

# Causal Inference over Longitudinal Data to Support Expectation Exploration

**Emre Kiciman**

Microsoft Research

emrek@microsoft.com

@emrek



should i **buy a fixer upper or new home?**

should i **rent or buy a home in 2014**

should i **get preapproved before house hunting**

should i **refinance my mortgage**

should i **get a divorce**

should i **break up with my boyfriend**

should i **pay off my mortgage**

should i **file bankruptcy**

should i **get a flu shot**

should i **lease or buy a car**

should i **join a gym**

should i **eat before or after working out**

should i **text him**

should i **buy a chromebook**

should i **quit my job**

should i **pop a blister**

should i **get bangs**

should i **wash my hair before coloring it**

should i **buy a house**

should i **cut my hair short**

should i **retire at 65**

should i **go to college**

should i **go to law school**

should i **take a multivitamin**

should i **text him or wait for him to text me**

should i **join the military**

should i **leave my husband**

should i **get married**

should i **pop a burn blister**

should i **see a doctor**

should i **consolidate my student loans**

should i **do cardio before or after weights**

should i **get a tattoo**

should i **workout with a cold**

should i **join the army**

should i **buy an xbox one**

should i **move to florida**

should i **have a third child**

should i **take creatine**

should i **get a dog**

should i **run everyday**

should i **be a teacher**

should i **become a real estate agent**

should i **quit drinking coffee**

should i **shave my head**

# Understanding Outcomes of Situations and Actions

Everyday, people find themselves in unfamiliar situations and consider the potential outcomes of actions.

- Individuals ask themselves (and Bing!) “What’ll happen if I *do that?*”
- Policy-makers ask “What happens when someone *does that?*”

Today, no virtual assistant or knowledge base answers these questions.

- Individuals rely on articles, coaches, friends’ advice, and gut instinct.
- Policy-makers rely on experiments, data analyses, and gut instinct.

# Capturing outcomes from longitudinal social media data

The information we need is already being recorded

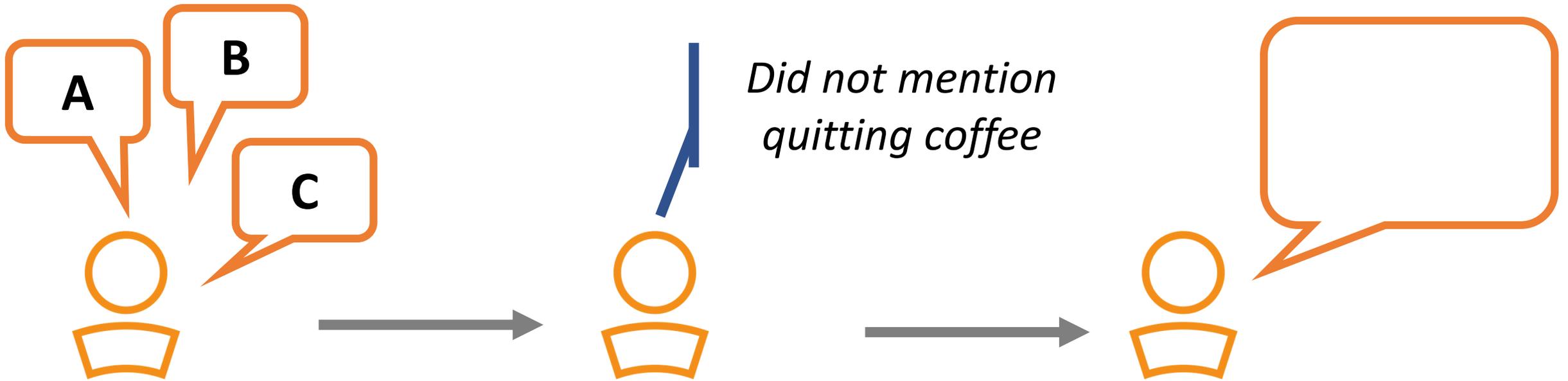
- Millions of people frequently and publicly report the actions they take and, over time, the outcomes they experience
- ... and they have been doing this for years

We can *compute* answers to people's questions about an experience

- Use causal inference methods to compare timelines of people who reported the experience to timelines of people who did not.
- Applies to varied data: Twitter, LinkedIn, search logs, financial, ...

**What happens after ...**





# A new analysis task for IR

*To answer “what if”, “should I”, and other expectation exploration tasks*

**Corpus: Rich, Individual-Level Longitudinal Data**

Today: Social media

Tomorrow: Many data sources

**Query: What happens after *experience T*?**

Need enough information to recognize “*people who did T*”

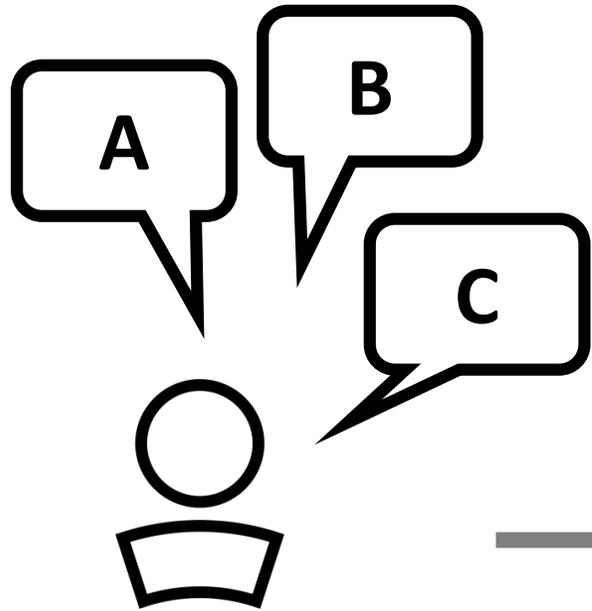
**Output: *Expectation Map* captures changes in outcomes over time**

Building block for many user experiences

# Rest of this talk

- Causal inference over longitudinal data
- Experiences applying to social sciences problems
- Applications to individual information needs

**Covariates/  
Confounds**



**Treatment**



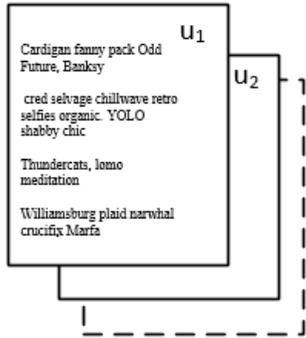
**Outcomes**



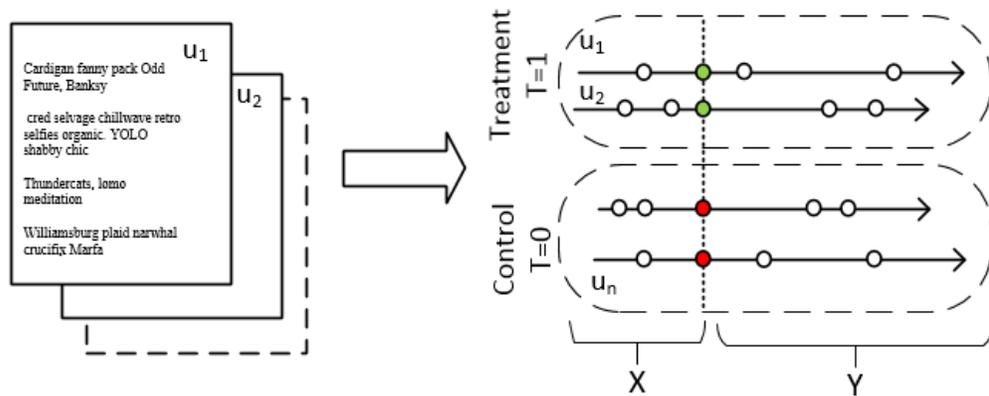
**Goal:** Estimate effect of a treatment  $T$  on an outcome  $Y$

Must break the dependence confounds  $X \rightarrow T$  (that is,  $T \perp\!\!\!\perp X$ )

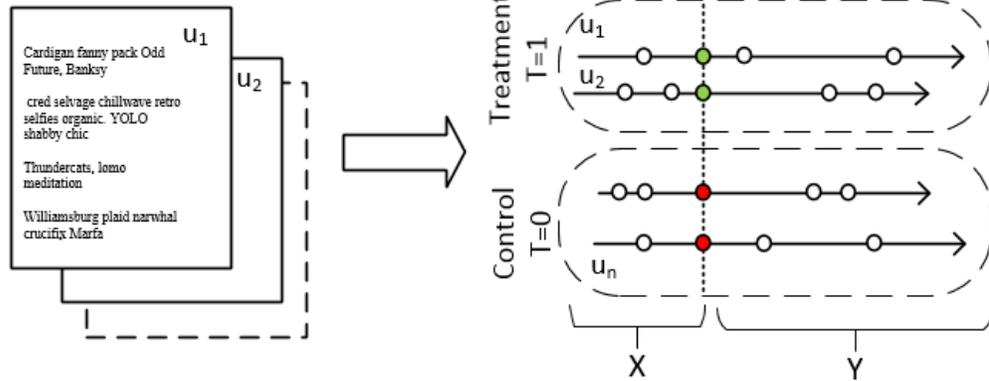
By construction, *randomized* experiments:  $T \perp\!\!\!\perp X$



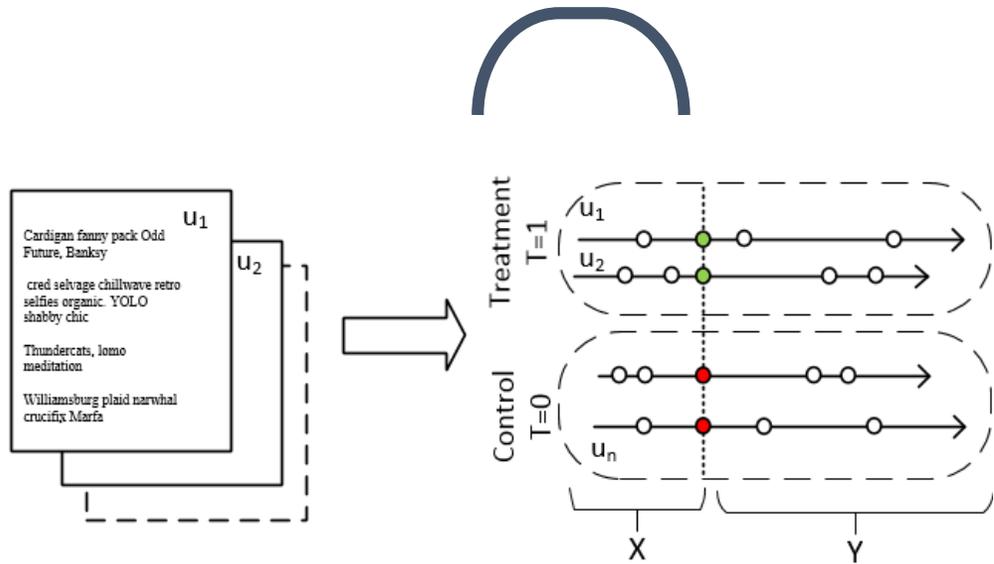
# 1. Build individual timelines of user activities



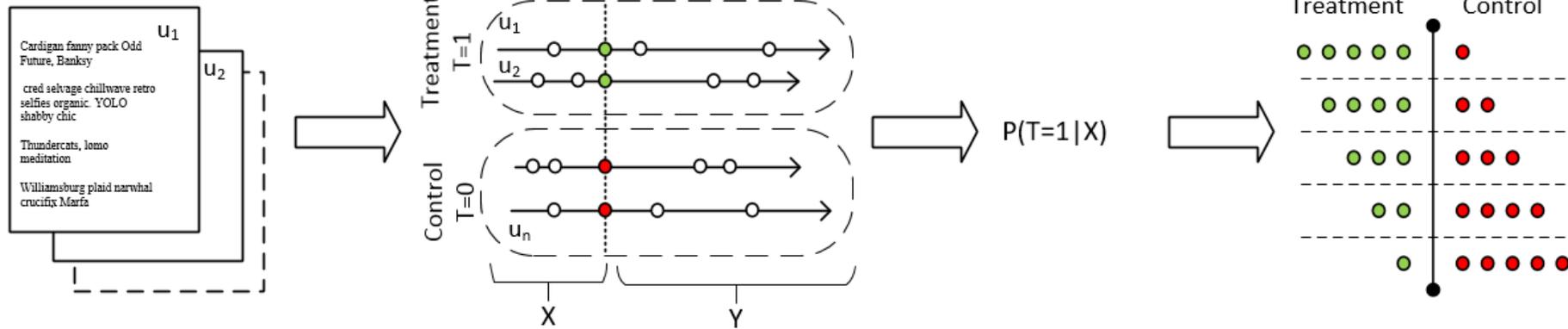
## 2. Treatment identification



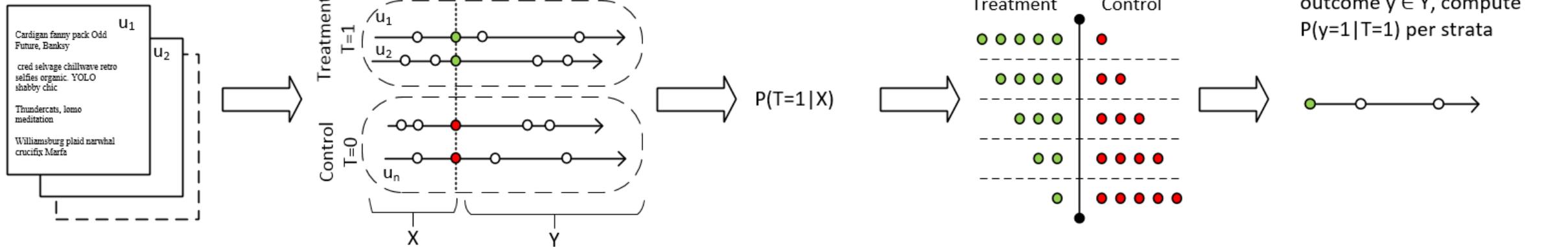
### 3. Extract covariates



## 4. Identify similar groups of users



## 5. Calculate population average outcomes



# Output: Expectation Map

Example treatment “ankle sprained”

Elapsed Time	Outcome Y	Effect
Day 1	crutches	50x
Day 1	play soccer	0.10x
...		
Day 42	crutches	1x

# Rest of this talk

- Causal inference over longitudinal data
- **Experiences applying to social sciences problems**
- Applications to individual information needs

# College

- Success in college predicts individual career success, career happiness and economic achievement.
- Drives macro-economic growth too.
- But **1 in 3 college students leaves without earning a degree**
- Many factors jeopardize college success: family responsibilities, financial pressures, individual behaviors





# Alcohol in College

- College students drink more than non-college peers
- Persistent public health issue: hangovers, lowered academic performance, DUI arrests, risky sexual behavior, sexual and other assaults.
- Excessive alcohol consumption is negatively associated with college success
- Studies of college behavior through self-report surveys and interviews; in-person or deliberate recruiting.

# What happens to students who mention alcohol more?

**Corpus:** 5 years of tweets by students starting college in Fall 2010

- 63k users, Aug 2010 – May 2015
- ~12 tweets/day per active person

**Treatment:** Mention alcohol keywords often during first semester

**Outcome:** Focus on topics known to be relevant to college success

**Table 3. Example tweets matching alcohol keywords**

---

@username hey! getting drunk is not the answer! but I will happily be by your side chipping in.

---

shouldn't have started drinking this wine :-)

---

beer in the fridge n I'm ready to go

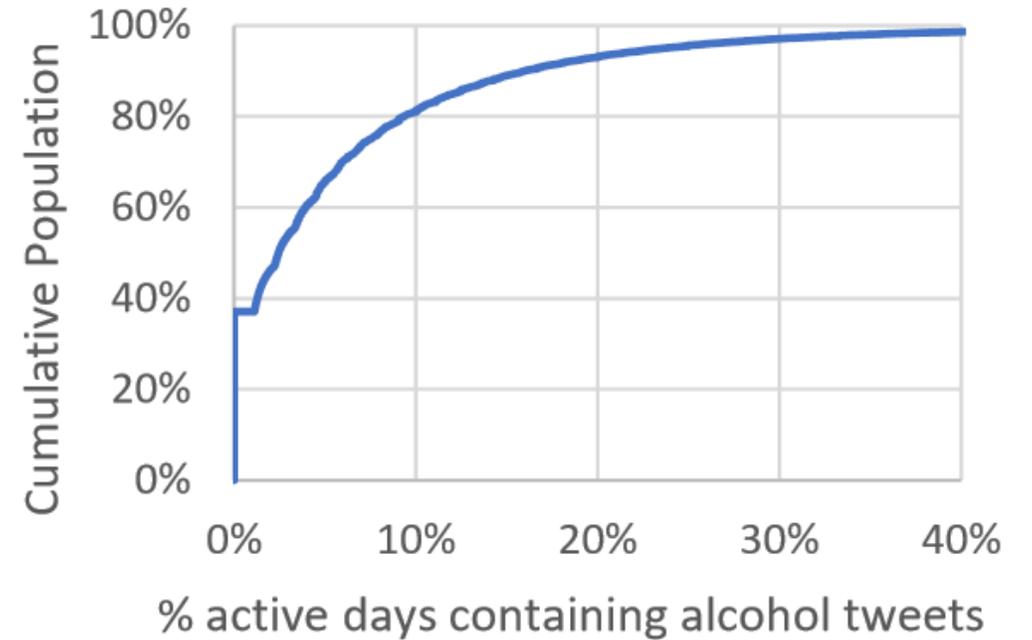
---

I hope the club gets snowed in and I'm left with all this vodka

---

need to do some serious pre drinking so my beer coat warms me up tonight

---



Identify alcohol keywords in Fall 2010 (9/15-12/15/2010)

Compare top 1/3 of alcohol mentioners (by % active days w/alcohol keywords)

... to bottom 1/3 of alcohol mentioners (no alcohol mentions)

# Outcome Topics

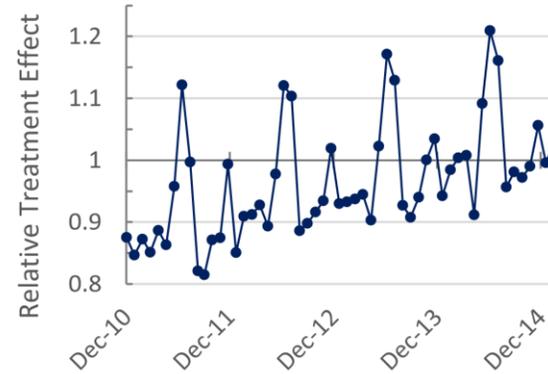
Table 2: Topics linked to college success

Concept	Seed words	Final topic words	Example Tweets
Peer group interaction	friend, boyfriend, girlfriend	boyfriend, buddy, roommate, bandmate, fiance, +20 <i>more</i>	Yup buddy I found my bandmate!!
Family responsibilities	mother, father, brother, sister	mother, father, brother, stepdad, grandmother, +21 <i>more</i>	Thankful for my little bro and mom I have a sister #fact
Study habits	study, library, homework	study, library, math, tutor, textbooks, worksheets, +49 <i>more</i>	anyone that wants to study for history we're in the library but anyways ima off to study
Financial pressures	debt, student loans, loans	wages, afford, utilities, tuition, evicted, fees, +28 <i>more</i>	finally my wages wooo @anon its all about money. Im in debt. dont want more loans
Legal/criminal challenges	police, cops, jail, parole	cops, police, restraining, probation, rehab, +15 <i>more</i>	meeting my parole officer cops pulling out breathalyzer f***k we drunk

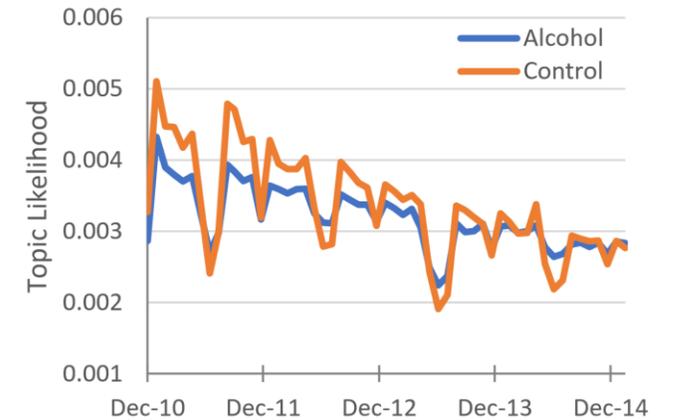
In addition, 195+ validated empath topics

# Academic effects

People in the Alcohol group were much less likely to mention studying over the next several years, and somewhat less likely over the entire time period



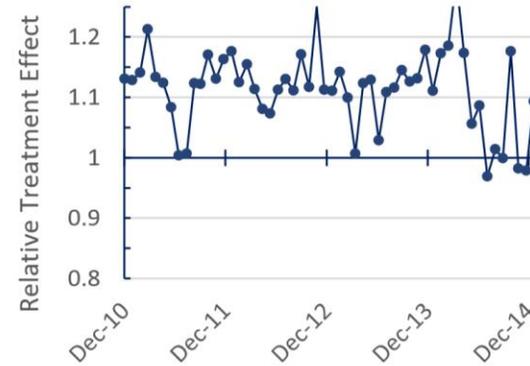
(a) Study Habits RTE



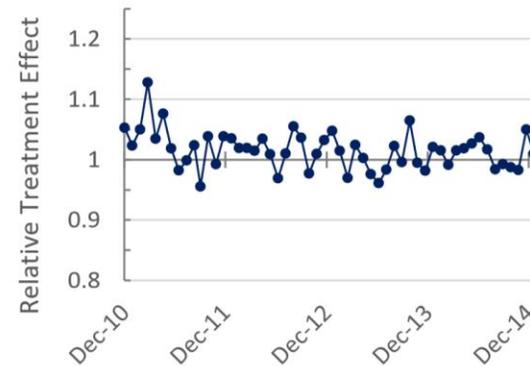
(b) Study Habits Outcomes

# Criminal and Financial effects

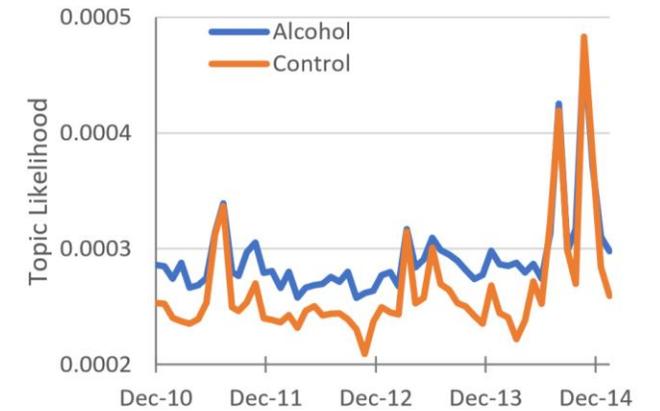
People in the Alcohol group were more likely to mention legal and criminal challenges through most of our study; and slightly more likely to mention financial pressures over time.



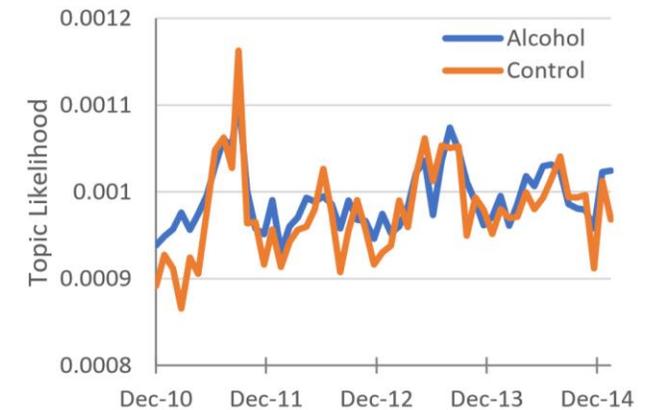
(a) Criminal RTE



(c) Financial RTE



(b) Criminal Outcomes



(d) Financial Outcomes

# Other analyses

- **Transitions from mental health discussions to suicidal ideation**
  - Explores heterogeneous effects [CHI16], and human assessments of balance [ICWSM17]
  - w/De Choudhury (GATech) et al.
- **Daily Behaviors and Sleep and Exercise Quality**
  - Combines query logs, location traces and MS Band data [JDSA 18]
  - w/Farajtabar, Nathan, and White
- **Conjunction of factors triggering waves of seasonal influenza**
  - Explores geographic flow, juxtaposition of multiple large-scale datasets, county-level matching [eLife 2018]
  - w/Chattopadhyay (UChicago), Elliott (UChicago), Shaman (Columbia), Rzhetsky (U Chicago)

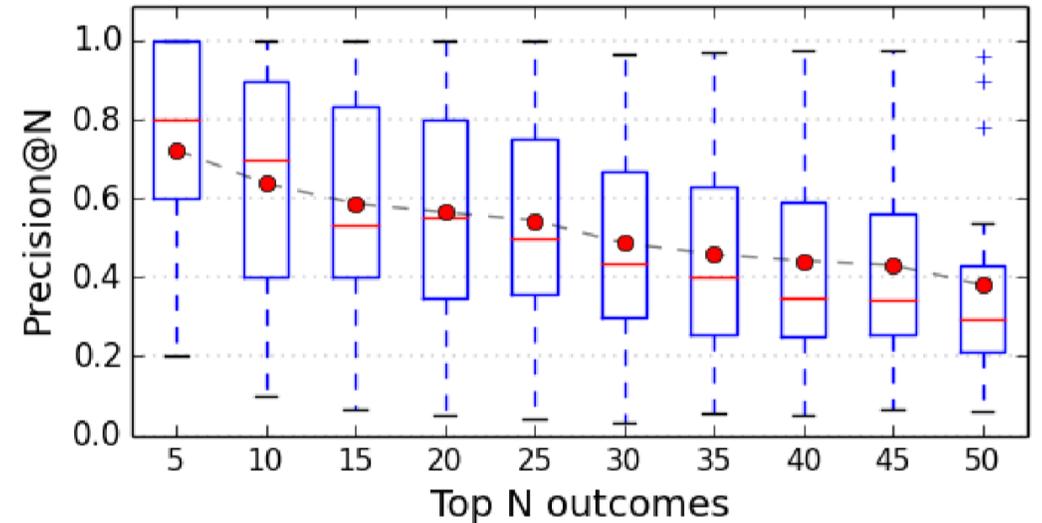
# Rest of this talk

- Causal inference over longitudinal data
- Experiences applying to social sciences problems
- **Applications to individual information needs**

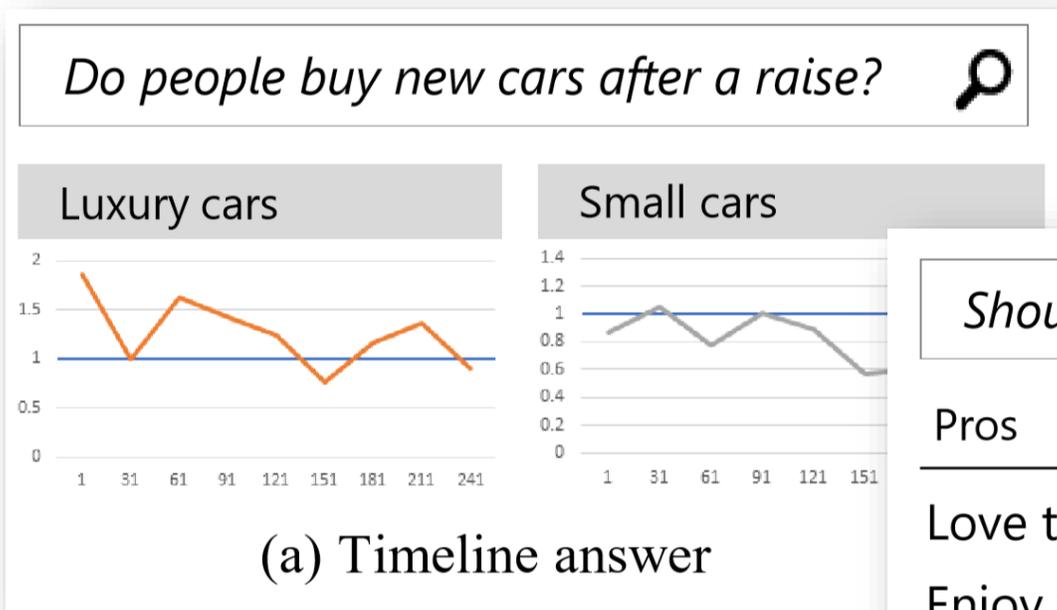
# Does this generalize?

- Calculate consequences of 39 experiences from business, health and society questions
  - E.g., people taking Prozac, getting a divorce, buying life insurance, investing money
  - Experiences selected from Bing search queries
- MTurk workers evaluate surface validity based on supporting social media messages, and online reference material

More broadly, key evaluation criteria must go beyond ***Correctness*** to include ***Interpretability and Usefulness***



# Building block for IR experiences



Should I get a dog? 🔍

Pros	Cons
Love the dog	Early wal
Enjoy walks	Scratche
...	...

(b) Pros / cons list

Hey, I sprained my ankle badly

When will I play football again?

People start to mention playing football after 8 weeks

(c) Conversational agent

# Opportunities for research

- Core methods: feature representation, data bias, performance
- Mapping from queries to treatment identification
  - And expanding class of treatments (multi-dose, network effects, ...)
- End-to-end task integration: Interpretability, representation, evaluation

# Summary

Everyone asks “what happens after ...”

- For decision support, for expectation setting, for exploration, ...

Causal inference over longitudinal, individual-level data can give answers

- A building block for new IR answers and experiences

Developed/validated methods in computational social science tasks

- Mental health; sleep and exercise quality; alcohol and college trajectories; ...

Now: Scale, interaction, interpretability to make broadly accessible

- Many exciting research opportunities

# Questions?

@emrek; [emrek@microsoft.com](mailto:emrek@microsoft.com)

Selected papers mentioned in the talk (all at <http://kiciman.org/> )

- Kiciman, Richardson, [Towards Decision Support and Goal Achievement: Identifying Action-Outcome Relationships from Social Media](#). KDD'15
- De Choudhury, Kiciman, Dredze, Coppersmith, Kumar, [Shifts to Suicidal Ideation from Mental Health Content in Social Media](#). CHI'16 *honorable mention*
- Olteanu, Varol, Kiciman, [Distilling the Outcomes of Personal Experiences: A Propensity-scored Analysis of Social Media](#). CSCW'17
- Kiciman, Counts, Gasser, [Using Longitudinal Social Media Analysis to Understand the Effects of Early College Alcohol Use](#). ICWSM'18
- **Kiciman, Thelin, [Answering \*What If, Should I\* and other Expectation Exploration Queries Using Causal Inference over Longitudinal Data](#). DESIRES'18**

Extra Slides

# Methodological Limitations of Social Media Analysis

- Population biases require measurement
- This analysis misses language meaning beyond tokens
  - Requires additional qualitative reading
- Studying changes in conversational behavior only
  - E.g., reporting biases

# Methodological Limitations of our Observational Study

- Assumes ignorability of unobserved confounds
- Assumes SUTVA (no network effects)
- Have not characterized heterogeneous effects
- Survivor bias
- Easiest to apply to point treatments