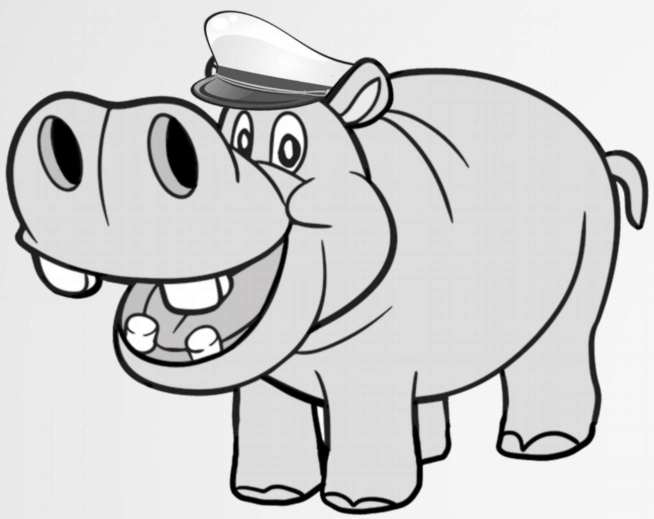


# Online Popularity under Promotion

ICWSM '17
Montréal, May 16th, 2017


Marian-Andrei Rizoiu    Lexing Xie

ComputationalMedia @ANU: http://cm.cecs.anu.edu.au

# The problem

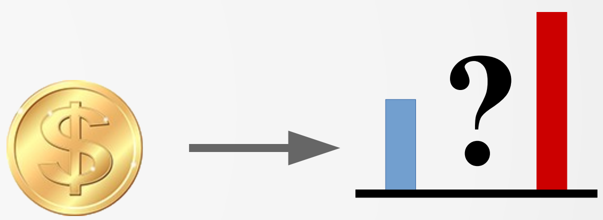
Cultural Markets seem to be unpredictable.

[Salganik et al Science‘06]

1. How well do promotions work?

[D. Watts ‘11] [Zarezade et al WSDM ‘17]

[Zhang et al. WSDM ‘14]



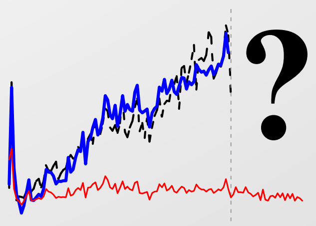

2. When should one promote?

[Chierichetti et al. SIAM Jour. Comp. ‘14]

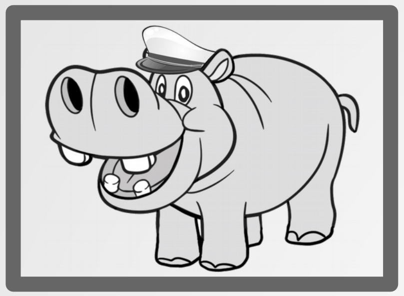
[Bollapragada et al. OR ‘04]


3. How to predict popularity?

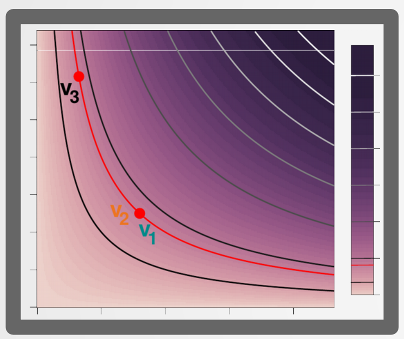
[Rizoiu et al WWW’17][Kobayashi et al ICWSM’16]



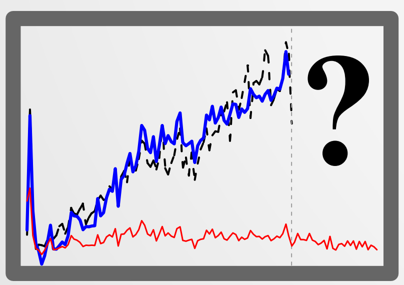
# Presentation outline



Modeling popularity with HIP



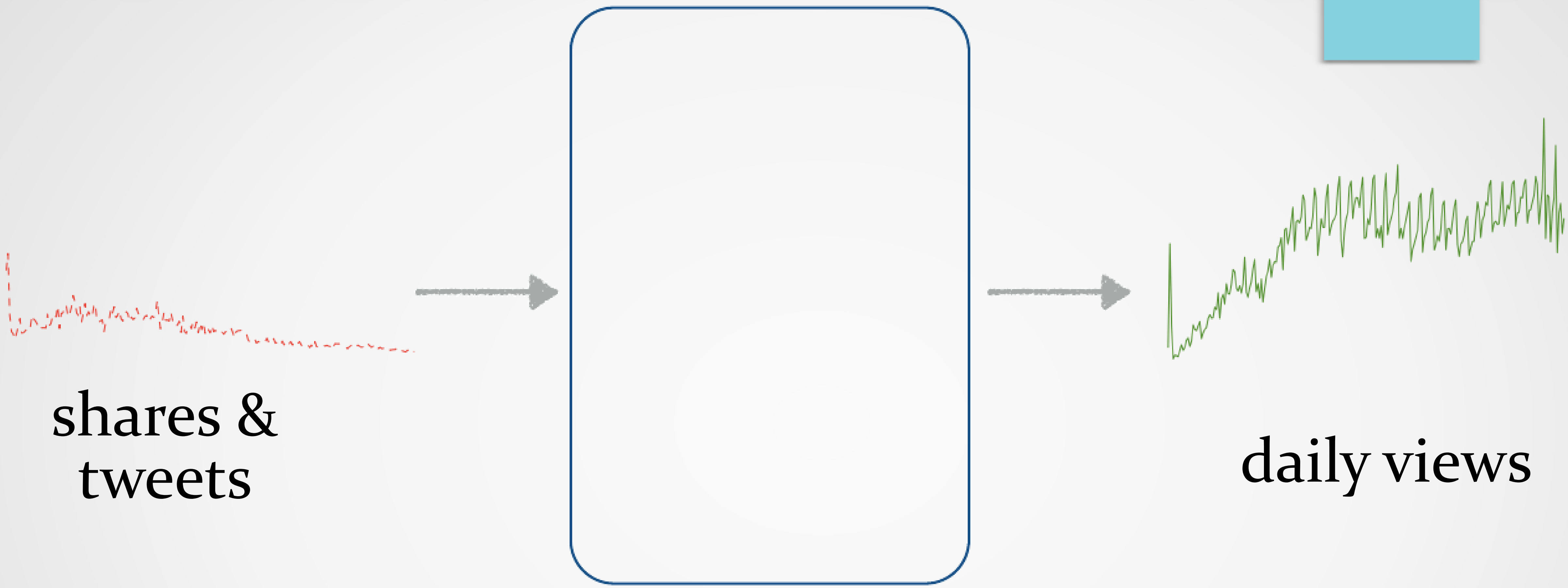
Content virality and maturity time




Forecasting popularity under promotion

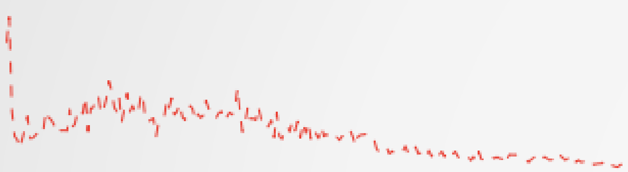
Promotions schedules and memory lengthening through promotion

# HIP: Linking promotion and popularity

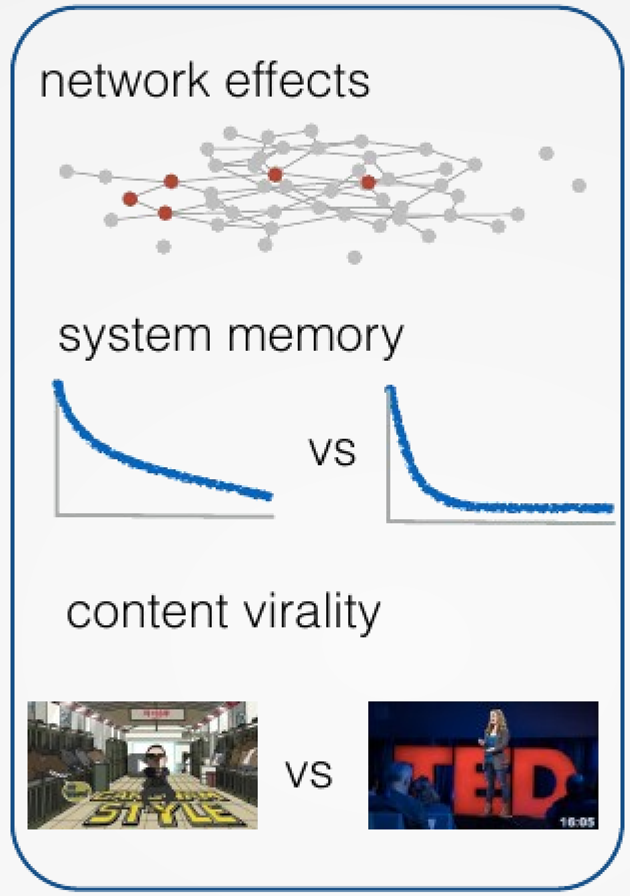


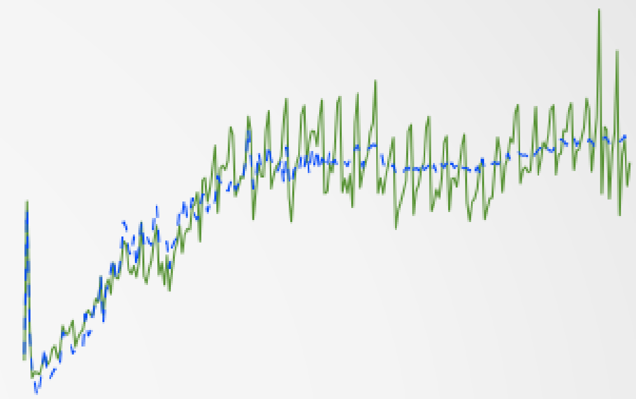

M.-A. Rizoiu, L. Xie, S. Sanner, M. Cebrian, H. Yu, and P. Van Hentenryck, "**Expecting to be HIP: Hawkes Intensity Processes for Social Media Popularity**", in Proc. International Conference on World Wide Web (**WWW '17**), Perth, Australia, pp. 735-744, 2017.

# HIP: Linking promotion and popularity



shares & tweets



daily views

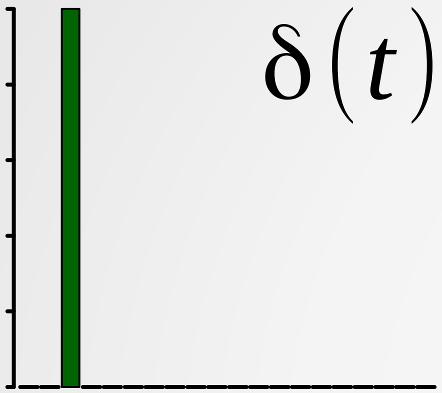
$$\xi(t) = \mu s(t) + C \int_0^t \xi(t-\tau)(\tau + c)^{-(1+\theta)} d\tau$$

popularity

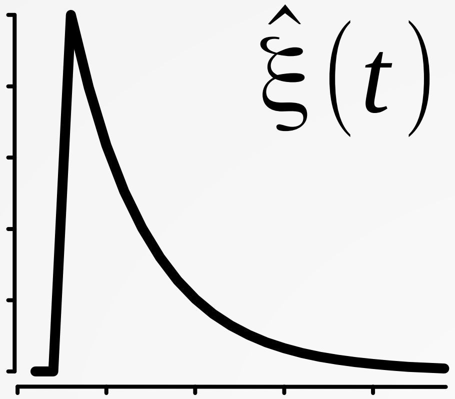
promotion

# HIP as a Linear Time-Invariant system

promotion

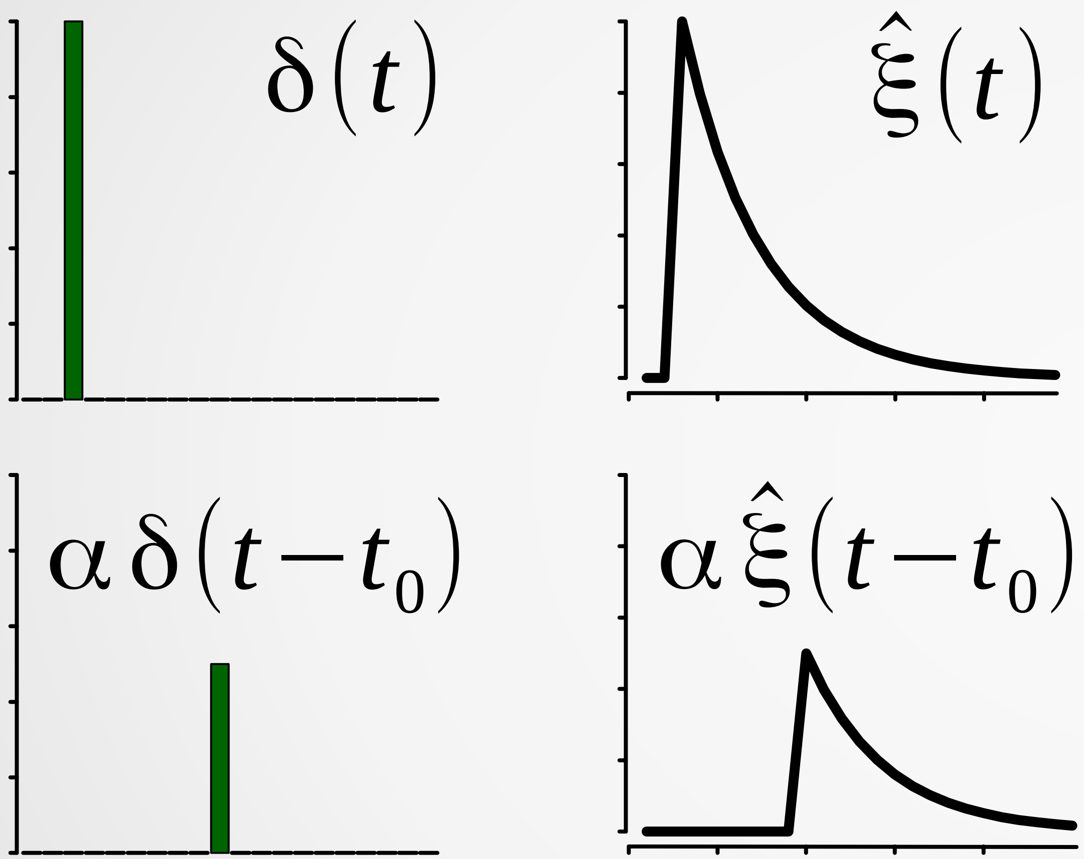



response

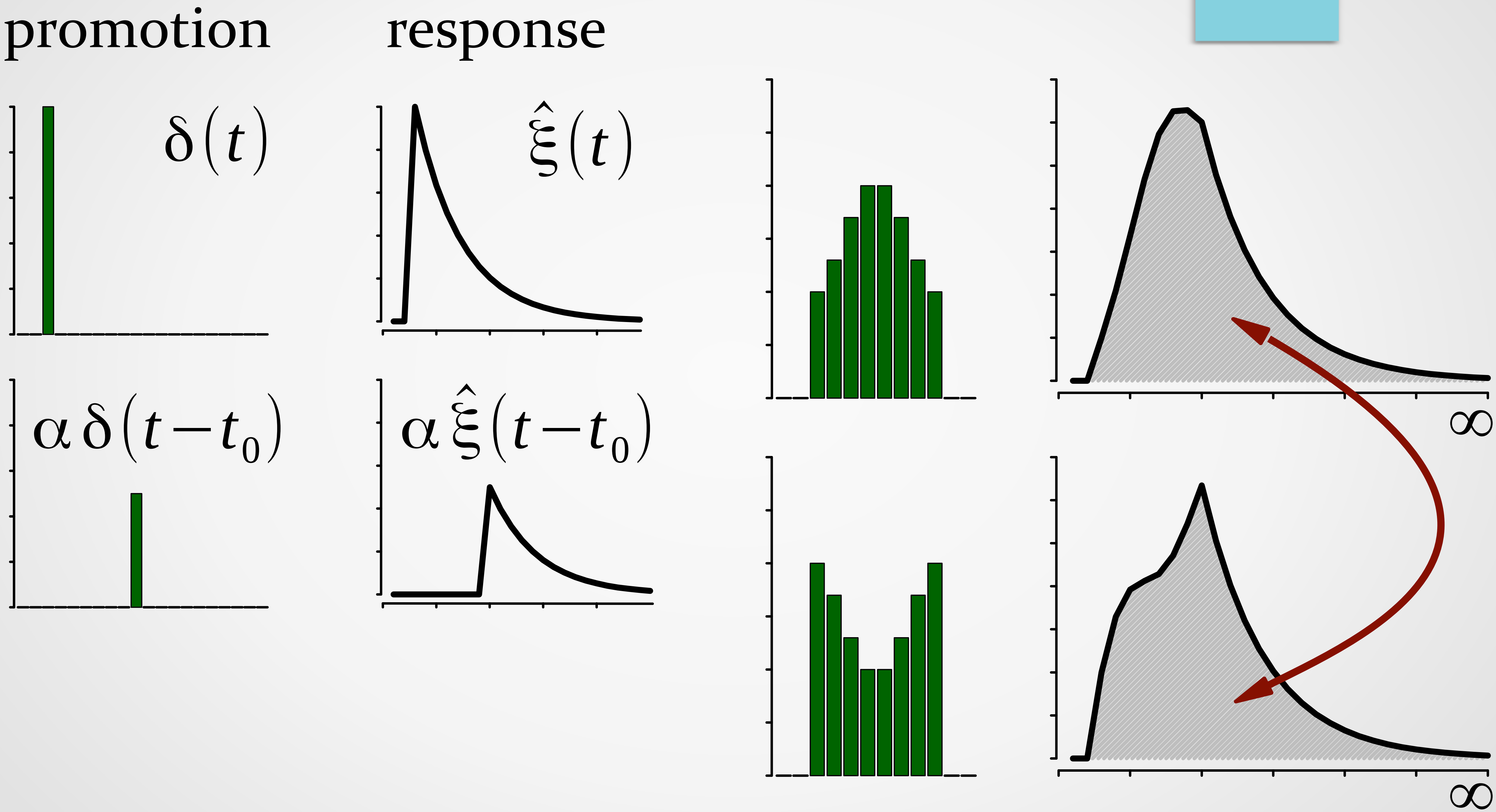


# HIP as a Linear Time-Invariant system

promotion response

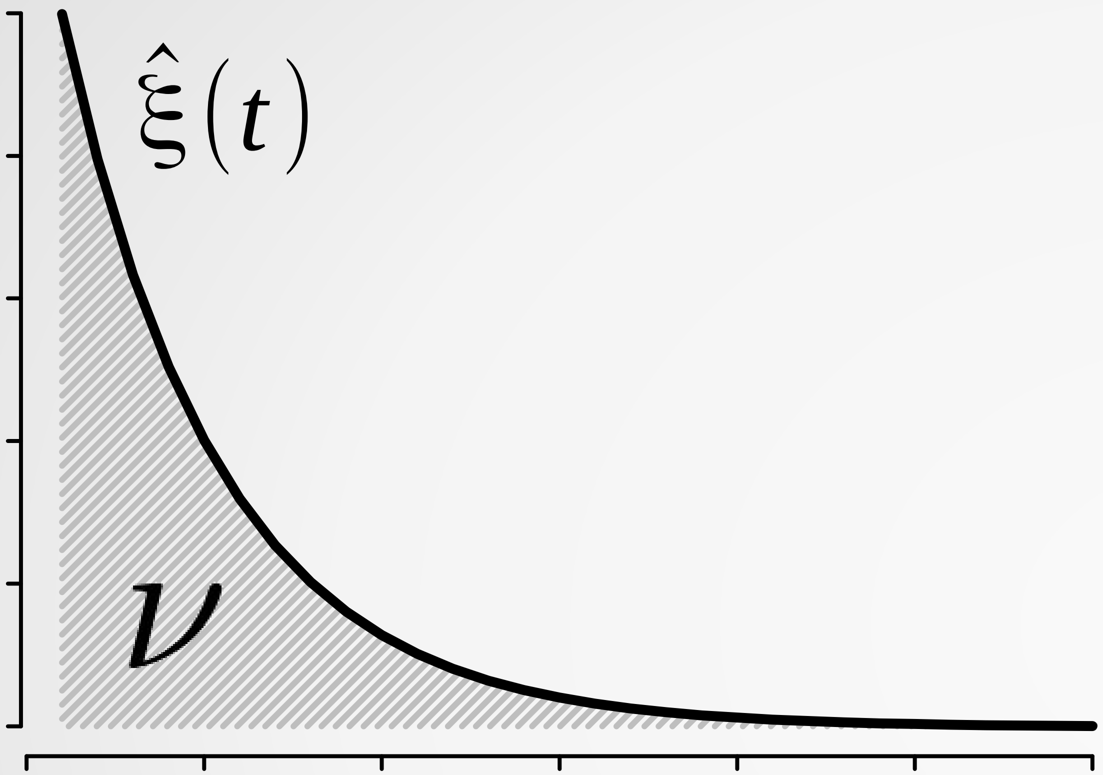


# HIP as a Linear Time-Invariant system




**Corollary:** same budget — same return

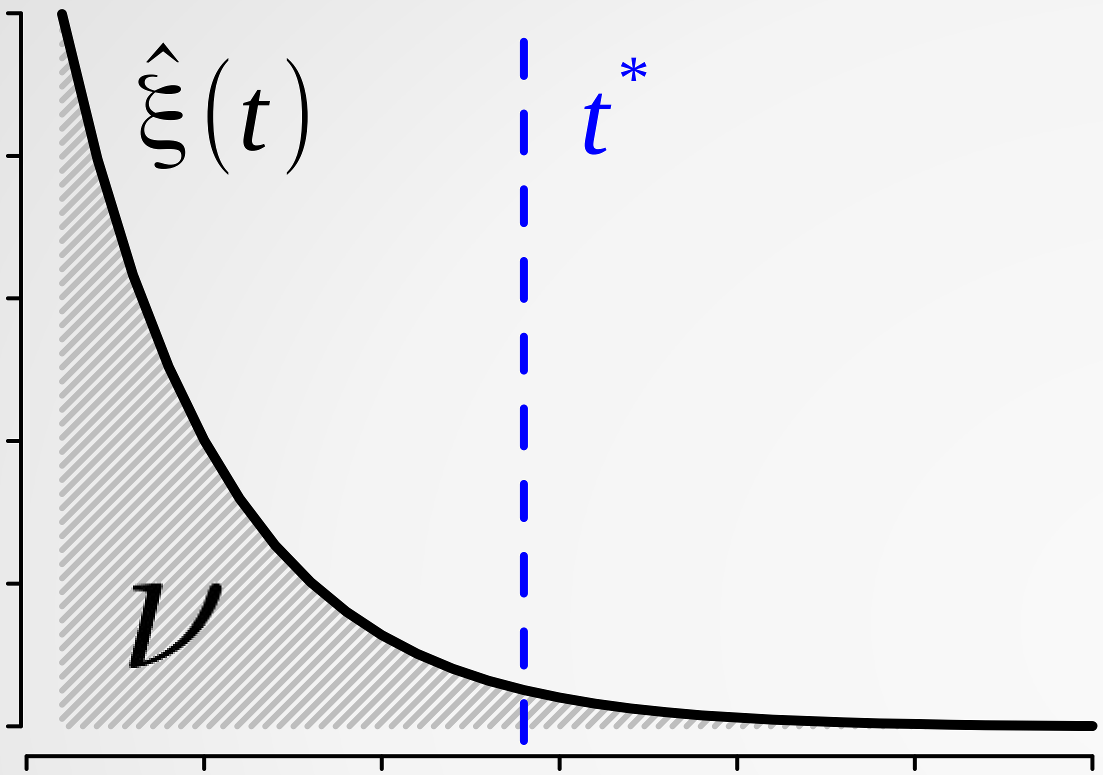
# Viral potential and maturity time




Viral potential score:

$$\nu = \int_0^\infty \hat{\xi}(t)dt$$

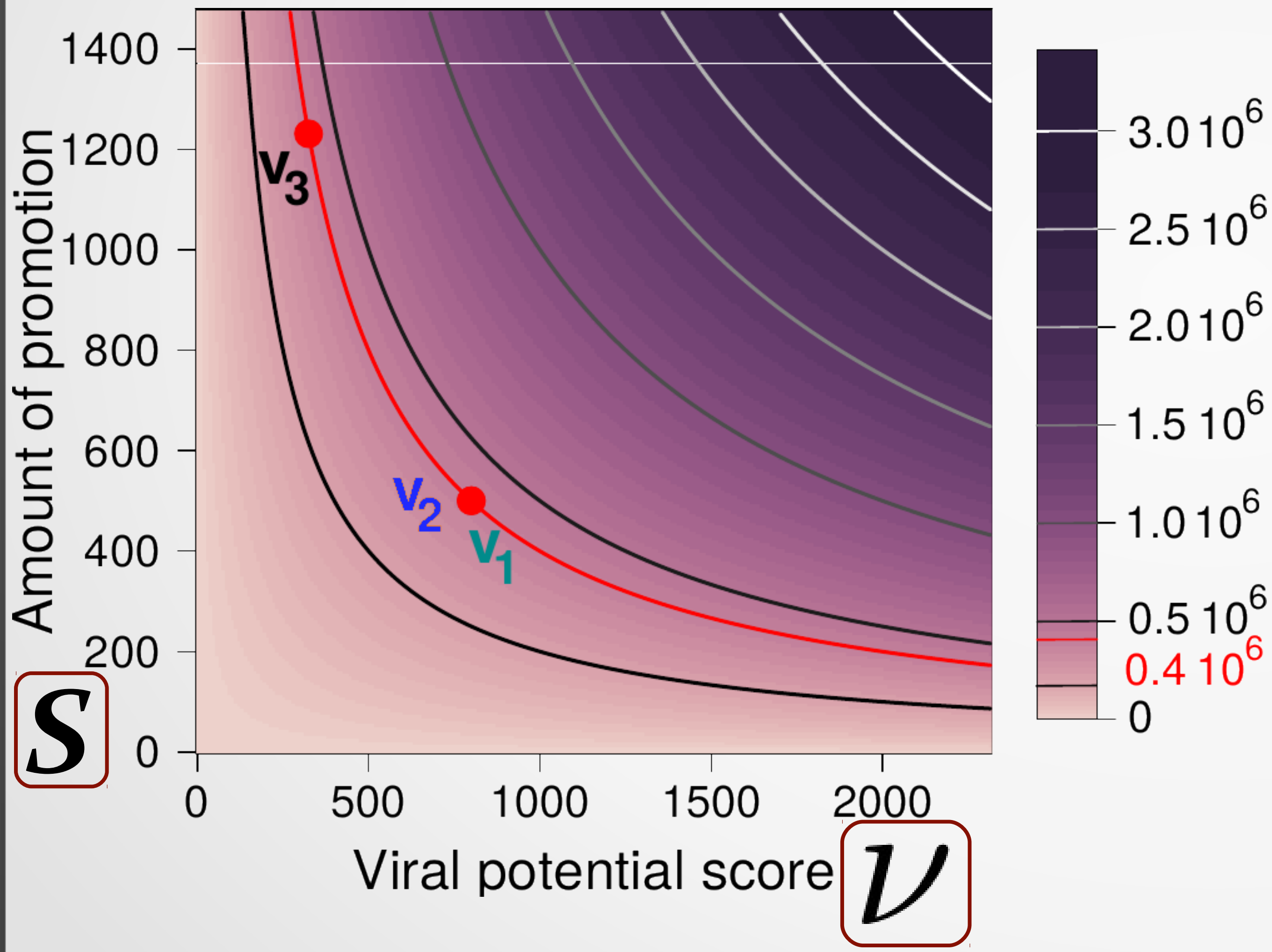
# Viral potential and maturity time



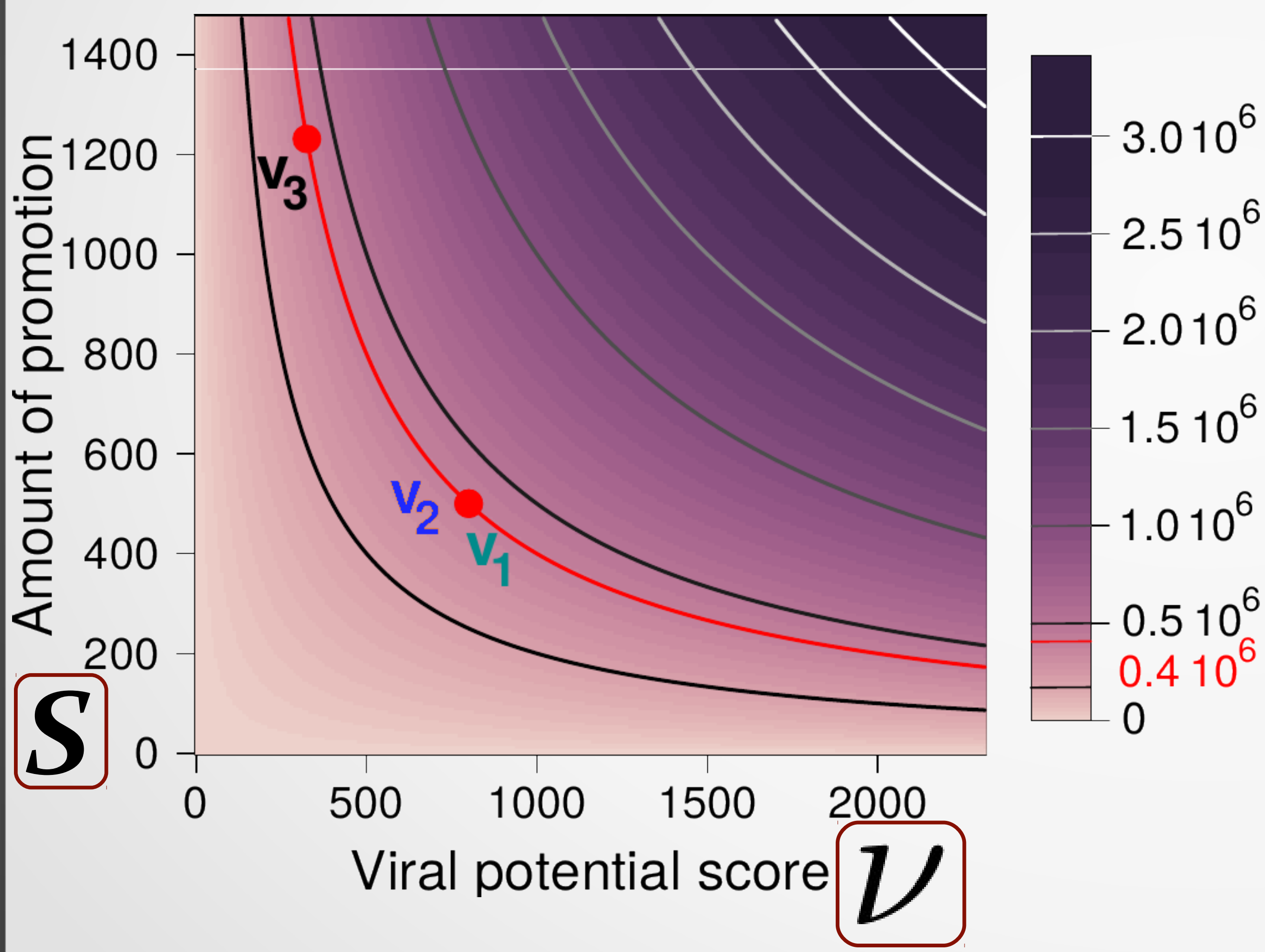

Viral potential score: $$\nu = \int_0^\infty \hat{\xi}(t)dt$$

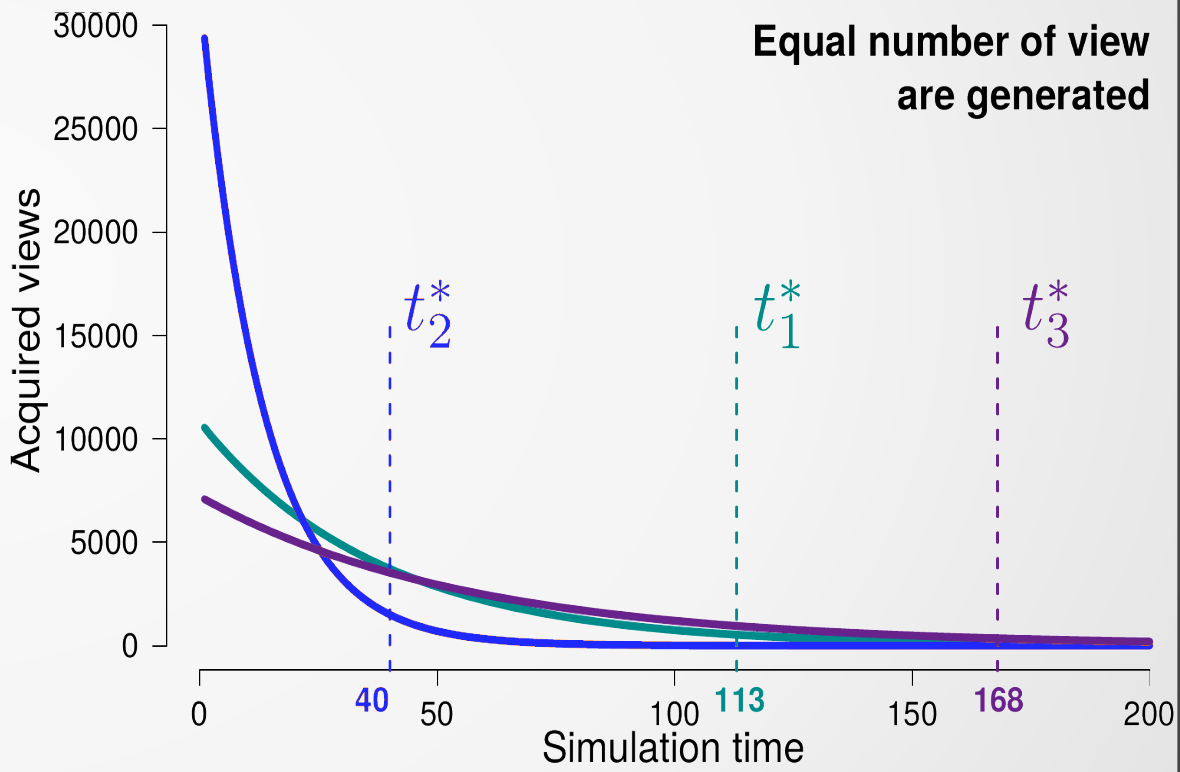
Maturity time: $$t^* = min\left\{t \geq 0 \middle| \int_0^t \hat{\xi}(t)dt \geq 0.95\nu\right\}$$

# Virality map

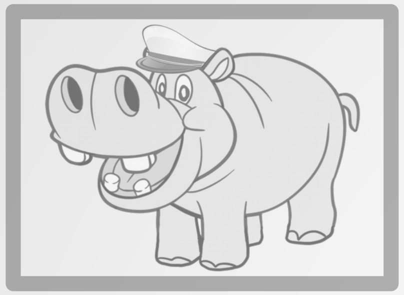


# Virality map

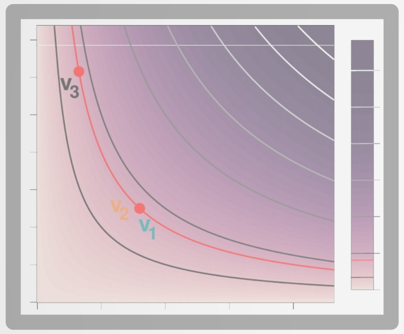


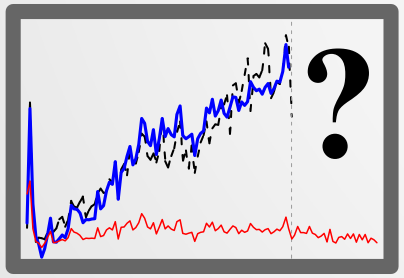
# Presentation outline



Modeling popularity with HIP



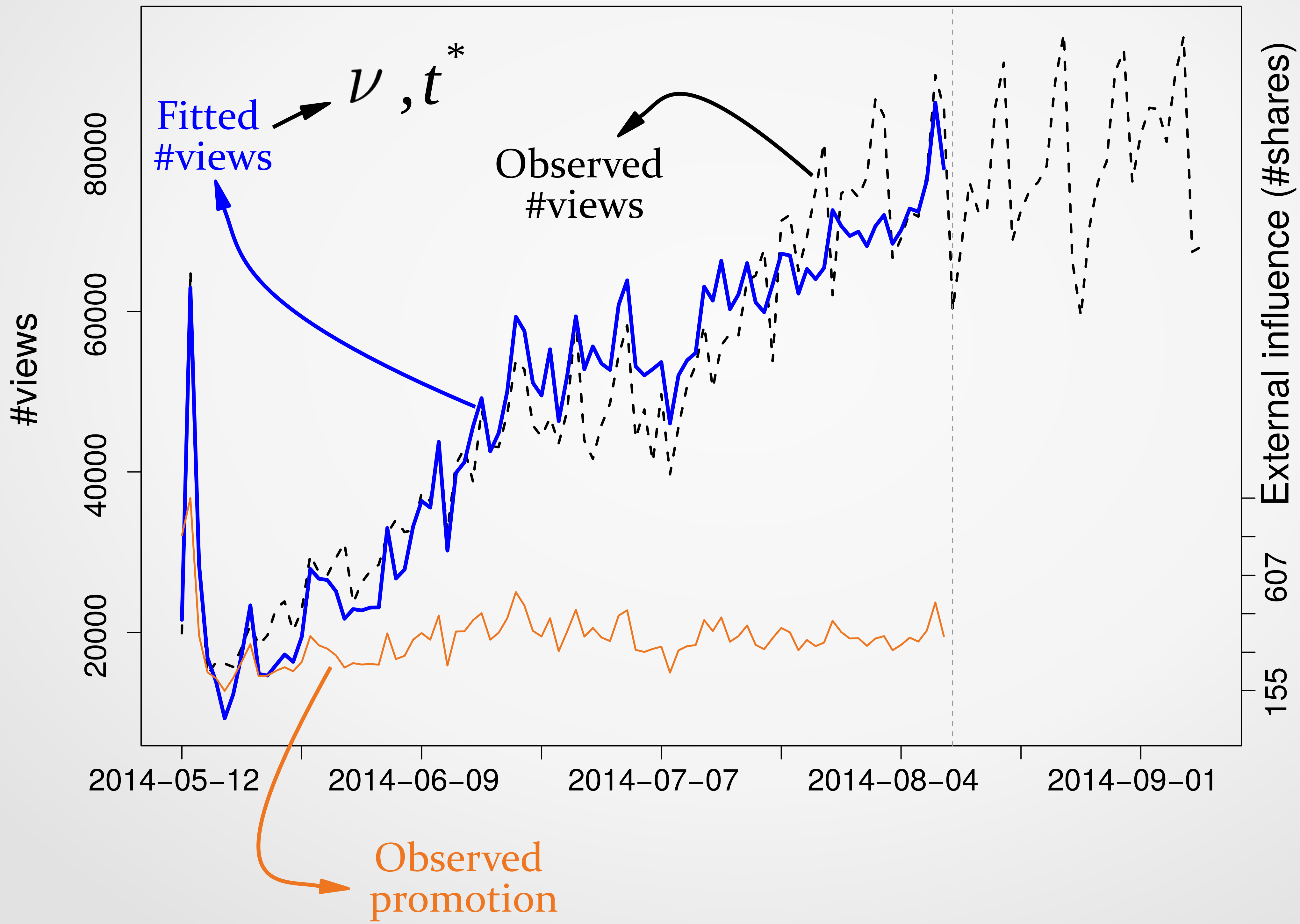
Content virality and maturity time



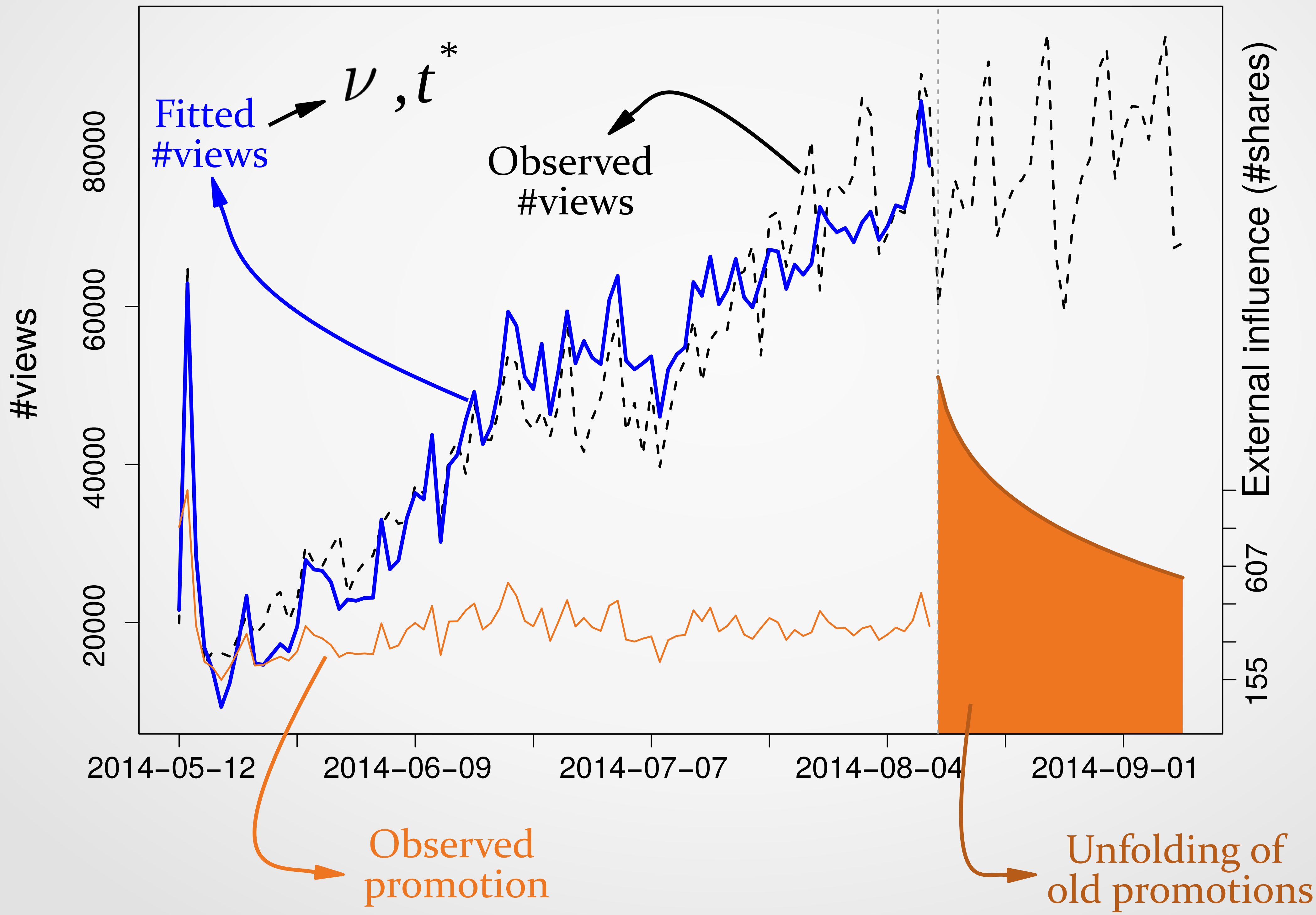

A progression of two problems relating to predicting popularity under promotion

Promotions schedules and memory lengthening through promotion

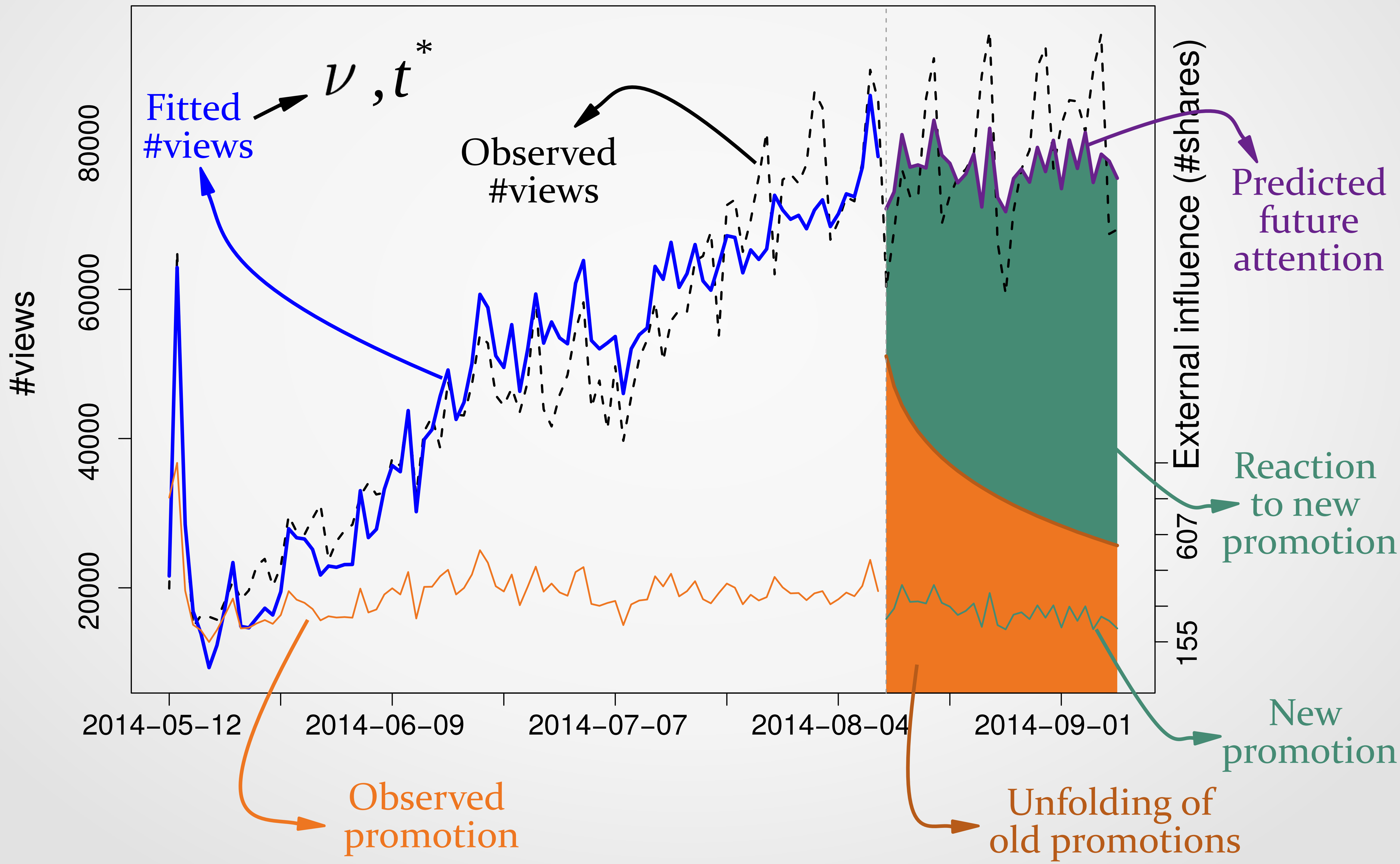
# Forecasting future views (1)



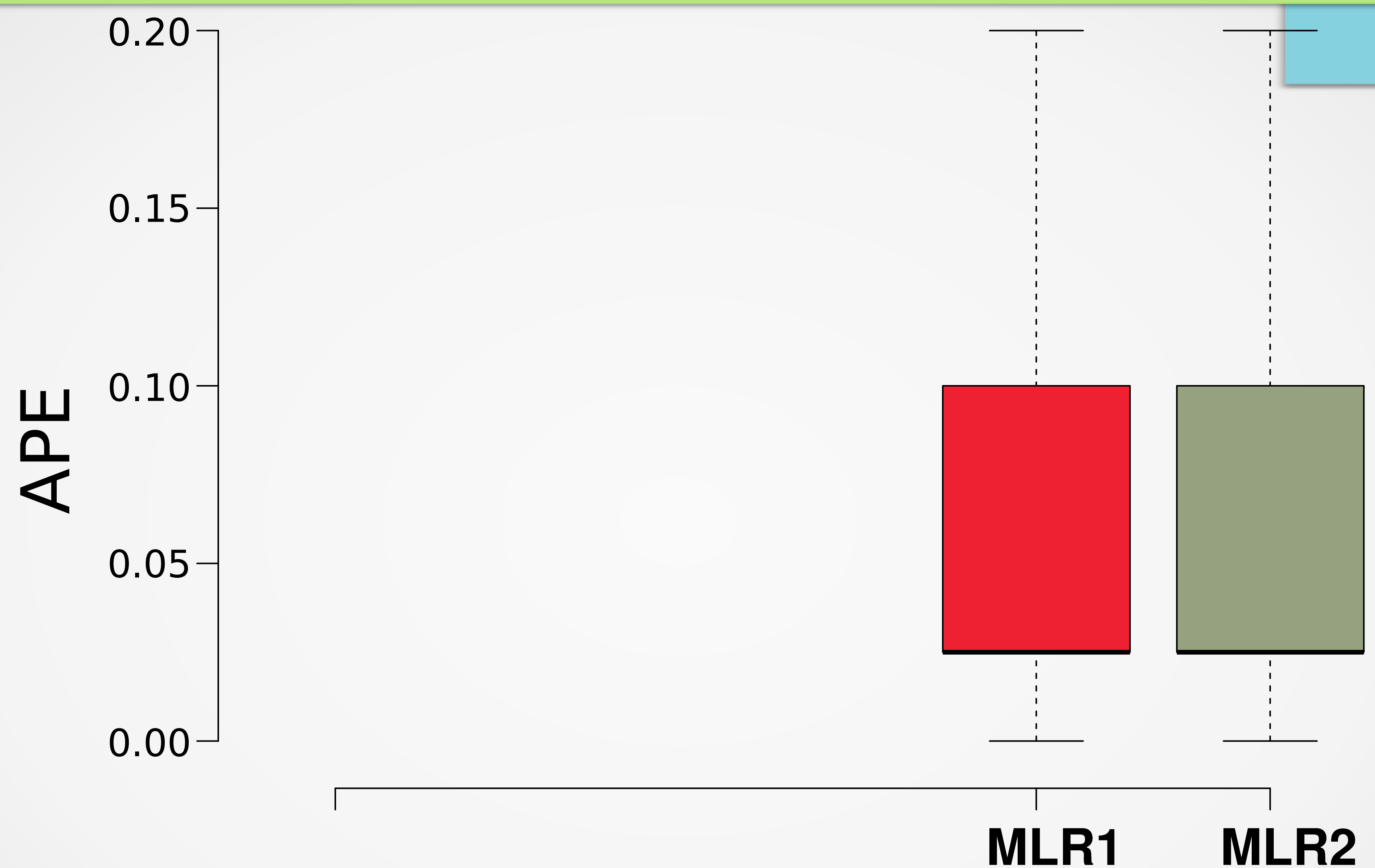
# Forecasting future views (1)



# Forecasting future views (1)



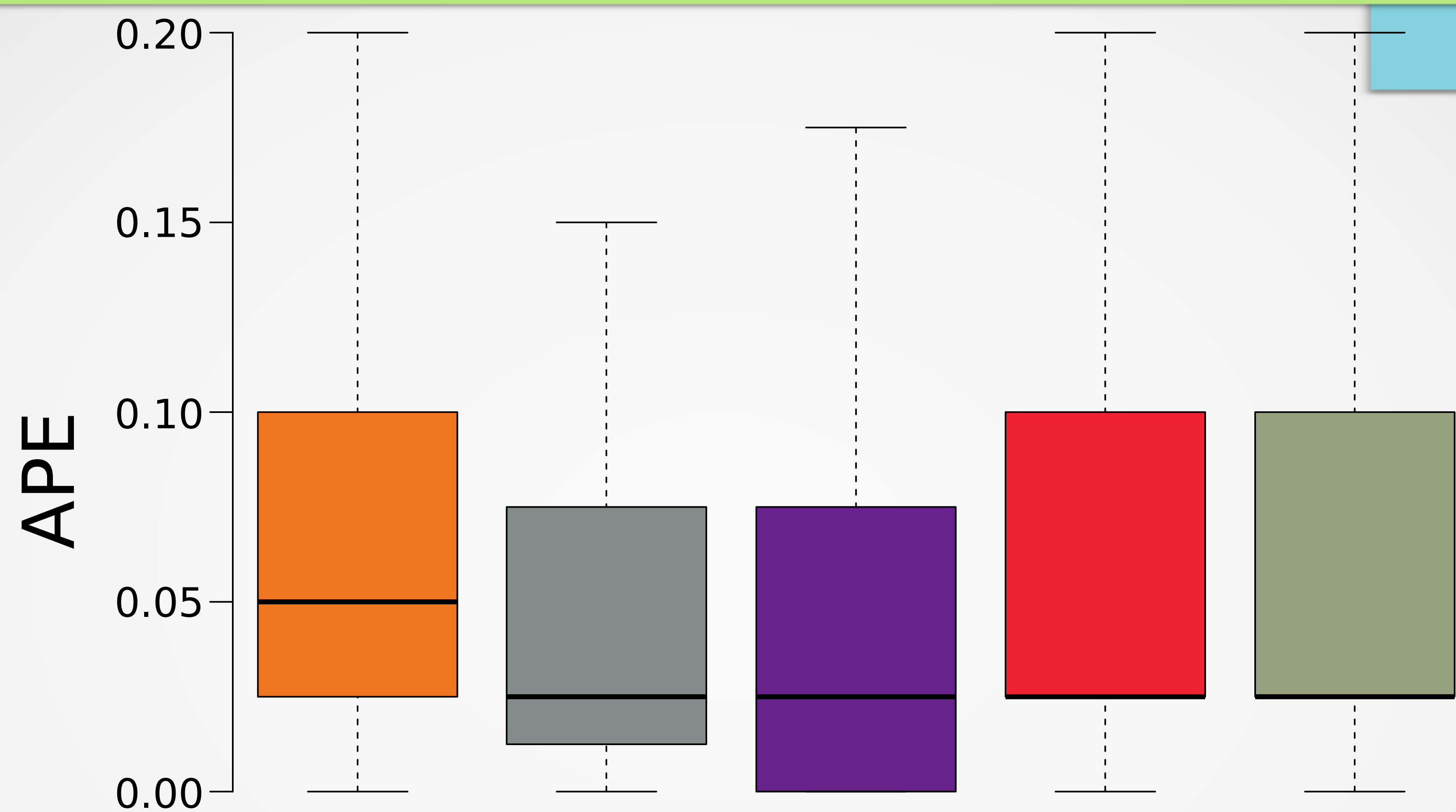
# Forecasting future views (2)




[Pinto et al WSDM'13]

| | MLR1 | MLR2 |
|---|---|---|
| History: | ✓ | ✓ |
| Viral potential: | | |
| Promo. volume: | | ✓ |
| Promo. timing: | | ✓ |

# Forecasting future views (2)

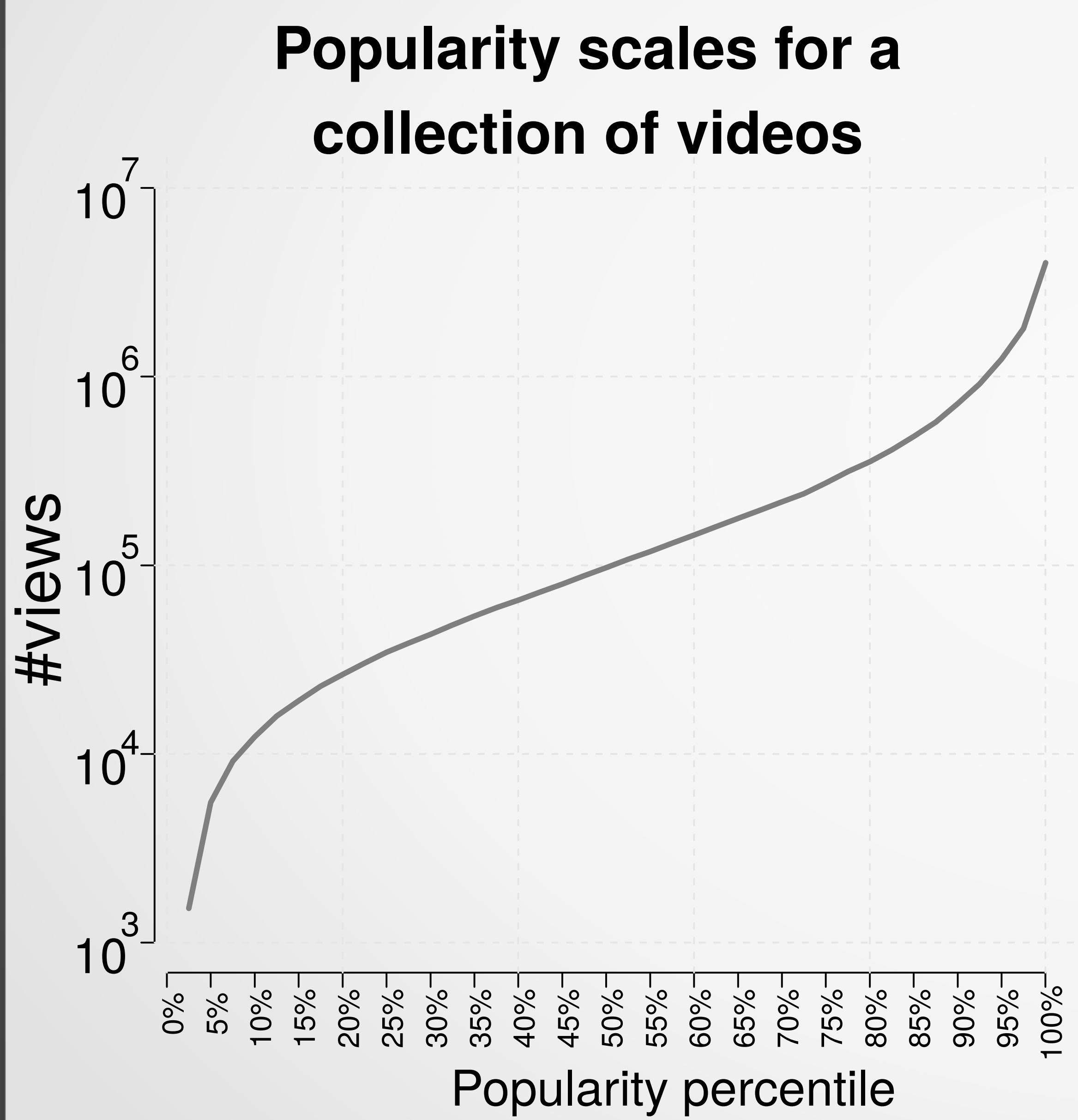



[Rizoiu et al WWW‘17]

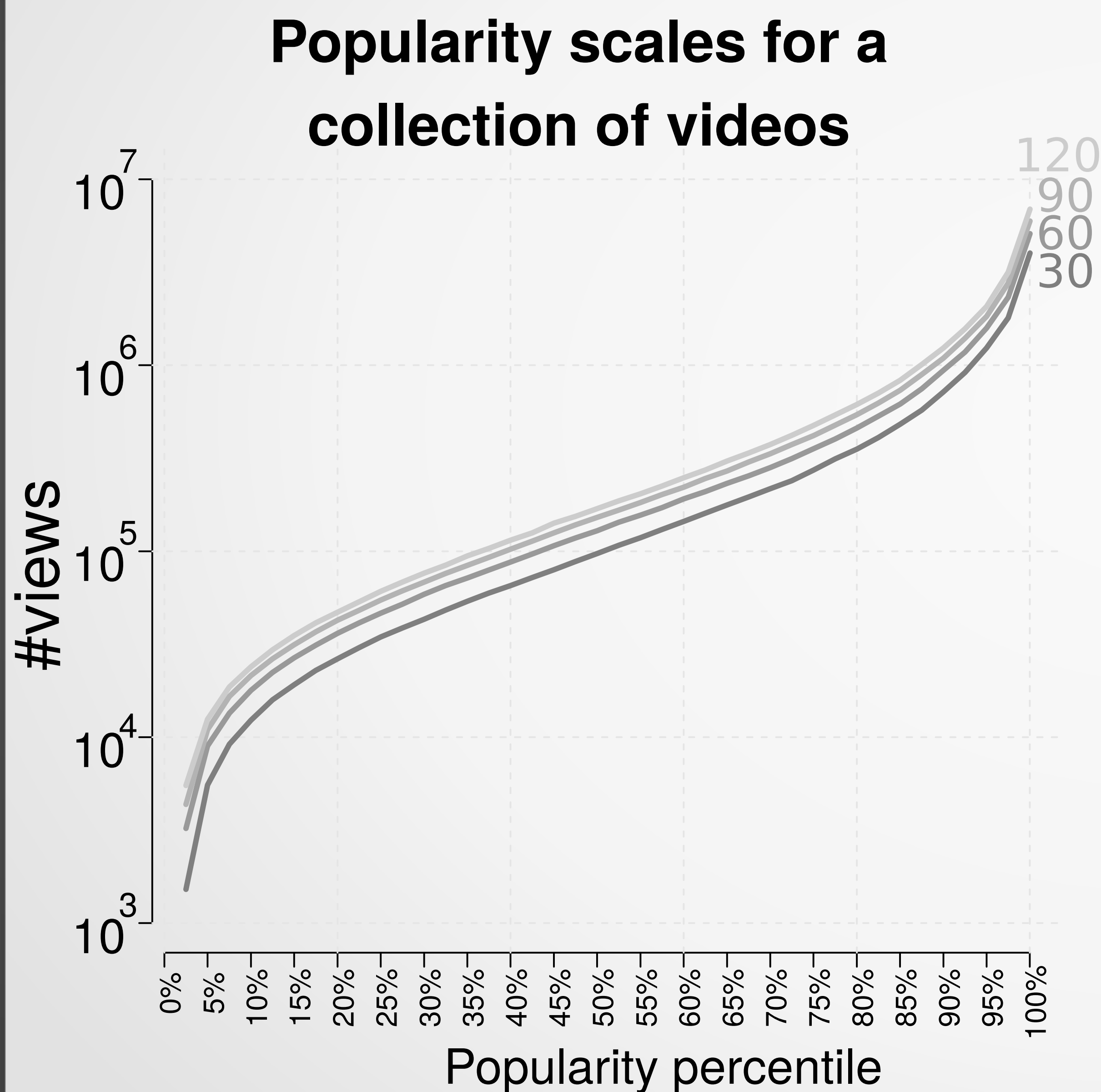
| | P1 | P2 | P3 | MLR1 | MLR2 |
|---|---|---|---|---|---|
| History: | ✓ | ✓ | ✓ | ✓ | ✓ |
| Viral potential: | ✓ | ✓ | ✓ | | |
| Promo. volume: | | ✓ | ✓ | | ✓ |
| Promo. timing: | | | ✓ | | ✓ |

[Pinto et al WSDM’13]

# Popularity scales over time

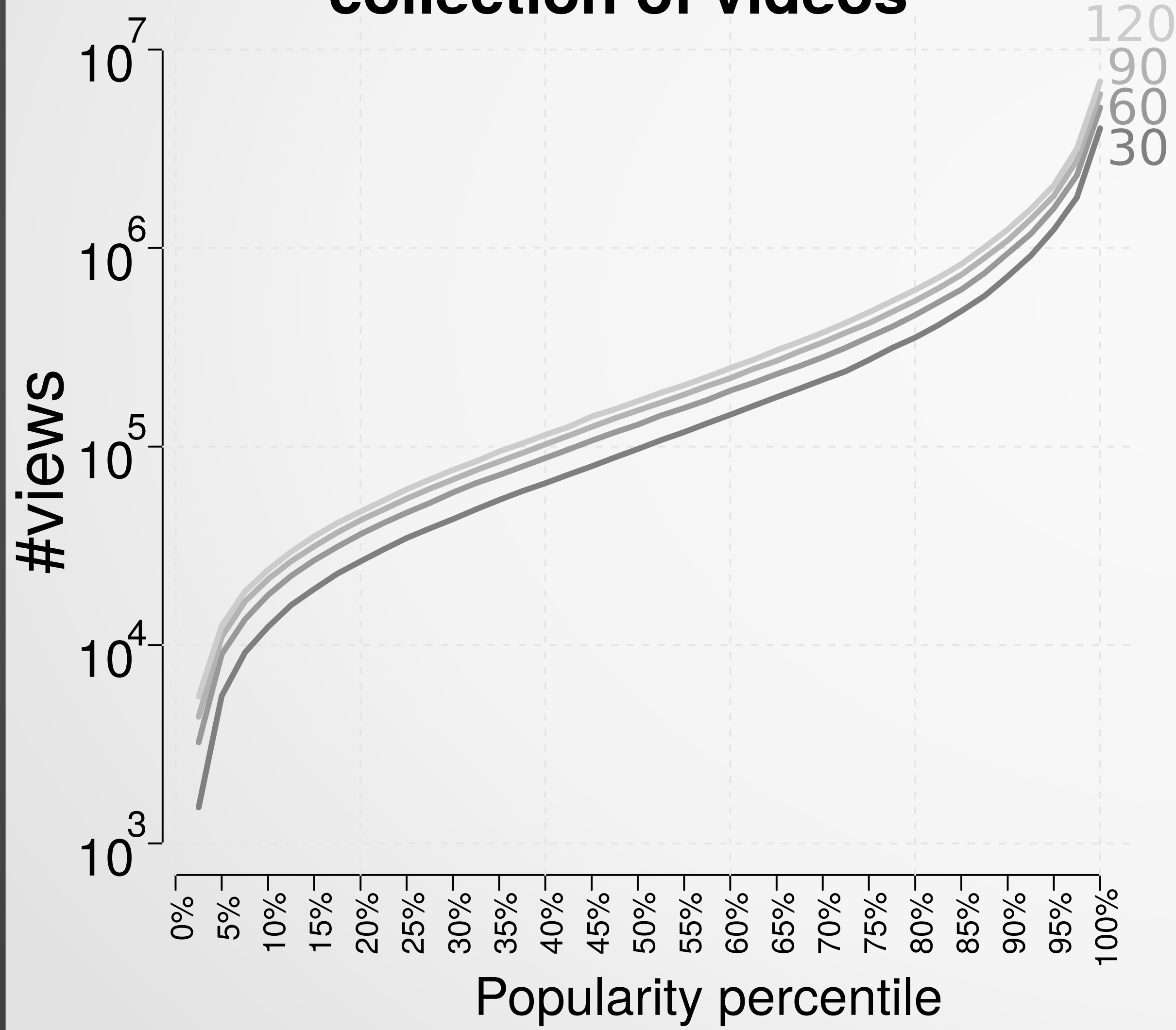


# Popularity scales over time

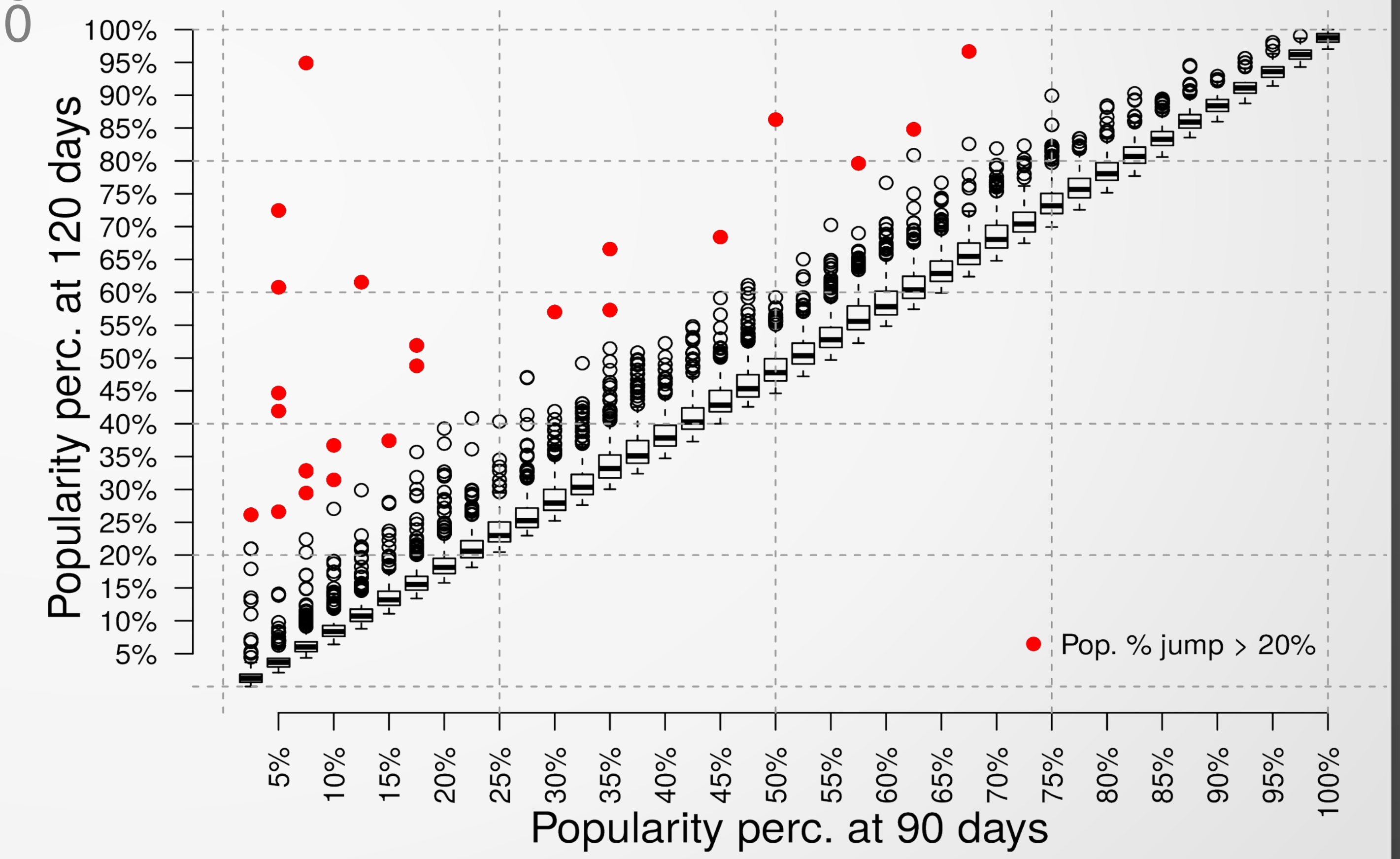


# Popularity scales over time

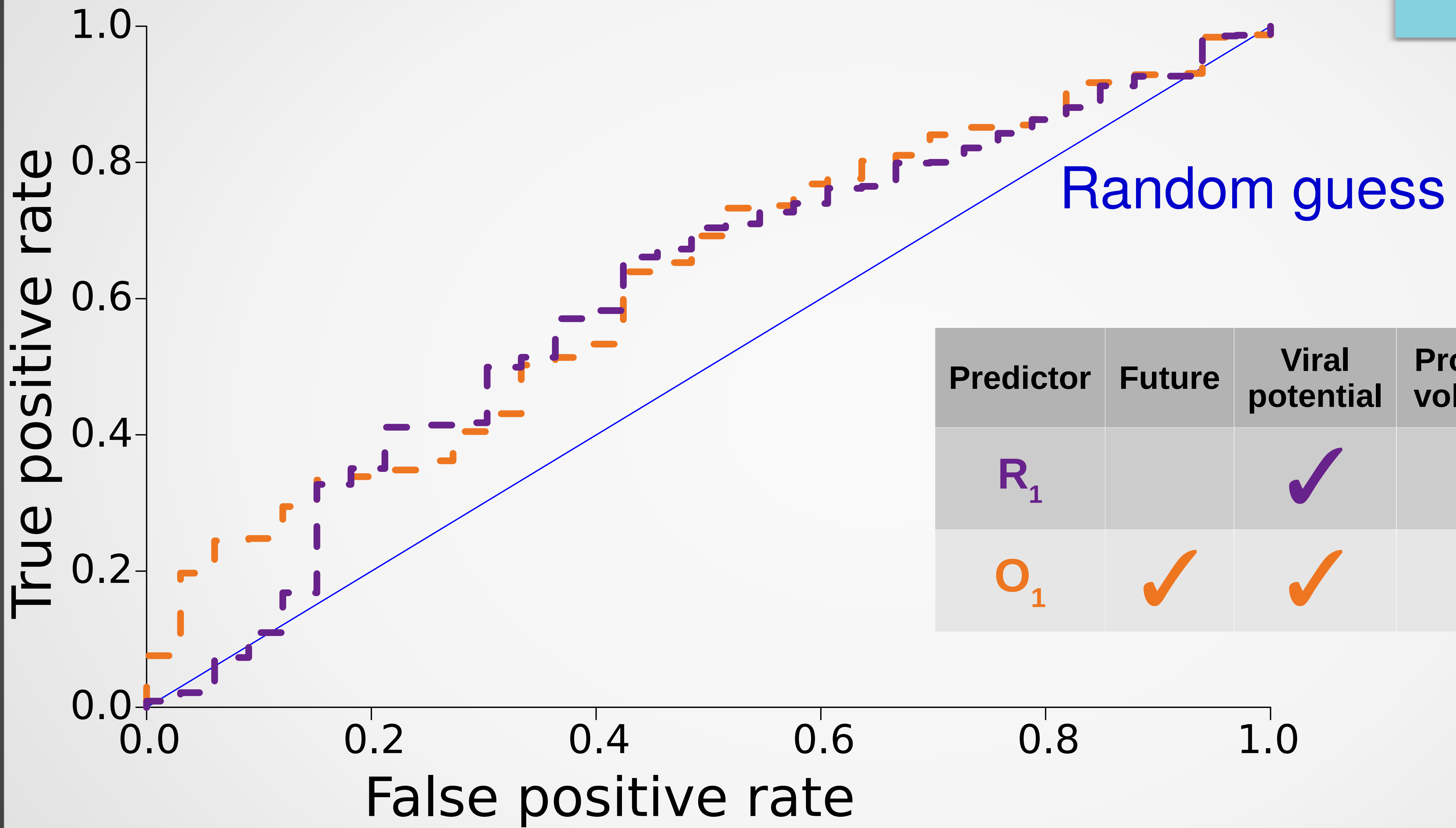
Popularity scales for a collection of videos




Individual video pop. % at 90 days vs. 120 days

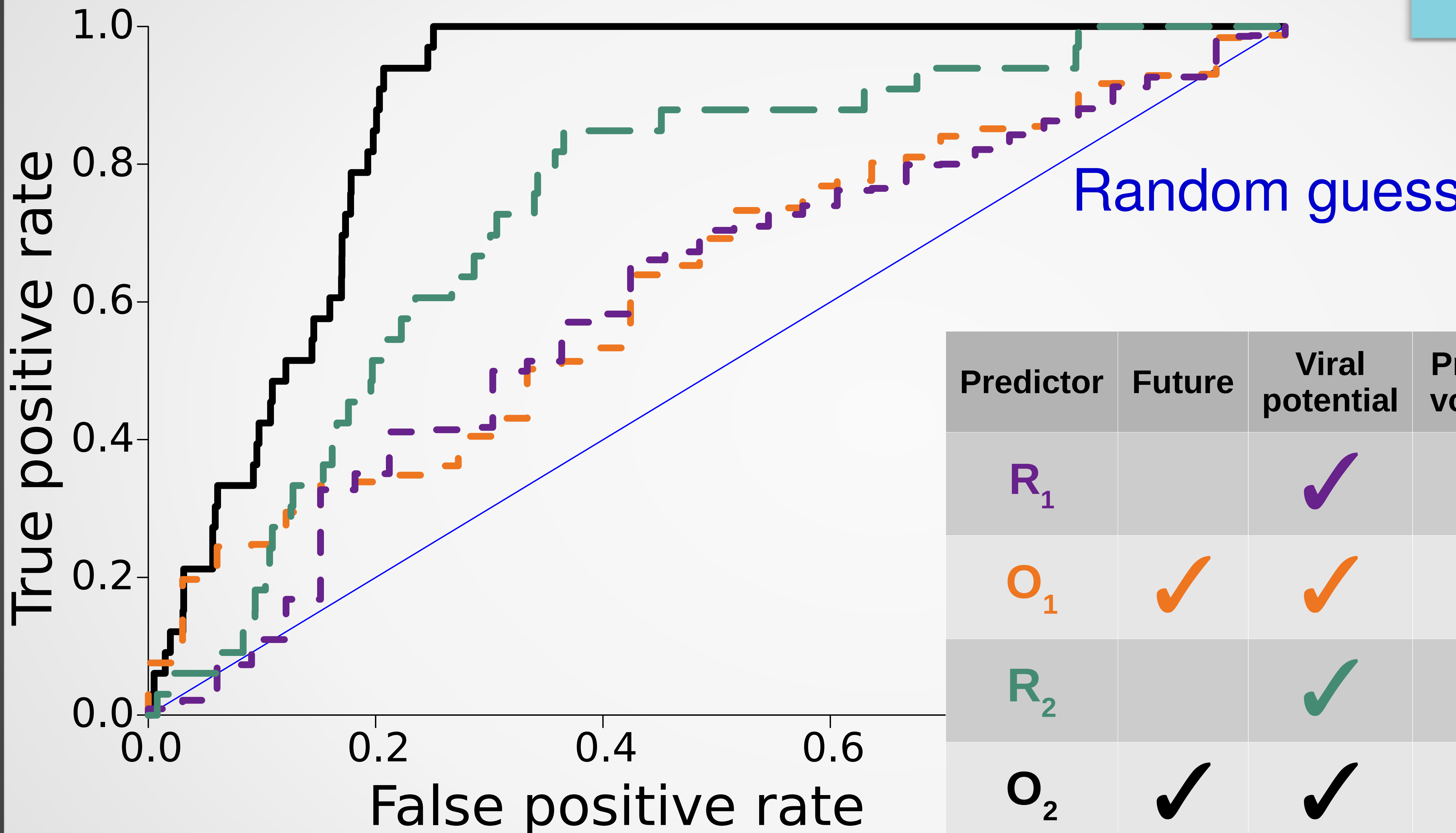


# ROC curves for videos that jump



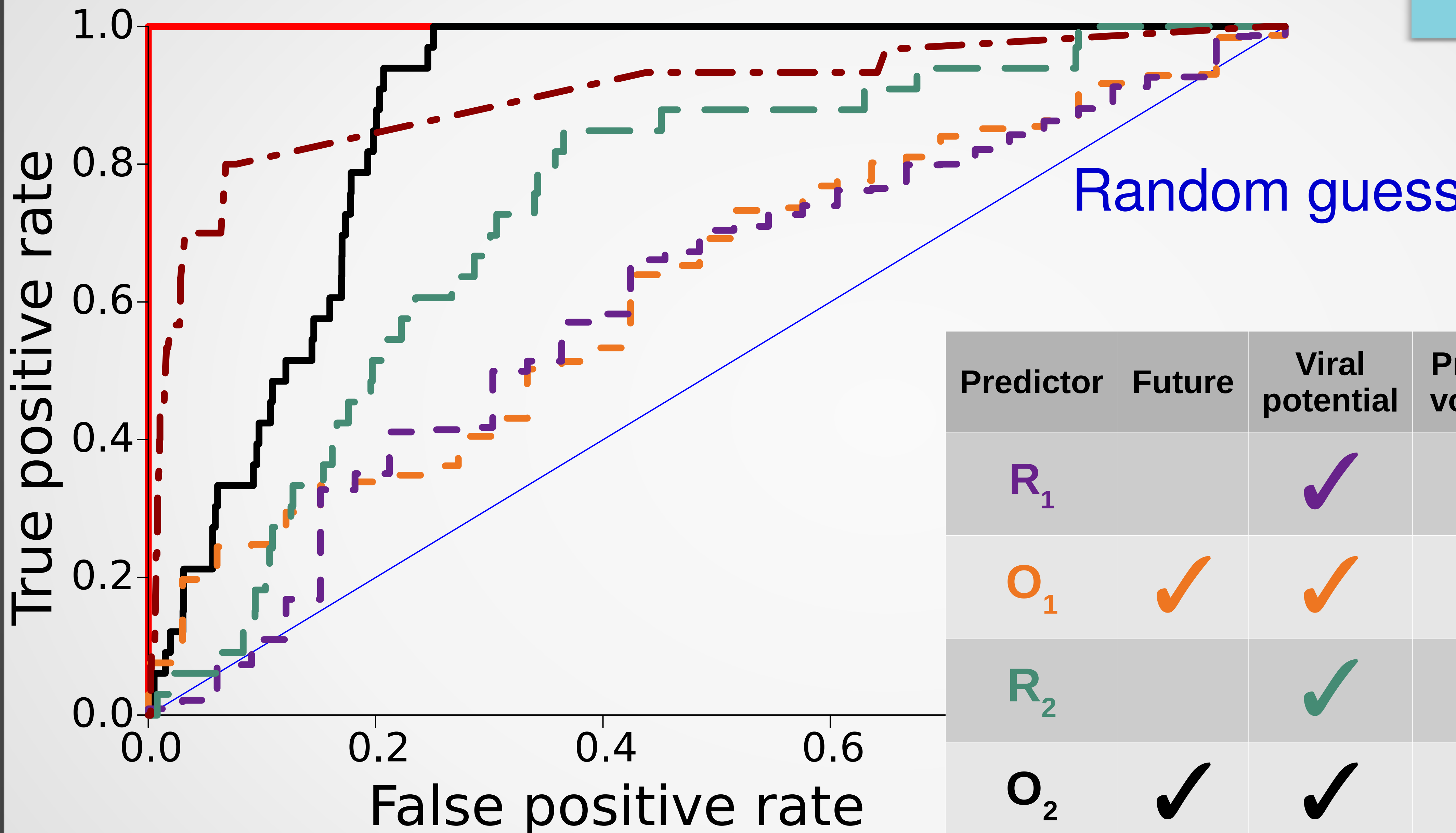

| Predictor | Future | Viral potential | Promo. volume | Pop. scale position |
|---|---|---|---|---|
| $R_1$ | | ✓ | | |
| $O_1$ | ✓ | ✓ | | |

# ROC curves for videos that jump



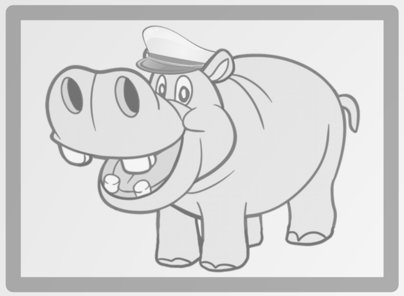

| Predictor | Future | Viral potential | Promo. volume | Pop. scale position |
|---|---|---|---|---|
| $R_1$ | | ✓ | | |
| $O_1$ | ✓ | ✓ | | |
| $R_2$ | | ✓ | ✓ | |
| $O_2$ | ✓ | ✓ | ✓ | |

# ROC curves for videos that jump

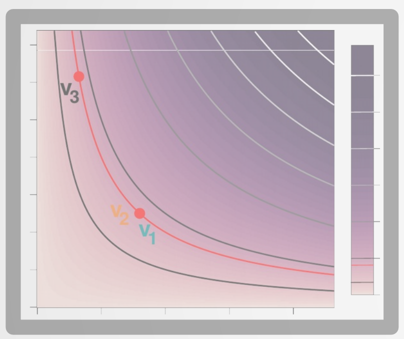



| Predictor | Future | Viral potential | Promo. volume | Pop. scale position |
|---|---|---|---|---|
| $R_1$ | | ✓ | | |
| $O_1$ | ✓ | ✓ | | |
| $R_2$ | | ✓ | ✓ | |
| $O_2$ | ✓ | ✓ | ✓ | |
| $R_3$ | | ✓ | ✓ | ✓ |
| $O_3$ | ✓ | ✓ | ✓ | ✓ |

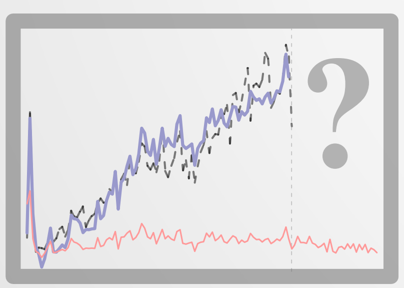
# Presentation outline



Modeling popularity with HIP



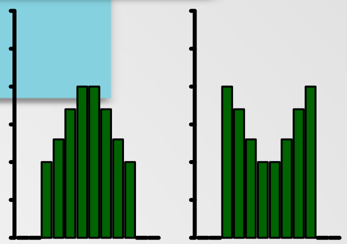
Content virality and maturity time




Forecasting popularity under promotion

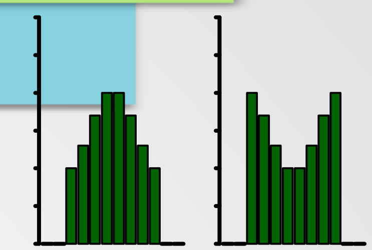
When does promotion timing matter?
Why do people prefer constant promotion?

# Designing promotion schedules

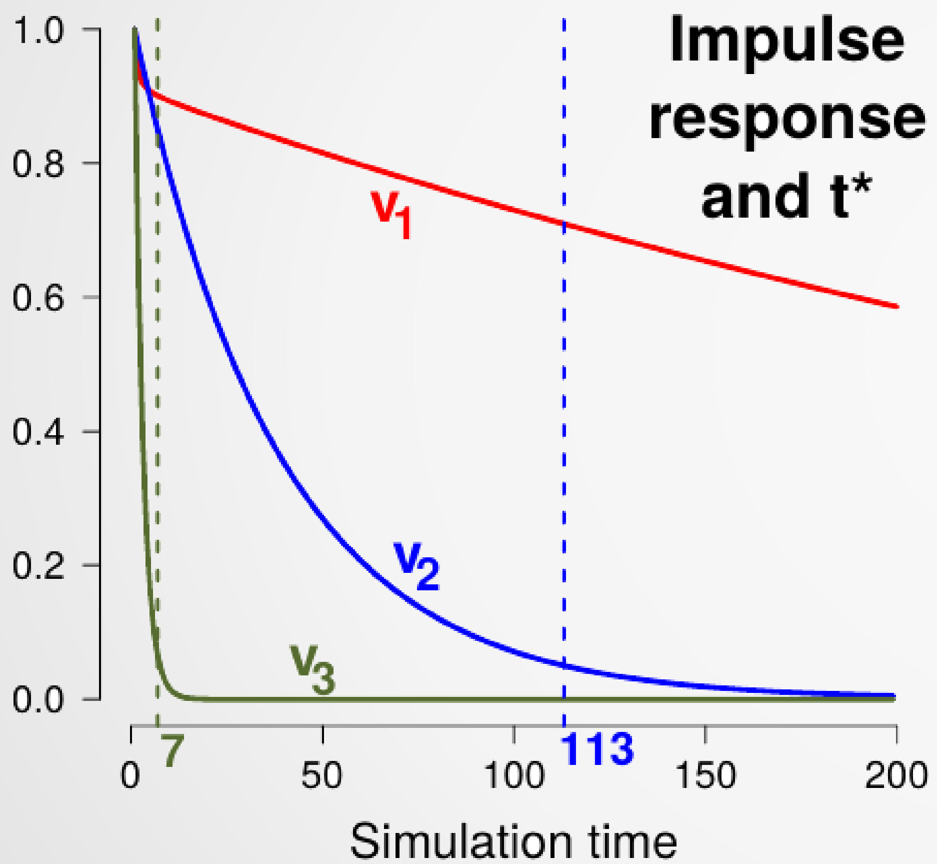


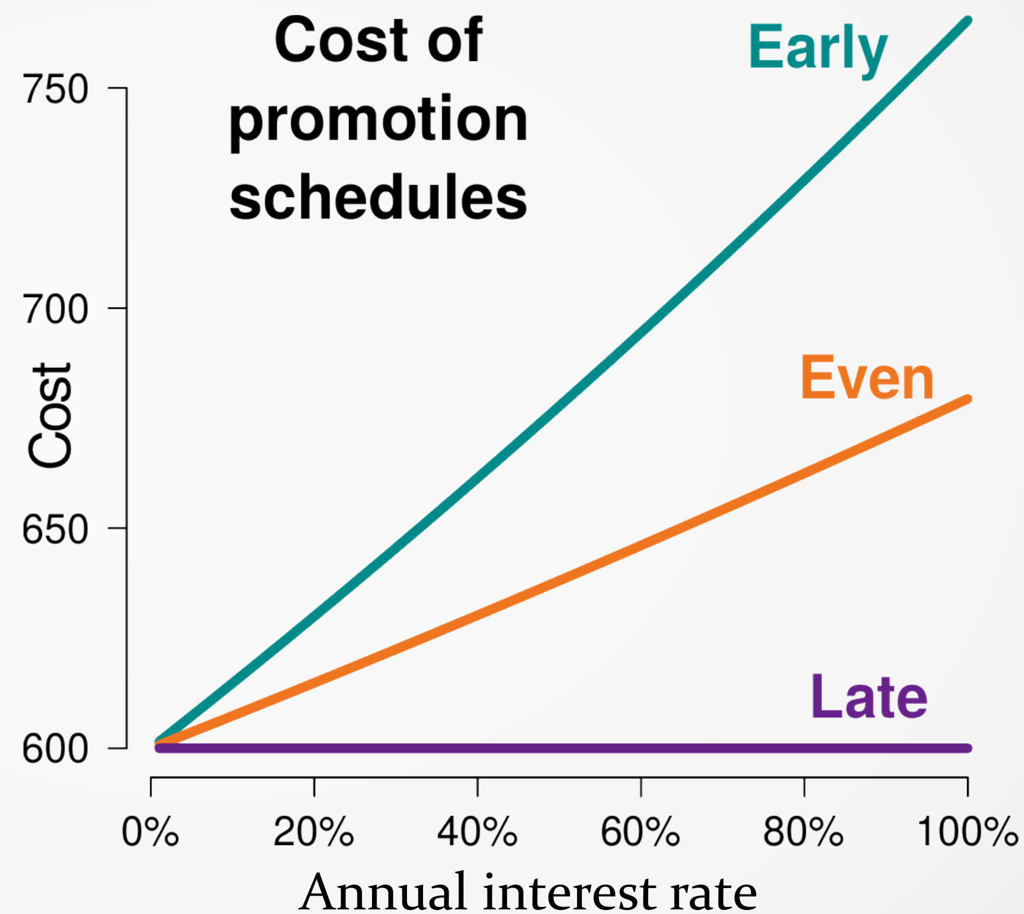
LTI corollary: **same budget, same return!**

# Designing promotion schedules

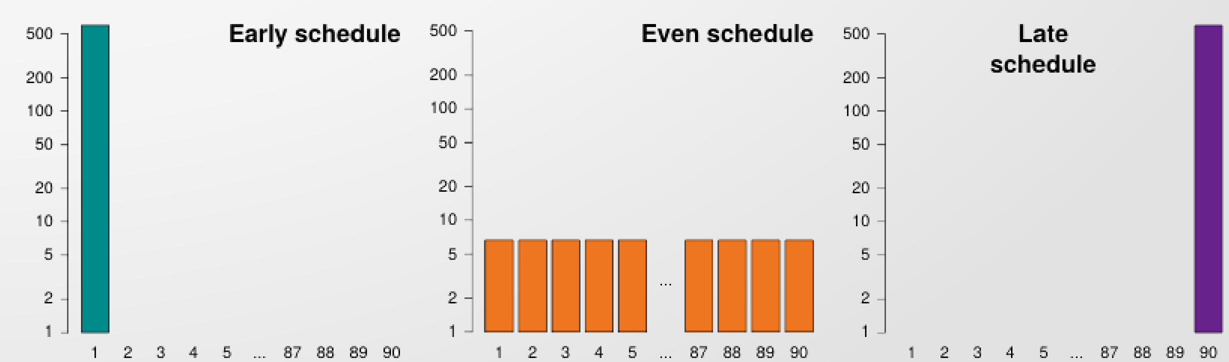


LTI corollary: **same budget, same return!**

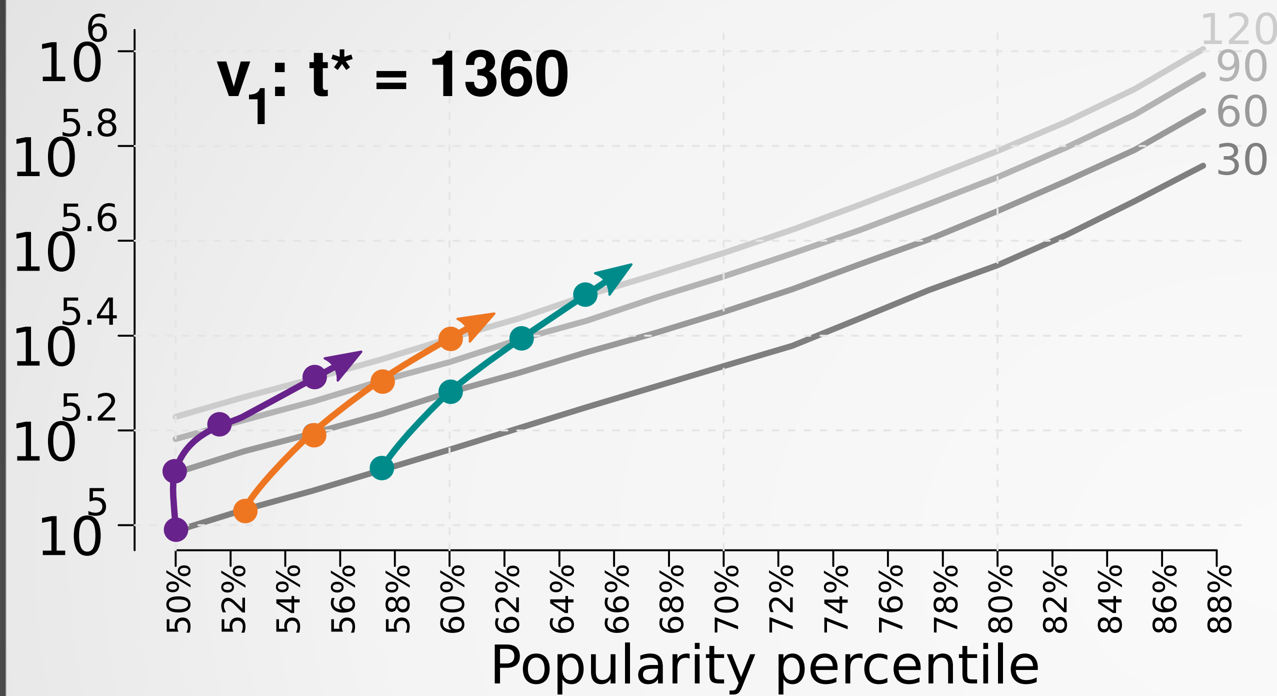


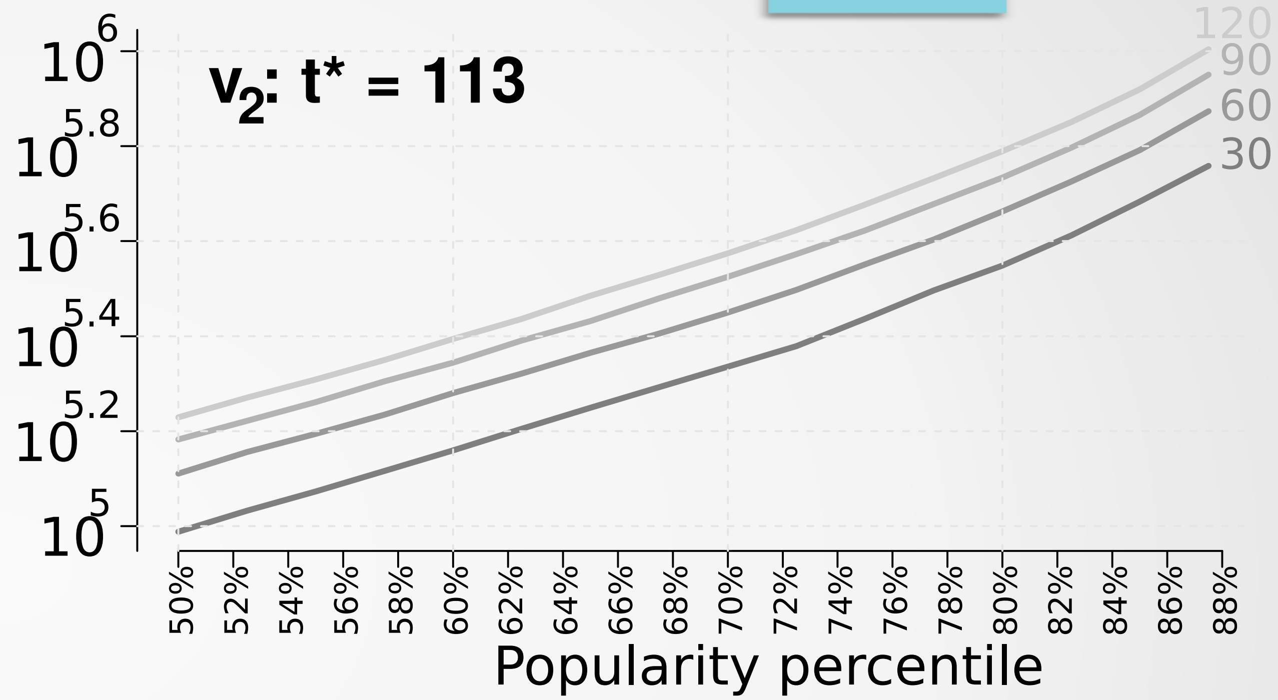



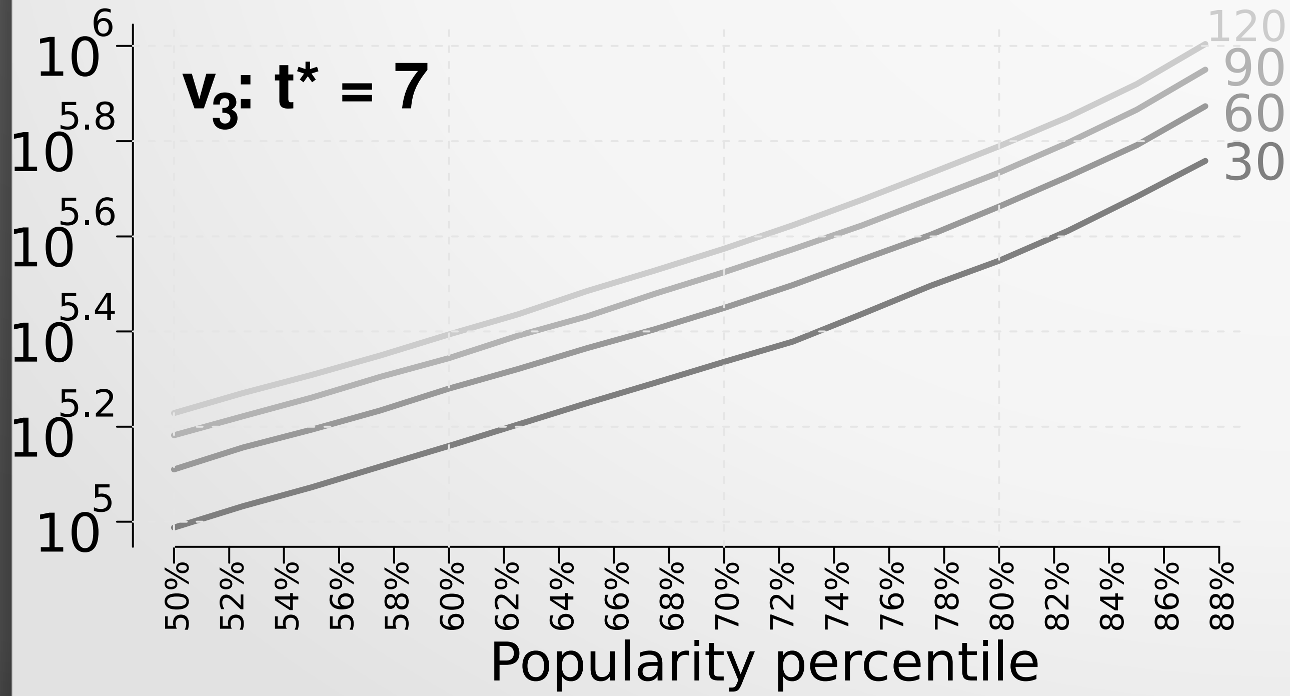

Compounding interest: $cost = (1+a)^k$

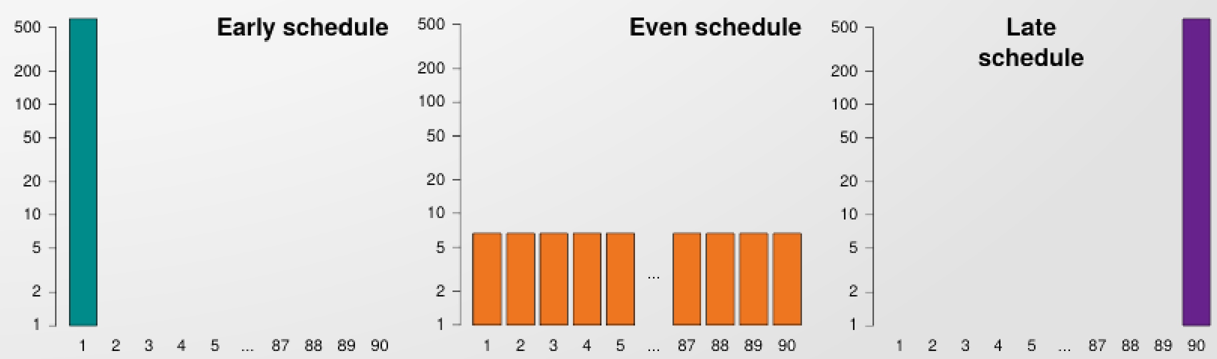


# Interplay of 2 temporal factors

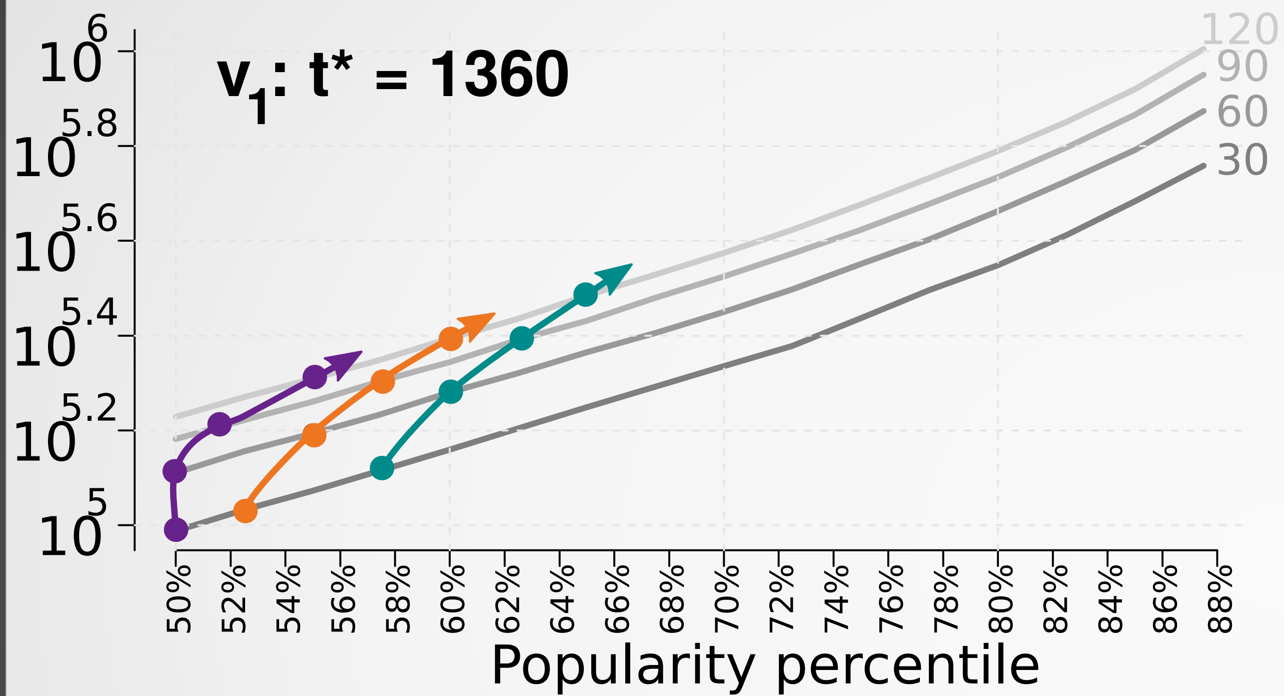


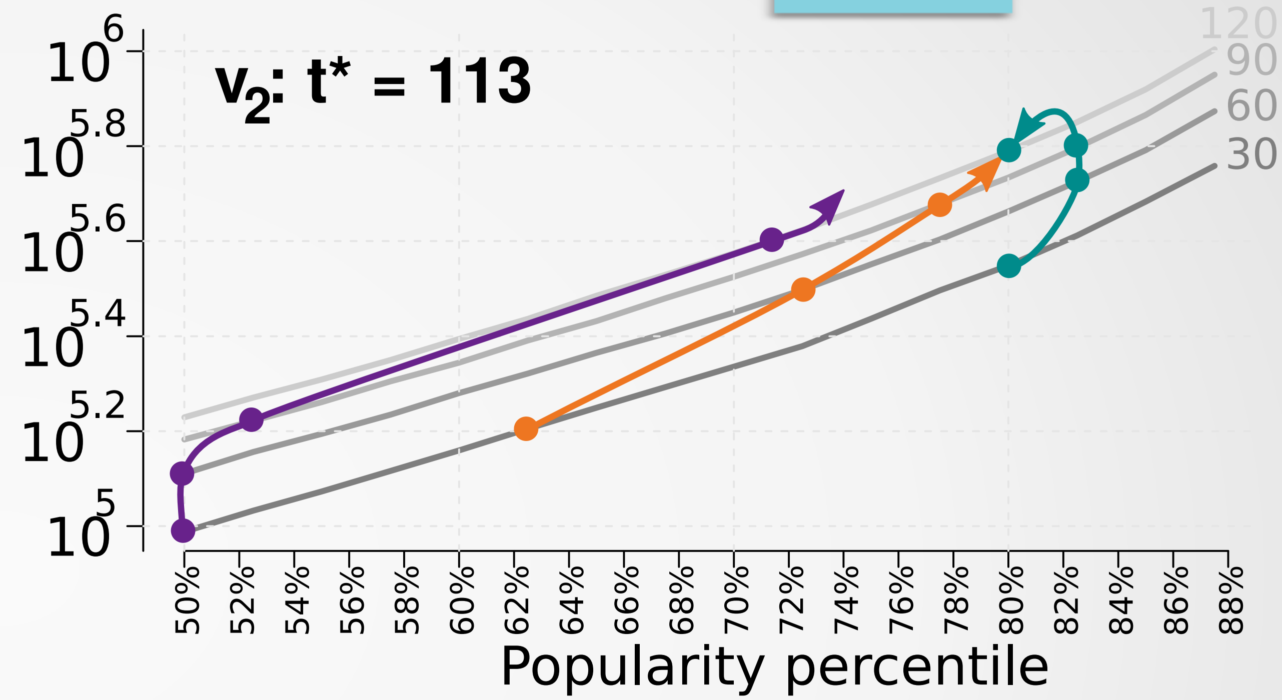



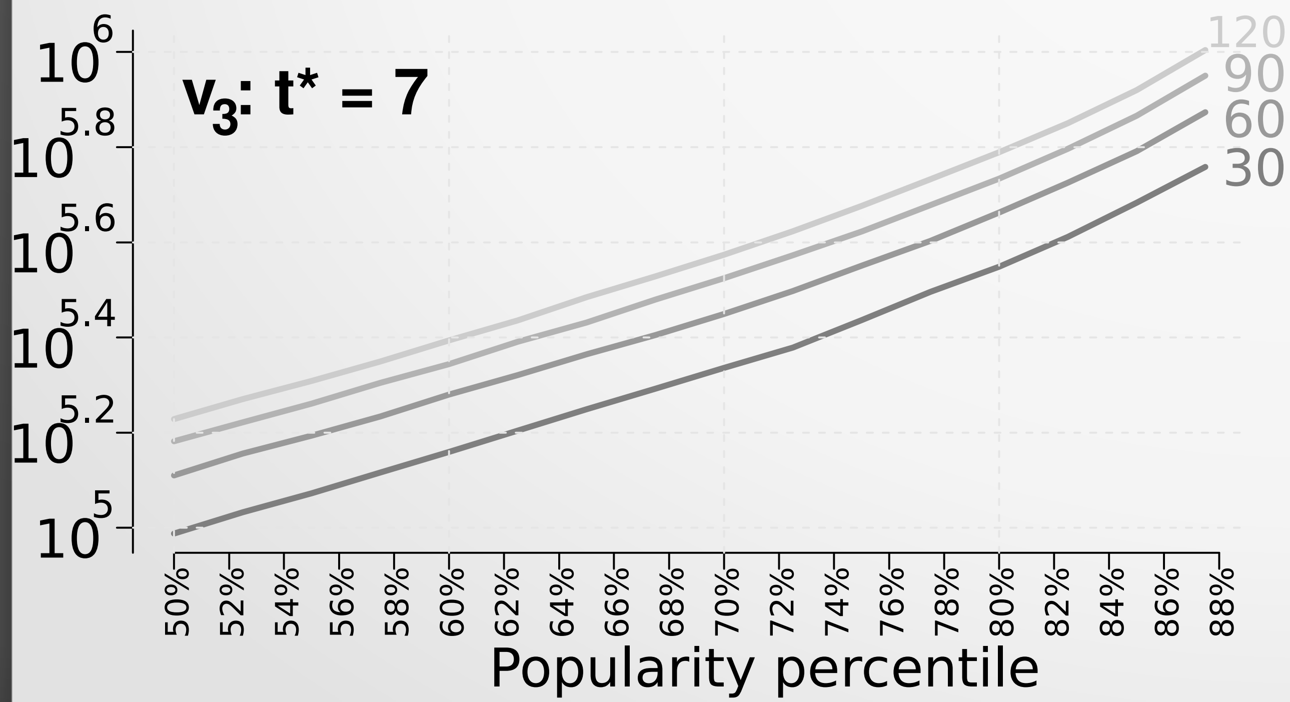



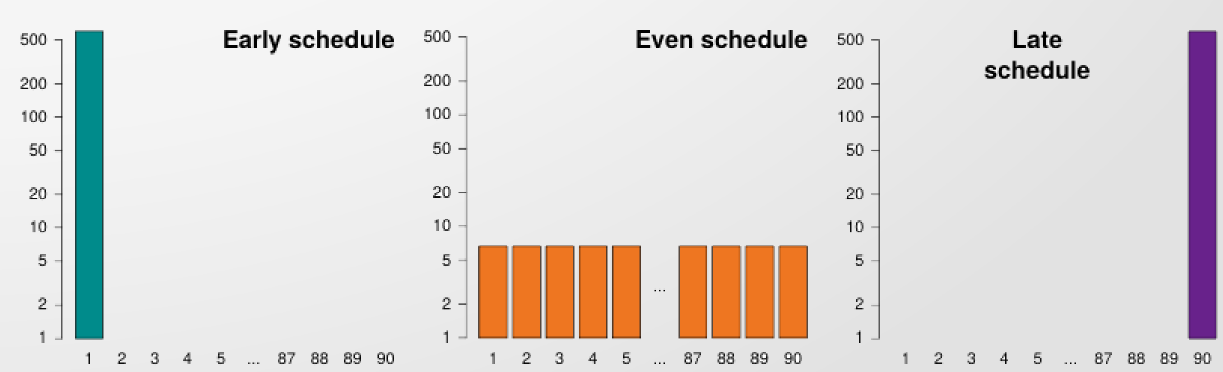



# Interplay of 2 temporal factors

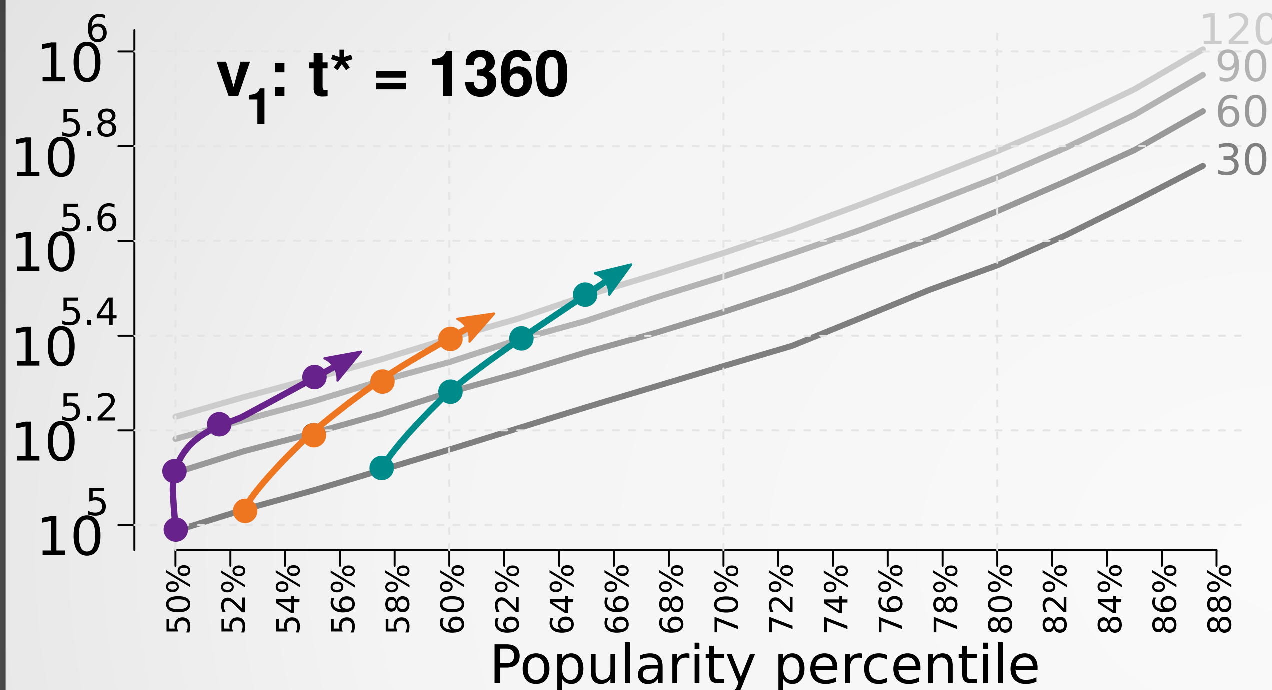


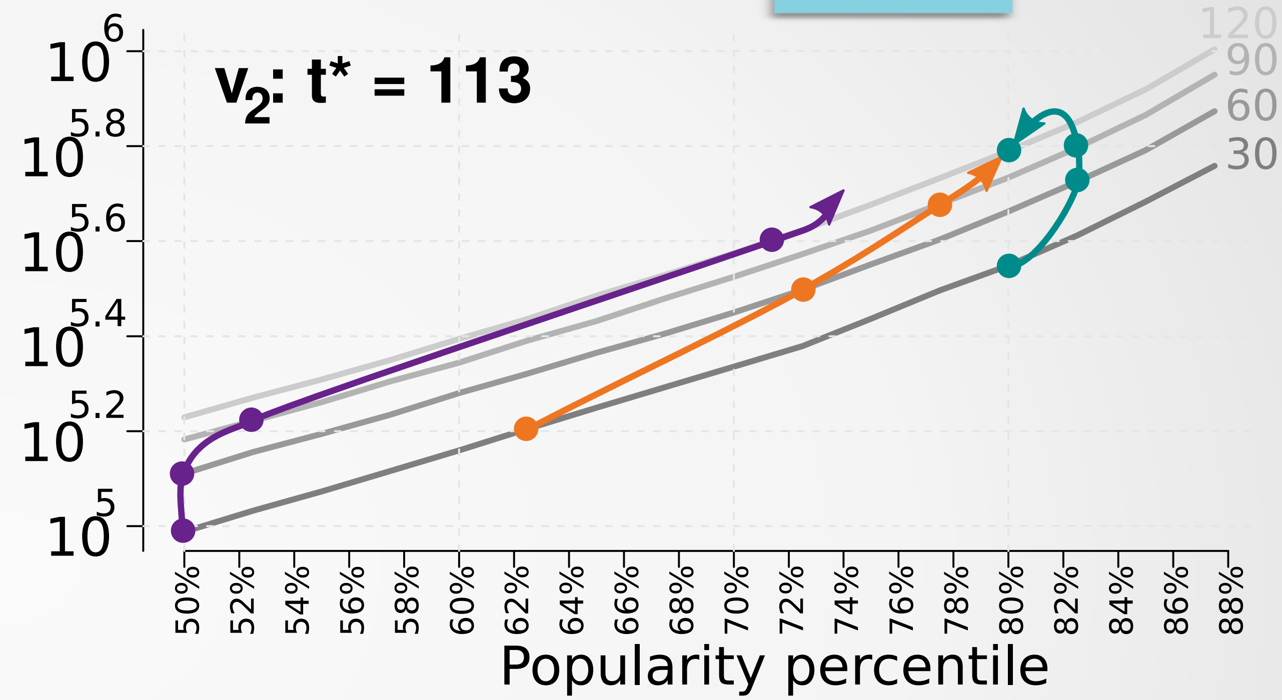



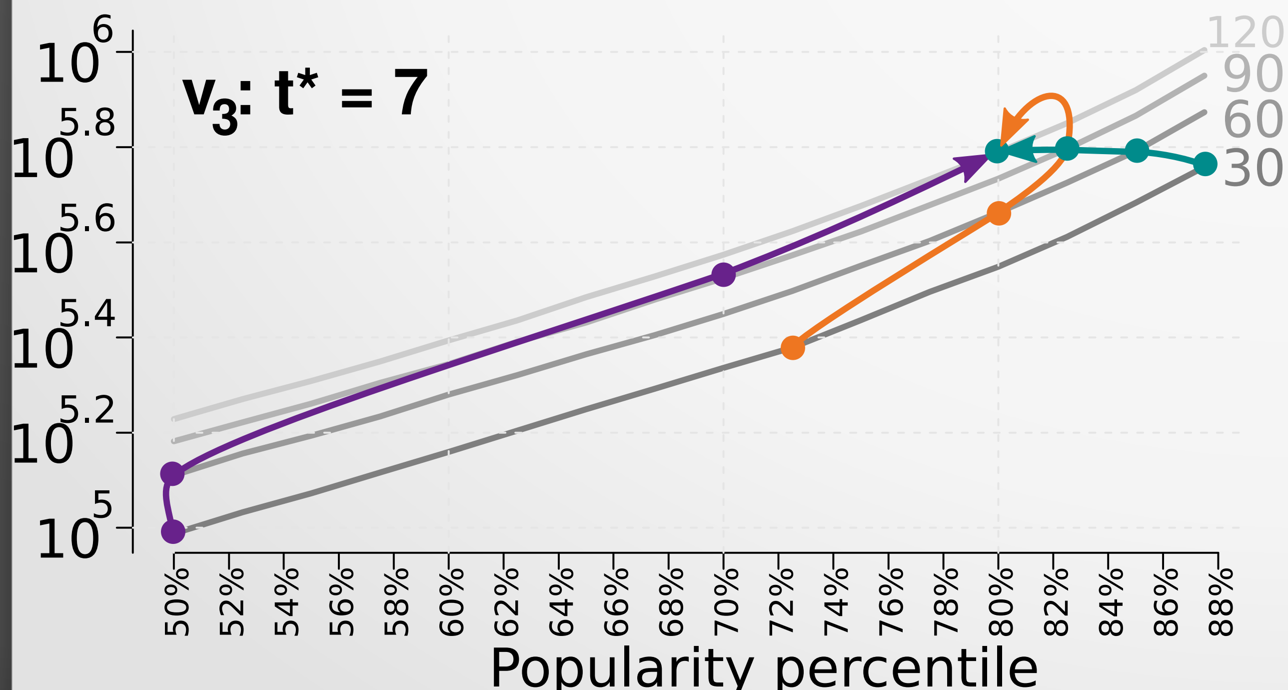



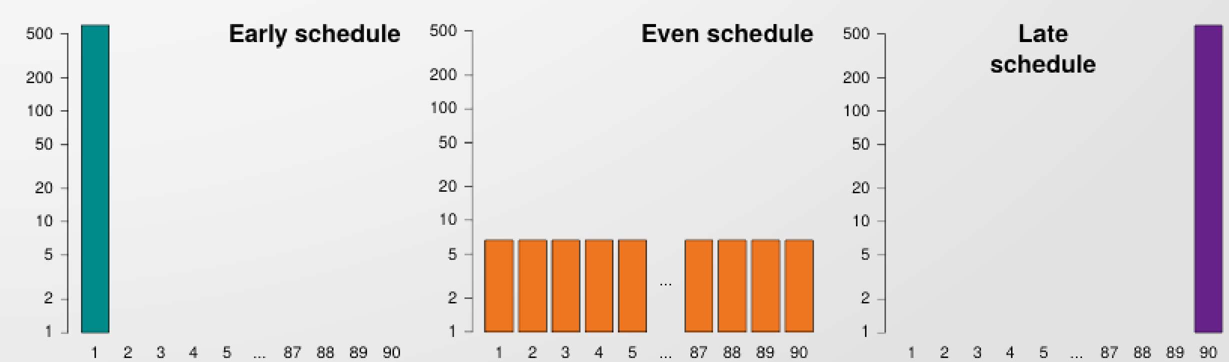



# Interplay of 2 temporal factors



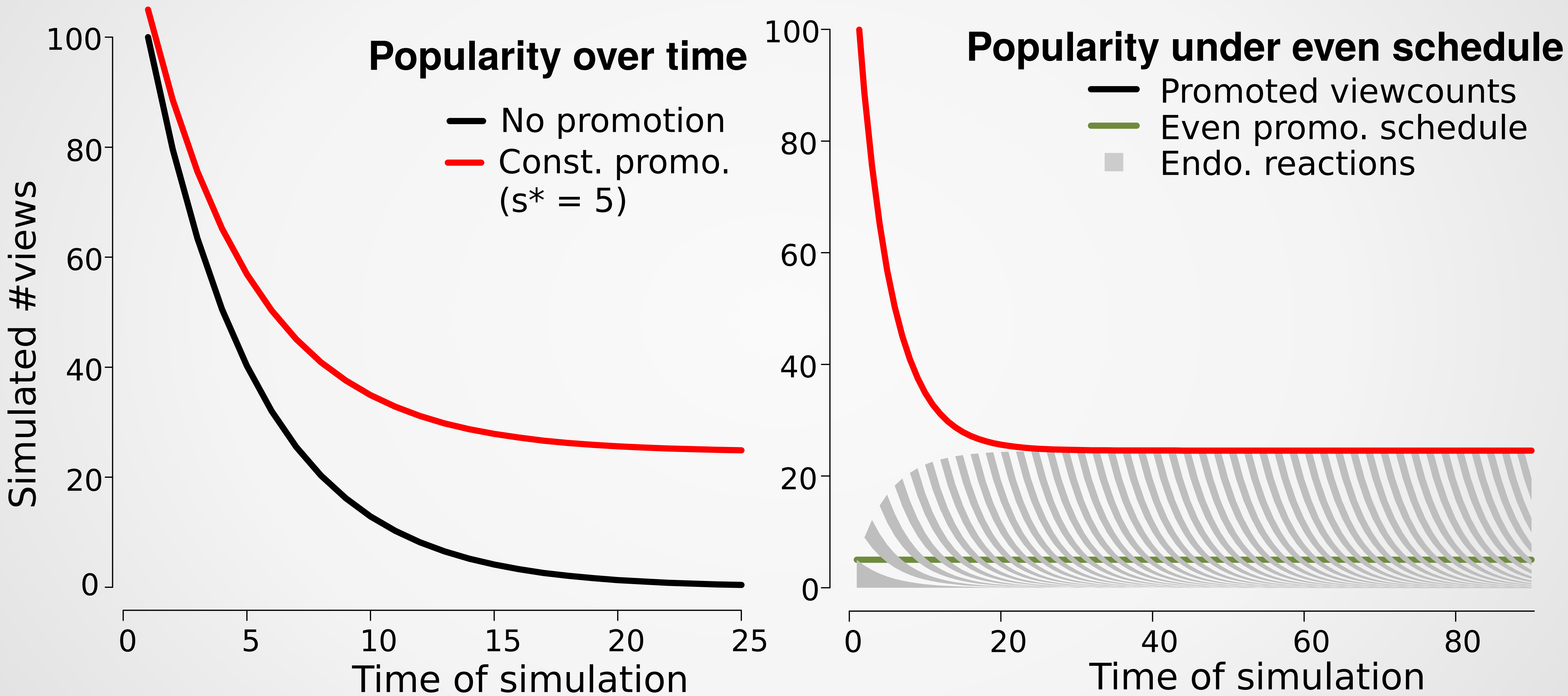



# Why is constant promotion desirable?

LTI corollary: **the effects of daily promotion add up over time!**

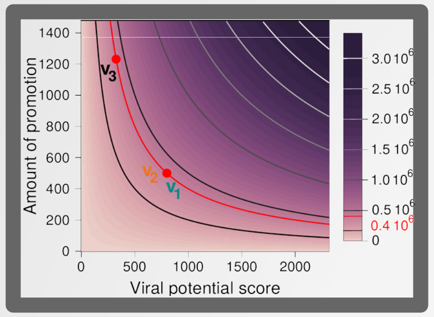
Explains why TV commercials appear at fixed intervals, every day.

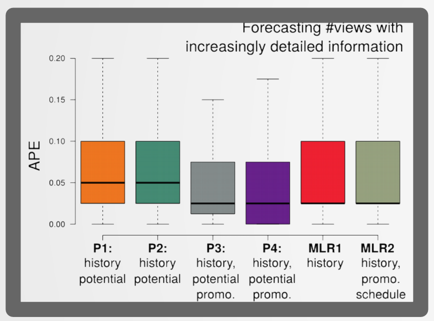
# Memory lengthening through promotion

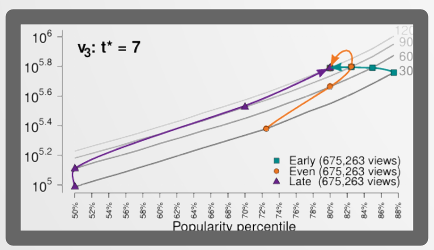



Constant promotion leads to an apparent **memory lengthening.**

# Summary



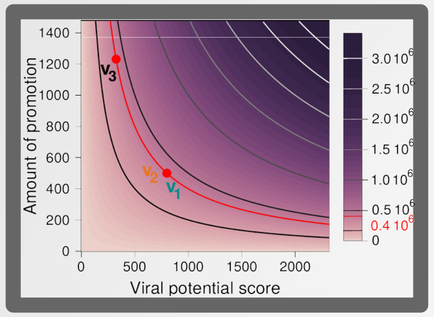




Two measures: *virality score* and *maturity time*

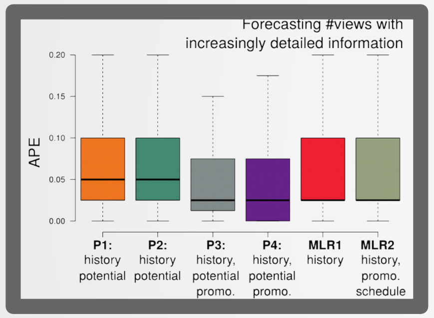
Important factors for forecasting popularity: *virality score*, *promotion volume* and *popularity scale position*

Maturity time influences the cost-effectiveness of promotion schedules

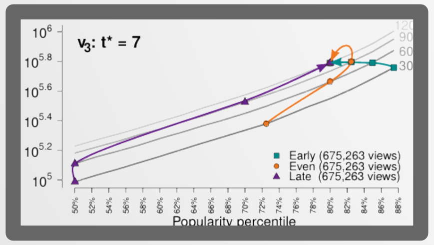
# Summary




Two measures: *virality score* and *maturity time*




Important factors for forecasting popularity: *virality score*, *promotion volume* and *popularity scale position*




Maturity time influences the cost-effectiveness of promotion schedules

**Limitations & future work:**

Average over network;

Reaction to past and future promotions is the same.

# Thank you!

## Links:

Papers, code, dataset
and interactive visualizer:

https://github.com/andrei-rizoiu/hip-popularity

*Referece*:

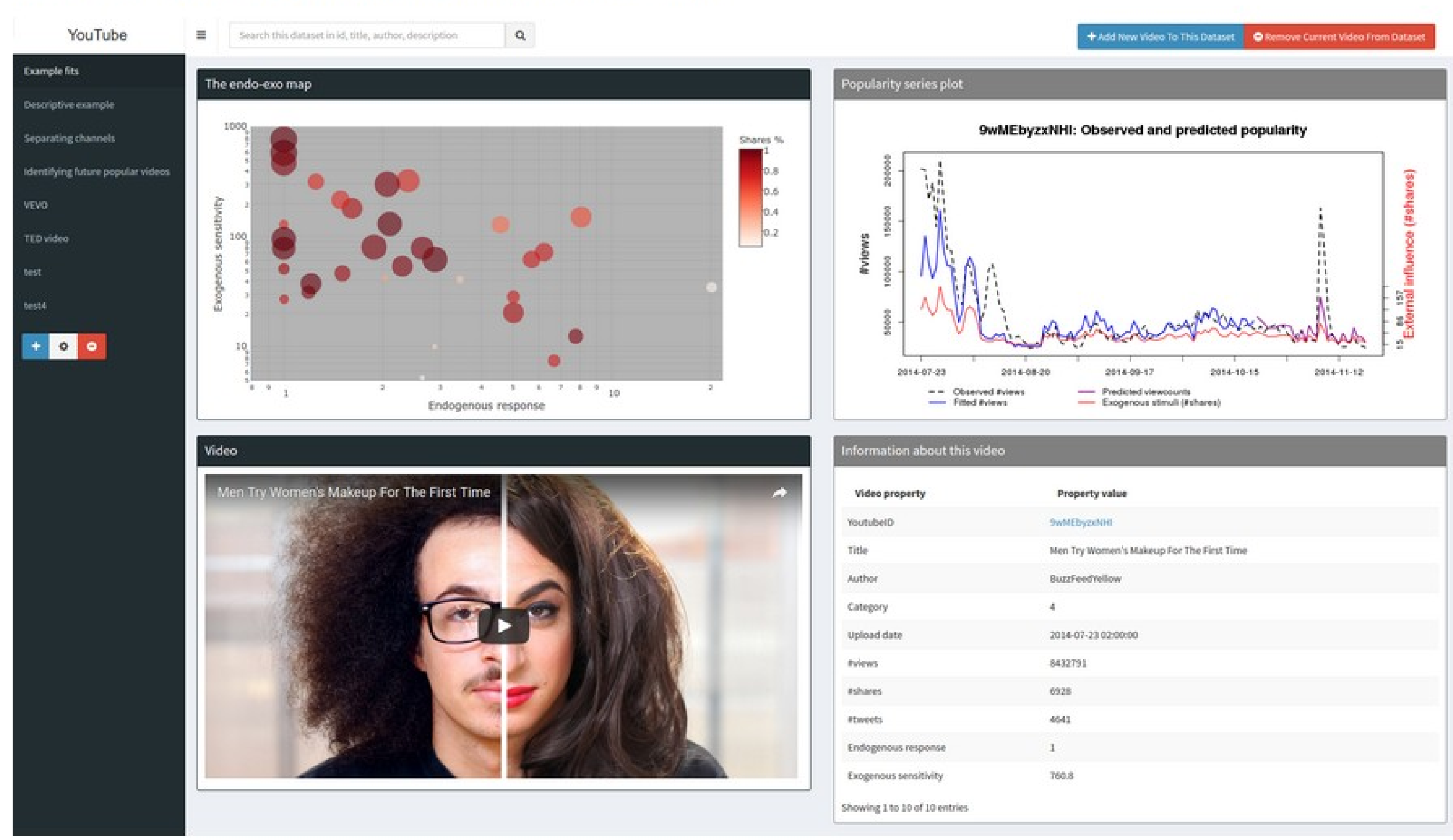
Rizoiu, M.-A., Xie, L., Sanner, S., Cebrian, M., Yu, H., & Van Hentenryck, P. (2017). **Expecting to be HIP: Hawkes Intensity Processes for Social Media Popularity**. In Proceedings of the *International Conference on World Wide Web 2017*, pp. 1-9. Perth, Australia. doi: 10.1145/3038912.3052650

pdf at arxiv with supplementary material

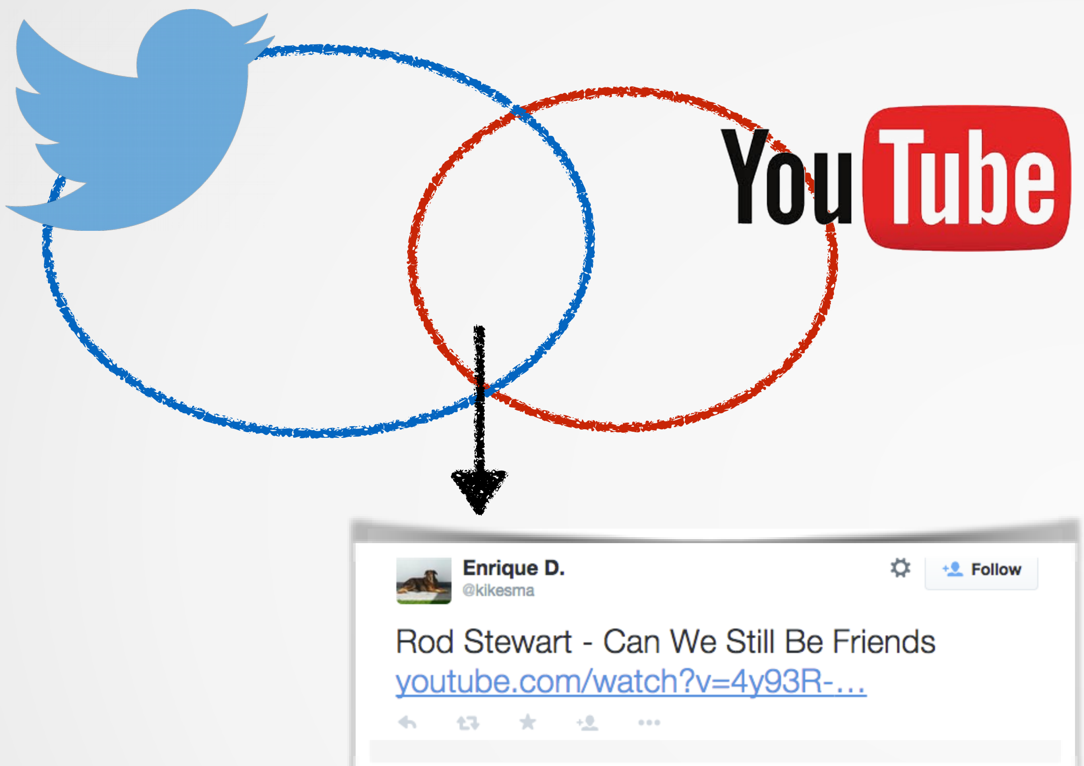
## HIP visualization system

This is an *interactive* visualization of the plots in the paper: the endo-exo map, observed and fitted popularity series and video metadata. It has additional visualizations of TED videos and VEVO musicians. Furthermore, it allows users to add and compare their own videos.

**(access the visualizer by clicking on the thumbnail below)**



# Supp: Twitted videos dataset




| Category | #vids | Category | #vids |
|---|---|---|---|
| Comedy | 865 | Music | 3549 |
| Education | 298 | News & Politics | 1722 |
| Entertainment | 2422 | Nonprofits & Activism | 333 |
| Film & Animation | 664 | People & Blogs | 1947 |
| Gaming | 882 | Science & Technology | 262 |
| Howto & Style | 180 | Sports | 614 |
| Total: | 13,738 | | |

2014.06 - 2014.12
1.061B tweets, 5.89M/day
64.3M users;
81.9M YouTube videos

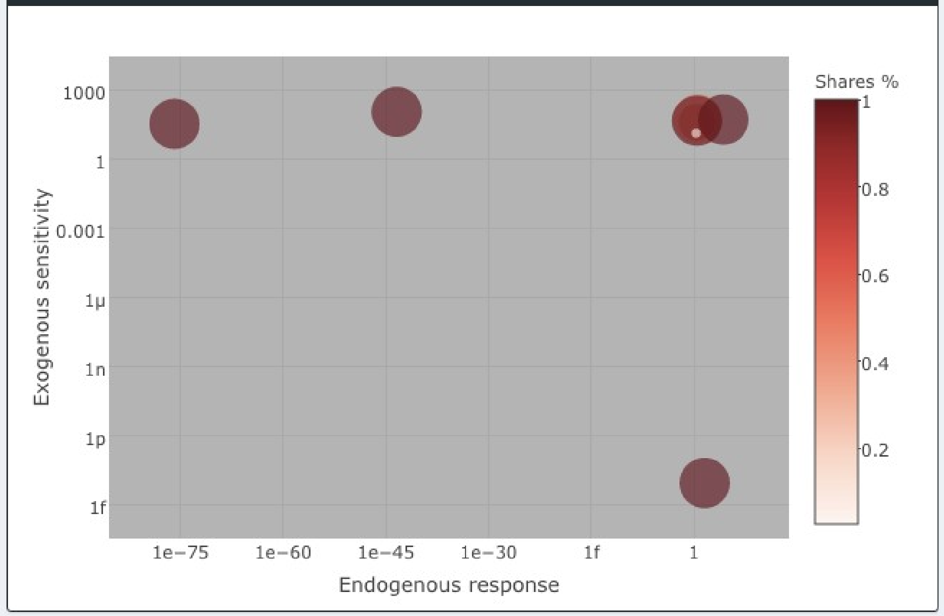
# Supp: when HIP fails the fitting (1)

**Relations between videos:**

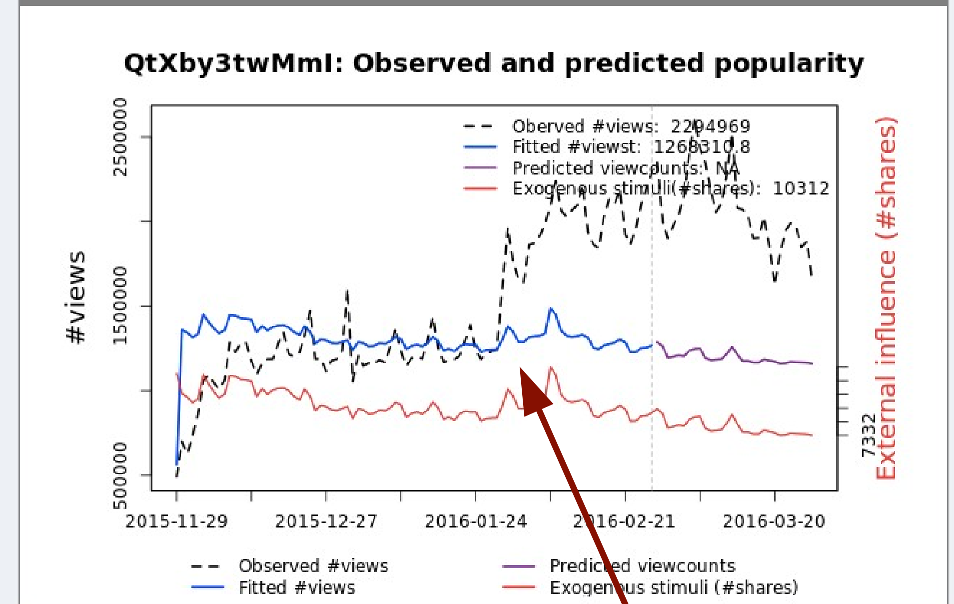
Search this dataset in id, title, author, descrip

+ Add New Video To This Dataset | Remove Current Video From Dataset

### The endo-exo map




### Popularity series plot




### Video


### Information about this video

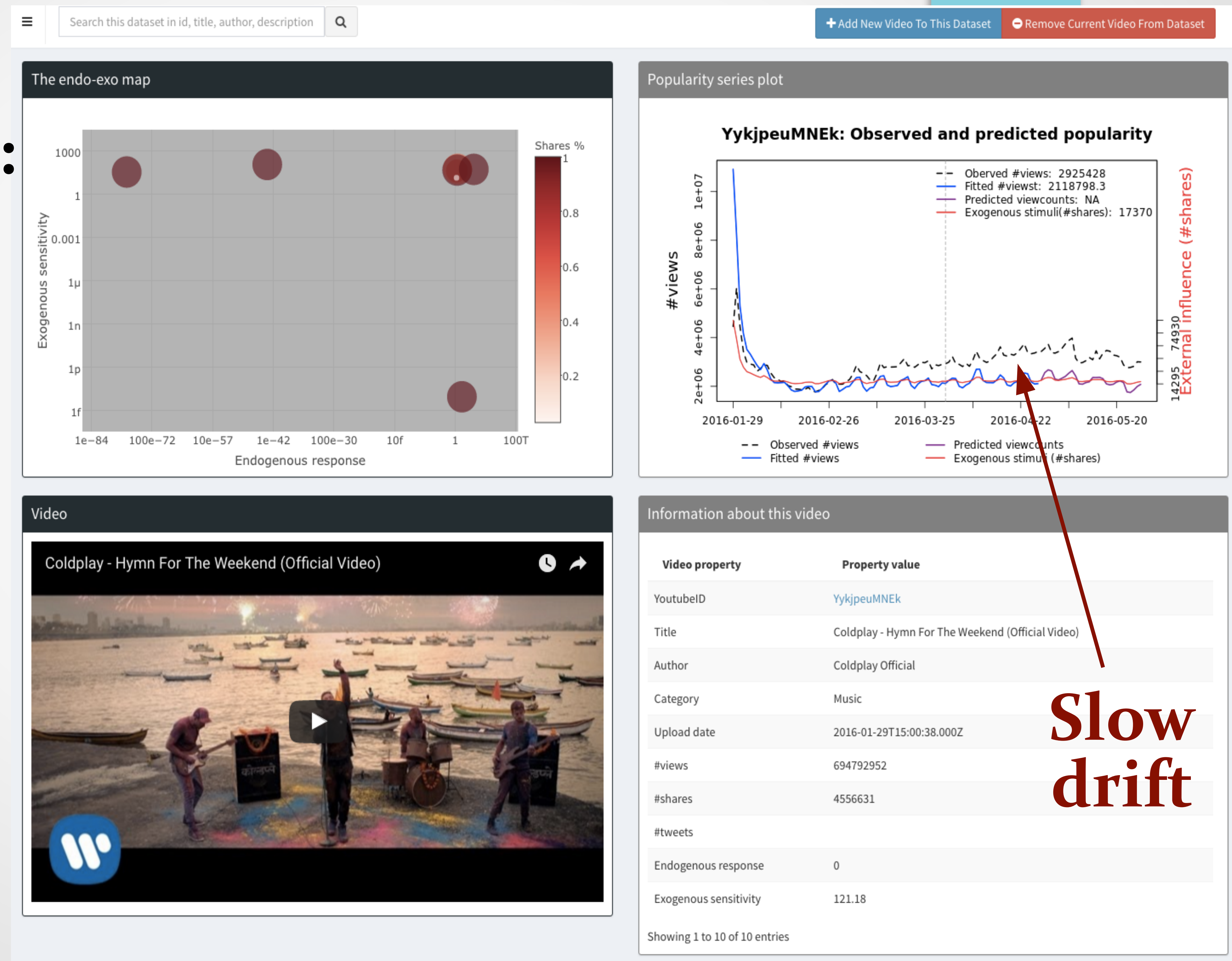
| Video property | Property value |
|---|---|
| YoutubeID | QtXby3twMmI |
| Title | Coldplay - Adventure Of A Lifetime (Official Video) |
| Author | Coldplay Official |
| Category | Music |
| Upload date | 2015-11-29T08:00:45.000Z |
| #views | 513645273 |
| #shares | 2655318 |
| #tweets | |
| Endogenous response | 0 |
| Exogenous sensitivity | 36.32 |

Showing 1 to 10 of 10 entries

**New video released**

# Supp: when HIP fails the fitting (2)

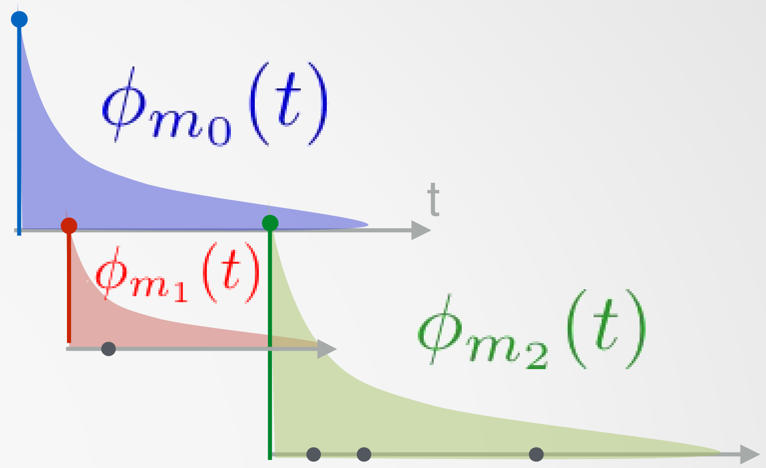
**Long term evolutions:**




**Slow drift**

# Supp: Hawkes Process [Hawkes '71]

$$\lambda(t) = \mu(t) + \sum_{t_i < t} \phi_{m_i}(t - t_i)$$




**Most state-of-the-art popularity prediction systems require observing individual events.**

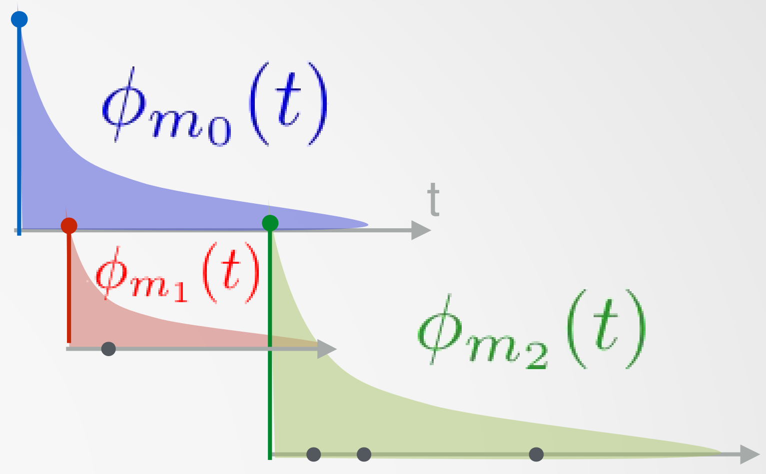
[Zhao et al KDD'15] [Shen et al AAAI'14]

[Farajtabar et al NIPS'15] [Mishra et al CIKM'16]

# Supp: Hawkes Process [Hawkes '71]

$$\lambda(t) = \mu(t) + \sum_{t_i < t} \phi_{m_i}(t - t_i)$$

the rate of 'daughter' events — content virality — user influence — memory



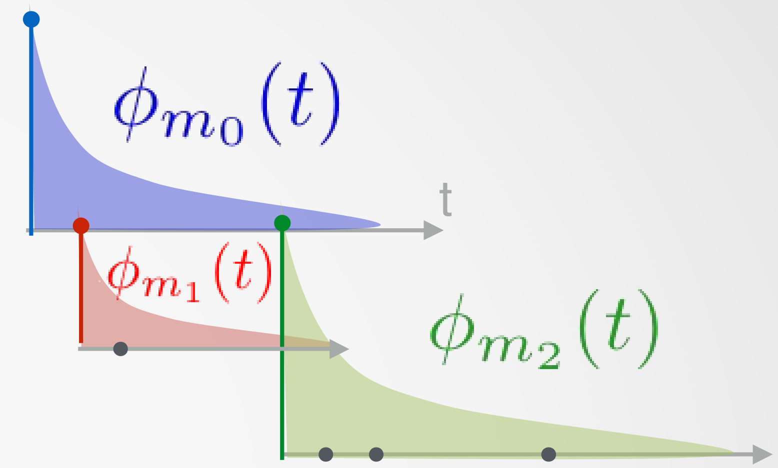

$$\phi_m(\tau) = \kappa\, m^{\beta}\, \hat{\tau}^{-(1+\theta)}$$

**Most state-of-the-art popularity prediction systems require observing individual events.**

[Zhao et al KDD'15] [Shen et al AAAI'14] [Farajtabar et al NIPS'15] [Mishra et al CIKM'16]

# Supp: Hawkes Intensity Process (HIP)

$$\lambda(t) = \mu(t) + \sum_{t_i < t} \phi_{m_i}(t - t_i)$$




the rate of 'daughter' events

content virality

user influence

memory

$$\phi_m(\tau) = \kappa\, m^\beta \hat{\tau}^{-(1+\theta)}$$

expected number of events

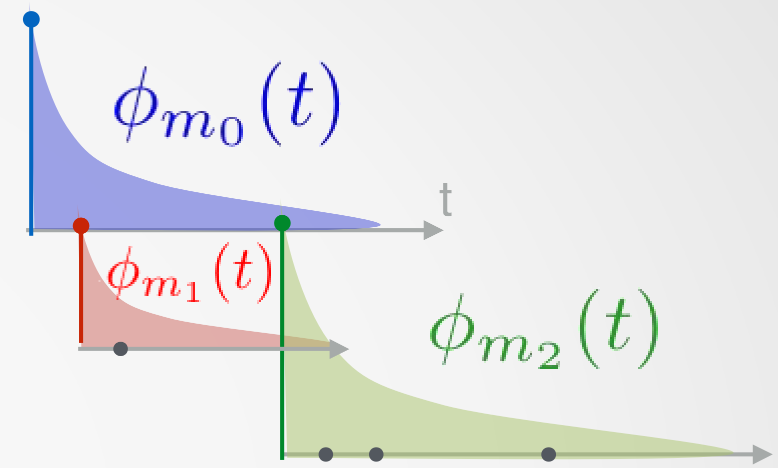
$$\xi(t) = \mu s(t) + C \int_0^t \xi(t-\tau)\hat{\tau}^{-(1+\theta)} d\tau$$

popularity

exogenous stimuli

# Supp: Hawkes Intensity Process (HIP)

$$\lambda(t) = \mu(t) + \sum_{t_i<t} \phi_{m_i}(t - t_i)$$




the rate of 'daughter' events — content virality — user influence — memory

$$\phi_m(\tau) = \kappa\, m^{\beta} \hat{\tau}^{-(1+\theta)}$$

expected number of events

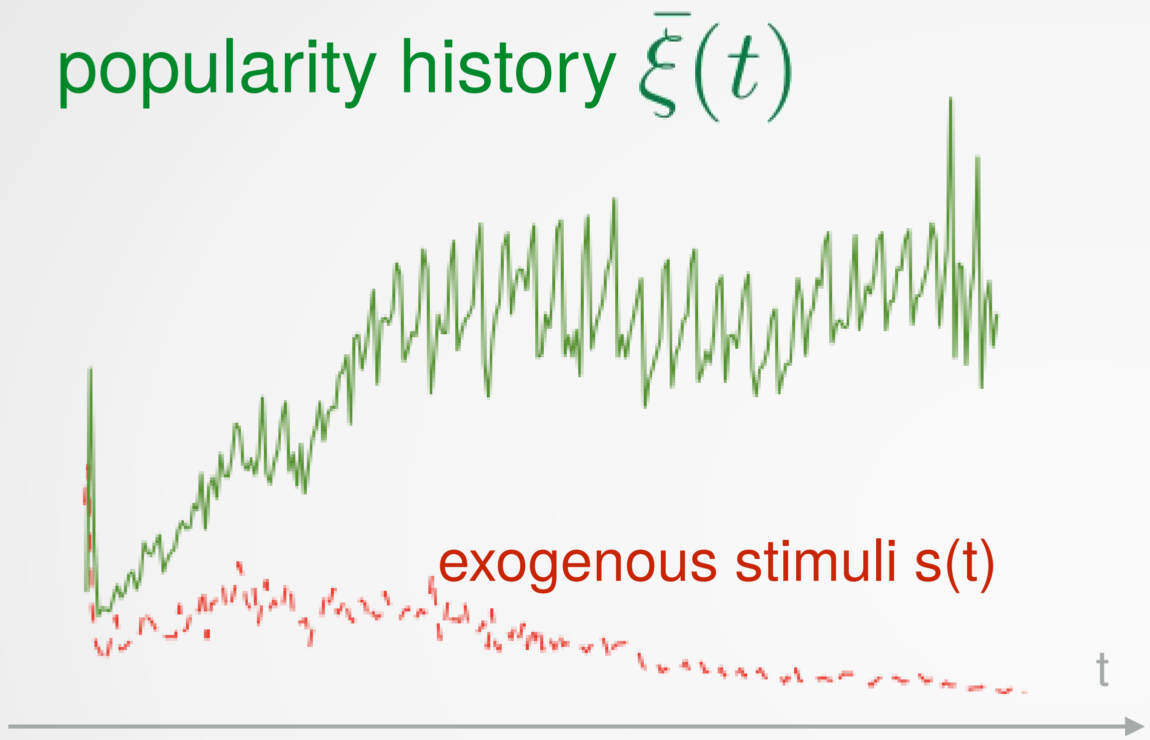
$$\xi(t) = \mu s(t) + \underbrace{C \int_0^t \xi(t-\tau)\hat{\tau}^{-(1+\theta)} d\tau}_{\text{endogenous reaction}}$$

popularity

exogenous sensitivity — exogenous stimuli

endogenous reaction

# Supp: Estimating the HIP model




$$\text{find } \{\mu, C, \theta, \ldots\}$$

$$\text{s.t.} \quad \min \sum_t l(\xi(t) - \bar{\xi}(t))$$

$$\xi(t) = \mu s(t) + \underbrace{C \int_0^t \xi(t-\tau)\hat{\tau}^{-(1+\theta)} d\tau}_{\text{endogenous reaction}}$$

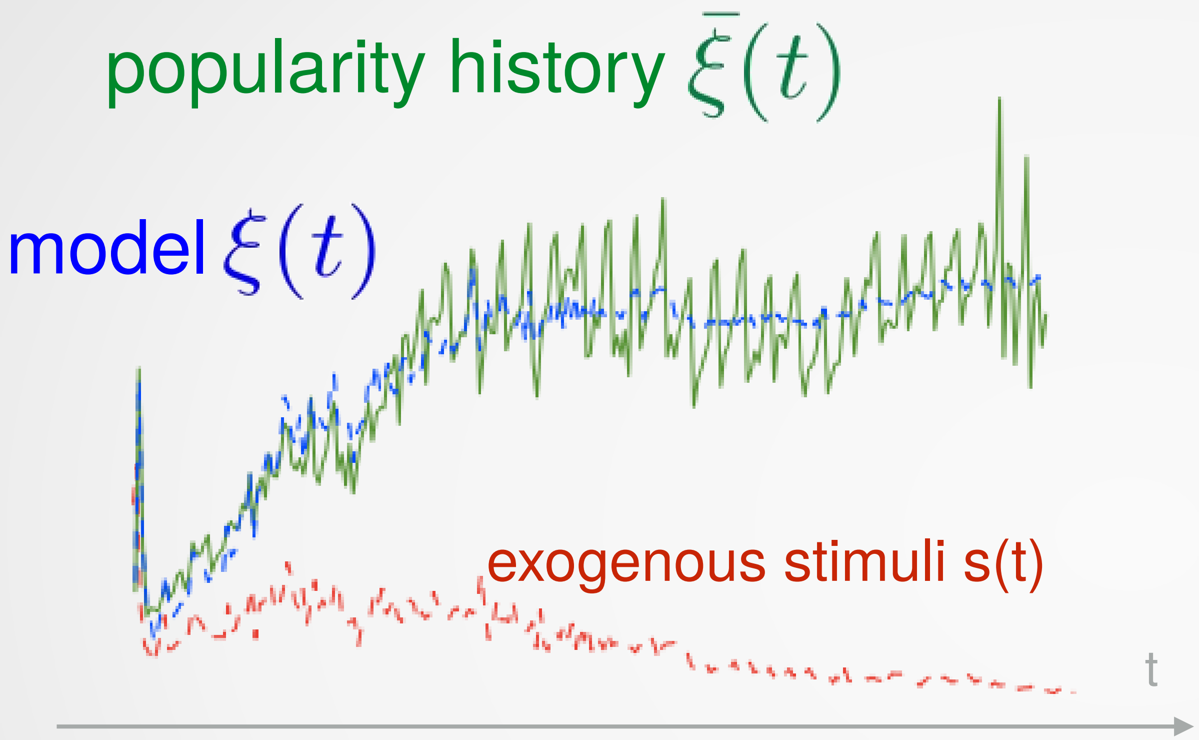
popularity

exogenous sensitivity

exogenous stimuli

endogenous reaction

# Supp: Estimating the HIP model




$$\text{find } \{\mu, C, \theta, \ldots\}$$

$$\text{s.t.} \quad \min \sum_t l(\xi(t) - \bar{\xi}(t))$$

$$\xi(t) = \mu s(t) + \underbrace{C \int_0^t \xi(t-\tau)\hat{\tau}^{-(1+\theta)} d\tau}_{\text{endogenous reaction}}$$

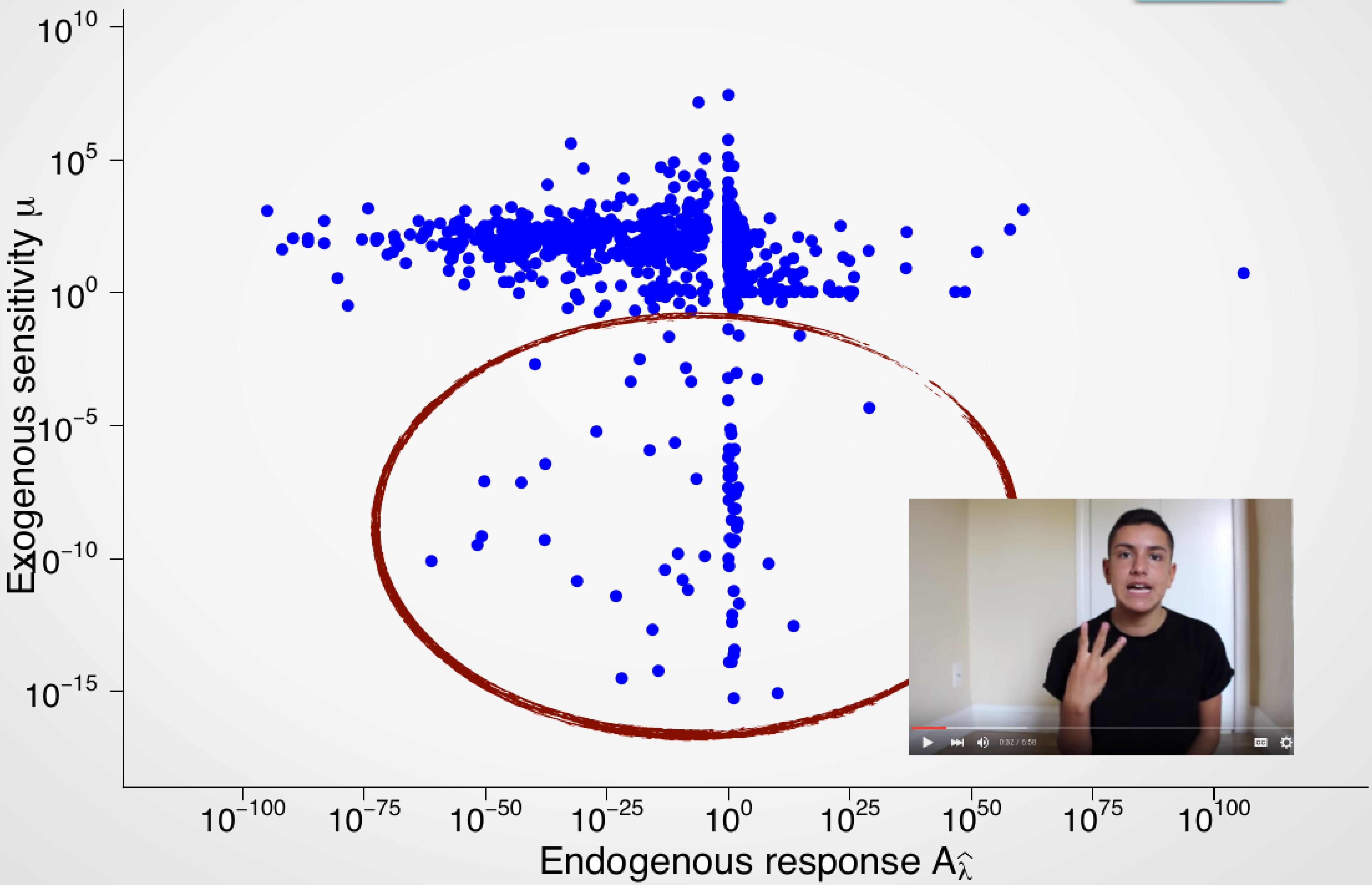
popularity

exogenous sensitivity

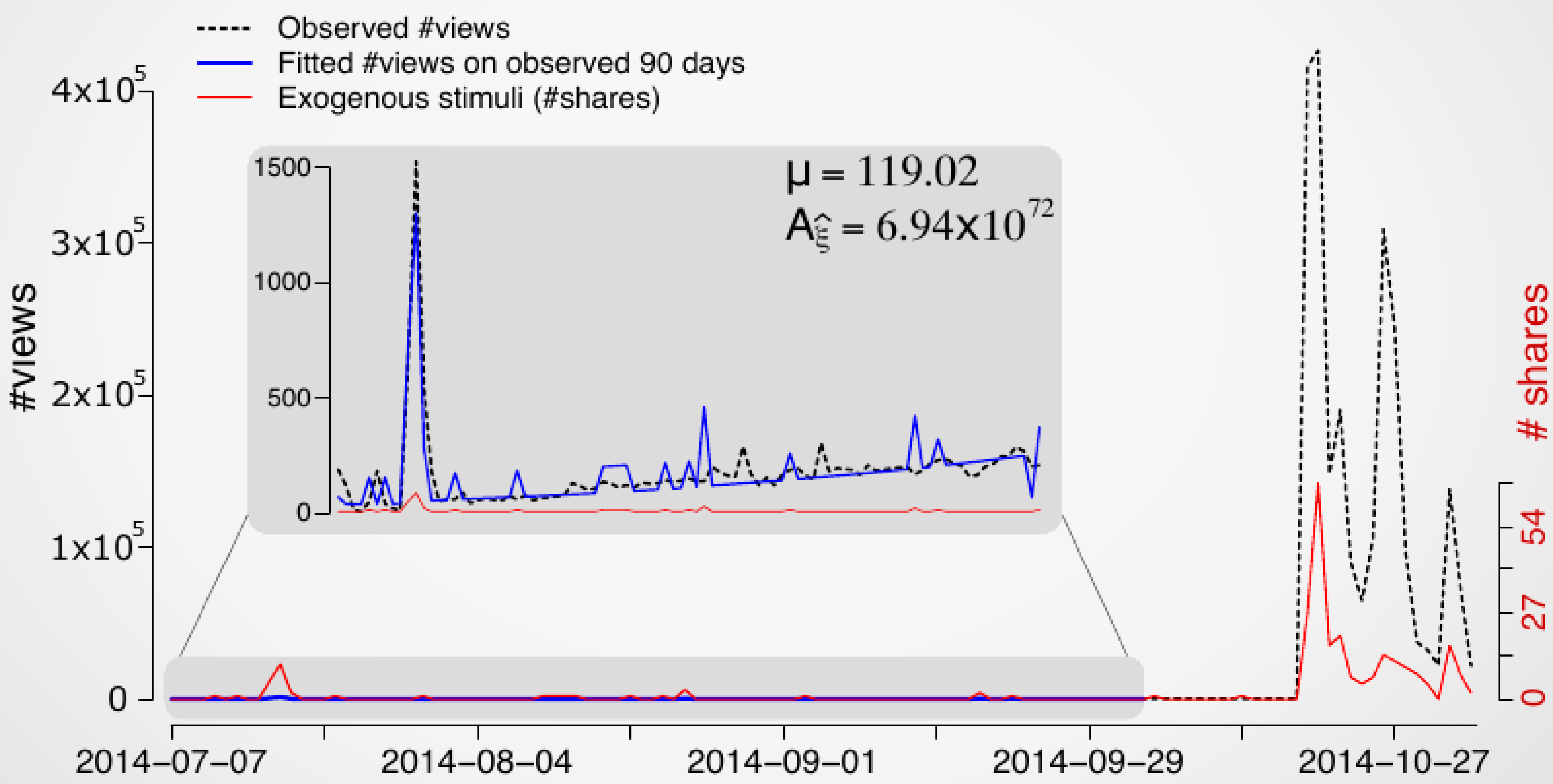
exogenous stimuli

endogenous reaction

# Supp: Un-promotable videos

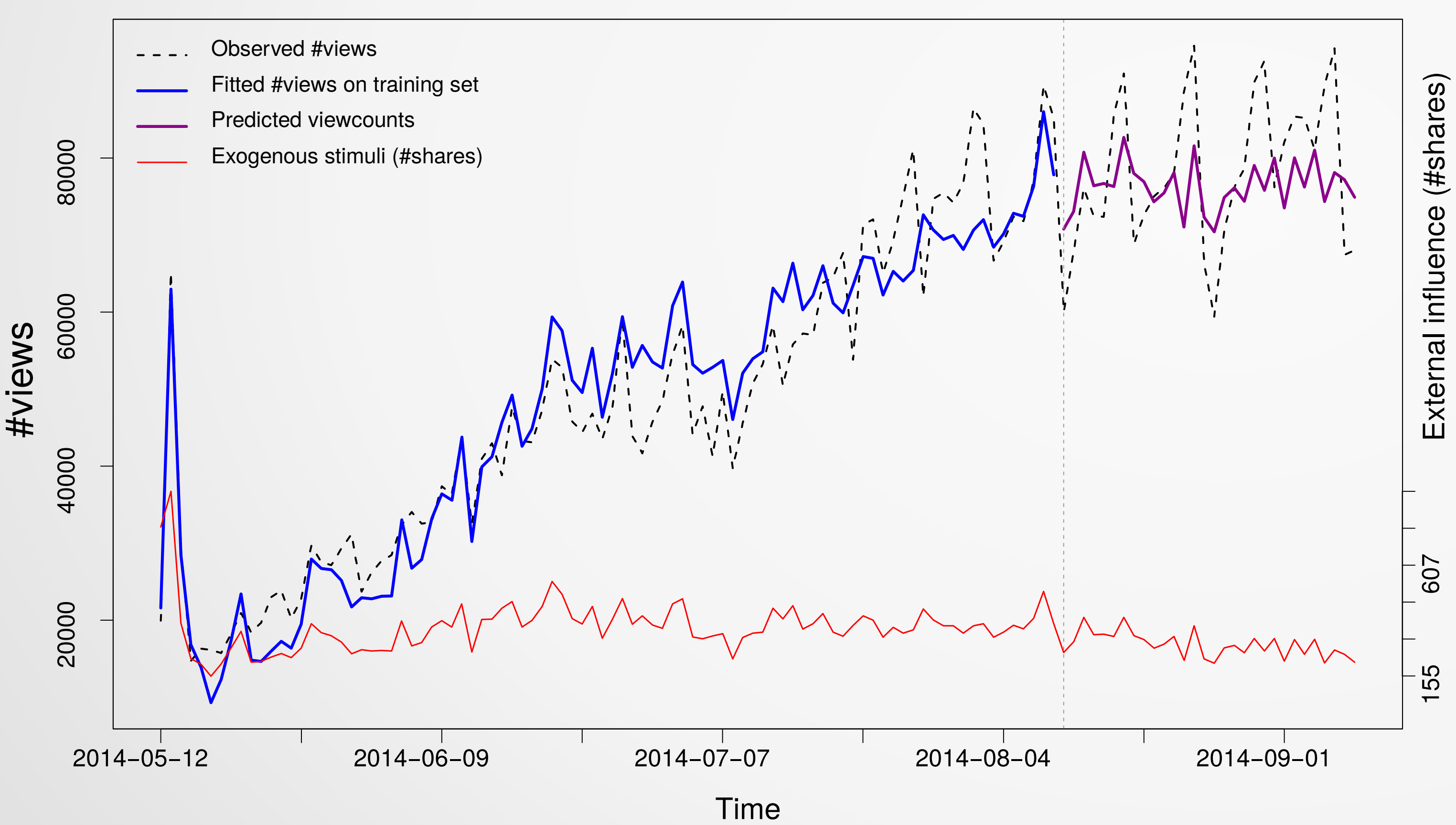


# Supp: “Potentially viral” video

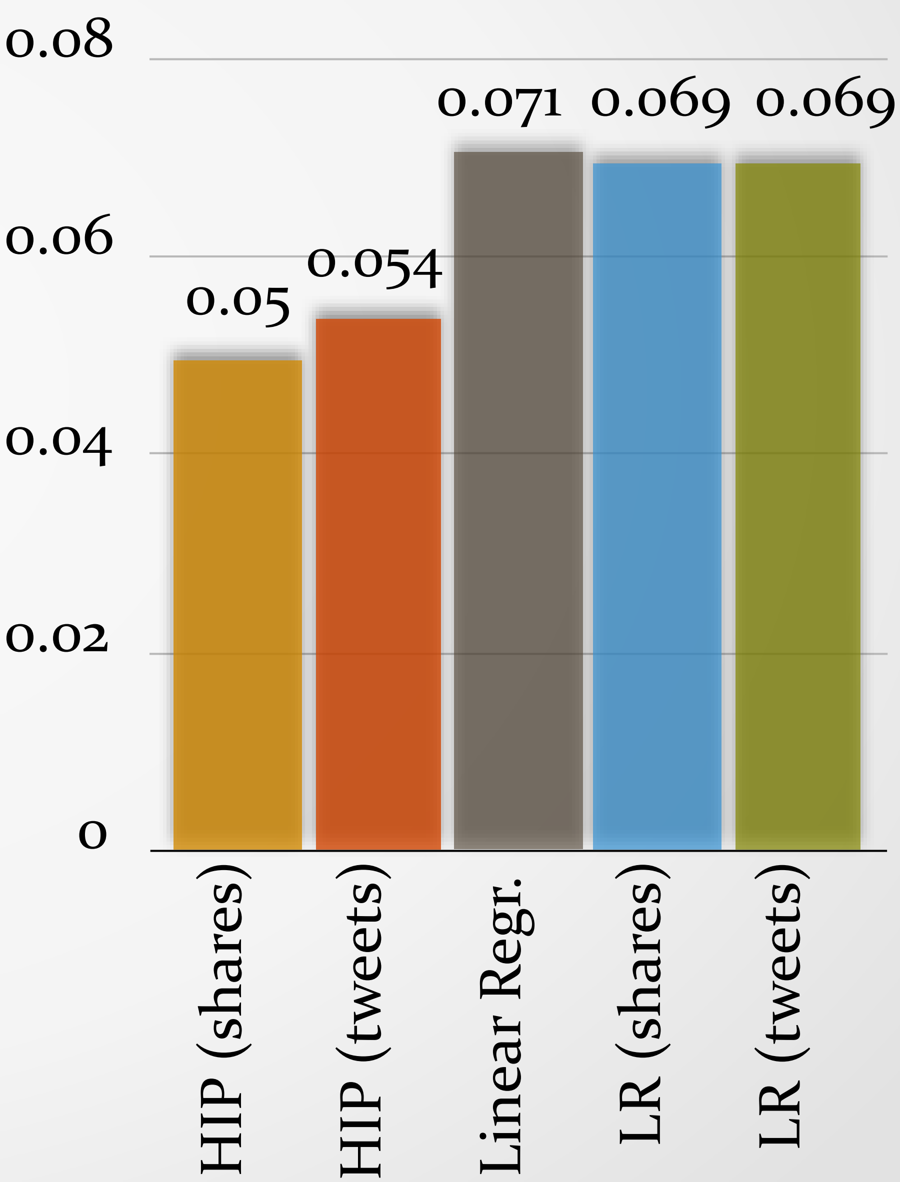


# Forecasting the effect of promotions

Observed and predicted popularity with confidence interval




average error in popularity percentile

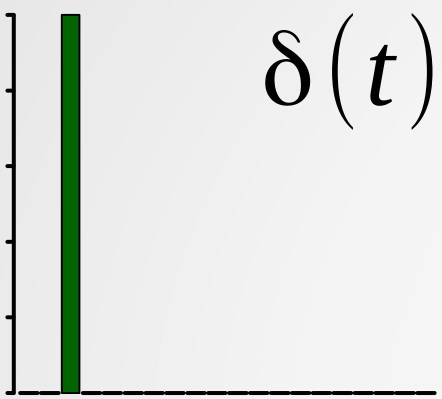



[Pinto et al WSDM'13]

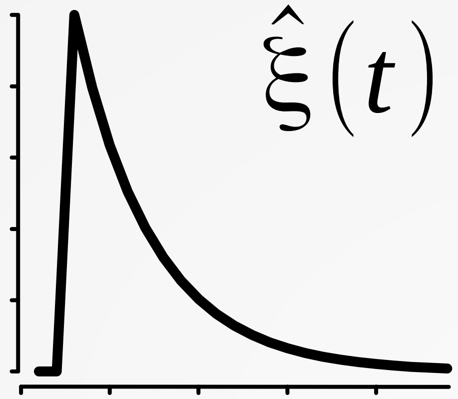
[Szabo & Huberman Comm. ACM'13] [Yu et al ICWSM'15]

# HIP as a Linear Time-Invariant system

promotion

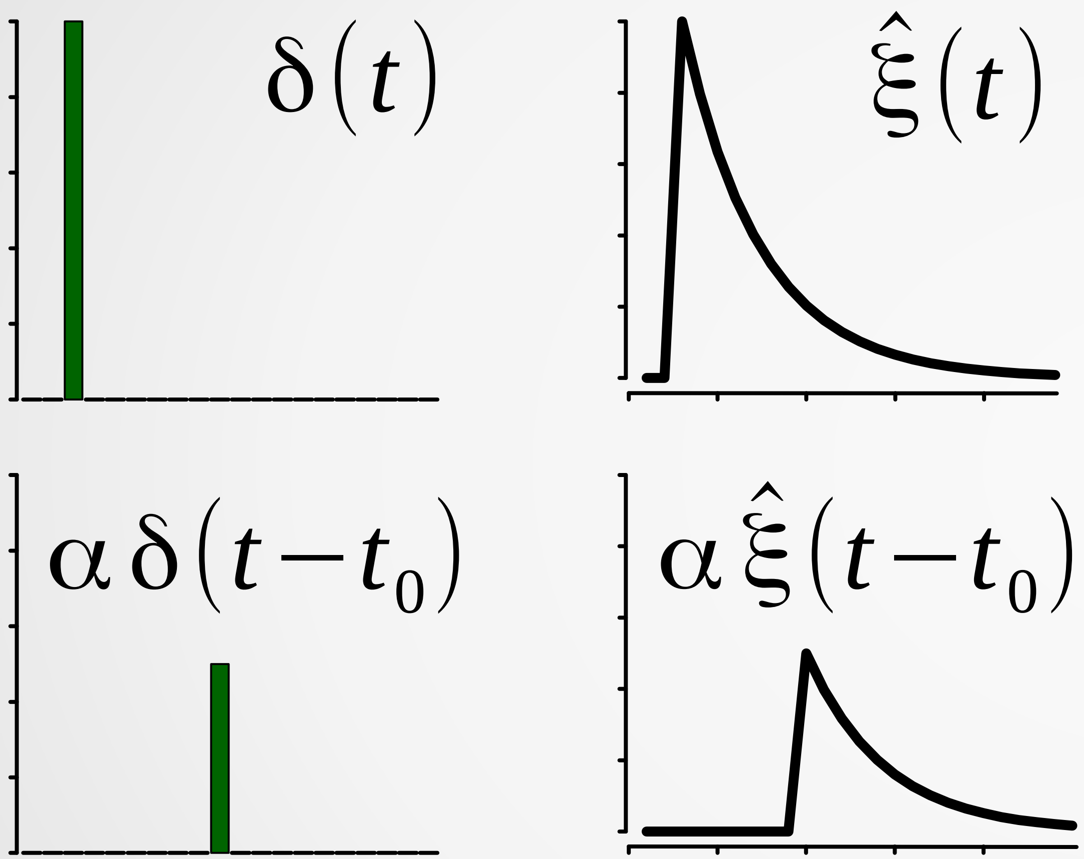



response



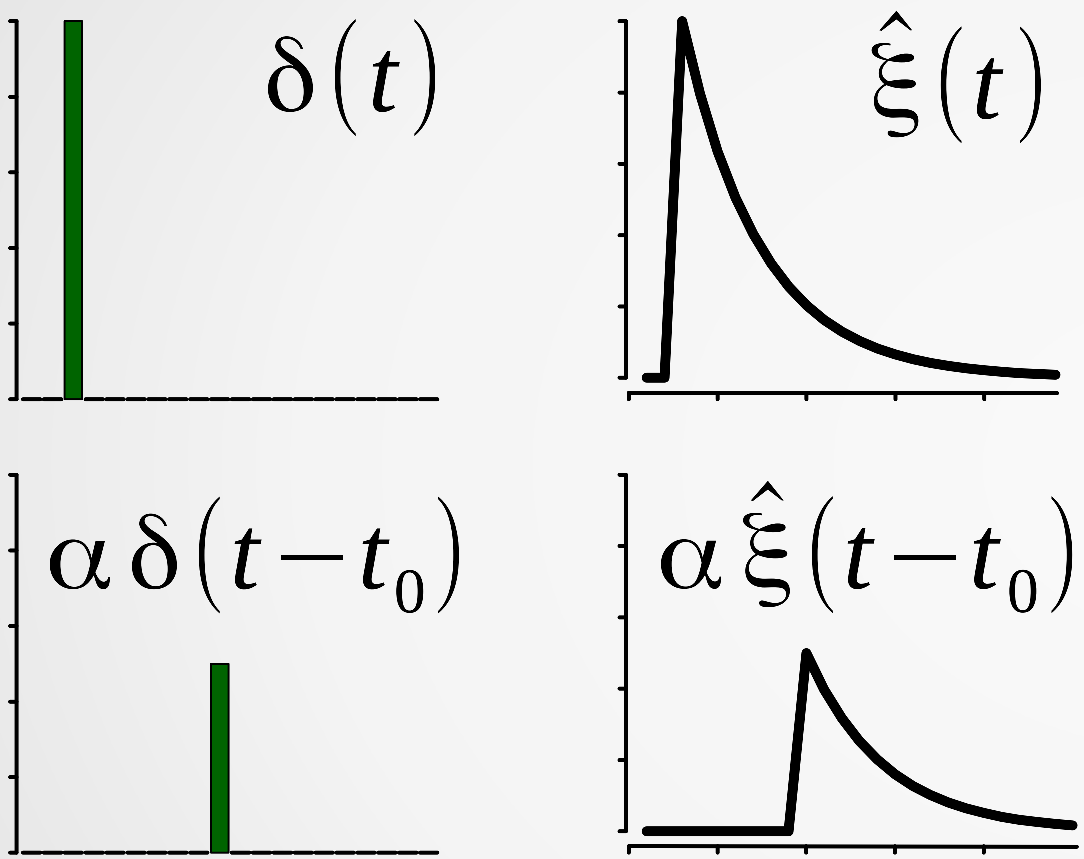
# HIP as a Linear Time-Invariant system

promotion response

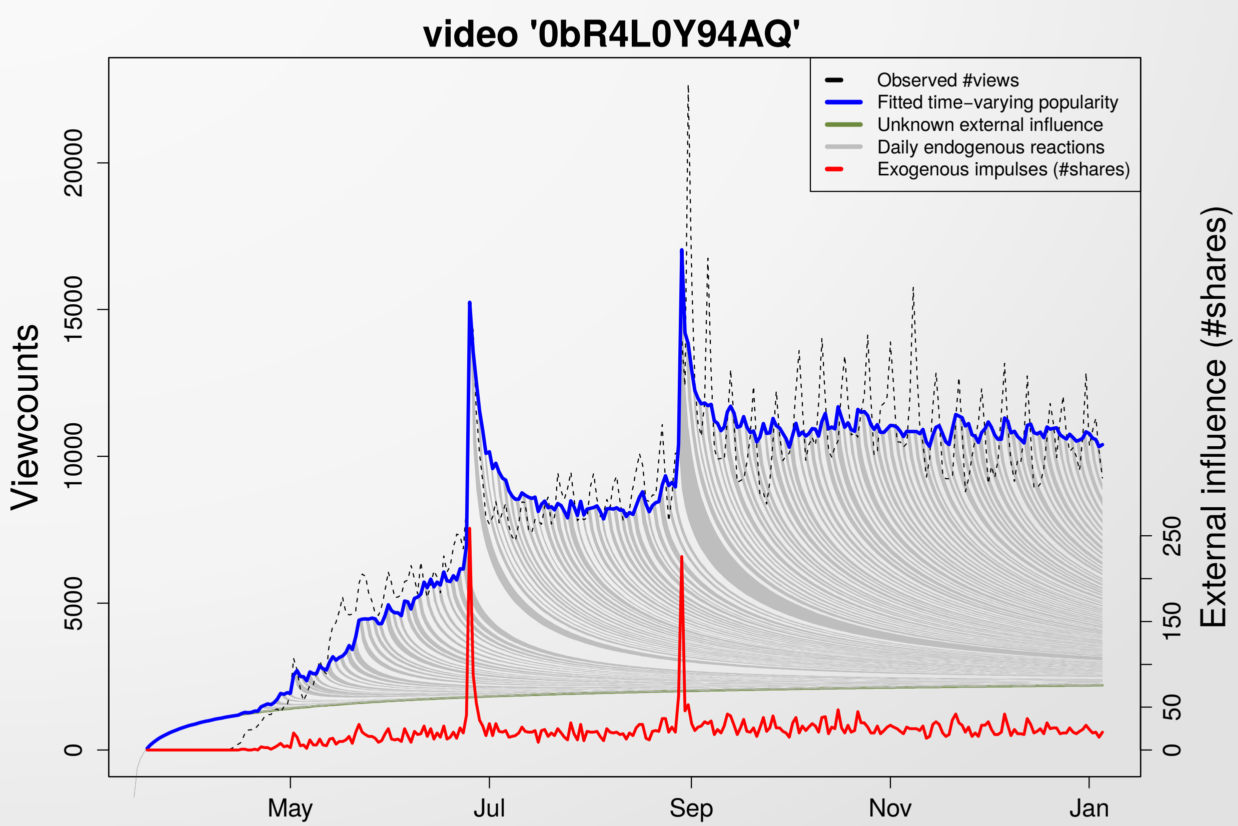


# HIP as a Linear Time-Invariant system

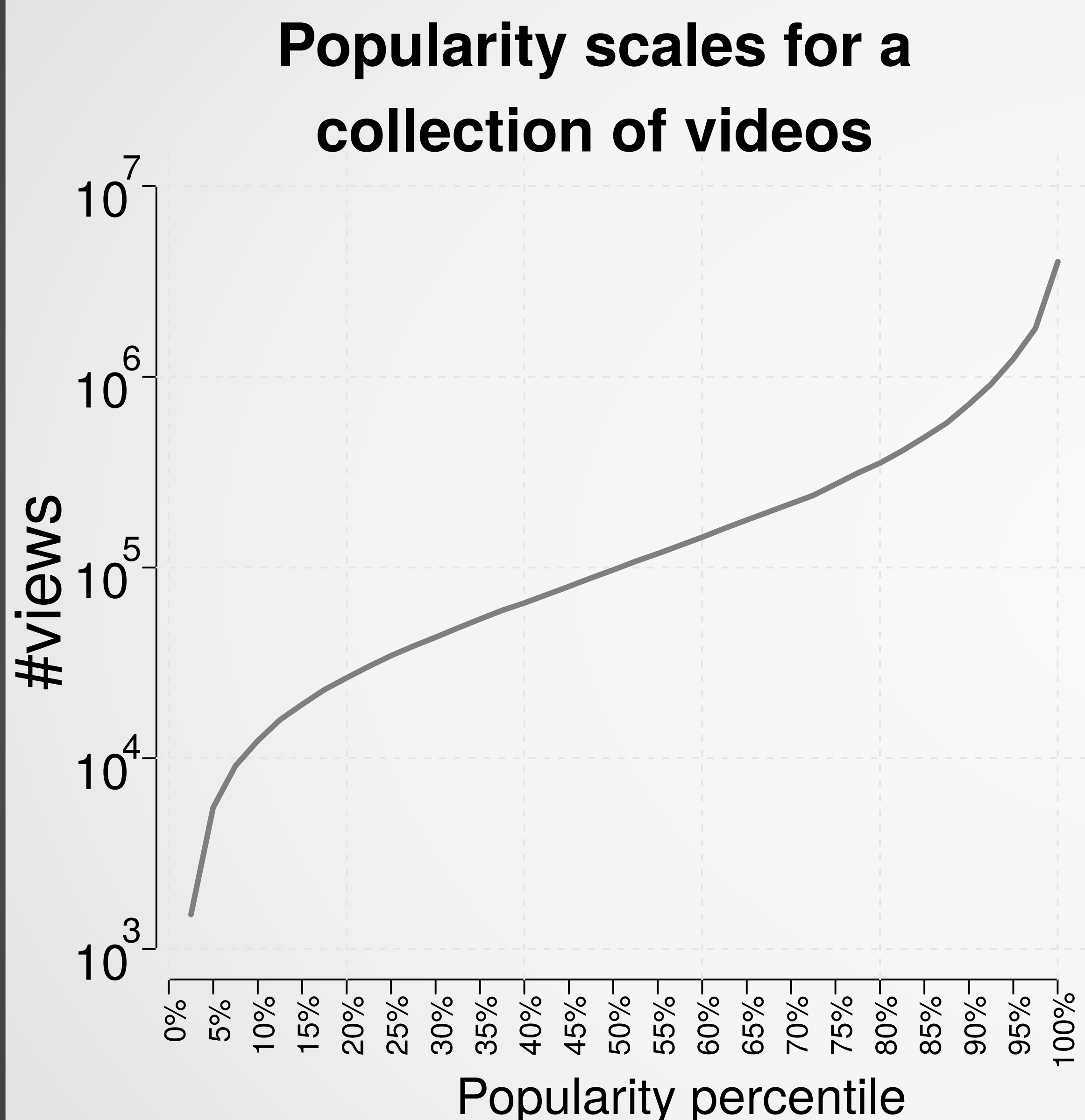
promotion    response



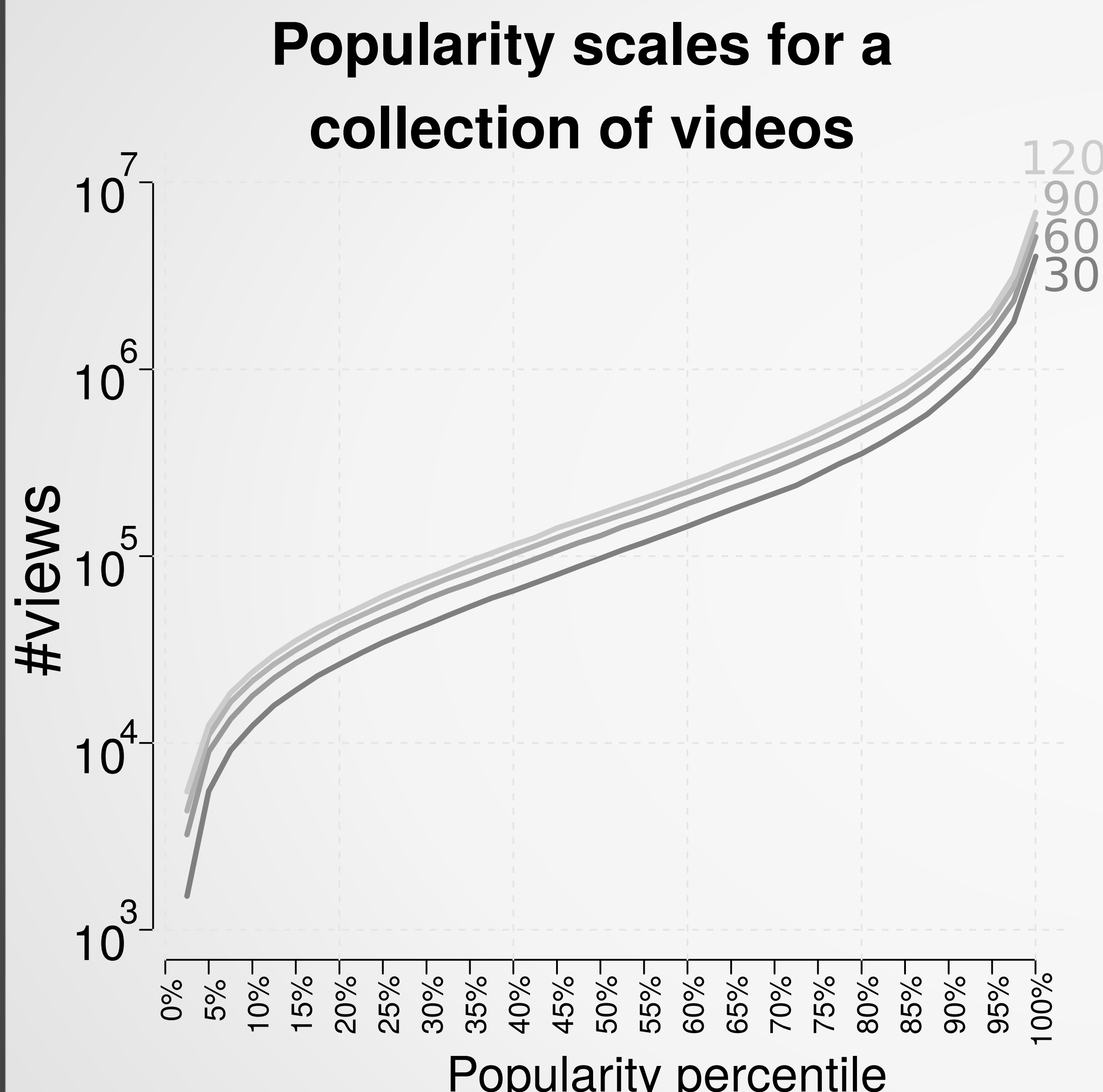

scale,
shift, add



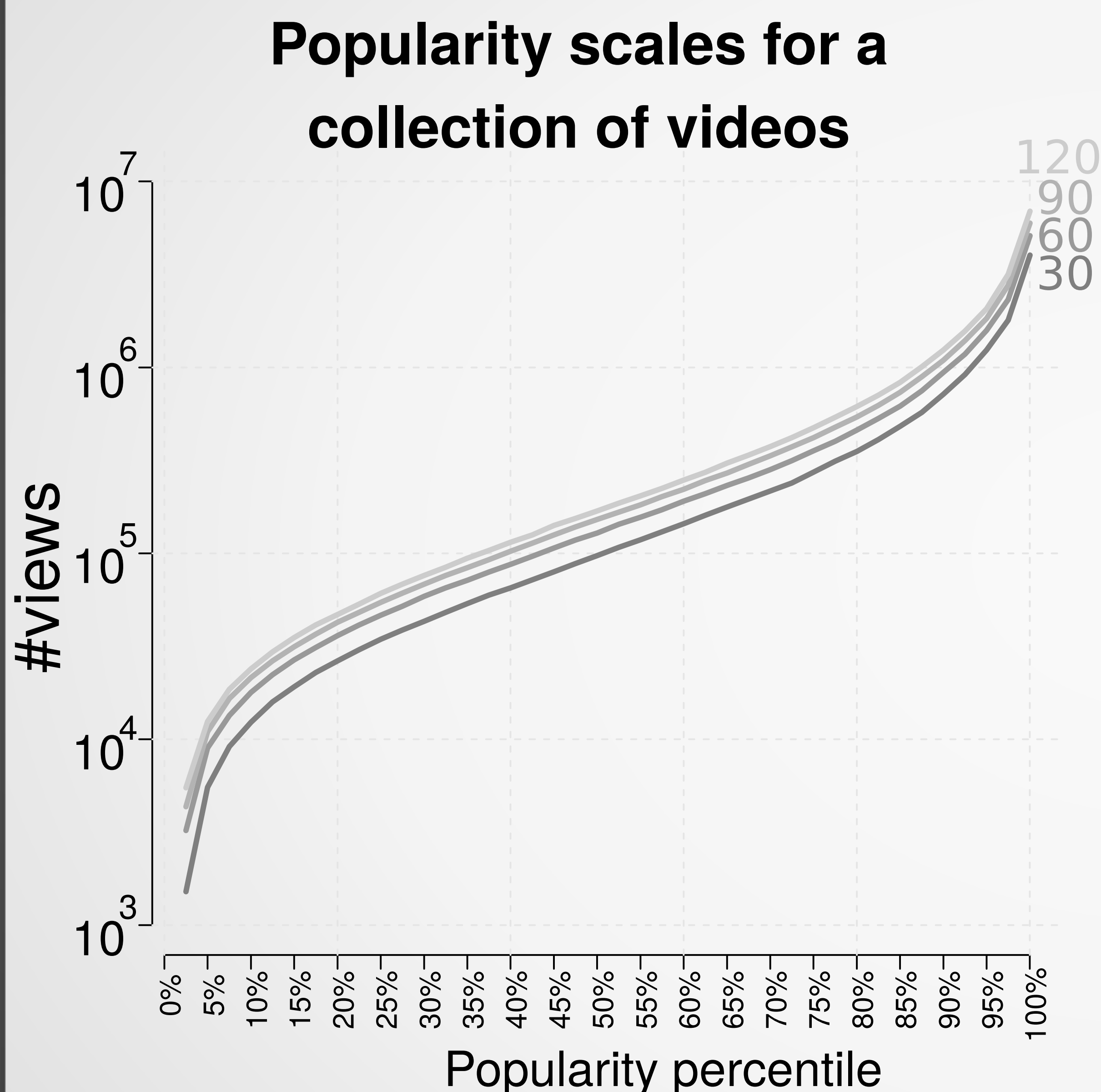
# Popularity scales over time

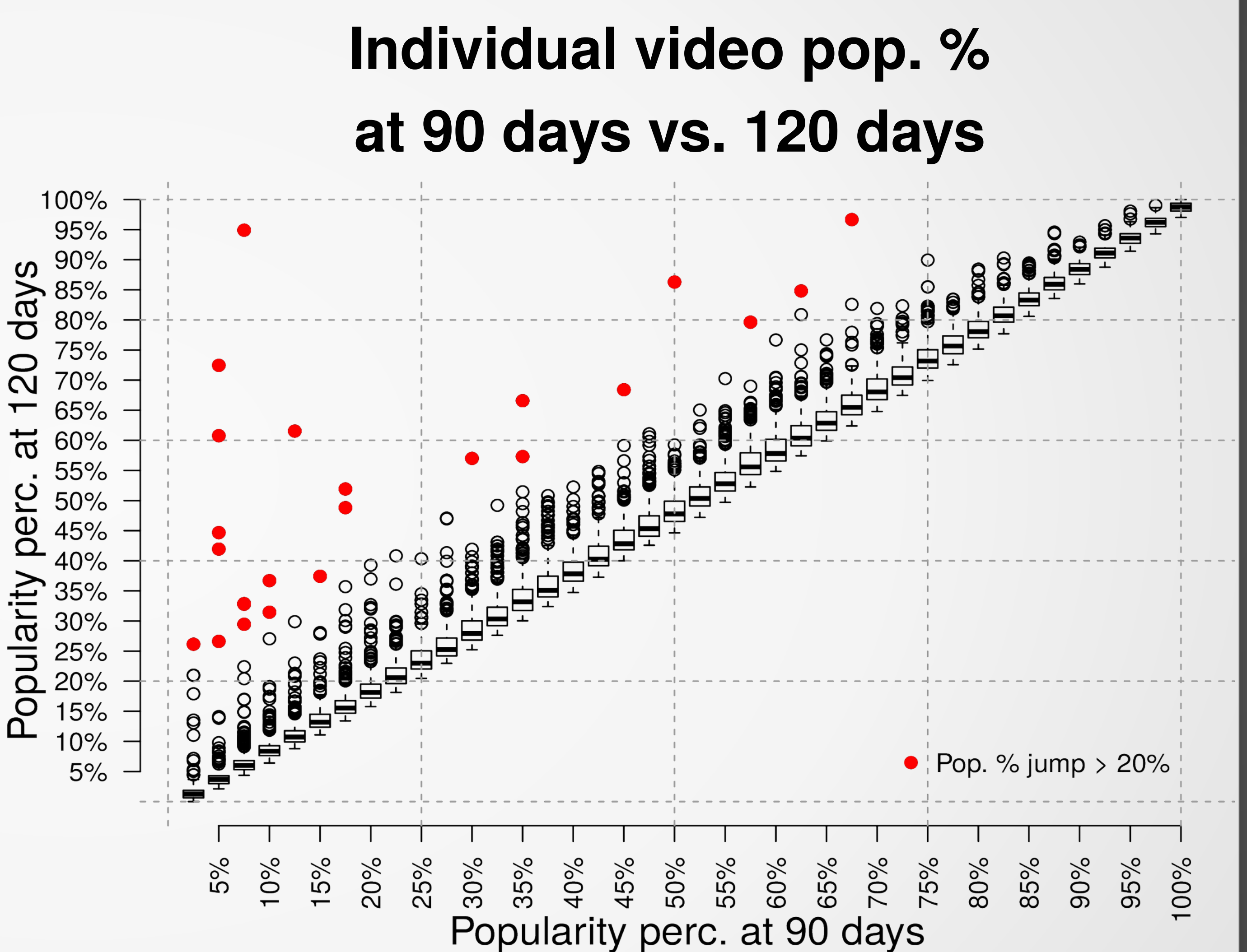


# Popularity scales over time



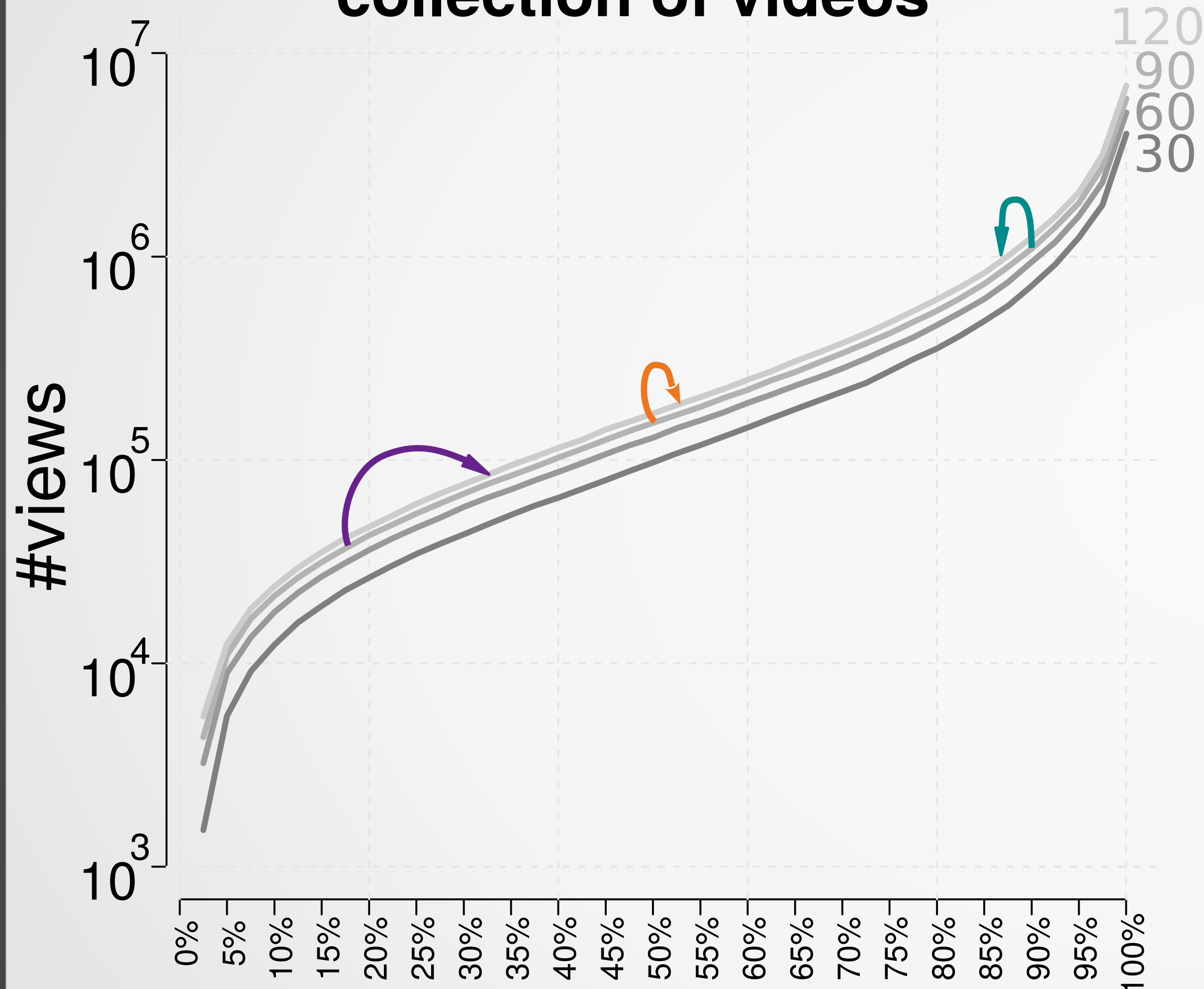
# Popularity scales over time



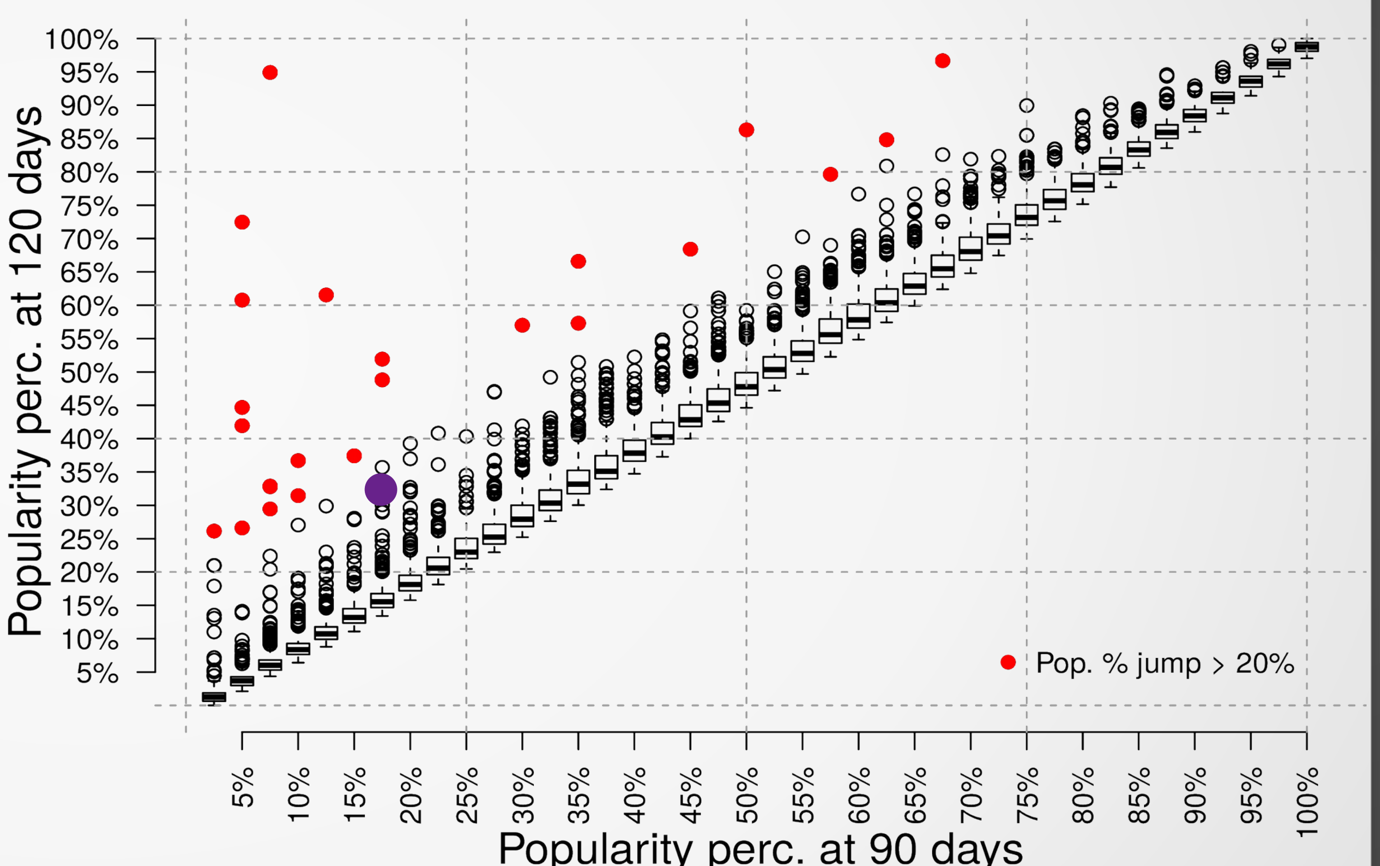



# Popularity scales over time

**Popularity scales for a collection of videos**




**Individual video pop. % at 90 days vs. 120 days**




**Impact of 40k views:**

start at 17.5% → +15%    start at 50% → +2.5%    start at 90% → -2.5%