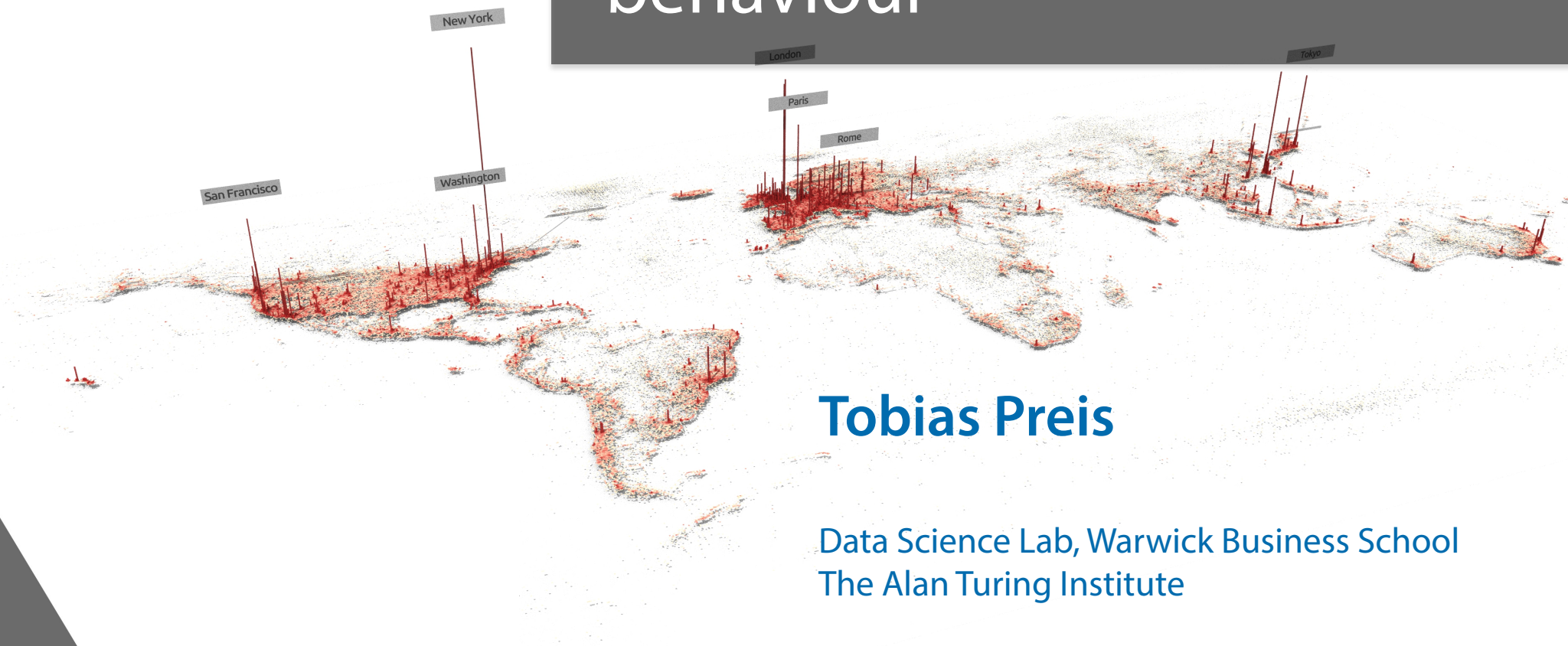


Using Google, Wikipedia, Flickr and Twitter to analyse and predict human behaviour

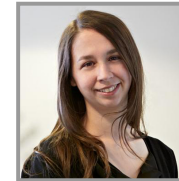


Tobias Preis

Data Science Lab, Warwick Business School
The Alan Turing Institute

Tobias.Preis@wbs.ac.uk
<http://www.tobiaspreis.com>

Data Science Lab



Key Objectives

1

Reducing cost and delay of existing measurements of behaviour

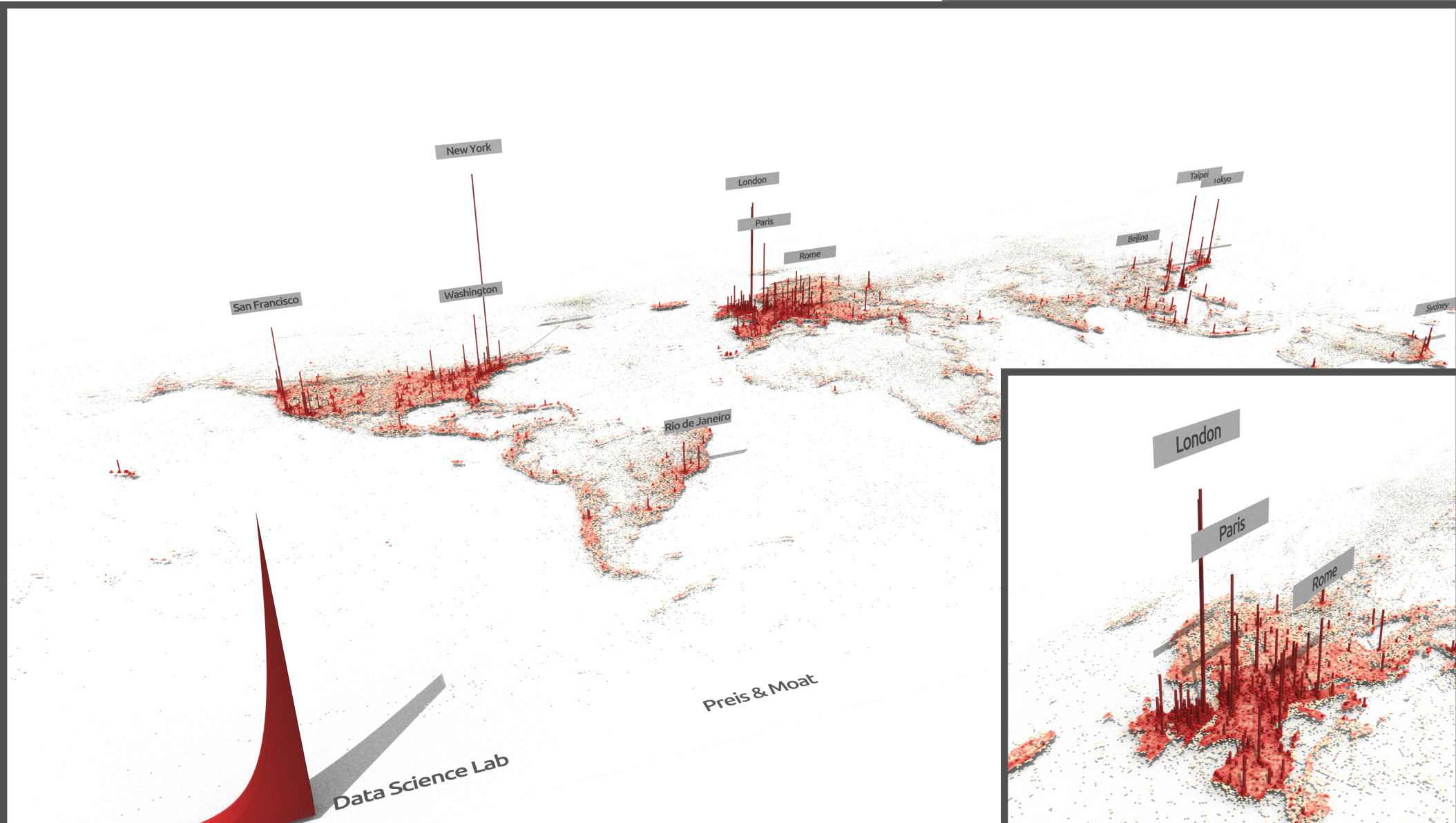
2

Measuring new quantities of behaviour

3

Predicting future behaviour

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



<http://www.bbc.co.uk/iplayer/episode/b06g7rs8/click-03102015>

A close-up, shallow depth-of-field photograph of a MacBook Pro. The screen at the top shows the Google logo in its multi-colored font. Below the screen, the 'MacBook Pro' text is visible on the dark bezel. The keyboard is in the foreground, with several keys in focus. The overall lighting is soft and warm.

Google

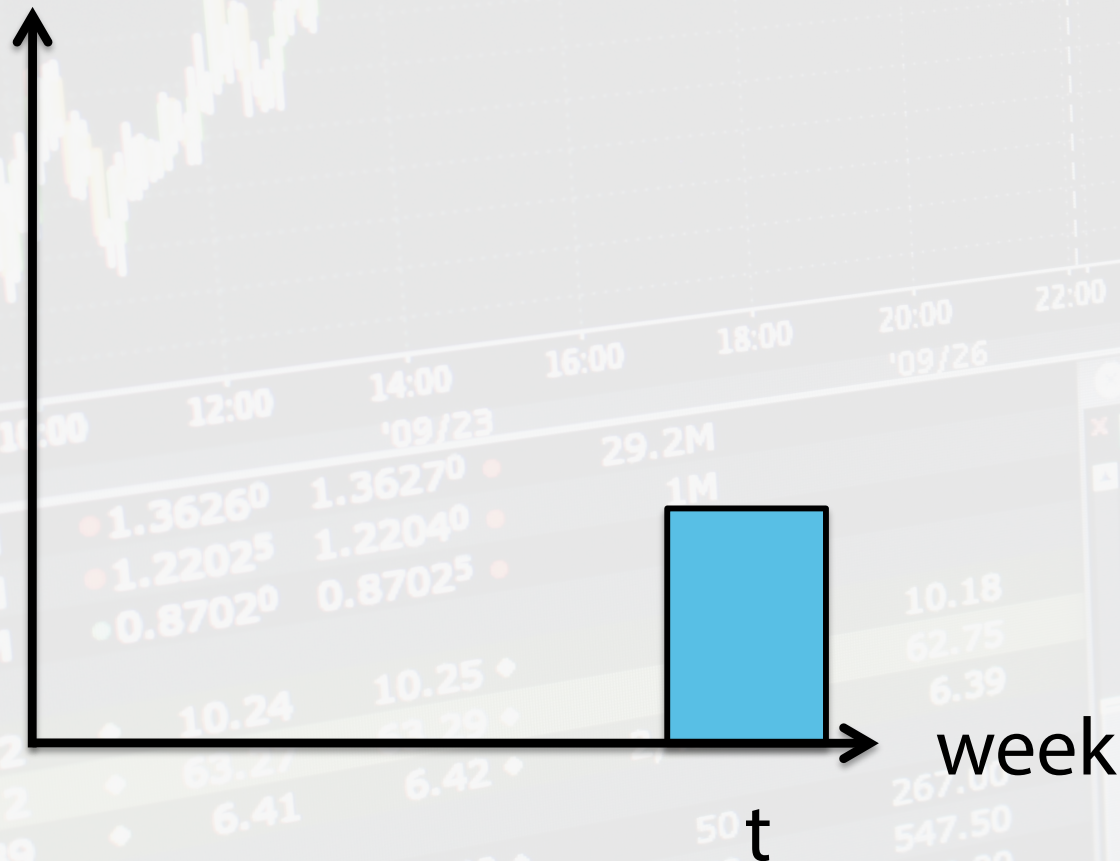
1

Google and the stock market

Hypothetical strategy

Moat et al. (2013);
Preis et al. (2013)

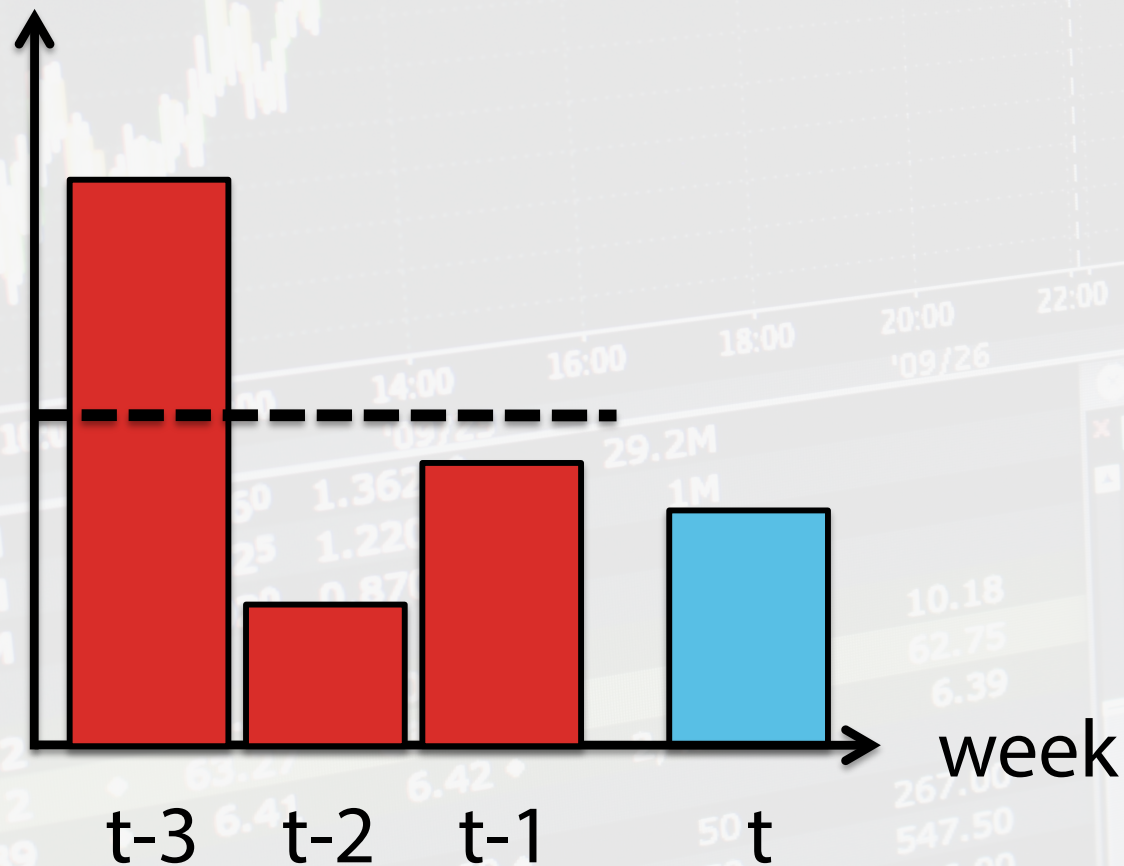
volume of *Google* searches
for keyword



Hypothetical strategy

Moat et al. (2013);
Preis et al. (2013)

volume of *Google* searches
for keyword

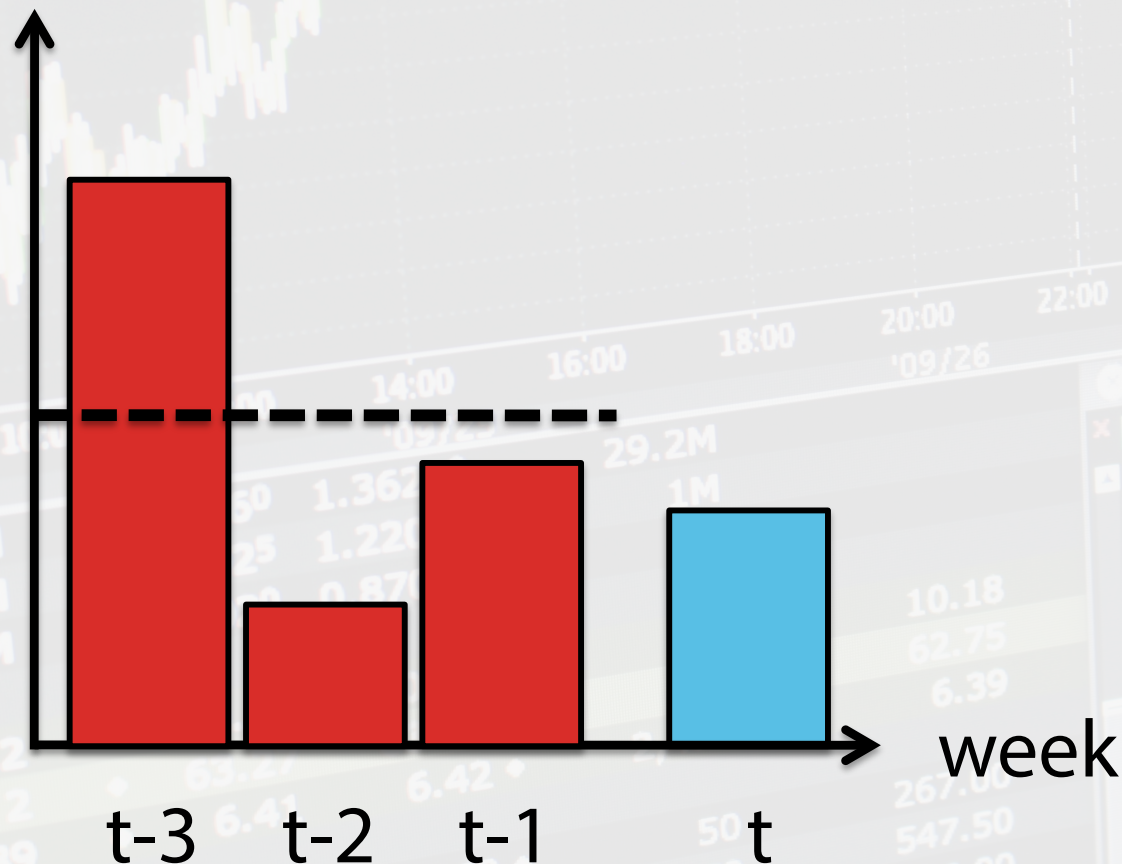


Hypothetical strategy

Moat et al. (2013);
Preis et al. (2013)

volume of *Google* searches
for keyword

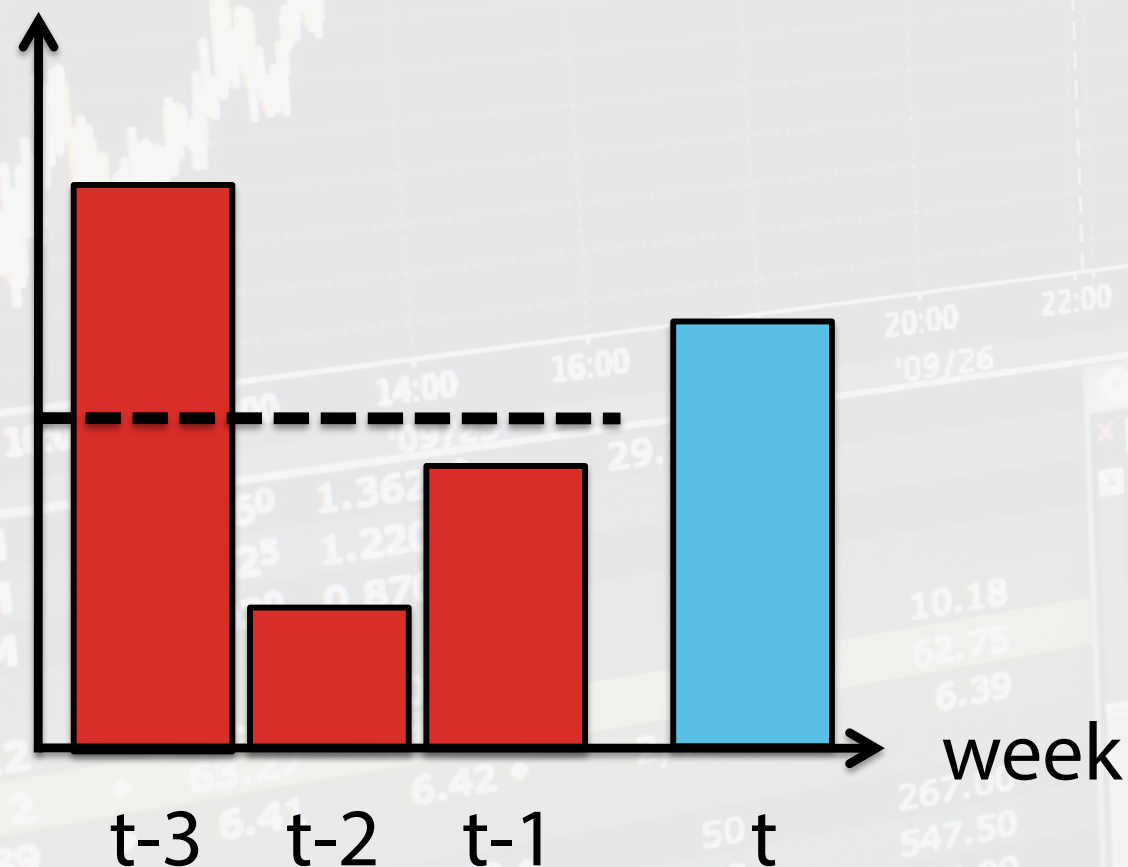
Search volume
decreased:
BUY stock
in week $t+1$



Hypothetical strategy

Moat et al. (2013);
Preis et al. (2013)

volume of *Google* searches
for keyword

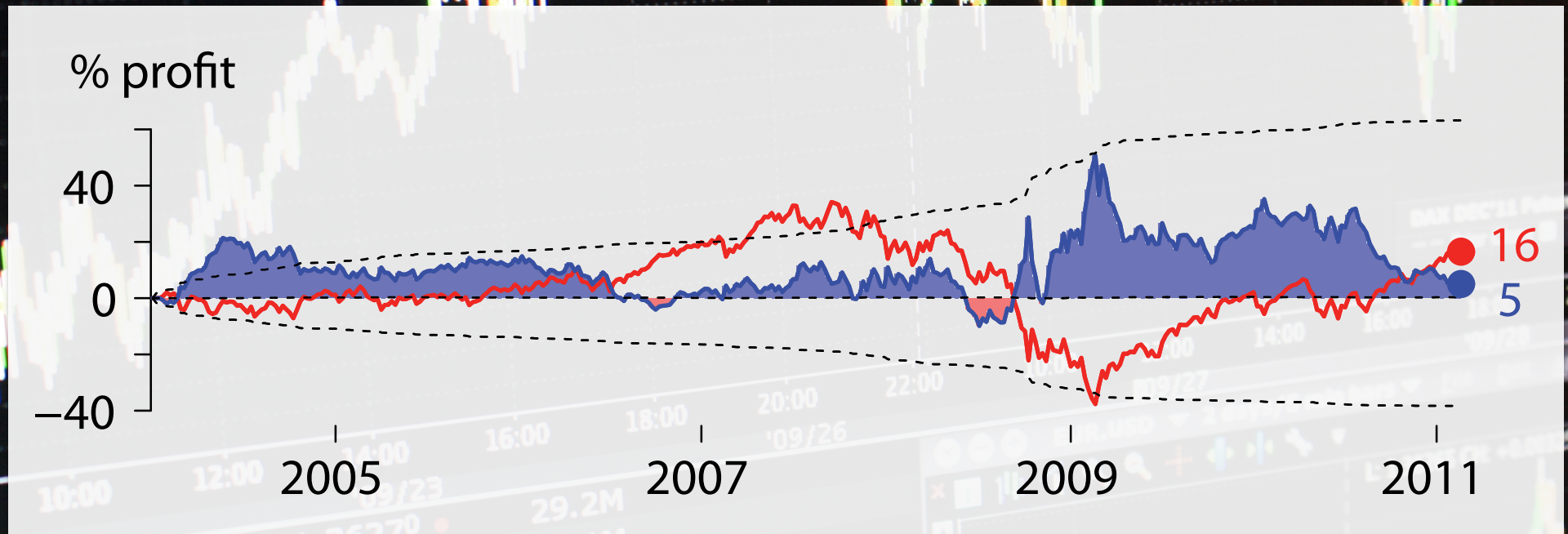


Search volume
decreased:
BUY stock
in week $t+1$

Search volume
increased:
SELL stock
in week $t+1$

Example: "culture"

Featured by:

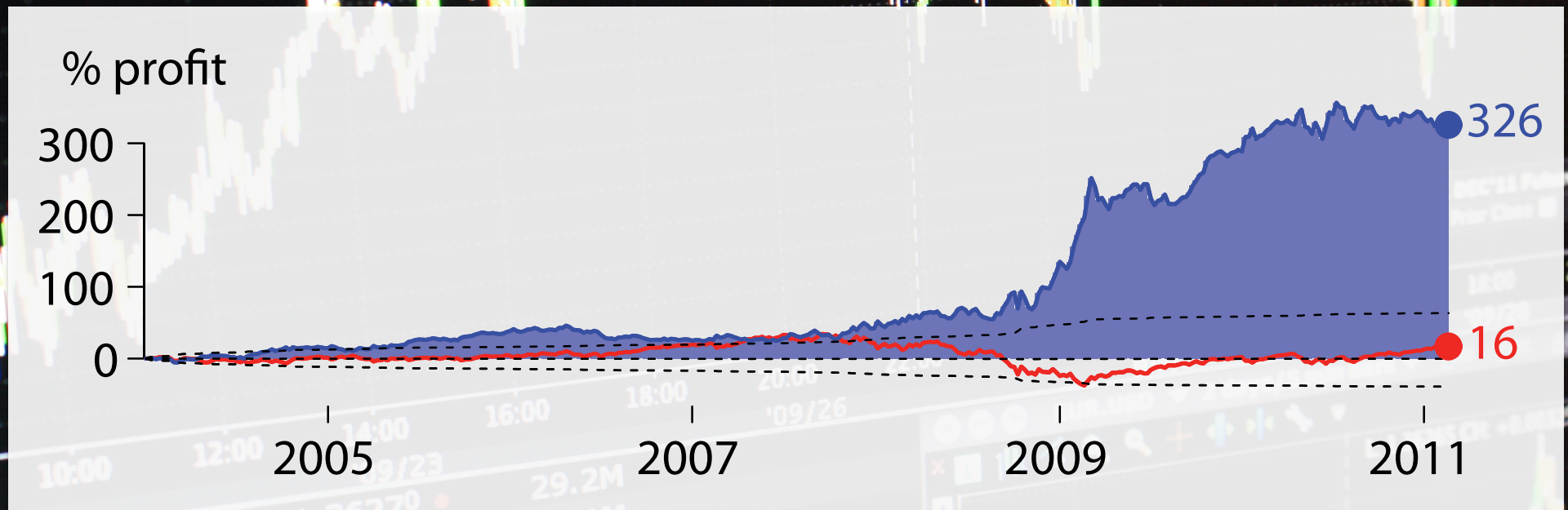


Preis, Moat
& Stanley
(2013)

"culture" trading strategy
buy and hold strategy
mean \pm 1 sd of random strategies

Example: "debt"

Featured by:

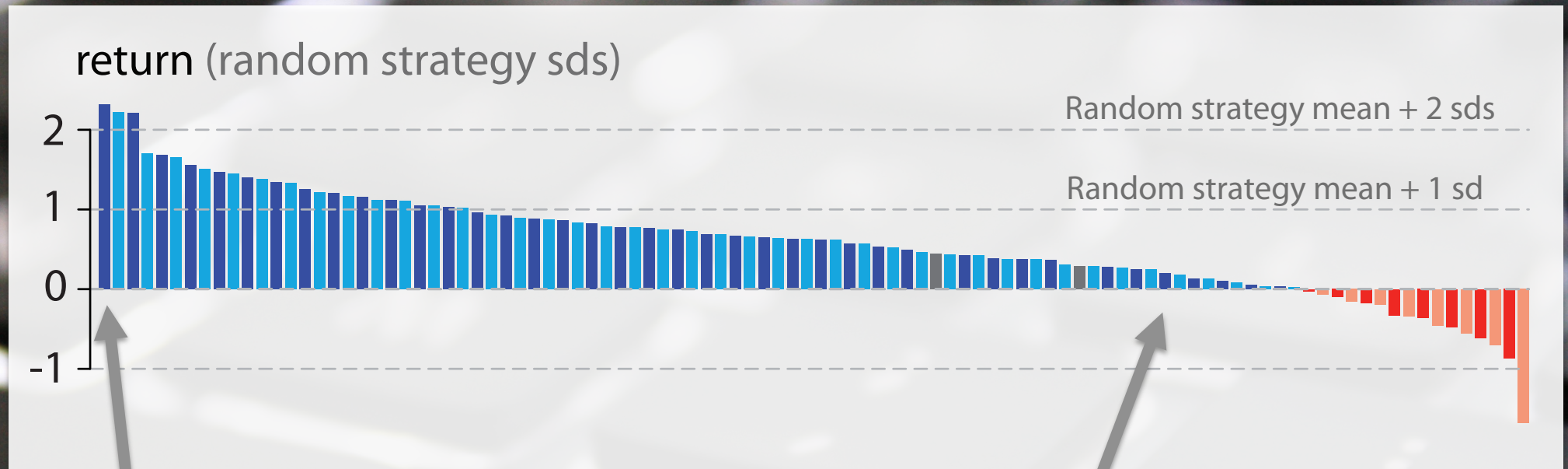


Preis, Moat
& Stanley
(2013)

— "debt" trading strategy
— buy and hold strategy
- - - mean \pm 1 sd of random strategies

<https://www.nature.com/articles/srep01684>

How different keywords perform

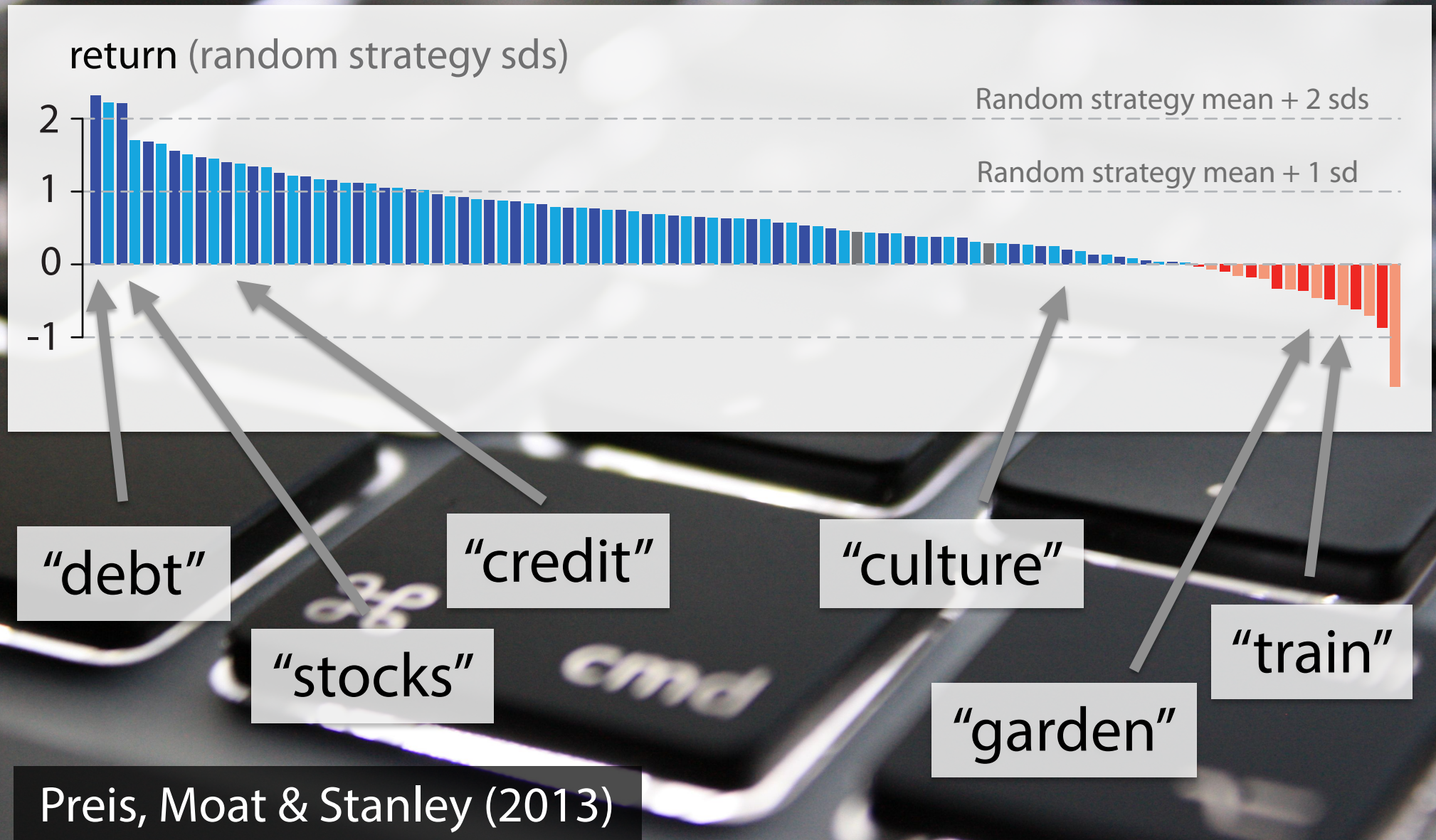


"debt"

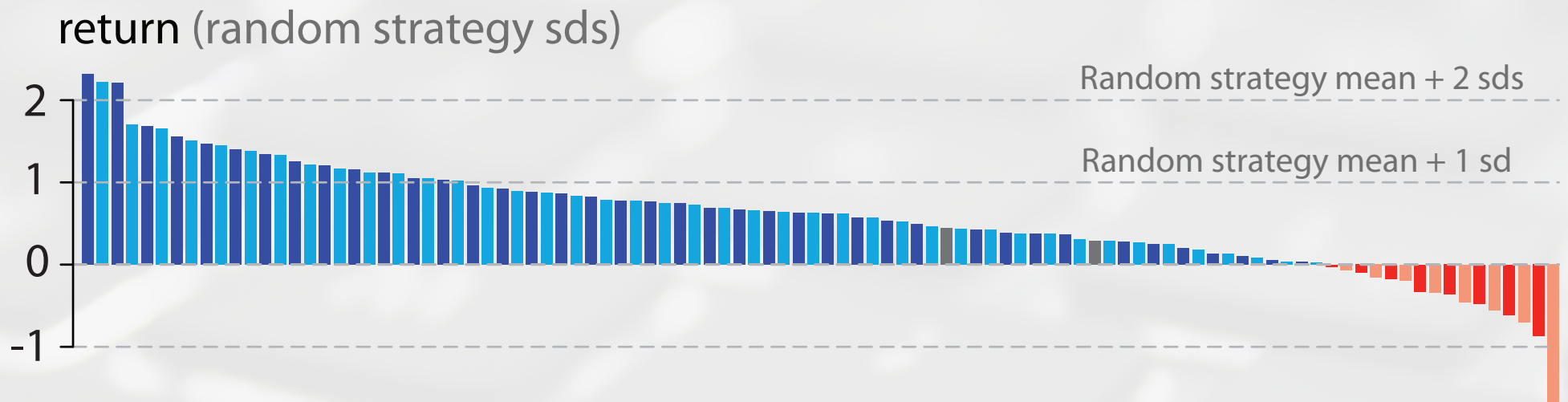
"culture"

Preis, Moat & Stanley (2013)

How different keywords perform



How different keywords perform



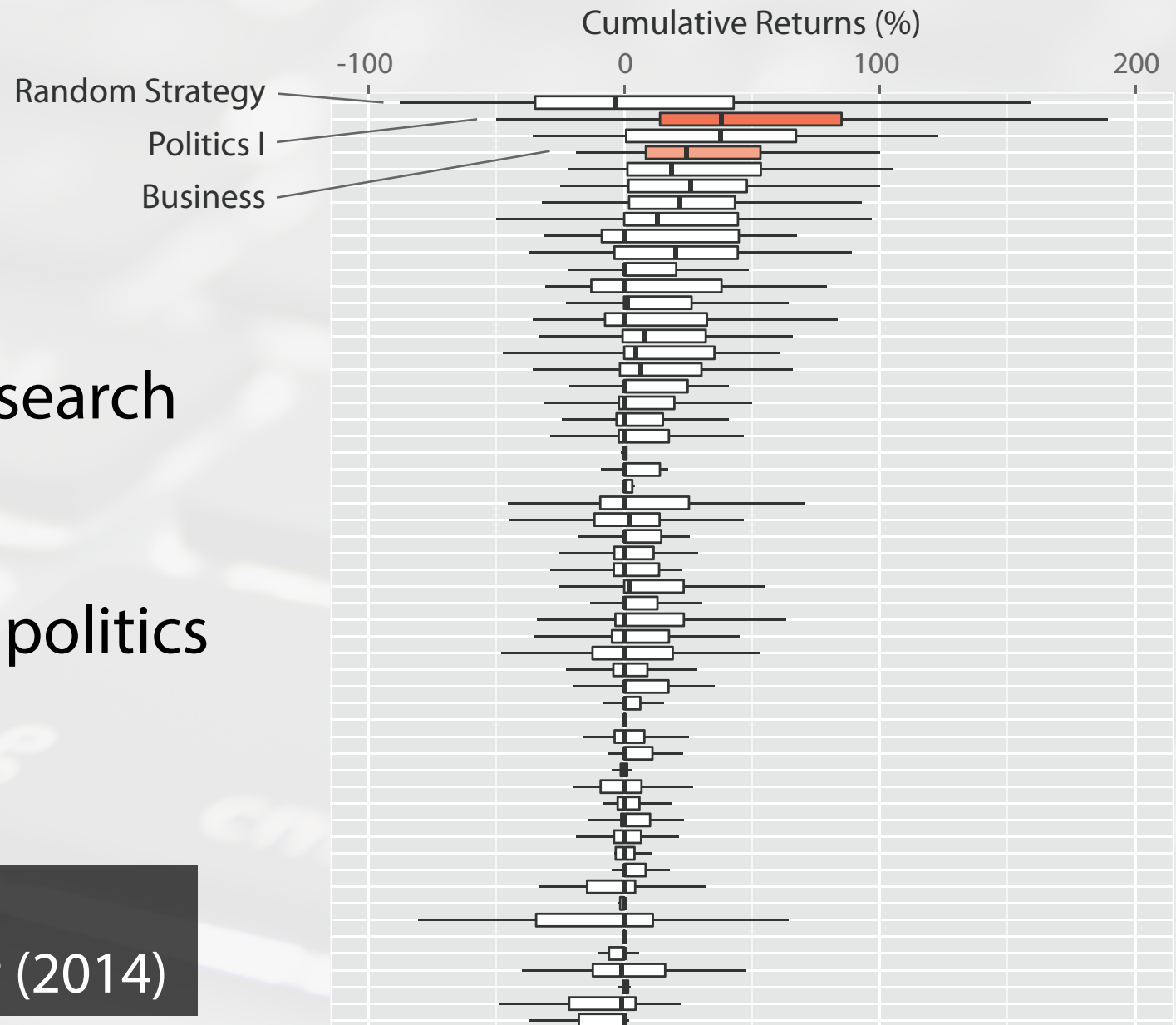
Financial relevance

$\frac{\# \text{ occurrences in FT}}{\# \text{ hits on Google}}$

Returns significantly correlated with indicator of financial relevance

Preis, Moat & Stanley (2013)

What is searched for before falls?



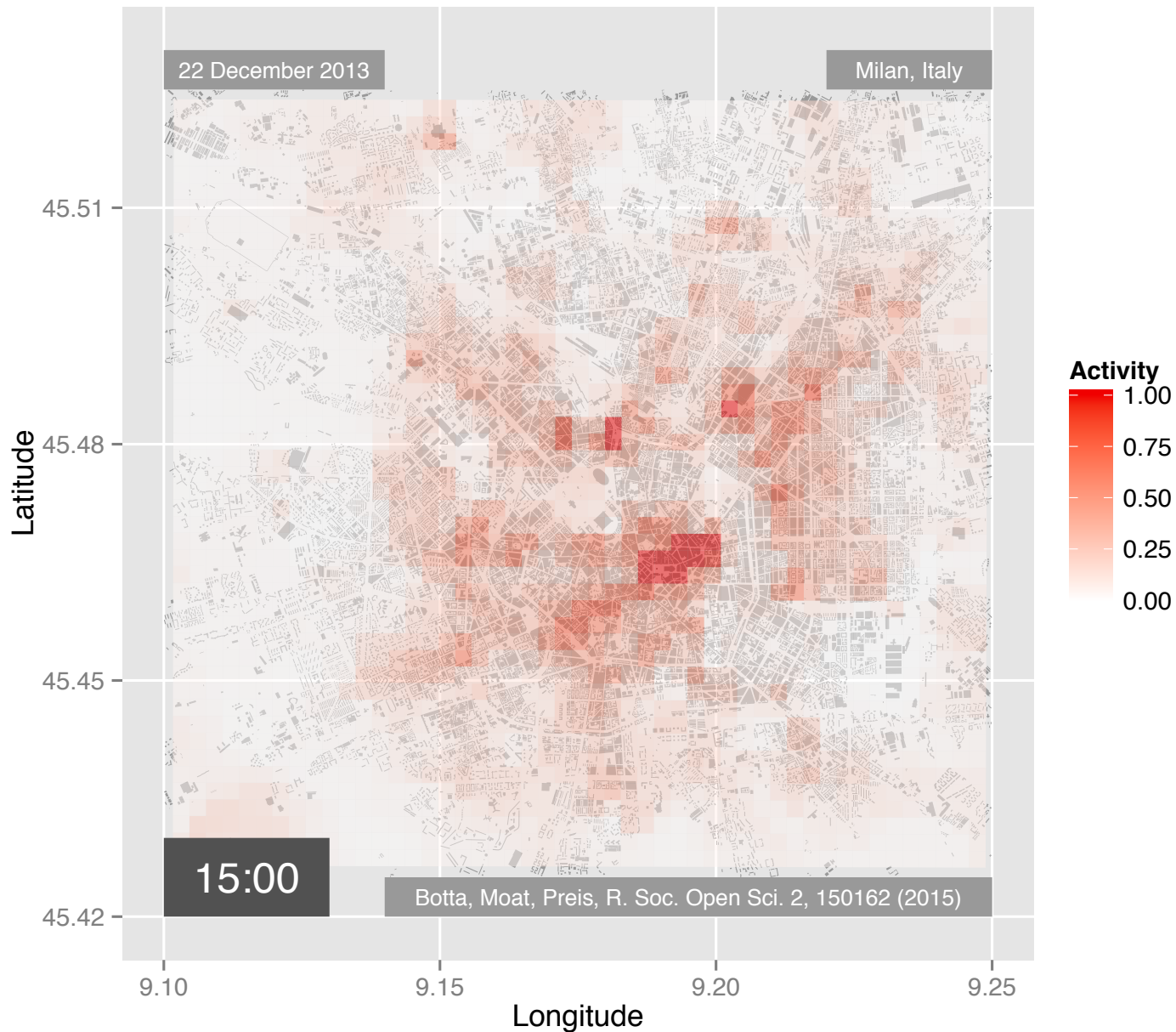
A large crowd of people is seated in a stadium, filling the bleachers. Many individuals are holding up their smartphones, capturing photos or videos of an event. The scene is brightly lit, and the atmosphere appears to be one of a significant gathering or concert.

Photo: Kris Krüg

2

Counting crowds

Mobile phone activity in Milan

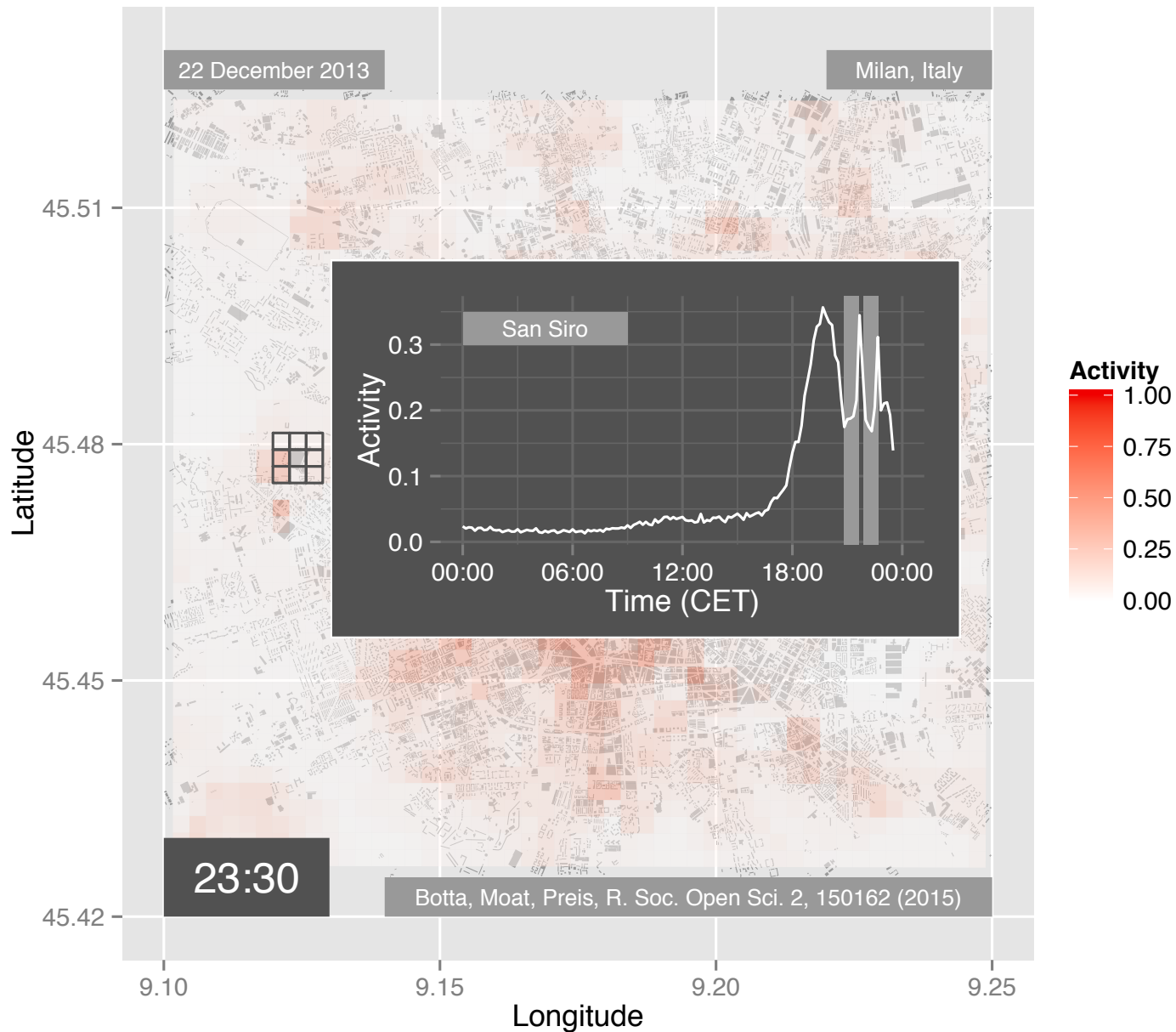


Featured by:

BBC

Science

Mobile phone activity in Milan

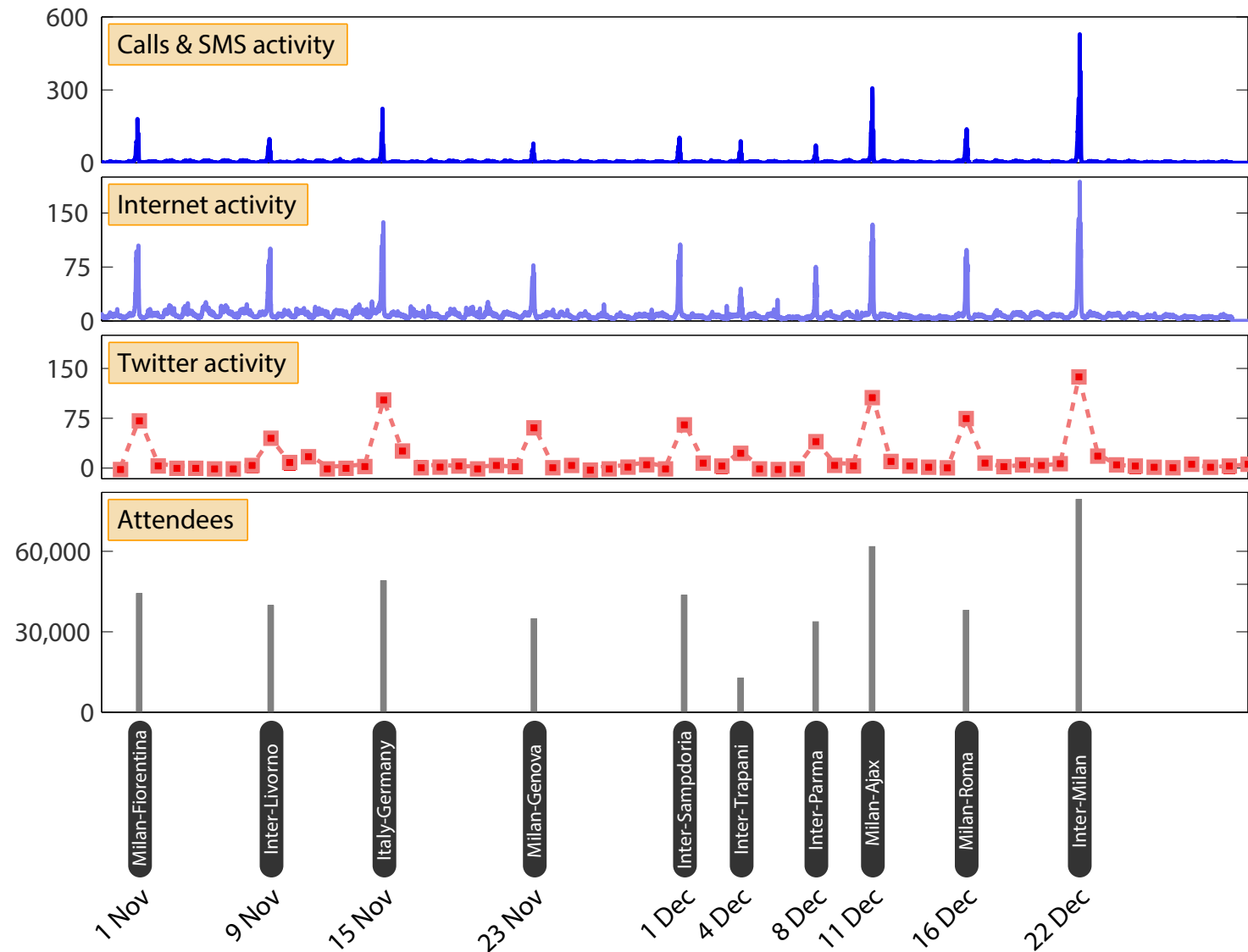


Featured by:

BBC

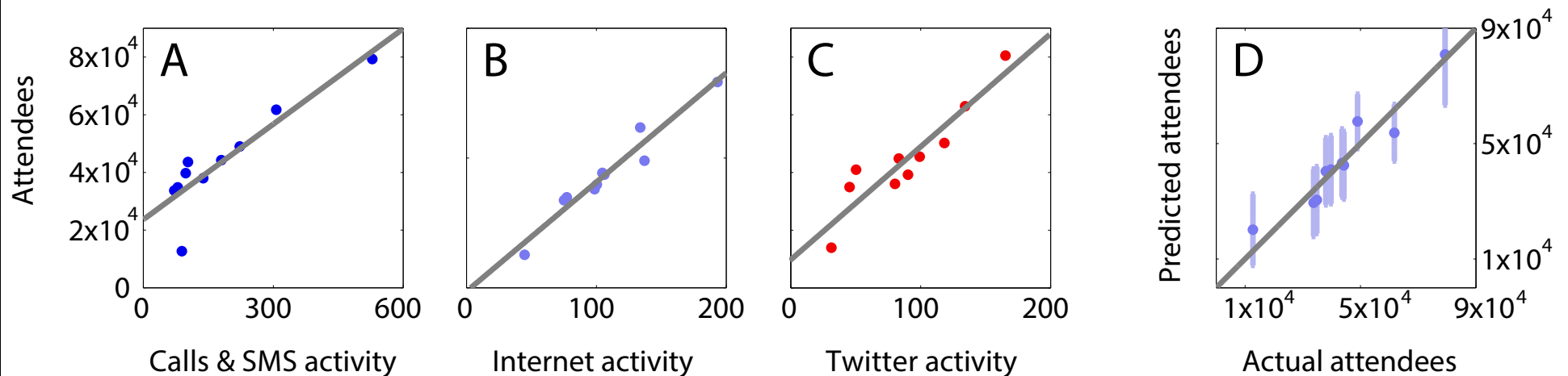
Science

Mobile phone and *Twitter* data



Mobile phone and *Twitter* data

San Siro stadium



Featured by:



Match attendance estimated
from phone activity

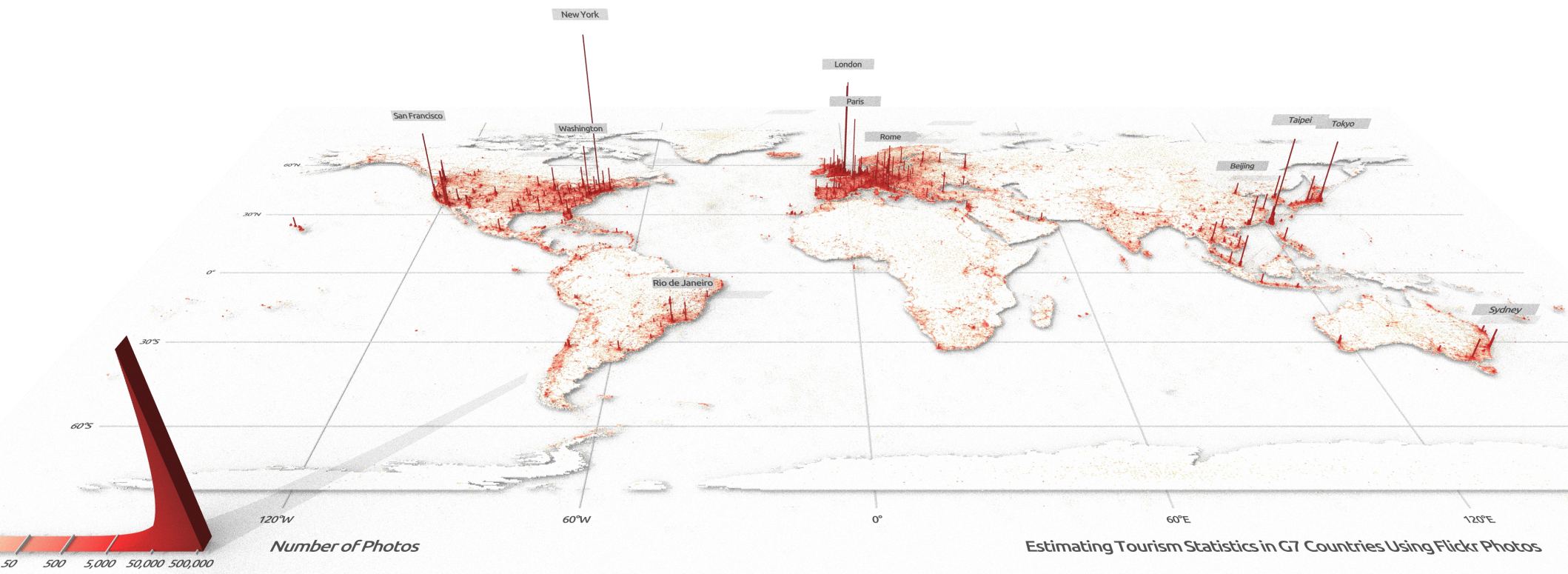
<http://rsos.royalsocietypublishing.org/content/royopensci/2/5/150162.full.pdf>

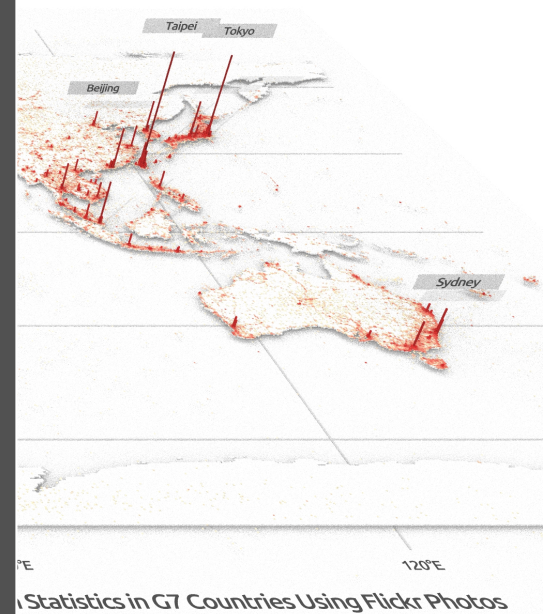
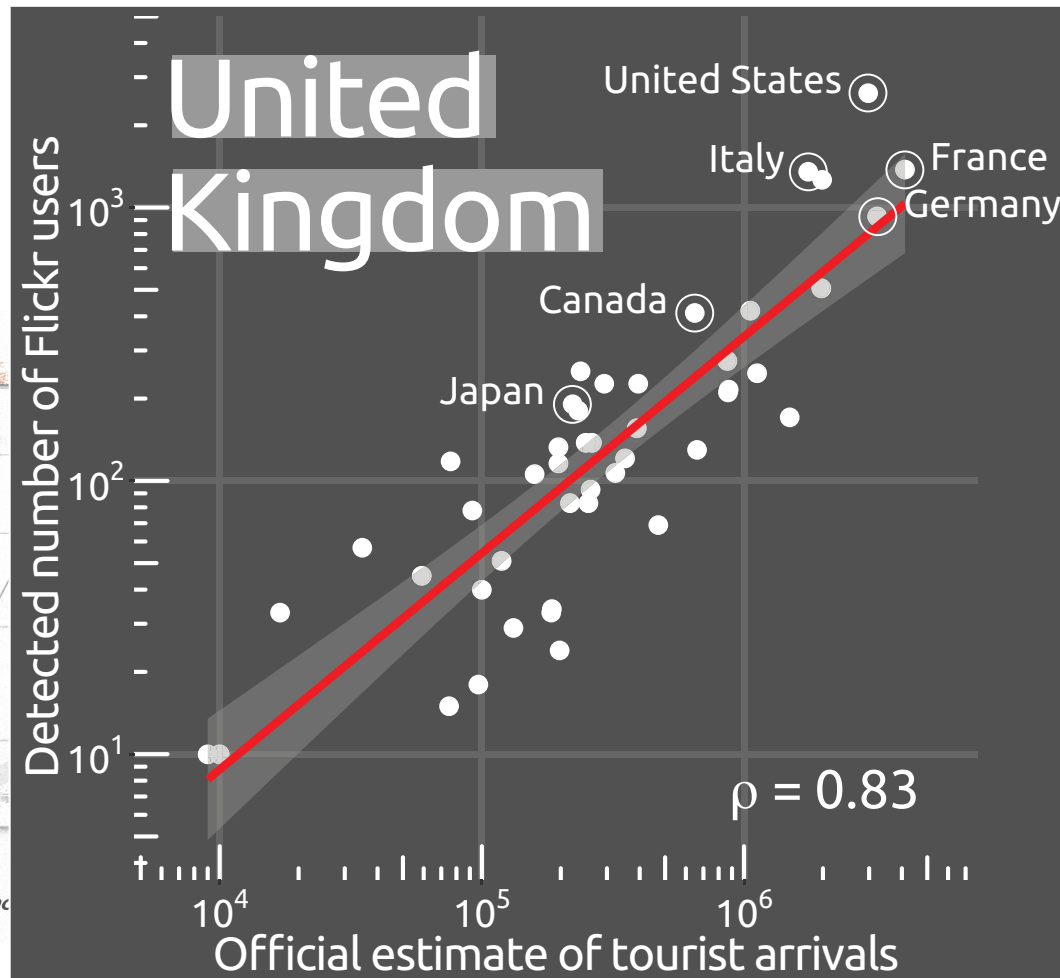
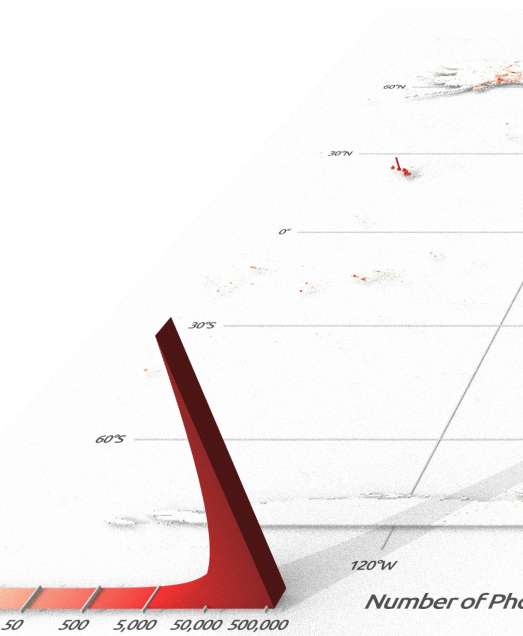
Botta, Moat and Preis (2015)



3

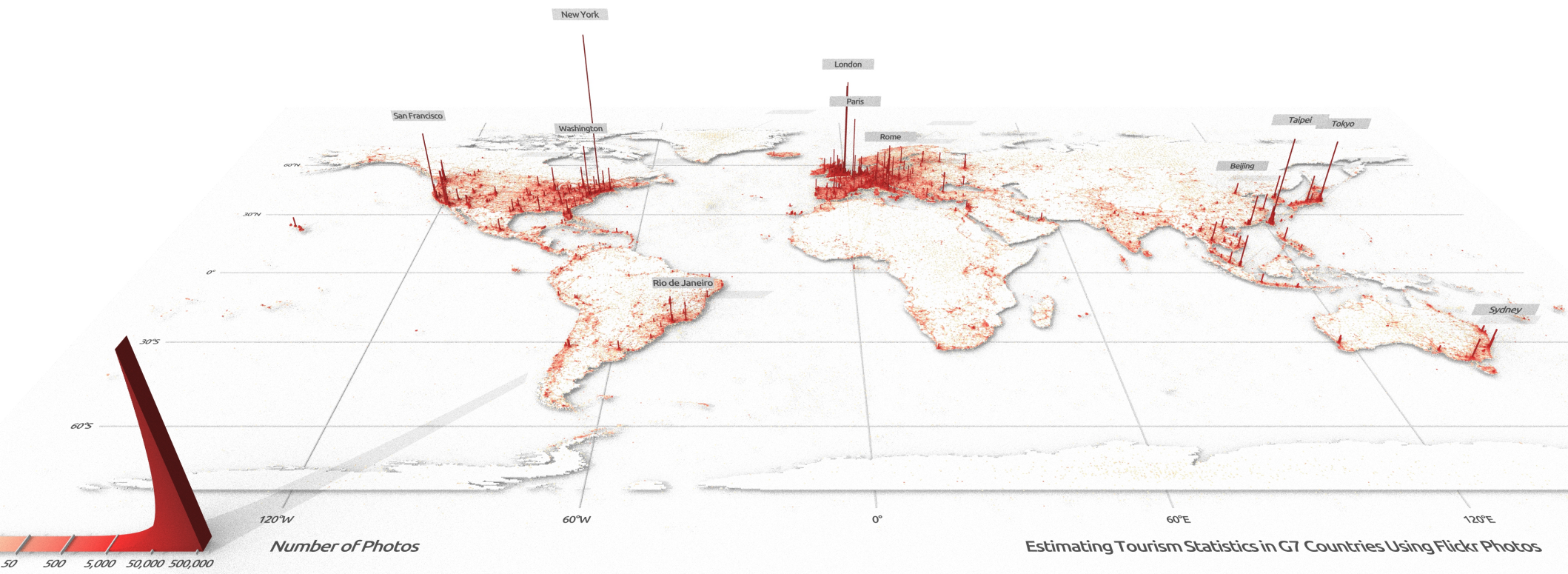
Flickr traces of travel flows

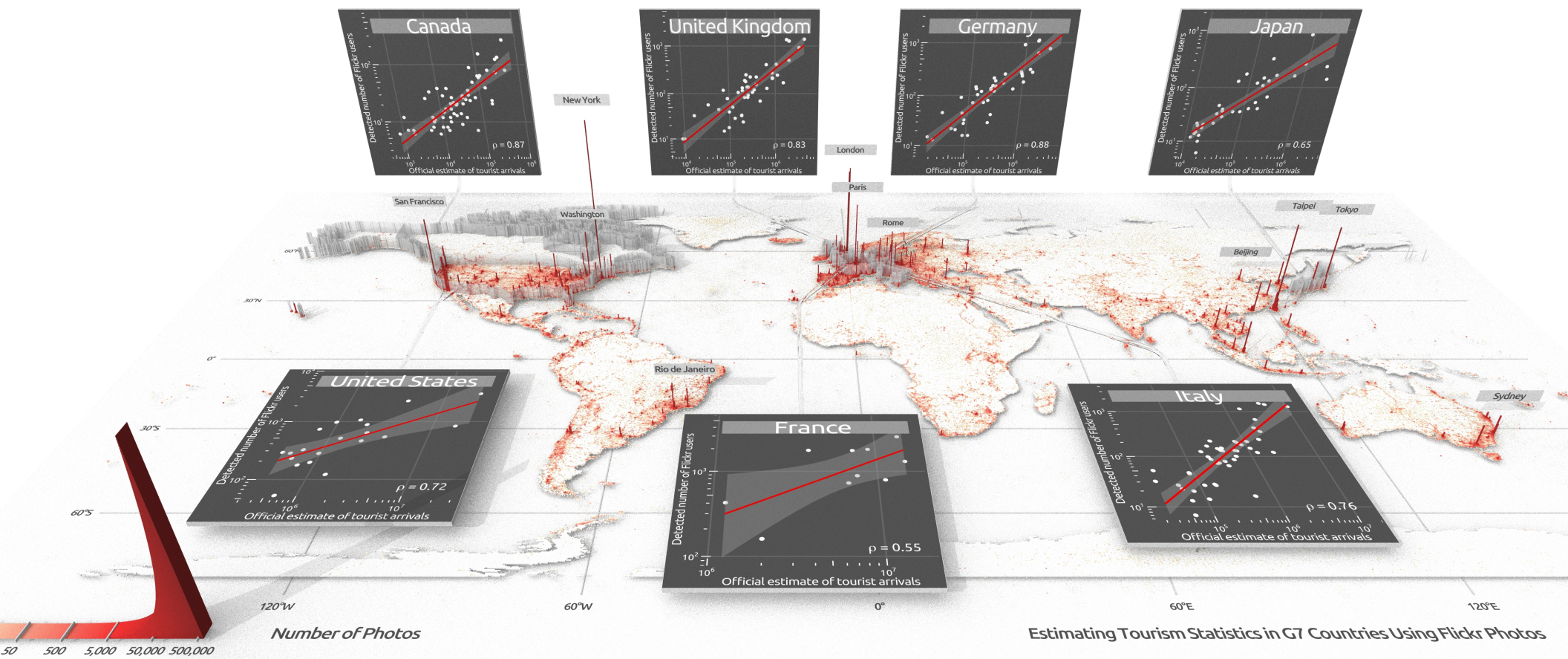




Statistics in G7 Countries Using Flickr Photos

Barchiesi, Moat, Alis, Bishop & Preis (2015)







Data from the Internet may help us
measure and even predict
human behaviour



Engineering and Physical Sciences
Research Council

