USC Viterbi
School of Engineering

Using Simpson's Paradox to Discover Interesting Patterns in Behavioral Data

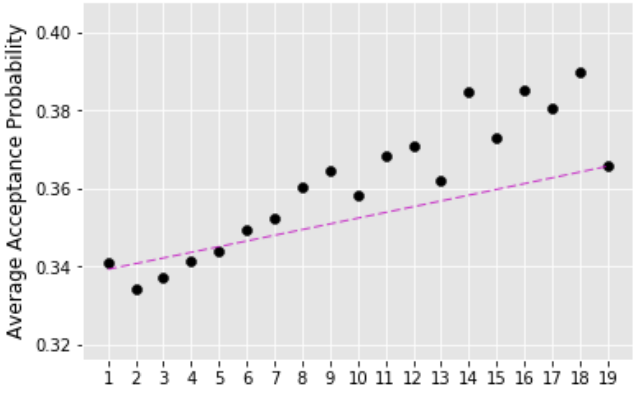
Kristina Lerman
USC Information Sciences Institute

Thanks to **Nazanin Alipourfard, Peter G. Fennell,**
IC2S2 2018

http://www.isi.edu/~lerman

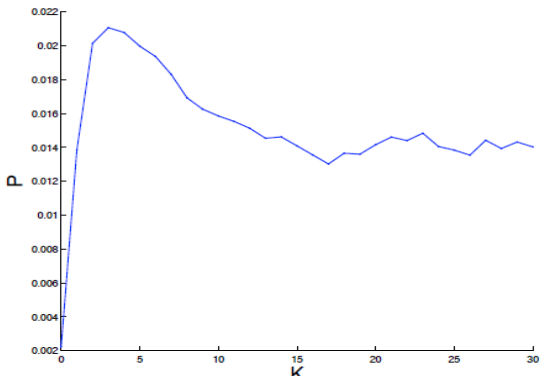
Can you trust the trend?

Do Stack Exchange users write better answers later in a session? [1]



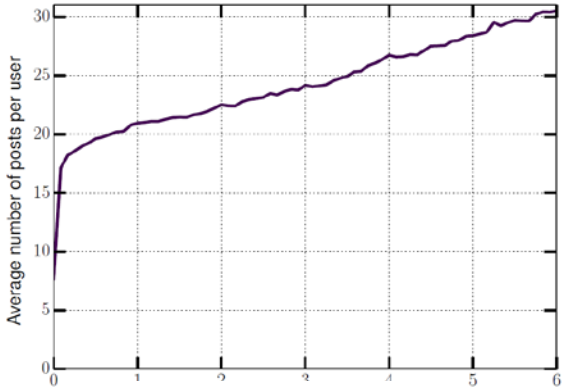
Answer position within session

Do additional exposures by friends suppress a Twitter user's use of a hashtag?[2]



Number of tweeting friends

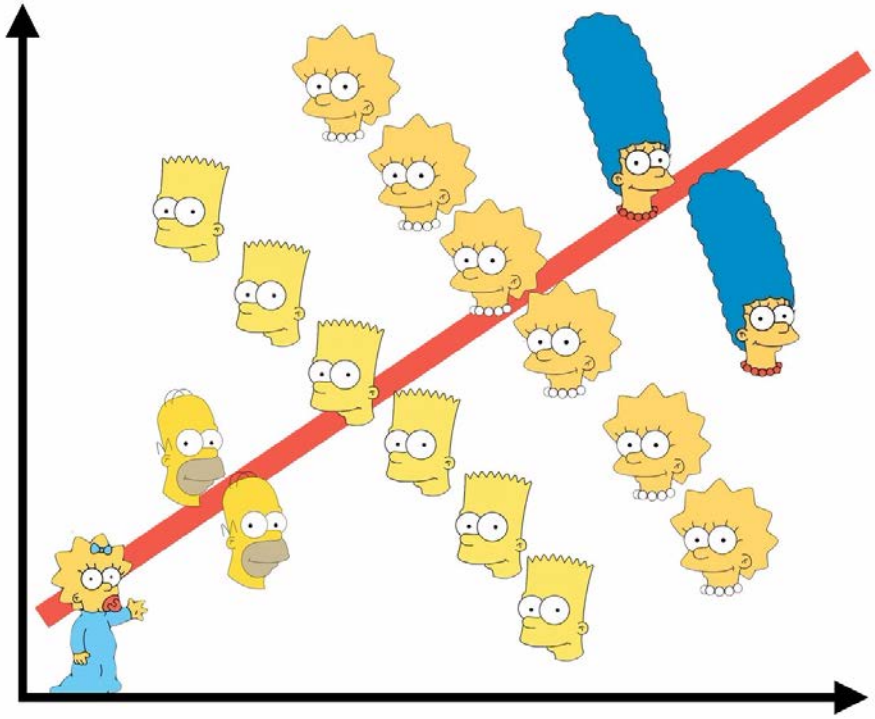
Do Reddit users become more active over time?[3]



Time since join date

1. Alipourfard, Fennell, & Lerman (2018) "Can you trust the trend: Discovering simpson's paradoxes in social data" in WSDM.
2. Romero, Meeder & Kleinberg (2011) "Differences in the Mechanics of Information Diffusion Across Topics" in *WWW*.
3. Barbosa et al. (2016) "Averaging gone wrong: Using time-aware analyses to better understand behavior." in *WWW*.

SIMPSON'S PARADOX



A TREND APPEARS IN DIFFERENT SUB-GROUPS OF DATA BUT DISAPPEARS OR REVERSES WHEN THESE SUB-GROUPS ARE COMBINED.*

* Simpson (1951). "The Interpretation of Interaction in Contingency Tables". *JRSS*

SIMPSON'S PARADOX, AN ILLUSTRATION

Which treatment should doctor recommend for kidney stones*?

	Treatment A	Treatment B
	78% (273/350)	**83%** (289/350)

*** Wikipedia**

SIMPSON'S PARADOX , AN ILLUSTRATION

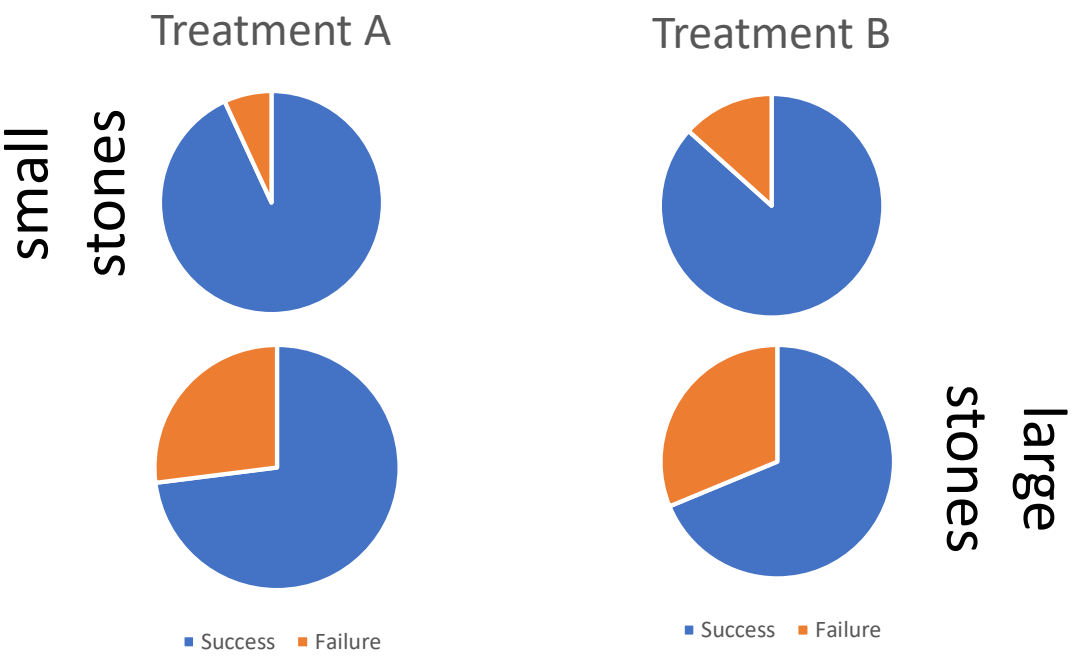
Which treatment should doctor recommend for kidney stones*? After accounting for the confounder—stone size—the best choice reverses.

	Treatment A	Treatment B
Small Stones	**93%** (81/87)	87% (234/270)
Large Stones	**73%** (192/263)	69% (55/80)
All	78% (273/350)	**83%** (289/350)

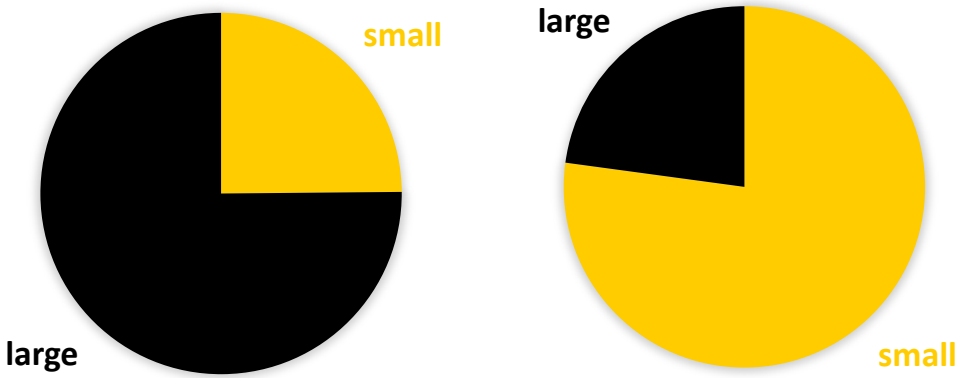
*** Wikipedia**

WHY SIMPSON'S REVERSAL OCCURS

- **Small stones are more easily treated, especially by treatment A**



- **But are overwhelmingly assigned to treatment B**



Central ideas and papers

- Simpson's paradox implies that functional differences exist within the population.

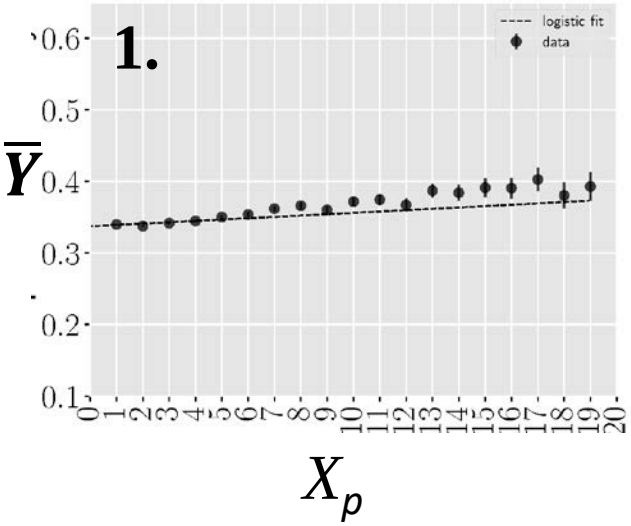
  E.g., small kidney stones are easier to treat

  - "Computational Social Scientist Beware: Simpson's Paradox in Behavioral Data", in *J. Computational Social* Science 2018
- We developed an algorithm to automatically identify Simpson's paradoxes data.
  - "Using Simpson's paradox to discover interesting behavioral patterns in data" in *ICWSM* 2018
  - "Can you Trust the Trend? Discovering Simpson's Paradoxes in Social Data" in *WSDM* 2018

Method to discover Simpson's paradoxes in data

For each covariate X_p

- Step 1: Estimate trend of outcome Y with respect to a covariate X_p



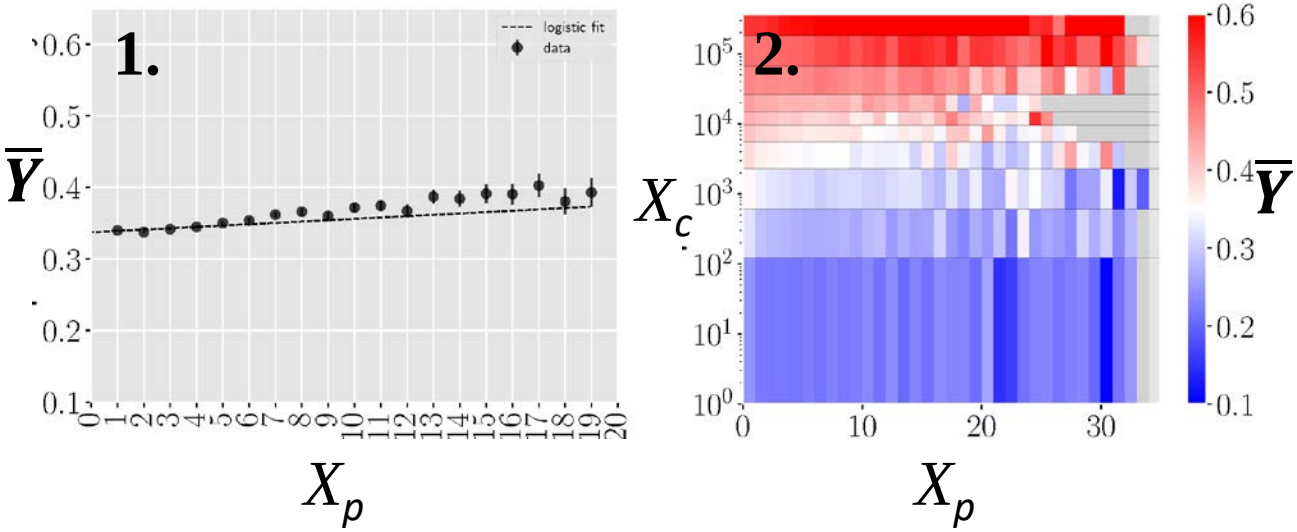
Method to discover Simpson's paradoxes in data

For each covariate X_p

- Step 1: Estimate trend of outcome Y with respect to a covariate X_p

  For each remaining covariate X_c

  - Step 2: Disaggregate data by conditioning on X_c



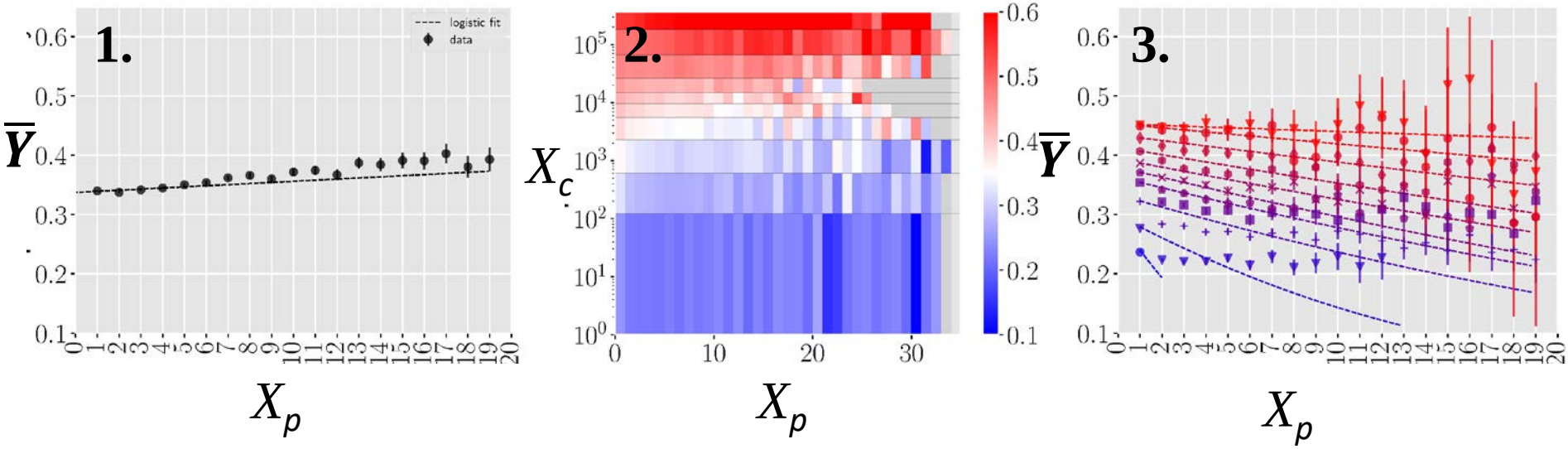
Method to discover Simpson's paradoxes in data

For each covariate X_p

- Step 1: Estimate trend of outcome Y with respect to a covariate X_p

  For each remaining covariate X_c

  - Step 2: Disaggregate data by conditioning on X_c
  - Step 3: Compare trends of the outcome within disaggregated bins to the trend in the aggregated data



Empirical validations

Problem solving
- 2 years

Outcome Y:
- Is the problem solved correctly on the first attempt?

11 features:
- Day, month, first five, number of problems solved, …

Question answering
- 6 years

Outcome Y:
- Will the answer be accepted as best answer by the asker?

19 features:
- words, code lines, session length, reputation, tenure, number of answers written, …

Language learning
- 2 weeks

Outcome Y:
- Are all words correctly recalled in a lesson

22 features:
- Day, month, session length, first five, distinct words, hour24, …

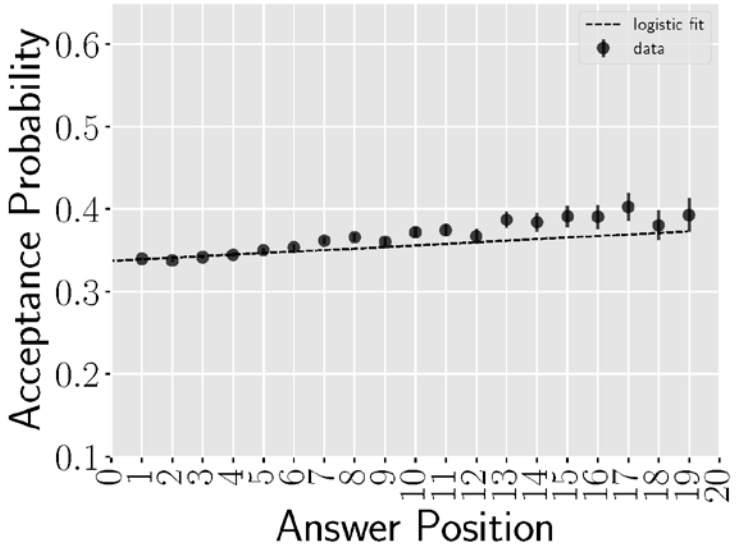
Stack Overflow – Simpson's pairs

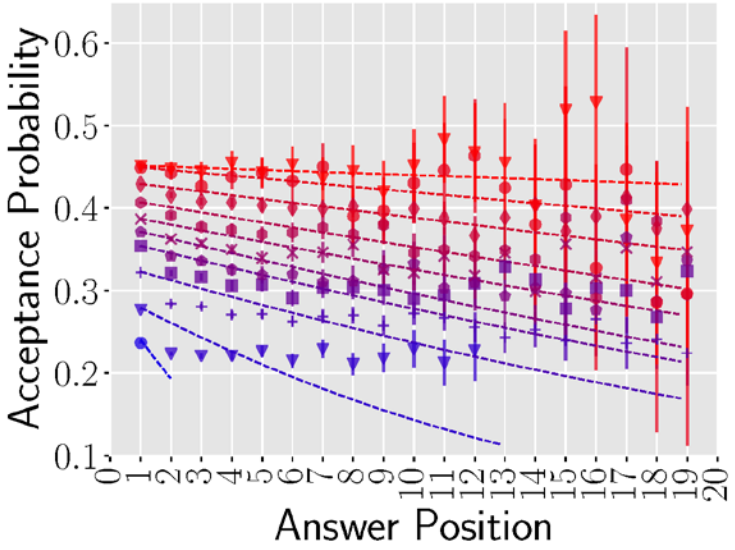
Pseudo-R^2	Covariate of Performance	Conditioning variable
0.03	Answer Position w/in Session	Experience (answers written)
0.03	Session Length	Experience (answers written)
0.02	Experience	Reputation
0.02	Answer Position w/in Session	Reputation
0.02	Session Length	Reputation
<0.01	Answer Position w/in Session	Session Length
<0.01	Time since Previous Answer	Answer Position w/in Session

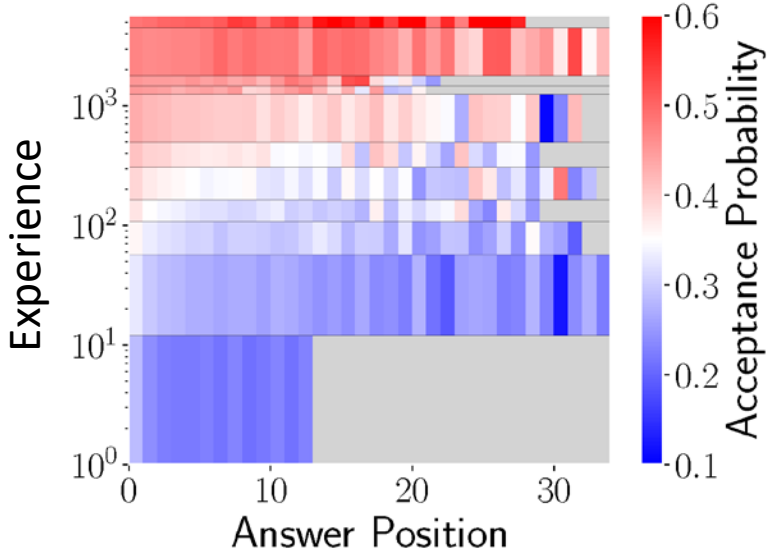
*

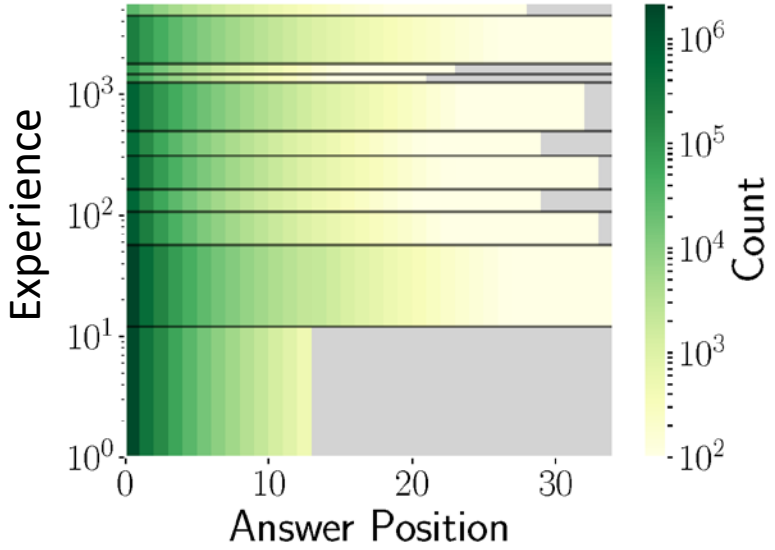
* Ferrara, et al. (2017) "Dynamics of content quality in collaborative knowledge production" in *ICWSM*.

Stack Overflow Simpson's Paradox





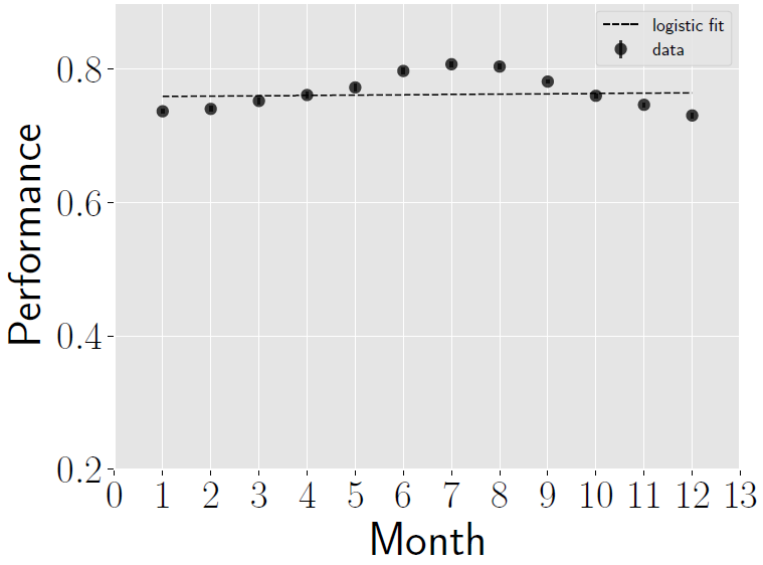


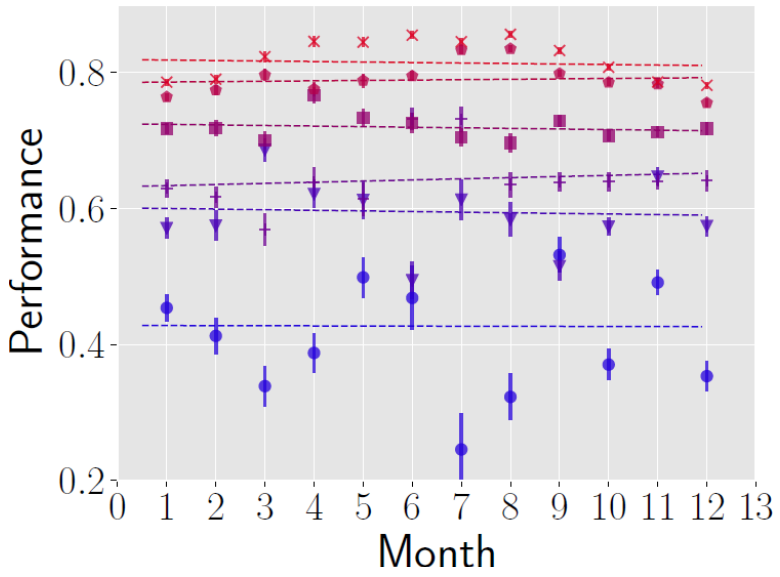


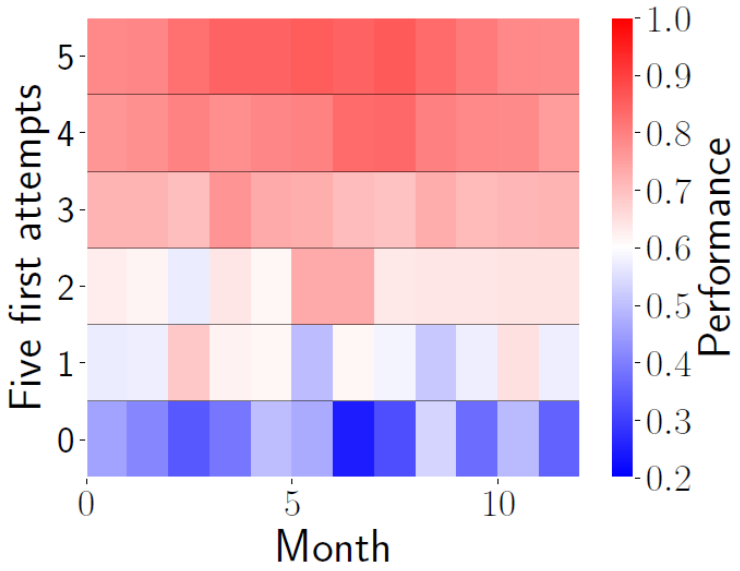
Khan Academy – Simpson's pairs

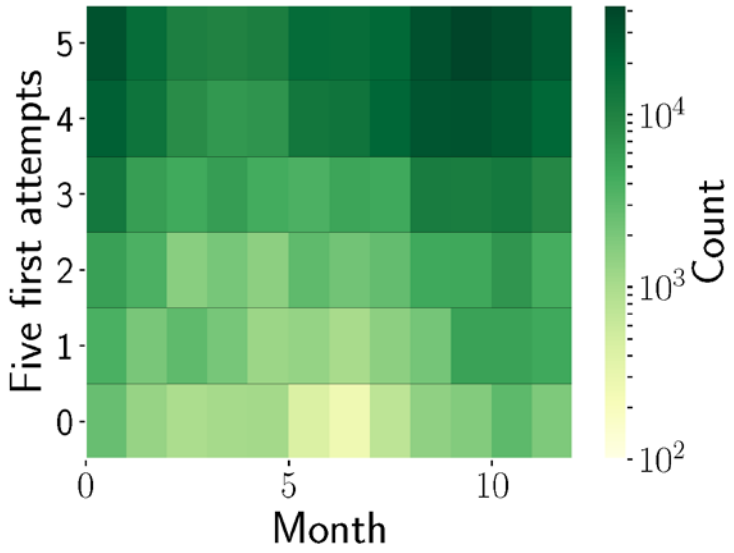
Pseudo-R^2	Covariate of Performance	Conditioning variable
0.04	Hour24	Problems Correct
0.03	Month	Five First Attempts
0.01	Month	Session Index
0.01	Month	Total Solve Time
0.02	Session Number	Number of Problems
0.01	Session Number	Tenure
0.01	Session Number	Total Solve Time

Khan Academy Simpson's Paradox



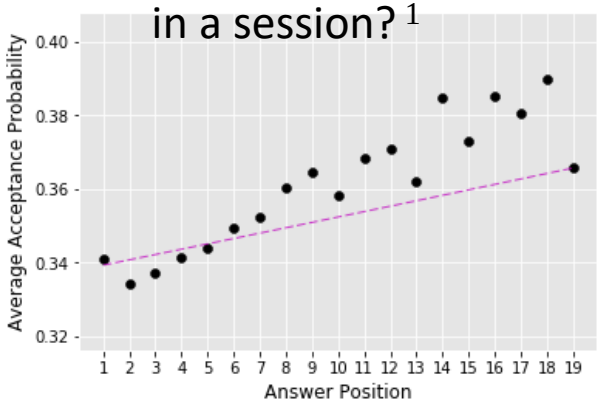




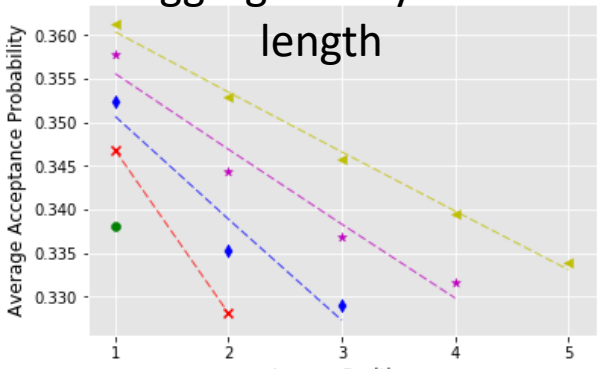


Can you trust the trend?

Do Stack Exchange users write better answers later in a session? [1]

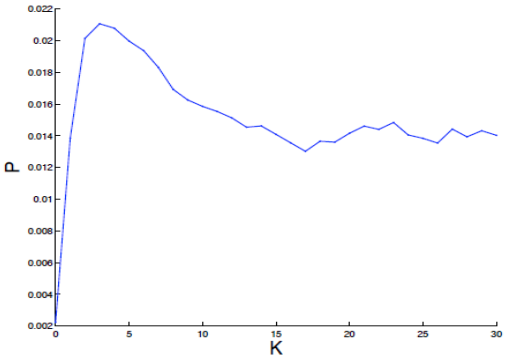


Disaggregated by session length

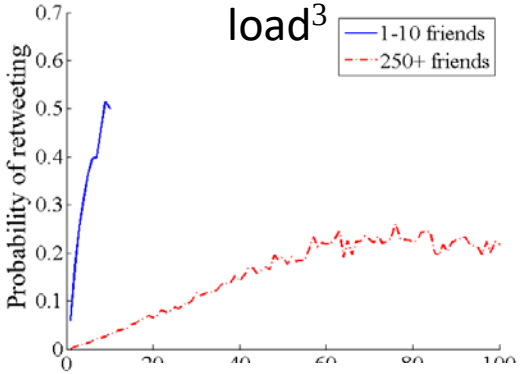


Answer position within session

Do additional exposures by friends suppress a Twitter user's response?[2]

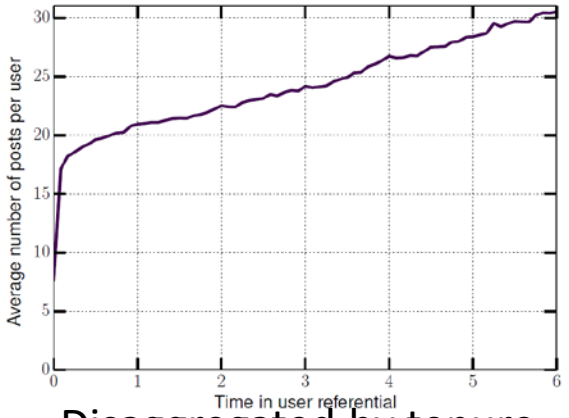


Disaggregated by cognitive load[3]

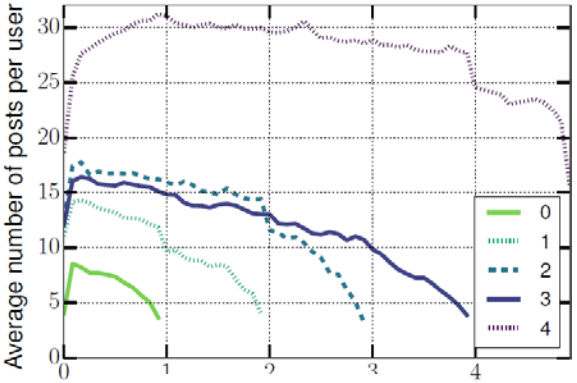


Number of tweeting friends

Do Reddit users become more active over time?[4]



Disaggregated by tenure



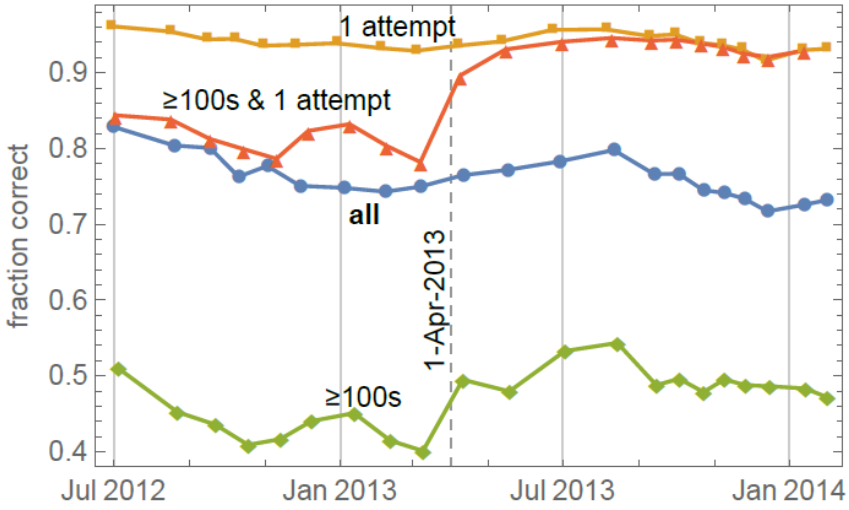
Time since join date

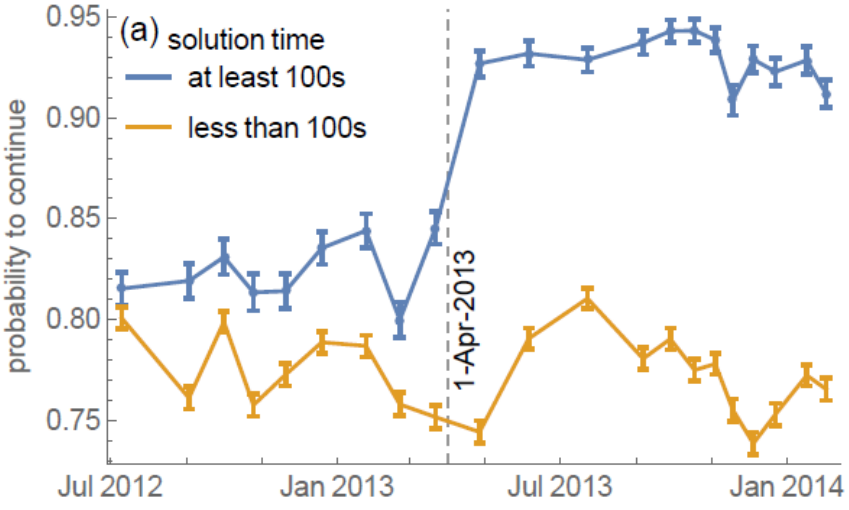
1. Alipourfard, et al. (2018) "Can you trust the trend: Discovering simpson's paradoxes in social data" in WSDM.
2. Romero et al. (2011) "Differences in the Mechanics of Information Diffusion Across Topics" in *WWW*.
3. Hodas & Lerman (2012) "How visibility and divided attention constrain social contagion", in SocialCom.]
4. Barbosa et al. (2016) "Averaging gone wrong: Using time-aware analyses to better understand behavior." in *WWW*.

Algorithm to find natural experiments in data

Performance over time in Khan Academy:
a) performance declines in aggregate; but
b) increases for 'slow' users. Subgroups automatically identified by our algorithm.

Slow users become more persistent: they are more likely to continue working on problems they got wrong on their first attempt.





Hypothesis: User interface change in Khan Academy in April 2013 made slow users more "gritty": more likely to continue working on a problem they got wrong. As a result, they performance on other problems increases.

To summarize

https://github.com/ninoch/Trend-Simpsons-Paradox/

- Simpson's paradox occurs when an association/trend observed in the subgroups disappears or reverses when the subgroups are combined into one.
- Algorithm to automatically identify subgroups with different trends
  - A tool for data-driven discovery
  - And to formulate new hypotheses about data.
- Algorithm available!

  https://github.com/ninoch/Trend-Simpsons-Paradox/
  - Works for binary outcomes, linear models

THANK YOU!

Sponsors

NSF: CIF-1217605

ARO: W911NF-15-1-0142, W911NF-16-1-0306

Questions?

lerman@isi.edu