Controversy on Social Media: Collective Attention, Echo Chambers, and Price of Bipartisanship

Gianmarco De Francisci Morales
ISI Foundation

with
Kiran Garimella
Aristides Gionis
Michael Mathioudakis

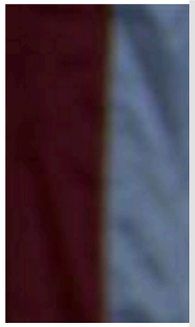
Controversy: from Latin contra (against)
vertere (turn) “turned against, disputed”

Amid Outcry, Trump Continues Campaign of Controversy

by **ALI VITALI**

Amid Outcry, Trump Continues Campaign of Controversy

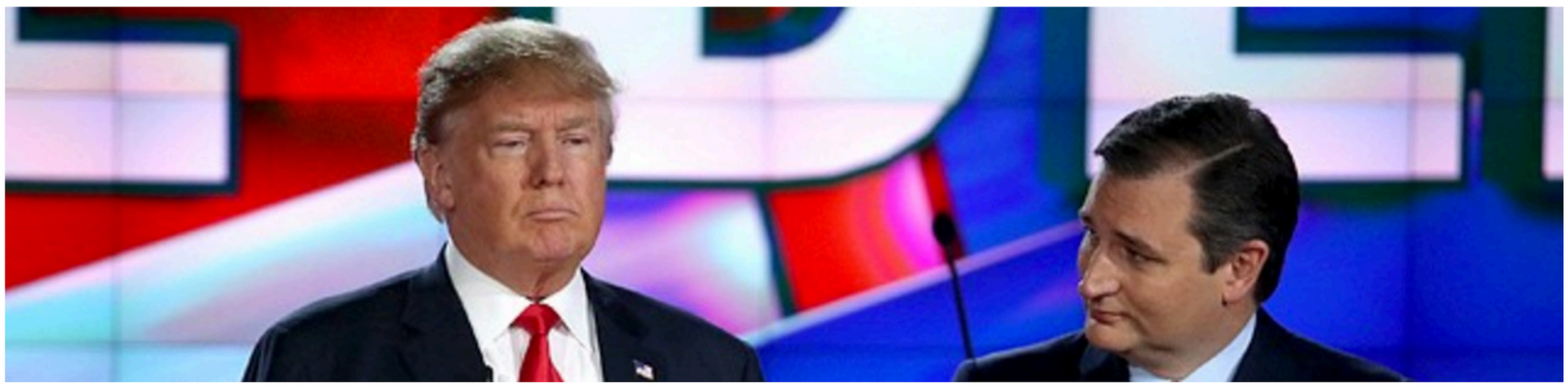
by ALI VITALI



January 12, 2016, 06:00 am

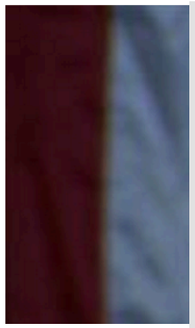
Trump puts Cruz off balance by fanning ‘birther’ controversy

By Alexander Bolton



Amid Outcry, Trump Continues Campaign of Controversy

by ALI VITALI

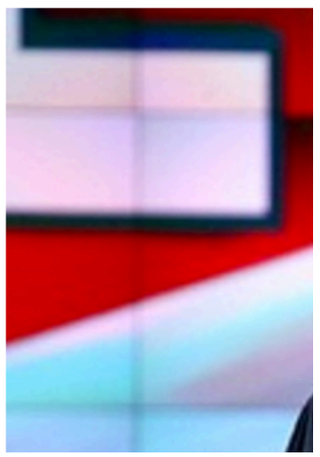


January 12, 2016, 06:00 am

Trump puts Cruz off balance by fanning 'birther' controversy

By Alexander Bolton

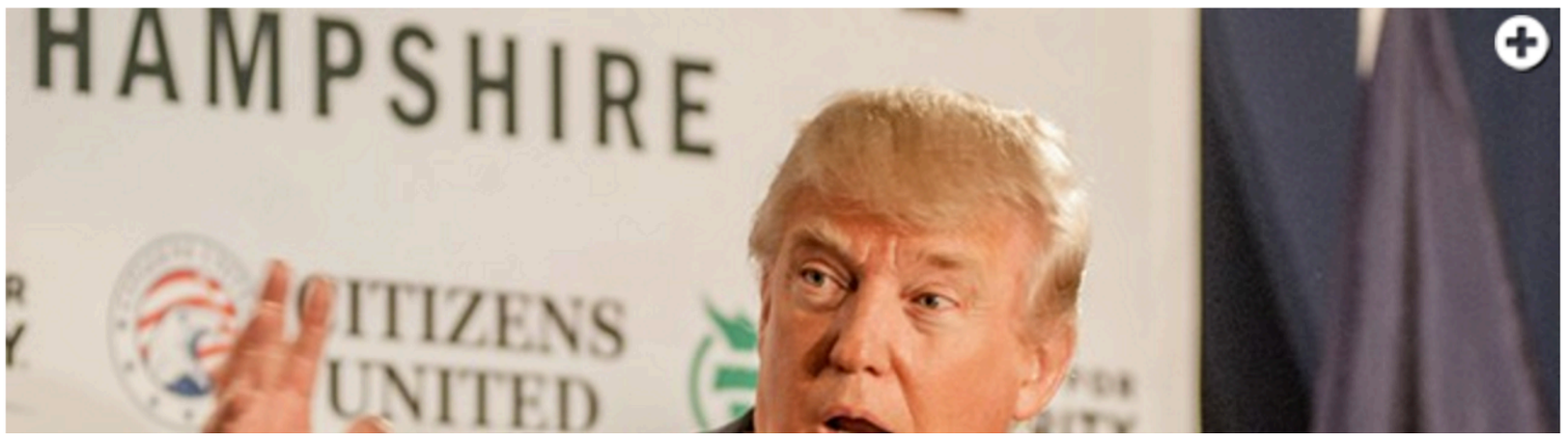
777 1

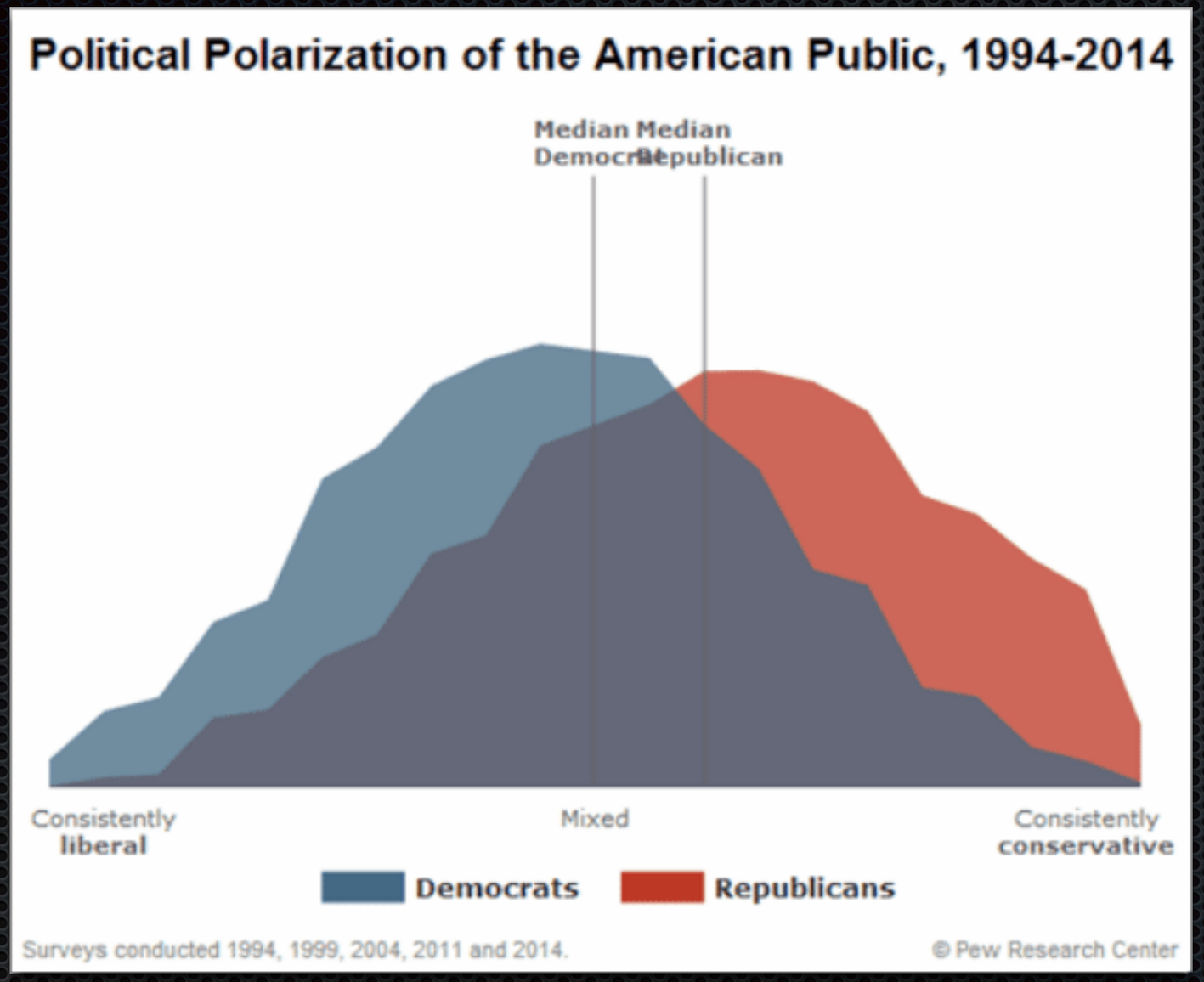


ELECTIONS / U.S. POLITICS

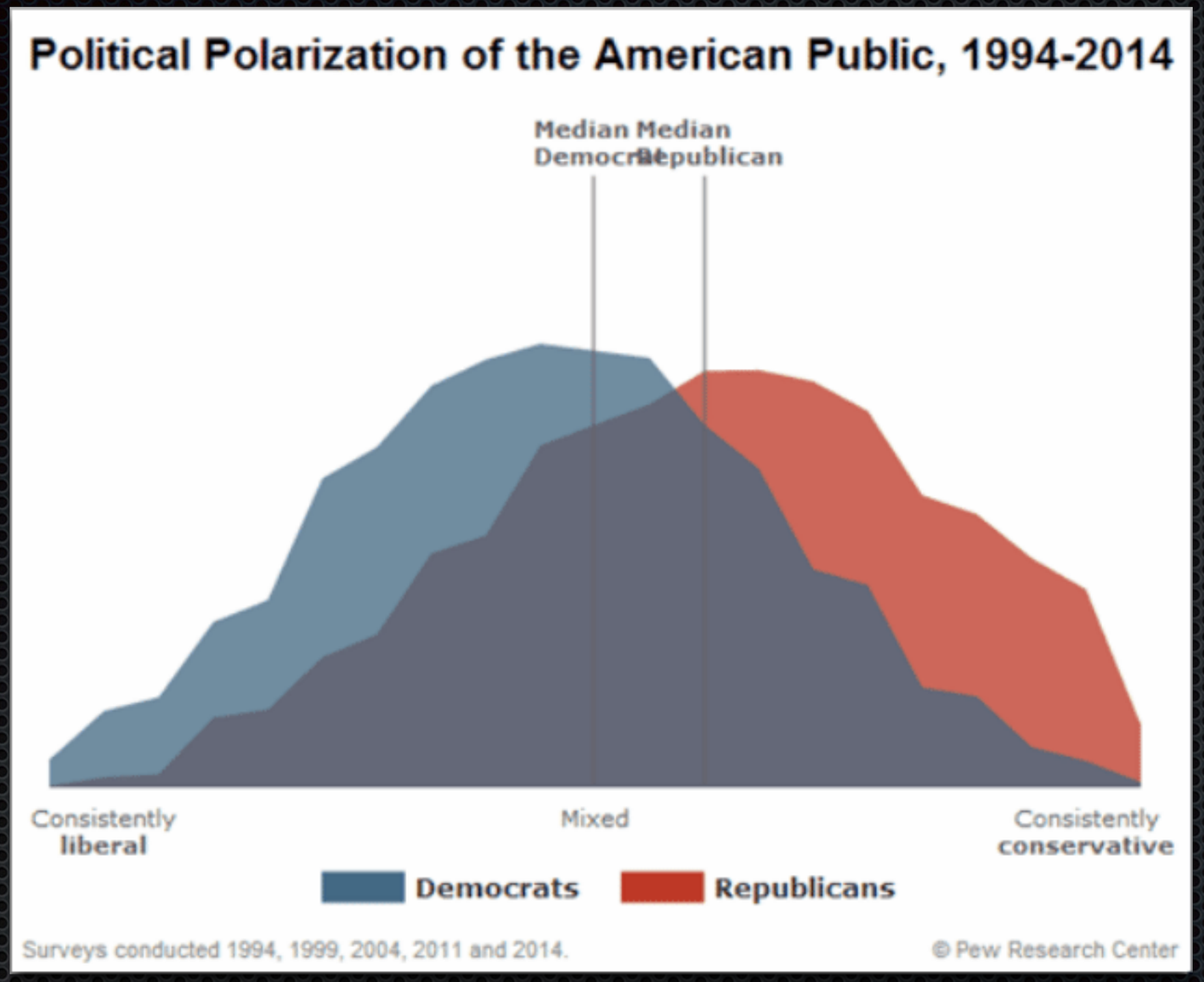
Controversy Arrives Well Ahead of Trump's Burlington Visit

POSTED BY TERRI HALLENBECK ON MON, JAN 4, 2016 AT 8:39 PM





The gap grows



The gap grows

Goal

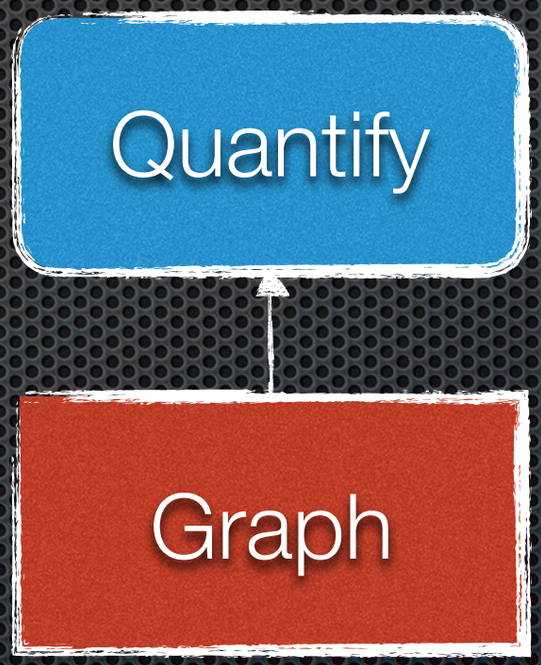
Understand how controversies unfold in social media

Outline

Outline

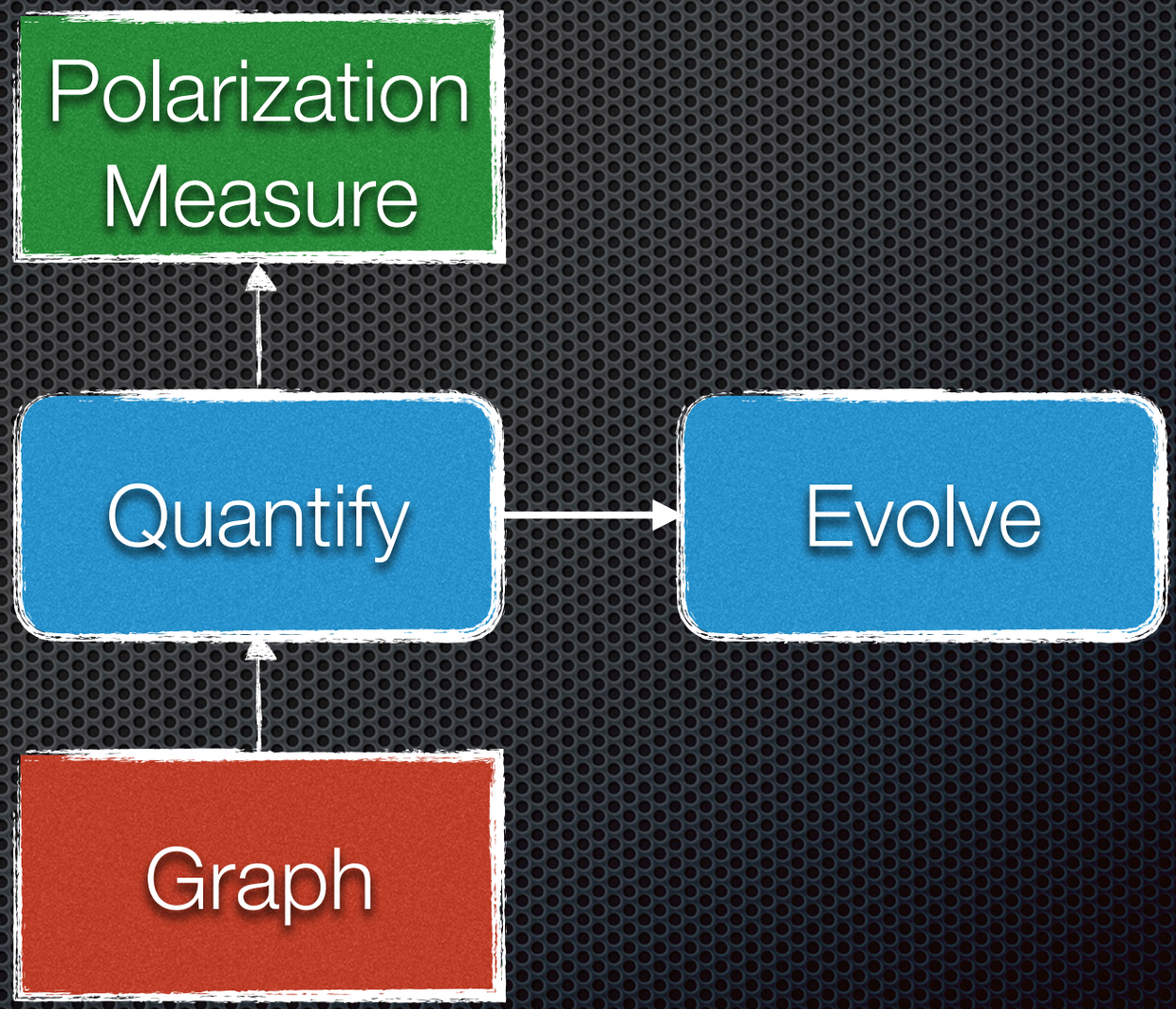
Quantify

Outline

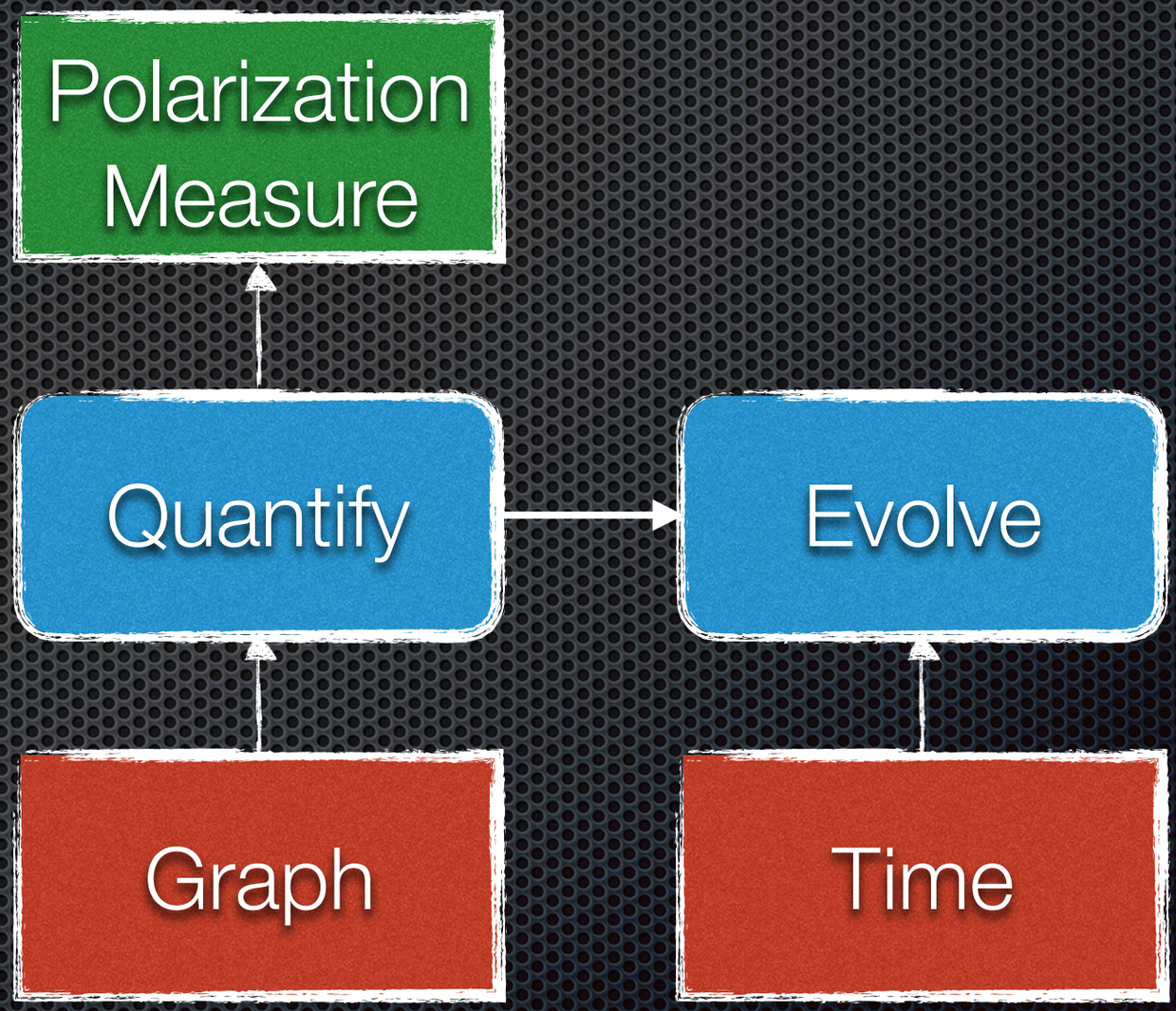


Outline

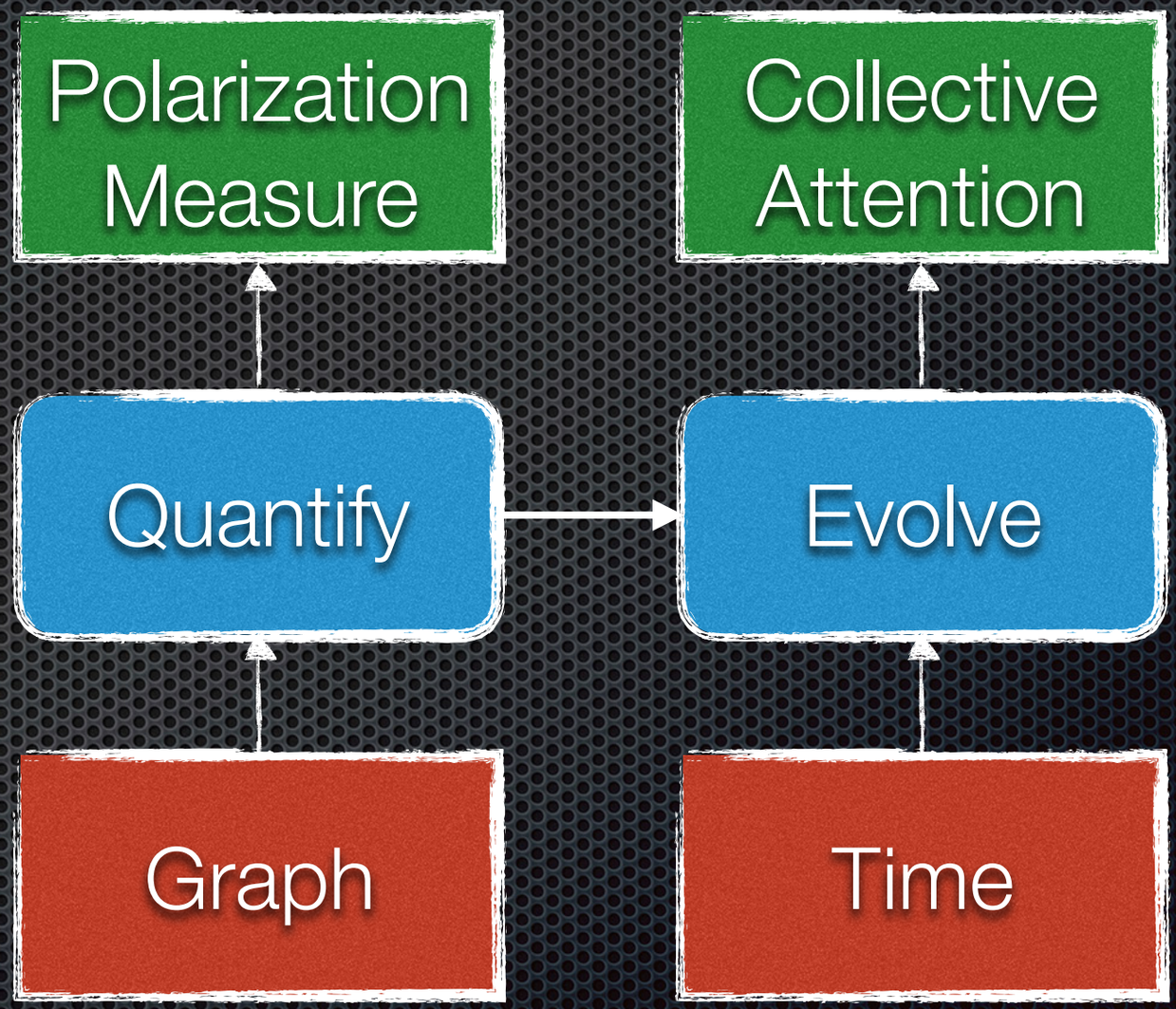
Outline



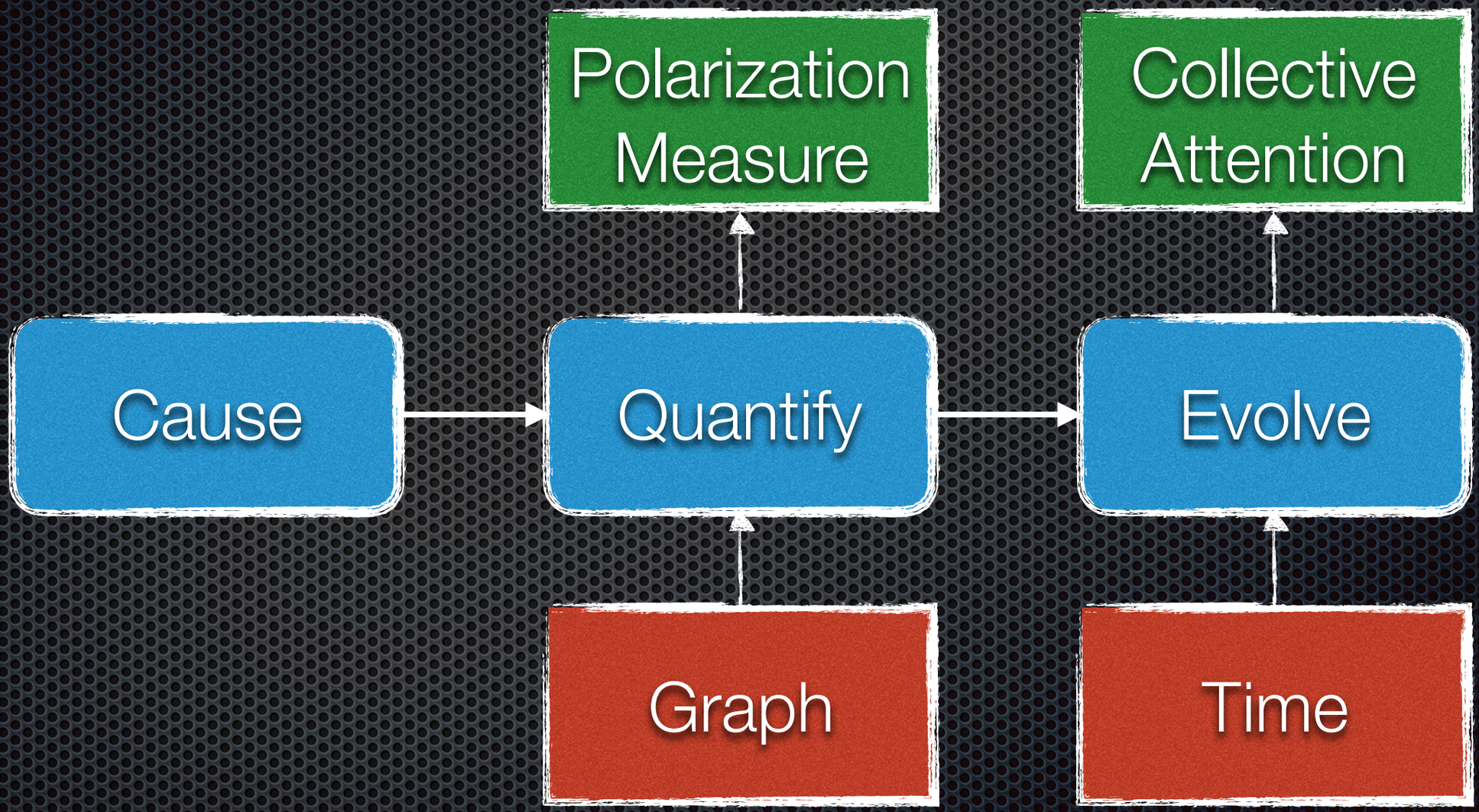
Outline



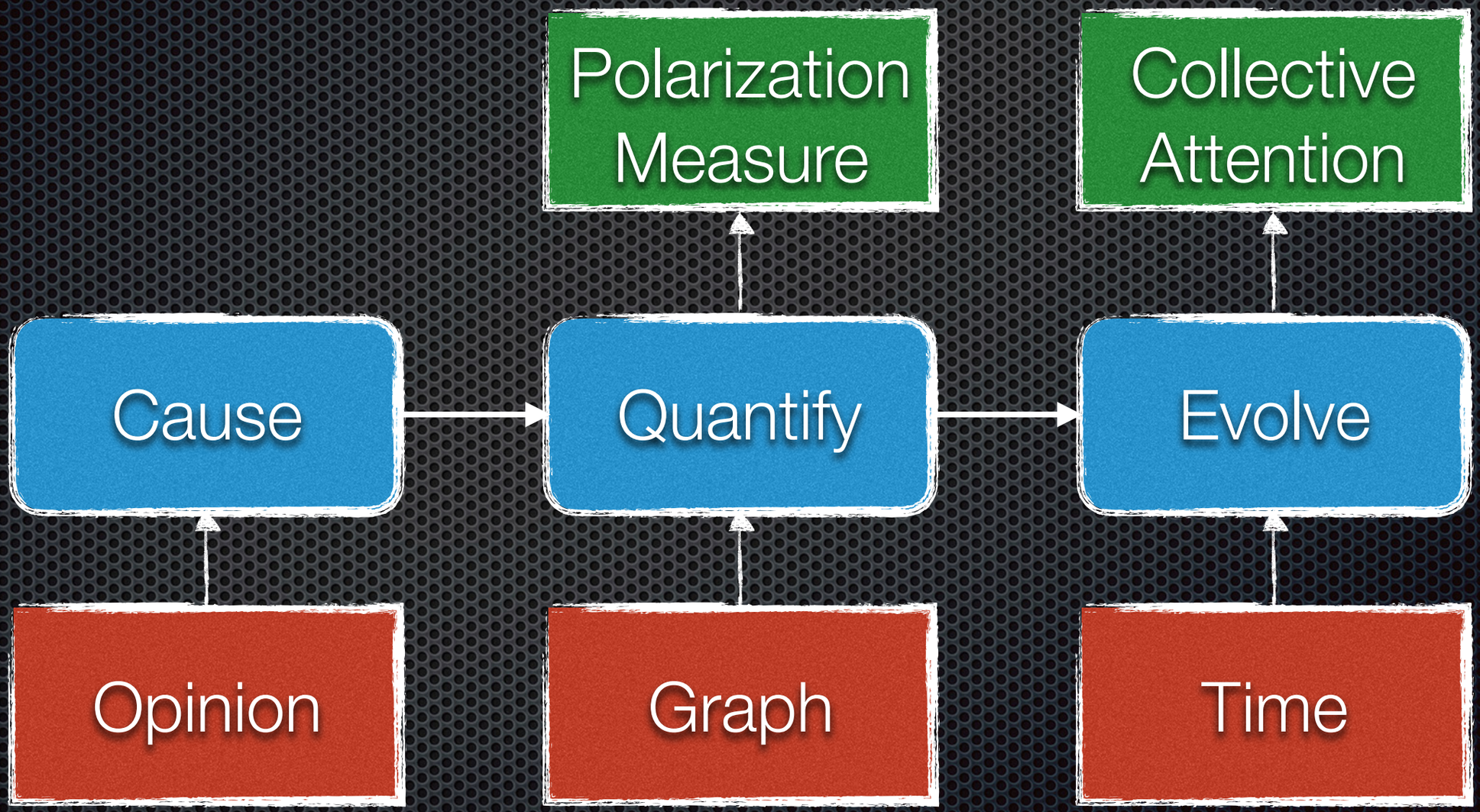
Outline


Polarization Measure
Collective Attention
Quantify
Evolve
Graph
Time

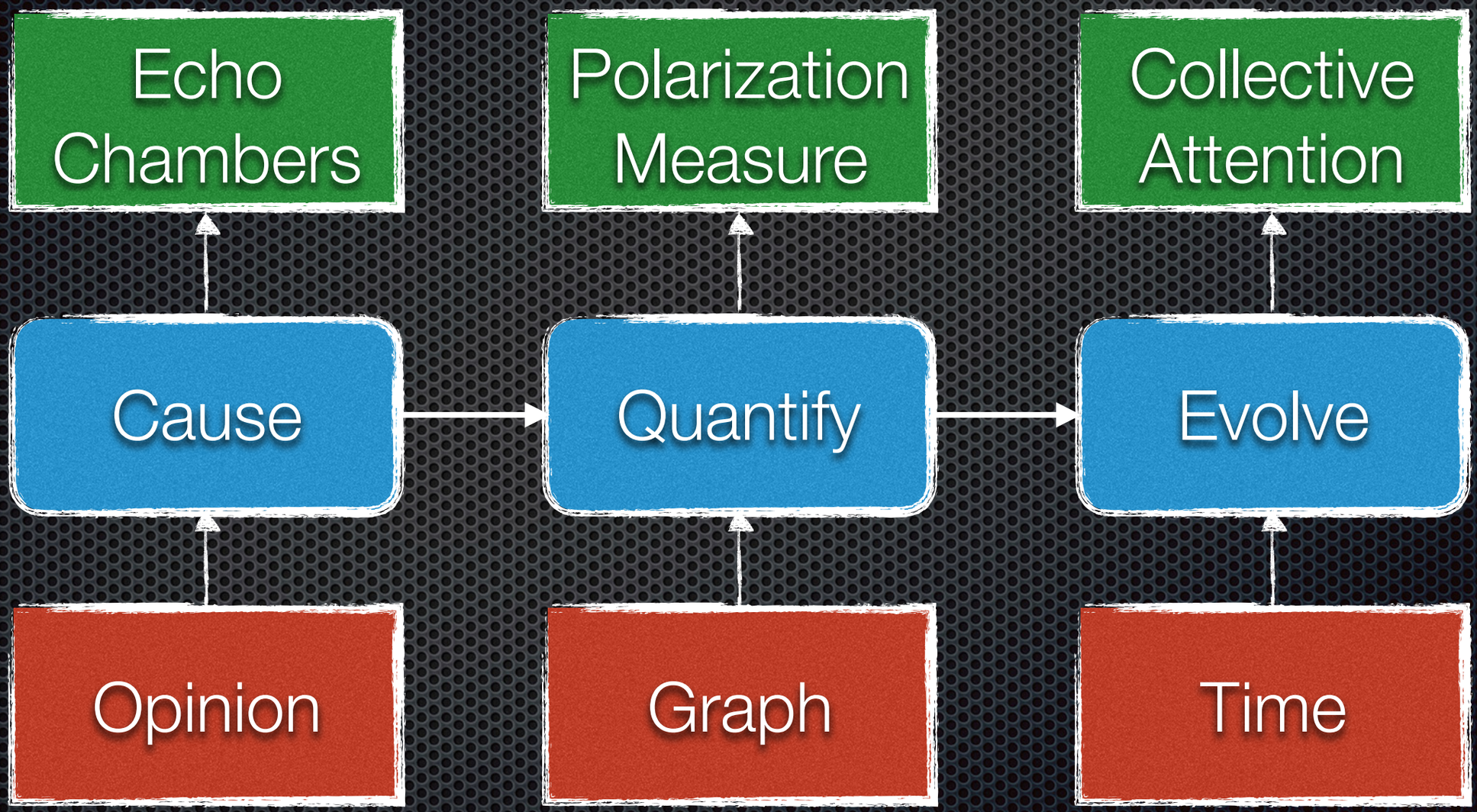
Outline


Polarization Measure
Collective Attention
Cause
Quantify
Evolve
Graph
Time

Outline

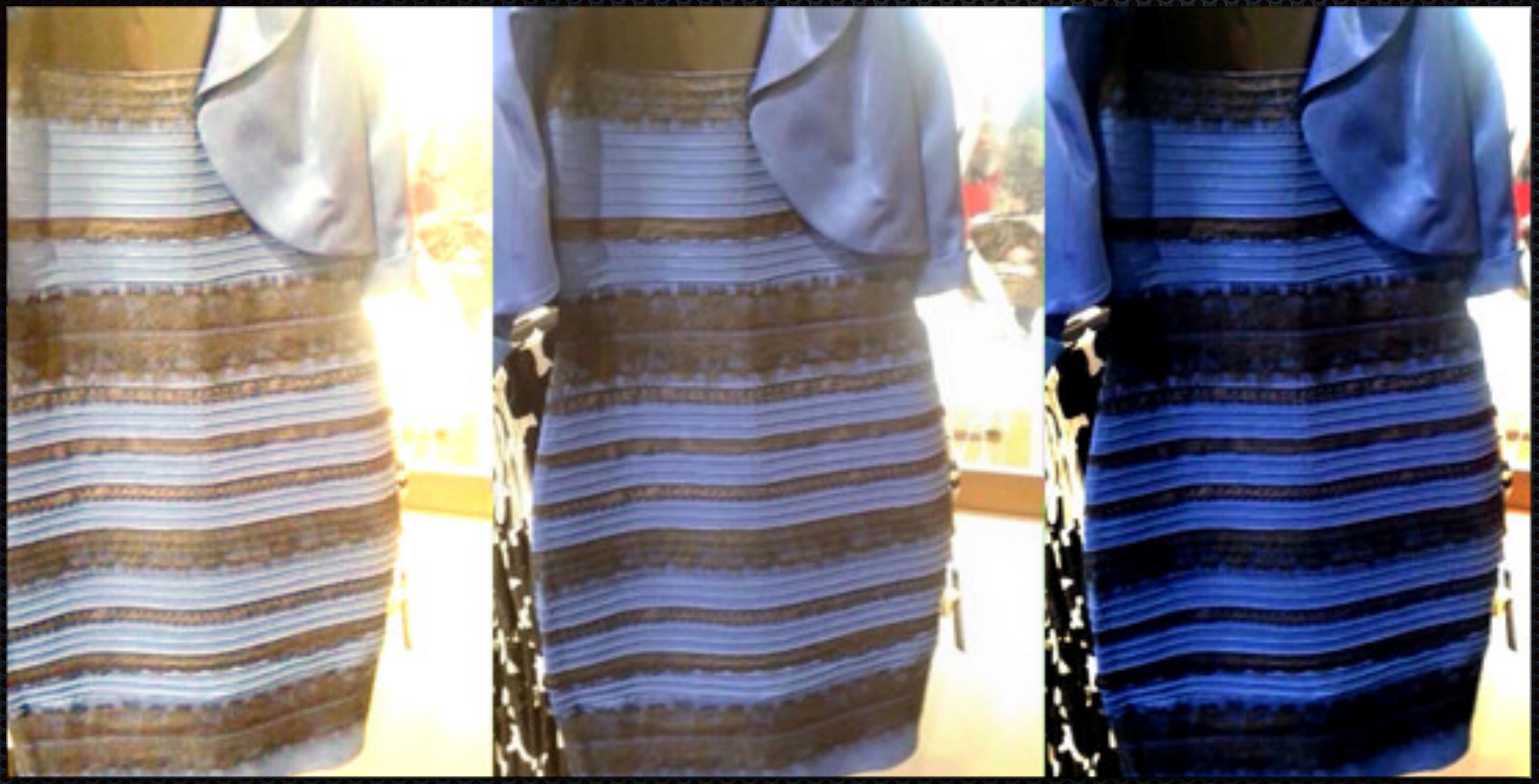

Polarization Measure
Collective Attention
Cause
Quantify
Evolve
Opinion
Graph
Time

Outline


Echo Chambers
Polarization Measure
Collective Attention
Cause
Quantify
Evolve
Opinion
Graph
Time

Quantifying Controversy in Social Media

WSDM 2016, TSC 2018



Black/Blue or White/Gold?

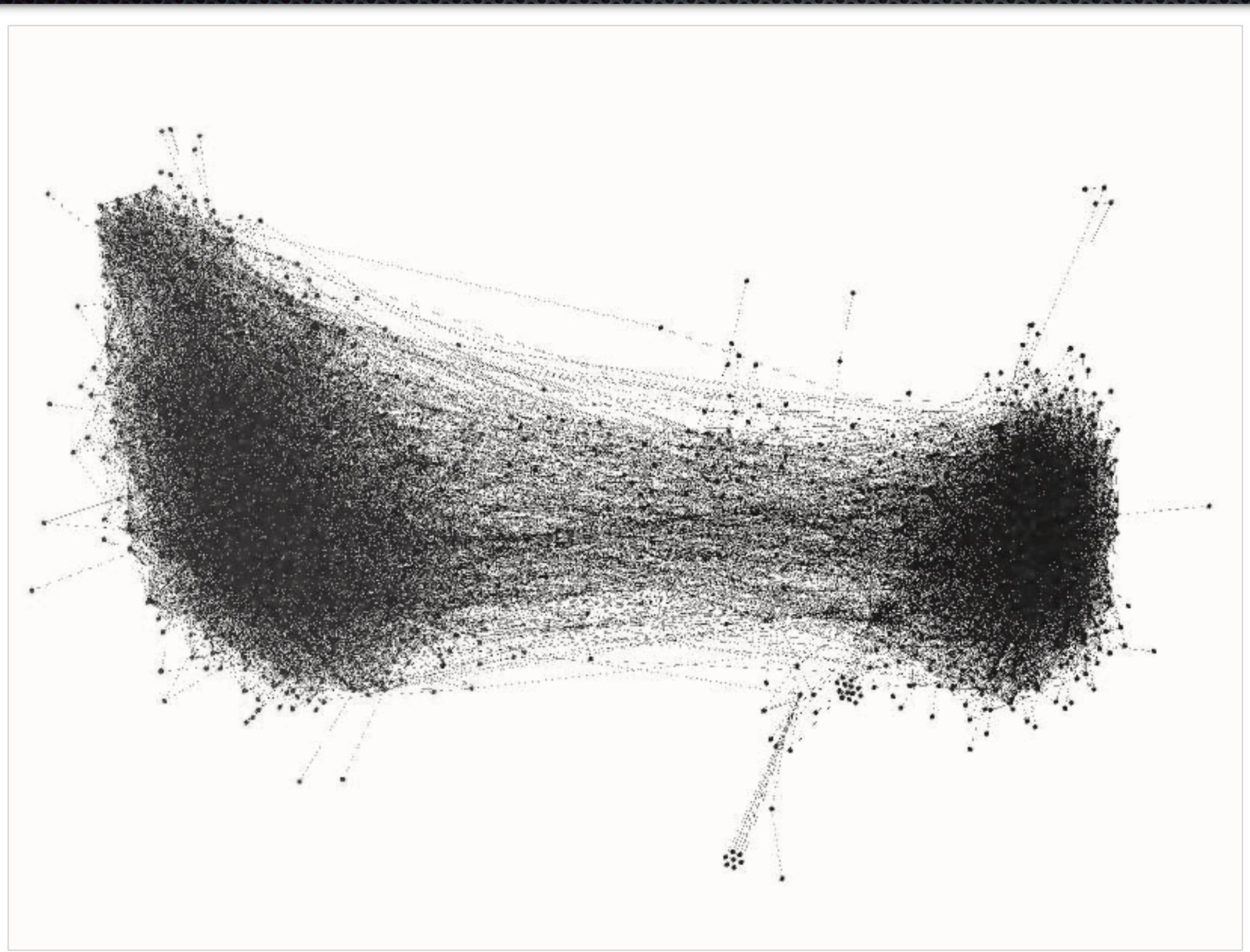
Desiderata

- In the wild
- Not necessarily political
- No domain knowledge
- Language independent
- Allows comparison

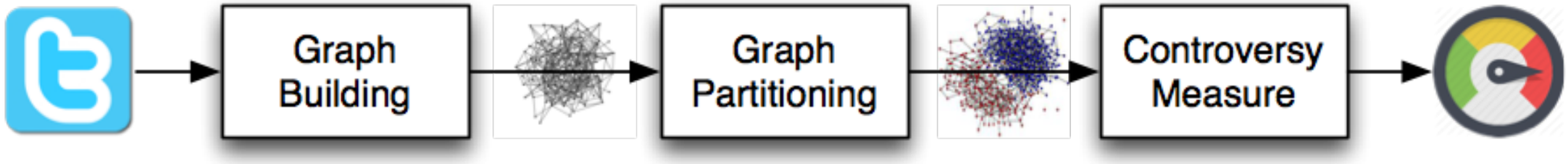
Problem Formulation

- Graph-based unsupervised formulation
- Conversation graph for a topic (endorsements)
- Find partition of graph (represents 2 sides)
- Measure distance between partitions (random walks)

Endorsement Graph



Pipeline

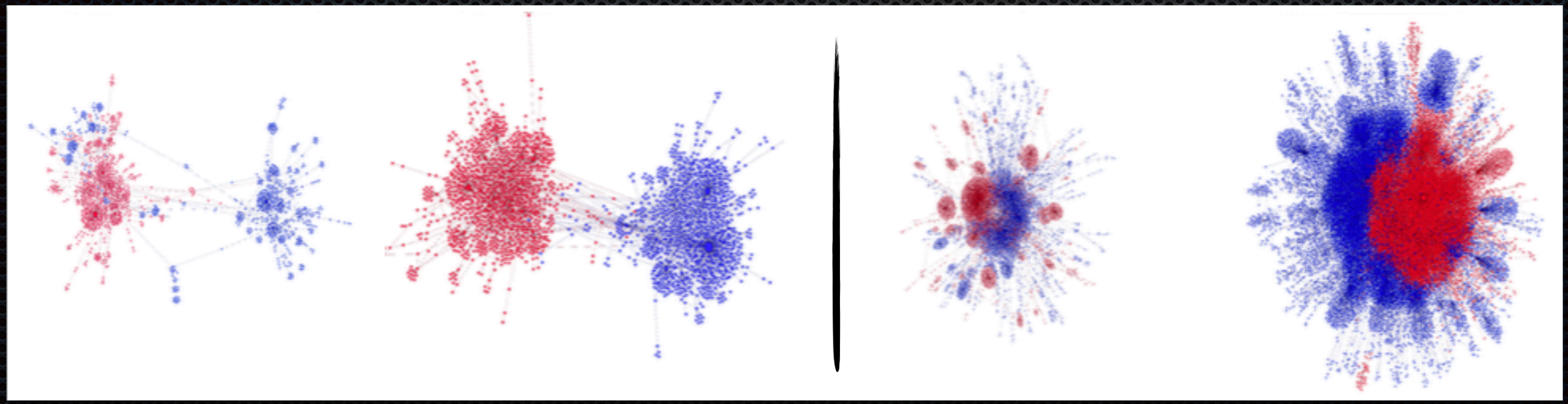


- **Retweets**
- Follow
- Mentions
- Content

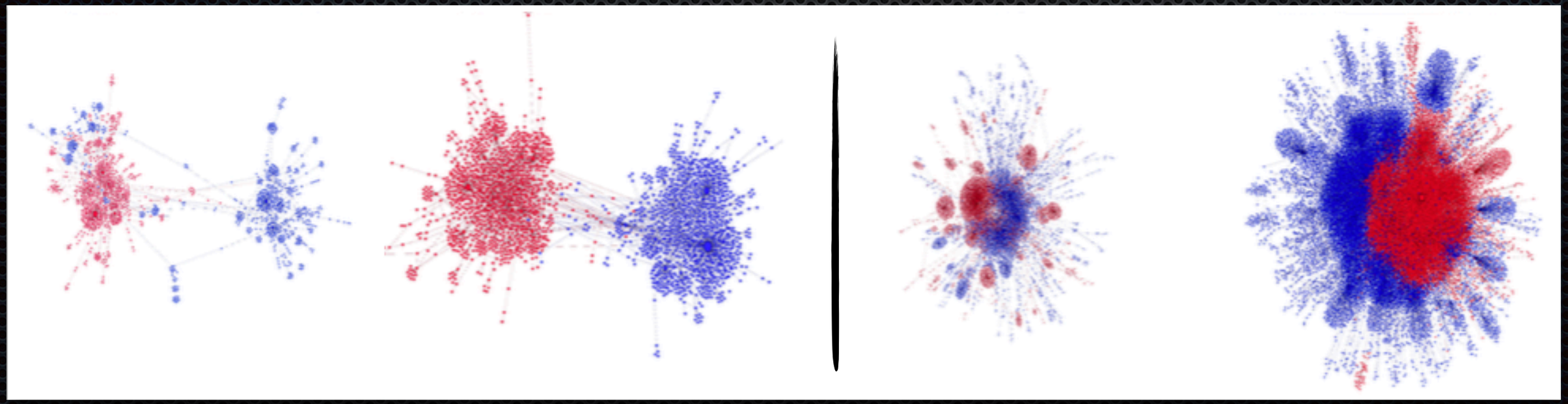
- **METIS**
- Spectral
- Label propagation

- **Random walk**
- Edge betweenness
- 2d embedding
- Sentiment variance

Example

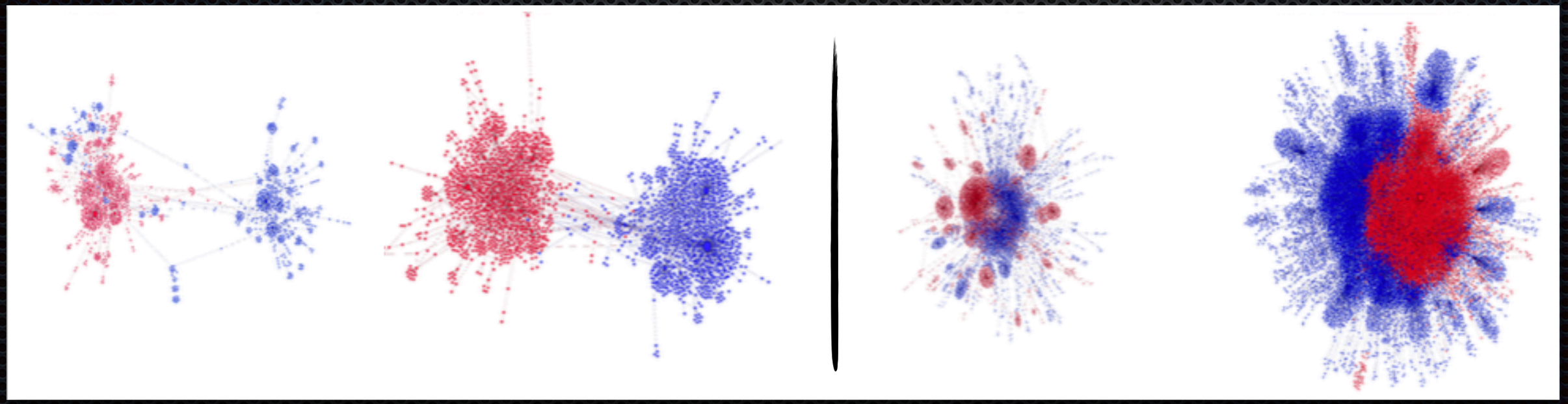


Example



#beefban #марш #sxsw #germanwings

Example



#beefban #марш #sxsw #germanwings

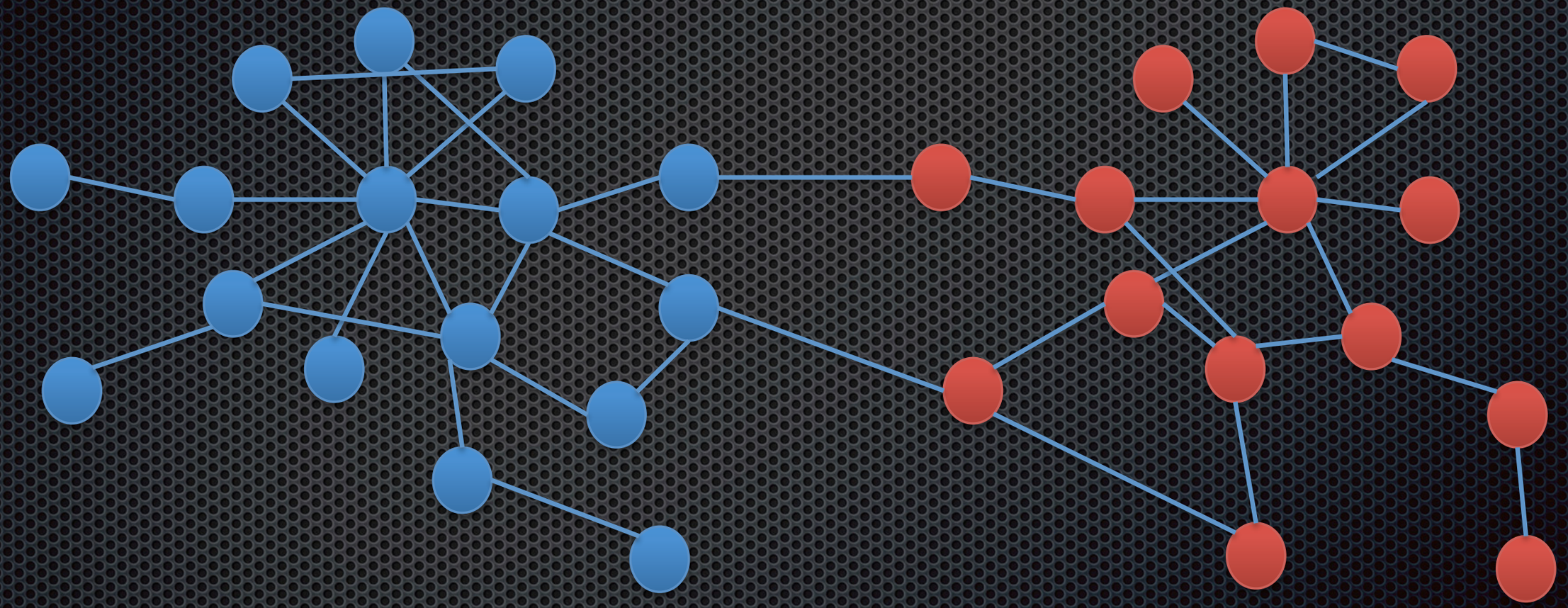
Controversial

Non controversial

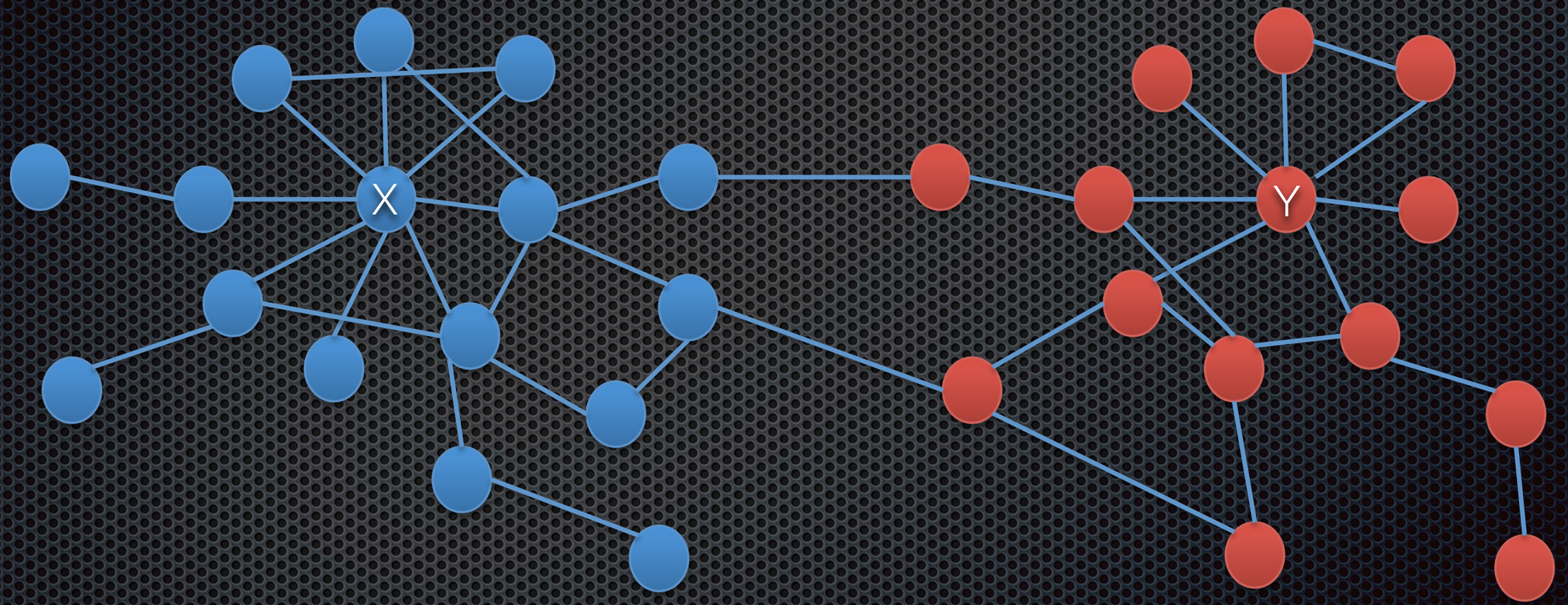
RWC Rationale

- Random Walk Controversy score
- Zaller's RAS model (Receive, Accept, Sample)
  "The Nature and Origins of Mass Opinion"
  - Response Axiom: "Individuals form opinions by averaging across the considerations that are immediately salient or accessible to them"
- Authoritative (influential) users with high degree set opinions
- Measure likelihood of user to be exposed to opinions from influential users on either side

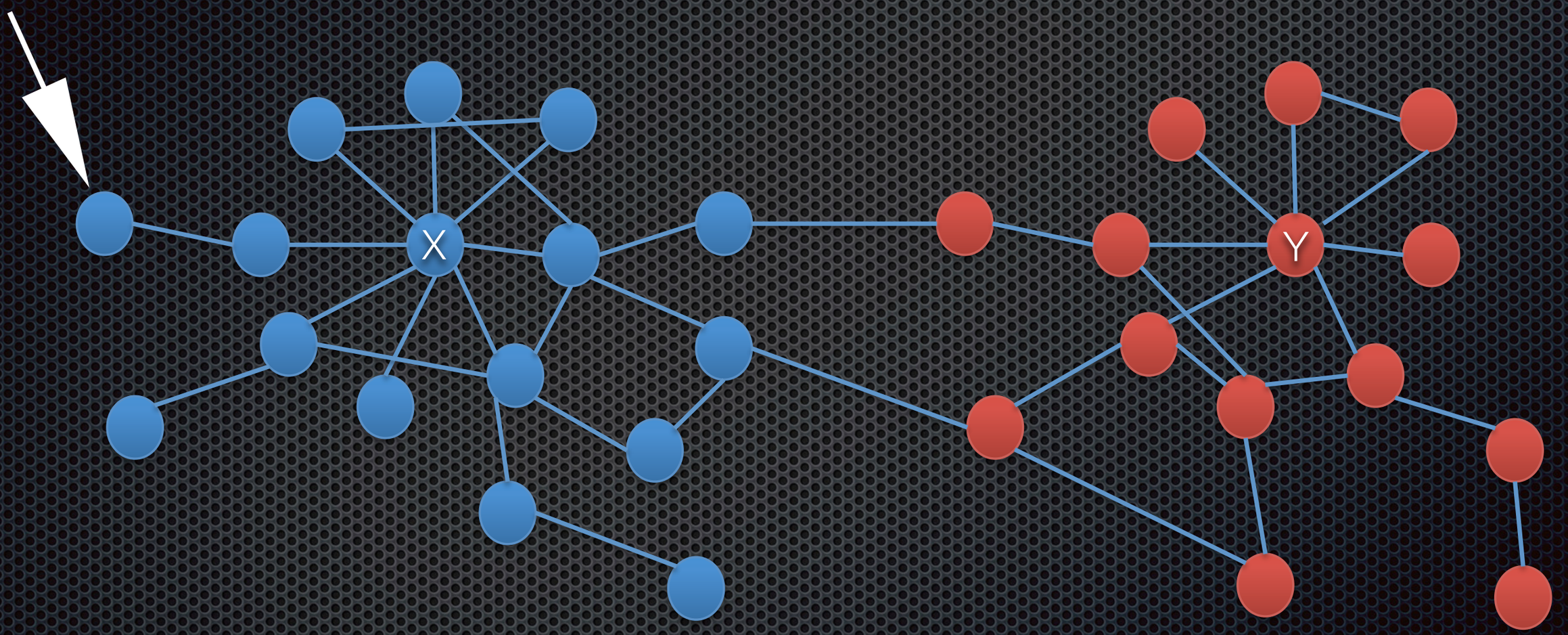
Random Walk



Random Walk



Random Walk



RWC Definition

Consider two random walks, one ending in partition X *and one ending in partition* Y , RWC *is the difference of the probabilities of two events: (i) both random walks started from the partition they ended in and (ii) both random walks started in a partition other than the one they ended in*

$$RWC = P_{XX}P_{YY} - P_{YX}P_{XY}$$

$$P_{AB} = Pr[\text{start in partition } A \mid \text{end in partition } B]$$

RWC Properties

- Probabilities are conditional on ending in either partition
  - Random walks end on either side with equal probability
  - Not skewed by size of each partition
    - Not skewed by total degree of vertices in each partition
- Close to 1 when probability of crossing sides low (high controversy)
- Close to 0 when probability of crossing comparable to that of staying (low controversy)

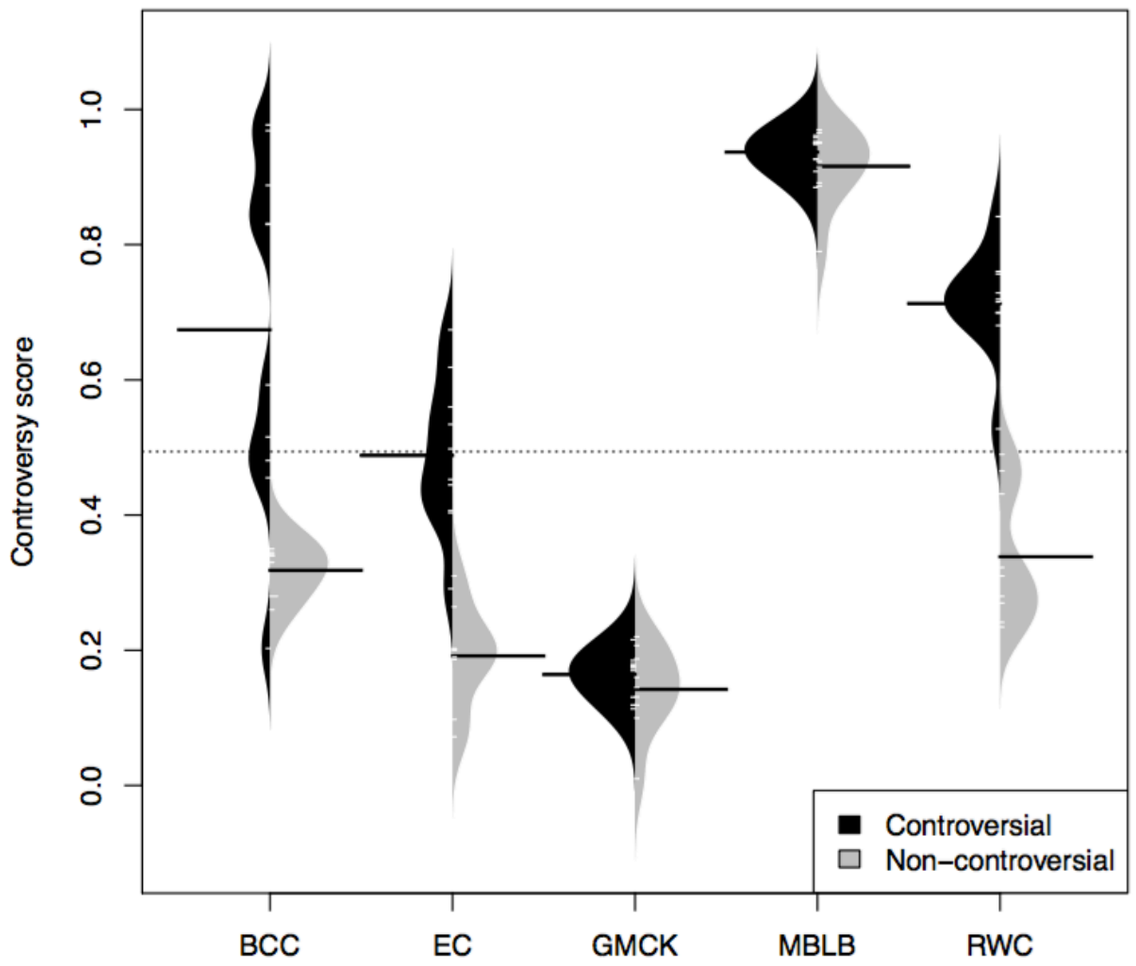


Figure 3: Controversy scores on *retweet* graphs of various controversial and non-controversial datasets

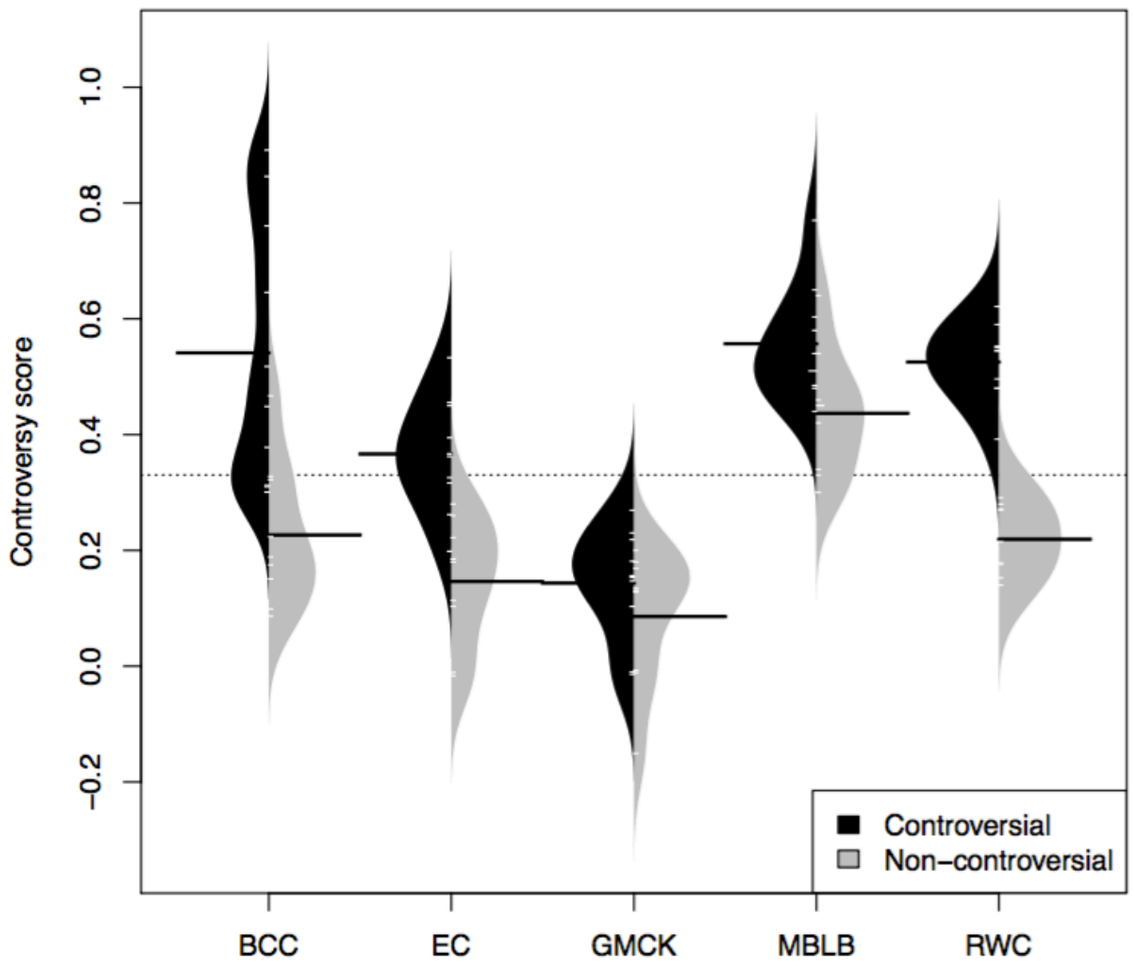
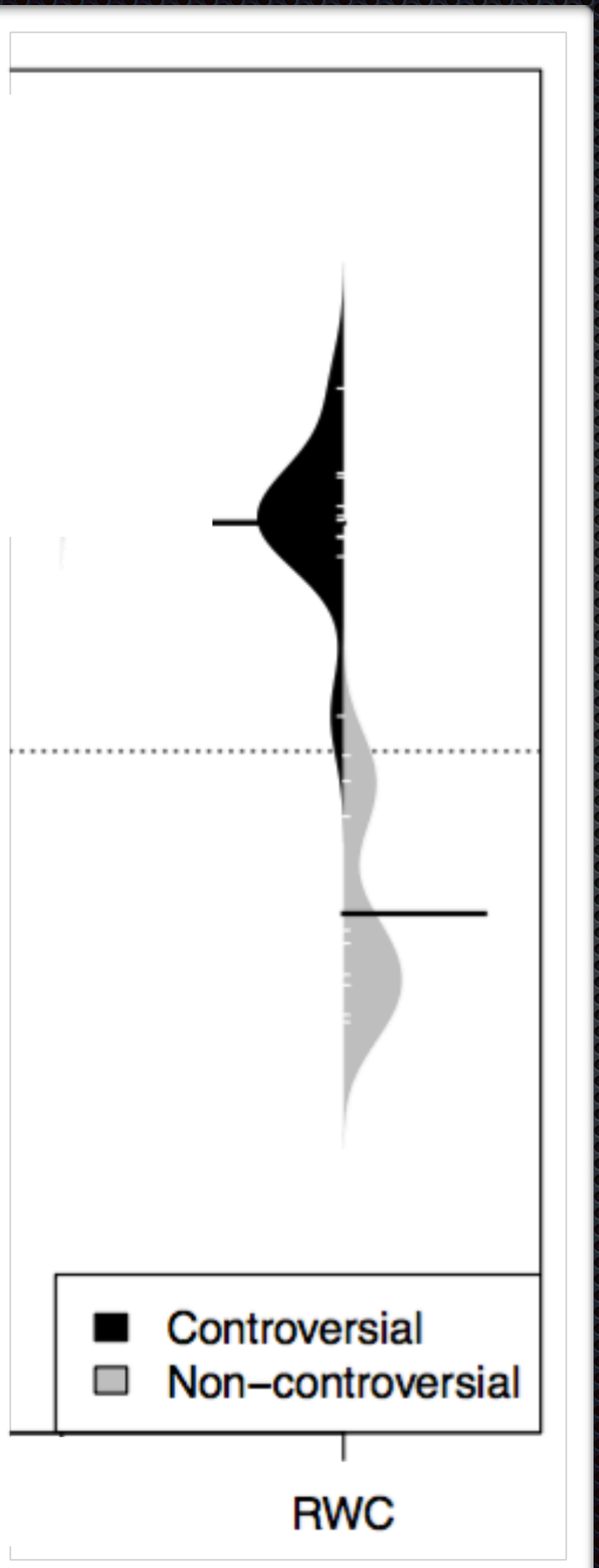
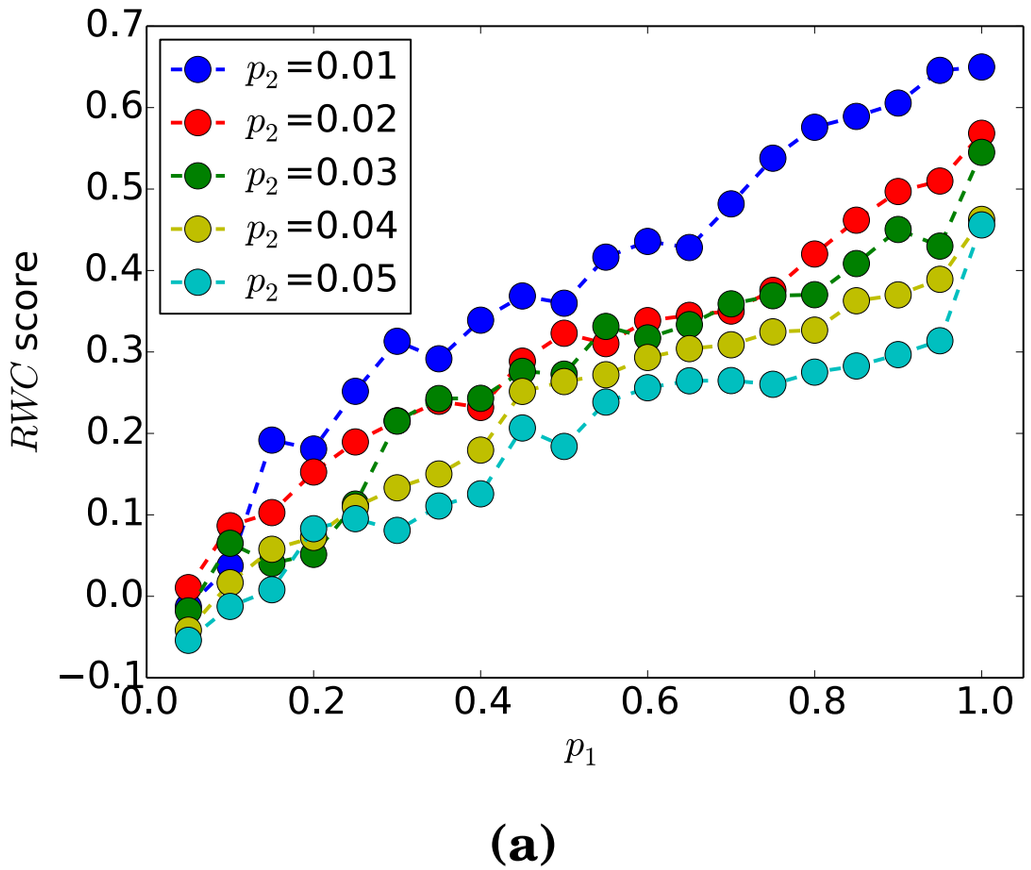


Figure 4: Controversy scores on *follow* graphs of various controversial and non-controversial datasets.

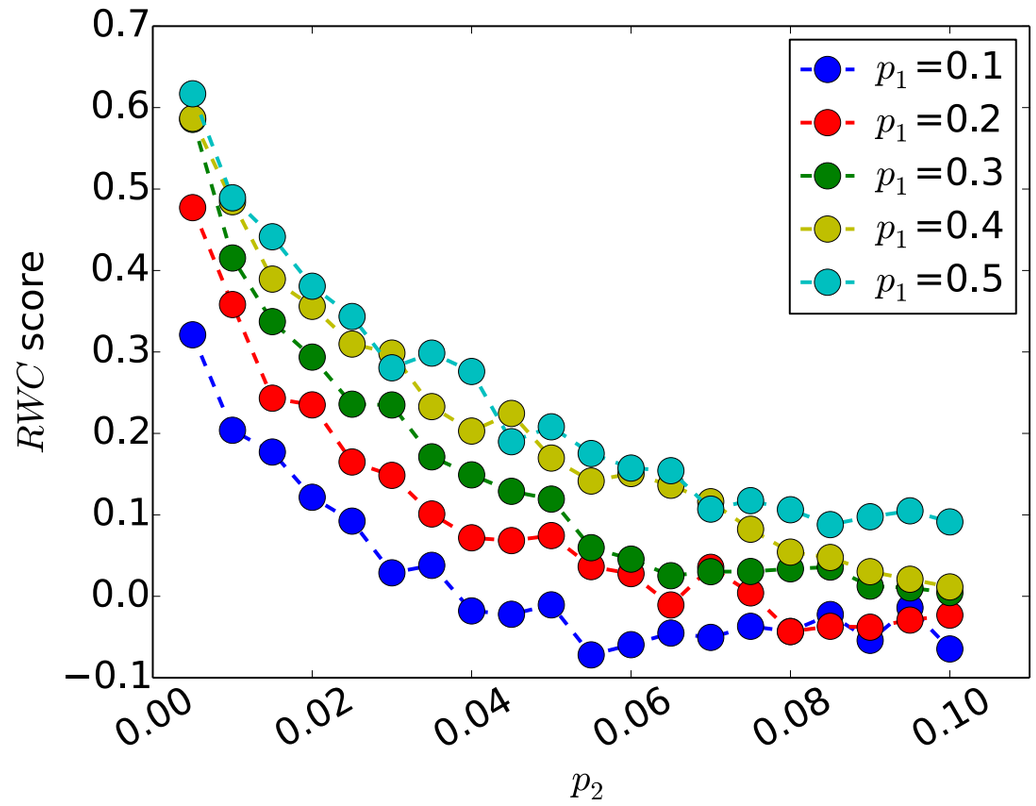
Controversy Detection

Controversy Score





(a)



(b)

Fig. 10: RWC scores for synthetic Erdös-Rényi graphs planted with two communities. p_1 is the intra-community edge probability, while p_2 is the inter-community edge probability.

Planted Synthetic Graphs

Summary

- RWC: a measure for how controversial a discussion on a topic is on social media
- Graph-based measure: no domain knowledge, language agnostic
- Intuitive semantics founded on opinion formation models
- Captures controversy better than state-of-the-art
- User-level polarization measure easy to derive

The Effect of Collective Attention on Controversial Debates on Social media

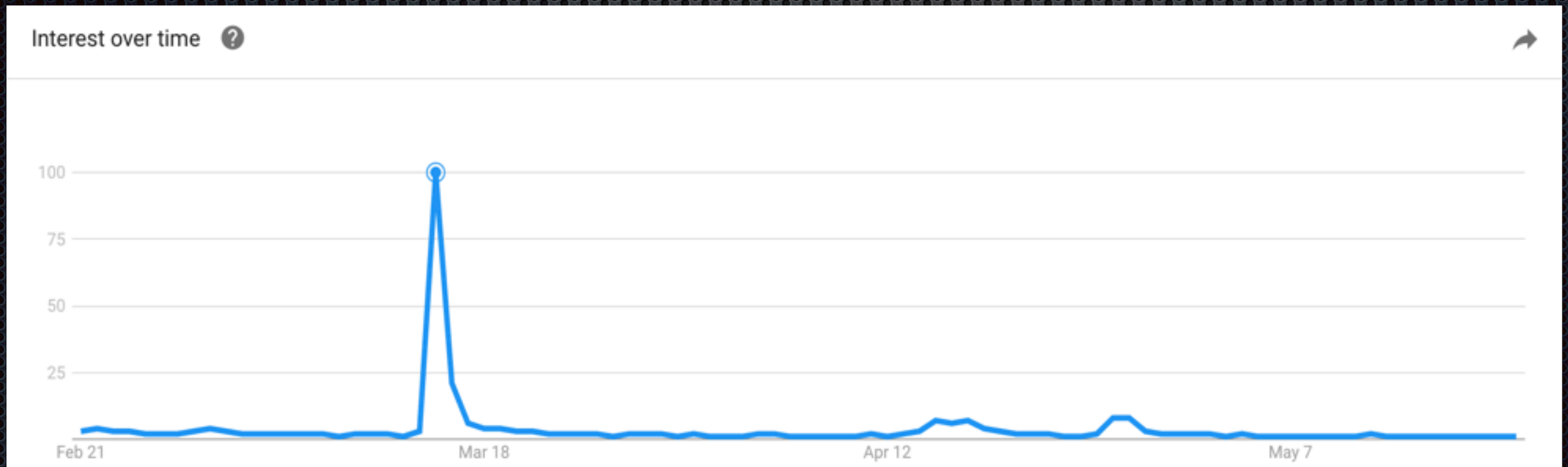
WebSci 2017 (Best Paper Award)

The Effect of Collective Attention on Controversial Debates on Social media

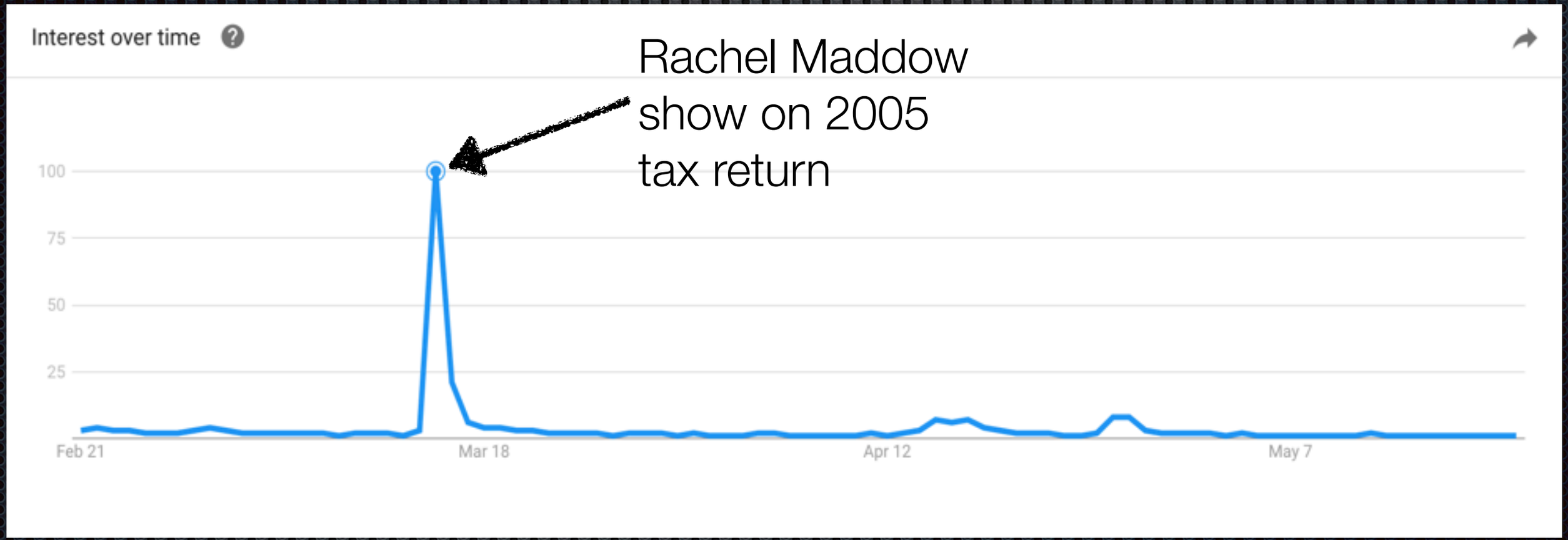
The Effect of Collective Attention on Controversial Debates on Social media

The Effect of Collective Attention on Controversial Debates on Social media

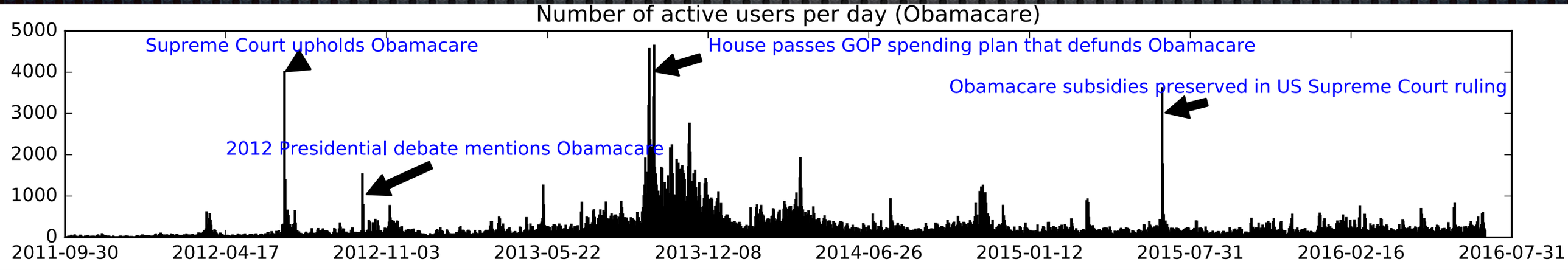
"Trump taxes" on Google



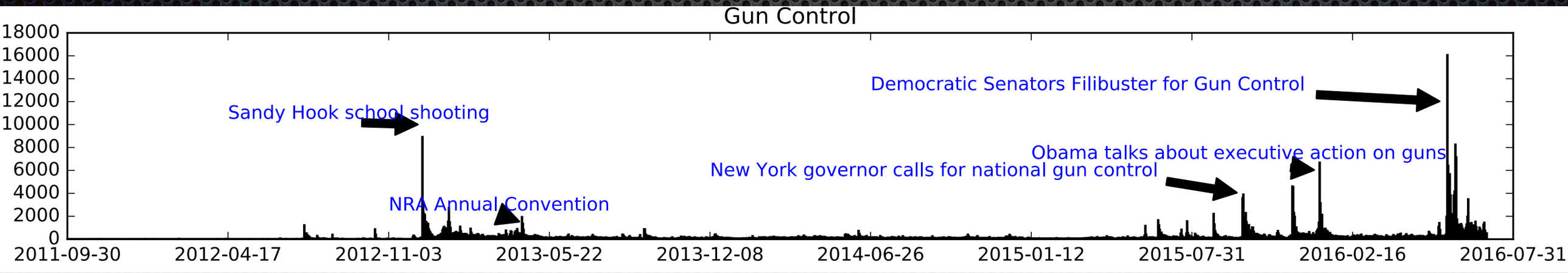
"Trump taxes" on Google



Obamacare on Twitter



Gun Control on Twitter



Literature so far

- Controversial debates examined in isolation
- As static snapshots

Contribution

- Controversial debates are dynamic
- They change with collective attention

- Analyze controversial debates over time
- Particularly when collective attention increases
  - When external ‘event’ happens

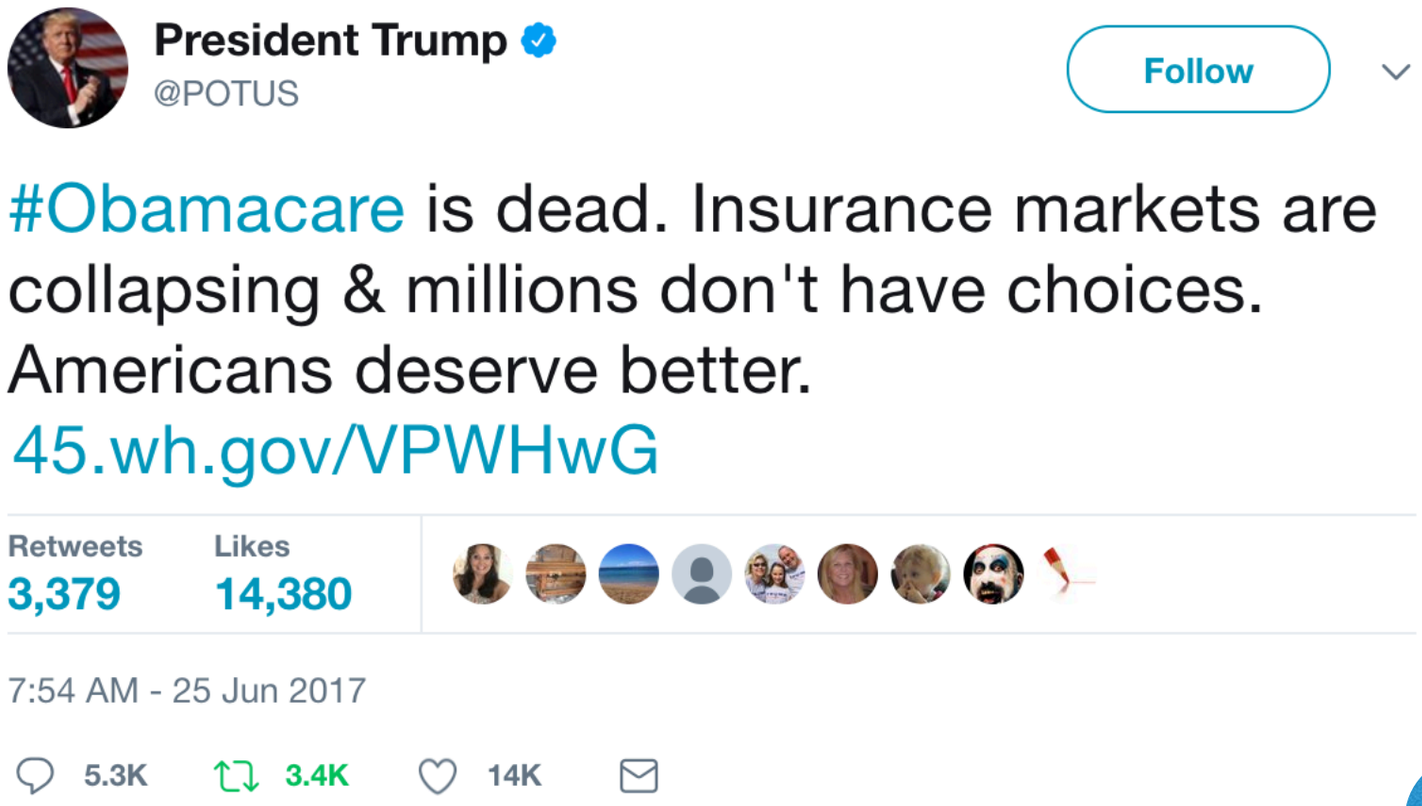
Data

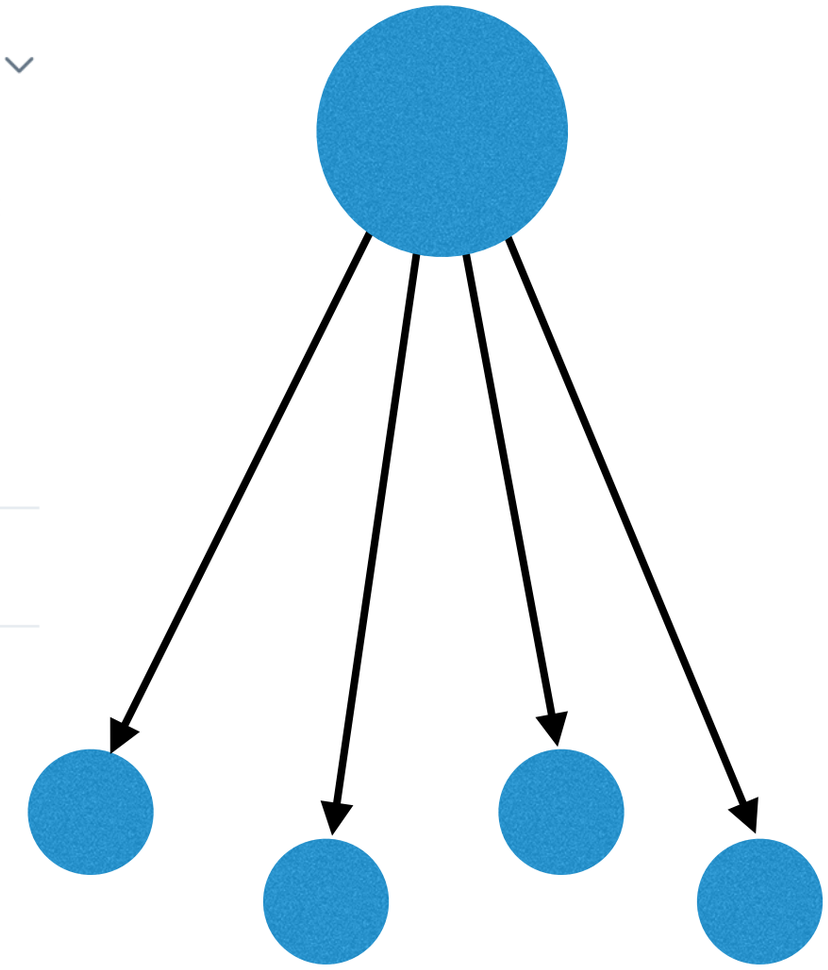
- Twitter
- 4 longitudinal polarized topics
  - Obamacare, Abortion, Gun control, Fracking
- 5 years (2011 -- 2016)
- Hundreds of thousands of users
- Millions of tweets

Definitions

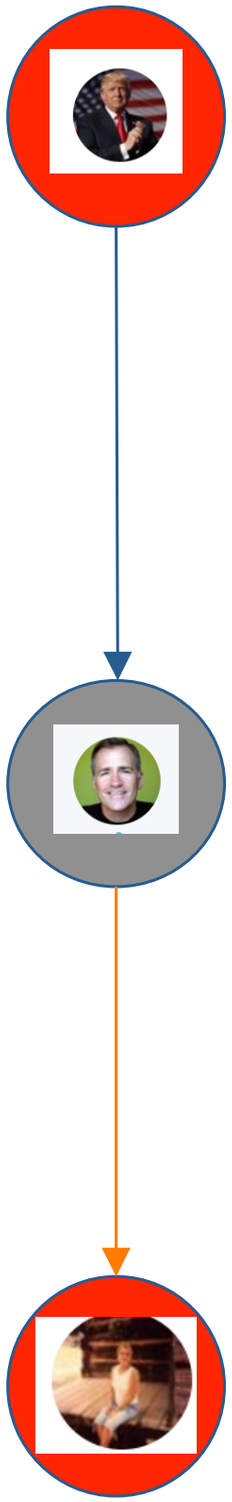
- Retweet Graph
- Reply Graph
- Core Users

Retweet Graph





Reply graph



President Trump ✓
@POTUS

Follow

#Obamacare is dead. Insurance markets are collapsing & millions don't have choices. Americans deserve better.
45.wh.gov/VPWHwG

Retweets	Likes
3,379	14,380

7:54 AM - 25 Jun 2017

5.3K 3.4K 14K

Tweet your reply

Mike Elgan ✓ @MikeElgan · 8h

Replying to @POTUS

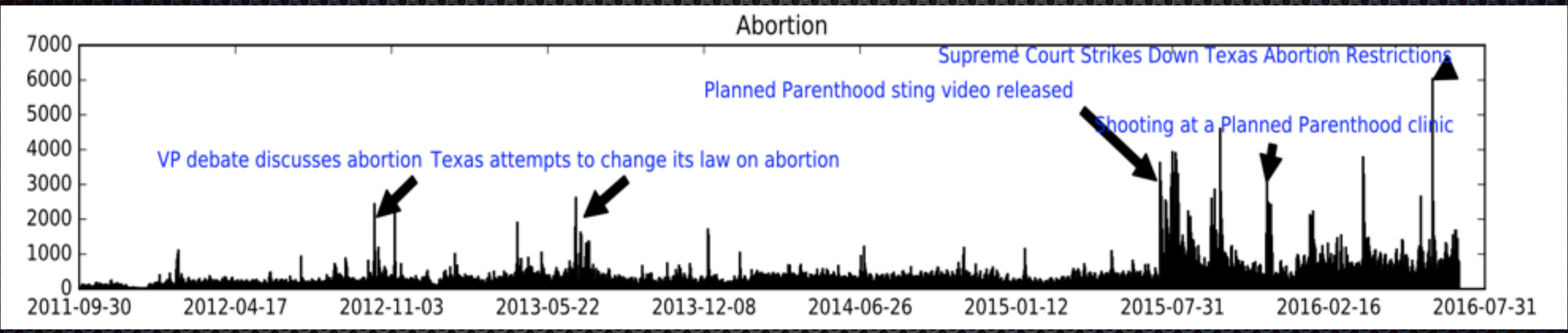
How could anyone be stupid enough to believe someone who never stops lying

Debra Martin @Soaper55 · 20h

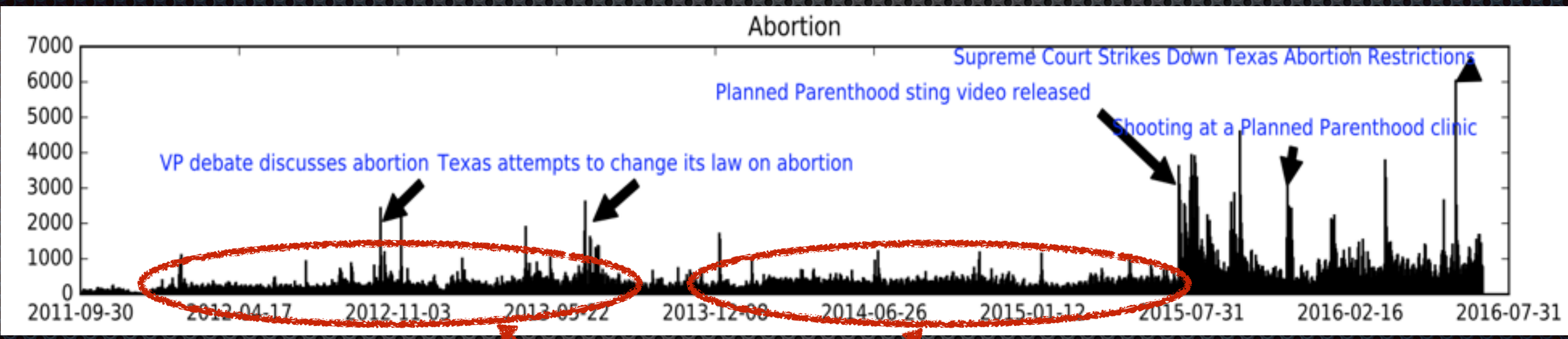
Maybe because he doesn't lie! You all really be trying to help America and Americans instead of holding us back into the death spiral.

204 5 151

Core

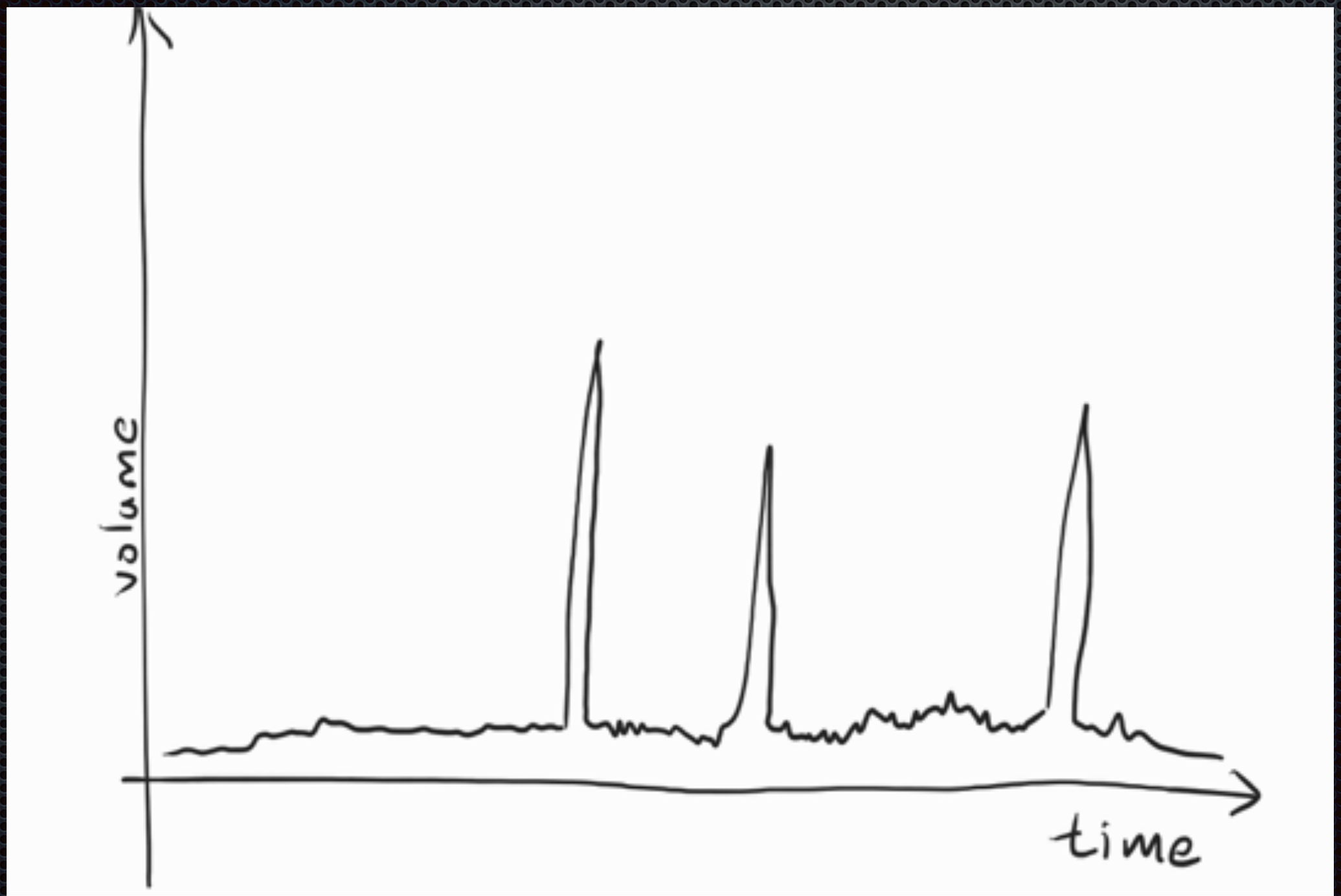

Abortion
Supreme Court Strikes Down Texas Abortion Restrictions
Planned Parenthood sting video released
Shooting at a Planned Parenthood clinic
VP debate discusses abortion
Texas attempts to change its law on abortion
7000
6000
5000
4000
3000
2000
1000
0
2011-09-30
2012-04-17
2012-11-03
2013-05-22
2013-12-08
2014-06-26
2015-01-12
2015-07-31
2016-02-16
2016-07-31

Core

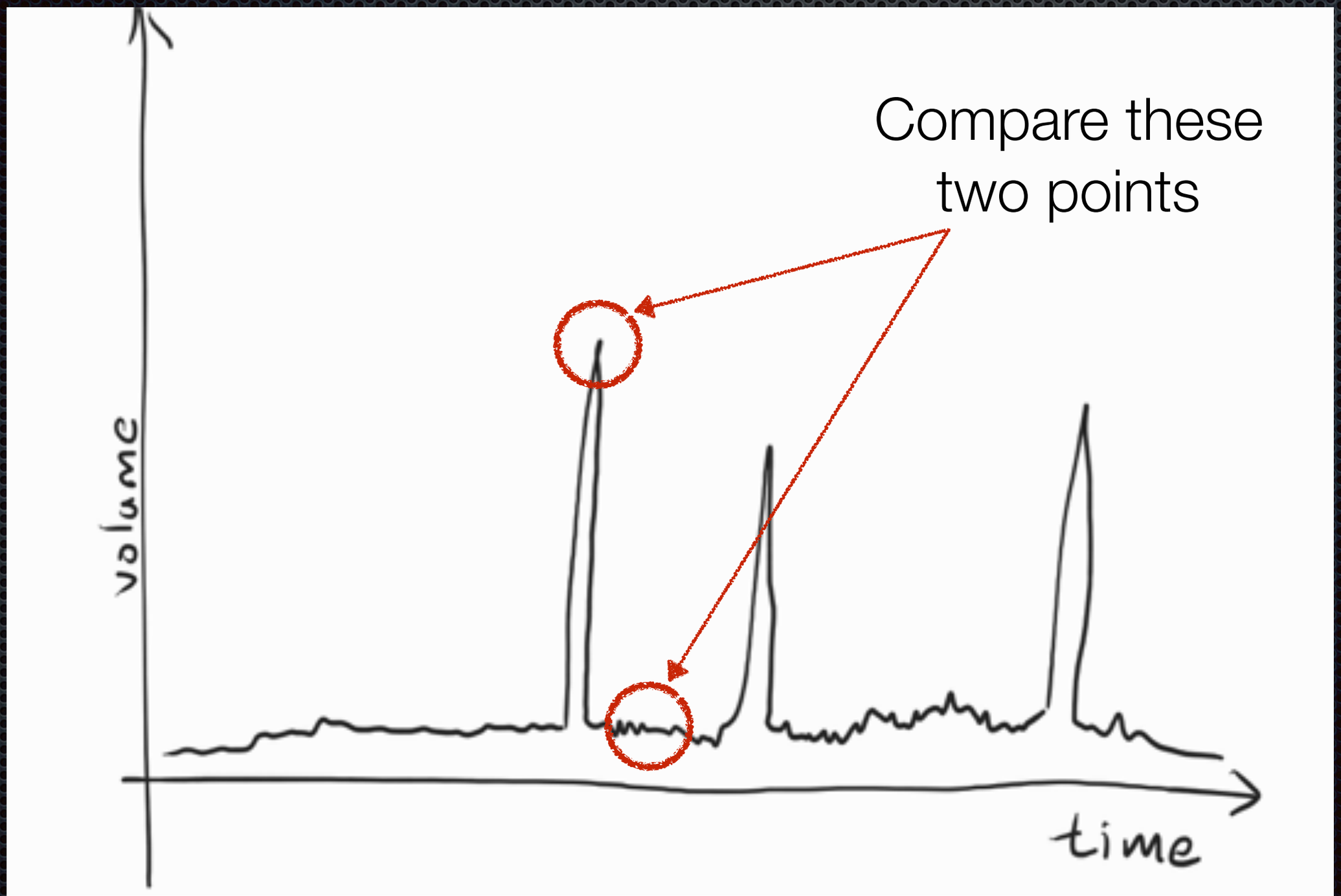


Core Users

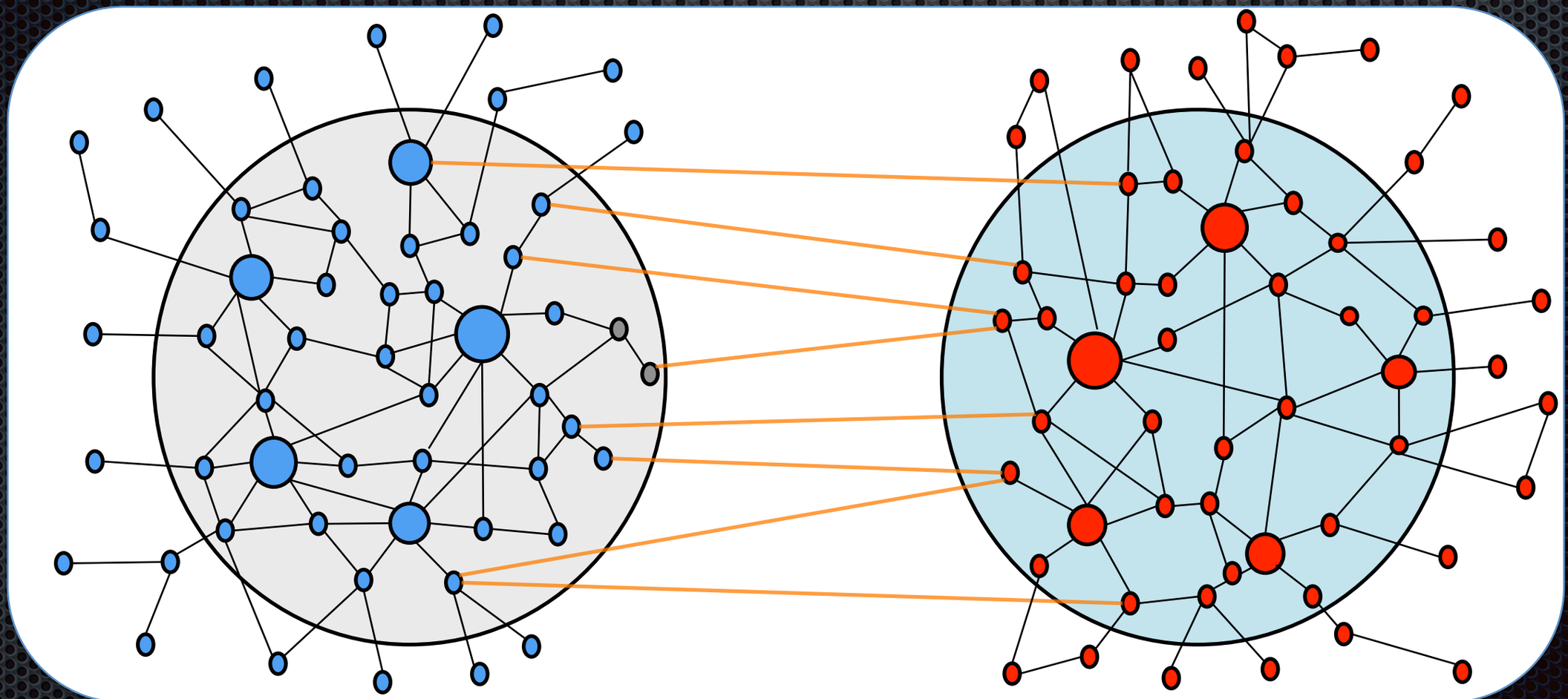
Experiments



Experiments

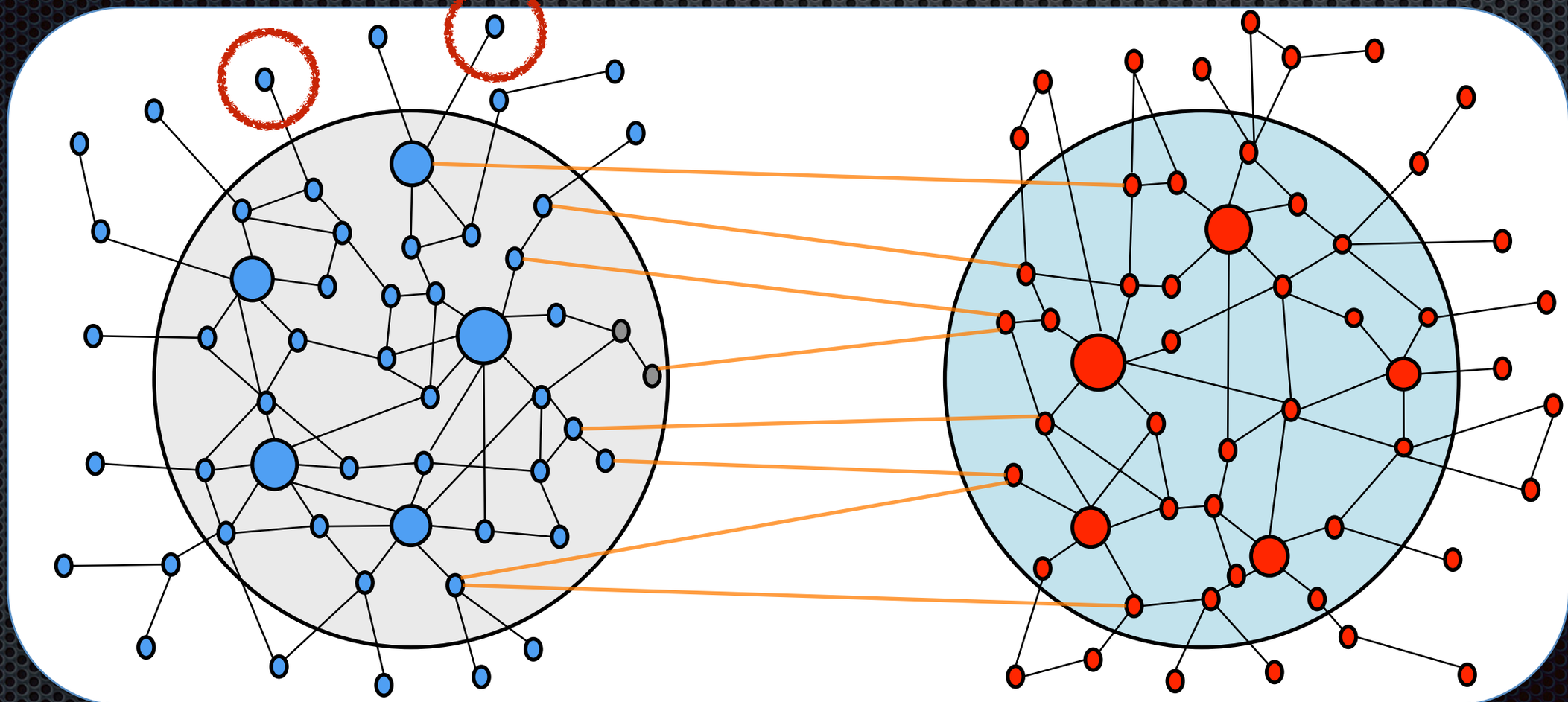


Retweet Graph



Retweet Graph

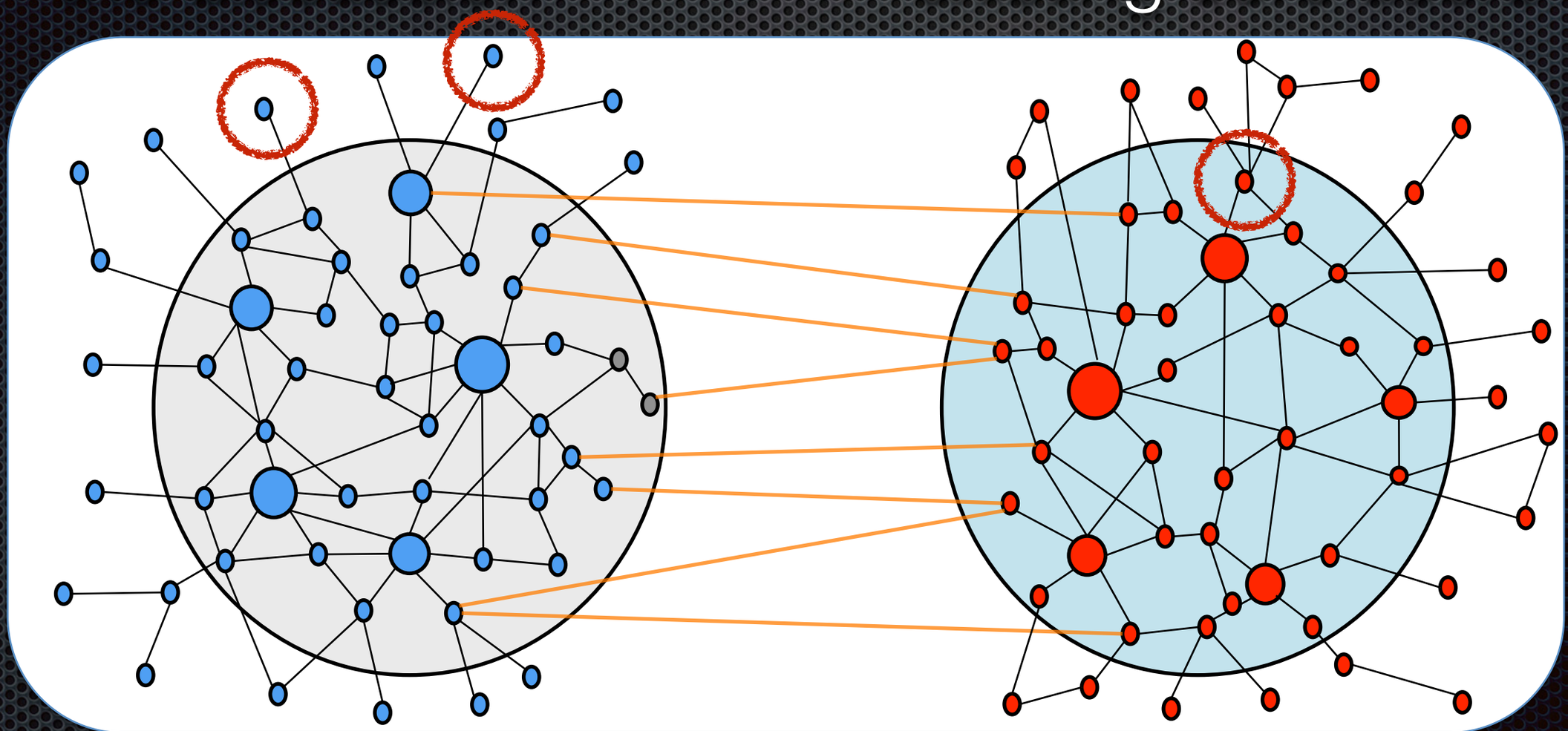
1) New users enter the discussion



Retweet Graph

1) New users enter the discussion

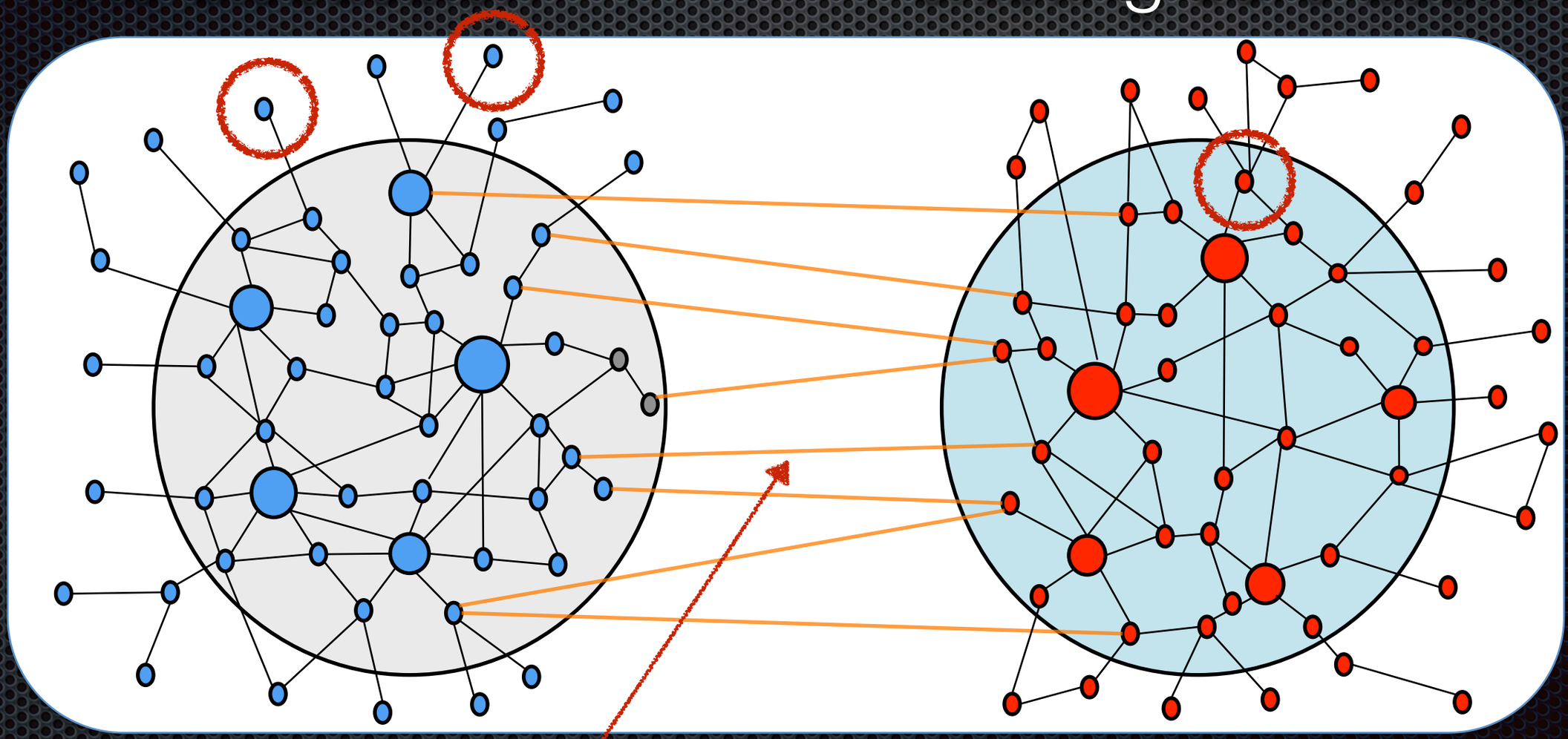
2) Most retweets to existing core users



Retweet Graph

1) New users enter the discussion

2) Most retweets to existing core users

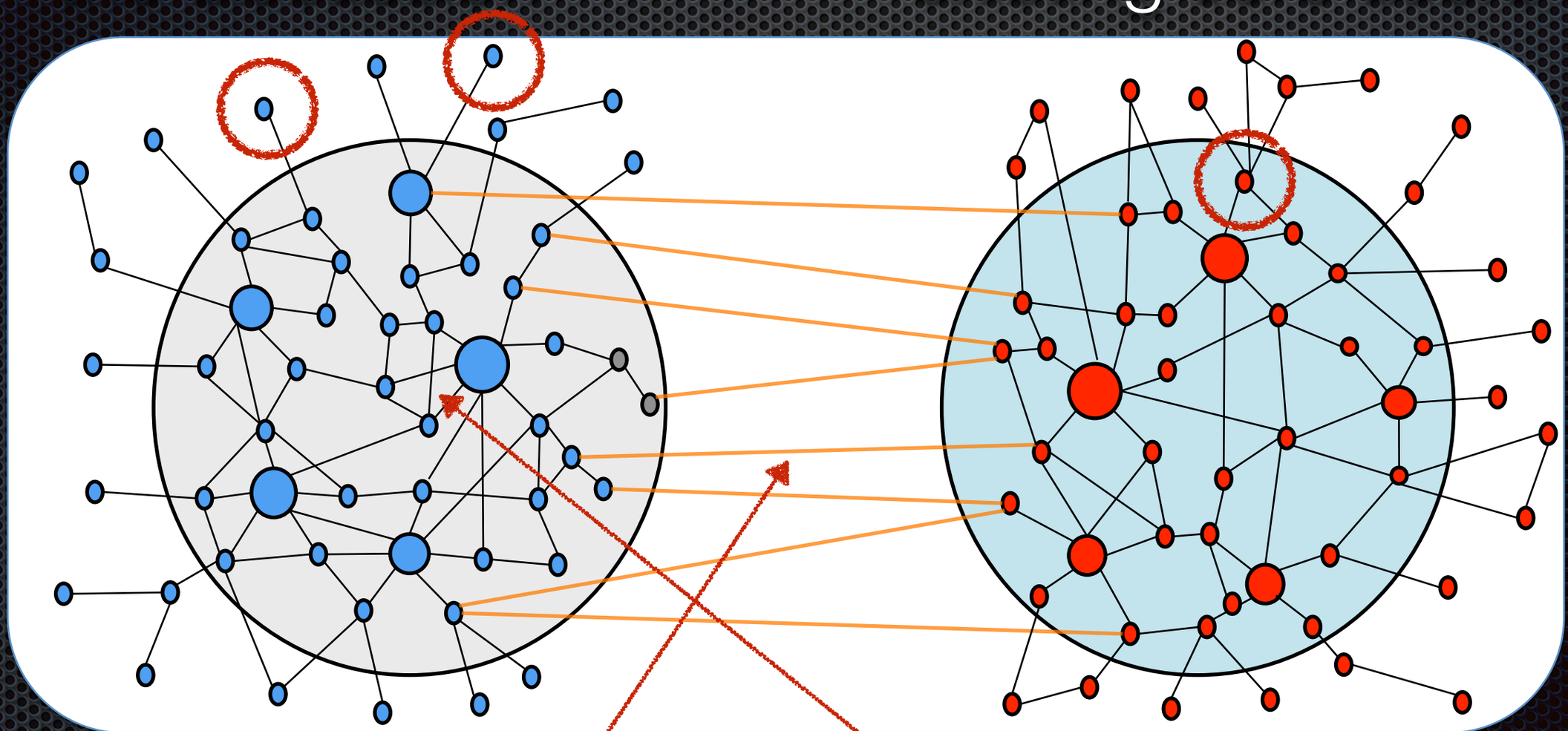


3) Cross-side retweets decrease

Retweet Graph

1) New users enter the discussion

2) Most retweets to existing core users



3) Cross-side retweets decrease

4) Within-side retweets increase

Controversy Measure

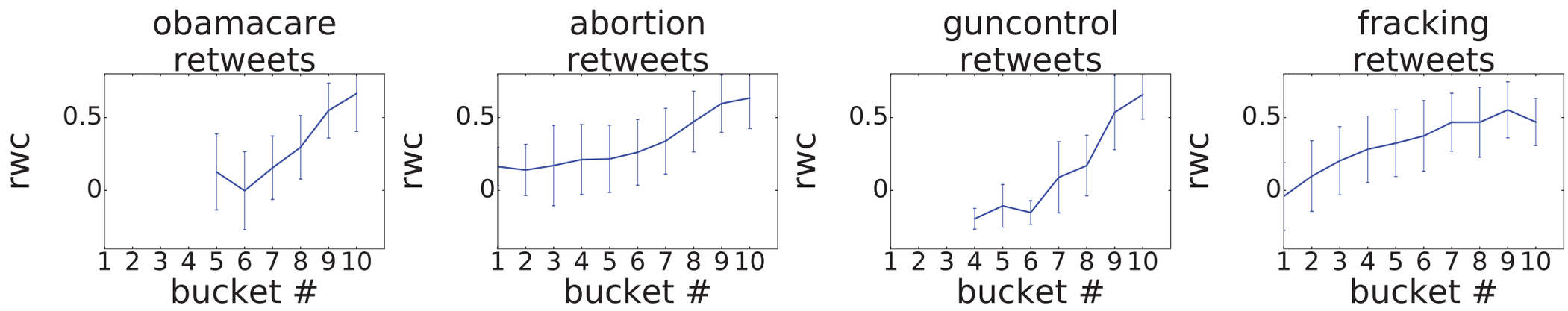
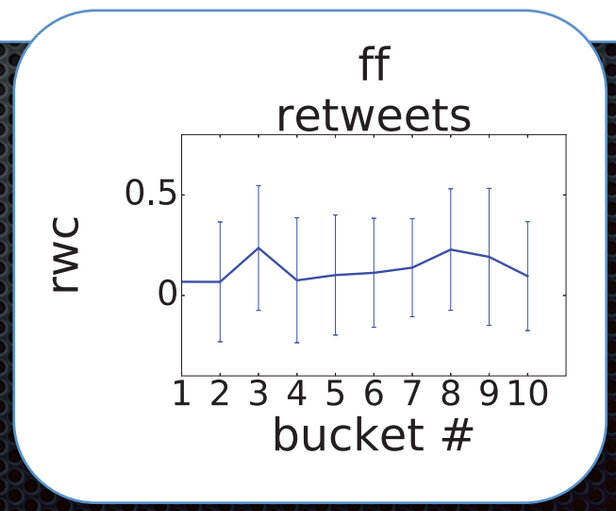


Figure 2: RWC score as a function of the activity in the retweet network. An increase in interest in the controversial topic corresponds to an increase in the controversy score of the retweet network.



Core-Periphery Openness

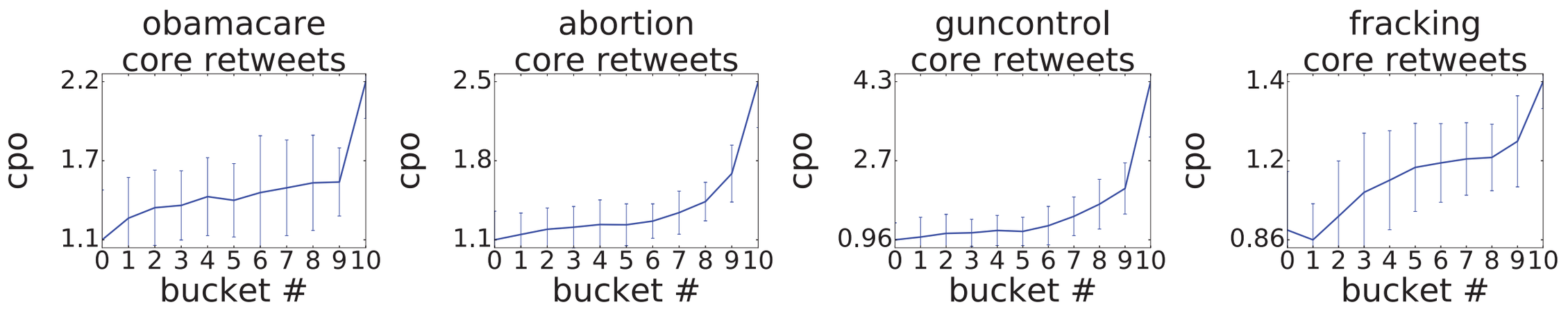
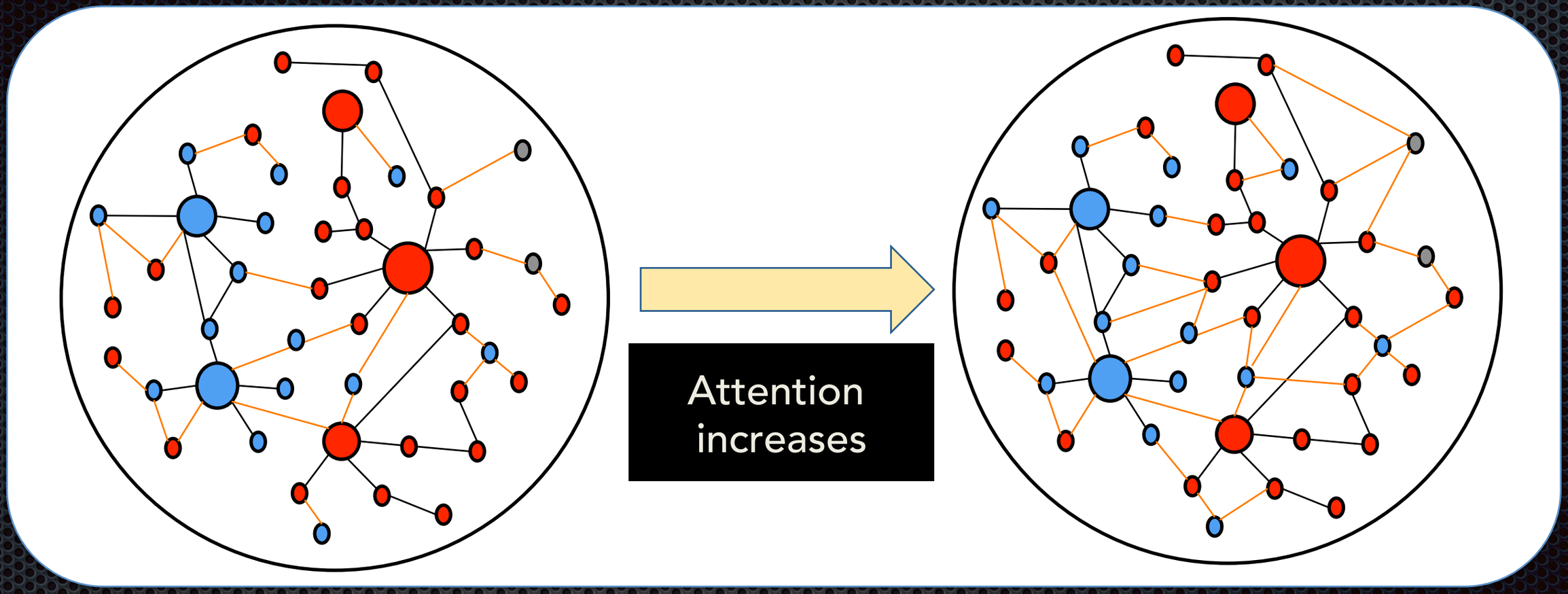


Figure 12: Core–periphery openness as a function of activity in the retweet network. As the interest increases, the number of core-periphery edges, normalized by the expected number of edges in a random network, increases. This suggests a propensity of periphery nodes to connect with the core nodes when interest increases.

Reply Graph



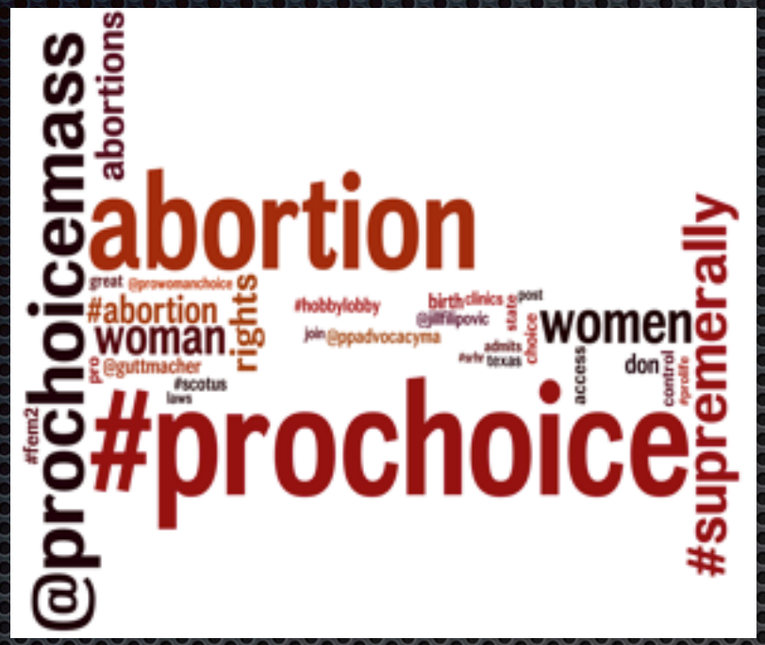
- Cross-side edges increase: more discussion

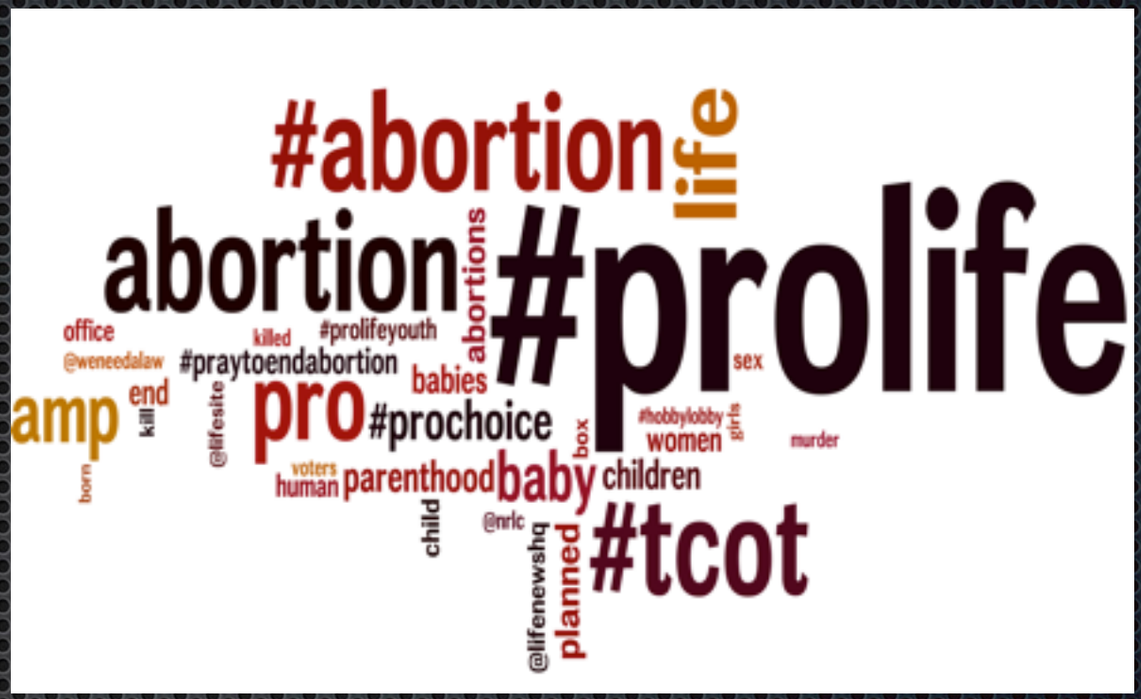
Content

Pro Choice

Pro Life

Normal Condition





Attention Increase

Content

Pro Choice

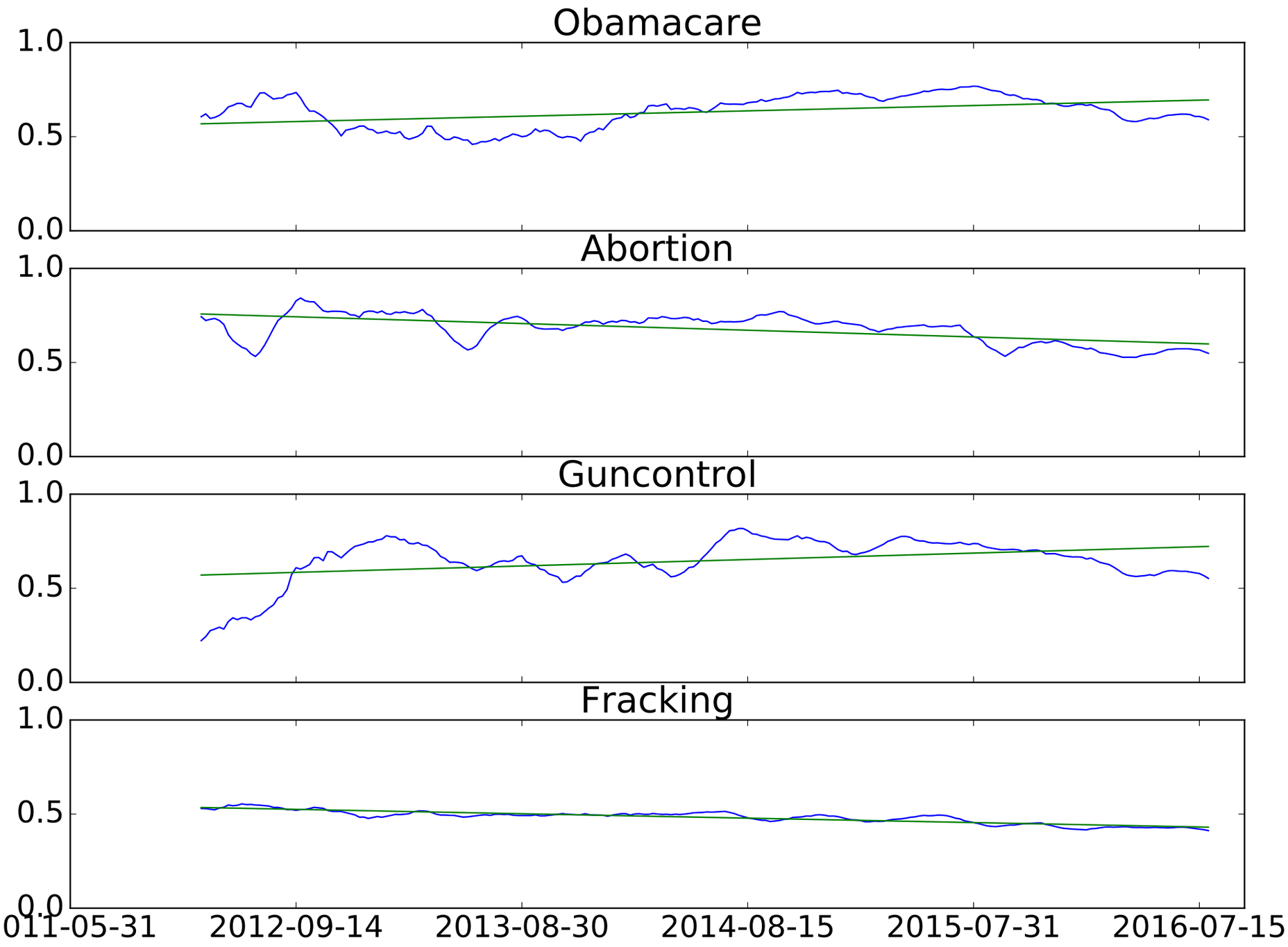
Pro Life

Normal Condition

Content becomes uniform across the sides

Attention Increase

Long-Term Polarization



Summary

- Controversial debates during external events
- Polarization increases
- Retweet graph becomes hierarchical (core-periphery)
- More replies across sides
- Content becomes more uniform
- Many more results in the paper!

Political Discourse on Social Media

Echo Chambers, Gatekeepers, and the Price of Bipartisanship

WWW 2018

Political Discourse on Social Media

- Characterized by heavy polarization
- Emergence of ***echo chambers*** ("Hear your own voice")
- Might hamper deliberative process in democracy
  - Lack of shared world view
- Concern expressed by former US Presidents, Facebook, Twitter, and more

Polarization Cause

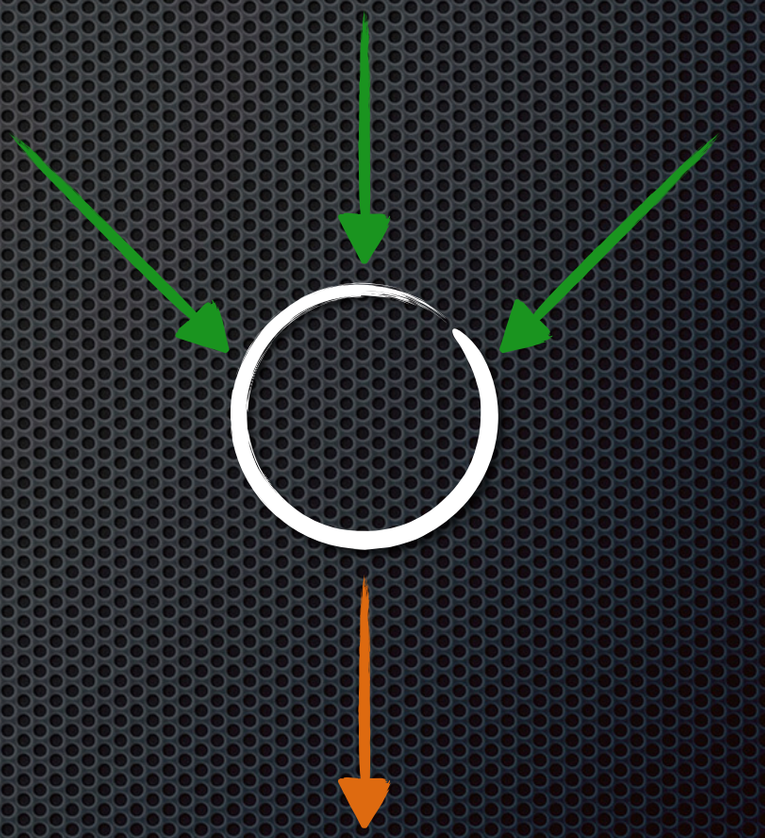
- Selective exposure?
  - People see only content that agrees with their pre-existing opinion
- Biased assimilation?
  - People pay more attention to content that agrees with their pre-existing opinion

Echo Chamber Definition

- Echo = opinion
- Chamber = network
- Joint content + network definition
- Echo chamber = political leaning of content that users receive from network agrees with that of content they share to the network

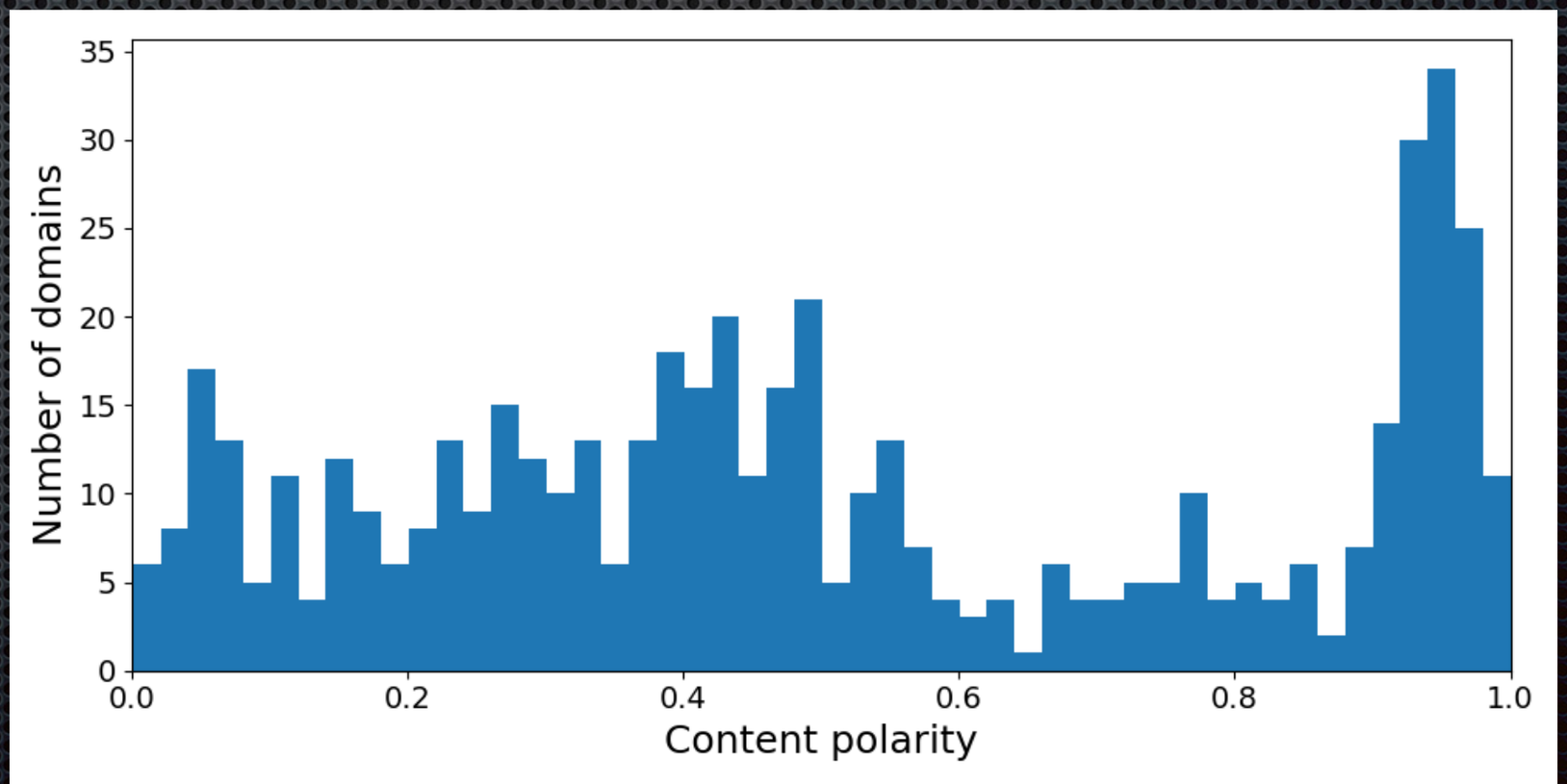
Production/Consumption

- Consumption
  - What you receive in your feed
  - What your followees tweet
- Production
  - What you tweet



Political Leaning Scores

- Based on source of the content (500 domains)
- Score derived by self-declared affiliation of sharers on FB
- FoxNews.com is aligned with conservatives (CP = 0.9),
HuffingtonPost.com is aligned with liberals (CP = 0.17)



Production/Consumption Scores

- Polarity scores based on “content” leaning (from source)
- Production score
  - Average political leaning of the content the user tweets
- Consumption score
  - Average political leaning of the content the user receives on their feed
  - Results of selection by the user

δ-partisanship

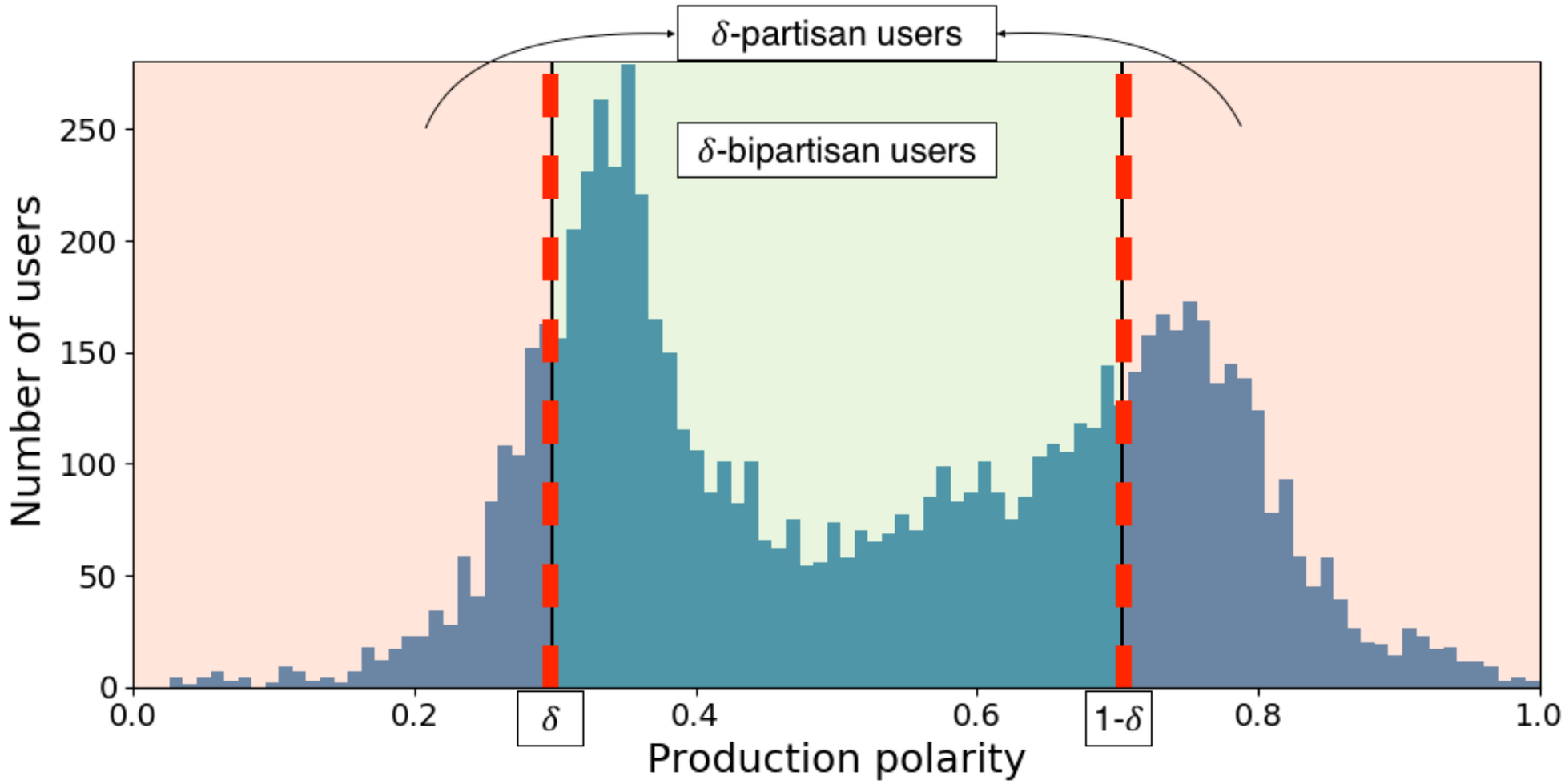


Figure 1: Example showing the definition of δ-partisan users. The dotted red lines are drawn at δ and 1-δ. Users on the left of the leftmost dashed red line or right of the rightmost one are δ-partisan.

δ-{partisan,consumer,gatekeeper}

- δ-partisan: produces content with polarity beyond δ
  - δ-bipartisan: produces content with polarity within δ
- δ-consumer: consumes content with polarity beyond δ
- δ-gatekeeper: δ-partisan but not δ-consumer
  - consumes from both sides but produces content aligned with only one side
  - blocks information flow towards its community

Network Measures

- Network-based latent-space user polarity
  - Based on following politicians with aligned ideology
- Network centrality (PageRank)
- Local clustering coefficient
- Retweet/favorite rates and volumes

Correlation

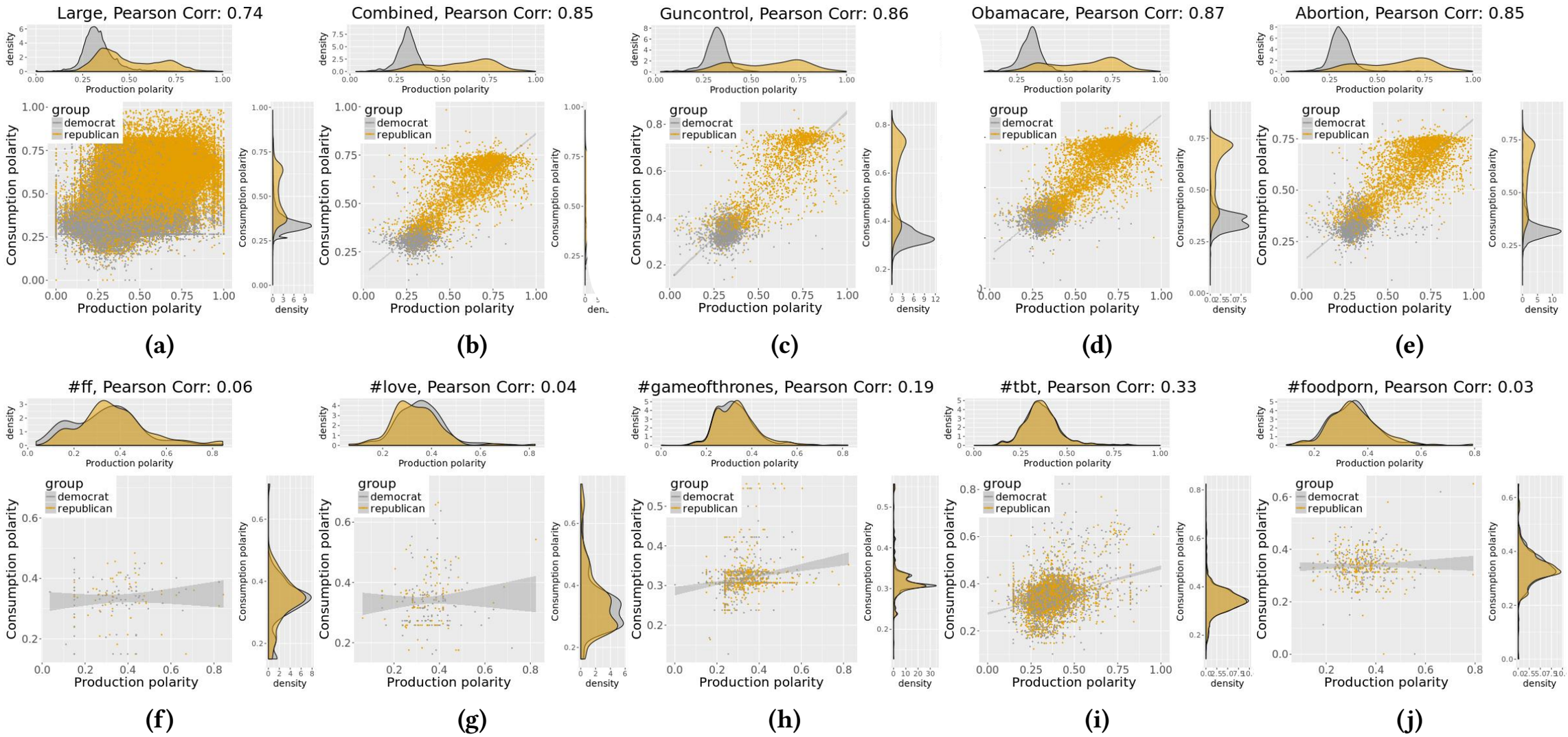
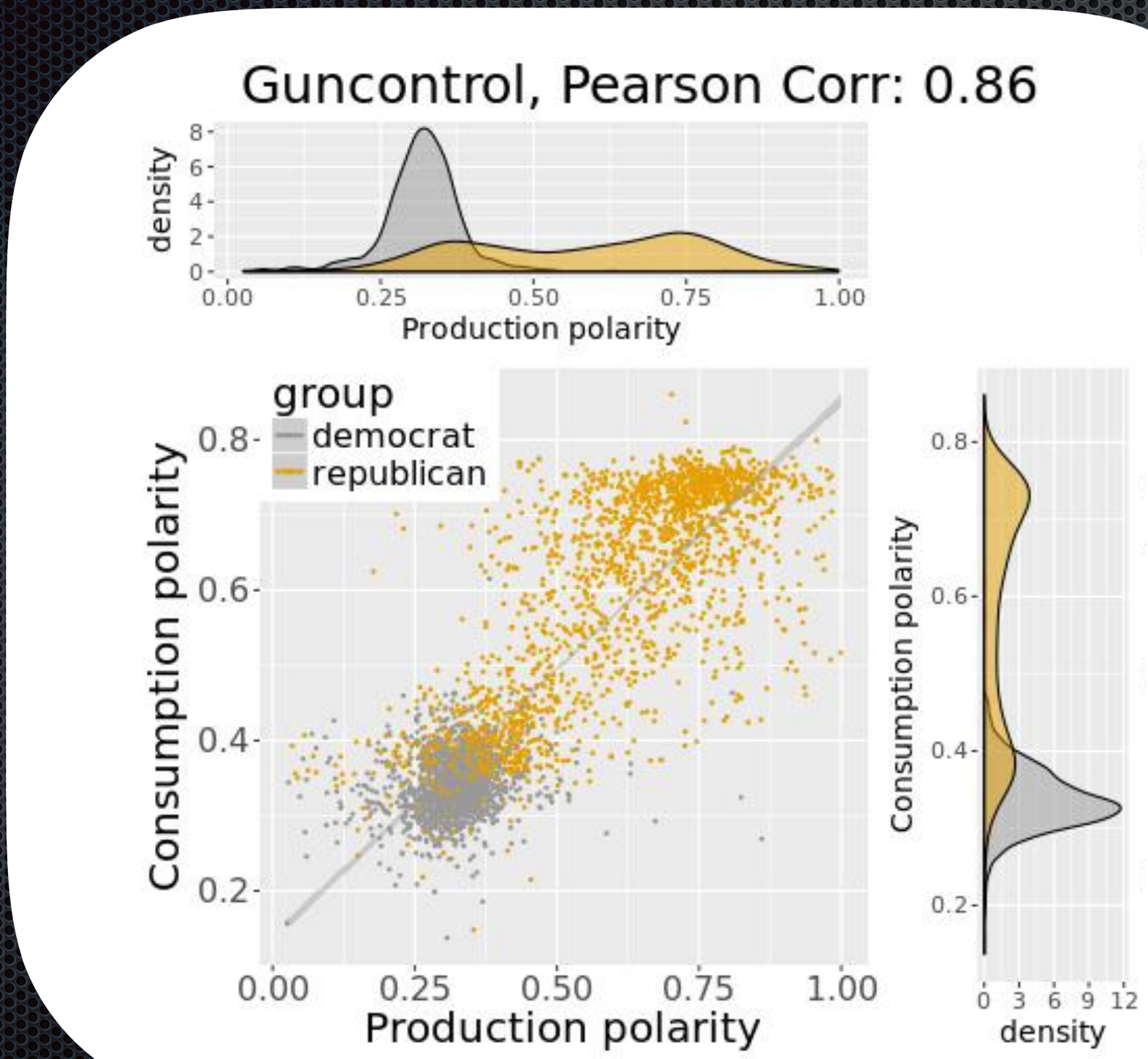


Figure 3: Distribution of production and consumption polarity, for Political (first row) and Non-Political (second row) datasets. The scatter plots display the production (x-axis) and consumption (y-axis) polarities of each user in a dataset. Colors indicate user polarity sign, following [6] (grey = democrat, yellow = republican). The one-dimensional plots along the axes show the distributions of the production and consumption polarities for democrats and republicans.

Correlation: Gun Control



Variance

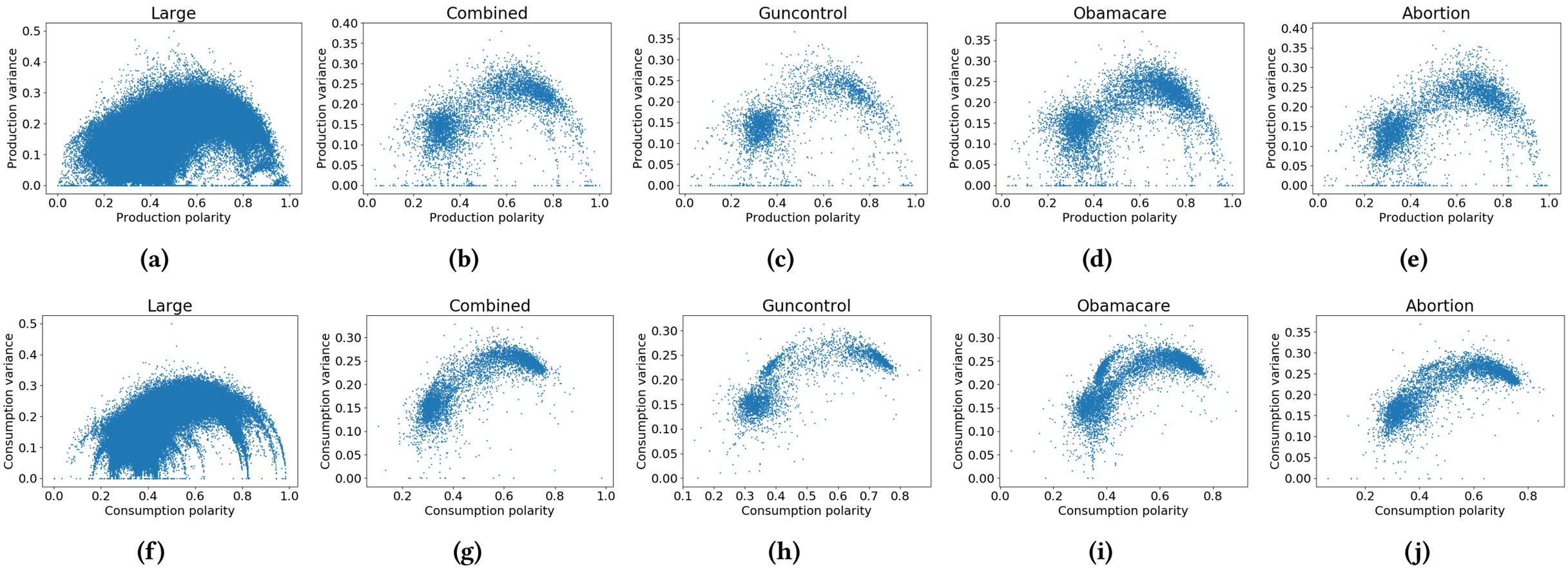
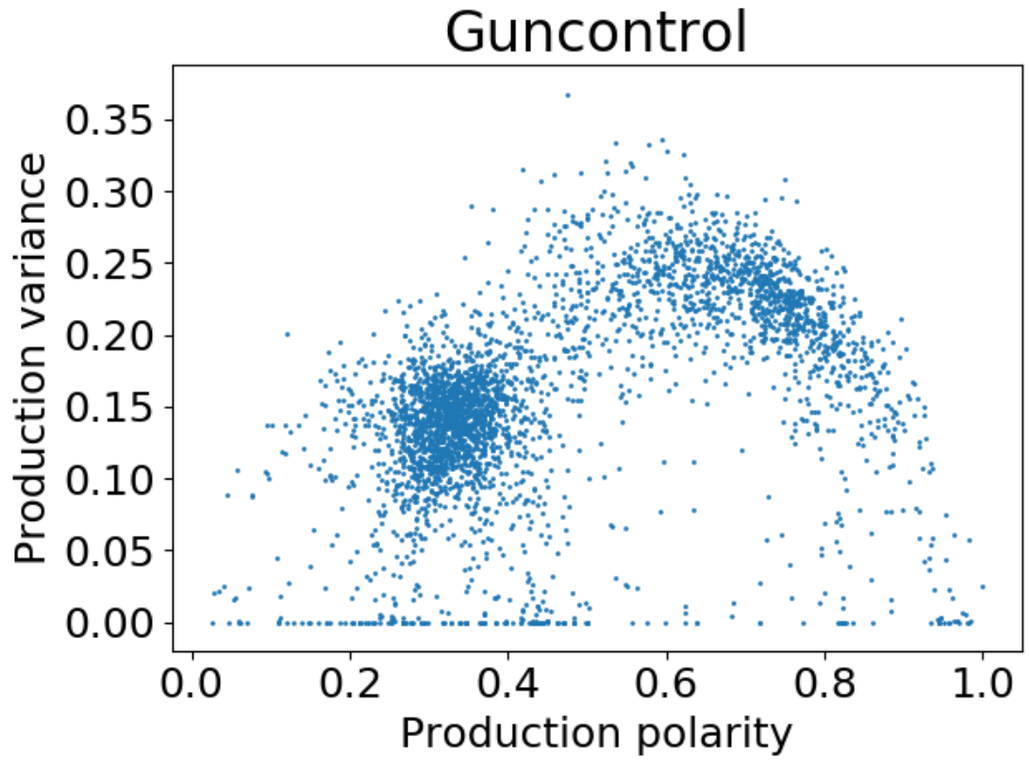
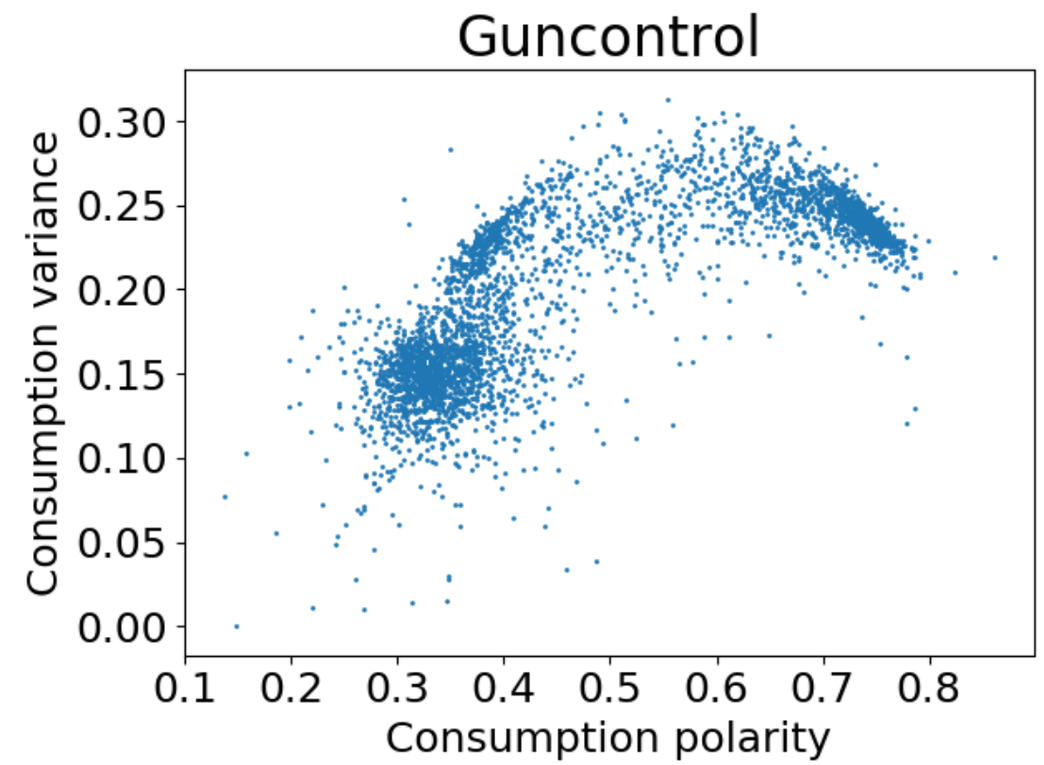


Figure 4: Top: Production polarity variance vs. production polarity (mean). Bottom: Consumption polarity variance vs. consumption polarity (mean).

Variance


Guncontrol
Production variance
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.0
0.2
0.4
0.6
0.8
1.0
Production polarity

(c)


Guncontrol
Consumption variance
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Consumption polarity

Price of Bipartisanship

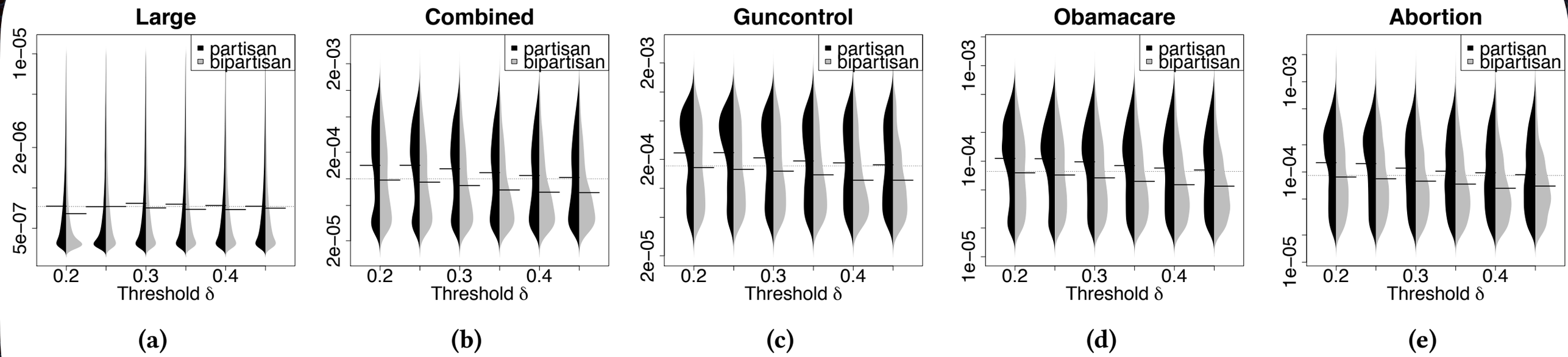
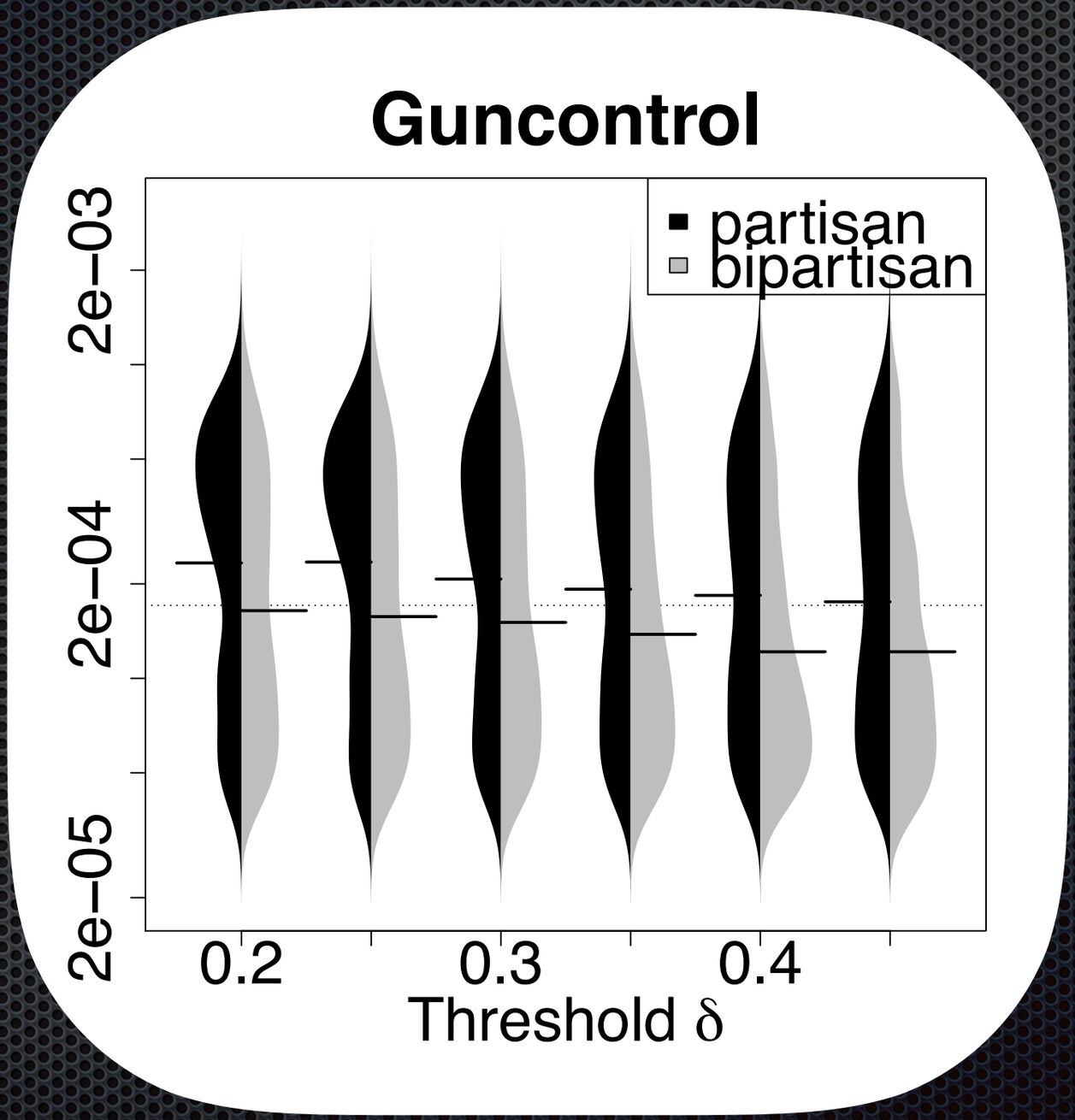


Figure 6: Pagerank for δ-partisan and δ-bipartisan users.

Price of Bipartisanship: PR



Partisans vs Bipartisans
Gatekeepers vs Non-gatekeepers

Features	Partisans	Gatekeepers
PageRank	✓	✓
clustering coefficient	✓	✓ (-)
user polarity	✓	✓ (-)
degree	✓	✓
retweet rate	✓	✗
retweet volume	✓	✗
favorite rate	✓	✗
favorite volume	✓	✗
# followers	✗	✗
# friends	✗	✗
# tweets	✗	✗
age on Twitter	✗	✗

Summary

- Find echo chambers in political discussion on Twitter
- Definition of echo chambers with two elements:
  - Content (echo) + Network (chamber)
- Data supports the selective exposure theory
- Bi-partisan users pay a price in terms of network centrality and content appreciation

Conclusions

- How do controversies unfold on social media?
- Measuring is the first step (RWC)
- Controversies are dynamic (time is an important factor)
  - Collective attention increases polarization
- Echo chambers associated with controversies
  - Evidence of selective exposure and price of bi-partisanship

What's next?

- Joint opinion formation + network generation model
- Adding data to opinion dynamics models
- Temporal dynamics of the process
- Application to other contexts (Reddit, Facebook)
- Interventions: can we do something about it?

Thanks!
Questions please!

@gdfm7

gdfm@acm.org