Platform effects in social media networks

Momin M. Malik[1] & Jürgen Pfeffer[2]

[1]Institute for Software Research, Carnegie Mellon University
[2]Bavarian School of Public Policy, Technical University of Munich

XXXVI Sunbelt Conference of the International Network for Social Network Analysis
Newport Beach, California
April 10, 2016, 11am
Social Media Networks: Challenges and Solutions, Sunday AM2

These slides available at:
http://mominmalik.com/sunbelt2016.pdf

Full paper forthcoming in *ICWSM* in May, pre-print at:
http://mominmalik.com/icwsm2016pre.pdf

Outline

1. Motivation

2. Theory

3. Method and Data

4. Findings

5. Conclusion

Ongoing theorizing

Photo credit: Chris Zimmerman, 2015, https://twitter.com/socialbeit/status/672106094882369536
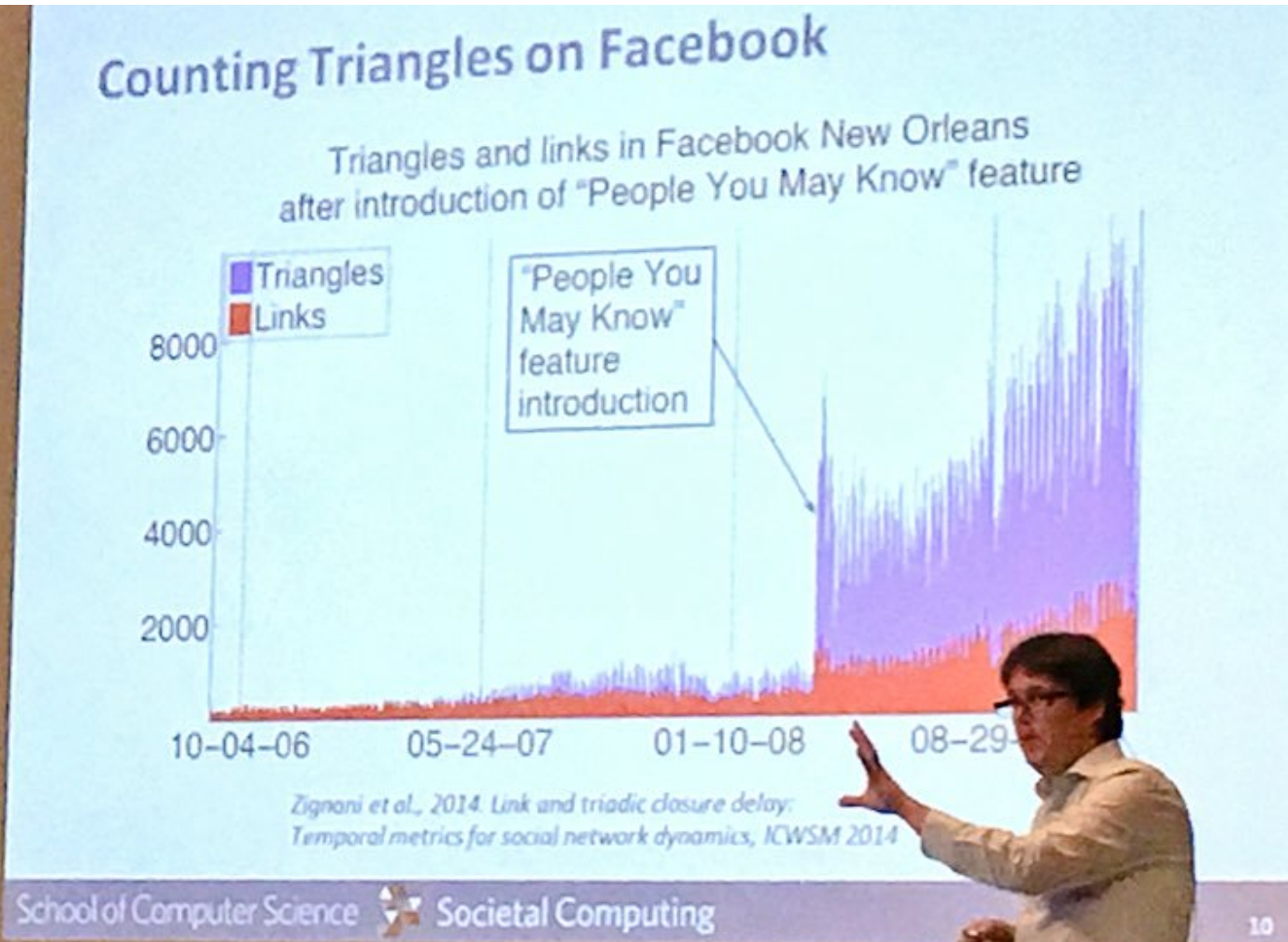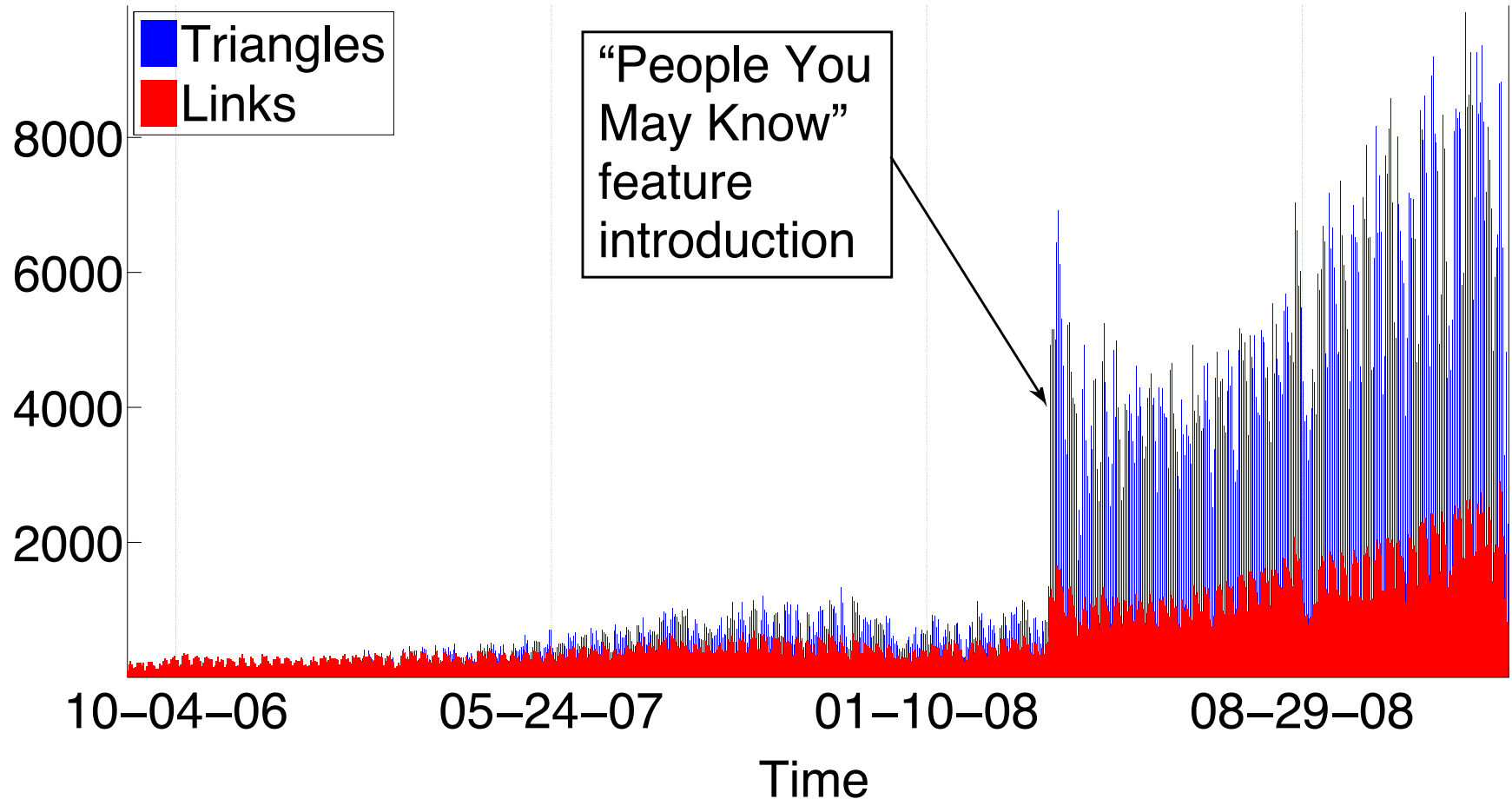
Adapted from fig. 2b of Zignani et al., 2014

Triangles and links in Facebook New Orleans

Hopes for new ways of measurement

"Disciplines are revolutionized by the development of novel tools: the telescope for astronomers, **the microscope for biologists**, the particle accelerator for physicists, and brain imaging for cognitive psychologists. **Social media provide a high-powered lens into the details of human behavior and social interaction** that may prove to be equally transformative."

Golder, S., & Macy, M. (2012). Social science with social media. *ASA footnotes*, *40*(1), 7.

A "cell theory" for networks?

"For example, **what does existing sociological network theory**, built mostly on a foundation of one-time "snapshot" data, typically with only dozens of people, **tell us about massively longitudinal data sets of millions of people**, including location, financial transactions, and communications? These vast, emerging data sets on how people interact **surely offer qualitatively new perspectives** on collective human behavior, but our current paradigms may not be receptive."

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., Alstyne, M. A. (2009) Computational social science. *Science*, *323*(5915), 721–723.

Cells *described* in 1665; *cell theory* in 1830s!
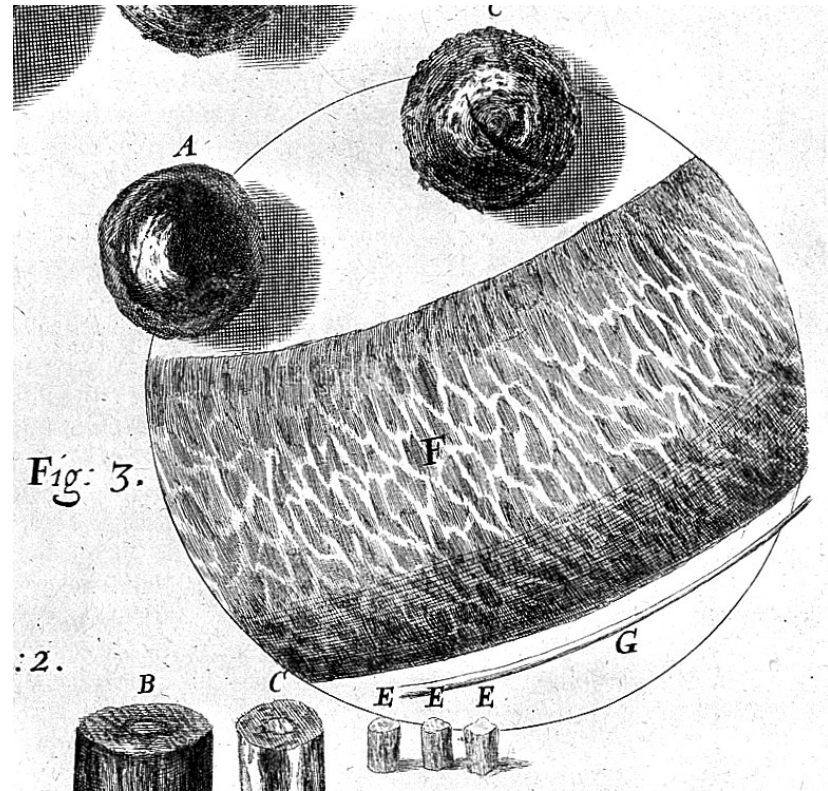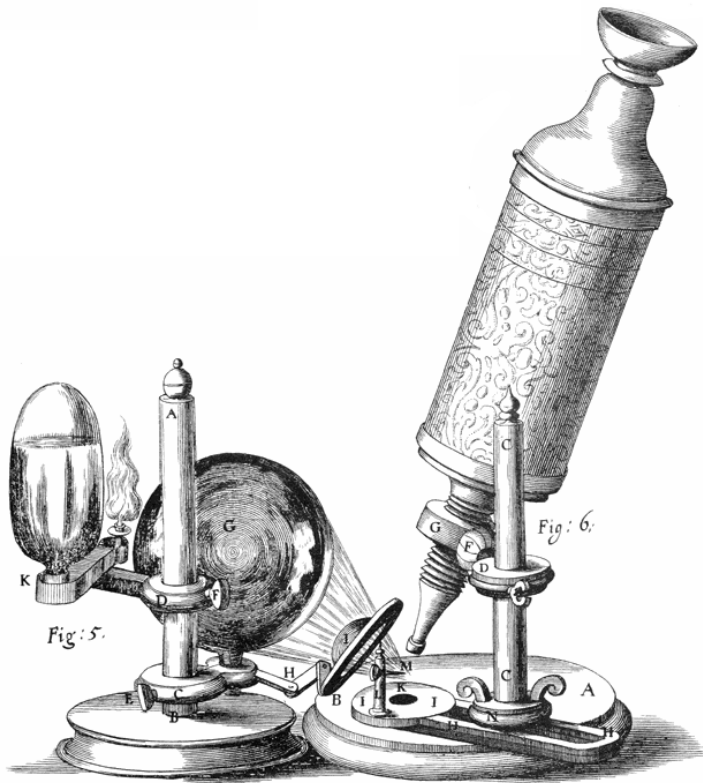
Hooke, R. (1665). *Micrographia: or some phyſiological deſcriptions of minute bodies made by magnifying glasses. With observations and inquiries thereupon*. London: J. Martyn and J. Allestry.

Szekely, F. (2011). Unreliable observers, flawed instruments, 'disciplined viewings': Handling specimens in early modern microscopy. *Parergon, 28*(1), 155–176.

Epstein, B. (2015). *The ant trap: Rebuilding the foundations of the social sciences*. Oxford University Press.

Laboratory for Optical and Computational Instrumentation. (2015). History of the light microscope. Microscopy Museum, University of Wisconsin-Madison. http://loci.wisc.edu/outreach/history-light-microscope

Understand the tool!

Think of differences between Twitter, Instagram, Facebook, Reddit, 4chan…

- Norms of use, platform-specific culture and behavior
- Adoption patterns, demographics of user base
- Access constraints, platform-side filtering

Gayo-Avello, D. (2011). Don't turn social media into another 'Literary Digest' Poll, *Communications of the ACM*, *54*(10), 121-128.
Tufecki, Z. (2014). Big questions for social media big data: Representativeness, validity, and other methodological pitfalls. ICWSM-14, 505-514.
Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior, Science, 346(6213), 1063-1064.
Harigttai, E. (2015). Is bigger always better? Potential biases of big data derived from social network sites. *Annals of the American Academy of Political and Social Science*, *659*, 63-76.

Remaining blind spot: Companies themselves

Social media platforms are not neutral utilities nor research environments

They are corporations, concerned about markets and their business model

They are constantly engineering and re-engineering platforms to *encourage desirable behavior* from their users

- Spending more time on the platform
- Giving up more data

Van Dijjck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press.
Gehl, R. W. (2014). *Reverse engineering social media: Software, culture, and political economy in new media capitalism*. Temple University Press.

We call the success of such engineering attempts *platform effects*.

Key question: *How much of observed behavior is due to platform effects?*

Our approach

Exploit data artifacts!

Two examples of *discontinuities* in previously collected data, where some platform change resulted in a change in the observed behavior

Apply *Regression Discontinuity* (RD) design: observational inference technique that estimates causal impact by fitting curves to either side of the discontinuity, and measuring the distance between them

(Time series methods that account for temporal autocorrelation aren't compatible with RD design; we use quantile regression to get fitted tolerance intervals)

Data

Not technically social media, but a nice discontinuity in Netflix Prize Data Set noticed by Koren (2009)

Facebook New Orleans data set crawled by Viswanath, Mislove, Cha and Gummadi (2009)

- Publically available profiles in FB New Orleans, ≈52% of all profiles there (multiple issues in boundary specifications, other issues in data)

- About 800,000 unique edges

Koren, Y. (2009). Collaborative filtering with temporal dynamics. KDD '09, 447–456.
Viswanath, B., Mislove, A., Cha, M., & Gummadi, K. P. (2009). On the evolution of user interaction in Facebook. WOSN '09, 37–42.

Netflix average daily ratings
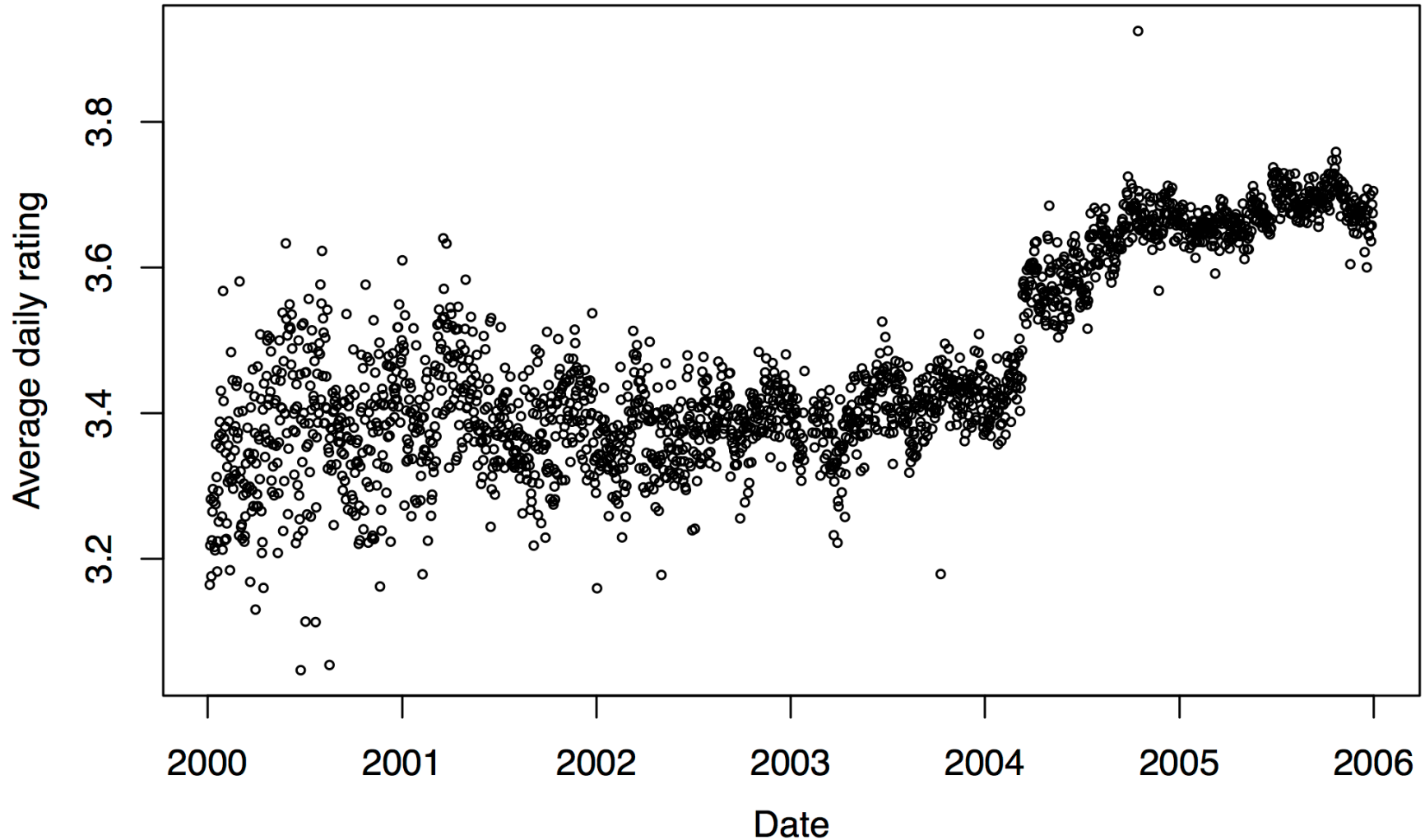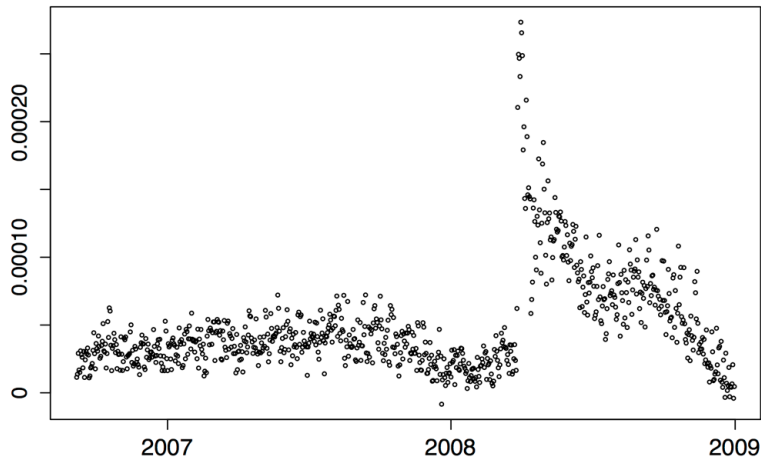
Facebook network statistic time series
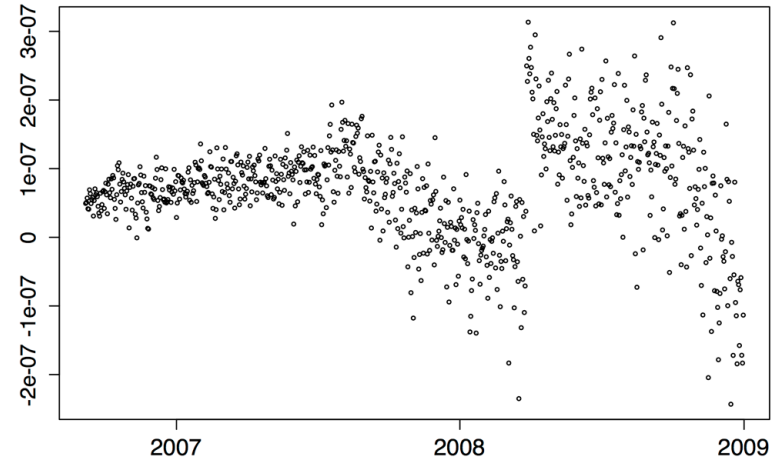

Daily added edges
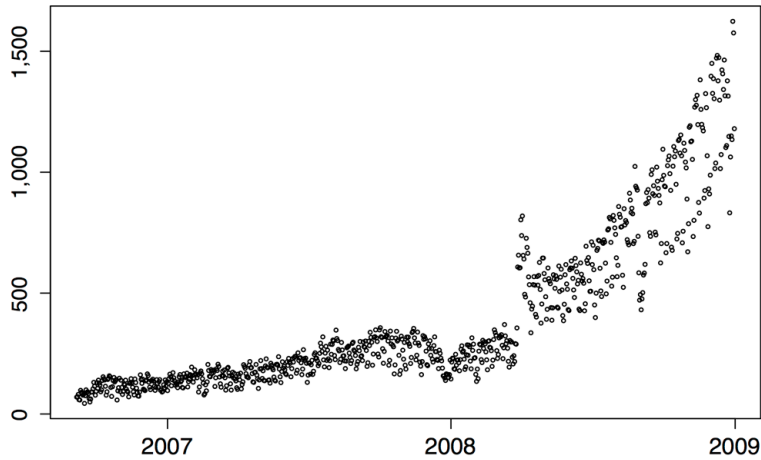
Daily added triangles
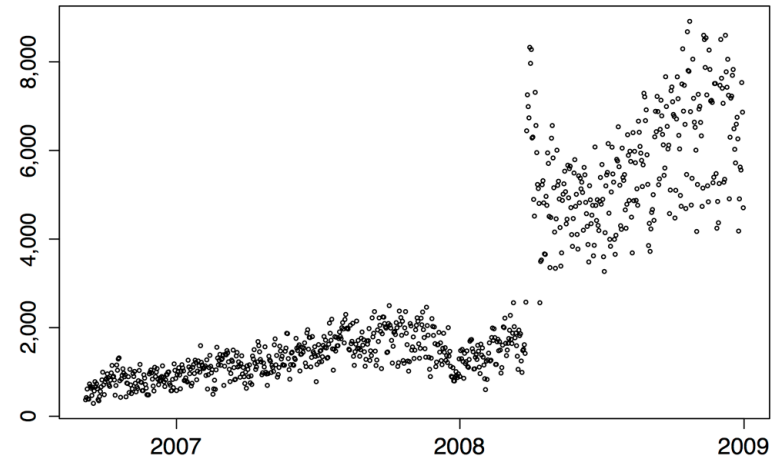
Daily change in transitivity
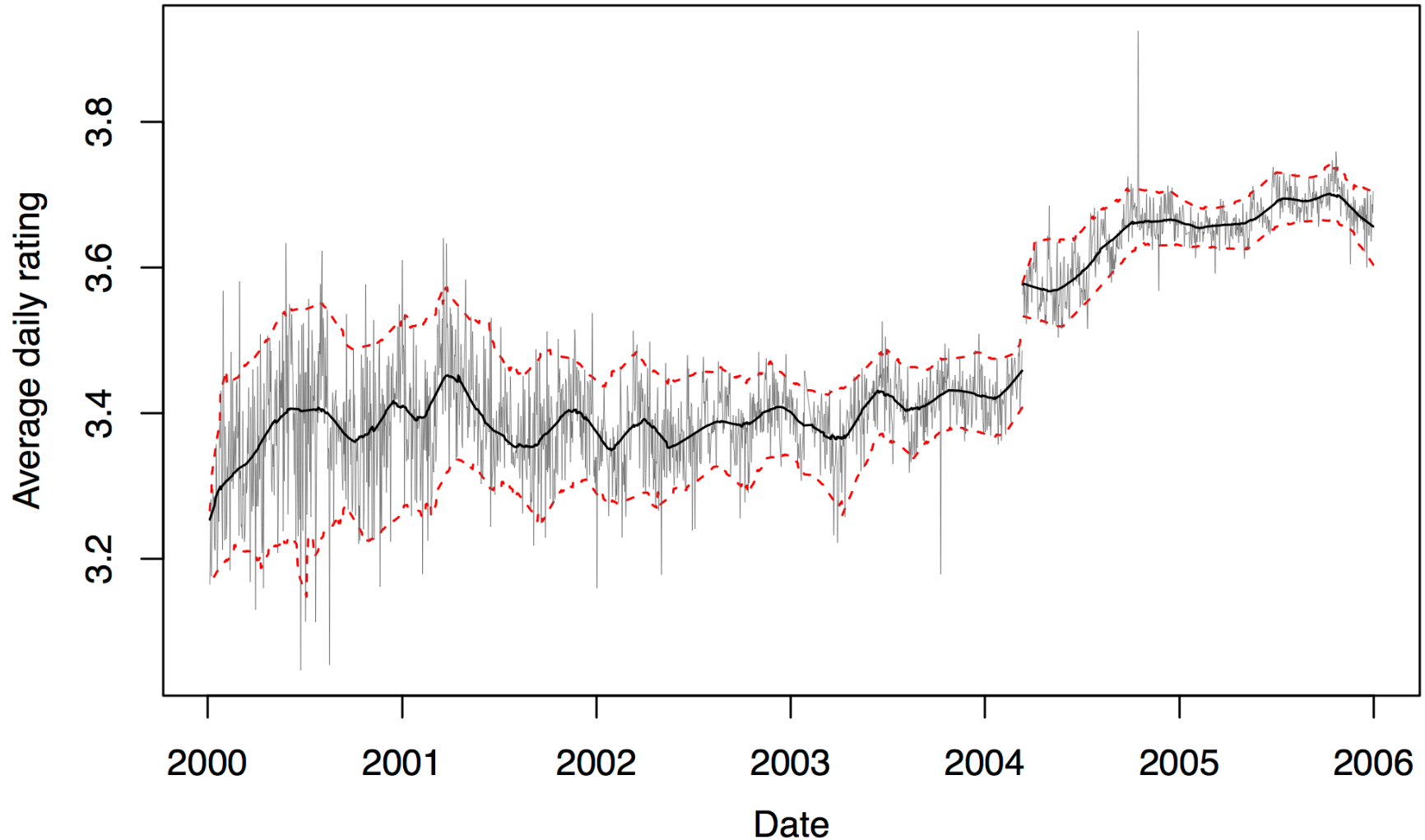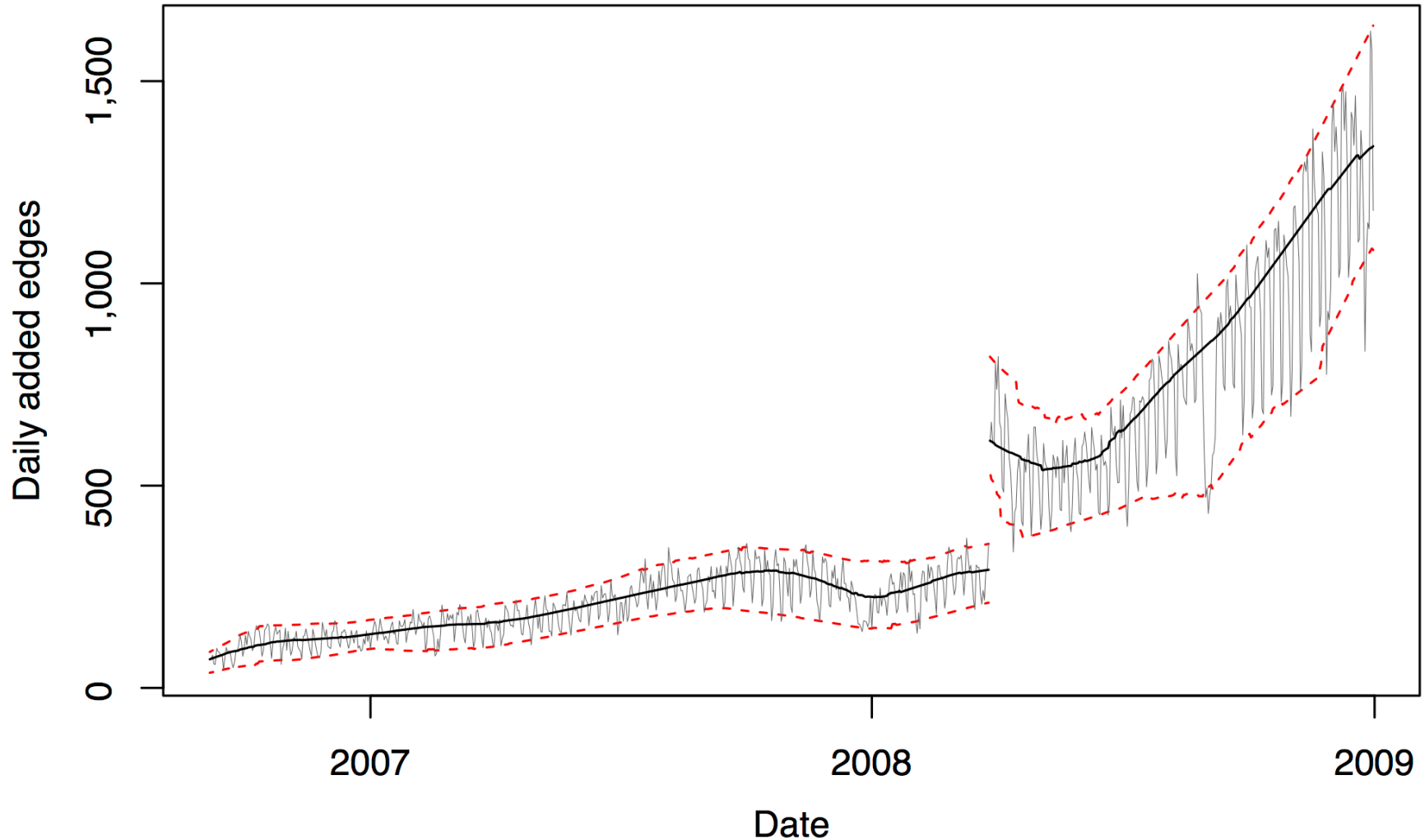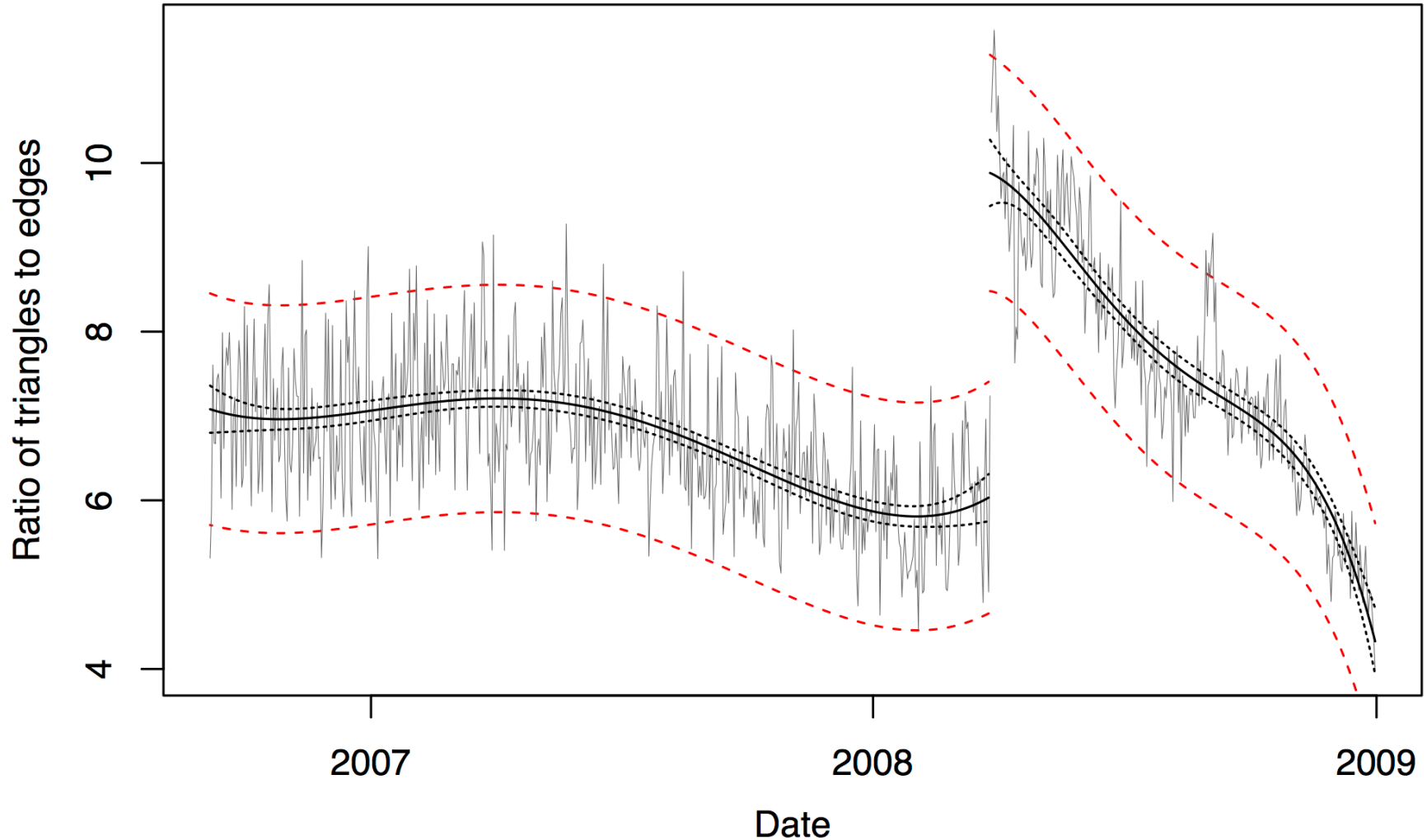
Daily change in density

Findings

Netflix: ≈ +.12 average daily rating (+3%)

Facebook: ≈ +300 edges per day (+200%)!

Facebook: ≈ +3.8 daily triangles per edge!

Limitations

Netflix:

- Source of shock unknown. Not social media or a network.

Facebook:

- Not making a network model; and, first differencing not enough to make the time series of observed network statistics stationary

- Major questions about data quality, although not critical

- Are social media ties the right thing to measure? (E.g., measure *interaction* instead?)

Conclusions

Study platform effects! They change what we will think is happening.

"Data artifacts" are not curiosities or annoyances, they can be deeply theoretically revealing about platform effects
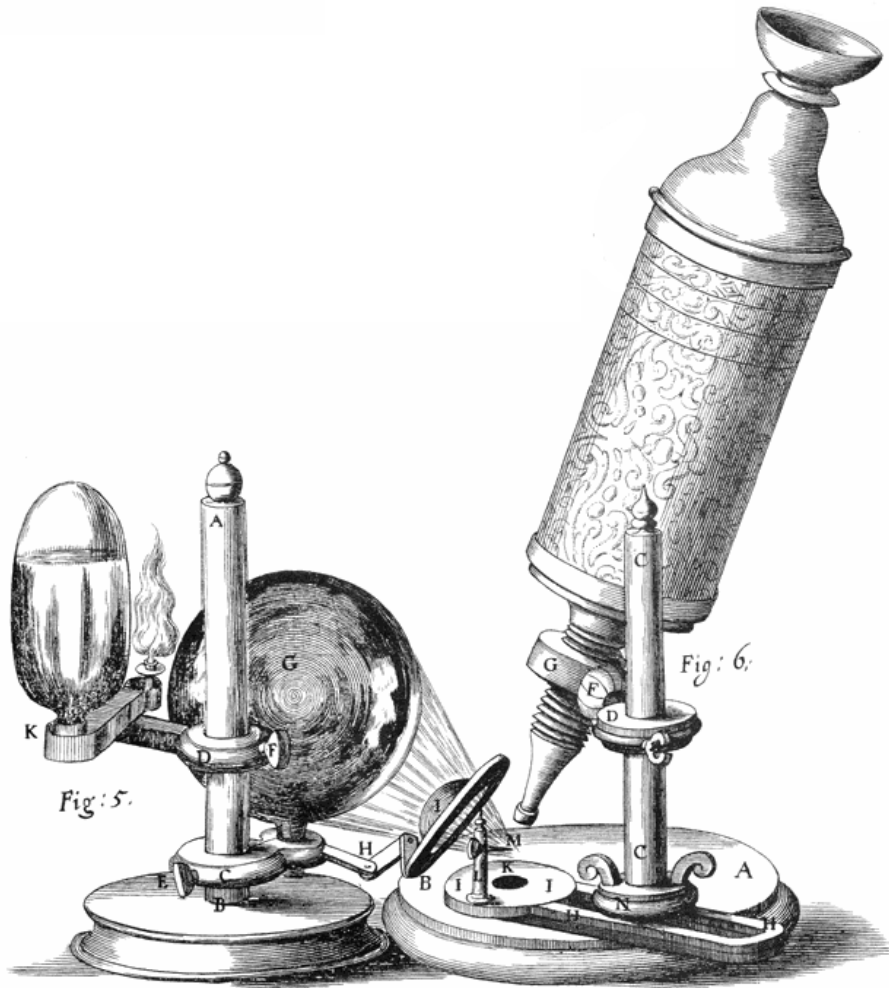
As external researchers, apply observational inference techniques to data artifacts for researching platform effects
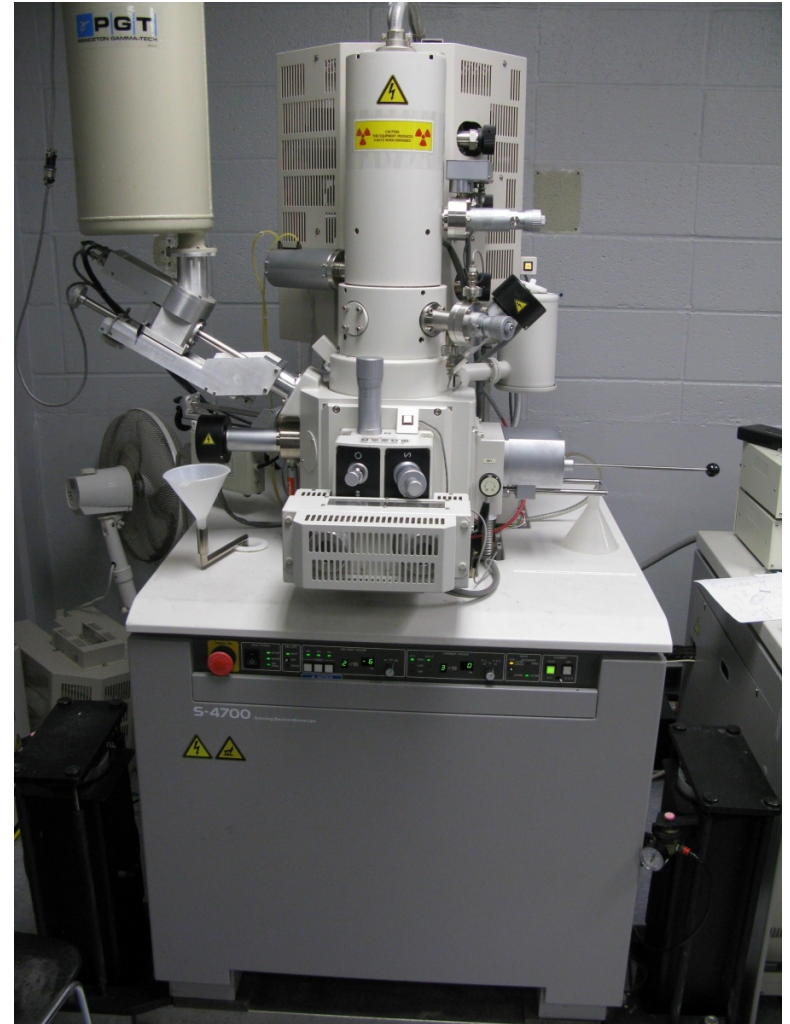
And…

Don't mistake what the instrument measures for the underlying phenomenon!

In order to know what/how to measure, we already have to have a pretty good idea of what we are looking for

Much of science involves improving the tools and how we use them, as we understand what we are trying to study

≠

Thank you! Questions?

momin@cmu.edu
http://mominmalik.com/sunbelt2016.pdf
http://mominmalik.com/icwsm2016pre.pdf