

# Expecting to be HIP: Hawkes Intensity Processes for Social Media Popularity

Perth, April 6<sup>th</sup>, 2017

Marian-Andrei RizoIU

Lexing Xie

Scott Sanner

Manuel Cebrian

Honglin Yu

Pascal Van Hentenryck

ComputationalMedia @ANU:

<http://cm.cecs.anu.edu.au>

# Popularity over time



My philosophy for a happy life | Sam Berns | TEDxMidAtlantic

TEDx Talks  
**TEDx** [Subscribe](#) 2,346,801 8,190,511

+ Add to [Share](#) [More](#) 👍 75,912 🗨️ 1,287

Video statistics Through May 12, 2015 ?

VIEWS	TIME WATCHED	SUBSCRIPTIONS DRIVEN	SHARES
8,190,550	85 years	18,065	28,720

Cumulative **Daily** ?



J.S. Bach - Brandenburg Concerto No.5 in D BWV1050 - Croatian Baroque Ensemble

Croatian Baroque Ensemble  
**Subscribe** 3,860 1,225,253

+ Add to [Share](#) [More](#) 👍 5,275 🗨️ 128

Video statistics Through May 12, 2015 ?

VIEWS	SHARES
1,225,397	3,870

Cumulative **Daily** ?



# Why popularity?

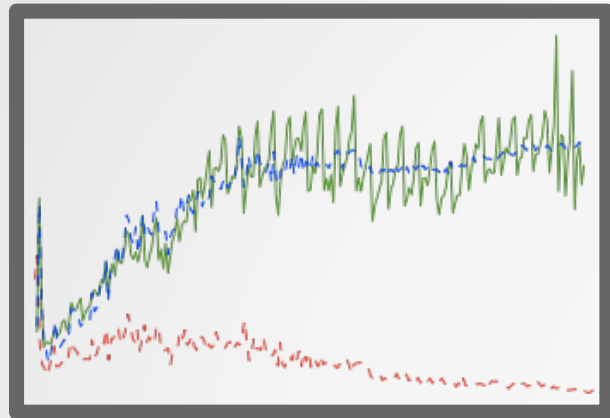
"The fundamental scarcity in the modern world is the scarcity of attention." — Herbert Simon

how do online memes become popular?  
can one predict? can one promote/demote?

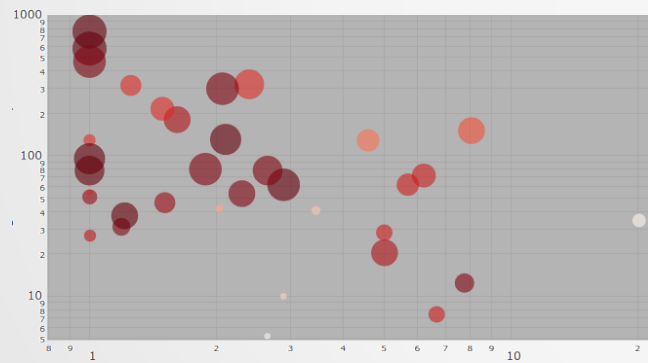
## Applications:

- manage information overload
- information dissemination for public good

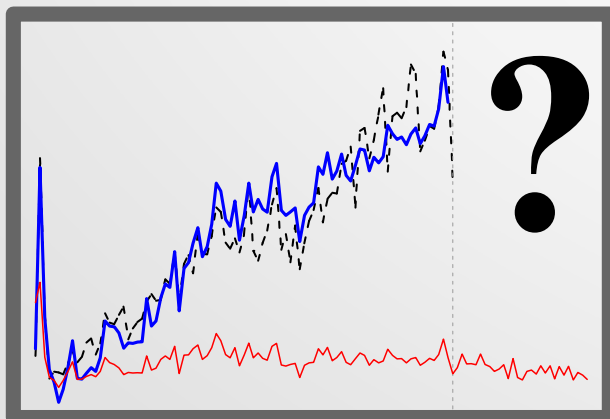
# Presentation outline



Design HIP and estimate it from data



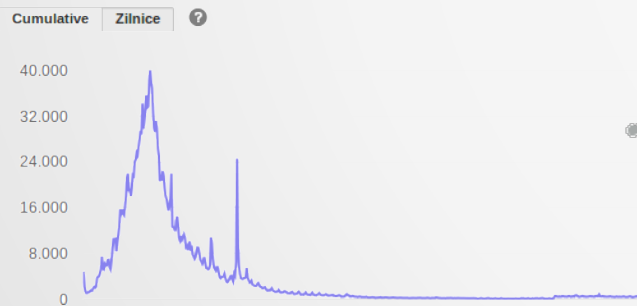
Explain popularity dynamics



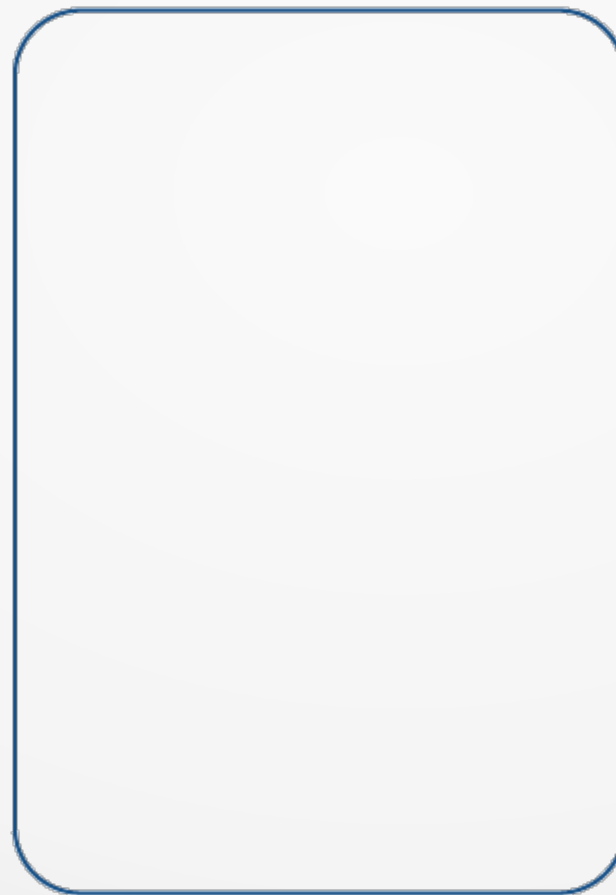
Forecast future popularity

# Linking exo-endo popularity

AFIȘĂRI	PE BAZĂ DE ABONAMENTE	NUMĂR DE DISTRIBUIRI
2.278.811,434	1.223.802	2.432.395



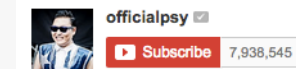
exogenous  
stimuli



endogenous  
response

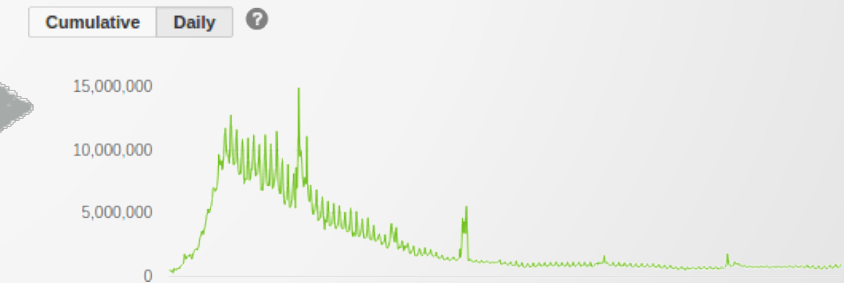


PSY - GANGNAM STYLE (강남스타일) M/V



2,321,368,075

VIEWS	SUBSCRIPTIONS DRIVEN	SHARES
2,278,812,248	1,223,802	2,432,395

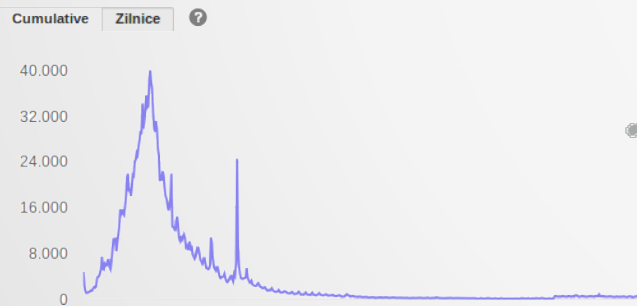


observed  
popularity

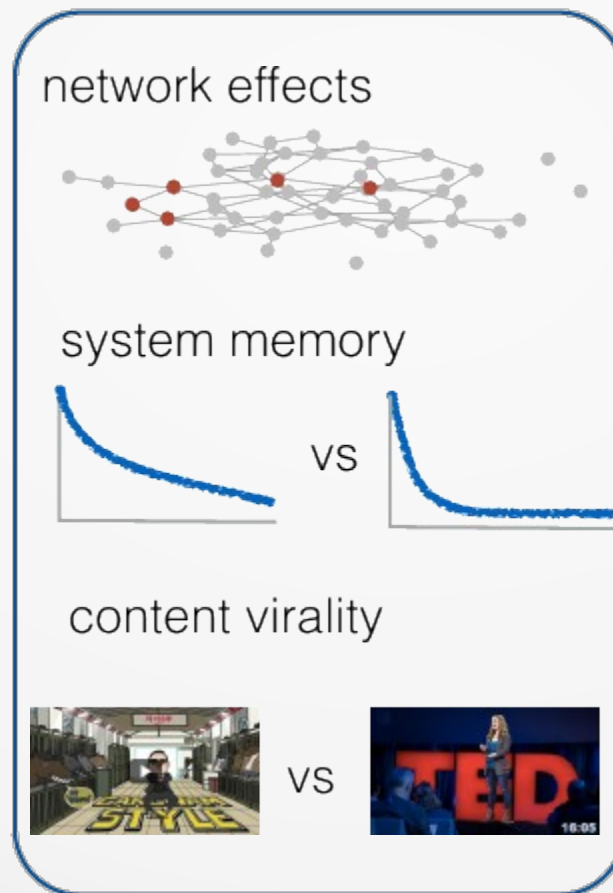


# Linking exo-endo popularity

AFIȘĂRI	PE BAZĂ DE ABONAMENTE	NUMĂR DE DISTRIBUIRI
2.278.811.434	1.223.802	2.432.395



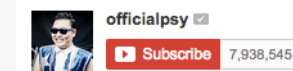
exogenous stimuli



endogenous response

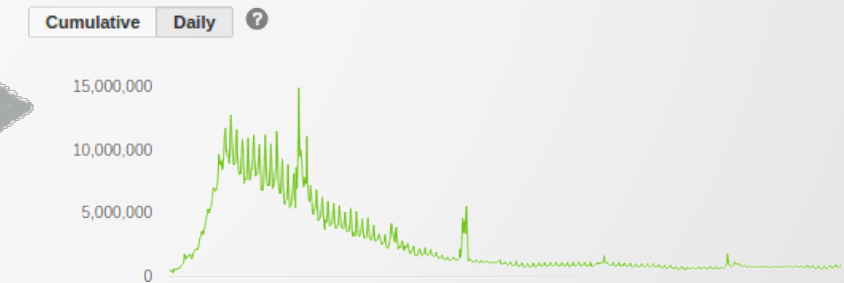


PSY - GANGNAM STYLE (강남스타일) M/V



2,321,368,075

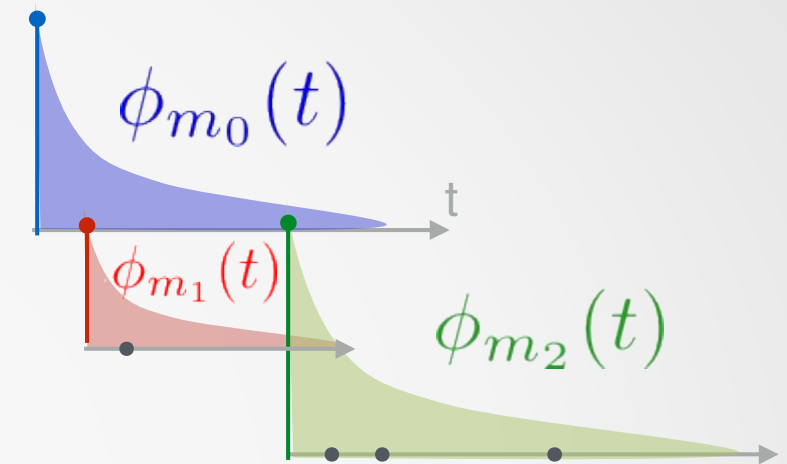
VIEWS	SUBSCRIPTIONS DRIVEN	SHARES
2,278,812,248	1,223,802	2,432,395



observed popularity

# 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.

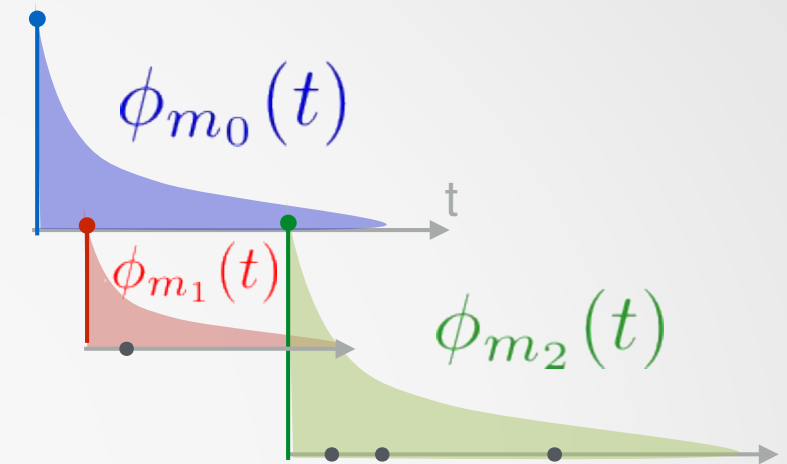
[Zhao et al KDD'15] [Shen et al AAAI'14]

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

# 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]

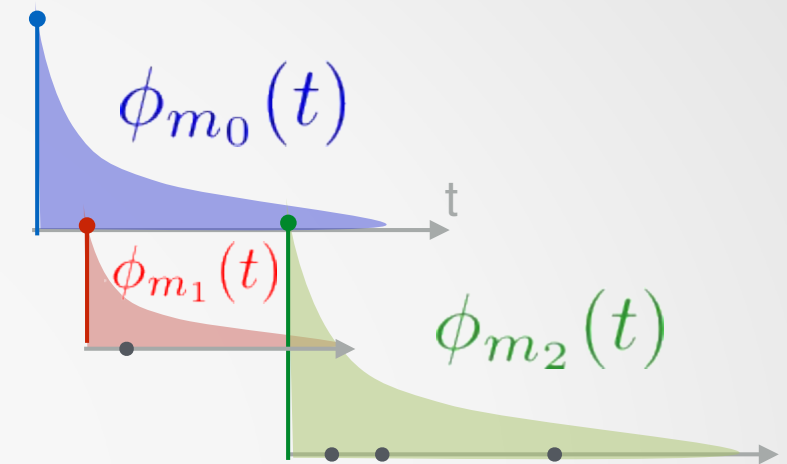
[Farajtabar et al NIPS'15] [Mishra et al CIKM'16]



# 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

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

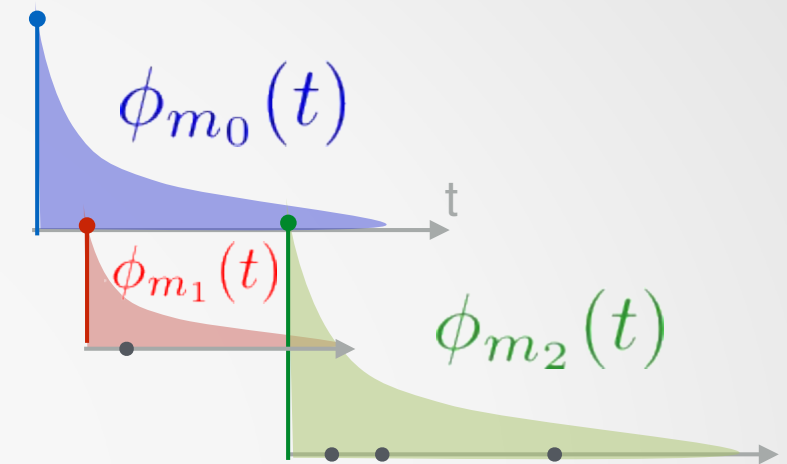
popularity

↓  
exogenous stimuli

# 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

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

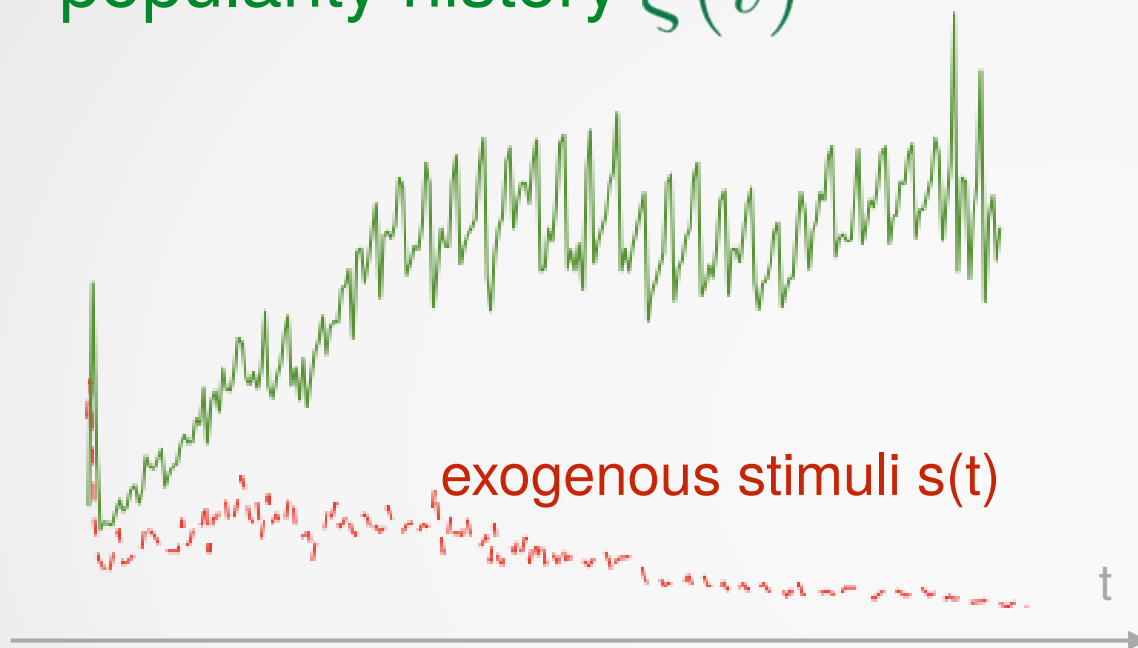
popularity

exogenous sensitivity    exogenous stimuli

endogenous reaction

# Estimating the HIP model

popularity history  $\bar{\xi}(t)$

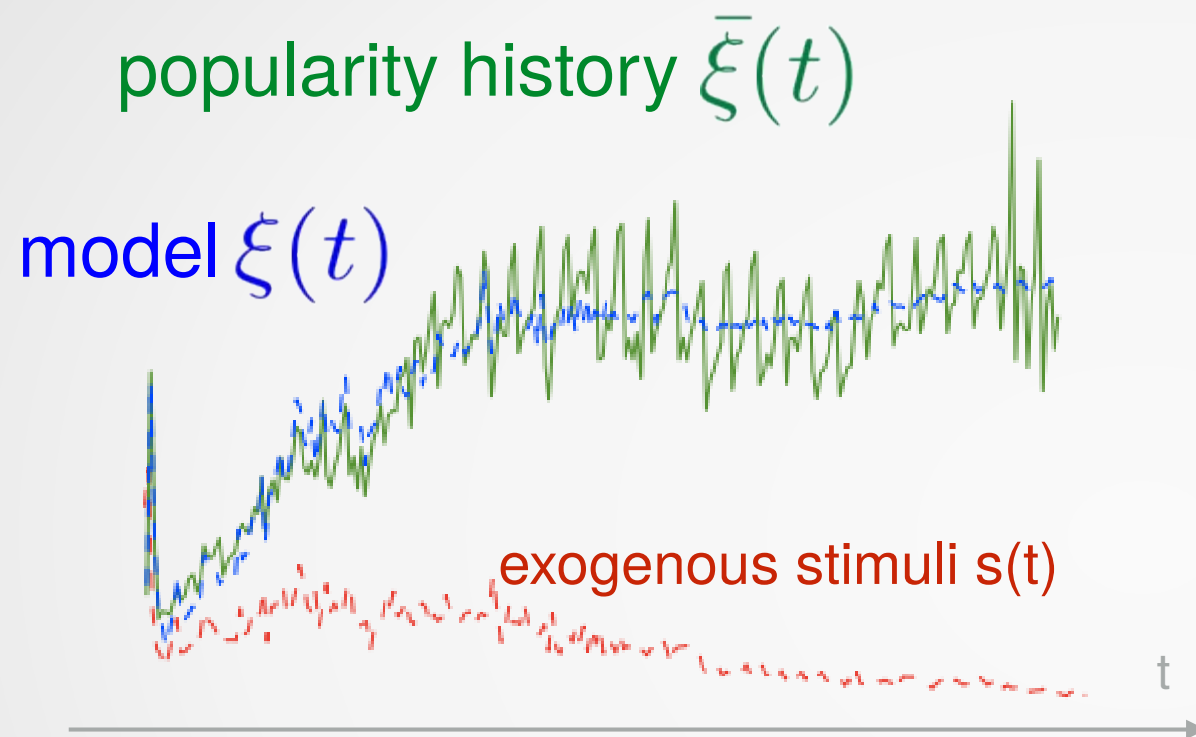


find  $\{\mu, C, \theta, \dots\}$

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

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

# Estimating the HIP model



find  $\{\mu, C, \theta, \dots\}$

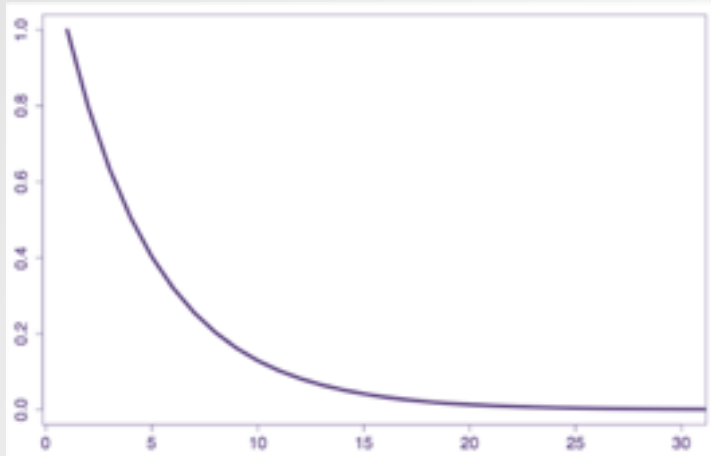
s.t.  $\min \sum_t l(\xi(t) - \bar{\xi}(t))$

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

popularity  $\swarrow$   $\downarrow$   $\underbrace{\hspace{10em}}_{\text{endogenous reaction}}$

exogenous sensitivity exogenous stimuli

# HIP as a Linear Time-Invariant system

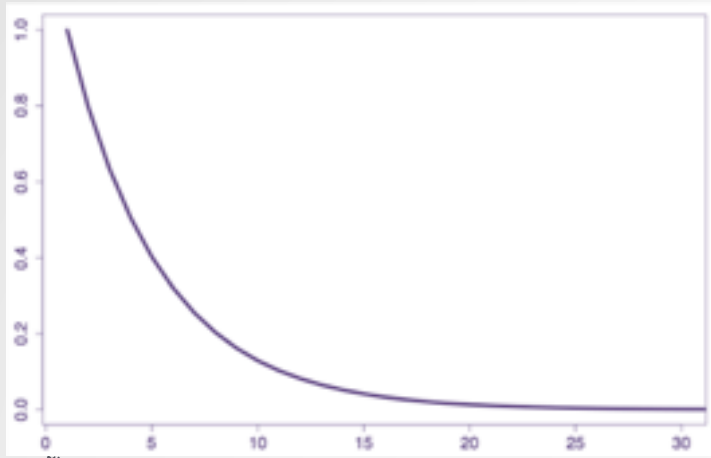


Impulse  
response

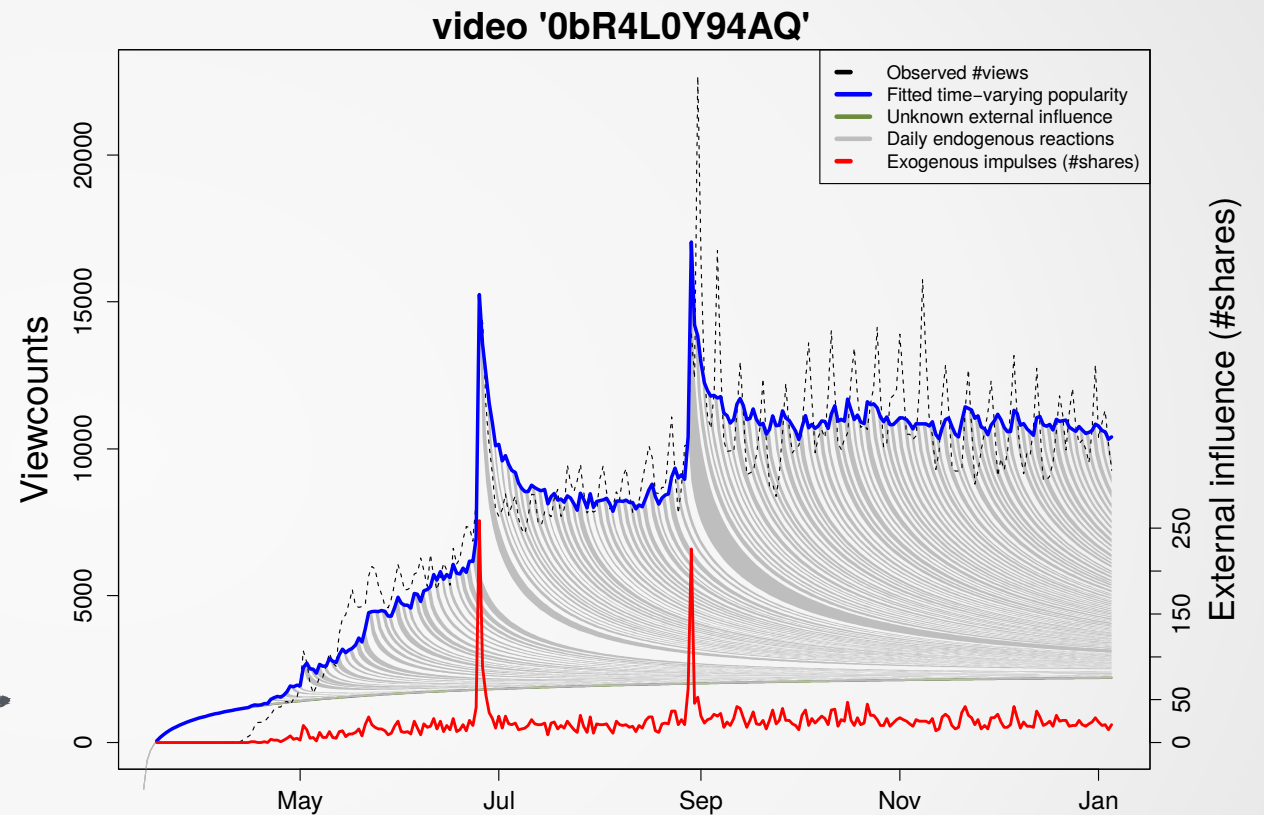
$$\xi(t) = \underbrace{\mu s(t)}_{\substack{\text{popularity} \\ \text{exogenous} \\ \text{sensitivity}}} + \underbrace{C \int_0^t \xi(t - \tau) \hat{\tau}^{-(1+\theta)} d\tau}_{\substack{\text{exogenous} \\ \text{stimuli}} \quad \text{endogenous} \\ \text{reaction}}$$



# HIP as a Linear Time-Invariant system



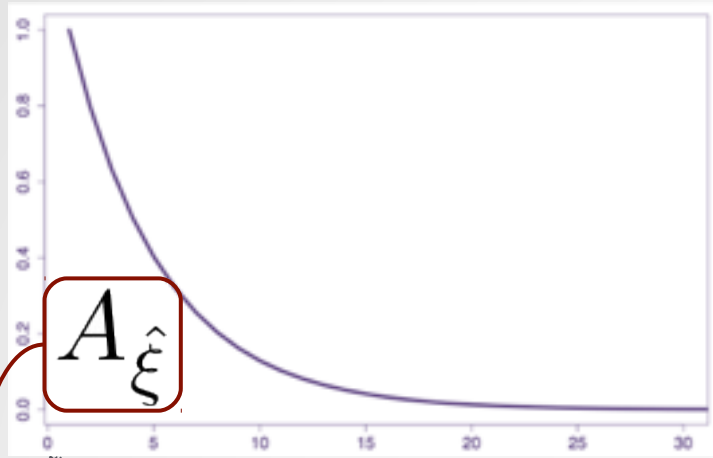
scale, shift,  
add



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

popularity  $\swarrow$   $\downarrow$  endogenous reaction  
 exogenous sensitivity exogenous stimuli

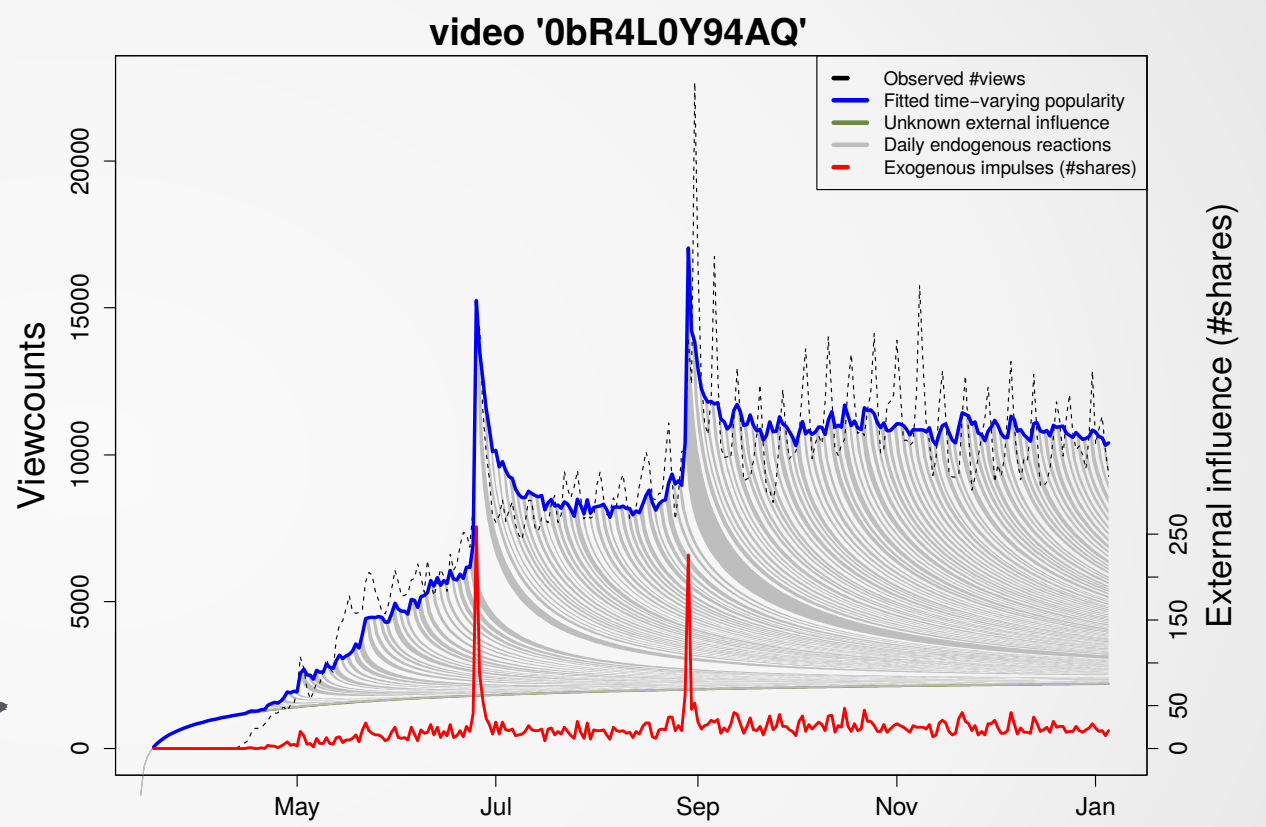
# HIP as a Linear Time-Invariant system



$A_{\hat{\xi}}$

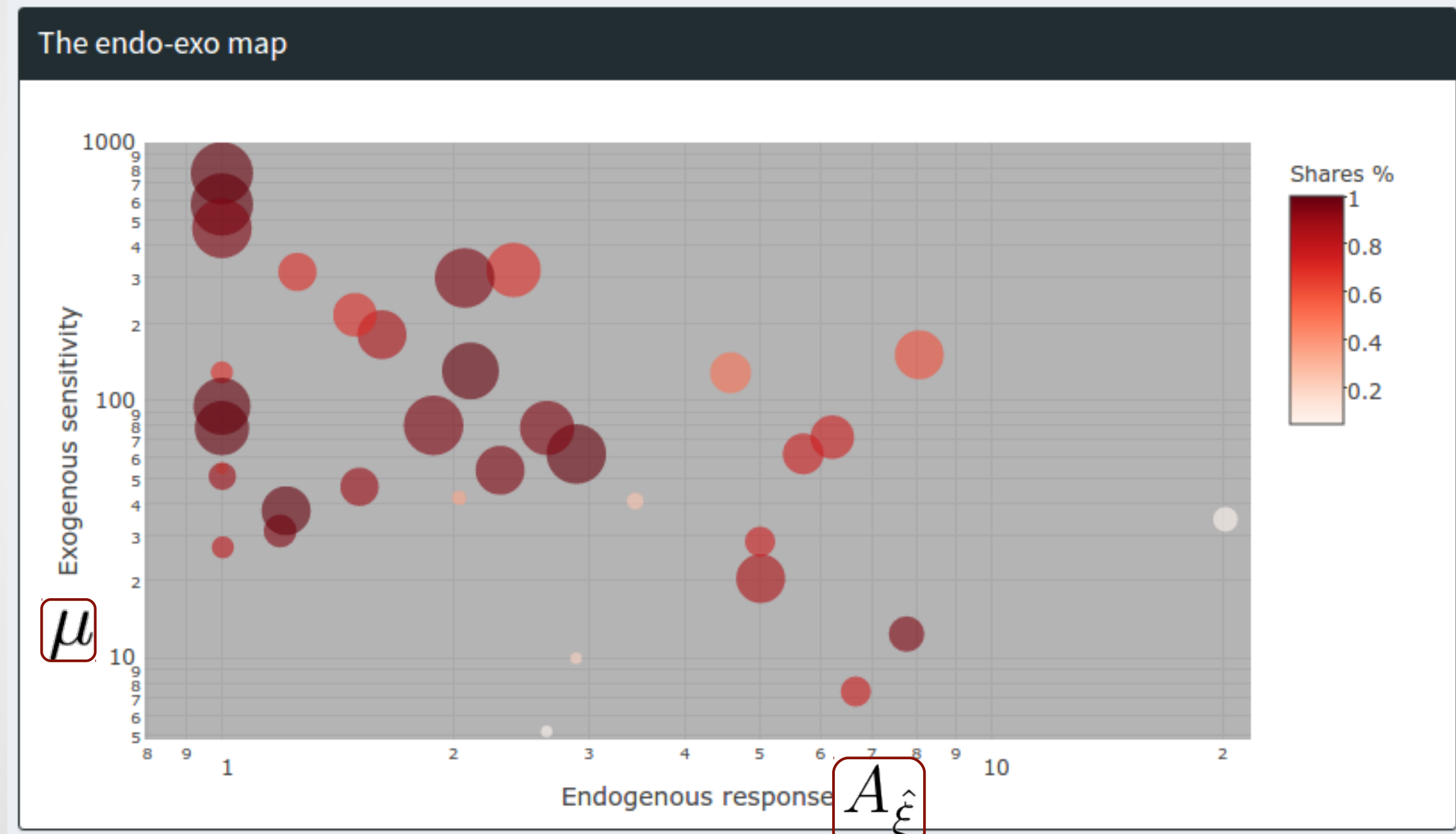
scale, shift,  
add

endogenous  
response



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

# The “endo-exo” map



# Explain popularity dynamics

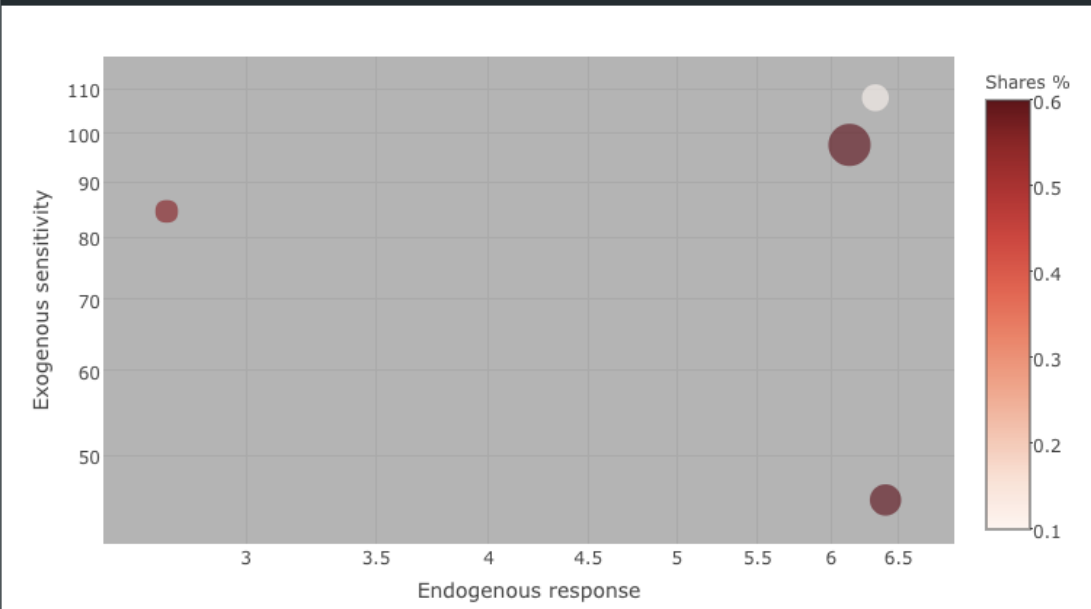
YouTube

Search this dataset in id, title, author, descrip

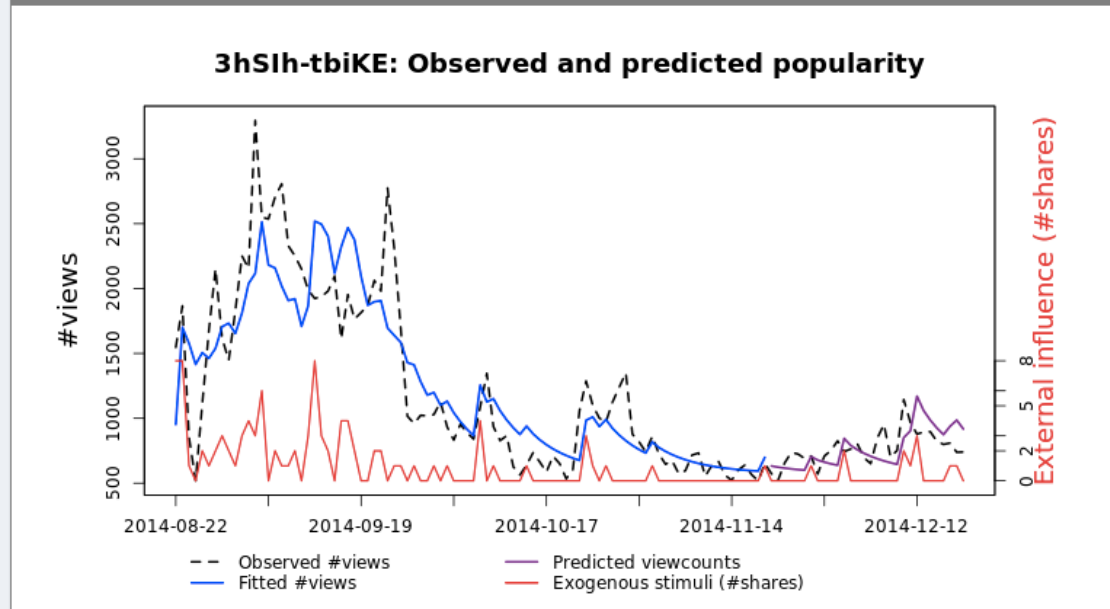
+ Add New Video To This Dataset

Remove Current Video From Dataset

The endo-exo map



Popularity series plot



Video



Information about this video

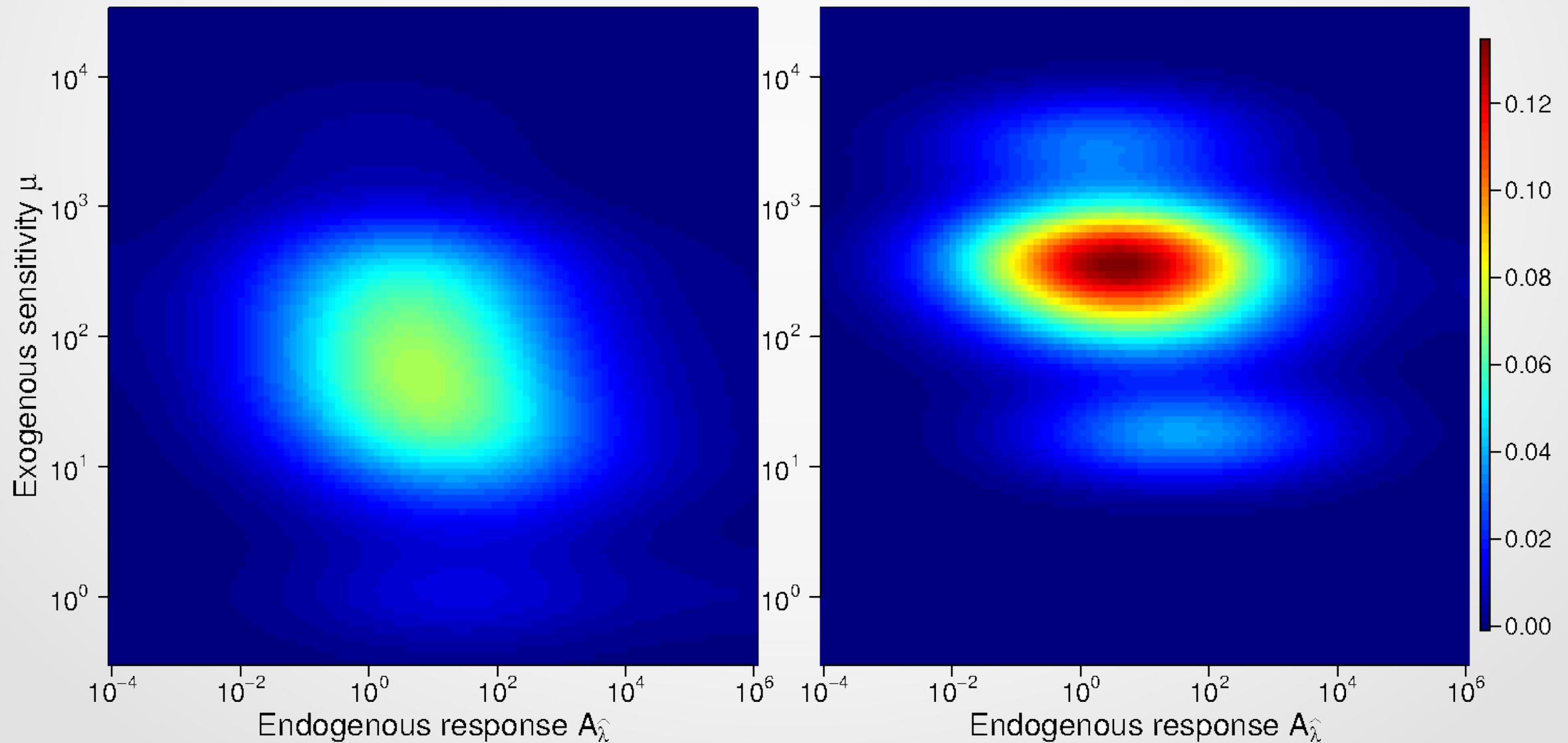
Video property	Property value
YoutubeID	3hSIh-tbiKE
Title	Agents Of S.H.I.E.L.D. - ASL Ice Bucket Challenge
Author	Agents of SHIELD Italia
Category	Film & Animation
Upload date	2014-08-22 02:00:00
#views	157595
#shares	117
#tweets	182
Endogenous response	6.32
Exogenous sensitivity	107.98

Showing 1 to 10 of 10 entries

# Explain popularity – all vs top 5%

Film and Animation:

more popular videos have higher sensitivity

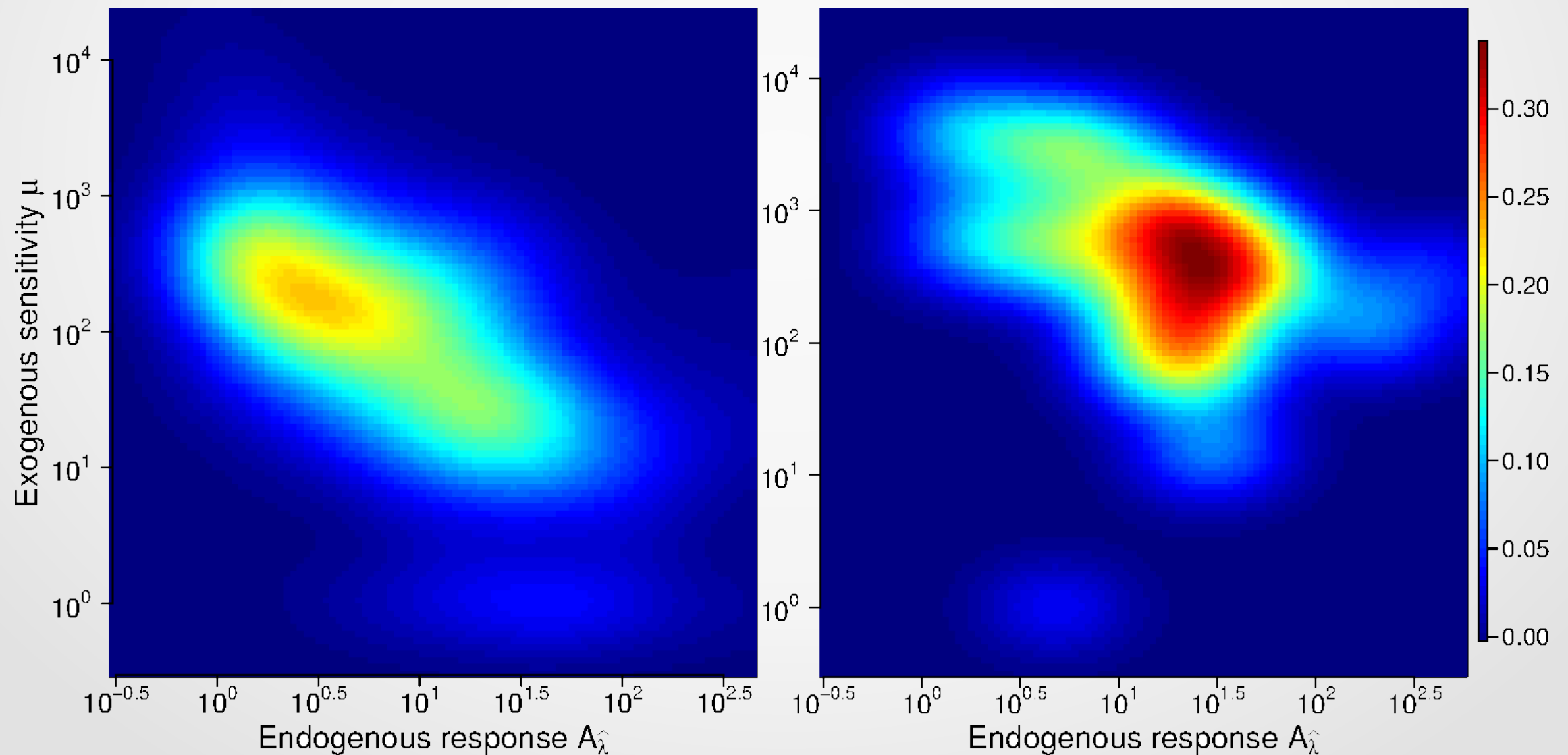




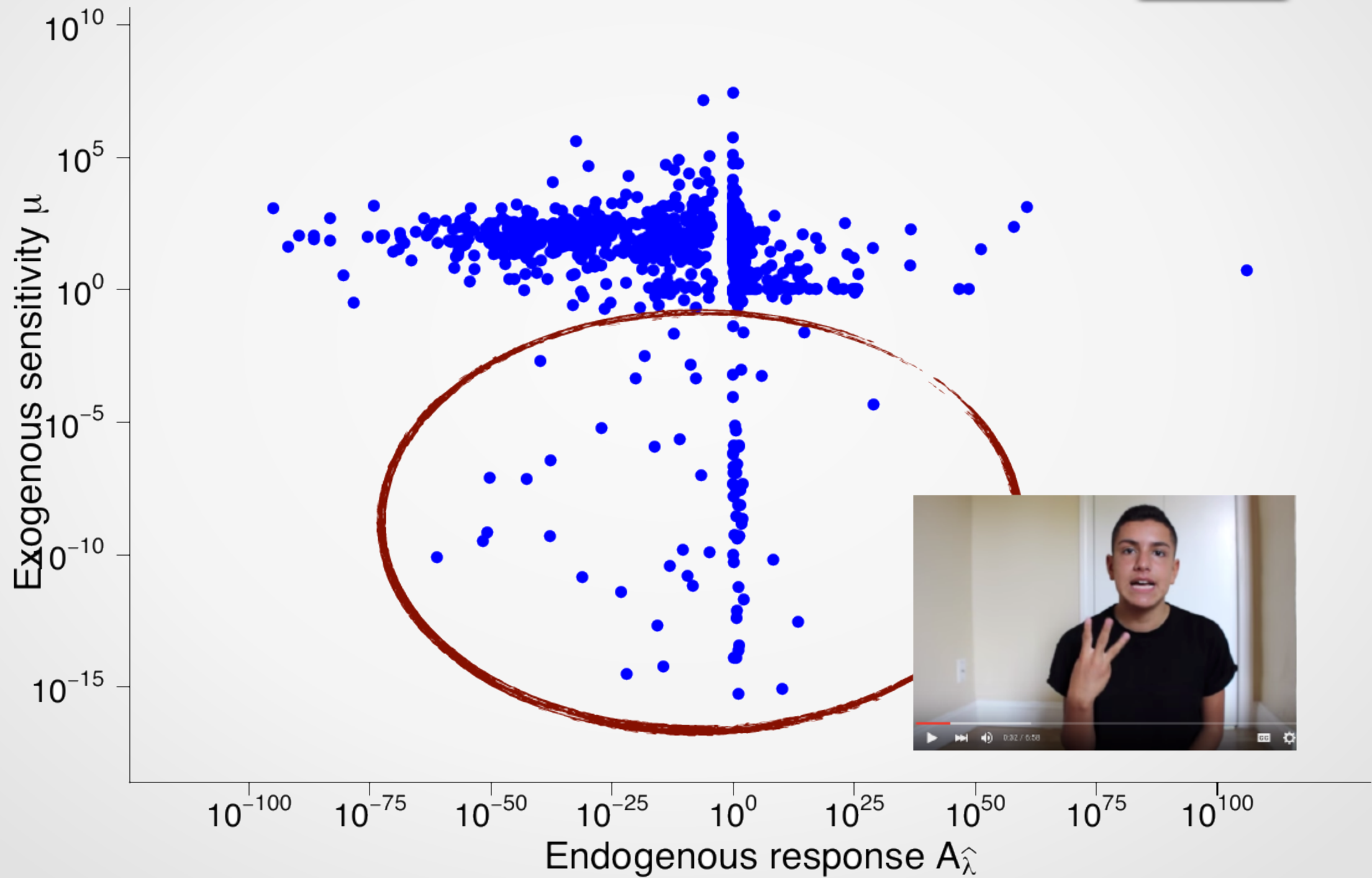
# Explain popularity – all vs top 5%

Games:

more popular videos have higher endogenous response

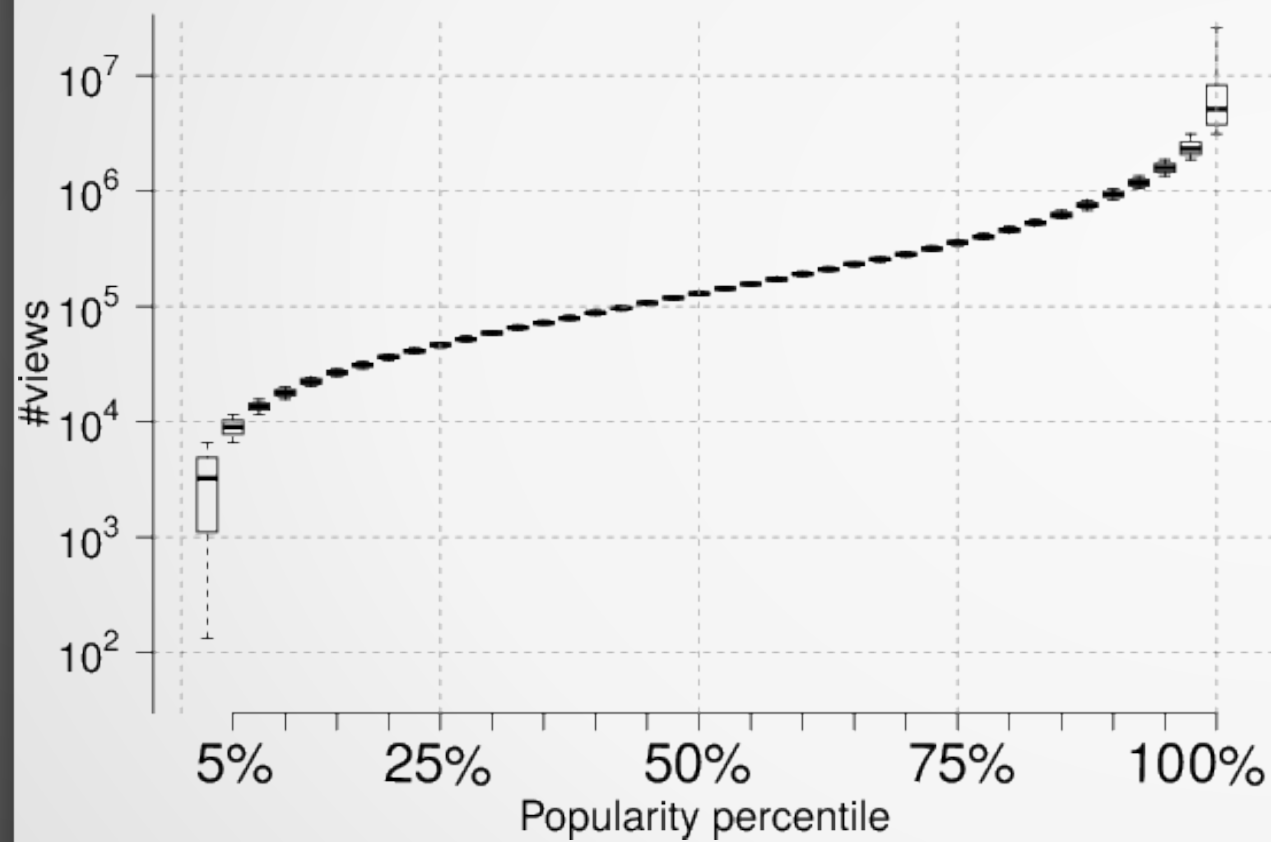


# Which videos are un-promotable?



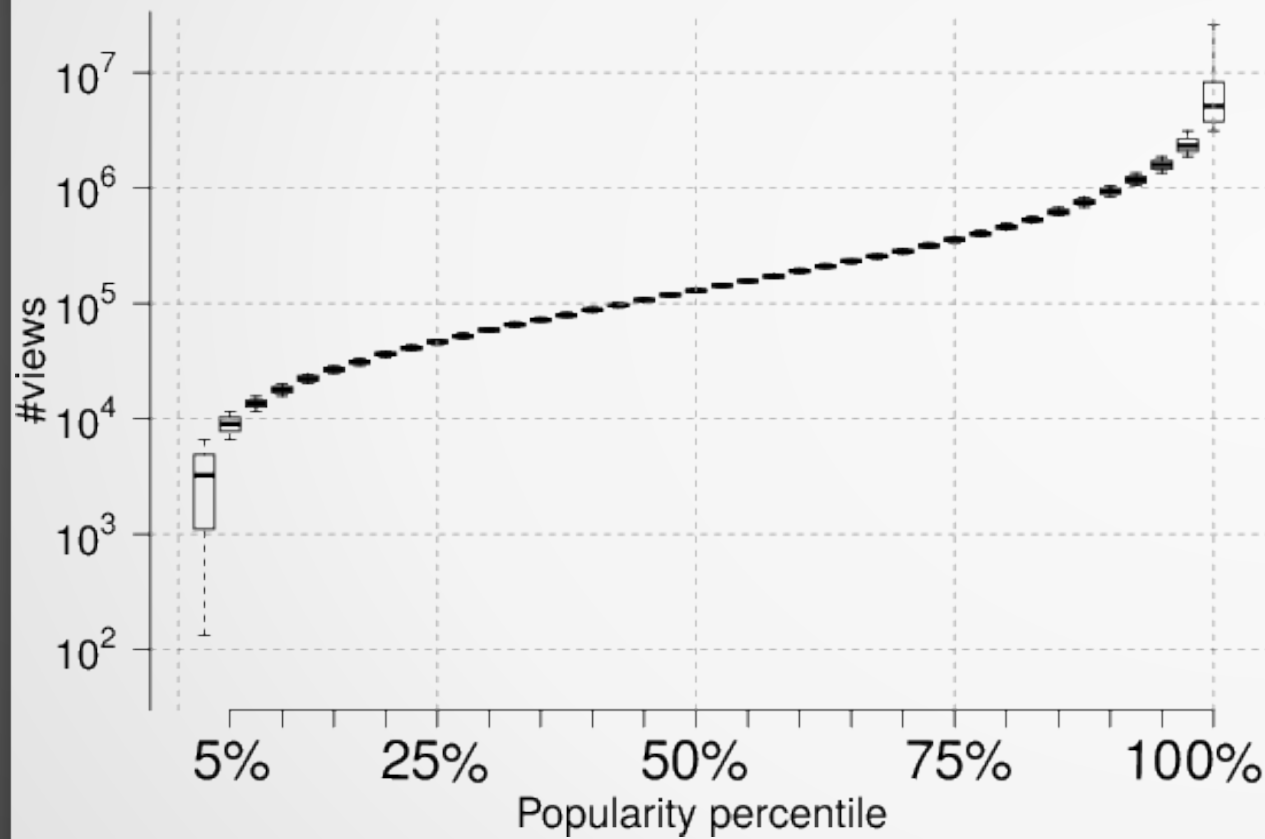
# Popularity scales over time

views: Popularity scale at 60 days

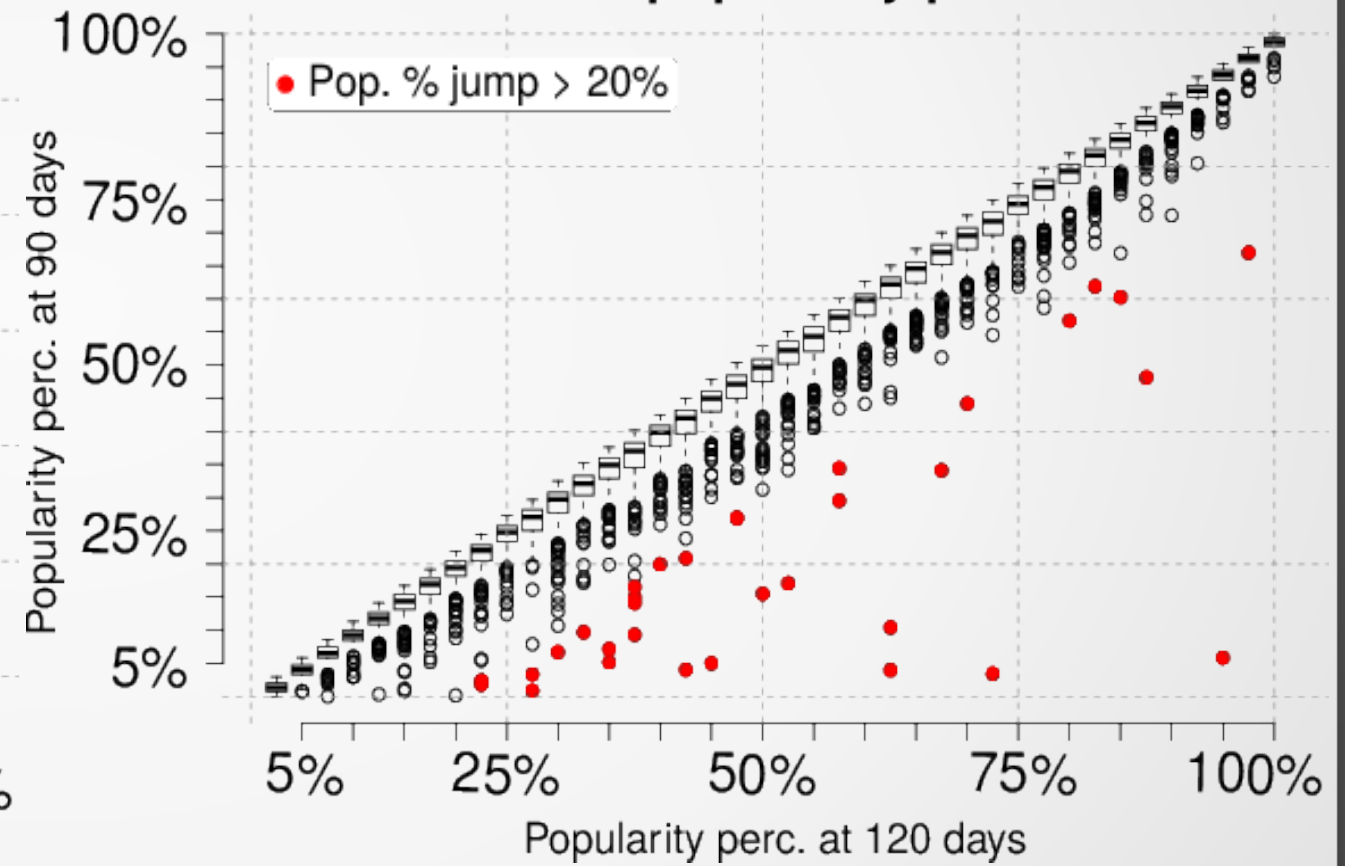


# Popularity scales over time

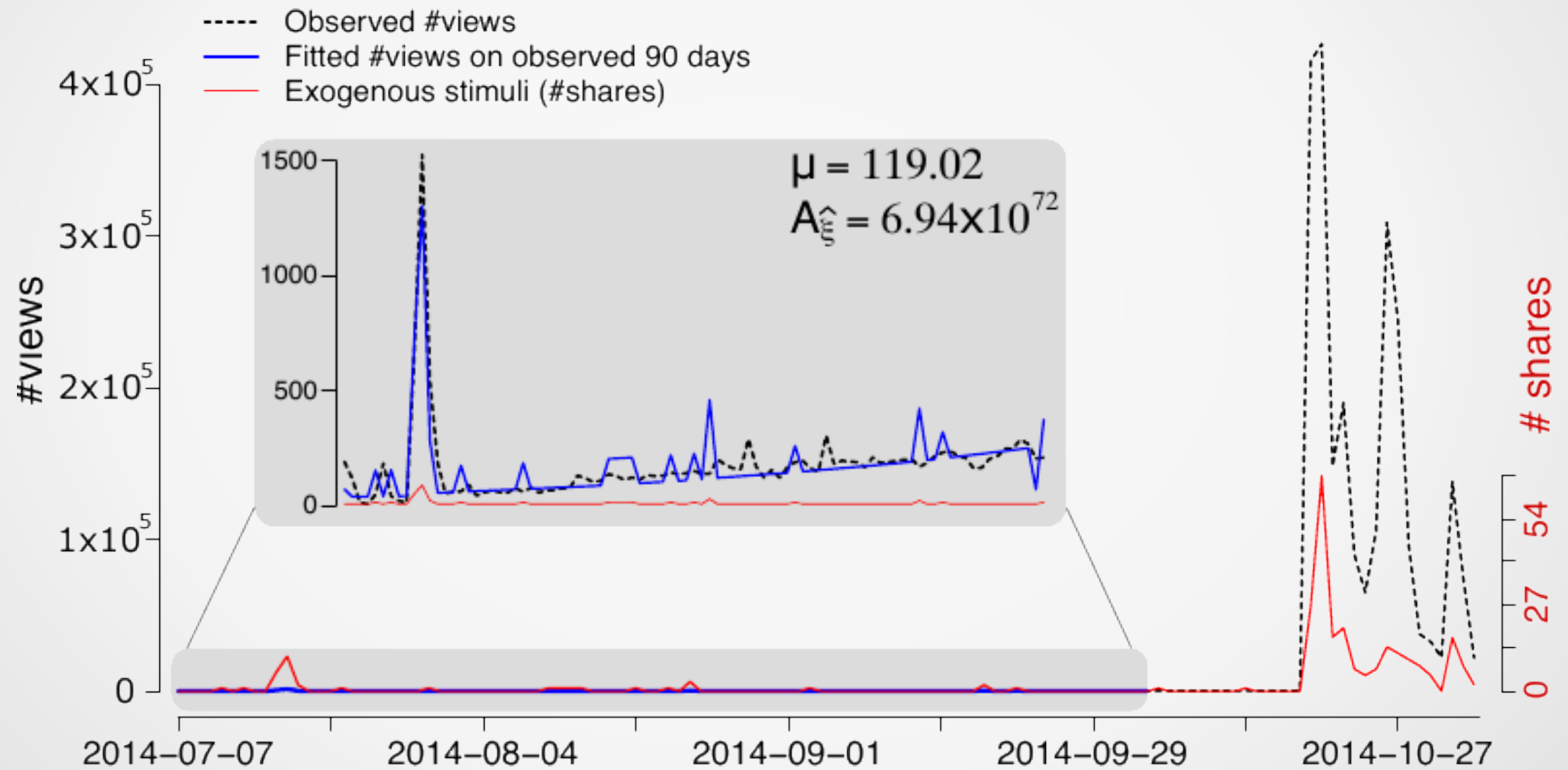
views: Popularity scale at 60 days



Evolution of popularity percentiles



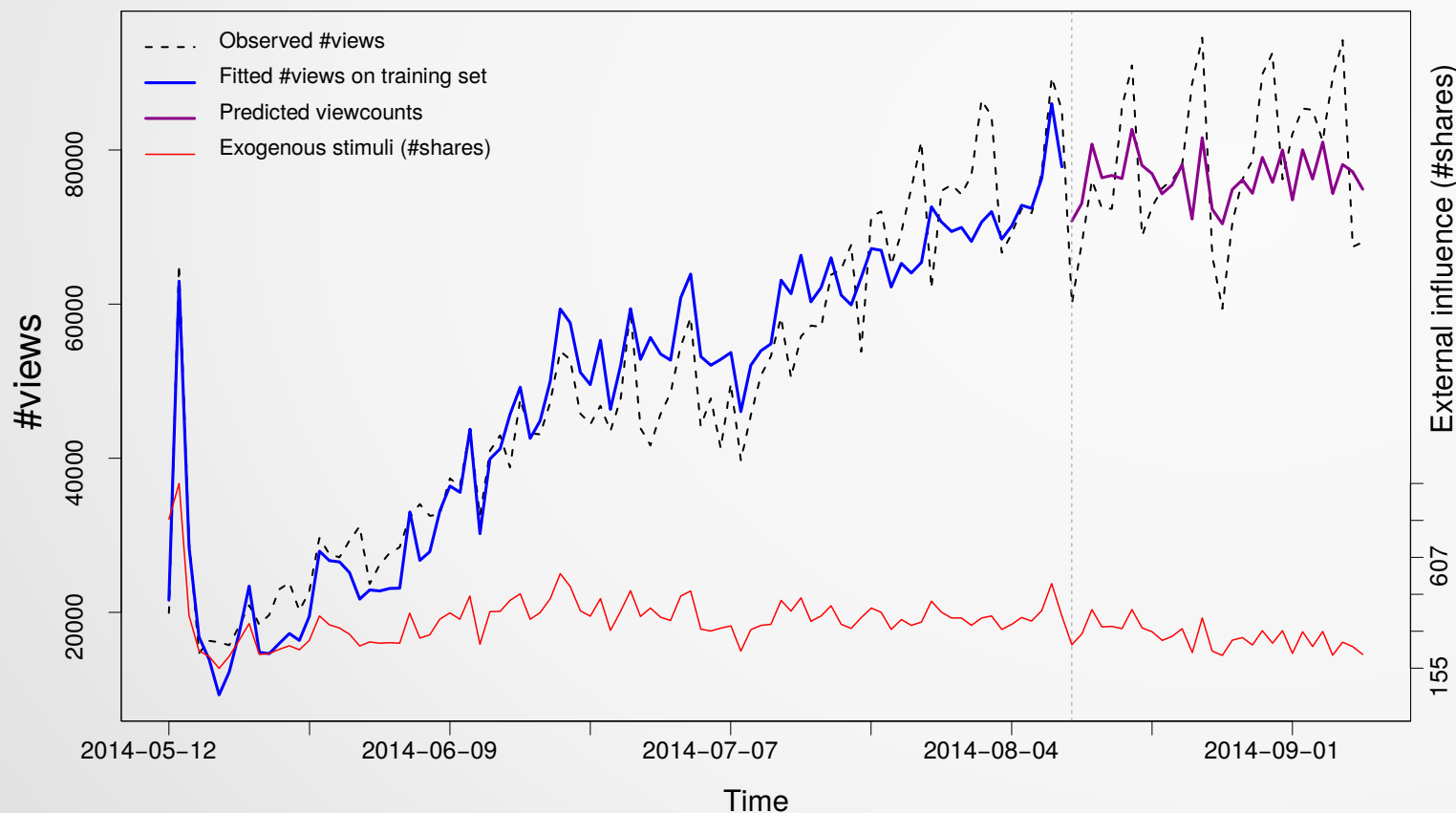
# “Potentially viral” video



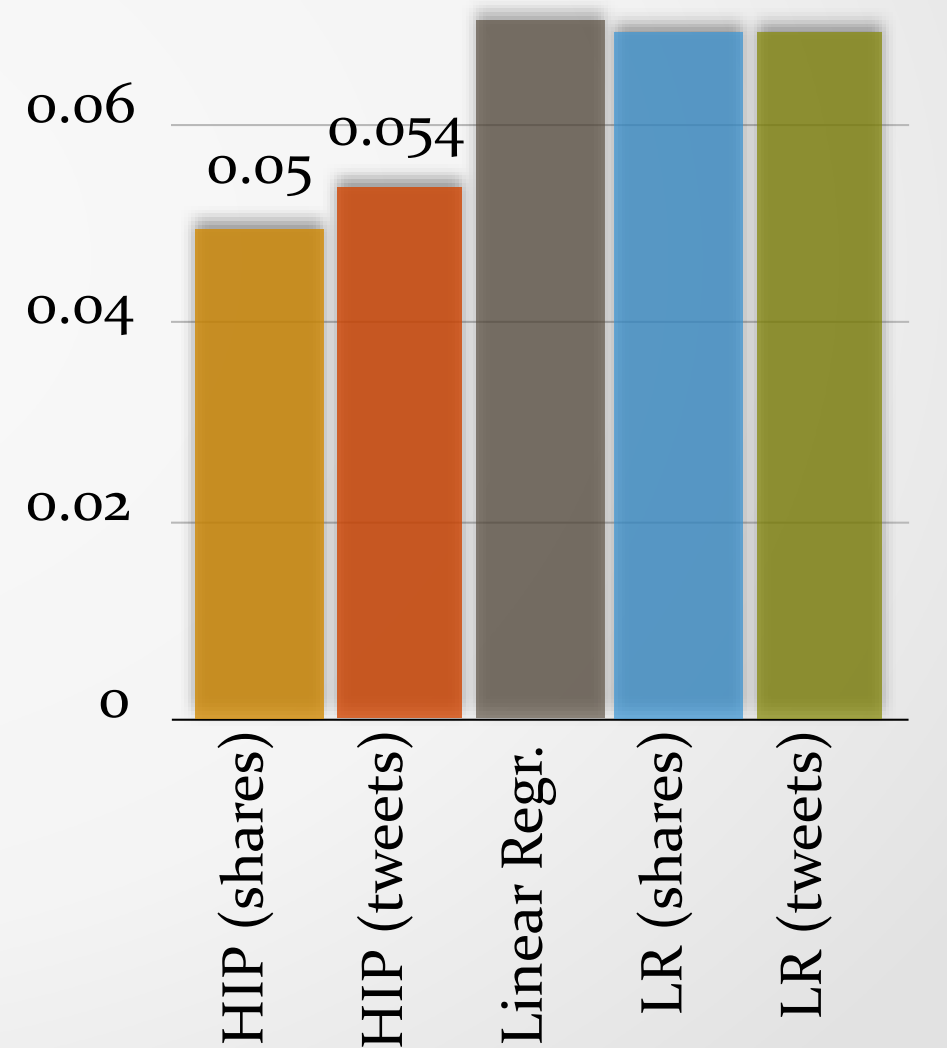


# Forecasting the effect of promotions

Observed and predicted popularity with confidence interval



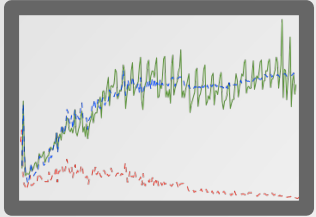
average error in popularity percentile



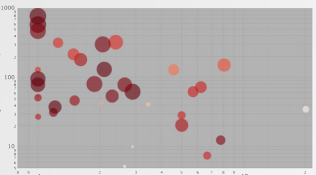
[Pinto et al WSDM'13]

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

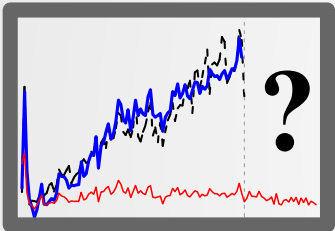
# Summary



HIP: a mathematical model linking promotion and popularity



Explain popularity dynamics and identify potentially viral videos



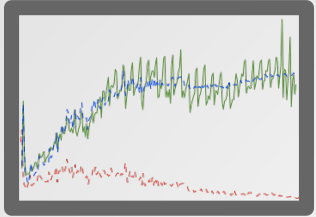
Forecast future popularity as a result of promotion.

## Next steps:

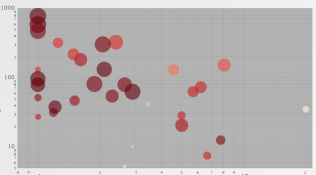
Predict popularity jumps, design promotion schedules

**To appear in ICWSM '17, Montréal, Canada**

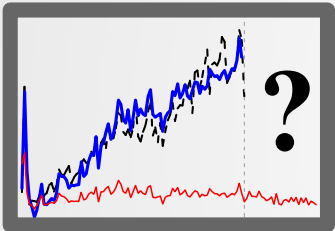
# Summary



HIP: a mathematical model linking promotion and popularity



Explain popularity dynamics and identify potentially viral videos



Forecast future popularity as a result of promotion.

## Next steps:

Predict popularity jumps, design promotion schedules

**To appear in ICWSM '17, Montréal, Canada**

**Limitations & future work:** unobserved sources of external influence, seasonality, network structure

# Thank you!

## Links:

Code, dataset and interactive visualizer:

<https://github.com/andrei-izoiu/hip-popularity>

Referece:

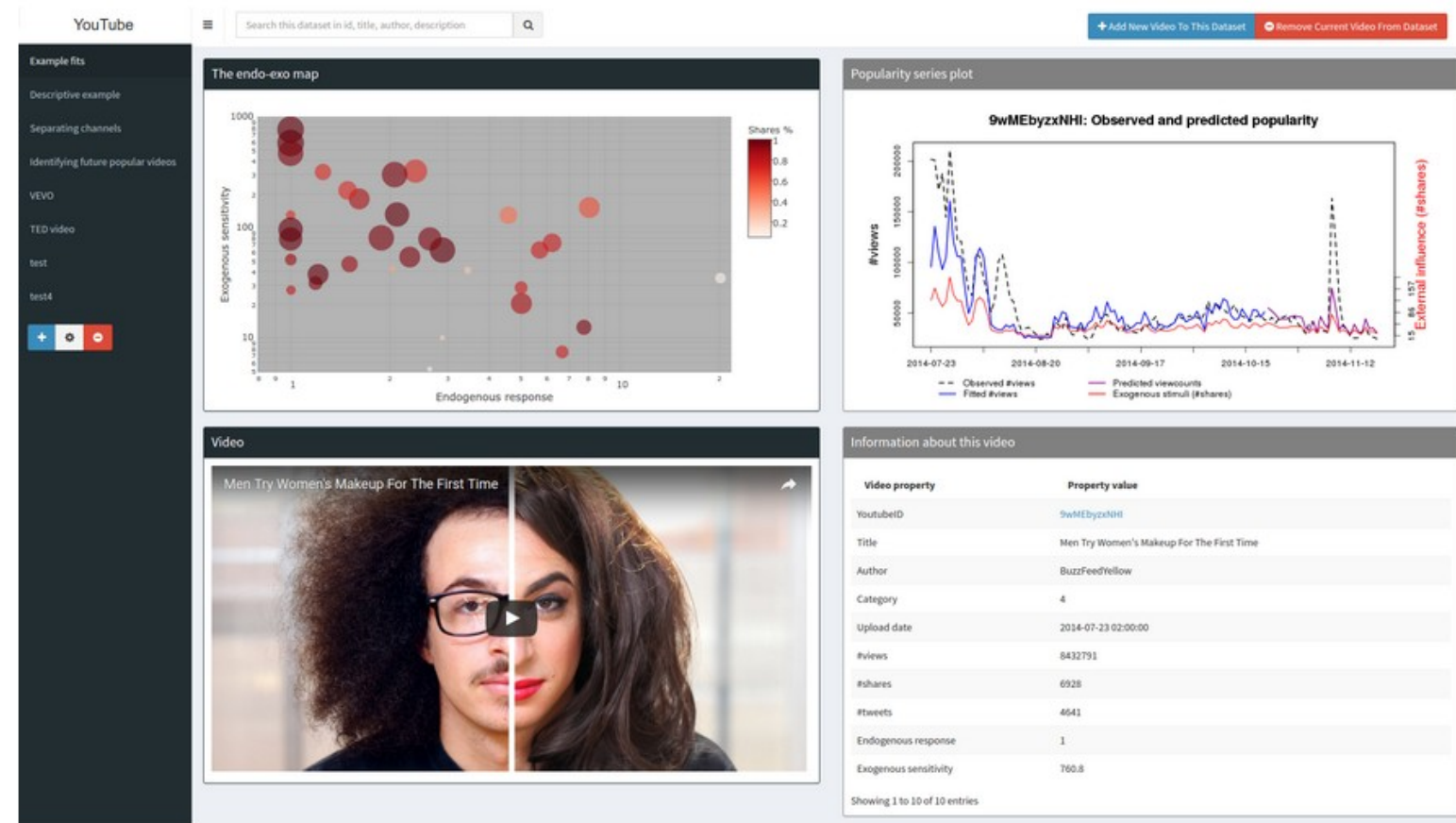
Rizolu, M.-A., Xie, L., Sanner, S., Cebrian, M., Yu, H., & Van Hentenryck, P. (2017). **Expecting to be HIP: Hawkes Intensity Processes for Social Media Popularity**. In Proceedings of the *International Conference on World Wide Web 2017*, pp. 1-9. Perth, Australia. doi: [10.1145/3038912.3052650](https://doi.org/10.1145/3038912.3052650)

[pdf at arxiv with supplementary material](#)

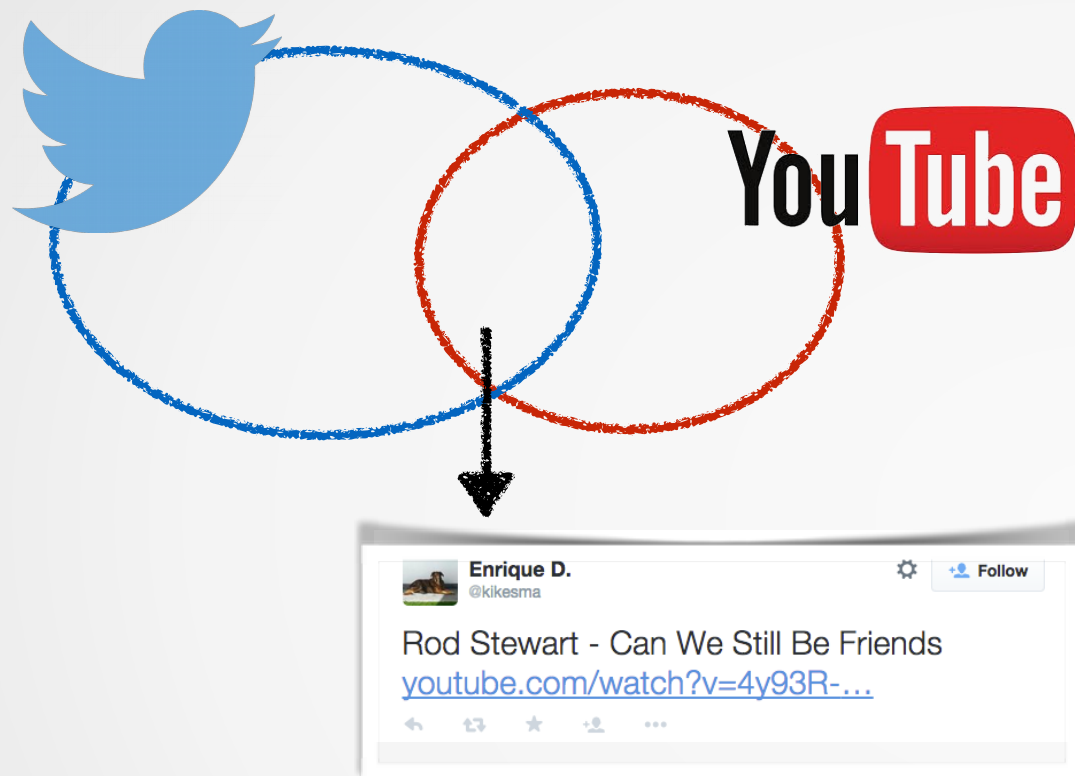
## HIP visualization system

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: Dataset



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

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	



# Supp: Prior work and gaps

## 1) Modeling popularity

power-law shapes [Crane & Sornette PNAS'08]

power-law decays with periodicity [Matsubara et al KDD'12]

collection of recurrence peaks [Cheng et al WWW'16]

How would popularity evolve under continuous external influence?

## 2) Explaining virality

diffusion history [Cheng et al WWW'14]

positive sentiment [Bakshy et al WSDM'11]

Can something go viral if promoted?

## 3) Predicting future popularity

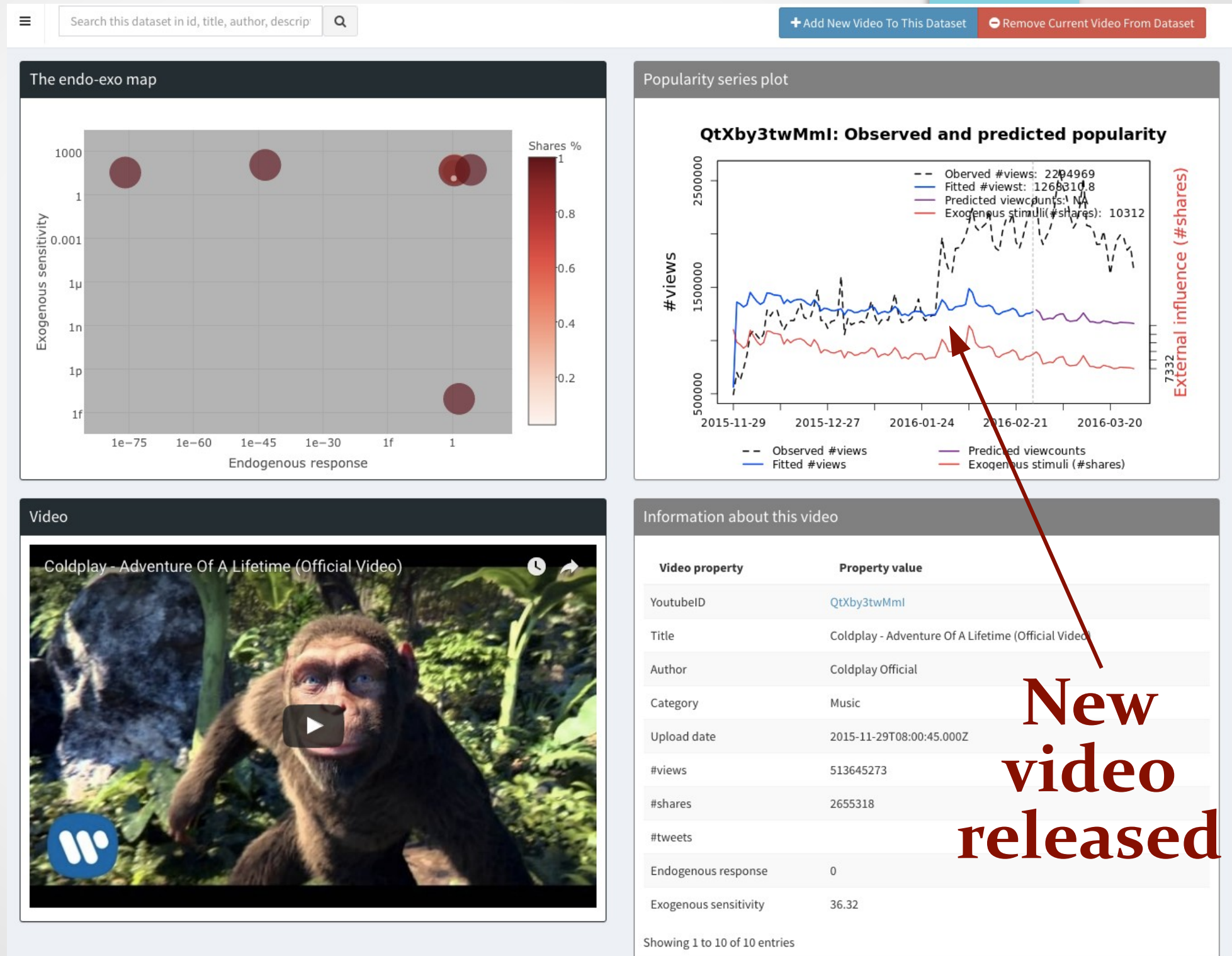
popularity history [Pinto et al WSDM'13] [Szabo and Huberman Comm.ACM 10]

timing features [Cheng et al WWW'14]

How to forecast future popularity given planned promotions?

# Supp: when HIP fails the fitting (1)

## Relations between videos:



**New video released**

# Supp: when HIP fails the fitting (2)

Long term evolutions:

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

### The endo-exo map

### Popularity series plot

**YkjpeuMNEk: Observed and predicted popularity**

-- Observed #views: 2925428  
— Fitted #viewst: 2118798.3  
— Predicted viewcounts: NA  
— Exogenous stimuli(#shares): 17370

### Video

Coldplay - Hymn For The Weekend (Official Video)

### Information about this video

Video property	Property value
YoutubelID	YkjpeuMNEk
Title	Coldplay - Hymn For The Weekend (Official Video)
Author	Coldplay Official
Category	Music
Upload date	2016-01-29T15:00:38.000Z
#views	694792952
#shares	4556631
#tweets	
Endogenous response	0
Exogenous sensitivity	121.18

Showing 1 to 10 of 10 entries

Slow drift

31 / 31