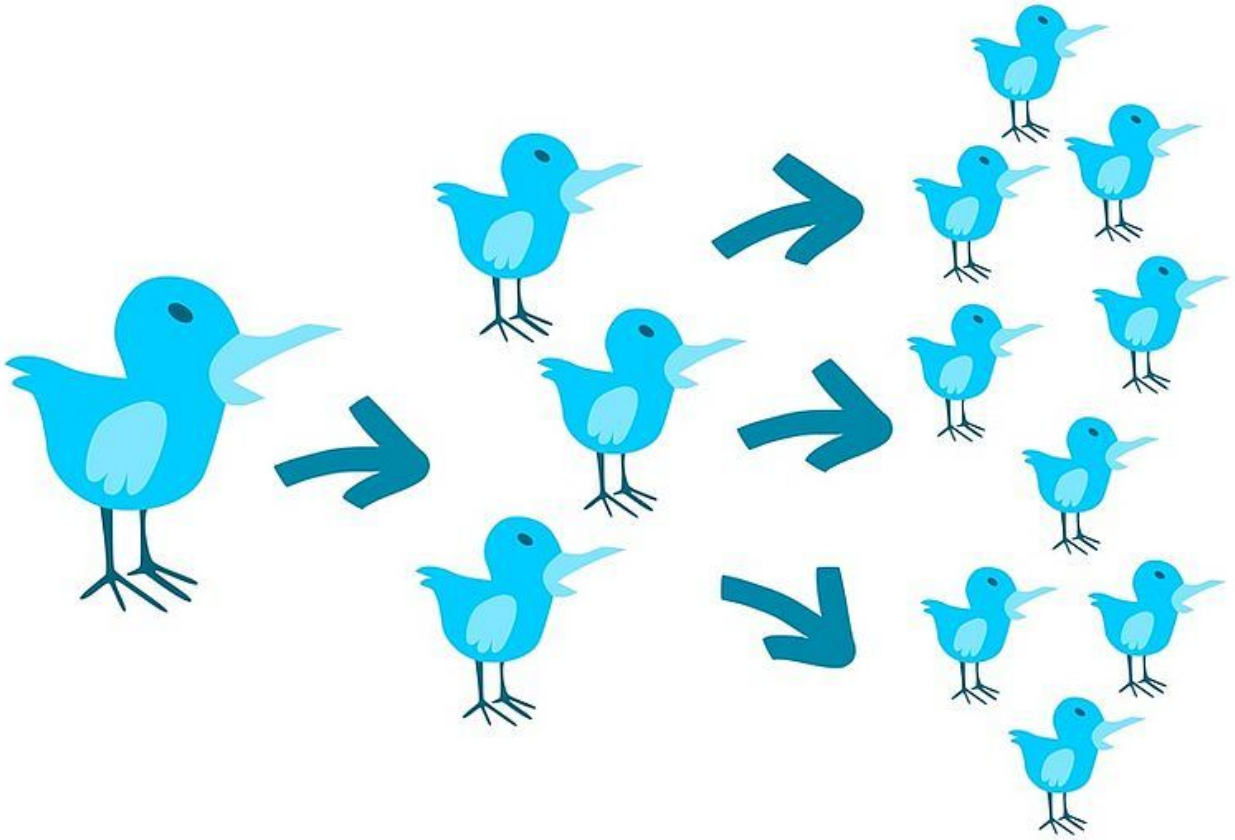


Slowing the Spread of Anti- Flu Shot Sentiment on Twitter

Marcus DeMaster, Shih Yu Chang, Allan Lo

ns Messages flu shot mal

James Chase @jason_prosser · Jan 18

More out of curiosity, I hear people say that the **flu shot makes** them **sick**. My son & I both get the shot. We've been lucky, no flu yet.

> **Marni Hughes** @marnihughesQ13
> Health department isn't telling us yet @jason_prosser But my son and I are recovering from the flu this week and we both got the shot.

2

Yuliii @yuli_barriosx3 · Jan 18

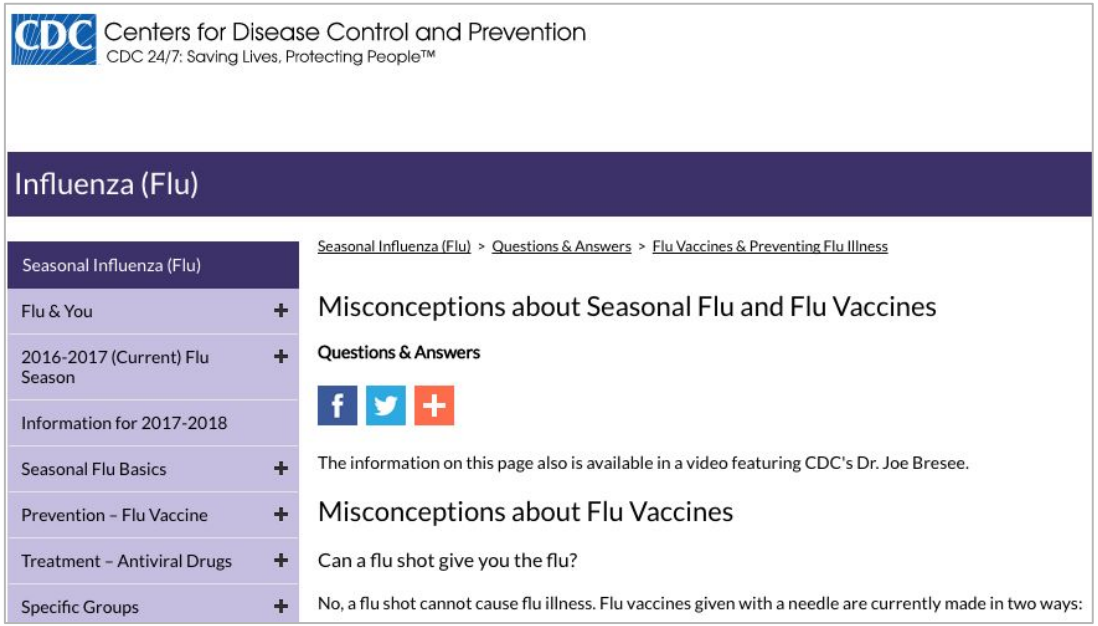
If this **flu shot makes** me **sick** Ima be sooo mad 😒

f#minor @wesleywyndam · Jan 13

I hate that getting a **flu shot makes** you **sick**

Experiment Question:

Does the CDC's flu shot page slow the spread of misconceptions on Twitter?



Designing the Administrator

- ❏ Friendly Avatar
- ❏ Healthy Living Enthusiast
- ❏ Non-bot Appearance
  - ❏ Unrelated tweets
  - ❏ Purchased followers

Issues with Initial Experiment Design

- Required Automation
  - NLP Classifier
  - Auto-replies
- Challenges
  - NLP package accuracy
  - Twitter spam blocker
  - US Flu season ending

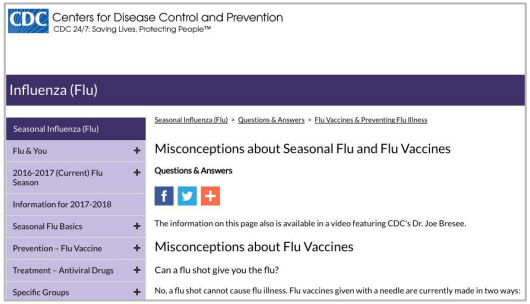
Final Experiment Design: ROXO

Search tweets with negative flu shot sentiment every 24 hours

Block Subjects on Follower Count

Randomize

Treatment



Placebo

Observe Retweet & Favorite Baselines

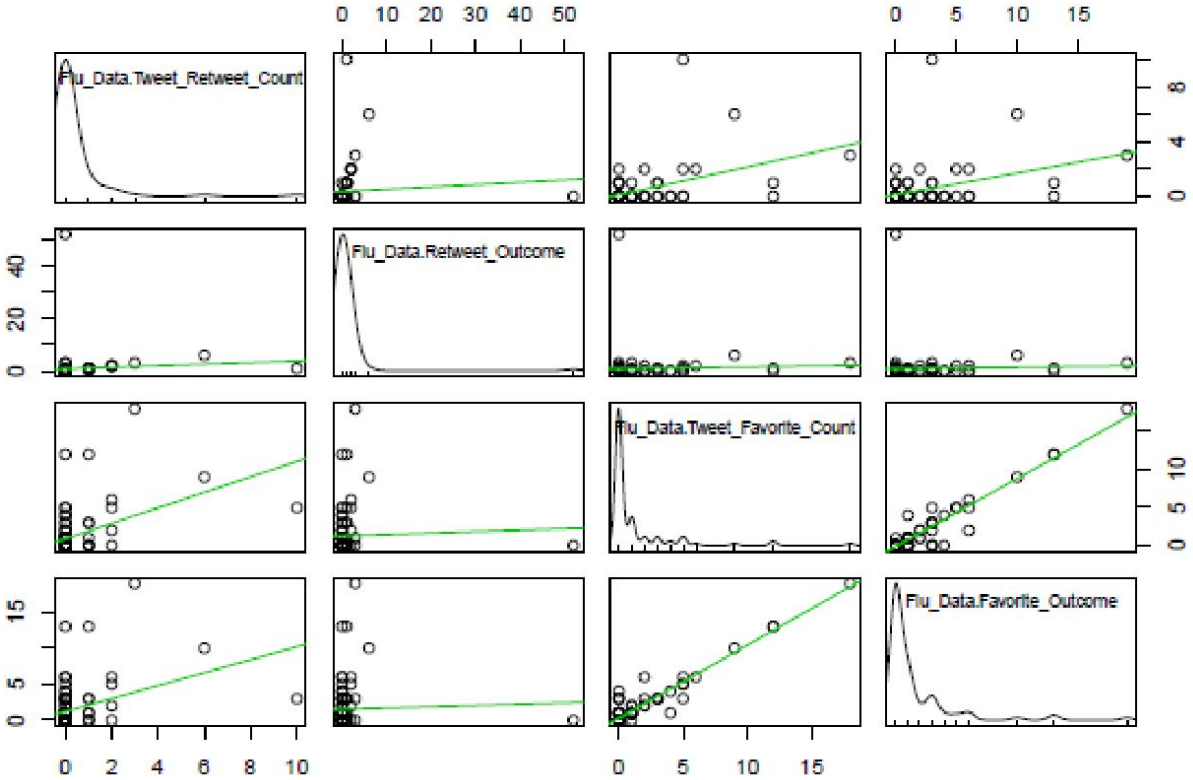
8-day ***Experiment*** Period

Observe Retweet & Favorite Outcomes

Collected Data

User_ID	User_Scre	User_Statu	User_Follo	User_Frier	User_Liste	User_Timezone	Tweet_ID	Tweet_Tex	Tweet_Retweet_Count	Tweet_Favorite_Count	Tweet_Cre	Assign_Ind	Assign_Da	Reply	Retweet_Outcome	Favorite_Outcom	Sex	Tweets_
7.39E+17	OuttaTime	108	3621	3705	34		8.49E+17	Yeah, so, n	0	1	########	1	4/3/2017		0	2	0	0.37762
2.59E+09	miggiato	7982	1715	637	19		8.49E+17	After being	0	3	########	0	4/3/2017	Here is some us	0	3	1	0.13387
7.66E+17	CaptainNe	23700	1575	1502	17	Midway Island	8.49E+17	"Safe	1	0	########	1	4/3/2017		0	0	1	0.07547
7.61E+08	Brian Halte	2337	838	1677	11	Pacific Time (US & Canada)	8.49E+17	You Don't	0	0	########	0	4/3/2017	Here is some us	0	0	0	0.15164
4.22E+08	carykayler	34285	779	323	3	Central Time (US & Canada)	8.49E+17	Some are f	0	0	########	0	4/3/2017	Here is some us	0	0	0	0.03279

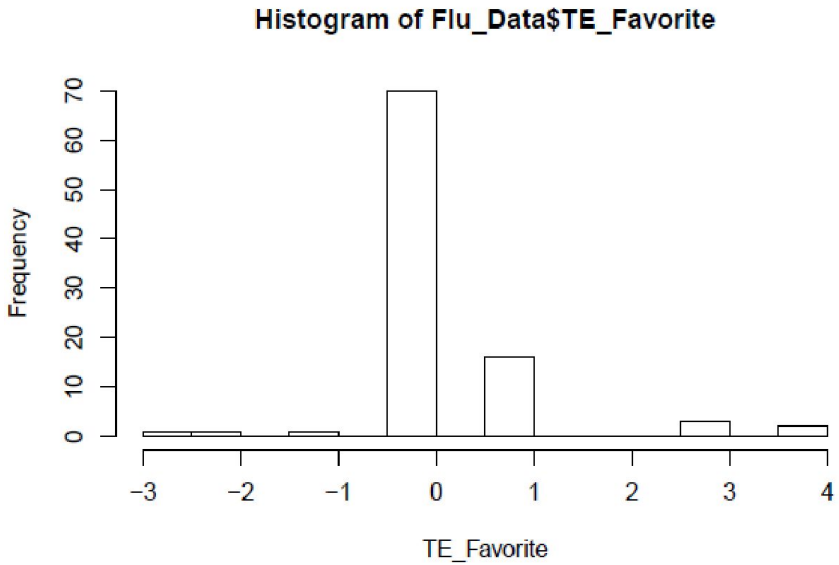
EDA (I)


Flu_Data.Tweet_Retweet_Count
Flu_Data.Retweet_Outcome
Flu_Data.Tweet_Favorite_Count
Flu_Data.Favorite_Outcome
0 10 20 30 40 50
0 5 10 15
0 20 40
0 4 8
0 5 10
0 5 15
0 2 4 6 8 10
0 5 10 15

EDA (II)

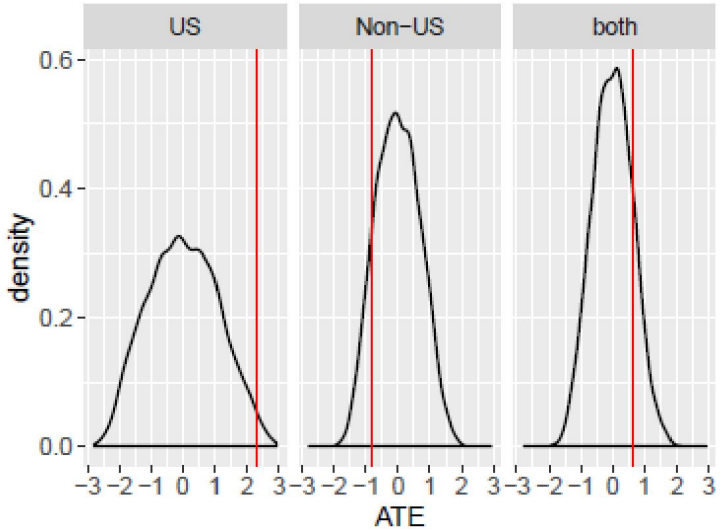
- Collected variables with numerical values have two characteristics:
  - right-skewed
  - outliers are far away from majority of data.
- Among variables User-Statuses-Count, User-Followers-Count, User-Friends-
- Count, User-Listed-Count, User-Followers-Count has most significant linear relationship with respect to Favorite-Outcome variable.
- By comparing data before and after experiments, Favorite-Outcome variable has much better trend with Tweet-Favorite-Count variable compared to Retweet-outcome and Tweet-Retweet-Count variable.

General ATE



- The individual treatment effect for each subject by favorite number
- For population treated with links
  - ATE = 0.53 (Retweet)
  - ATE = 0.29 (Favorite)
- Random assignment ATE for CDC treated links
  - ATE = 0 (Retweet)
  - ATE = 0.47 (Favorite)
- If we apply t-test, such difference is not statistically significant at the 0.05 level for Retweet measure, same for Favorite measure.

Location Effect : U.S. v.s. Non-U.S.



- For CDC treatment:
  - ATE = 2.30 at U.S.
  - ATE = - 0.85 at Non-U.S.
- We have to note that ATE value calculated from pool 0.6716826 different from ATE value obtained from combined one 0.6576287 .
- This is a biased estimate because the probability of being assigned to the treatment group varies by block (area): in US this probability is 17 / 45 = 37 . 8%, while in Non-US the probability of being assigned to the treatment group is 26 / 49 = 53 . 1%.
- Besides, the number of Favorites is lower on average in Non-US, so the overall treatment effect calculated this way is larger than it actually is.
- Therefore, if outcomes were higher in the treatment group, it might reflect differences between US and Non-US rather than a treatment effect.

CACE Effect

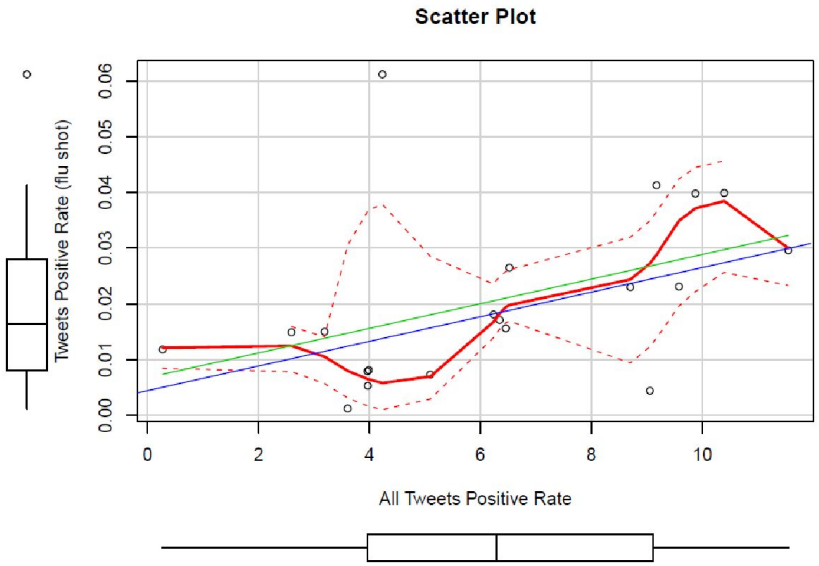
Treat. Ass.	Treated	No.#	Response
Baseline	NO	20	35%
Treatment	Yes	33	52%
Treatment	NO	10	34%
Placebo	Yes	40	33%
Placebo	No	11	36%

- Assume 80% of targets really get treated.
- For treatment and placebo groups, their compliance rate are not statistically different each other with p-value = 0.85 with t-test.
- The difference in the Never-Takers response rate between the treatment and placebo groups is not statistically significant since p = 0.77.
- CACE
  - 0.15 for receiving weblinks.
  - 0.19 among compilers.

Retweet v.s. Favorite

- Retweet_Outcome ~ Assign_Ind + Tweet_Retweet_Count + Tweet_favorite_Count (CDC = -1.0, p-value =0.4, negative effect with weaker statistical significance)
- Favorite_Outcome ~ Assign_Ind + Tweet_Retweet_Count + Tweet_favorite_Count (CDC = 0.3, p-value = 0.1, positive effect with stronger statistical significance)
- Favorite_Outcome ~ Assign_Ind*Sex + Tweet_Retweet_Count + Tweet_Favorite_Count
  - the estimated effect of CDC is 0.4027 with p-value 0.1.
  - CDC has stronger causal effect for MALE.

Gibbs Sampling for Missing New Tweets (I)



- During our experimental procedure, however, it is not easy to collect such data due to time constraints or other natural restrictions.
- The basic idea is that we first collect other available variable, the **tweets positive rate (ALL) g**enerated by a subject from his/her ALL previous Tweets averaged by days.
- The positive rate of new Tweets with flu shot topic is determined by linear model and **tweets positive rate (ALL).**

Gibbs Sampling for Missing New Tweets (II)

- We assume we have paired data . We wish to find the posterior distributions of the coefficients \beta_0 (the intercept), \beta_1 (the gradient) and of the precision \tau, which is the reciprocal of the variance.

$$y_i \sim \mathcal{N}(\beta_0 + \beta_1 x_i, 1/\tau)$$

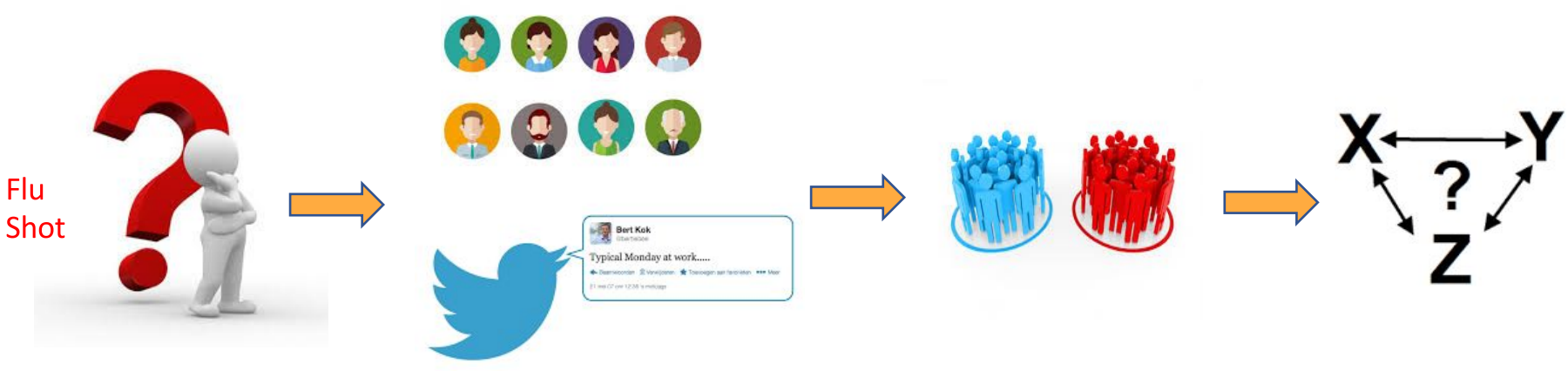
- The massive advantage of Gibbs sampling over other MCMC methods is that no tuning parameters are required!

1. Pick some initial $\theta_2^{(i)}$.
2. Sample $\theta_1^{(i+1)} \sim p(\theta_1 | \theta_2^{(i)}, x)$
3. Sample $\theta_2^{(i+1)} \sim p(\theta_2 | \theta_1^{(i+1)}, x)$

The estimated effect of CDC is -1.0 with p-value 0.38.

Future Improvements

- ❑ AMT-labeled NLP training set
- ❑ Reply-bot circumventing Twitter spam rules
- ❑ Test different value engagement strategies
  - ❑ Authority
  - ❑ Objectivity
  - ❑ Empathy
- ❑ Other public health outreach topics?
  - ❑ Childhood Immunizations
  - ❑ Antibiotics Use
  - ❑ Epidemics (Zika,Ebola, HIV, Malaria)



User	Tweets with Flu Shot	Control/Treatment	Read CDC/ No Read CDC	Increased Retweets number (outcome)	New tweets about flu shot after treatment (outcome)
api.search screen_name, user.name	api.search, tweets can do sentimental analysis	Group assignment for collected users.	Compliance or Non Compliance indicator	api.get_status	api.user_timeline