

Quantifying the complex world we inhabit using big data

Tobias Preis

Tobias.Preis@wbs.ac.uk

A map of the world built only from
GPS locations of *Flickr* photos

Preis & Moat (2013)

Internet data opens up opportunities for
new dynamic, adaptive economic models

Decisions

Decisions

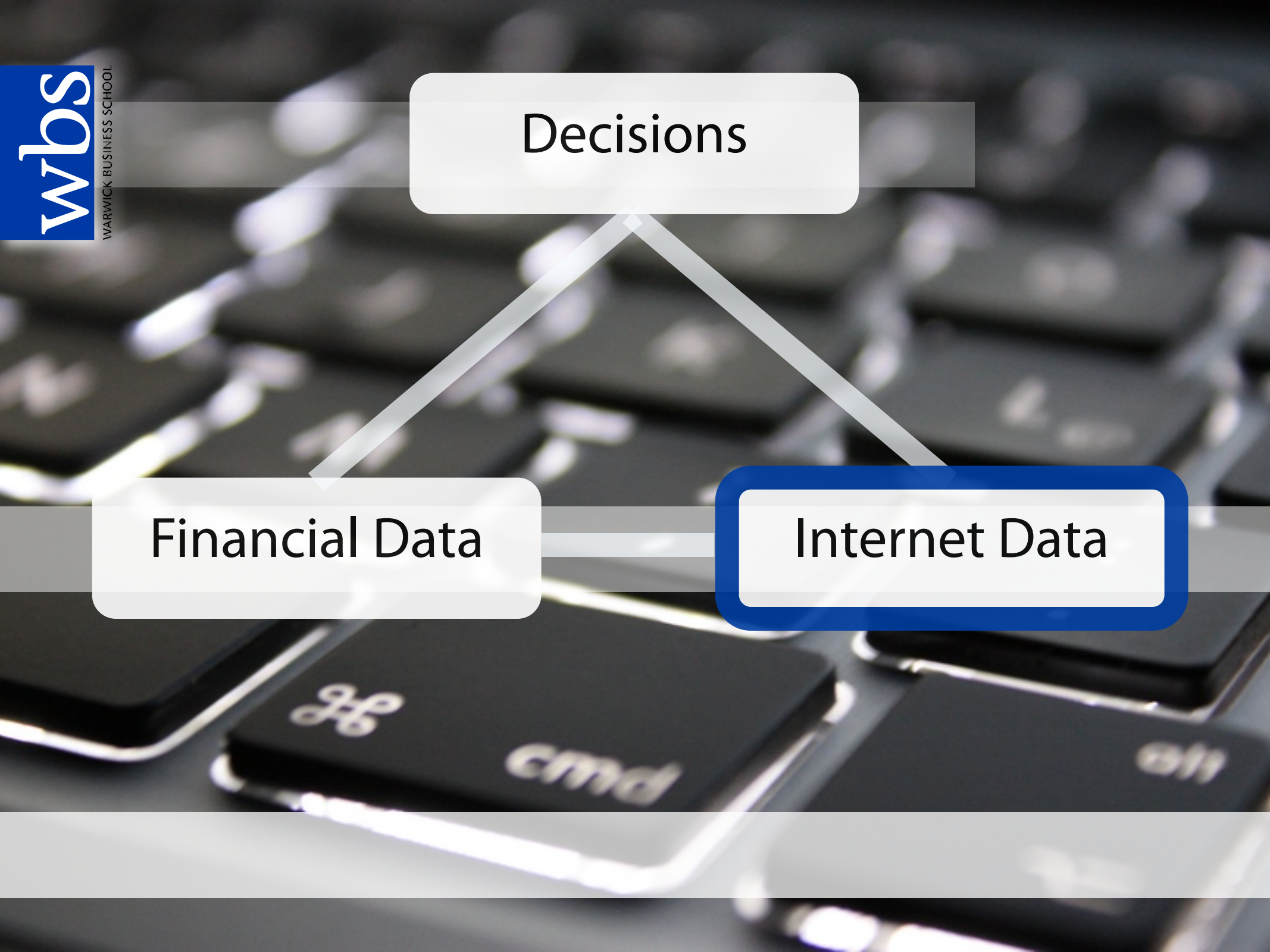
Financial Data



Decisions

Financial Data

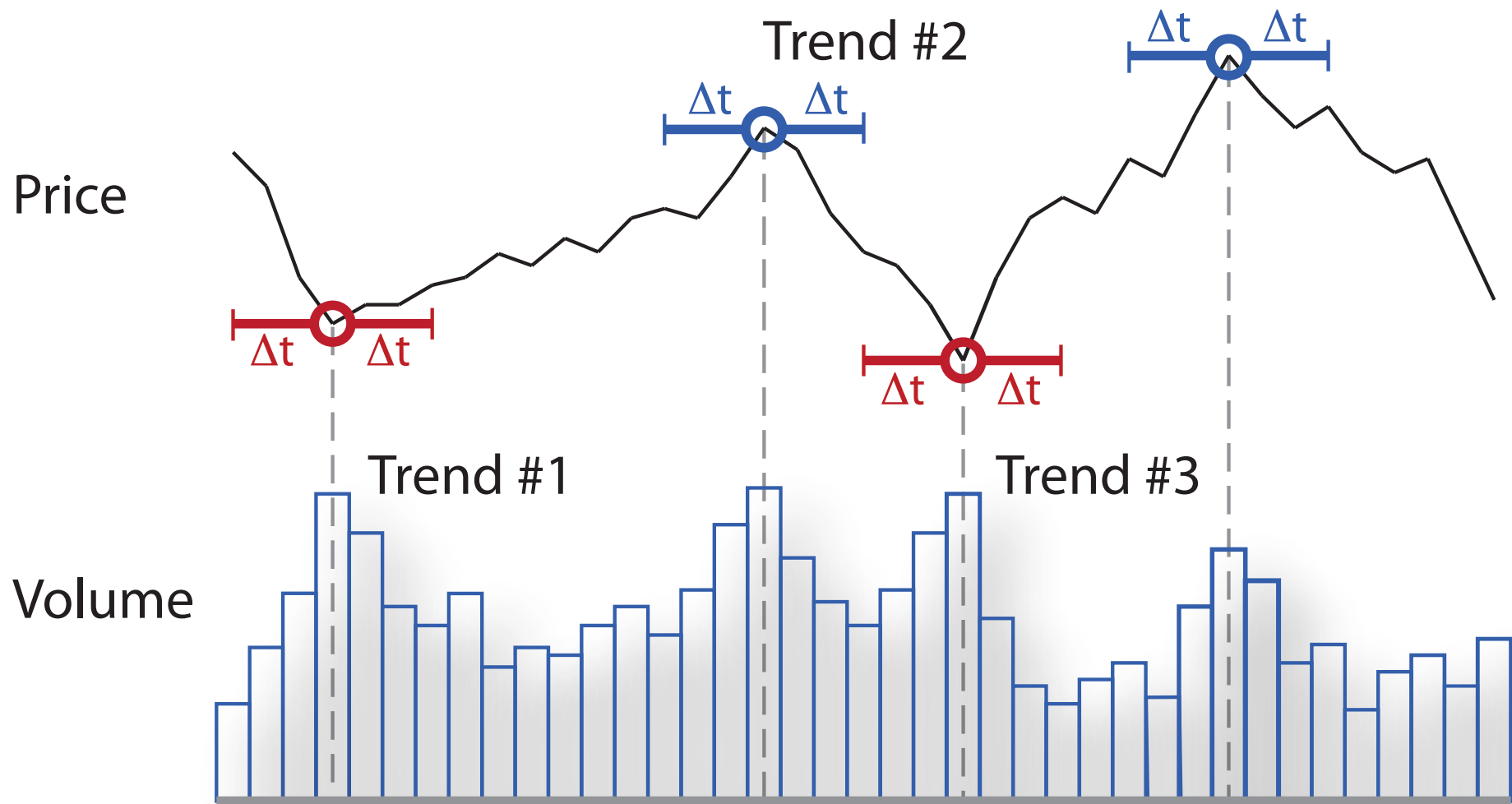
Internet Data



Can we quantify bubbles?

2155.00	-3.00	-0.14%	118	2175.00	2175.00	1.75%	9750	9750
17:32:30	-17.00	-0.78%	121	2175.00	2175.00	1.75%	9750	9750
27.065	-1.220	-4.31%	5,080	27.065	27.075	2.00%	2000	2000
1.905	-0.080	-4.03%	57,112	1.905	1.908	2.00%	2000	2000
11171	+51	0.46%	1	11171	11172	4	1000	1000
1.3621	-0.0018	-0.13%	9	1.3620	1.3621	10	1000	1000
1.3627 ⁸	+0.0034 ⁰	0.25%	4M	1.3627 ⁸	1.3629 ⁸	20	20	20
1.2205 ⁸	+0.0014 ⁵	0.12%	3M	1.2205 ⁸	1.2206 ⁸	20	20	20
0.8700 ⁸	+0.0007 ⁰	0.08%	1M	0.8700 ⁸	0.8700 ⁸	90	90	90
10.27	-0.21	-2.00%	221	10.26	10.27	20	20	20
63.20	+0.42	0.67%	2	63.21	63.24	20	20	20
6.43	-0.05	-0.77%	1,098	6.42	6.43	807	807	807
308.20	+12.70	4.30%	154	305.40	309.80	10	10	10
547.50	+13.90	2.60%	94	516.30	522.30	10	10	10
	+32.80	10.49%	50	323.70	328.30	10	10	10
		0.24%	94	341.30	345.80	10	10	10
			104					

Quantifying Stock Market Bubbles and Crashes

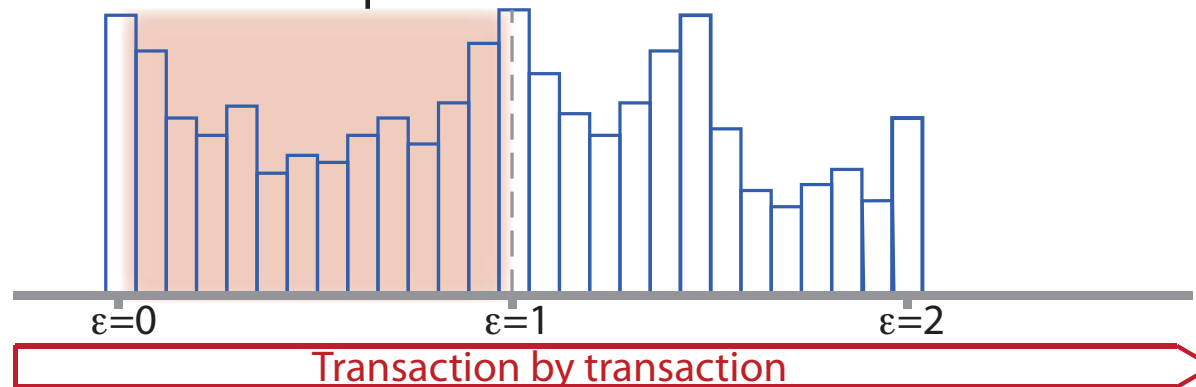


Tobias Preis and H. Eugene Stanley, Journal of Statistical Physics World **138**, 431-446 (2010)

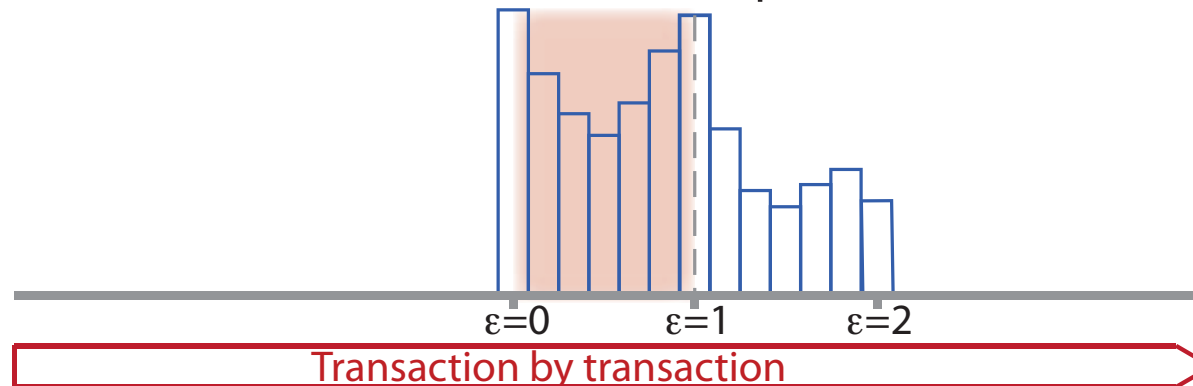
Quantifying Stock Market Bubbles and Crashes

Extraction of trend sequences

Volume sequence of trend #1



Volume sequence of trend #2

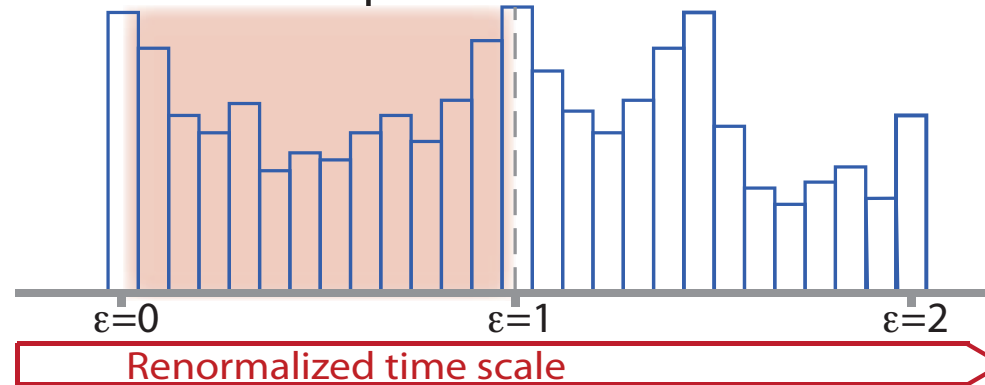


Tobias Preis and H. Eugene Stanley, Journal of Statistical Physics World **138**, 431-446 (2010)

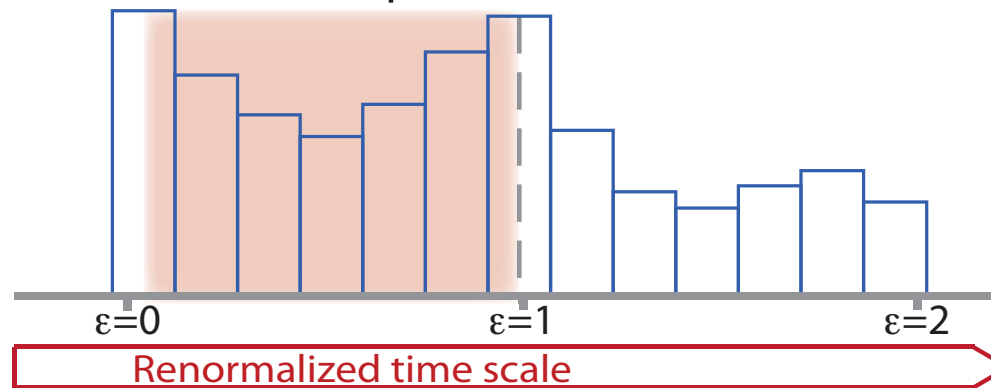
Quantifying Stock Market Bubbles and Crashes

Renormalization of trend sequences

Volume sequence of trend #1



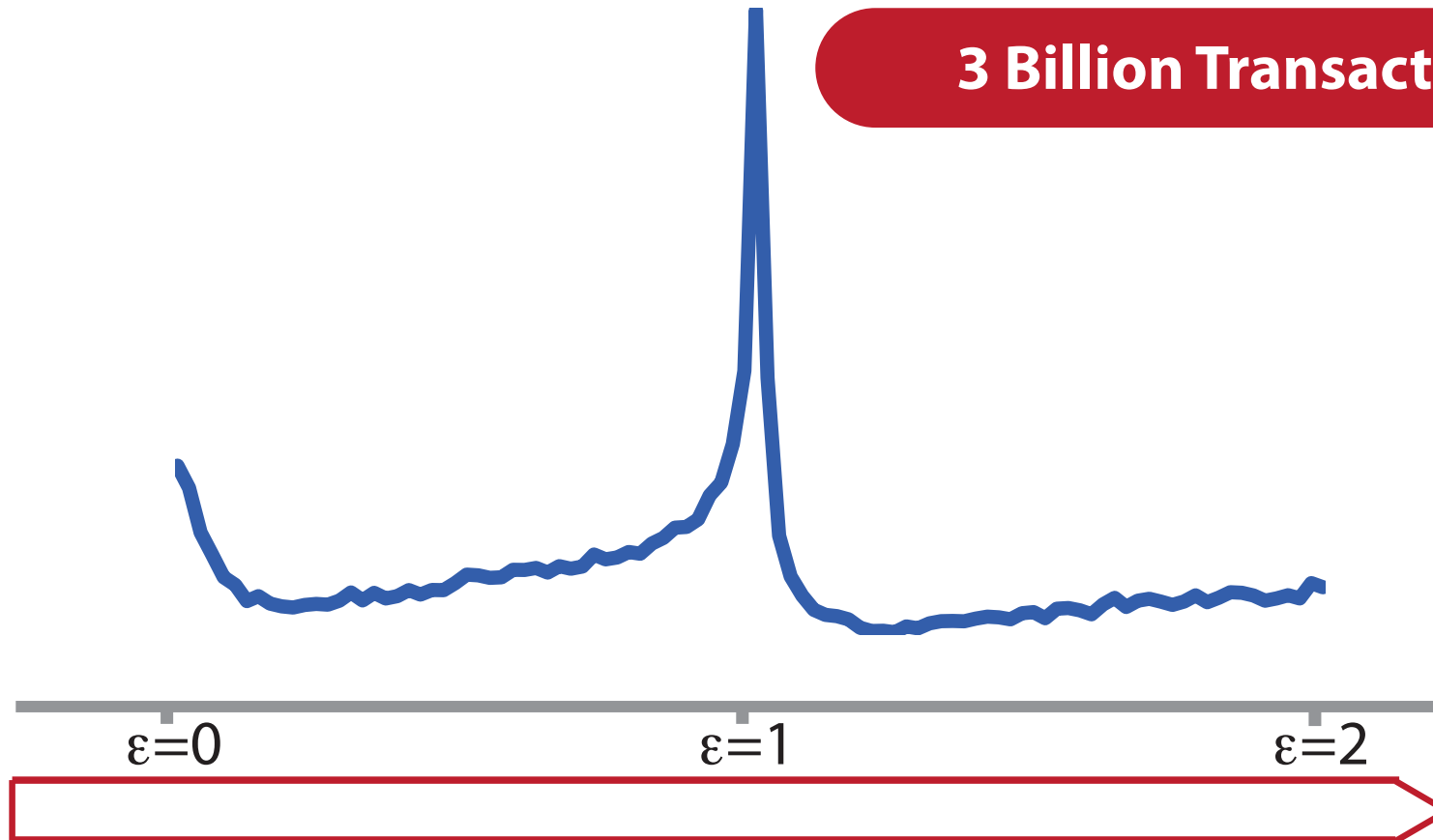
Volume sequence of trend #2



Tobias Preis and H. Eugene Stanley, Journal of Statistical Physics World **138**, 431-446 (2010)

A Law For Bubbles and Crashes

Averaged volume sequence for FDAX time series



Tobias Preis et al., PNAS **108**, 7674-7678 (2011)

Covered by

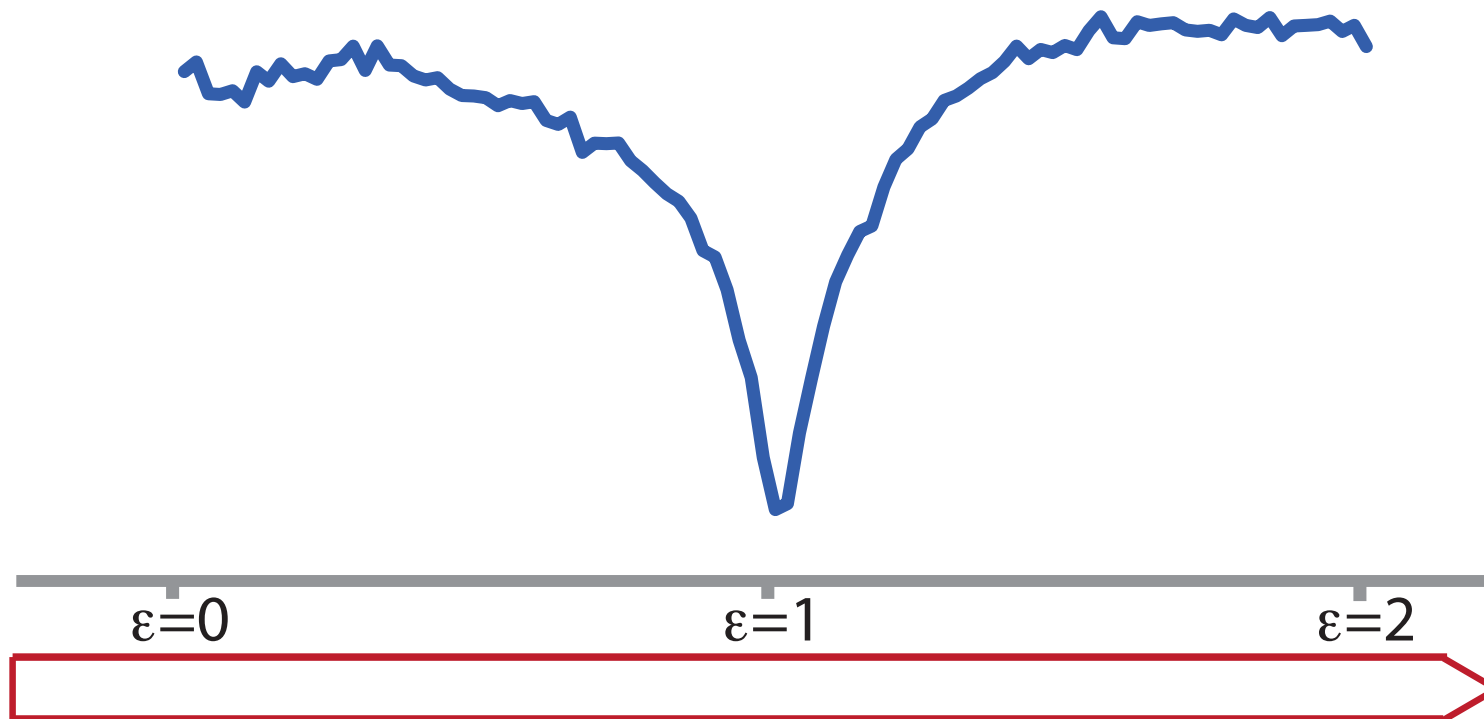
FINANCIAL TIMES
DEUTSCHLAND

THE WALL STREET JOURNAL

WIRED

A Law For Bubbles and Crashes

Averaged time interval sequence for FDAX time series



Tobias Preis et al., PNAS **108**, 7674-7678 (2011)

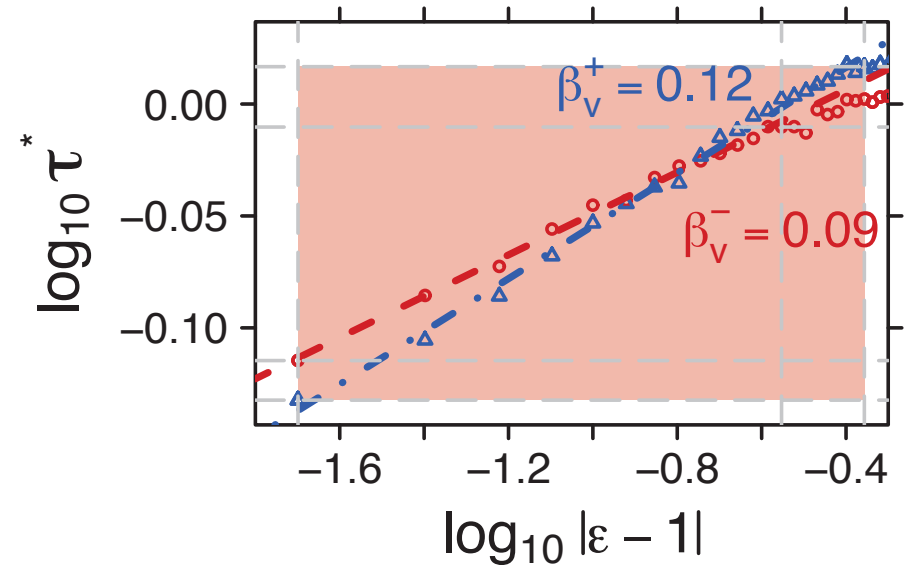
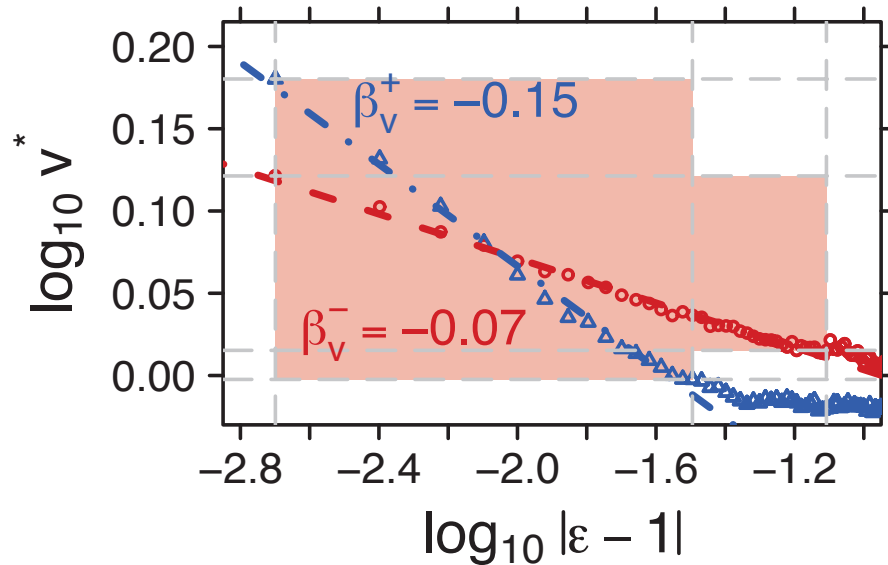
Covered by

FINANCIAL TIMES
DEUTSCHLAND

THE WALL STREET JOURNAL

WIRED

A Law For Bubbles and Crashes



Tobias Preis et al., PNAS **108**, 7674-7678 (2011)

Covered by

FINANCIAL TIMES
DEUTSCHLAND

THE WALL STREET JOURNAL

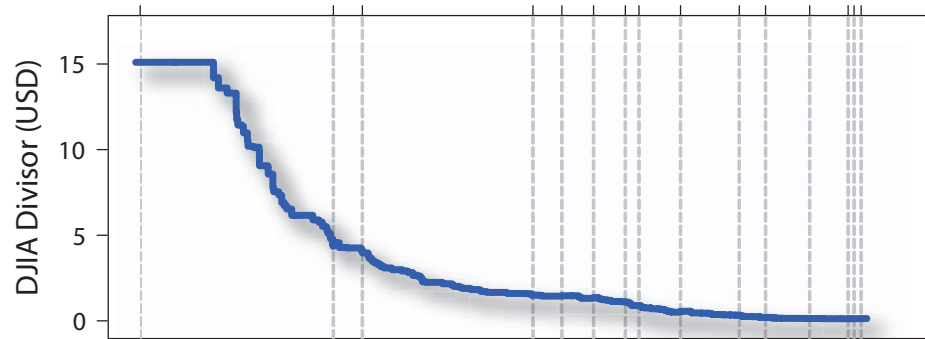
WIRED

Are Correlations Constant?

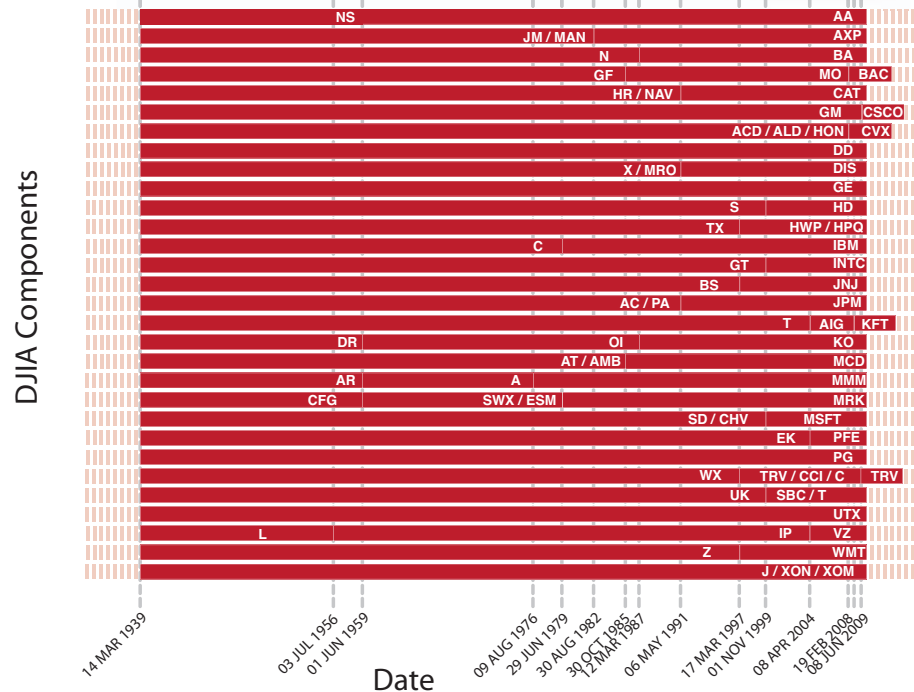


Quantifying Stock Market Correlations

A



B



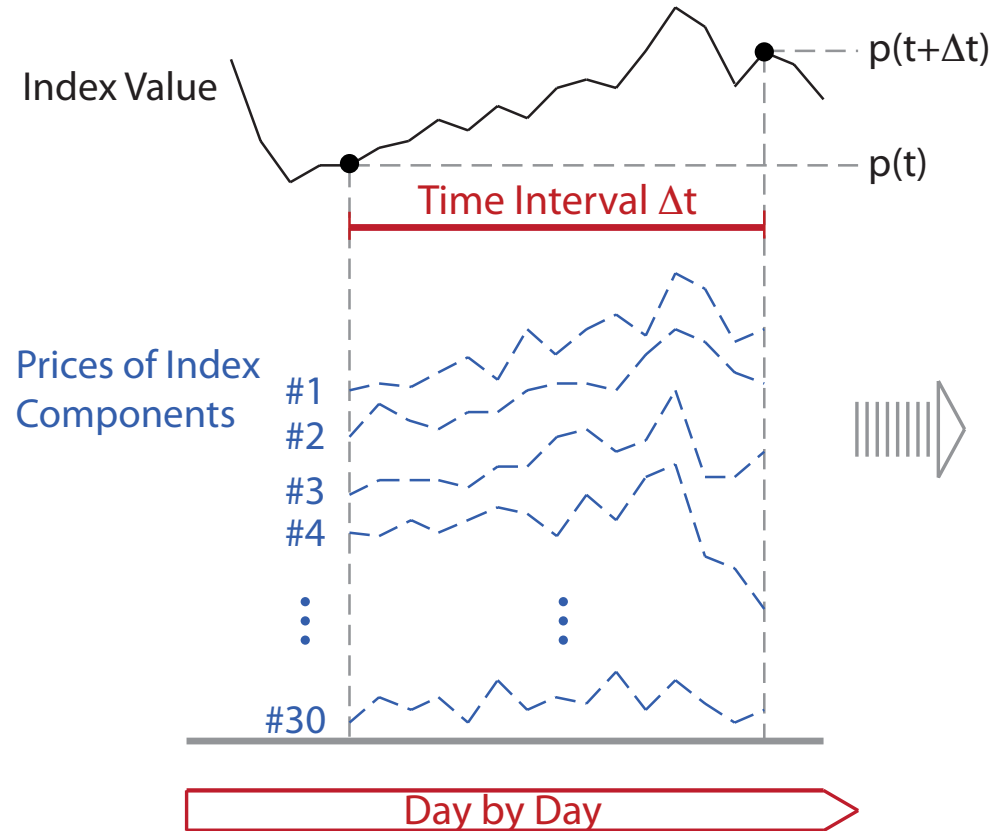
- 71 years of daily stock prices
- 30 DJIA index components
- Changes are explicitly considered

Preis, Kenett, Stanley, Helbing & Ben Jacob,
Scientific Reports 2, 752 (2012)

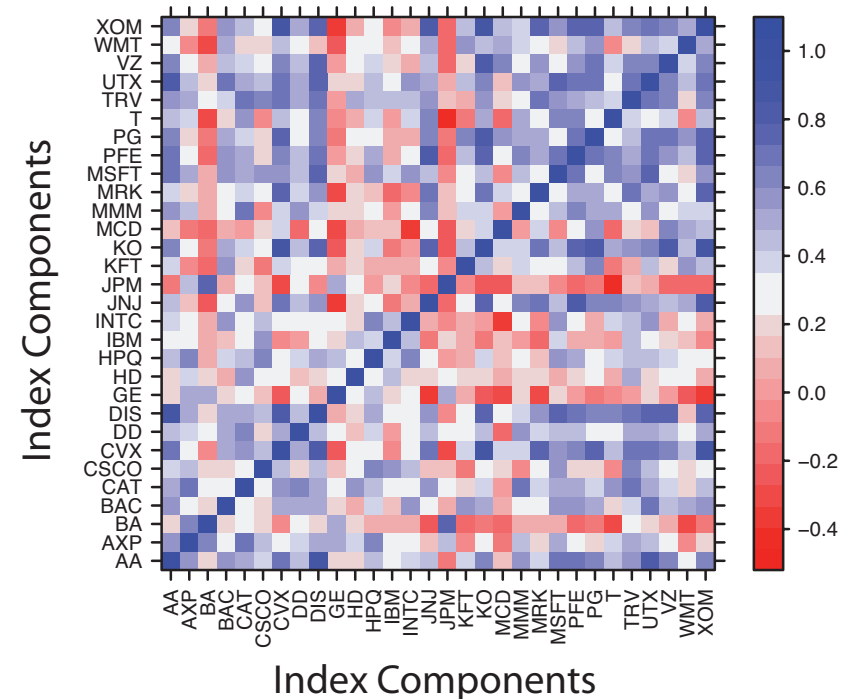
<http://www.nature.com/srep/2012/121018/srep00752/full/srep00752.html>

Quantifying Stock Market Correlations

A Dow Jones Industrial Average and its Components

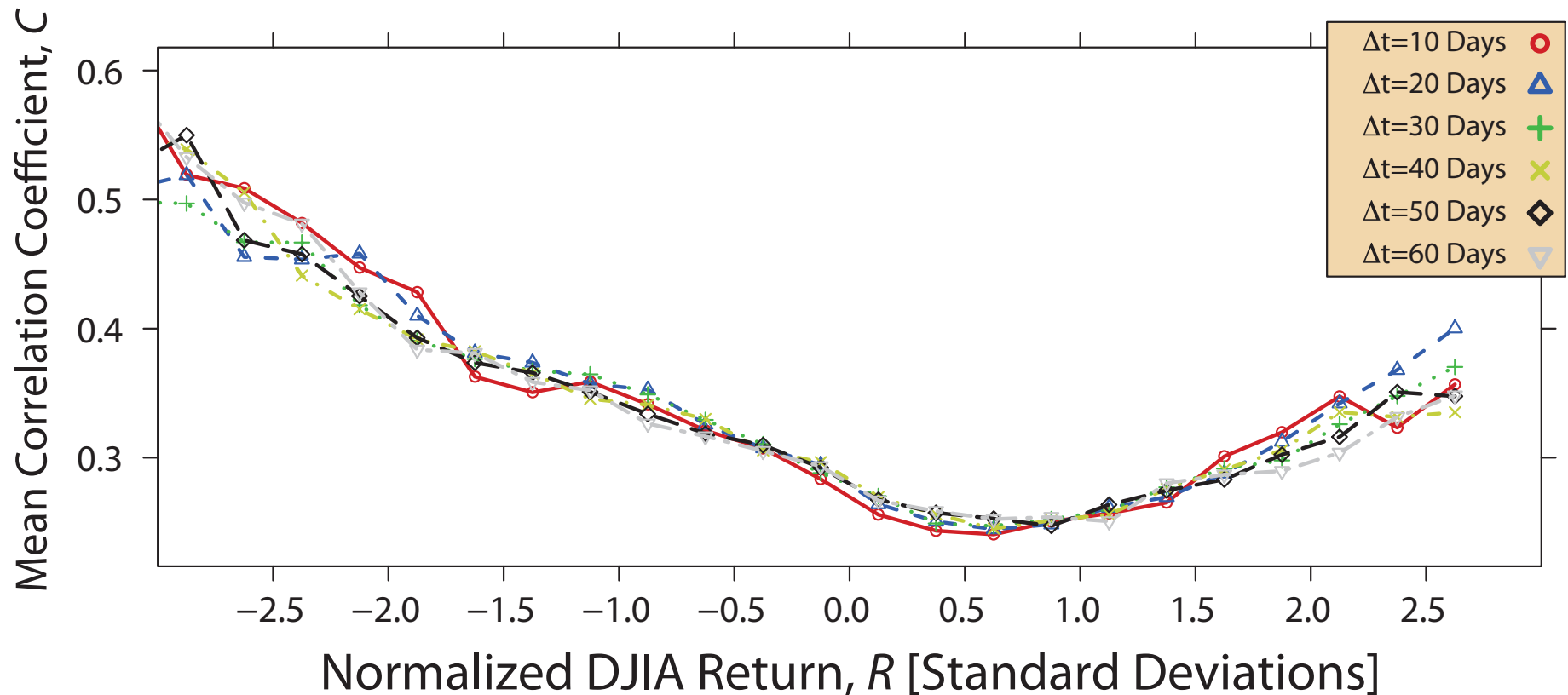


B Matrix of Pearson Correlation Coefficients



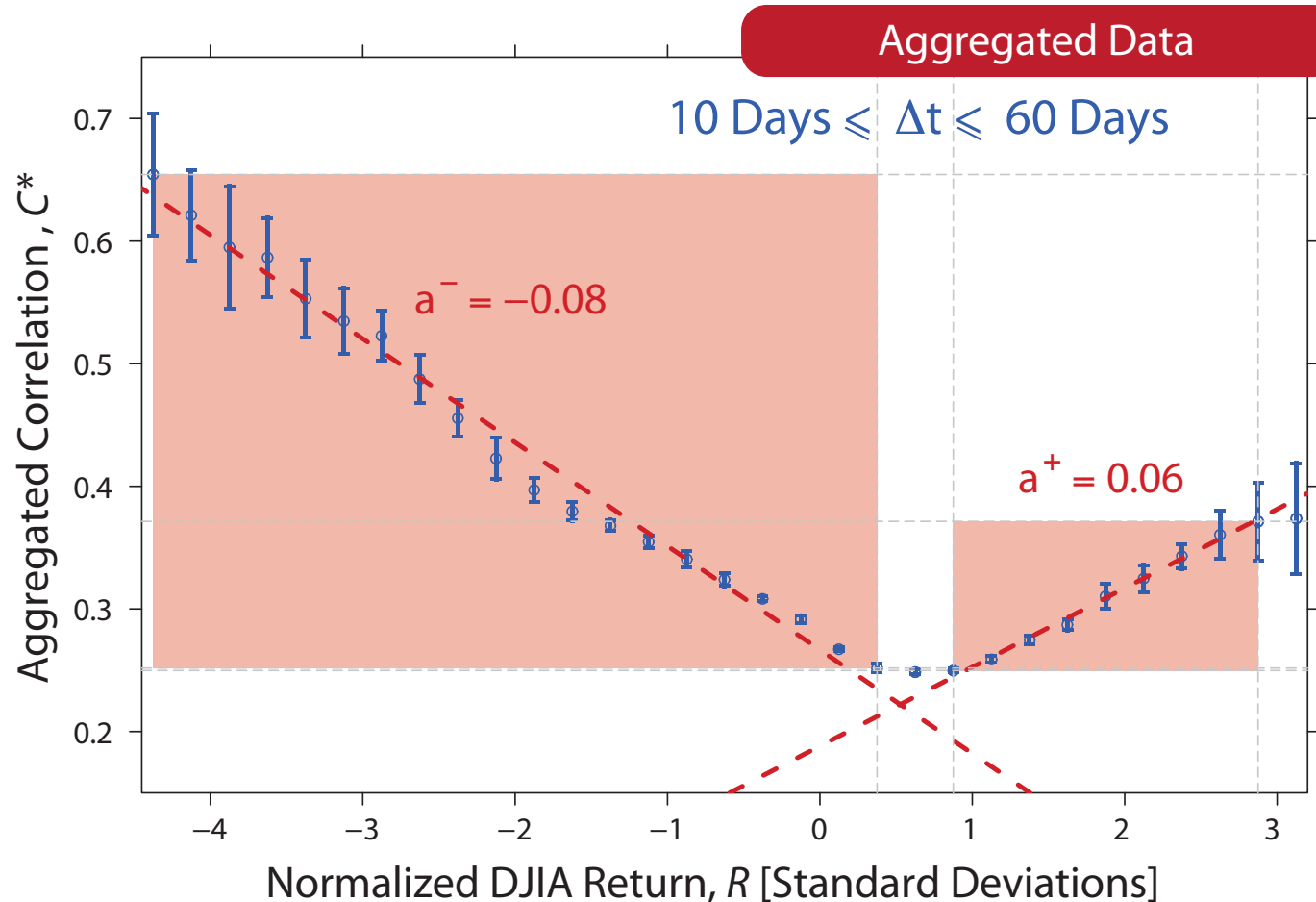
Preis, Kenett, Stanley, Helbing & Ben Jacob, *Scientific Reports* 2, 752 (2012)
<http://www.nature.com/srep/2012/121018/srep00752/full/srep00752.html>

Quantifying Stock Market Correlations



Preis, Kenett, Stanley, Helbing & Ben Jacob, *Scientific Reports* 2, 752 (2012)
<http://www.nature.com/srep/2012/121018/srep00752/full/srep00752.html>

Quantifying Stock Market Correlations



Preis, Kenett, Stanley, Helbing & Ben Jacob, *Scientific Reports* **2**, 752 (2012)
<http://www.nature.com/srep/2012/121018/srep00752/full/srep00752.html>

Behaviour Captured by *Google*



Trends

Web Search Interest: warwick. Worldwide, 2004 - present.



Explore trends

Hot searches

Search terms

Warwick

+ Add term

Other comparisons

Limit to

Web Search

Worldwide

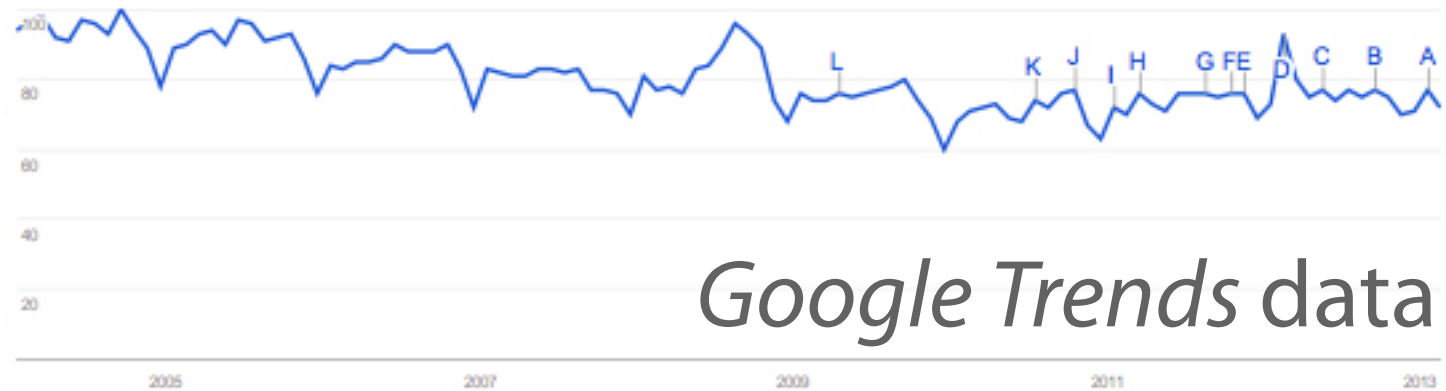
2004 - present

All Categories

Interest over time

The number 100 represents the peak search volume

☒ News headlines ☐ Forecast



Google Trends data

Embed

Regional interest



0 100

Region | City

Related terms

Top

Rising

warwick university	100	
warwick ri	90	
the warwick	65	
warwick castle	55	
warwick hotel	50	
warwick school	45	
dionne warwick	40	
university of warwick	35	
warwick avenue	30	
warwick ny	25	

Can *Google* Predict Flu Spreading?

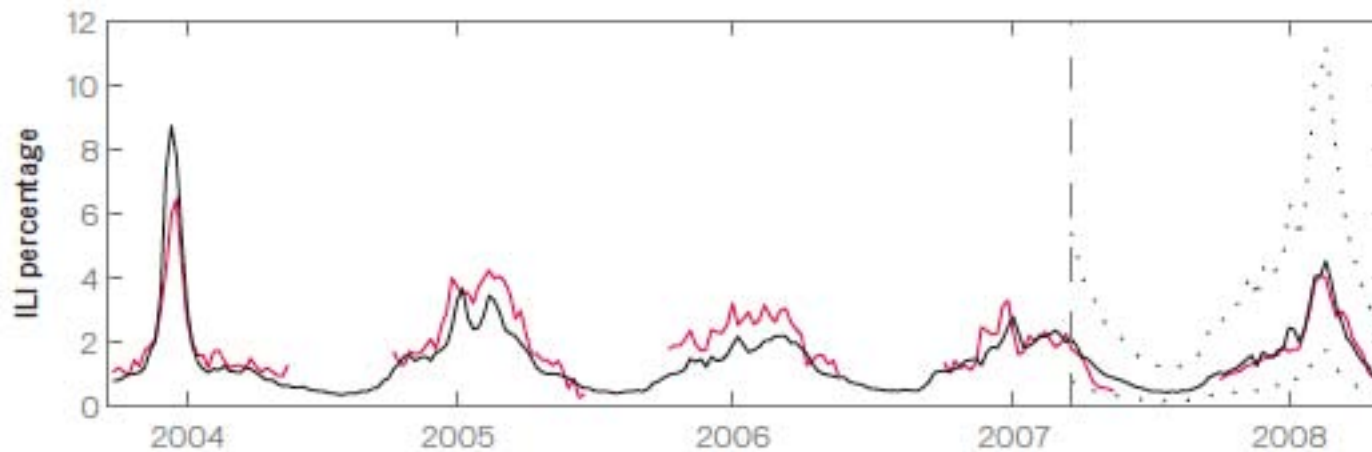


Figure 2: A comparison of model estimates for the Mid-Atlantic Region (black) against CDC-reported ILI percentages (red), including points over which the model was fit and validated. A correlation of 0.85 was obtained over 128 points from this region to which the model was fit, while a correlation of 0.96 was obtained over 42 validation points. 95% prediction intervals are indicated.

Ginsberg, Mohebbi, Patel, Brammer, Smolinski & Brilliant, *Nature* **368**, 1012 (2009)
<http://dx.doi.org/10.1038/nature07634>

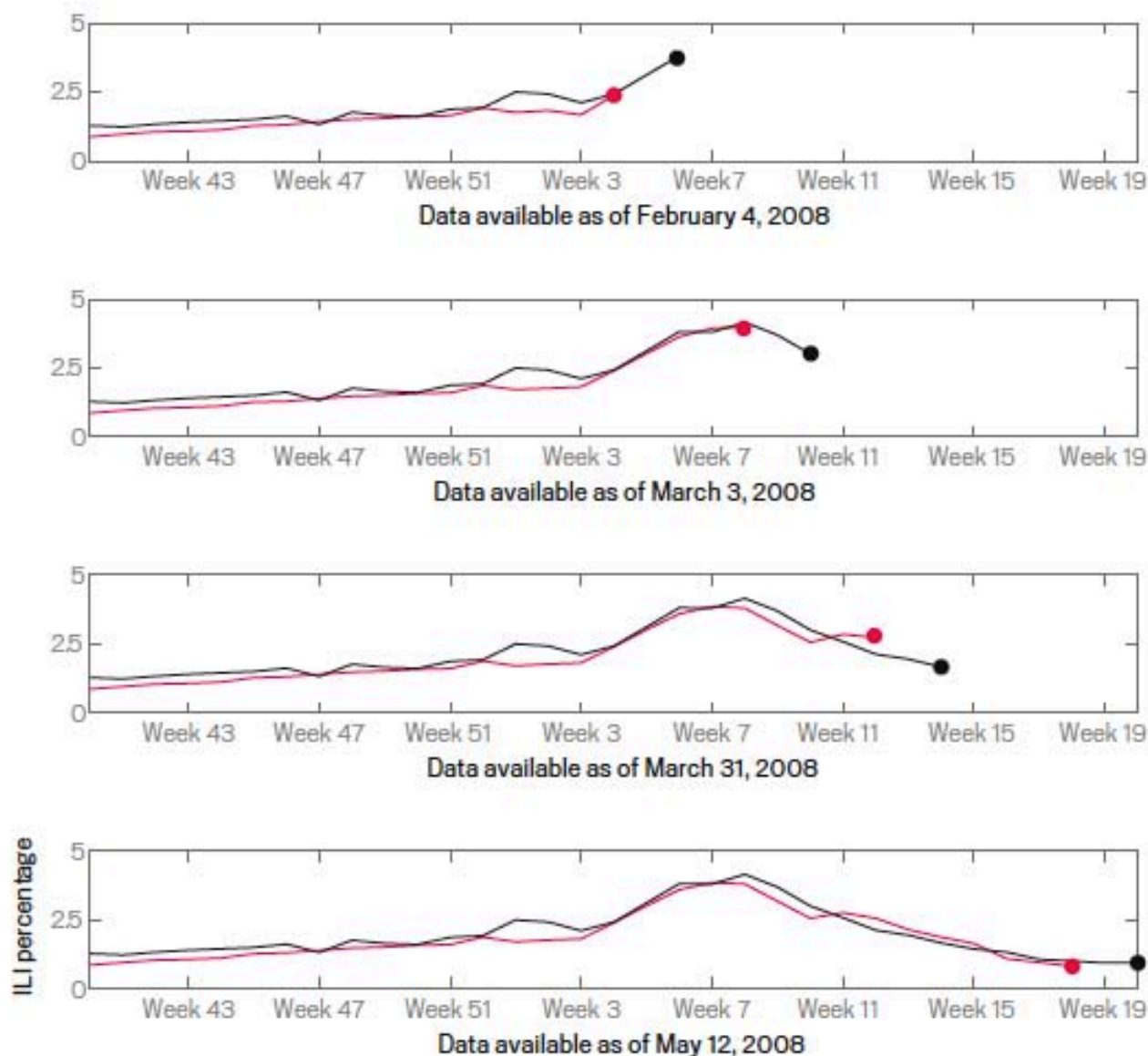


Figure 3: ILI percentages estimated by our model (black) and provided by CDC (red) in the Mid-Atlantic region, showing data available at four points in the 2007-2008 influenza season. During week 5, we detected a sharply increasing ILI percentage in the Mid-Atlantic region; similarly, on March 3, our model indicated that the peak ILI percentage had been reached during week 8, with sharp declines in weeks 9 and 10. Both results were later confirmed by CDC ILI data.

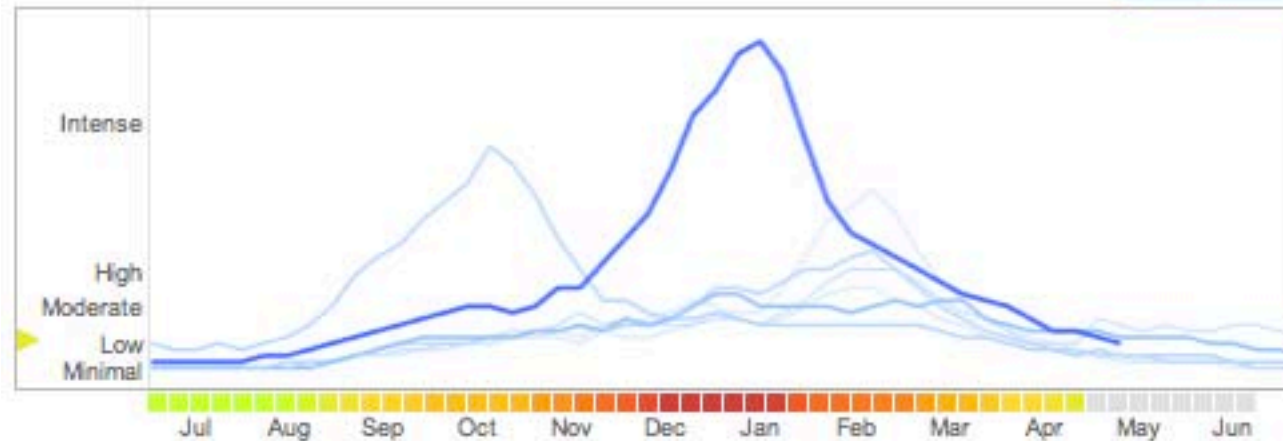
Ginsberg, Mohebbi, Patel, Brammer, Smolinski & Brilliant, *Nature* **368**, 1012 (2009)
<http://dx.doi.org/10.1038/nature07634>

Explore flu trends - United States

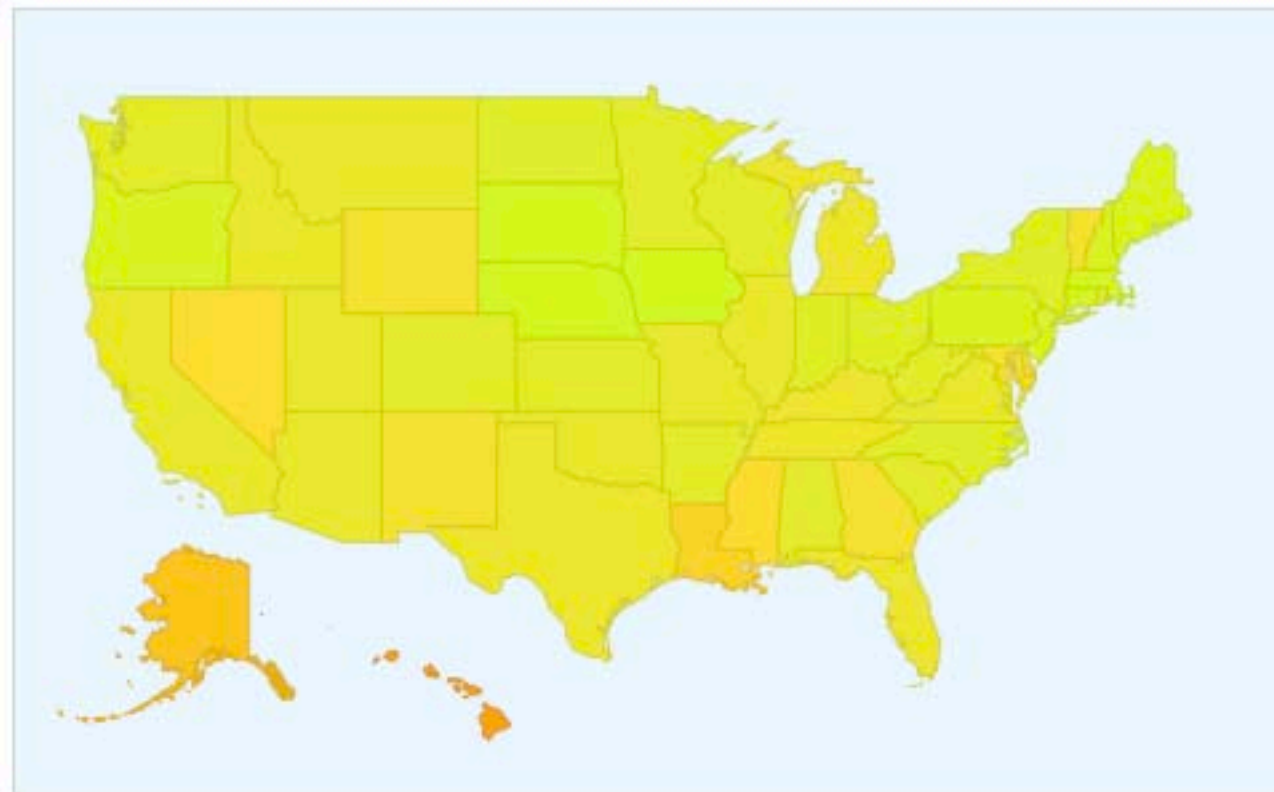
We've found that certain search terms are good indicators of flu activity. Google Flu Trends uses aggregated Google search data to estimate flu activity. [Learn more »](#)

National

● 2012-2013 ● Past years ▼



States | [Cities](#) (Experimental)



Google Flu Trends

<http://www.google.org/flutrends>

Ginsberg, Mohebbi, Patel,
Brammer, Smolinski &
Brilliant, *Nature* **368**, 1012
(2009)

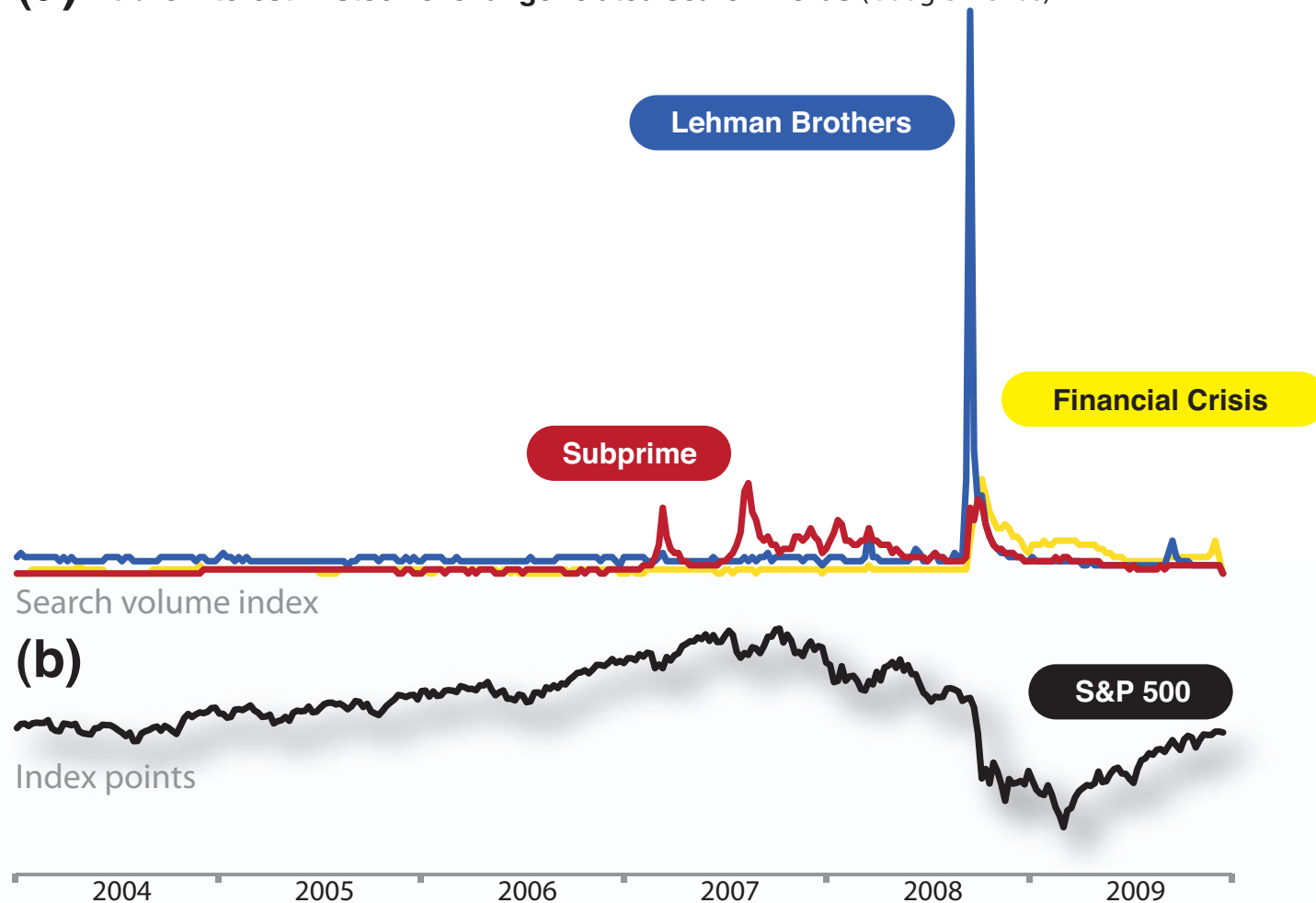
[http://dx.doi.org/10.1038/
nature07634](http://dx.doi.org/10.1038/nature07634)

What *Google* Knows About Markets



Google Search Query Data

(a) Public interest in stock exchange related search words (Google Trends)



Preis, Reith & Stanley, *Phil. Trans. R. Soc. A* 368, 5707 (2010)

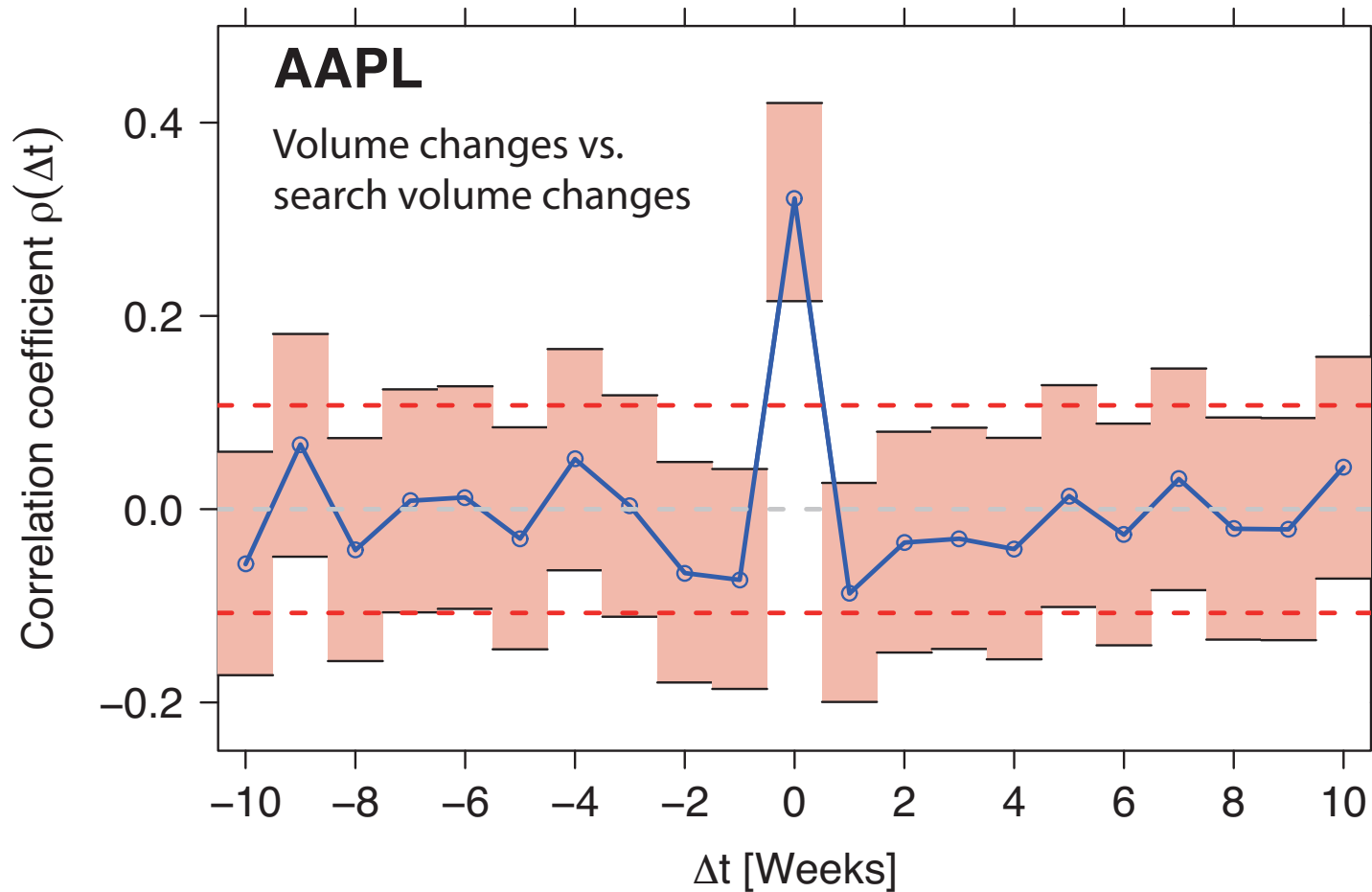
Covered by

TIME

CNN

Science

Google Search Query Data



Preis, Reith & Stanley, *Phil. Trans. R. Soc. A* **368**, 5707 (2010)

Covered by

TIME

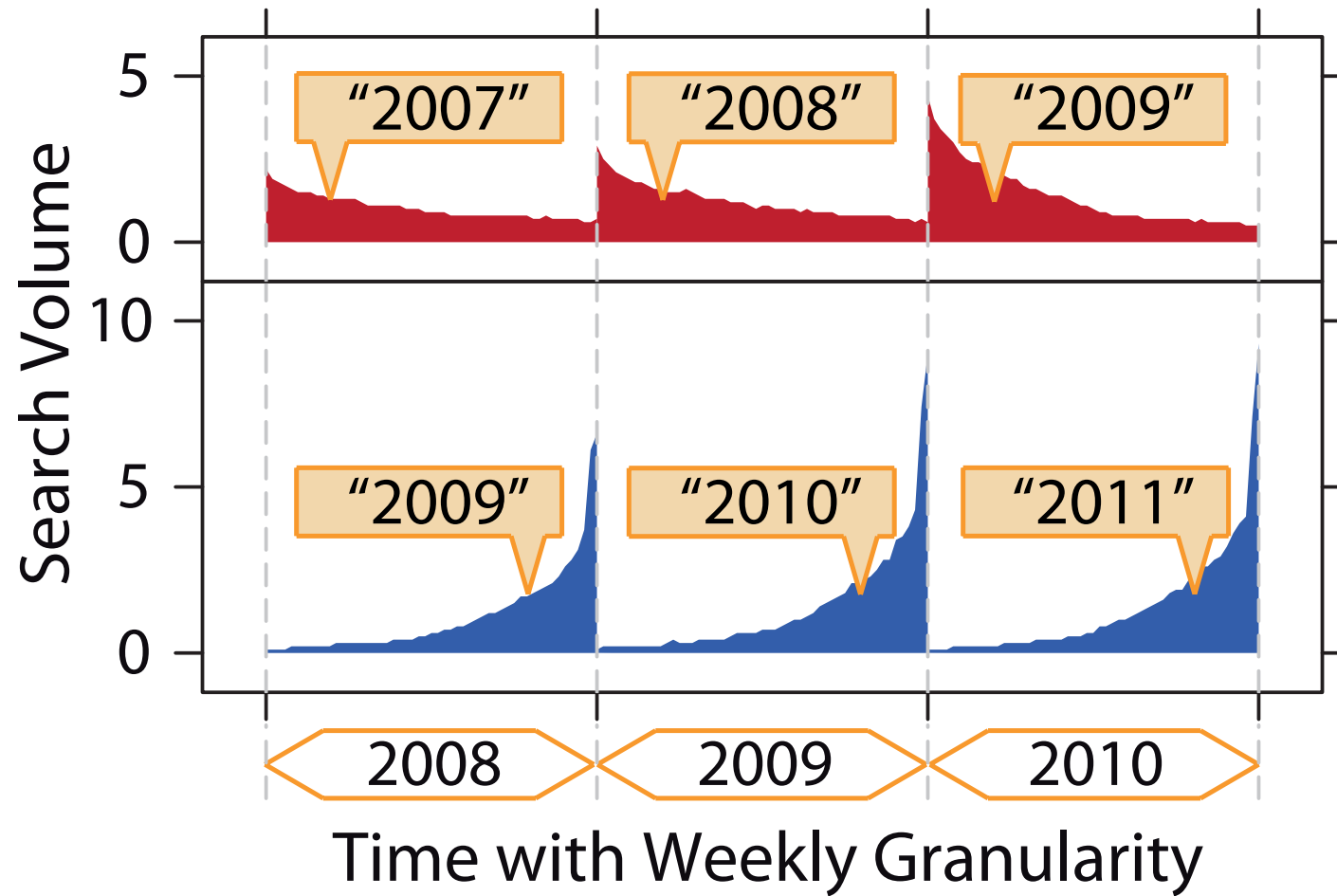
CNN

Science

What *Google* Knows About Economies



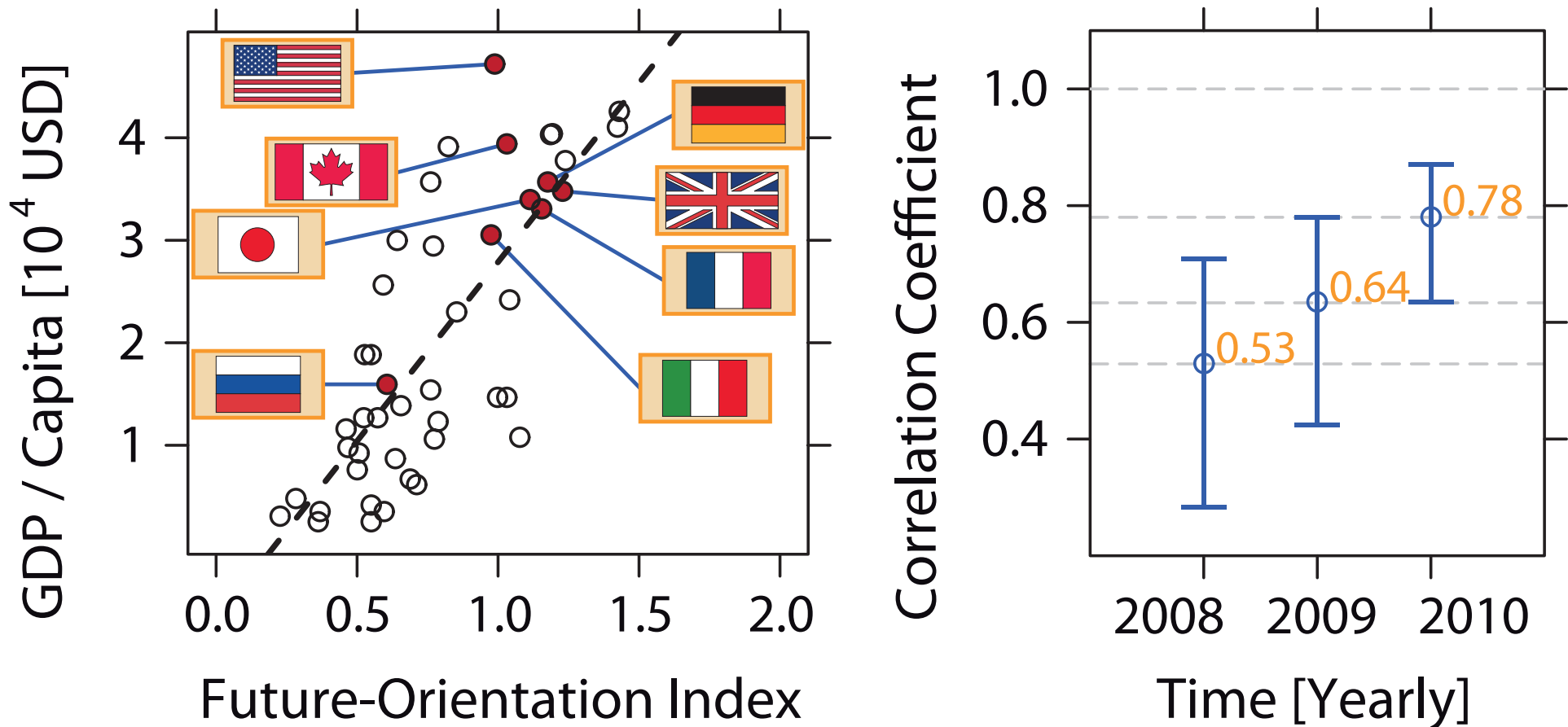
Quantifying the Advantage of Looking Forward



Preis, Moat, Stanley & Bishop, *Scientific Reports* 2, 350 (2012)

www.nature.com/srep/2012/120405/srep00350/full/srep00350.html

Quantifying the Advantage of Looking Forward



Preis, Moat, Stanley & Bishop, *Scientific Reports* 2, 350 (2012)

www.nature.com/srep/2012/120405/srep00350/full/srep00350.html

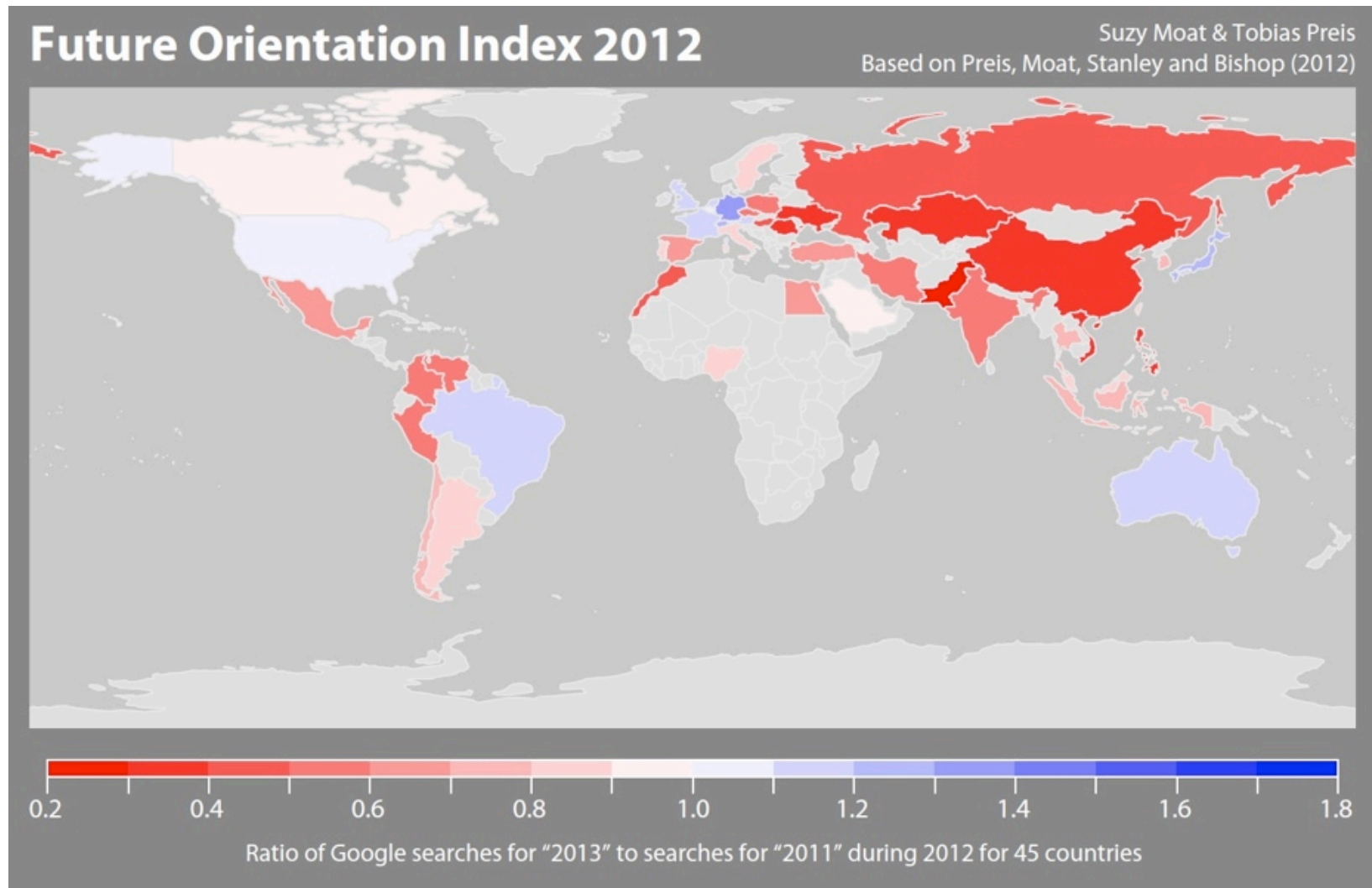
Covered by

NewScientist

THE WALL STREET JOURNAL



Quantifying the Advantage of Looking Forward



Covered by

MailOnline

LA STAMPA.it

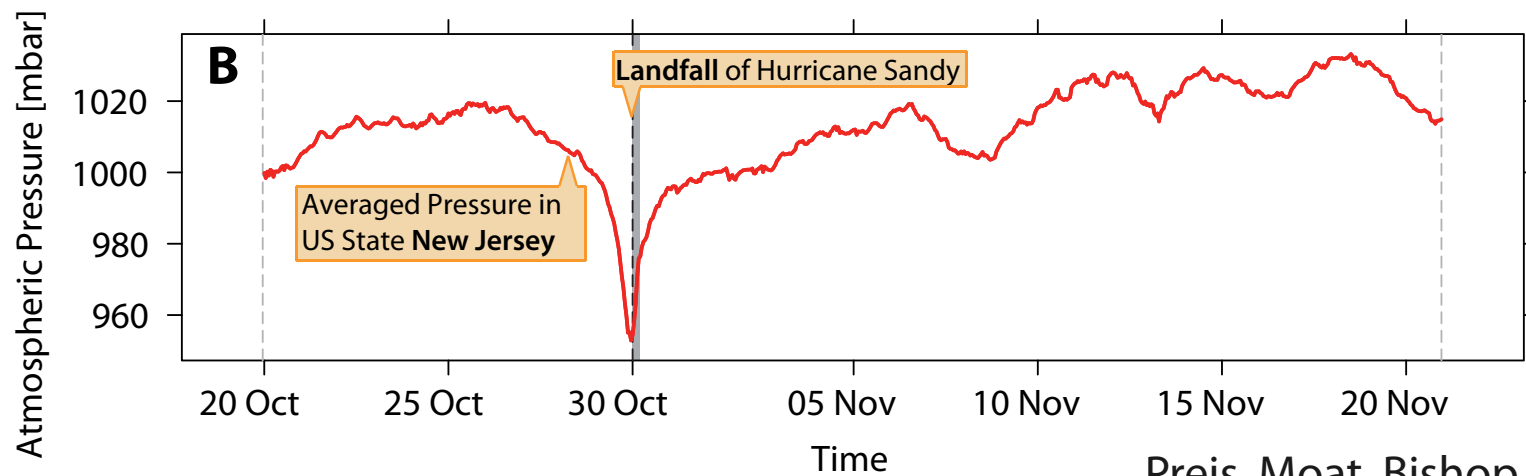
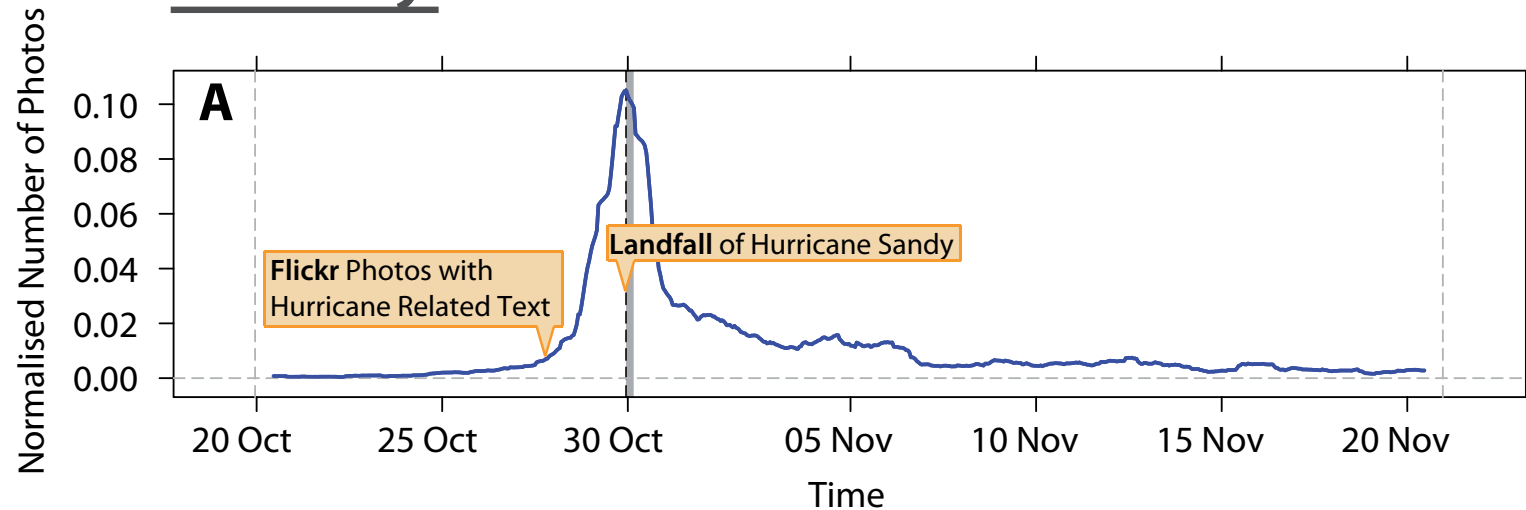
**Bloomberg
Businessweek**

Flickr World Map 2012



31.6 Million Photographs

Hourly Flickr data and *Hurricane Sandy*



Preis, Moat, Bishop, Treleven & Stanley
Under Review

Summary

1. Large technological systems have become a central part of our society.
2. These systems form a world wide sensor network, measuring human behaviour on a massive scale.
3. We can use this data to:
 - measure what is happening in our world right now