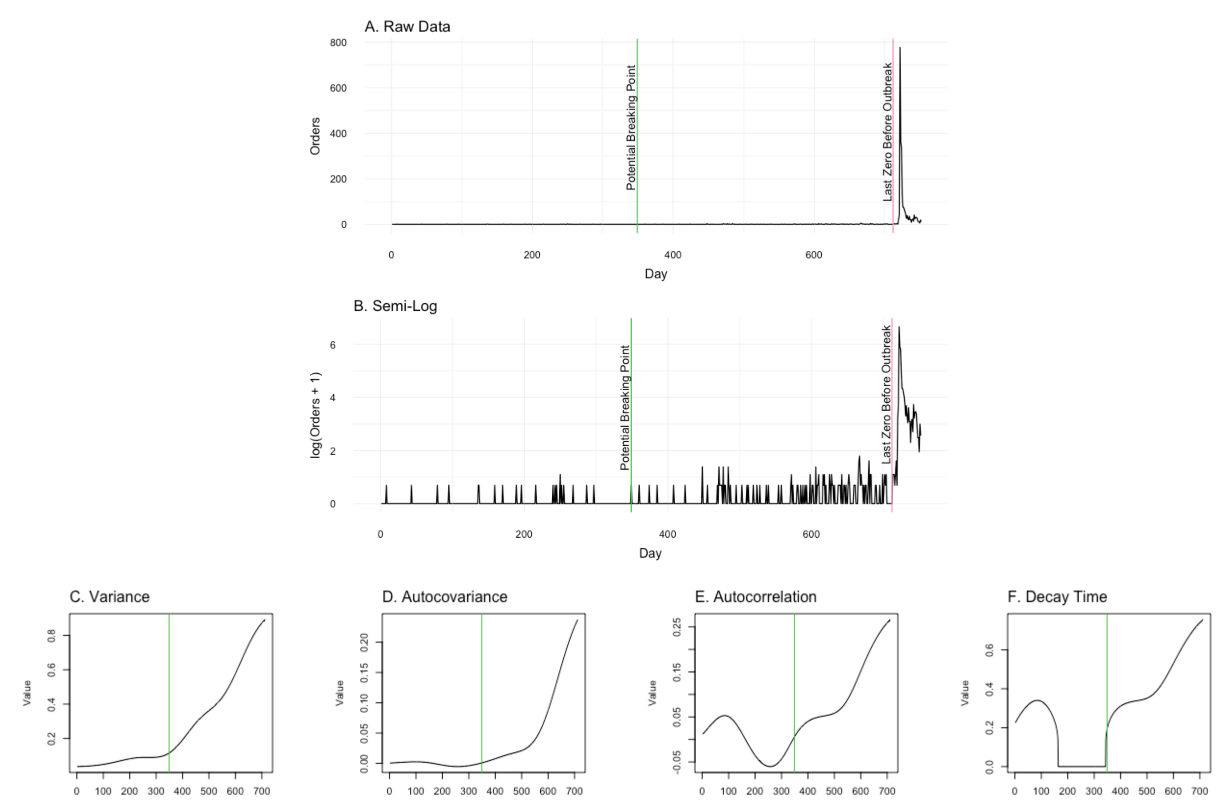


Fig. 7: The data set was truncated at the last observed instance of zero orders before the maximum; this happened to be at day 712 (pink line). Optimal early warning signal results were empirically deduced with a rolling window bandwidth of 10%. CSD is clearly detectable in all four EWS. A potential tipping point (green line) occurred in the middle of the time series, at the first instance of an order after a ~50-day quiet period. It seems to be a reasonable starting point for observable critical slowing down, especially for variance **(7C)** and decay time **(7F).**

Simulations

Deterministic

Figure 2: With accumulating buzz

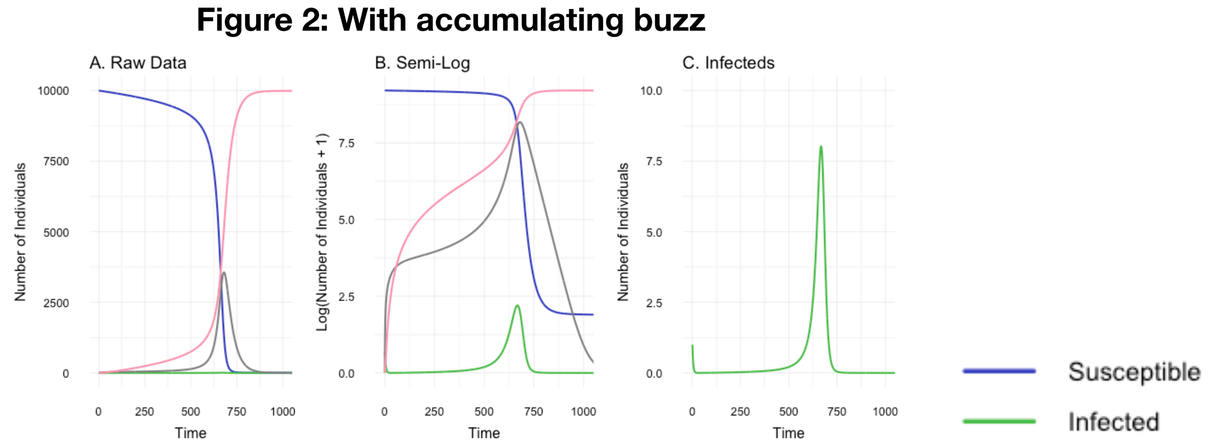
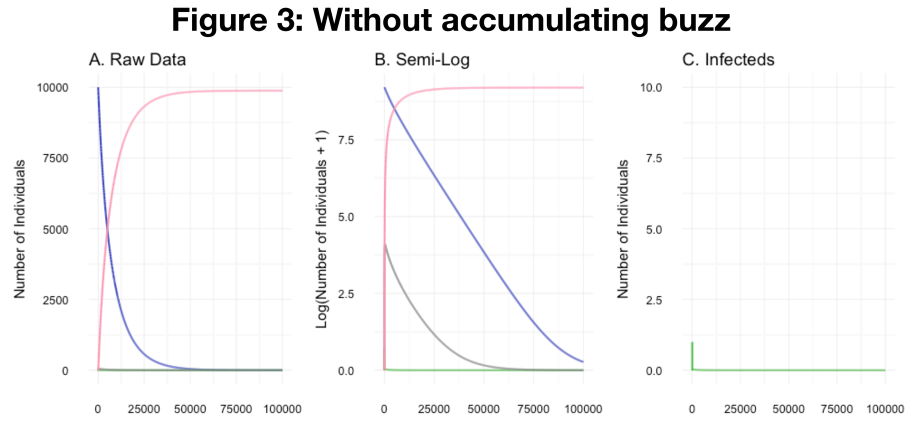


Figure 3: Without accumulating buzz



Figs. 2, 3: The SIBR ordinary differential equation set was solved (R package deSolve) with initial values of N=10,000 total individuals, S=9,999 and I=1. β_I=0.99, β_B=0.0099, ϕ=0.015, γ=0.26, α=0.026 and ε=0.0001 to yield R_0=0.44. An epidemic only occurs in the infected class (**2C, 3C**) when the likelihood of S-I state-switching is driven forward by the accumulation of buzz.

Stochastic

Figure 4: With accumulating buzz

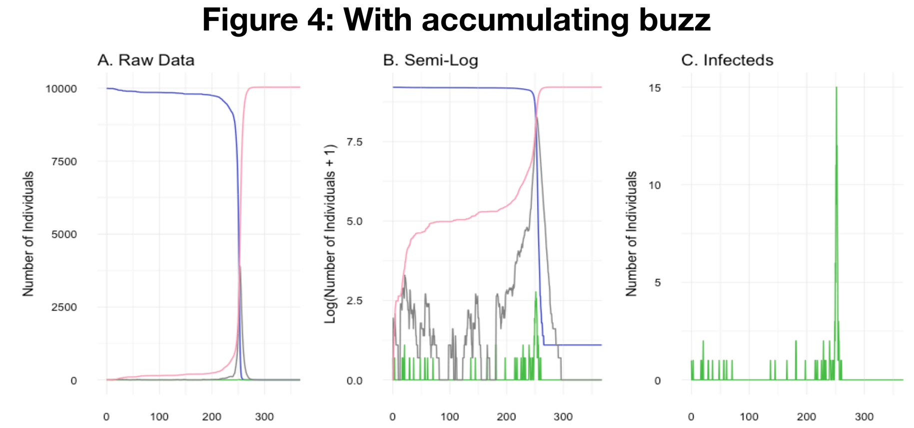
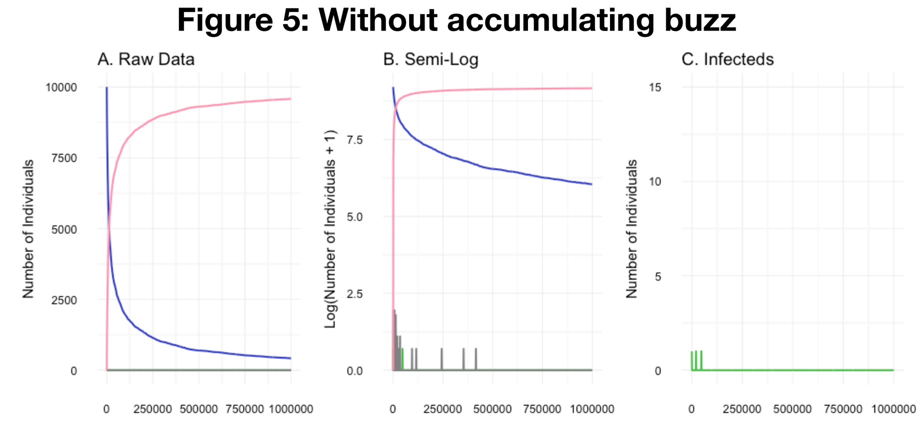


Figure 5: Without accumulating buzz



Figs. 4, 5: The SIBR model was stochastically simulated to better estimate probable outcomes. The transition matrix and associated rates for the Gillespie simulation algorithm (R package adaptivetau) were created from **Fig. 1** and **Table 1.** Initial values and parameters were set in tandem with the deterministic simulations (see left). Again, supercriticality was achieved only with the inclusion of accumulating buzz.

Performance Analysis

EWS can be interpreted as binary classifiers with performance measured by specificity and sensitivity. We studied how four potential signals of CSD could differentiate system behavior closer to versus further from an epidemic peak. The AUC statistics of receiver-operating characteristic (ROC) curves were analyzed for 81 parameter sets with 100 stochastic simulations per set and various bandwidths. For consistency in isolating the pre-spike interval, the outbreak cutoff was defined as the last observed instance of zero infected individuals before the maximum was hit. The remaining time series was divided into CSD(-) and CSD(+) groups based on the calculated δ_{crit} for its parameter set. Kendall's τ was used as a rank correlation coefficient.

Figure 6: AUC performance analysis of potential EWS by bandwidth

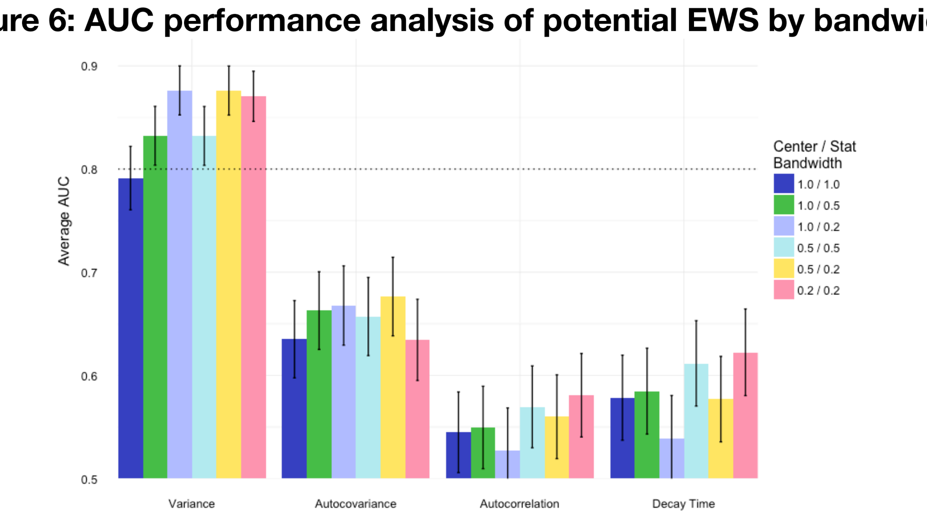


Fig. 6: For each simulation, the SIBR model was assigned initial values of N=10,000 total individuals, S=9,999 and I=1. β_I=0.69-0.71, β_B=0.0069-0.0071, ϕ=0.015, γ=0.24-0.26, α=0.024-0.026 and ε=0.0001 to yield R_0=0.31-0.34. An AUC cutoff of 0.8 was selected as indication of a good classifier. Variance performed highest and most consistently of the EWS evaluated when the statistical bandwidth was equal to 0.2.

Conclusions

① Accumulating buzz can push a social contagion system from sub- to supercritical **(Figs. 2-5).**
② Variance was the highest-performing and most consistent EWS for appropriately detecting CSD (AUC > 0.85) at low statistic bandwidths. **(Fig. 6).**
③ CSD is clearly exhibited in real-world consumer fads with potentially explainable "breaking points" **(Fig. 7).**
④ EWS of consumer behavior can be used to inform predictions of economic interest.

Future directions: Conduct performance analysis of EWS with changing transmissibility parameters.

Acknowledgements

We thank Jonathan Lillie and Amanda Lillie for the Etsy data; Eamon O'Dea and Paige Miller for AUC code contributions; and Paige Miller, Hollis Jenkins, Joseph Walker, John Vinson, Suzanne O' Regan and the Drake Lab for their support. This research was funded by a grant from the National Institutes of Health.

References

[1]Scheffer et al. *Science* **2012,** *338,* 344–348.
[2]Strang & Macy. *Am. J. Sociol.* **2001,** *107,* 147-182.
[3]Com. Dis. Surv. Centre & Com. Dis. Scot. Unit. *BMJ* **1978,** *6112,* 587.
[4]Bikhchandani et al. *J. Econ. Perspect* **1998,** *12,* 151-170.
[5]Diekmann et al. *J Math Biol* **1990,** *28,* 365-382.
Cabbage patch hat photo courtesy of The Lillie Pad.