

How Censorship in China Allows Government Criticism but Silences Collective Expression¹

Gary King, Jennifer Pan, Molly Roberts

Institute for Quantitative Social Science
Harvard University

(talk at University of California, Berkeley, 12/3/2012)

¹Forthcoming, *American Political Science Review*

Chinese Censorship

Chinese Censorship

- The largest selective suppression of human expression in history

Chinese Censorship

- The largest selective suppression of human expression in history
- Press freedom: 187th of 197 countries, but by far the largest effort

Chinese Censorship

- The largest selective suppression of human expression in history
- Press freedom: 187th of 197 countries, but by far the largest effort
- Content filtering: Removing individual posts **by hand**

Chinese Censorship

- The largest selective suppression of human expression in history
- Press freedom: 187th of 197 countries, but by far the largest effort
- Content filtering: Removing individual posts **by hand**
- Estimated workforce involved:

Chinese Censorship

- The largest selective suppression of human expression in history
- Press freedom: 187th of 197 countries, but by far the largest effort
- Content filtering: Removing individual posts **by hand**
- Estimated workforce involved:
 - Total government censors: $\approx 200,000$

Chinese Censorship

- The largest selective suppression of human expression in history
- Press freedom: 187th of 197 countries, but by far the largest effort
- Content filtering: Removing individual posts **by hand**
- Estimated workforce involved:
 - Total government censors: $\approx 200,000$
 - Internal censors at individual sites like Sina.com: **up to $\approx 1,000$**

Chinese Censorship

- The largest selective suppression of human expression in history
- Press freedom: 187th of 197 countries, but by far the largest effort
- Content filtering: Removing individual posts **by hand**
- Estimated workforce involved:
 - Total government censors: $\approx 200,000$
 - Internal censors at individual sites like Sina.com: **up to $\approx 1,000$**
 - Internet police: **20,000–50,000**

Chinese Censorship

- The largest selective suppression of human expression in history
- Press freedom: 187th of 197 countries, but by far the largest effort
- Content filtering: Removing individual posts **by hand**
- Estimated workforce involved:
 - Total government censors: $\approx 200,000$
 - Internal censors at individual sites like Sina.com: **up to $\approx 1,000$**
 - Internet police: **20,000–50,000**
 - 50 cent party members (*wumao dang*): **250,000–300,000**

Chinese Censorship

- The largest selective suppression of human expression in history
- Press freedom: 187th of 197 countries, but by far the largest effort
- Content filtering: Removing individual posts **by hand**
- Estimated workforce involved:
 - Total government censors: $\approx 200,000$
 - Internal censors at individual sites like Sina.com: **up to $\approx 1,000$**
 - Internet police: **20,000–50,000**
 - 50 cent party members (*wumao dang*): **250,000–300,000**
- (Other examples: Iran, Saudia Arabia, Syria, Vietnam, etc.)

Our Main Point

Our Main Point

- **Prevailing Wisdom**: the goal of censorship: to stop government criticism

Our Main Point

- **Prevailing Wisdom**: the goal of censorship: to stop government criticism
- **A BBC Reporter** thought he found a contradiction in Chinese Premier Wen Jiabao's farewell speech:

Our Main Point

- **Prevailing Wisdom**: the goal of censorship: to stop government criticism
- **A BBC Reporter** thought he found a contradiction in Chinese Premier Wen Jiabao's farewell speech:

"[Premier Wen] claimed that China's leaders welcomed criticism expressed on the internet. It is 'natural for that to happen', he said.

And yet he presides over a government that routinely censors the internet — and has imprisoned many who have given voice to their criticism."

Our Main Point

- **Prevailing Wisdom**: the goal of censorship: to stop government criticism
- **A BBC Reporter** thought he found a contradiction in Chinese Premier Wen Jiabao's farewell speech:

"[Premier Wen] claimed that China's leaders welcomed criticism expressed on the internet. It is 'natural for that to happen', he said.

And yet he presides over a government that routinely censors the internet — and has imprisoned many who have given voice to their criticism."

- **We show**:

Our Main Point

- **Prevailing Wisdom:** the goal of censorship: to stop government criticism
- **A BBC Reporter** thought he found a contradiction in Chinese Premier Wen Jiabao's fairwell speech:

"[Premier Wen] claimed that China's leaders welcomed criticism expressed on the internet. It is 'natural for that to happen', he said.

And yet he presides over a government that routinely censors the internet — and has imprisoned many who have given voice to their criticism."

- **We show:**
 - There is no contradiction; Wen Jiabao was sincere

Our Main Point

- **Prevailing Wisdom:** the goal of censorship: to stop government criticism
- **A BBC Reporter** thought he found a contradiction in Chinese Premier Wen Jiabao's farewell speech:

"[Premier Wen] claimed that China's leaders welcomed criticism expressed on the internet. It is 'natural for that to happen', he said.

And yet he presides over a government that routinely censors the internet — and has imprisoned many who have given voice to their criticism."

- **We show:**
 - There is no contradiction; Wen Jiabao was sincere
 - China allows plenty of criticism in social media

Our Main Point

- **Prevailing Wisdom:** the goal of censorship: to stop government criticism
- **A BBC Reporter** thought he found a contradiction in Chinese Premier Wen Jiabao's farewell speech:

"[Premier Wen] claimed that China's leaders welcomed criticism expressed on the internet. It is 'natural for that to happen', he said.

And yet he presides over a government that routinely censors the internet — and has imprisoned many who have given voice to their criticism."

- **We show:**
 - There is no contradiction; Wen Jiabao was sincere
 - China allows plenty of criticism in social media
 - Censorship is not "leaky & imperfect"; our understanding of it is

Our Main Point

- **Prevailing Wisdom:** the goal of censorship: to stop government criticism
- **A BBC Reporter** thought he found a contradiction in Chinese Premier Wen Jiabao's farewell speech:

"[Premier Wen] claimed that China's leaders welcomed criticism expressed on the internet. It is 'natural for that to happen', he said.

And yet he presides over a government that routinely censors the internet — and has imprisoned many who have given voice to their criticism."

- **We show:**
 - There is no contradiction; Wen Jiabao was sincere
 - China allows plenty of criticism in social media
 - Censorship is not "leaky & imperfect"; our understanding of it is
 - The goal: stopping collective action, clipping social ties

Our Main Point

- **Prevailing Wisdom**: the goal of censorship: to stop government criticism
- **A BBC Reporter** thought he found a contradiction in Chinese Premier Wen Jiabao's fairwell speech:

"[Premier Wen] claimed that China's leaders welcomed criticism expressed on the internet. It is 'natural for that to happen', he said.

And yet he presides over a government that routinely censors the internet — and has imprisoned many who have given voice to their criticism."

- **We show**:
 - There is no contradiction; Wen Jiabao was sincere
 - China allows plenty of criticism in social media
 - Censorship is not "leaky & imperfect"; our understanding of it is
 - The goal: stopping collective action, clipping social ties
 - (They also censor 2 other smaller categories we'll talk about)

Our Data Source

Our Data Source



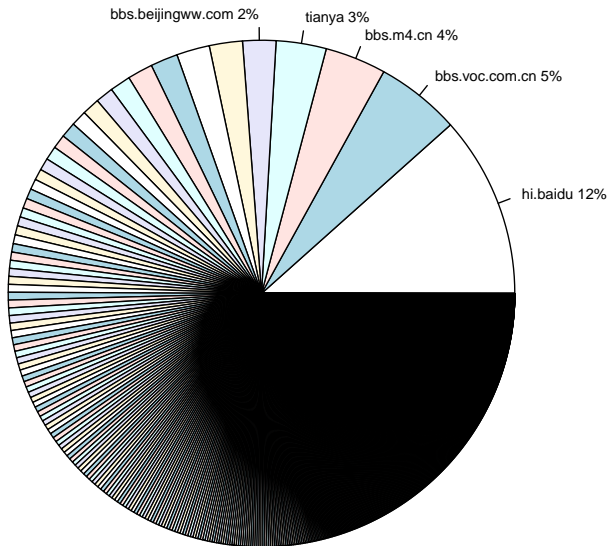
crimson hexagon

KNOW MORE. KNOW WHY. KNOW HOW.

Chinese Social Media: Fractured over 1,400+ sites



Chinese Social Media: Fractured over 1,400+ sites



(Plus 59% from Sina Blog)

Posts Collected and Analyzed

Posts Collected and Analyzed

- Collect 3,674,698 social media posts in 85 topic areas over 6 months

Posts Collected and Analyzed

- Collect 3,674,698 social media posts in 85 topic areas over 6 months
- Random sample: 127,283

Posts Collected and Analyzed

- Collect 3,674,698 social media posts in 85 topic areas over 6 months
- Random sample: 127,283
- (Repeat design; Total analyzed: 11,382,221)

Posts Collected and Analyzed

- Collect 3,674,698 social media posts in 85 topic areas over 6 months
- Random sample: 127,283
- (Repeat design; Total analyzed: 11,382,221)
- ↷ For each post (on a timeline in one of 85 content areas):

Posts Collected and Analyzed

- Collect 3,674,698 social media posts in 85 topic areas over 6 months
- Random sample: 127,283
- (Repeat design; Total analyzed: 11,382,221)
- ↪ For each post (on a timeline in one of 85 content areas):
 - Download content the instant it appears

Posts Collected and Analyzed

- Collect **3,674,698** social media posts in 85 topic areas over 6 months
- Random sample: **127,283**
- (Repeat design; Total analyzed: **11,382,221**)
- ↪ For each post (on a timeline in one of 85 content areas):
 - Download content the instant it appears
 - (Carefully) revisit each later to determine if it was censored

Posts Collected and Analyzed

- Collect **3,674,698** social media posts in 85 topic areas over 6 months
- Random sample: **127,283**
- (Repeat design; Total analyzed: **11,382,221**)
- ↪ For each post (on a timeline in one of 85 content areas):
 - Download content the instant it appears
 - (Carefully) revisit each later to determine if it was censored
 - Use computer-assisted methods of text analysis (some existing, some new, all adapted to Chinese)

Censorship is not Ambiguous: Example Error Page



The screenshot shows the top of a forum page for '金羊社区' (Golden Sheep Community) with the URL 'bbs.ycwb.com'. The page features a search bar, navigation tabs for '首页', '论坛', '博客', '家园', and '排行榜', and a search button. Below the navigation is a search input field with the placeholder text '请输入搜索内容' and a search button labeled '帖子 - 搜索'. A red arrow points to a yellow-bordered error message box that reads: '抱歉, 指定的主题不存在或已被删除或正在被审核' (Sorry, the specified topic does not exist, has been deleted, or is being investigated) and includes a link: '【[点击这里返回上一页](#)】' (Click here to return to the previous page).

Sorry, the host you were looking for does not exist, has been deleted, or is being investigated

Censorship is not Ambiguous: BBS Error Page

404 ERROR

The page you requested is temporarily down. How about you go look at another page.



你访问的页面暂时找不到了哦，去看看别的页面吧。

Jingjing, one of China's cartoon internet police

[返回首页](#) [反馈错误](#)

[关于我们](#) | [主页制作](#) | [客户服务](#) | [人才招聘](#) | [信息管理](#) | [业务联系](#) | [有奖新闻](#) | [网站地图](#)

Copyright(c) (2007-2012)New Silkroad Online. All rights reserved.

[新ICP备07500354号]互联网违法和不良信息举报中心 有害信息举报中心 互动频道举报奖励办法

版权所有：中国电信股份有限公司新疆分公司 技术支持：800-8930169 电话：0991-4667780 传真：(0991)4682953

电子信箱：edit@mail.xj.cninfo.net 文化部互联网游戏运营许可证 编号：文网文(2010)054号

增值电信业务经营许可证A2.B1.B2-20090001 网络文化经营许可证080626 广播电视节目制作经营许可证编号：(新)字第051号



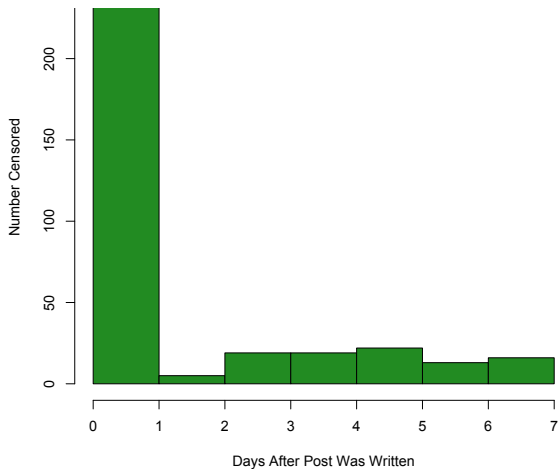
The Censors are Fast; Our Automated Methods are Faster

The Censors are Fast; Our Automated Methods are Faster

Example: Shanghai Subway Crash

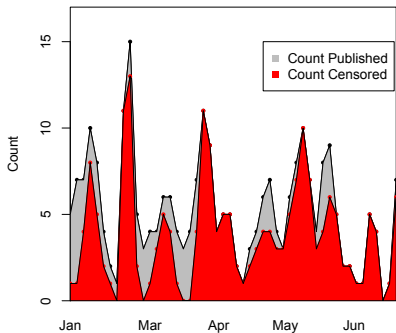
The Censors are Fast; Our Automated Methods are Faster

Example: Shanghai Subway Crash

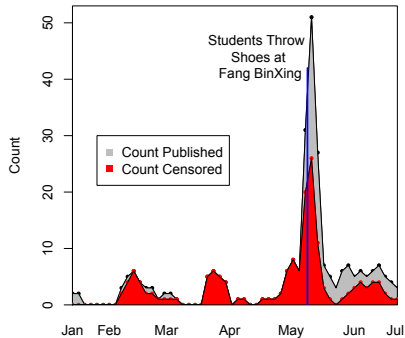


For 2 Rare Topics: Constant Censorship Effort

For 2 Rare Topics: Constant Censorship Effort



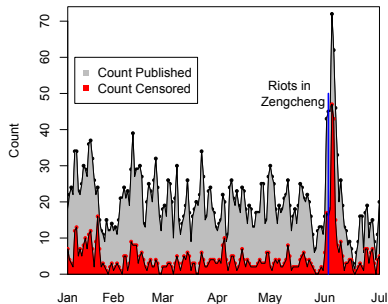
Pornography



Criticism of the Censors

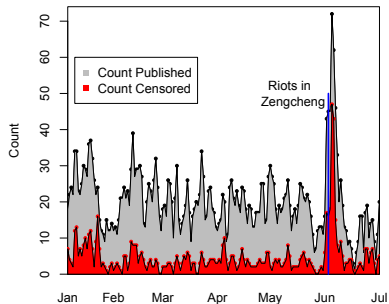
All other topics: Censorship & Post Volume are “Bursty”

All other topics: Censorship & Post Volume are “Bursty”

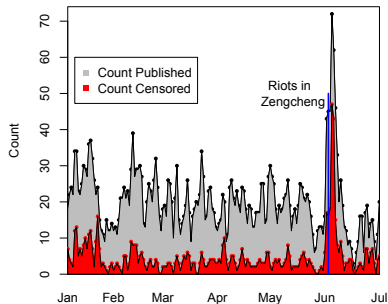


All other topics: Censorship & Post Volume are “Bursty”

- Unit of analysis:

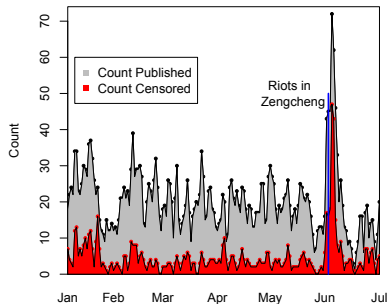


All other topics: Censorship & Post Volume are “Bursty”



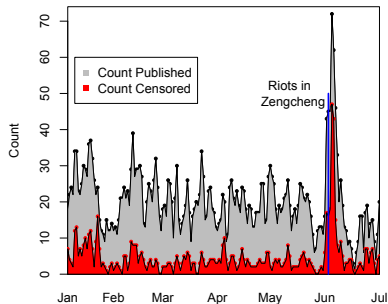
- Unit of analysis:
 - volume burst

All other topics: Censorship & Post Volume are “Bursty”



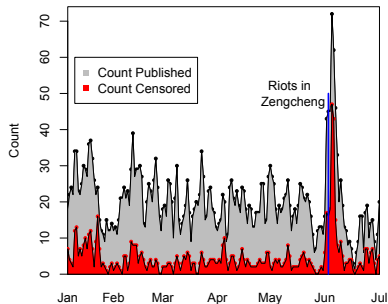
- Unit of analysis:
 - volume burst
 - (≈ 3 SDs greater than baseline volume)

All other topics: Censorship & Post Volume are “Bursty”



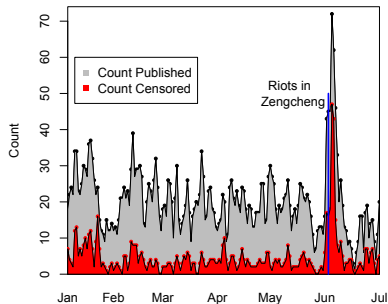
- Unit of analysis:
 - volume burst
 - (≈ 3 SDs greater than baseline volume)
- We monitored **85 topic areas** (Jan–July 2011)

All other topics: Censorship & Post Volume are “Bursty”



- Unit of analysis:
 - volume burst
 - (≈ 3 SDs greater than baseline volume)
- We monitored 85 topic areas (Jan–July 2011)
- Found 87 volume bursts in total

All other topics: Censorship & Post Volume are “Bursty”



- Unit of analysis:
 - volume burst
 - (≈ 3 SDs greater than baseline volume)
- We monitored 85 topic areas (Jan–July 2011)
- Found 87 volume bursts in total
- Identified real world events associated with each burst

The Goals of Chinese Censorship

The Goals of Chinese Censorship

Theories. The Purpose of Censorship is to:

- 1 Stop criticism of the state

The Goals of Chinese Censorship

Theories. The Purpose of Censorship is to:

- 1 Stop criticism of the state
- 2 Stop collective action

The Goals of Chinese Censorship

Theories. The Purpose of Censorship is to:

- 1 Stop criticism of the state
- 2 Stop collective action

Either or both could be right or wrong.

The Goals of Chinese Censorship

Theories. The Purpose of Censorship is to:

- 1 Stop criticism of the state *Wrong*
- 2 Stop collective action

Either or both could be right or wrong.

The Goals of Chinese Censorship

Theories. The Purpose of Censorship is to:

- 1 Stop criticism of the state *Wrong*
- 2 Stop collective action *Right*

Either or both could be right or wrong.

Our Central Hypothesis: The Rules Censors Follow

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events
- 3 Select events that could (or did) generate collective action

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events
- 3 Select events that could (or did) generate collective action
- 4 Censor all posts within bursts associated with these events

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events
- 3 Select events that could (or did) generate collective action
- 4 Censor all posts within bursts associated with these events
- 5 Judge individual posts for association with events, not for the posts' collective action potential (or for criticism of the state)

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events
- 3 Select events that could (or did) generate collective action
- 4 Censor all posts within bursts associated with these events
- 5 Judge individual posts for association with events, not for the posts' collective action potential (or for criticism of the state)

⇒ The government censors all posts (regardless of how critical or supportive of the state) in volume bursts associated with events with collective action potential

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events
- 3 Select events that could (or did) generate collective action
- 4 Censor all posts within bursts associated with these events
- 5 Judge individual posts for association with events, not for the posts' collective action potential (or for criticism of the state)

⇒ The government censors all posts (regardless of how critical or supportive of the state) in volume bursts associated with events with collective action potential

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events
- 3 Select events that could (or did) generate collective action
- 4 Censor all posts within bursts associated with these events
- 5 Judge individual posts for association with events, not for the posts' collective action potential (or for criticism of the state)

↪ The government censors **all posts** (regardless of how critical or supportive of the state) in volume bursts associated with events with collective action potential

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events
- 3 Select events that could (or did) generate collective action
- 4 Censor all posts within bursts associated with these events
- 5 Judge individual posts for association with events, not for the posts' collective action potential (or for criticism of the state)

↪ The government censors **all posts** (regardless of how critical or supportive of the state) in **volume bursts** associated with events with collective action potential

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events
- 3 Select events that could (or did) generate collective action
- 4 Censor all posts within bursts associated with these events
- 5 Judge individual posts for association with events, not for the posts' collective action potential (or for criticism of the state)

⇒ The government censors **all posts** (regardless of how critical or supportive of the state) in **volume bursts** associated with **events** with collective action potential

Our Central Hypothesis: The Rules Censors Follow

- 1 Monitor many broad topic areas
- 2 Detect volume bursts; identify associated real world events
- 3 Select events that could (or did) generate collective action
- 4 Censor all posts within bursts associated with these events
- 5 Judge individual posts for association with events, not for the posts' collective action potential (or for criticism of the state)

↪ The government censors **all posts** (regardless of how critical or supportive of the state) in **volume bursts** associated with **events** with **collective action potential**

Our Evidence

Our Evidence

3 increasingly specific tests based on

Our Evidence

3 increasingly specific tests based on

- 1 Post volume

Our Evidence

3 increasingly specific tests based on

- 1 Post volume
- 2 Nature of the event generating each volume burst

Our Evidence

3 increasingly specific tests based on

- 1 Post volume
- 2 Nature of the event generating each volume burst
- 3 Content of the posts

Test 1: Post Volume

Test 1: Post Volume

- Begin with our **87 volume bursts** in 85 topics areas

Test 1: Post Volume

- Begin with our **87 volume bursts** in 85 topics areas
- For each burst, calculate change in % censorship inside to outside each volume burst within topic areas – **censorship magnitude**

Test 1: Post Volume

- Begin with our **87 volume bursts** in 85 topics areas
- For each burst, calculate change in % censorship inside to outside each volume burst within topic areas – **censorship magnitude**
- If goal of censorship is to stop collective action, **we expect:**

Test 1: Post Volume

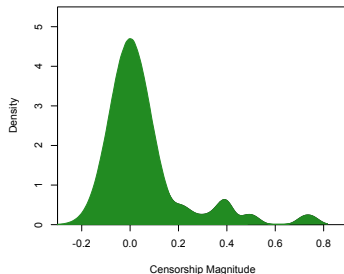
- Begin with our **87 volume bursts** in 85 topics areas
 - For each burst, calculate change in % censorship inside to outside each volume burst within topic areas – **censorship magnitude**
 - If goal of censorship is to stop collective action, **we expect:**
- 1 On average, % censored should increase during volume bursts

Test 1: Post Volume

- Begin with our **87 volume bursts** in 85 topics areas
- For each burst, calculate change in % censorship inside to outside each volume burst within topic areas – **censorship magnitude**
- If goal of censorship is to stop collective action, **we expect:**
 - 1 On average, % censored should increase during volume bursts
 - 2 Some bursts (associated with politically relevant events) should have much higher censorship

Test 1: Post Volume

- Begin with our **87 volume bursts** in 85 topics areas
- For each burst, calculate change in % censorship inside to outside each volume burst within topic areas – **censorship magnitude**
- If goal of censorship is to stop collective action, **we expect**:
 - 1 On average, % censored should increase during volume bursts
 - 2 Some bursts (associated with politically relevant events) should have much higher censorship



Test 2: Nature of Event Generating Increases in Volume

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

① Collective Action Potential

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

① Collective Action Potential

- protest or organized crowd formation outside the Internet

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

① Collective Action Potential

- protest or organized crowd formation outside the Internet
- individuals who have organized or incited collective action on the ground in the past;

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

① Collective Action Potential

- protest or organized crowd formation outside the Internet
- individuals who have organized or incited collective action on the ground in the past;
- topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

① Collective Action Potential

- protest or organized crowd formation outside the Internet
- individuals who have organized or incited collective action on the ground in the past;
- topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.

② Criticism of censors

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

① Collective Action Potential

- protest or organized crowd formation outside the Internet
- individuals who have organized or incited collective action on the ground in the past;
- topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.

② Criticism of censors

③ Pornography

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

① Collective Action Potential

- protest or organized crowd formation outside the Internet
- individuals who have organized or incited collective action on the ground in the past;
- topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.

② Criticism of censors

③ Pornography

④ (Other) News

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

① Collective Action Potential

- protest or organized crowd formation outside the Internet
- individuals who have organized or incited collective action on the ground in the past;
- topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.

② Criticism of censors

③ Pornography

④ (Other) News

⑤ Government Policies

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

- 1 ~~Collective Action Potential~~ **CENSORED**
 - protest or organized crowd formation outside the Internet
 - individuals who have organized or incited collective action on the ground in the past;
 - topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.
- 2 Criticism of censors
- 3 Pornography
- 4 (Other) News
- 5 Government Policies

Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

- 1 ~~Collective Action Potential~~ **CENSORED**
 - protest or organized crowd formation outside the Internet
 - individuals who have organized or incited collective action on the ground in the past;
 - topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.
- 2 ~~Criticism of censors~~ **CENSORED**
- 3 Pornography
- 4 (Other) News
- 5 Government Policies


Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

- 1 ~~Collective Action Potential~~ **CENSORED**
 - protest or organized crowd formation outside the Internet
 - individuals who have organized or incited collective action on the ground in the past;
 - topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.
- 2 ~~Criticism of censors~~ **CENSORED**
- 3 ~~Pornography~~ **CENSORED**
- 4 (Other) News
- 5 Government Policies






Test 2: Nature of Event Generating Increases in Volume

Event classification (each category can be +, -, or neutral comments about the state)

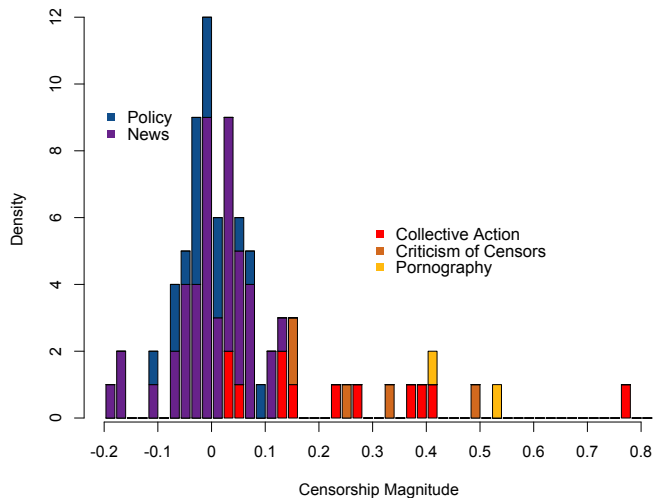
- 1 ~~Collective Action Potential~~ **CENSORED**
 - protest or organized crowd formation outside the Internet
 - individuals who have organized or incited collective action on the ground in the past;
 - topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.
- 2 ~~Criticism of censors~~ **CENSORED**
- 3 ~~Pornography~~ **CENSORED**
- 4 (Other) News 
- 5 Government Policies

Test 2: Nature of Event Generating Increases in Volume

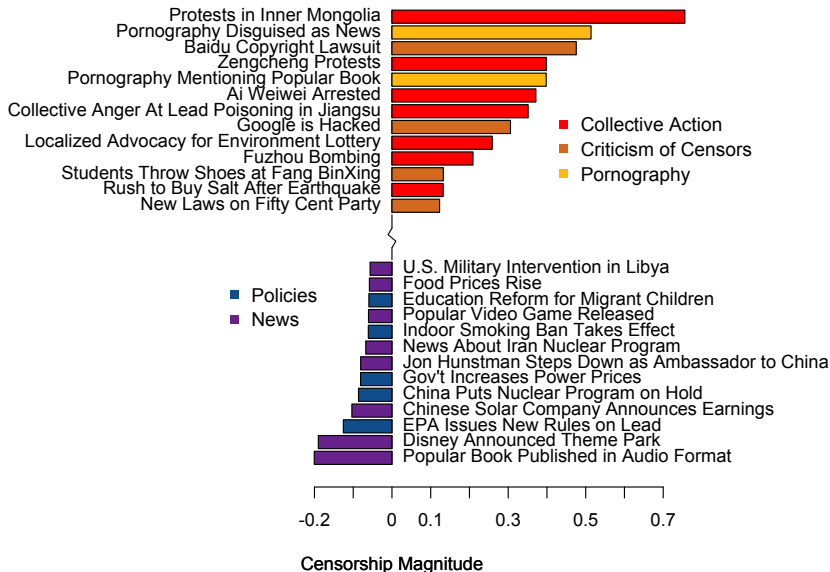
Event classification (each category can be +, -, or neutral comments about the state)

- 1 ~~Collective Action Potential~~ 
 - protest or organized crowd formation outside the Internet
 - individuals who have organized or incited collective action on the ground in the past;
 - topics related to nationalism or nationalist sentiment that have incited protest or collective action in the past.
- 2 ~~Criticism of censors~~ 
- 3 ~~Pornography~~ 
- 4 (Other) News 
- 5 Government Policies 

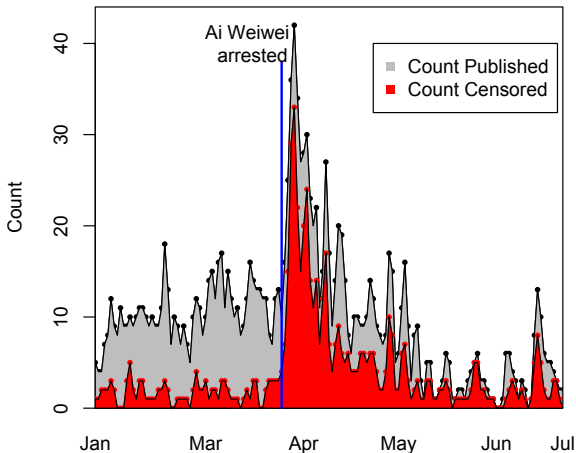
What Types of Events Are Censored?



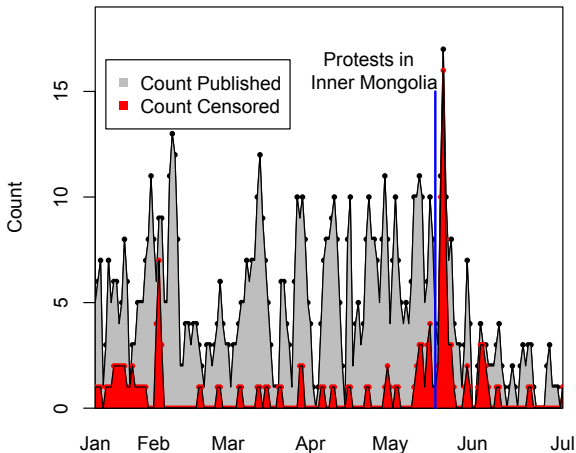
What Types of Events Are Censored?



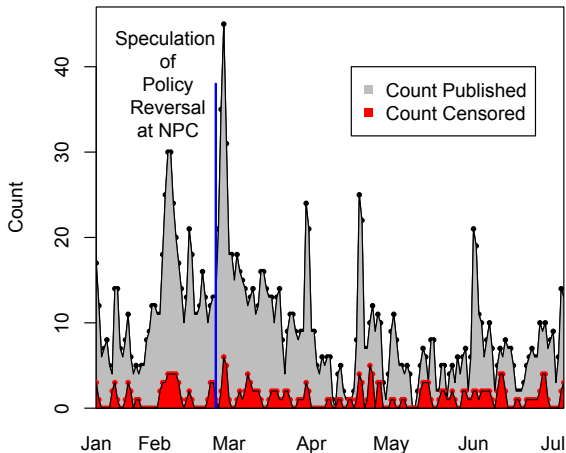
Censoring Collective Action: Ai Weiwei's Arrest



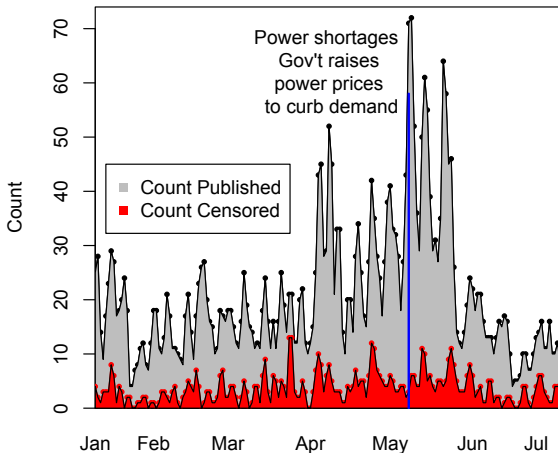
Censoring Collective Action: Protests in Inner Mongolia



Low Censorship on One Child Policy



Low Censorship on News: Power Prices



Test 3: Content of Posts

Test 3: Content of Posts

We adapted “ReadMe” (Hopkins-King) methodology to the content of censored posts in Chinese

Test 3: Content of Posts

We adapted “ReadMe” (Hopkins-King) methodology to the content of censored posts in Chinese

- Standardize encoding of Chinese text

Test 3: Content of Posts

We adapted “ReadMe” (Hopkins-King) methodology to the content of censored posts in Chinese

- Standardize encoding of Chinese text
- Eliminate punctuation and stop words

Test 3: Content of Posts

We adapted “ReadMe” (Hopkins-King) methodology to the content of censored posts in Chinese

- Standardize encoding of Chinese text
- Eliminate punctuation and stop words
- Experiment with segmenting characters into words

Test 3: Content of Posts

We adapted “ReadMe” (Hopkins-King) methodology to the content of censored posts in Chinese

- Standardize encoding of Chinese text
- Eliminate punctuation and stop words
- Experiment with segmenting characters into words
- Calculate the quantity of interest (% censored by category)

Test 3: Content of Posts

We adapted “ReadMe” (Hopkins-King) methodology to the content of censored posts in Chinese

- Standardize encoding of Chinese text
- Eliminate punctuation and stop words
- Experiment with segmenting characters into words
- Calculate the quantity of interest (% censored by category)
 - Code a validation set

Test 3: Content of Posts

We adapted “ReadMe” (Hopkins-King) methodology to the content of censored posts in Chinese

- Standardize encoding of Chinese text
- Eliminate punctuation and stop words
- Experiment with segmenting characters into words
- Calculate the quantity of interest (% censored by category)
 - Code a validation set
 - ReadMe estimates: % of posts in each category

Test 3: Content of Posts

We adapted “ReadMe” (Hopkins-King) methodology to the content of censored posts in Chinese

- Standardize encoding of Chinese text
- Eliminate punctuation and stop words
- Experiment with segmenting characters into words
- Calculate the quantity of interest (% censored by category)
 - Code a validation set
 - ReadMe estimates: % of posts in each category
 - Quantity of interest: % of posts censored in a category

Test 3: Content of Posts

We adapted “ReadMe” (Hopkins-King) methodology to the content of censored posts in Chinese

- Standardize encoding of Chinese text
- Eliminate punctuation and stop words
- Experiment with segmenting characters into words
- Calculate the quantity of interest (% censored by category)
 - Code a validation set
 - ReadMe estimates: % of posts in each category
 - Quantity of interest: % of posts censored in a category

$$P(\text{Censored}|\text{Category}) = \frac{P(\text{Category}|\text{Censored})P(\text{Censored})}{P(\text{Category})}$$

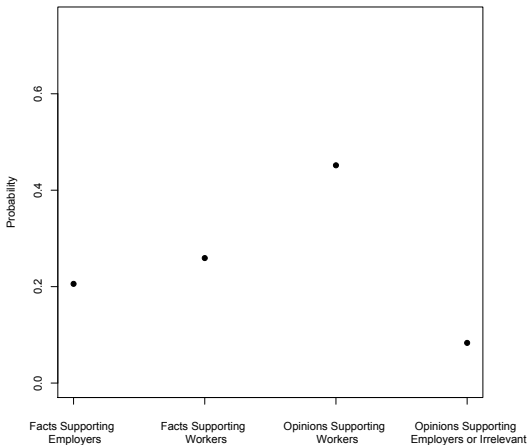
“ReadMe” Algorithm Validated in Chinese

“ReadMe” Algorithm Validated in Chinese

Example: Labor Strikes, 2010 (Training set: 100; Test set: 900)

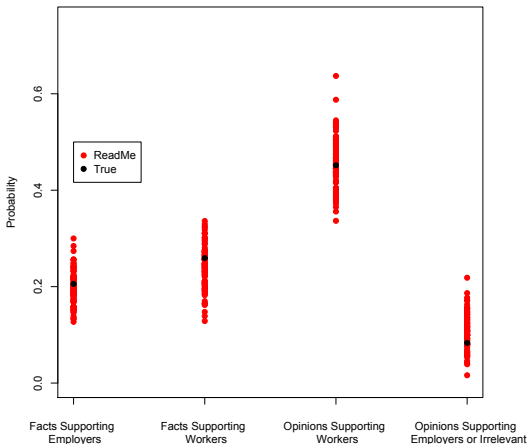
“ReadMe” Algorithm Validated in Chinese

Example: Labor Strikes, 2010 (Training set: 100; Test set: 900)



“ReadMe” Algorithm Validated in Chinese

Example: Labor Strikes, 2010 (Training set: 100; Test set: 900)



Test 3: Content of Censored and Uncensored Posts

Test 3: Content of Censored and Uncensored Posts

Predictions from Two Theories of Censorship



Predictions of State Criticism Theory

		Collective Action	
		yes	no
State Critique	yes	CENSORED	CENSORED
	no	Free	Free

Test 3: Content of Censored and Uncensored Posts

Predictions from Two Theories of Censorship

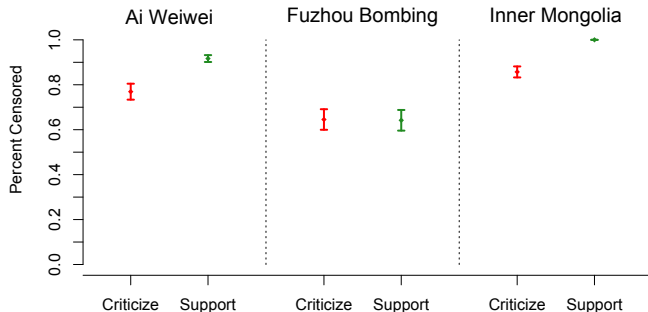
Predictions of Collective Action Theory

		Collective Action	
		yes	no
State Critique	yes		Free
	no		Free

With Collective Action, Posts that Support or Criticize the State are Censored

With Collective Action, Posts that Support or Criticize the State are Censored

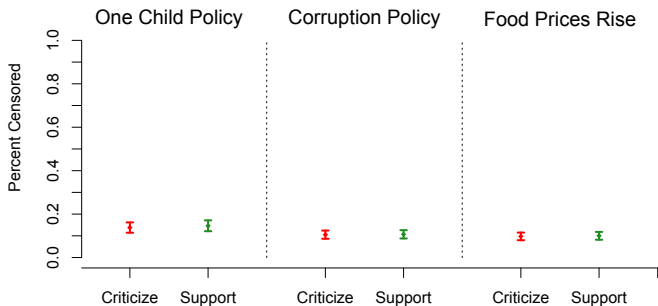
		Collective Action	
		yes	no
State Critique	yes	CENSORED	Free
	no	CENSORED	Free



Without Collective Action, Posts that Support *or* Oppose the State are NOT Censored



Without Collective Action, Posts that Support or Oppose the State are NOT Censored

		Collective Action	
		yes	no
State Critique	yes	CENSORED	Free
	no	CENSORED	Free



Uncensored Posts (w/o Collective Action Potential) Critical of the State

Uncensored Posts (w/o Collective Action Potential) Critical of the State

		Collective Action	
		yes	no
State Critique	yes		Free
	no		Free

Uncensored Posts (w/o Collective Action Potential) Critical of the State

This is a city government [Yulin City, Shaanxi] that treats life with contempt, this is government officials run amuck, a city government without justice, a city government that delights in that which is vulgar, a place where officials all have mistresses, a city government that is shameless with greed, a government that trades dignity for power, a government without humanity, a government that has no limits on immorality, a government that goes back on its word, a government that treats kindness with ingratitude, a government that cares nothing for posterity. . .

这是一个漠视生命的市政府[陕西省榆林市]、一个官员横行的市政府、一个没有公正的市政府，一个低级趣味的市政府，一个包二奶的市政府，一个为钱不要脸的市政府，一个为个权不要人格的市政府，一个没有血性的市政府，一个没有道德低线的市政府，一个出尔反尔的市政府，一个忘恩负义的市政府，一个不要子孙后代的市政府，一个什么怪事都出的市政府，一个什么的市政府，只要你想到的就有…

Censored Post (with Collective Action Potential) Supporting the State

Censored Post (with Collective Action Potential) Supporting the State

		Collective Action	
		yes	no
State Critique	yes		Free
	no		Free

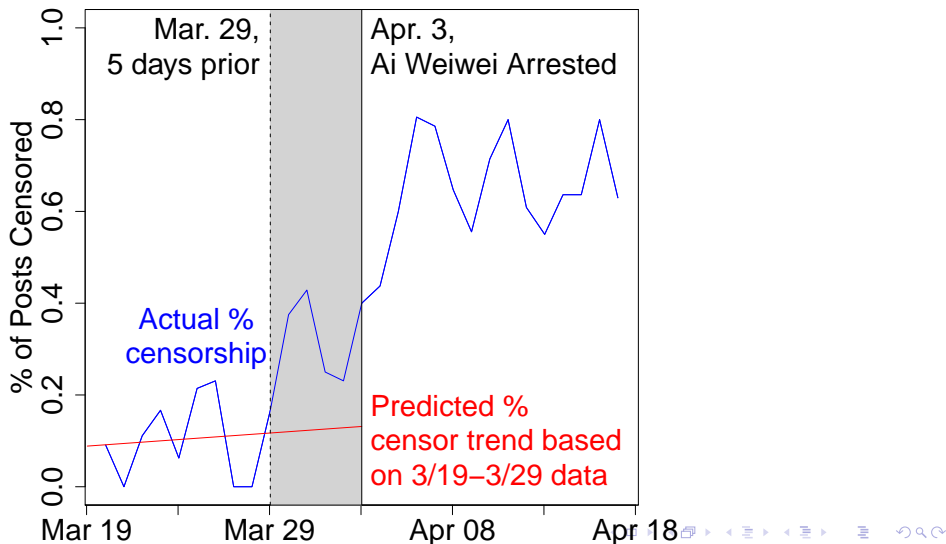
Censored Post (with Collective Action Potential) Supporting the State

The bombing led not only to the tragedy of his death but the death of many government workers. Even if we can verify what Qian Mingqi said on Weibo that the building demolition caused a great deal of personal damage, we should still condemn his extreme act of retribution. . . . The government has continually put forth measures and laws to protect the interests of citizens in building demolition. And the media has called attention to the plight of those experiencing housing demolition. The rate at which compensation for housing demolition has increased exceeds inflation. In many places, this compensation can change the fate of an entire family.

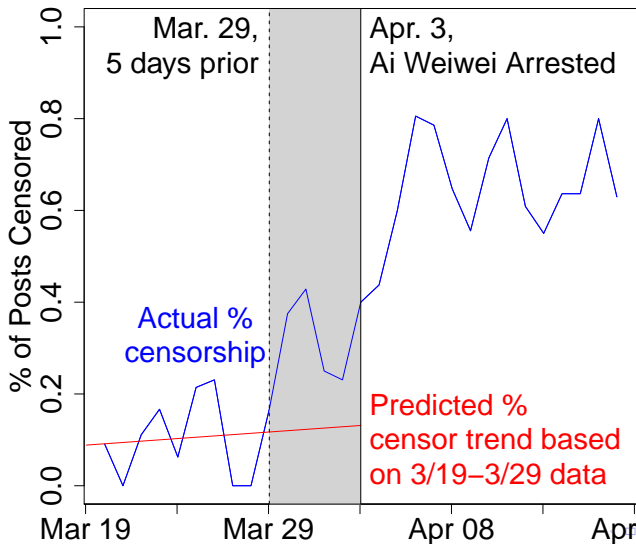
爆炸案造成他本人和多名政府工作人员死伤的悲剧，即使钱明奇在微博里所称拆迁造成的个人损失是属实的，我们也应谴责他的极端报复行为. . . . 政府在连续出台保护被拆迁者利益的政策法规，媒体也在为公平对待被拆迁者大声疾呼，各地拆迁补偿款的上升速度，大多高于商品房售价上升的速度，在不少地方，补偿款已经足以改变一个家庭的命运。

Censorship Pre-empting Collective Action: Ai Weiwei's Arrest

Censorship Pre-empting Collective Action: Ai Weiwei's Arrest

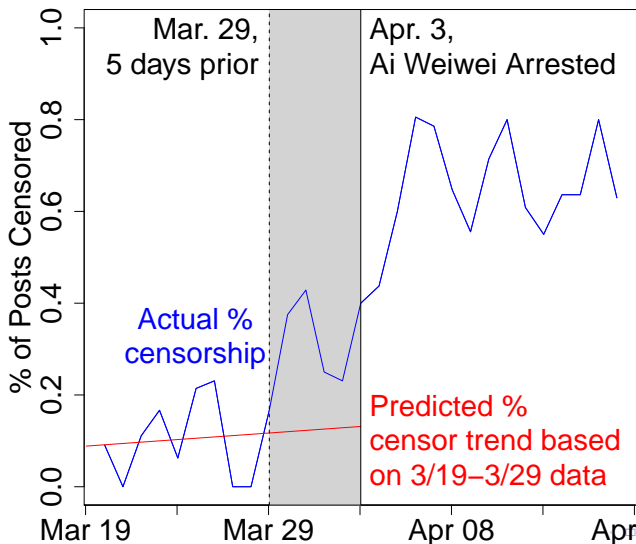


Censorship Pre-empting Collective Action: Ai Weiwei's Arrest



Placebo Test:

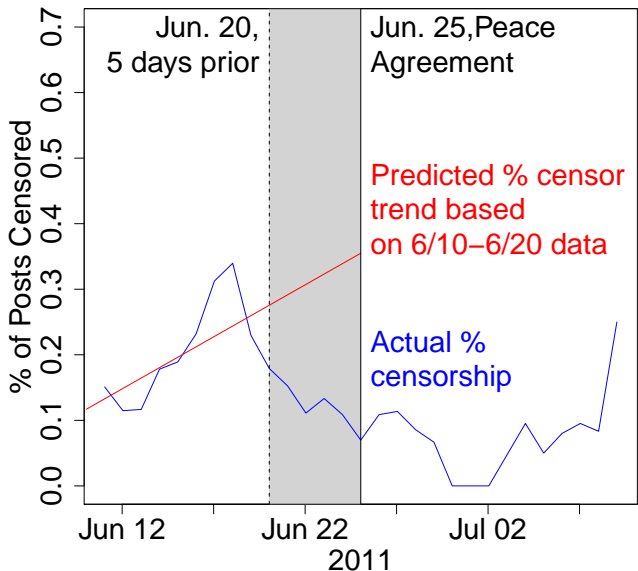
Censorship Pre-empting Collective Action: Ai Weiwei's Arrest



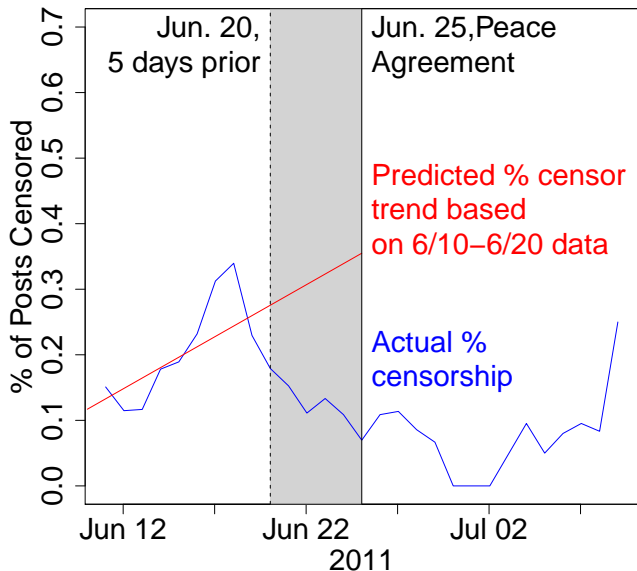
Placebo Test: Most extreme of all effects

Predicting the South China Sea Peace Agreement

Predicting the South China Sea Peace Agreement

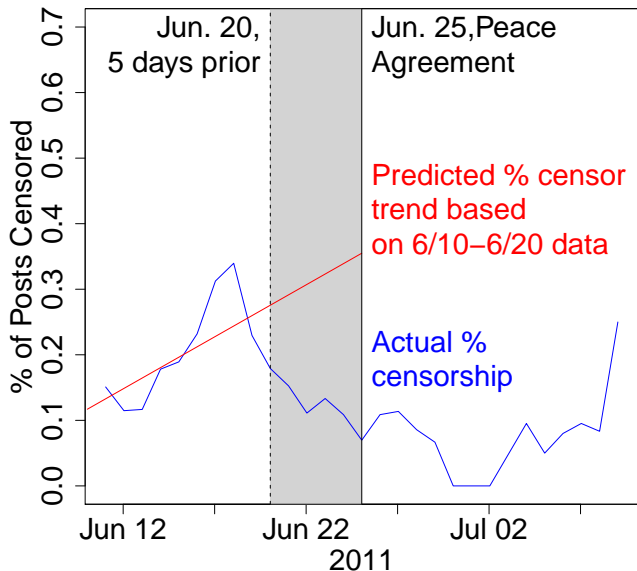


Predicting the South China Sea Peace Agreement



Placebo Test:

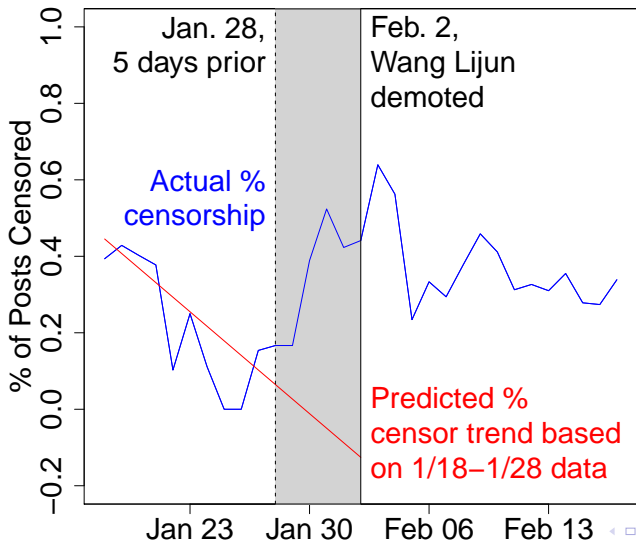
Predicting the South China Sea Peace Agreement



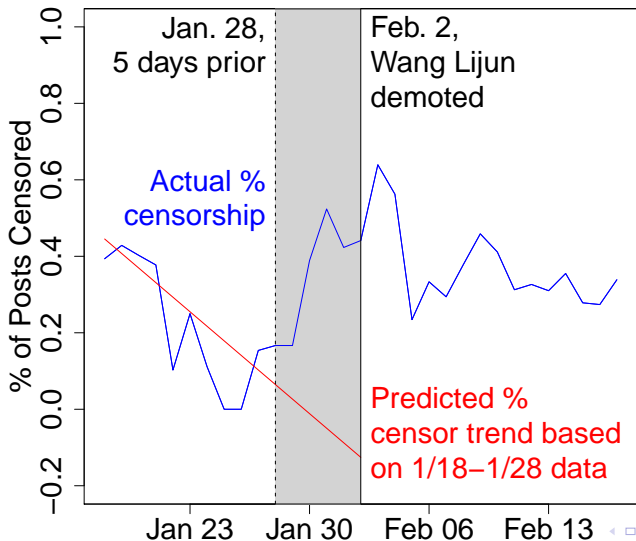
Placebo Test: Most extreme of all effects

Censorship Pre-empting Collective Action: Wang Lijun's Demotion

Censorship Pre-empting Collective Action: Wang Lijun's Demotion

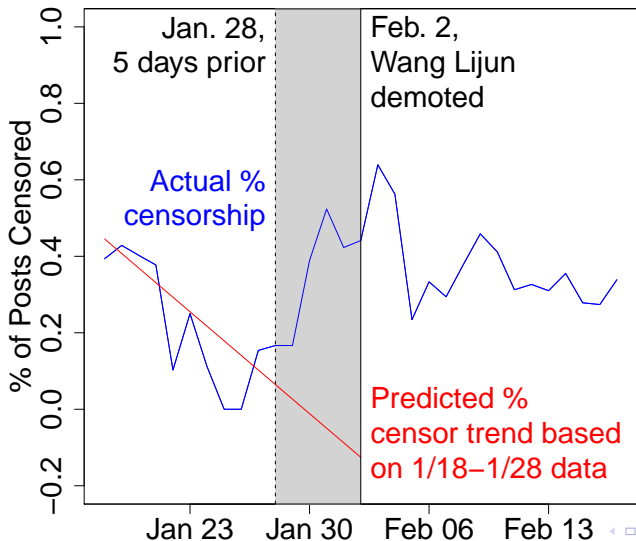


Censorship Pre-empting Collective Action: Wang Lijun's Demotion



Placebo Test:

Censorship Pre-empting Collective Action: Wang Lijun's Demotion



Placebo Test: Most extreme of all effects

Concluding Remarks

Concluding Remarks

- The Chinese Censorship Program:

Concluding Remarks

- The Chinese Censorship Program:
 - Stops collective action, not criticism

Concluding Remarks

- The Chinese Censorship Program:
 - Stops collective action, not criticism
 - Reveals government intentions; predicts actions

Concluding Remarks

- The Chinese Censorship Program:
 - Stops collective action, not criticism
 - Reveals government intentions; predicts actions
- Our methodology:

Concluding Remarks

- The Chinese Censorship Program:
 - Stops collective action, not criticism
 - Reveals government intentions; predicts actions
- Our methodology:
 - Automated text analysis methods validated in Chinese (without translation)

Concluding Remarks

- The Chinese Censorship Program:
 - Stops collective action, not criticism
 - Reveals government intentions; predicts actions
- Our methodology:
 - Automated text analysis methods validated in Chinese (without translation)
 - Enables continuous time measurement of Chinese government censorship activities

Concluding Remarks

- The Chinese Censorship Program:
 - Stops collective action, not criticism
 - Reveals government intentions; predicts actions
- Our methodology:
 - Automated text analysis methods validated in Chinese (without translation)
 - Enables continuous time measurement of Chinese government censorship activities
 - Reveals information about state action when traditional media is silent

Concluding Remarks

- The Chinese Censorship Program:
 - Stops collective action, not criticism
 - Reveals government intentions; predicts actions
- Our methodology:
 - Automated text analysis methods validated in Chinese (without translation)
 - Enables continuous time measurement of Chinese government censorship activities
 - Reveals information about state action when traditional media is silent
 - Might be applicable to other countries

For more information



gking.harvard.edu



people.fas.harvard.edu/~jjpan



scholar.harvard.edu/mroberts