Eller College of Management

Does Government Surveillance Give Twitter the Chills?

Laura Brandimarte

Joint work with:

Edward McFowland III (University of Minnesota)

Sriram Somanchi (University of Notre Dame)

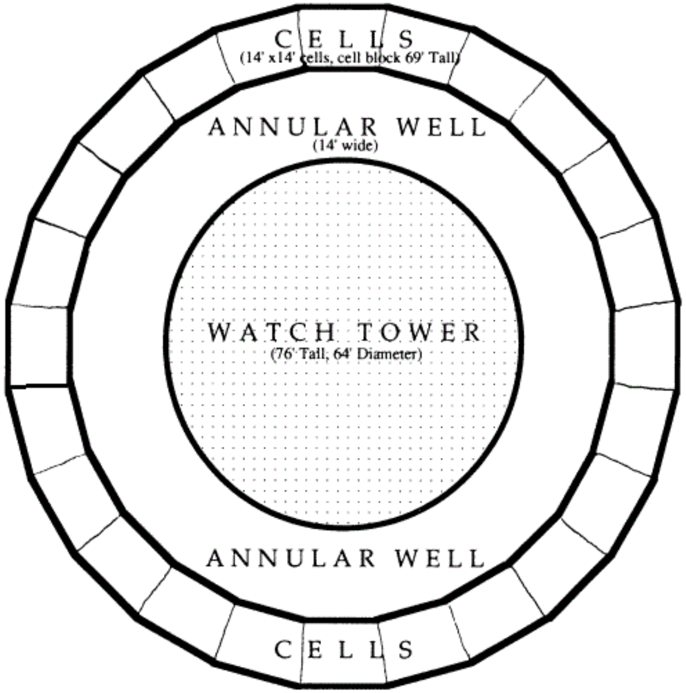
Uttara Ananthakrishnan (Carnegie Mellon University)

Global Governance B. A. – University of Rome Tor Vergata

October 9, 2017

Government Surveillance

- Constant collection of vast quantities of data about US and non-US citizens, mostly unbeknownst to them
- Big Brother is an old idea
  - Bentham's Panopticon (18th century)



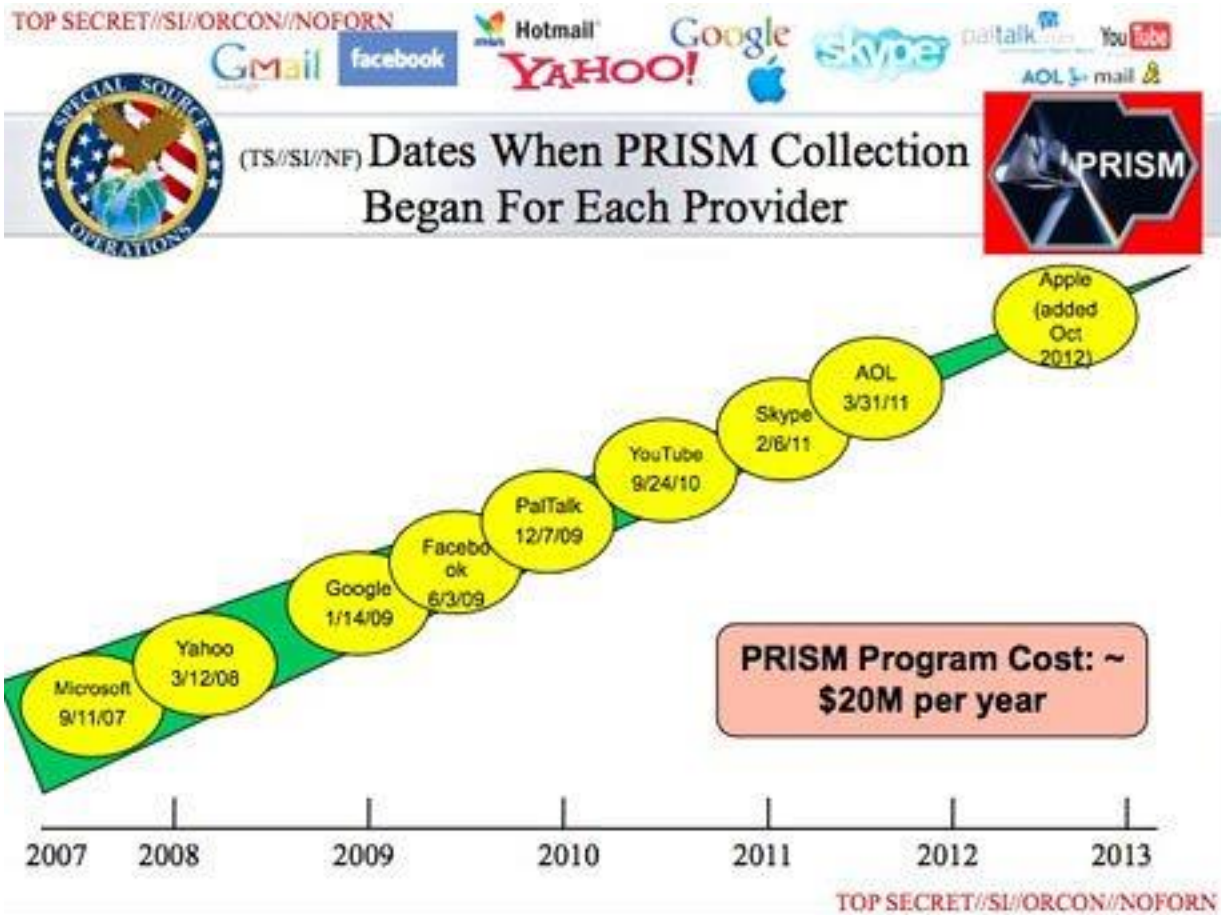
  - George Orwell, 1984

Government Surveillance

- We all probably knew that the Government has been watching us – how can they guarantee any safety and security otherwise?

- What we may not have expected is the extent to which surveillance is implemented
- Until CITIZENFOUR came along…

Glenn Greenwald's article on The Guardian, 6-6-2013



NSA’s programs

- PRISM
  - E.g., NSA uses a Google cookie (PREF), probably obtained through a court order under Foreign Intelligence Surveillance Act, to identify targets to attack (hack)

NSA's programs

Our Approach

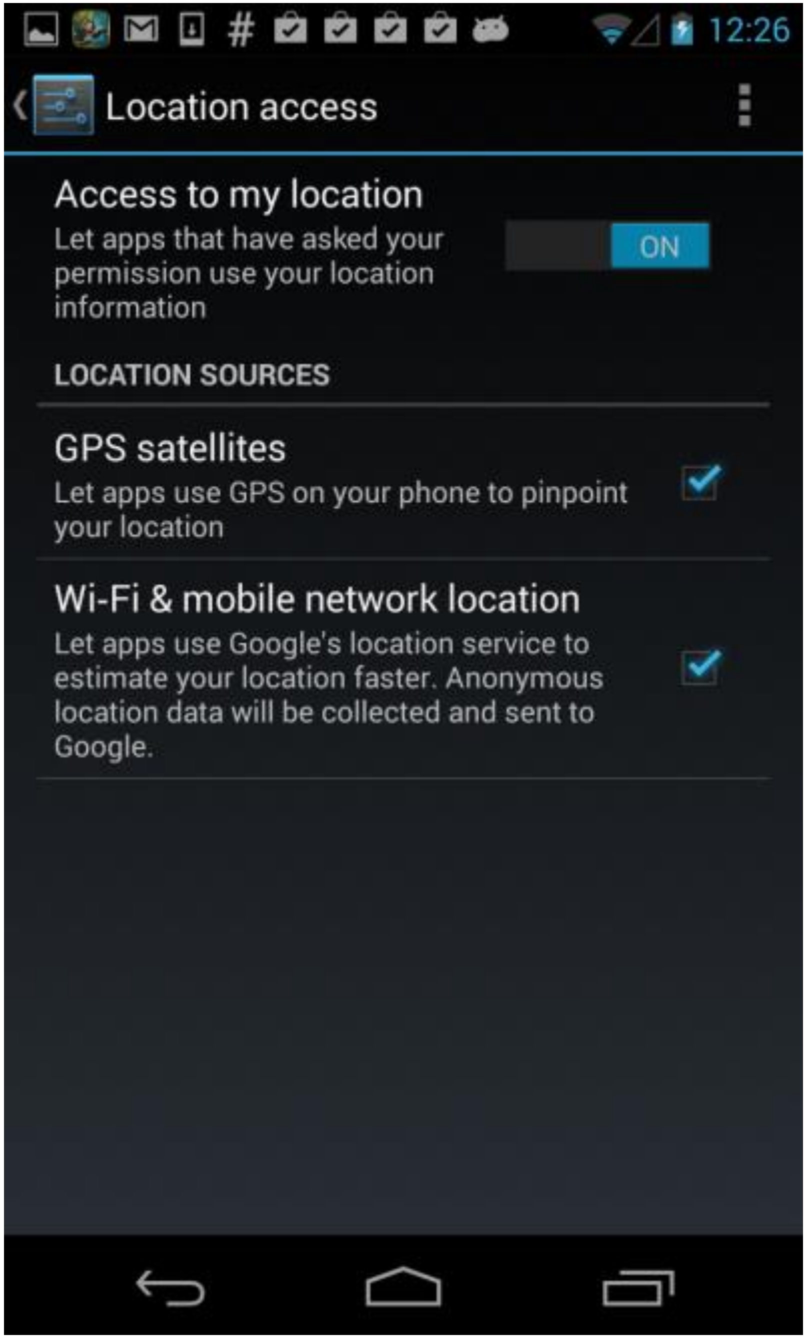
- Tracked target's converged communications and CNE accesses.
- Monitored passive internet traffic; created automated processes where possible (XKS ANCHORMAN, Workflows, Fingerprints).
- Provided TAO/GCHQ with WLLids/DSL accounts, Cookies, GooglePREFIDs to enable remote exploitation.
- Partnered with NGA and R4 to confirm locations and USRP equipment based on collected photographs.
- Drove CNE collection and partnered with TAO to increase USRP specific endpoint accesses.
- Provided knowledge to interagency partners for potential on the ground survey options and FBI-led intelligence guiding efforts.

NSA’s programs

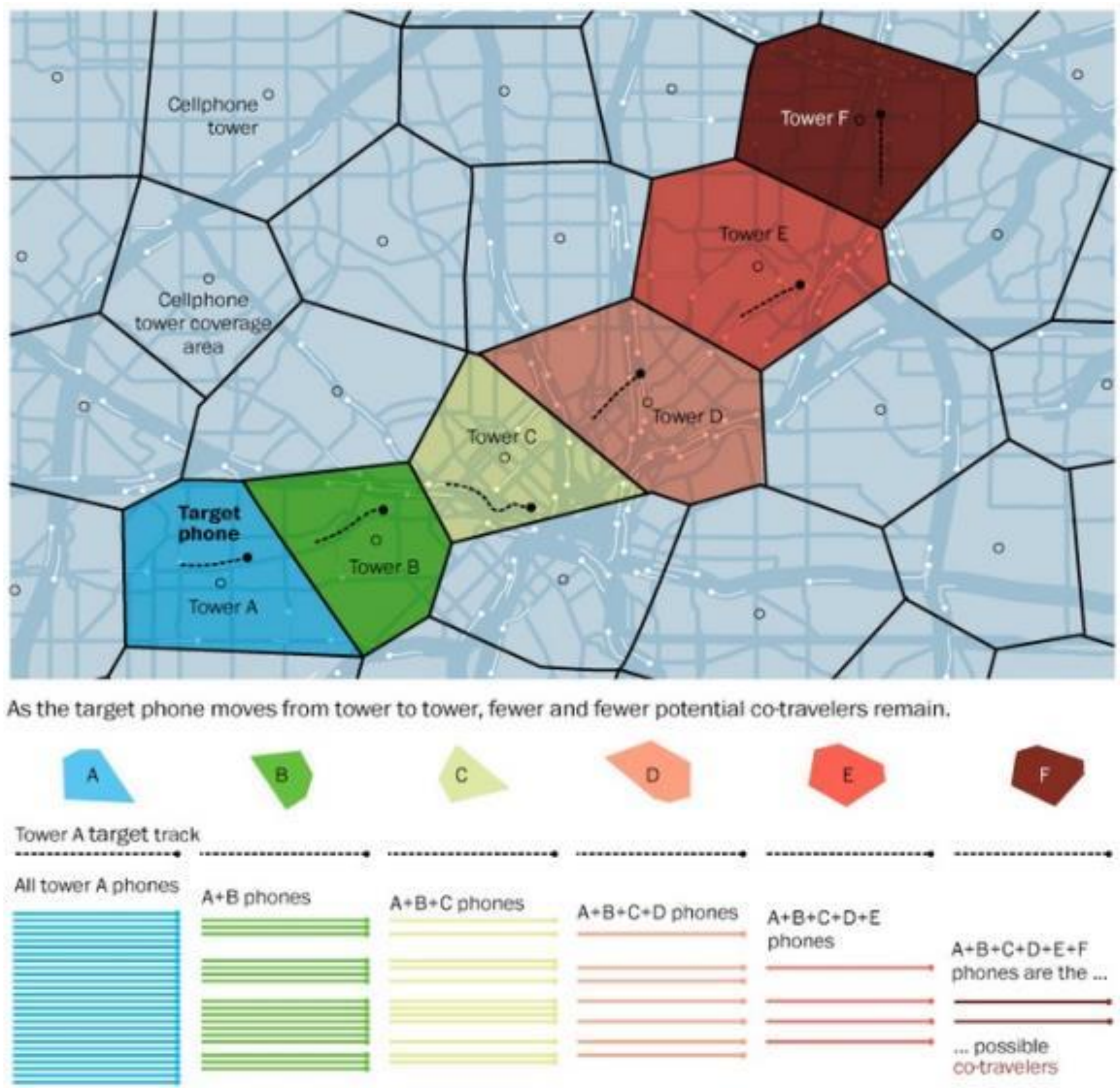
- PRISM
  - E.g., NSA uses a Google cookie (PREF), probably obtained through a court order under Foreign Intelligence Surveillance Act, to identify targets to attack (hack)
  - NSA uses DoubleClick’s cookies and “undercover nodes” to identify Tor users when they migrate to non-anonymous browsers
    - Tor Stinks (?!?)
  - NSA uses meta-data

NSA's programs

- HAPPYFOOT
  - E.g., NSA uses geo-location data used for mobile targeted ads in order to determine if US or non-US citizens are geographically close to and moving around with suspects


12:26
Location access
Access to my location
Let apps that have asked your permission use your location information
ON
LOCATION SOURCES
GPS satellites
Let apps use GPS on your phone to pinpoint your location
Wi-Fi & mobile network location
Let apps use Google's location service to estimate your location faster. Anonymous location data will be collected and sent to Google.

Eller College of Management



News World news Surveillance

Merkel phone tapping fair game under international law, says ex-MI6 deputy

Nigel Inkster says interception of German chancellor's calls by NSA might be judged 'politically unwise'

Richard Norton-Taylor
theguardian.com, Tuesday 18 February 2014 11.40 EST

Angela Merkel, the German chancellor. Photograph: Johannes Eisele/AFP/Getty Images

Intercepting the telephone calls of Angela Merkel would have been "politically unwise" and "certainly illegal under German law", according to a former senior British secret intelligence officer.

However, he says that under international law, tapping into the German chancellor's telephone conversations "would appear to be fair game".

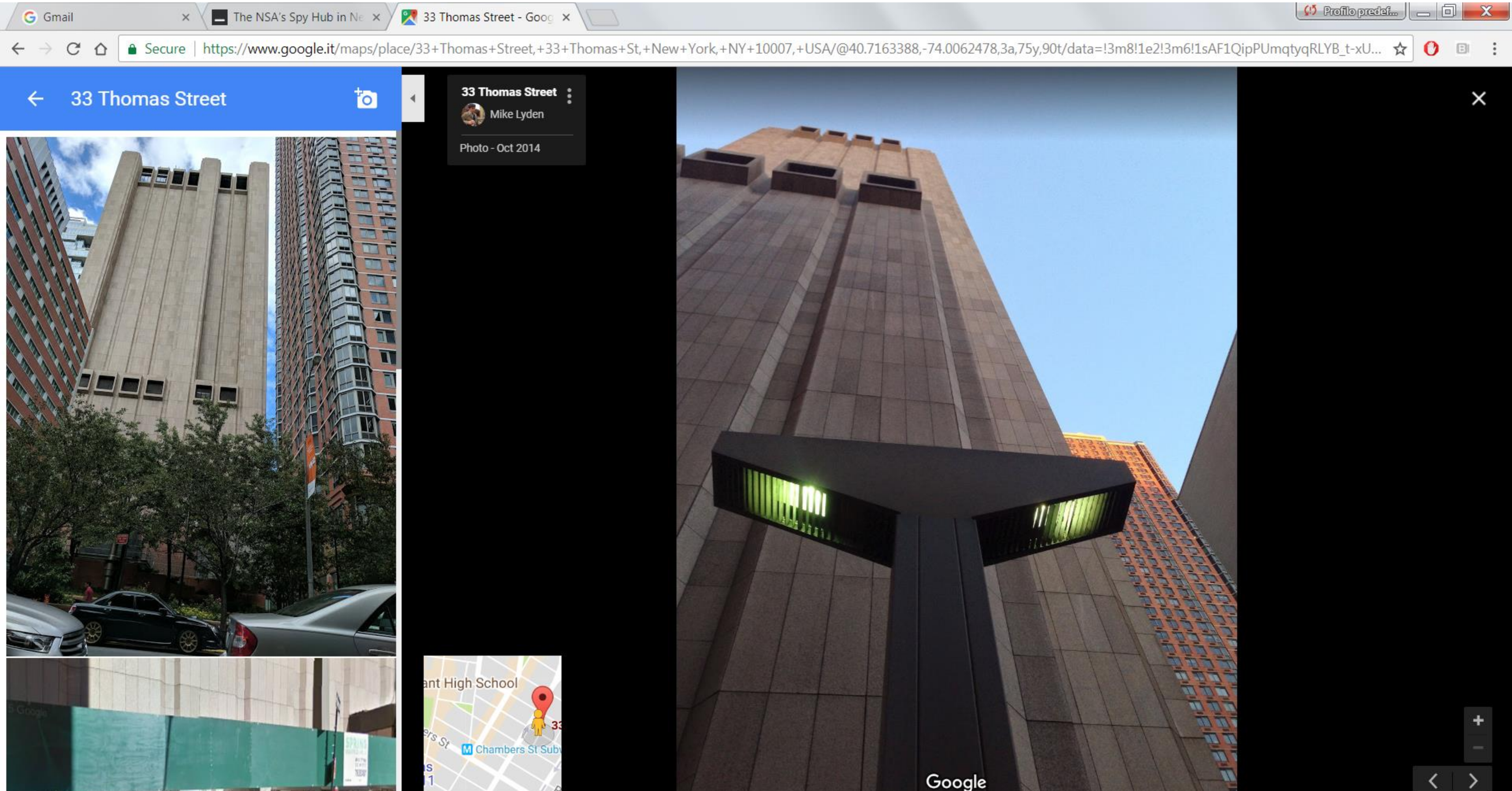
NSA’s programs

- BULLRUN: clandestine classified decryption program, partly run in collaboration with technology companies, as part of “multipronged effort” to weaken the encryption used in commercial software
  - E.g., “influencing and weakening encryption standards, by obtaining master keys, either by agreement, by force of law, or by computer network exploitation,” or by hardware-accelerated decryption for brute-force attacks

Eller College of Management

NSA's programs

- Project X



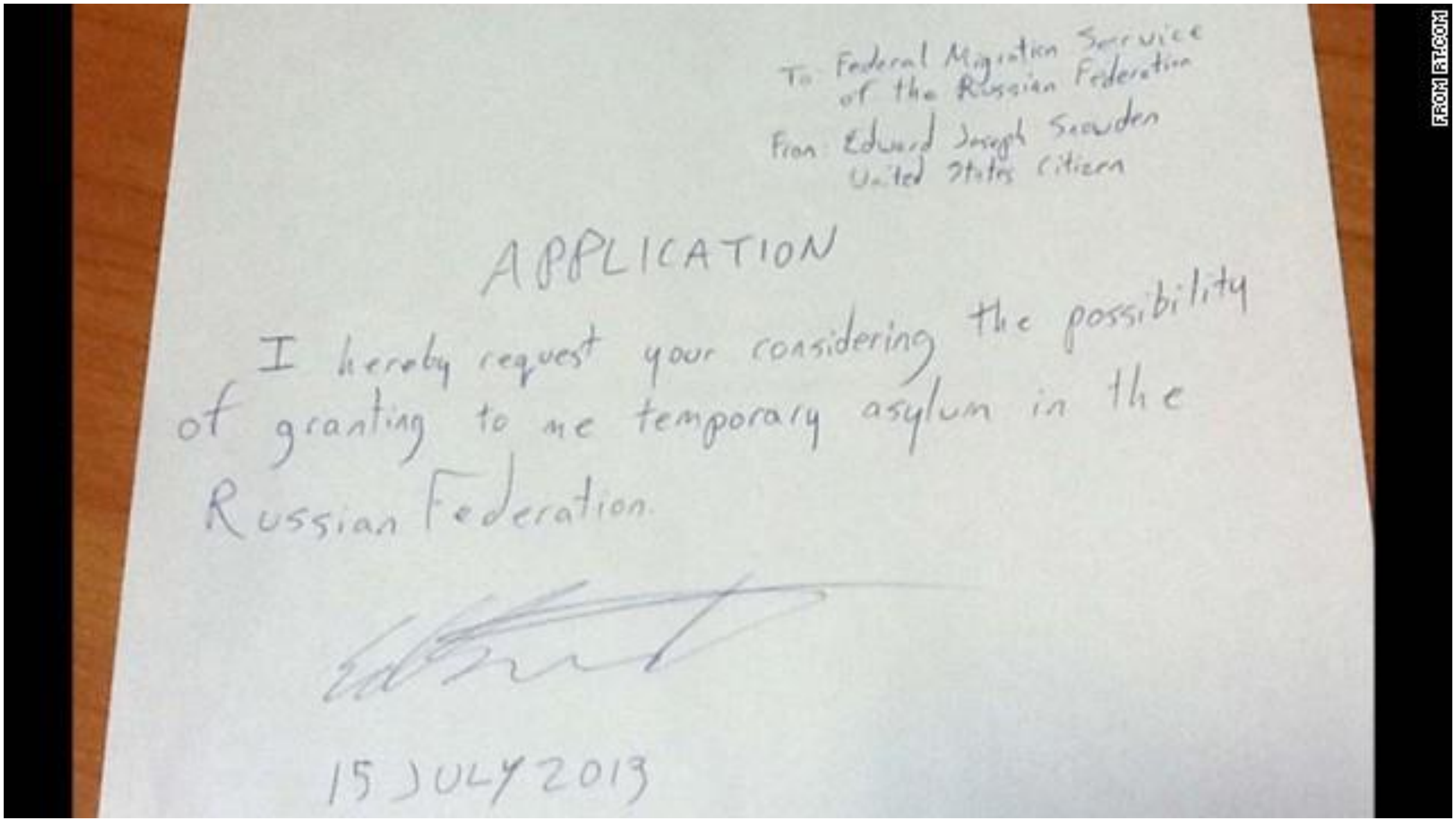
NSA's programs

- Project X
- As of July 2010, the NSA had obtained at least 40 court orders for spying under the BLARNEY program, allowing the agency to monitor communications related to multiple countries, companies, and international organizations
  - Among the approved targets were the IMF, the World Bank, the Bank of Japan, the EU, the UN, and at least 38 different countries, including U.S. allies such as Italy, Japan, Brazil, France, Germany, Greece, Mexico, and Cyprus

Ed Snowden reveals his identity, 6-11-2013

- 12-minute video

Ed Snowden flees to Russia, July 2013



To: Federal Migration Service
of the Russian Federation

From: Edward Joseph Snowden
United States Citizen

APPLICATION

I hereby request your considering the possibility of granting to me temporary asylum in the Russian Federation.

15 JULY 2013

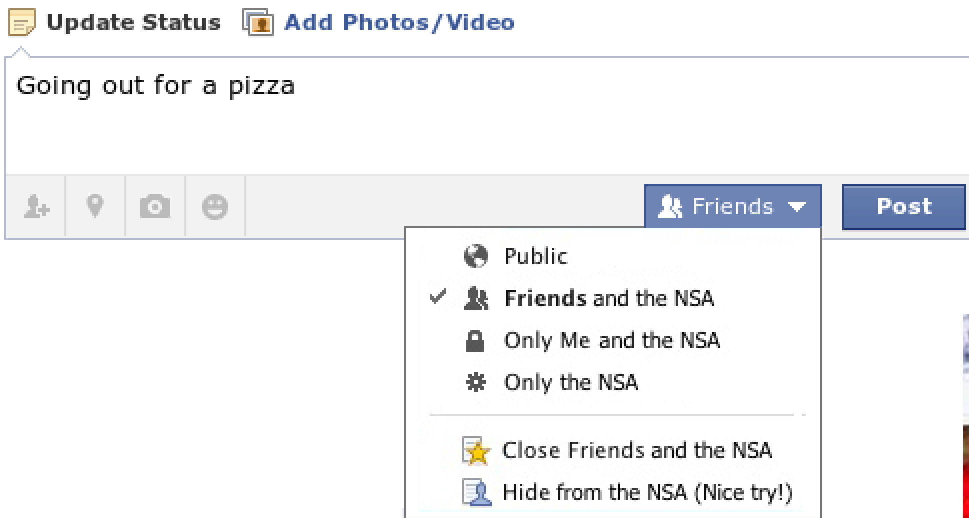
Eller College of Management

Ed Snowden flees to Russia, July 2013

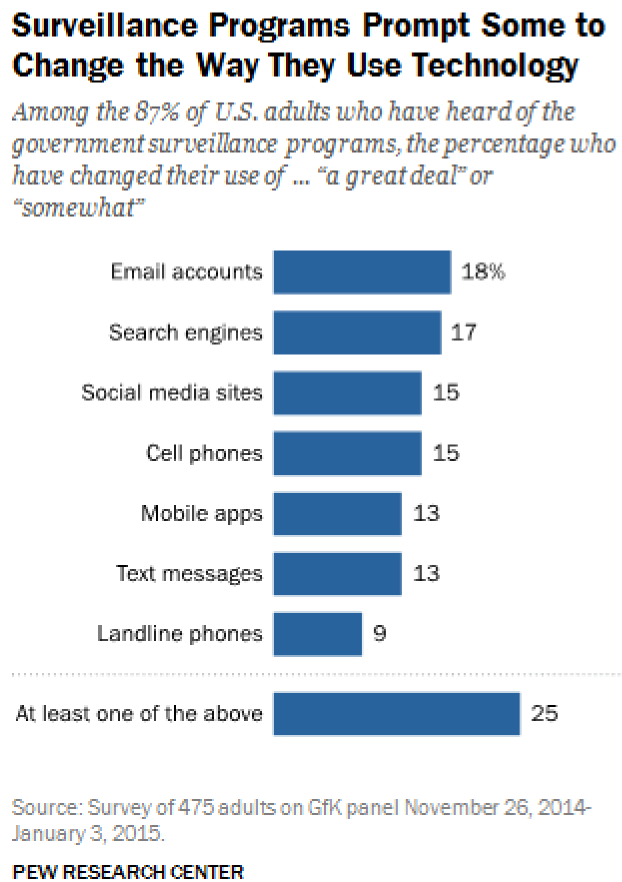

// A FILM BY ACADEMY AWARD® WINNER:
OLIVER STONE
THE ONLY SAFE PLACE IS ON THE RUN
SNOWDEN
JOSEPH GORDON-LEVITT // SHAILENE WOODLEY
// DIRECTED BY:
OLIVER STONE
// SCREENPLAY BY:
KIERAN FITZGERALD & OLIVER STONE
// IN THEATERS SEPTEMBER 16

Eller College of Management

How did the public react?



How did the public react?



How did the public react on Twitter?

Research questions

- Did awareness of Government surveillance programs affect the way people express themselves on Twitter?
  - Are they more or less inhibited when it comes to publicly discussing sensitive topics? Do they self-censor?
- If there is an effect, where is it more pronounced – in the States or abroad? And where exactly in the States?
- If there is an effect, is it a long-term or a short-term one?
  - Inhibitory effects of video-surveillance are relatively short-lived (Oulasvirta, 2012)

Previous attempts of estimating chilling effects of Government surveillance

- Marthews & Tucker (2014)
  - Google searches
- Stoycheff (2016)
  - Fictional Facebook post scenario
- Preibusch (2015)
  - Bing searches and TOR usage
- Penney (2016)
  - Wikipedia articles views

Previous attempts of estimating chilling effects of Government surveillance

- Marthews & Tucker (2014)
  - Google searches
- Stoycheff (2016)
  - Fictional Facebook post scenario
- Preibusch (2015)
  - Bing searches and TOR usage
- Penney (2016)
  - Wikipedia articles views

Sensitive vs. non-sensitive keywords

- Marthews & Tucker (2014)
  - list of words monitored by DHS in social media
    - e.g., explosion, anthrax, flu, pork…
  - list of words considered embarrassing if known to third parties (according to Mturk Workers)
    - e.g., AIDS, torrent…
  - list of most visited websites
- Our “treatment” group is similar, but our “control group” is different
  - List of food-related words borrowed from the literature (Abbar, Mejova& Weber, 2015)

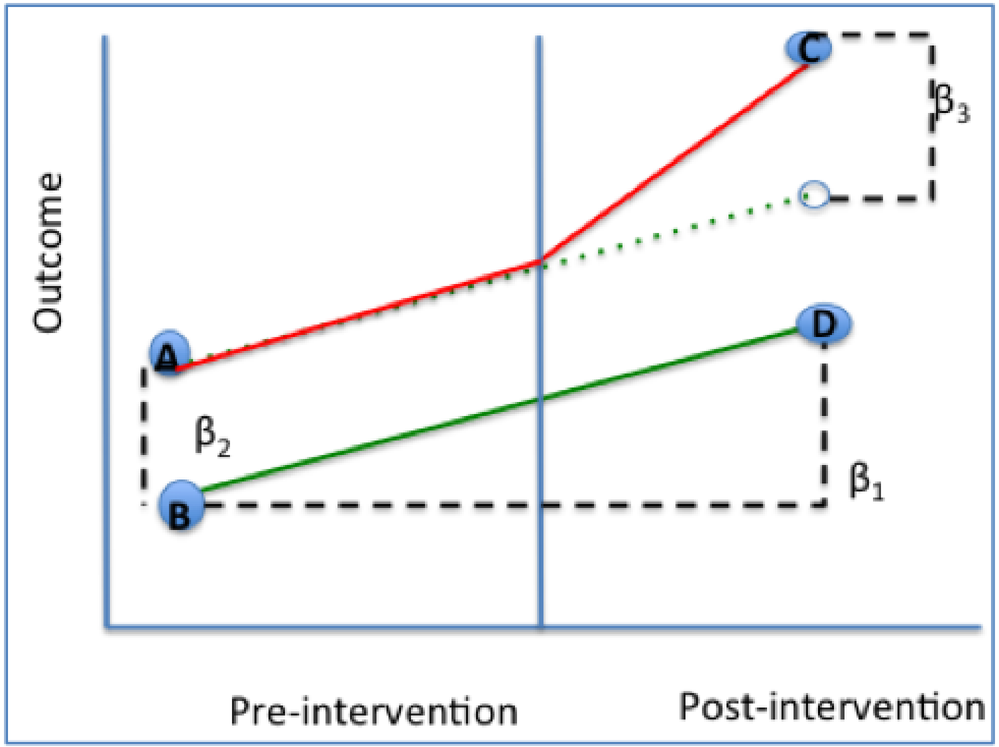
Approach

- Collect a pseudo-random sample (10%) of all Tweets from 2013
- Search for specific “sensitive” keywords
- Chec ve keyw to “non
- Test ’s reve

Statistical Machine Learning techniques for anomaly detection

Econometric techniques for causal inference (diff-in-diff models)

Methodology



Econometric model

$$Tweet\_volume_{ijt} = \beta_1 \, DHS_i \times Post\_Prism_t + \gamma_i + \delta_j + \theta_t$$

$$Tweet\_volume_{ijt} = \beta_1 \, DHS_i \times Post\_Prism_t + \beta_2 \, Popularity_i + \beta_3 \, Popularity_i \times DHS_i + \beta_4 \, Popularity_i \times Post\_Prism_t + \beta_5 \, Popularity_i \times DHS_i \times Post\_Prism_t + \gamma_i + \delta_j + \theta_t$$

Methodology

- Multivariate Linear Time Subset Scanning algorithm (MVLTSS, Neill, McFowland, & Zheng, 2013)

Methodology

- **Multivariate** Linear Time Subset Scanning algorithm
- Integrates information from different data streams, or many keywords

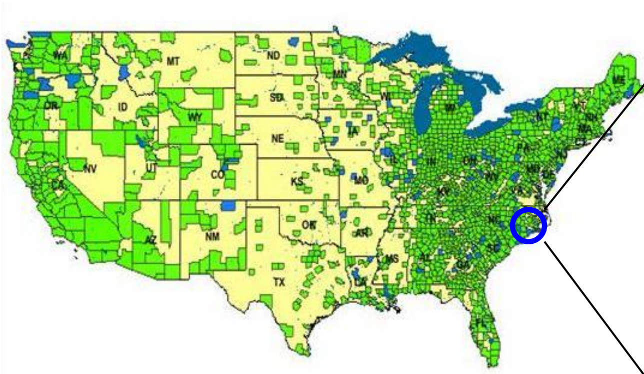
Methodology

- Multivariate **Linear Time** Subset Scanning algorithm
- Efficient method: search speed only grows linearly with the number of observations, rather than exponentially

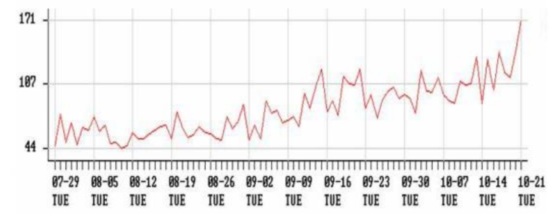
Methodology

- Multivariate Linear Time **Subset Scanning** algorithm
- Searches over subsets of the data

Methodology



Spatial time series data from spatial locations s_i (e.g. States)



Time series of counts $c_{i,m}^t$ for each State s_i for each data stream d_m (keyword).

<u>Self-censoring detection</u>

d_1 = influenza

d_2 = chemical

d_3 = gas

d_4 = screening

(etc.)

<u>Compare hypotheses</u>:

$H_1(D, S, W)$

D = subset of streams
S = subset of locations
W = time duration

vs. H_0: no events occurring

Eller College of Management

Methodology

- Compare observed data to expected (historical) data, and identify statistically significant anomalies in the observed data at a given time

Observed

	Keyword 1	Keyword 2	Keyword 3
State A	Count_{1A}	Count_{2A}	Count_{3A}
State B	Count_{1B}	Count_{2B}	Count_{3B}

Expected

	Keyword 1	Keyword 2	Keyword 3
State A	Baseline_{1A}	Baseline_{2A}	Baseline_{3A}
State B	Baseline_{1B}	Baseline_{2B}	Baseline_{3B}

- Find the most anomalous subset(s) of locations

Methodology

Observed

	Keyword 1	Keyword 2	Keyword 3
State A	$Count_{1A}$	$Count_{2A}$	$Count_{3A}$
State B	$Count_{1B}$	$Count_{2B}$	$Count_{3B}$

In this simple case, finding the most anomalous subset is relatively easy – we only have $O(2^6)$ possible subsets.
Imagine doing this for ~800 keywords, 50 States, and 52 weeks…
How to efficiently discover most anomalous subsets?

Methodology

- Use a priority function G to **order** locations: $G = f\left(\frac{Observed\ Count}{Expected\ Count}\right)$
- Use a scoring function to **assign a score** to subsets of cells that appear in the top k positions, as k=1..N
- We can ignore all other 2^N - N subsets because we can guarantee that the highest scoring subset will be one of those that we do evaluate
  - Intuition: for example, subset composed of 1st, 2nd, and 3rd ordered location is guaranteed to be more anomalous than subset composed of 1st, 2nd, and 7th location
- This is LTSS

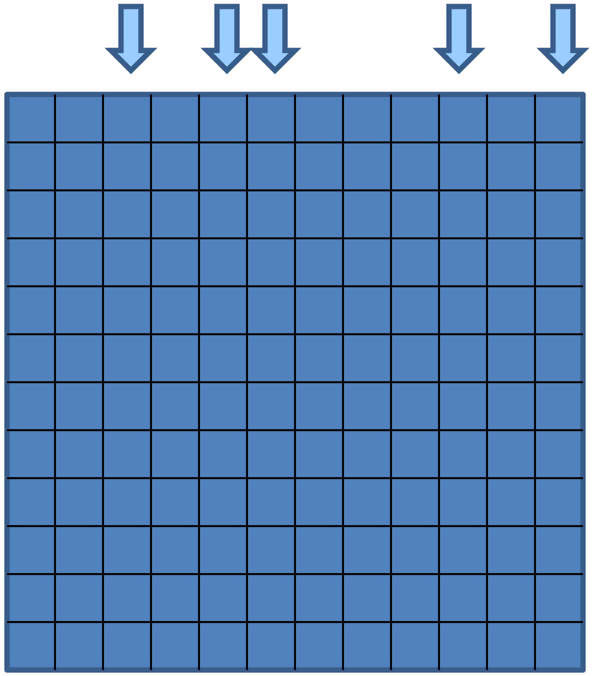
Methodology

- If we only had 1 keyword:
- Sort locations from highest to lowest priority…
- …then search over groups consisting of the top-k highest priority locations, for k = 1..N
- The highest priority subset is guaranteed to be one of these

Methodology

With many keywords:

1. Start with a randomly chosen subset of streams

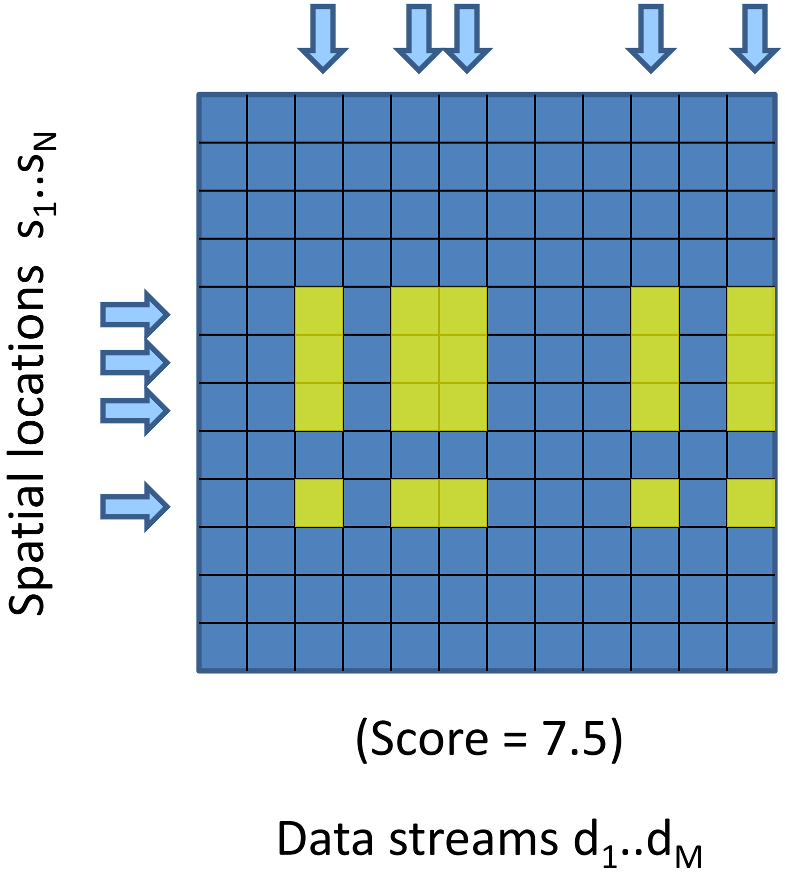


Eller College of Management

Methodology

With many keywords:

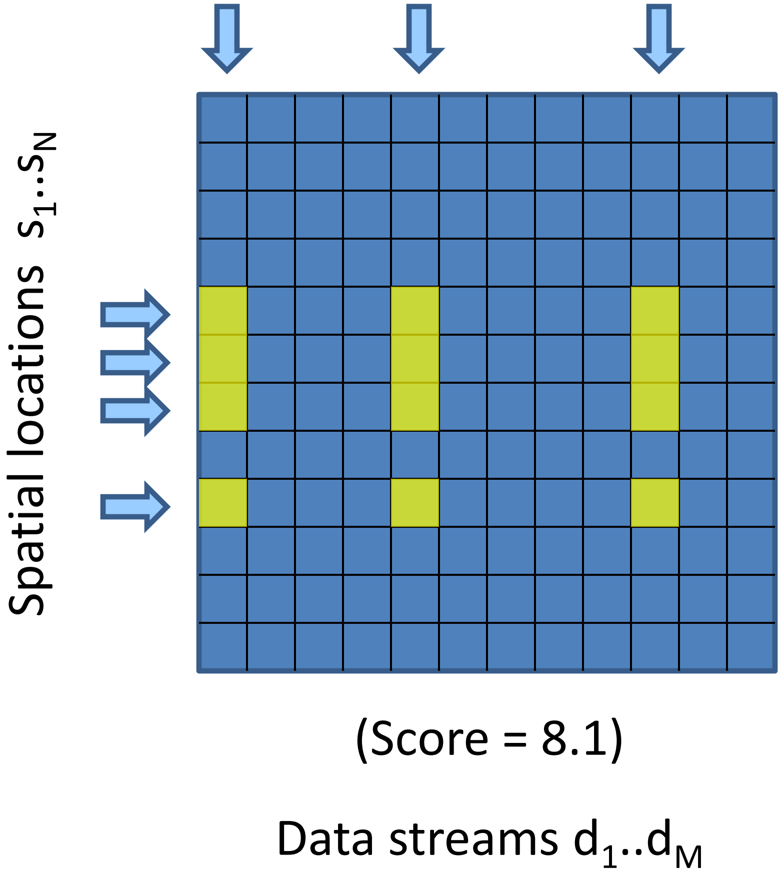
1. Start with a randomly chosen subset of streams
2. Use LTSS to efficiently find the highest-scoring subset of locations for the given streams



Methodology

With many keywords:

1. Start with a randomly chosen subset of streams
2. Use LTSS to efficiently find the highest-scoring subset of locations for the given streams
3. Use LTSS to efficiently find the highest-scoring subset of streams for the given locations

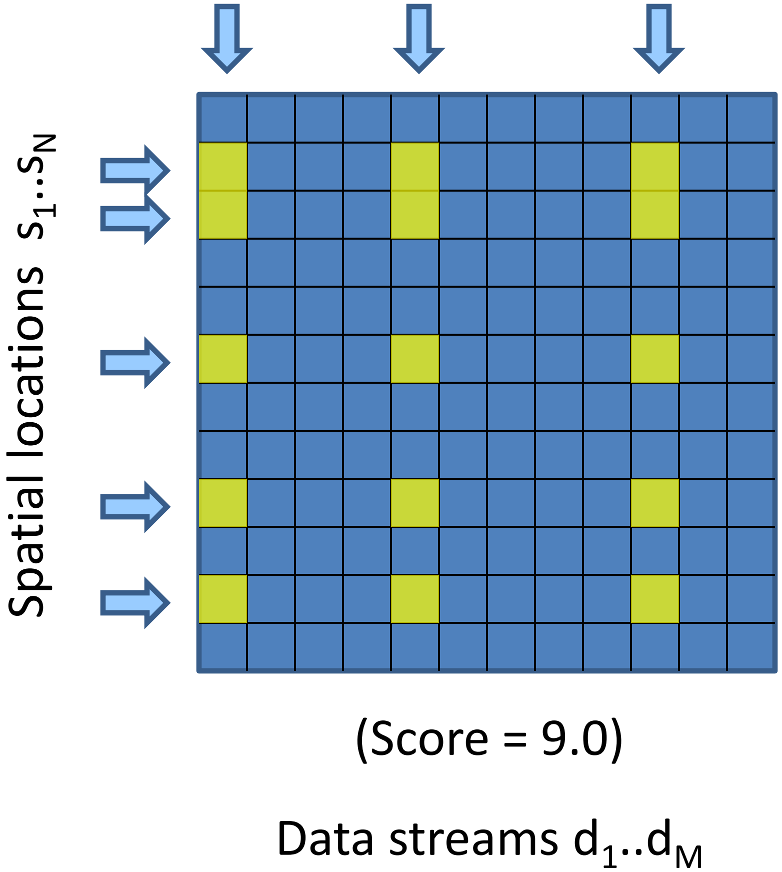


Eller College of Management

Methodology

With many keywords:

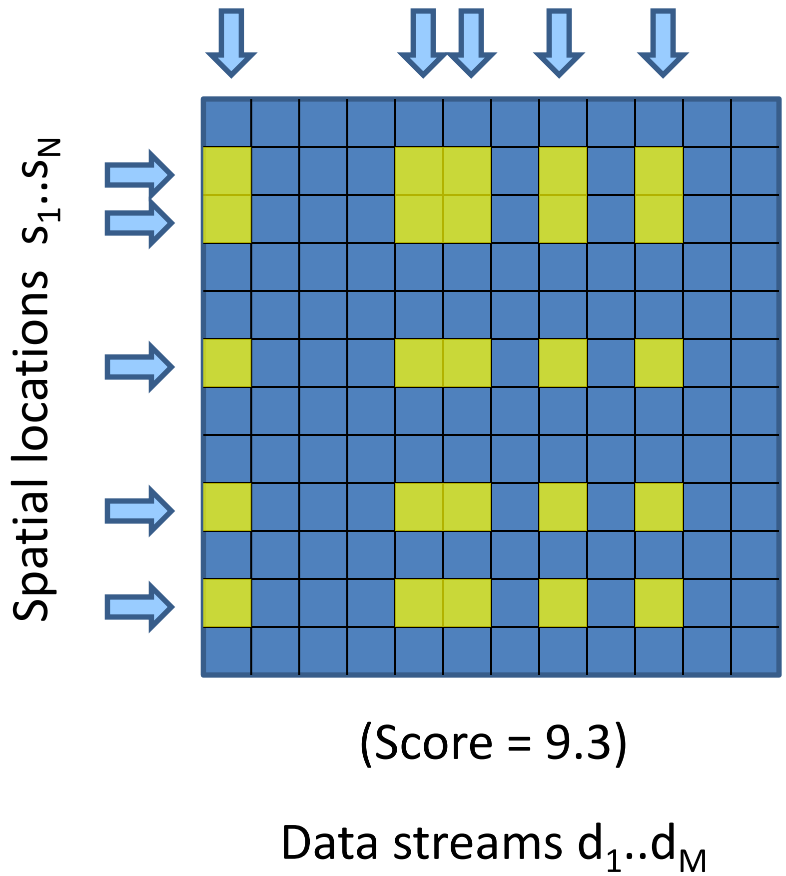
1. Start with a randomly chosen subset of streams
2. Use LTSS to efficiently find the highest-scoring subset of locations for the given streams
3. Use LTSS to efficiently find the highest-scoring subset of streams for the given locations
4. Iterate steps 2-3 until convergence



Methodology

With many keywords:

1. Start with a randomly chosen subset of streams
2. Use LTSS to efficiently find the highest-scoring subset of locations for the given streams
3. Use LTSS to efficiently find the highest-scoring subset of streams for the given locations
4. Iterate steps 2-3 until convergence

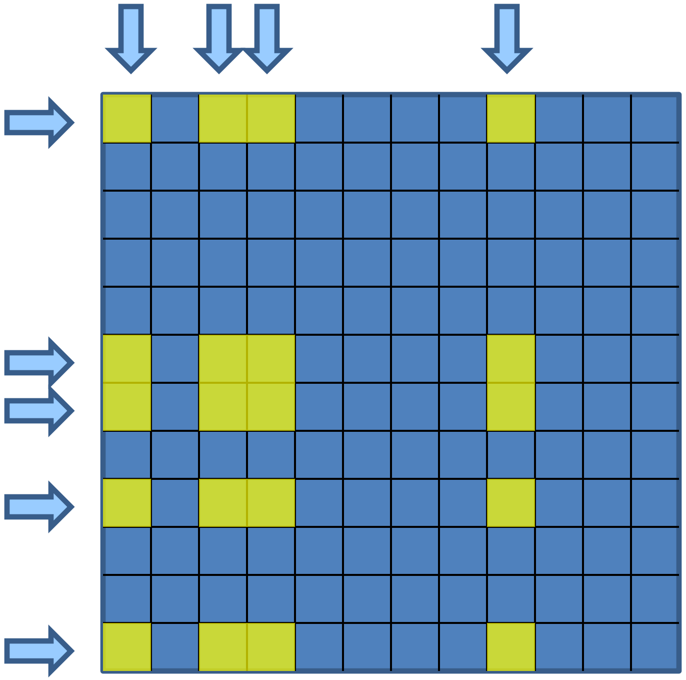


Methodology

With many keywords:

1. Start with a randomly chosen subset of streams
2. Use LTSS to efficiently find the highest-scoring subset of locations for the given streams
3. Use LTSS to efficiently find the highest-scoring subset of streams for the given locations
4. Iterate steps 2-3 until convergence
5. Repeat steps 1-4 for many random restarts

Spatial locations $s_1..s_N$



(Score = 11.0)

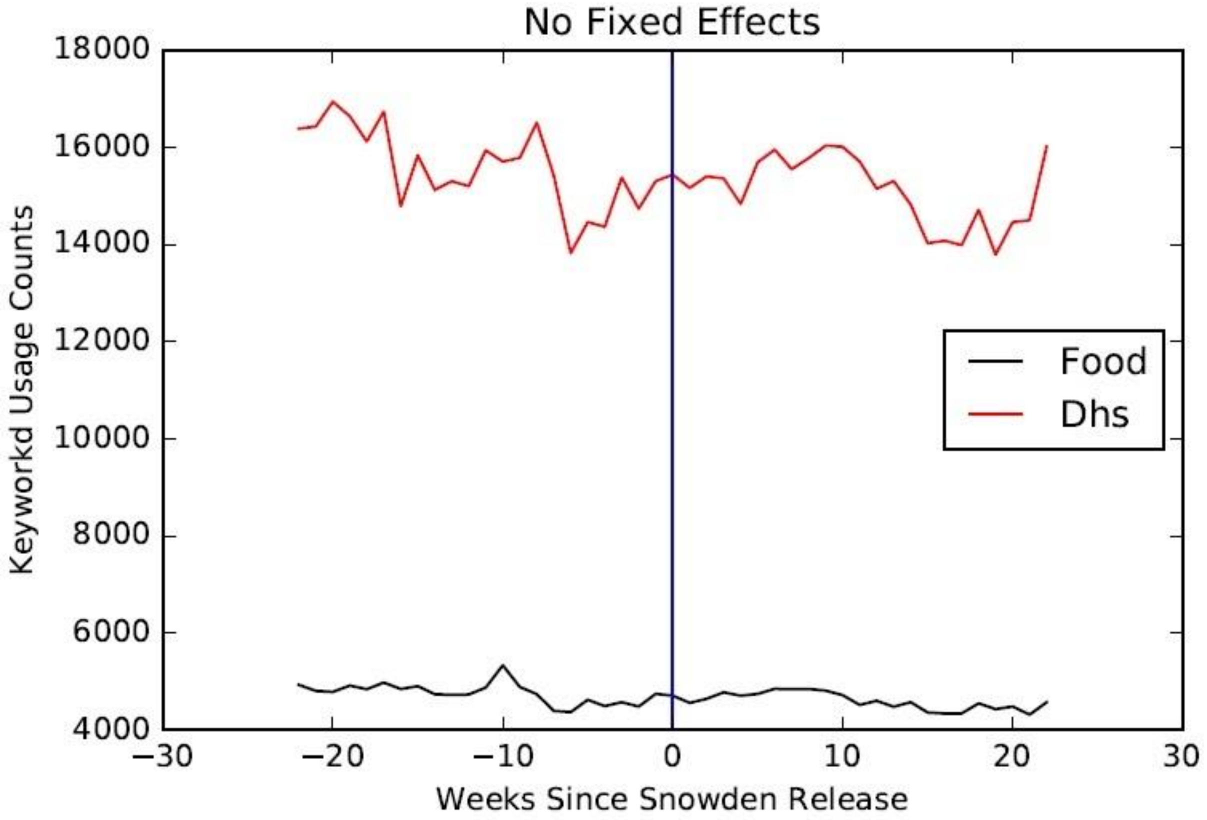
Data streams $d_1..d_M$

Data

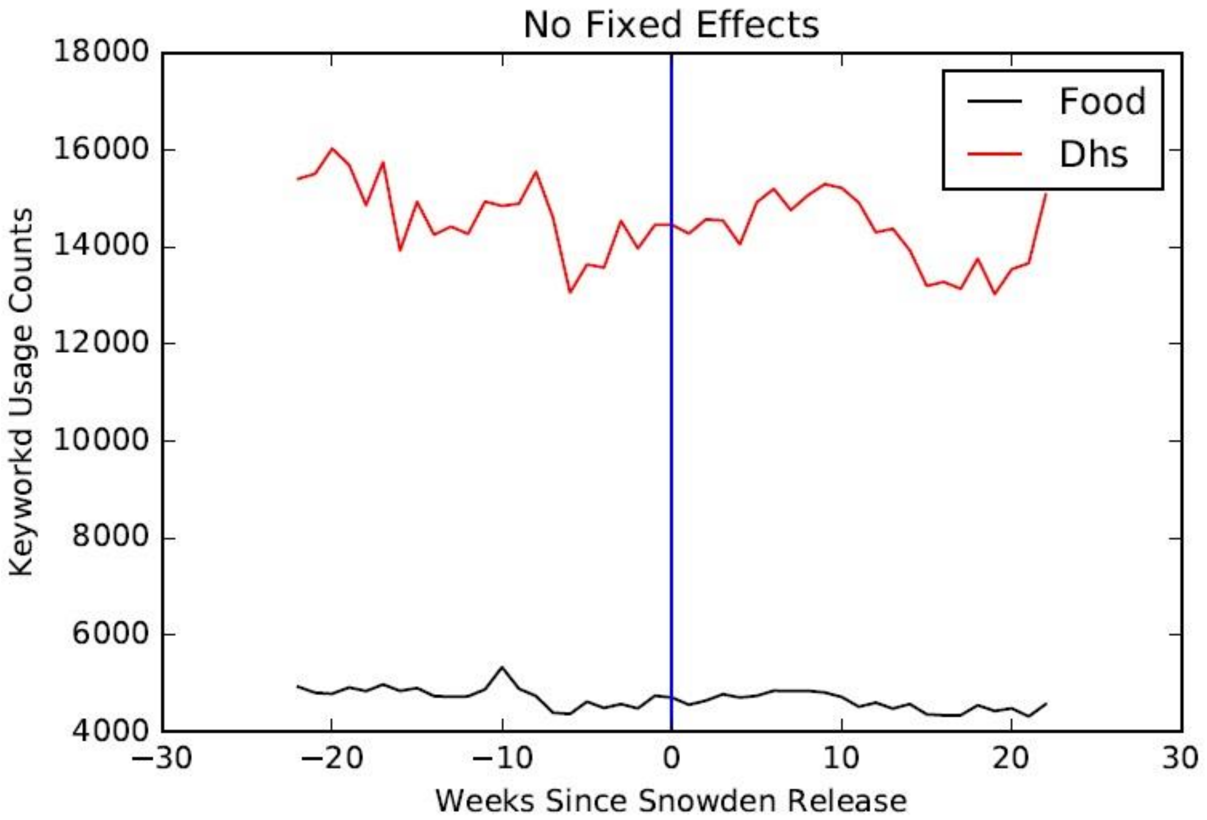
- 12 months **(~80 TB)** of data **(18 bln Tweets)**: 22 weeks before and 29 weeks after June 6, 2013
  - week -22 starts Jan 2
  - week 0: June 6-12
  - week 29 starts Dec 26

Eller College of Management

Preliminary results



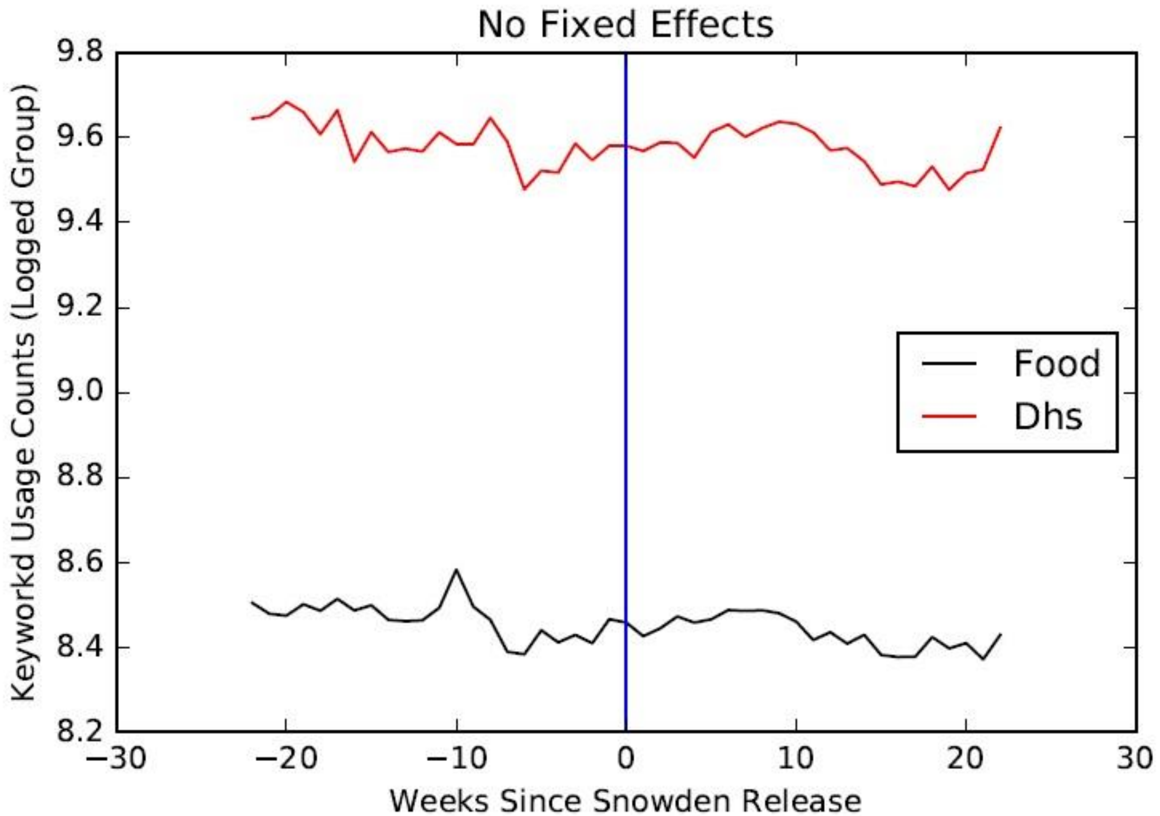
Preliminary results – no spike words



DHS spike words: 'colombia','help','mexico', 'north korea', 'power', 'recovery', 'sick','snow','storm','tornado','typhoon','watch','who'
Food spike words: 'chocolate', 'egg', 'pumpkin'

Eller College of Management

Preliminary results – logs



Preliminary results – MVLTSS

Effect most significant over weeks 6 through 12 after the first revelations, and specifically in thirty-six US States – “red states” unaffected

Preliminary results – econometrics approach

Table 2. Regression coefficients on logged volume of keywords. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) No Popularity
DHS*Post_Prism	-0.0103**
	(0.00407)
Popularity	
Popularity*Post_Prism	
Popularity*DHS	
Popularity*Post_Prism*DHS	
US	
Non-US	
US*Post-Prism	
Non-US*Post-Prism	
US*DHS	
Non-US*DHS	
US*DHS*Post-Prism	
Non-US*DHS*Post-Prism	
Blue_State*Post_Prism	
Blue_State*DHS	
Blue_State*Post_Prism*DHS	
Keyword F.E.	Yes
State F.E.	Yes
Week F.E.	Yes
Observations	4,473,200
R-squared	0.727

Preliminary results – econometrics approach

Table 2. Regression coefficients on logged volume of keywords. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) No Popularity	(2) Popularity - Overall
DHS*Post_Prism	-0.0103**	-0.00656**
	(0.00407)	(0.00313)
Popularity		0.140***
		(0.00986)
Popularity*Post_Prism		0.0516***
		(0.00478)
Popularity*DHS		-0.0121
		(0.0145)
Popularity*Post_Prism*DHS		-0.00749
		(0.00459)
US		
Non-US		
US*Post-Prism		
Non-US*Post-Prism		
US*DHS		
Non-US*DHS		
US*DHS*Post-Prism		
Non-US*DHS*Post-Prism		
Blue_State*Post_Prism		
Blue_State*DHS		
Blue_State*Post_Prism*DHS		
Keyword F.E.	Yes	Yes
State F.E.	Yes	Yes
Week F.E.	Yes	Yes
Observations	4,473,200	4,473,200
R-squared	0.727	0.733

Preliminary results – econometrics approach

Table 2. Regression coefficients on logged volume of keywords. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) No Popularity	(2) Popularity - Overall	(3) Popularity - US	(4) Popularity - Non-US	(5) Popularity - No info
DHS*Post_Prism	-0.0103**	-0.00656**	-0.00801***	0.0147	0.0461
	(0.00407)	(0.00313)	(0.00297)	(0.0147)	(0.0276)
Popularity		0.140***	0.118***	0.556***	0.873***
		(0.00986)	(0.00988)	(0.0218)	(0.0139)
Popularity*Post_Prism		0.0516***	0.0476***	0.173***	0.129***
		(0.00478)	(0.00463)	(0.0157)	(0.0121)
Popularity*DHS		-0.0121	-0.00764	-0.0940***	-0.157***
		(0.0145)	(0.0146)	(0.0295)	(0.0216)
Popularity*Post_Prism*DHS		-0.00749	-0.00649	-0.0253*	-0.0408**
		(0.00459)	(0.00466)	(0.0146)	(0.0167)
US					
Non-US					
US*Post-Prism					
Non-US*Post-Prism					
US*DHS					
Non-US*DHS					
US*DHS*Post-Prism					
Non-US*DHS*Post-Prism					
Blue_State*Post_Prism					
Blue_State*DHS					
Blue_State*Post_Prism*DHS					
Keyword F.E.	Yes	Yes	Yes	Yes	Yes
State F.E.	Yes	Yes	Yes	-	-
Week F.E.	Yes	Yes	Yes	Yes	Yes
Observations	4,473,200	4,473,200	4,304,400	84,400	84,400
R-squared	0.727	0.733	0.591	0.933	0.973

Preliminary results – econometrics approach

Table 2. Regression coefficients on logged volume of keywords. Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

VARIABLES	(1) No Popularity	(2) Popularity - Overall	(3) Popularity - US	(4) Popularity - Non-US	(5) Popularity - No info	(6) Location Interaction	(7) Political affiliation
DHS*Post_Prism	-0.0103**	-0.00656**	-0.00801***	0.0147	0.0461	0.0295***	-0.00557**
	(0.00407)	(0.00313)	(0.00297)	(0.0147)	(0.0276)	(0.00345)	(0.00259)
Popularity		0.140***	0.118***	0.556***	0.873***	0.140***	0.118***
		(0.00986)	(0.00988)	(0.0218)	(0.0139)	(0.00993)	(0.00988)
Popularity*Post_Prism		0.0516***	0.0476***	0.173***	0.129***	0.0516***	0.0476***
		(0.00478)	(0.00463)	(0.0157)	(0.0121)	(0.00656)	(0.00467)
Popularity*DHS		-0.0121	-0.00764	-0.0940***	-0.157***	-0.0121	-0.00764
		(0.0145)	(0.0146)	(0.0295)	(0.0216)	(0.0144)	(0.0146)
Popularity*Post_Prism*DHS		-0.00749	-0.00649	-0.0253*	-0.0408**	-0.00749	-0.00649
		(0.00459)	(0.00466)	(0.0146)	(0.0167)	(0.00545)	
US						-5.329***	
						(0.0899)	
Non-US						-4.037***	
						(0.0509)	
US*Post-Prism						0.117***	
						(0.00534)	
Non-US*Post-Prism						0.183***	
						(0.0279)	
US*DHS						0.272	
						(0.163)	
Non-US*DHS						0.258**	
						(0.0975)	
US*DHS*Post-Prism						-0.0370***	
						(0.00413)	
Non-US*DHS*Post-Prism						-0.0236***	
						(0.00865)	
Blue_State*Post_Prism							0.0114***
							(0.00140)
Blue_State*DHS							-0.00905
							(0.00729)
Blue_State*Post_Prism*DHS							-0.0046***
							(0.00170)
Keyword F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State F.E.	Yes	Yes	Yes	-	-	-	Yes
Week F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,473,200	4,473,200	4,304,400	84,400	84,400	4,473,200	4,304,400
R-squared	0.727	0.733	0.591	0.933	0.973	0.717	0.591

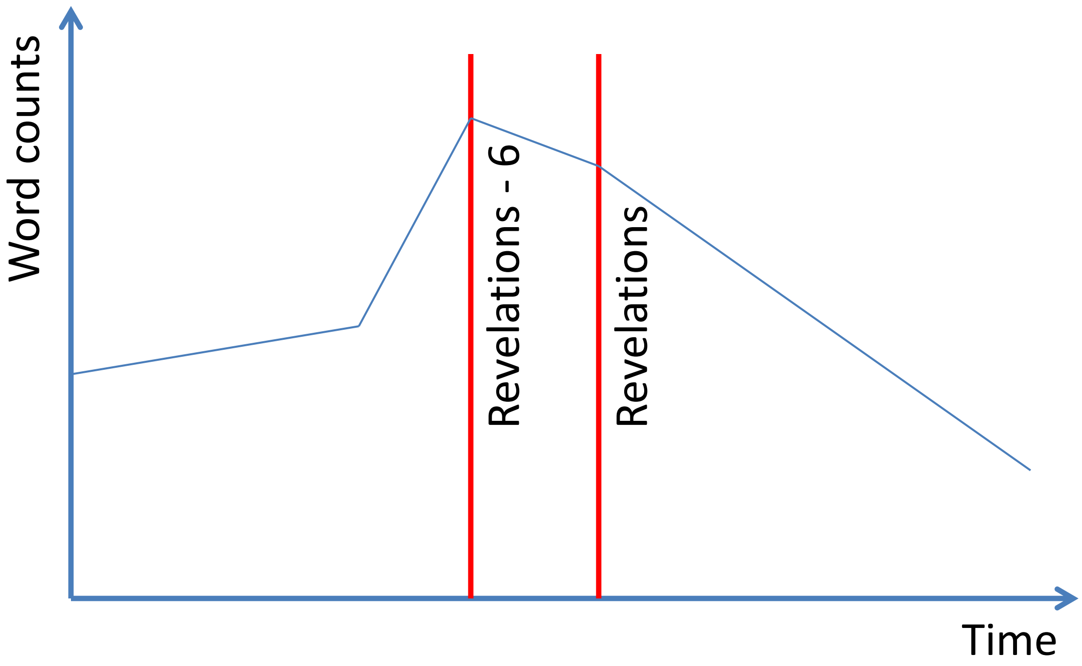
Preliminary results – econometrics approach

Table 3. Falsification test.

VARIABLES	(1) ltweet_volume
DHS*Post_Prism	
DHS*Fake_Post_Prism	-0.00591*
	(0.00312)
Popularity	0.139***
	(0.00970)
Popularity*Fake_Post_Prism	0.0518***
	(0.00462)
Popularity*DHS	-0.0117
	(0.0143)
Popularity*Fake_Post_Prism*DHS	-0.00802*
	(0.00462)
Observations	4,562,664
R-squared	0.733

Eller College of Management

Preliminary results – econometrics approach



To-do

- Specific subset of words (e.g., not considering weather-related keywords)
- Specific locations
- Train MVLTSS on a subset of data (to avoid overfitting) and estimate causal modal on the complement

Implications

- Perception of security facilitates free speech, which is not only an ethical principle: it also has practical economic consequences
- Surveillance may harm the US economy
  - The very existence of most virtual communities hinges upon its members being active contributors (e.g., online health communities)
- The role of social media as a means for social and political organization, or for support at times of crisis, can only be maintained in (perceived) safe environments
  - Online communities discussing health-related issues
  - Arab spring

Thanks for listening!

Questions?

lbrandimarte@email.arizona.edu