Event-Driven Analysis of Crowd Dynamics in the *Black Lives Matter* Online Social Movement

Hao Peng
haopeng@umich.edu
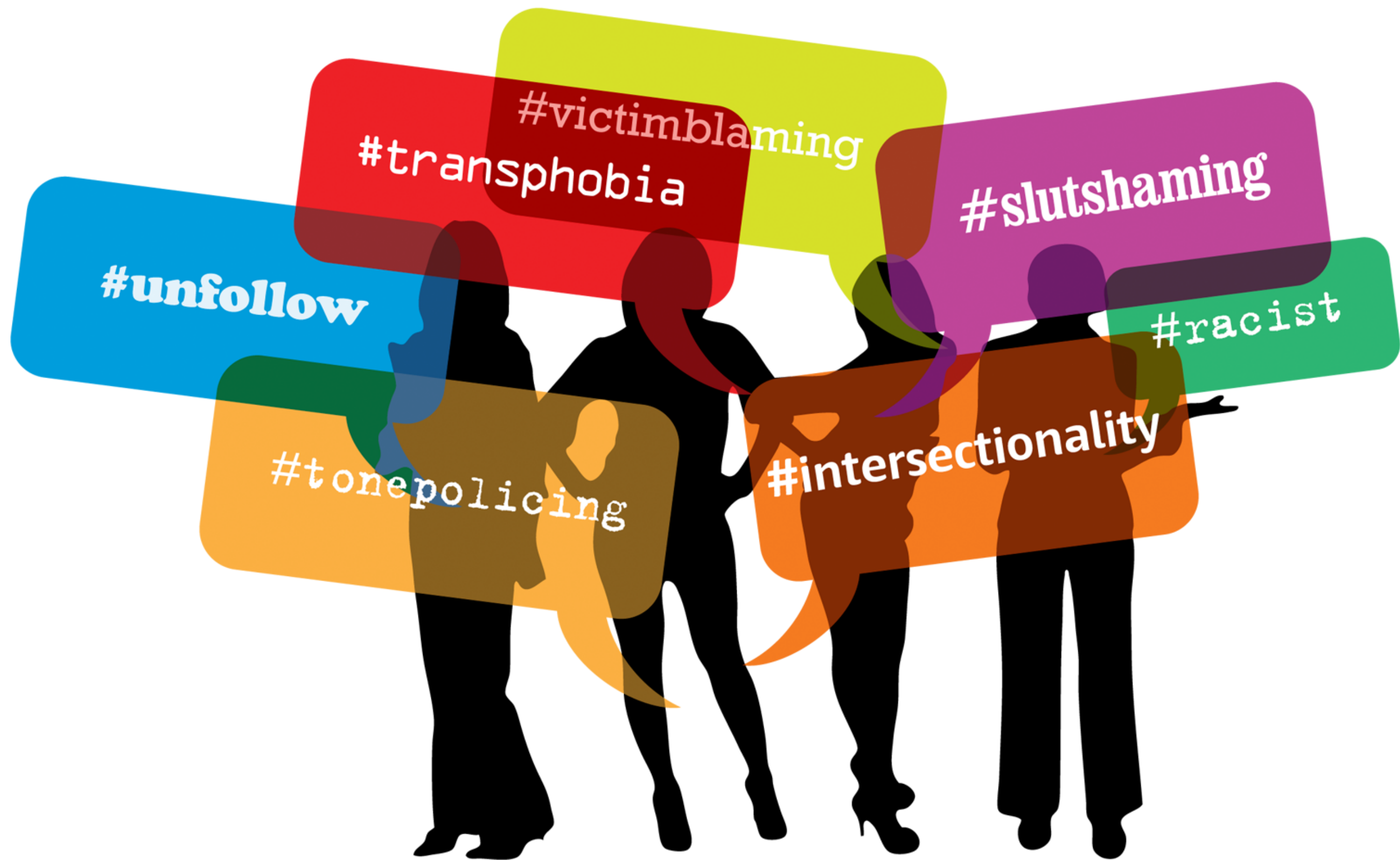
School of Information
University of Michigan

Protest for Catalonia's independence from Spain

[Bilbao, Spain, 09/16/2017]

Online Social Movements

- Online and offline scenarios are often **intertwined.**
- Online participation can be **triggered** by offline events.
- Offline protests can be **organized** in online platforms.

Related Work

- Arab Spring 2010; Egyptian Revolution 2011; Occupy Wall Street 2011
- Information diffusion; Recruitment process; Movement framing

  (Starbird and Palen, 2012; González-Bailón et al., 2013; Stewart et al., 2017)

- Research show that crowd behavior are often linked to key offline events.

  (Spiro et al., 2016; De Choudhury et al., 2016; Varol et al., 2014)
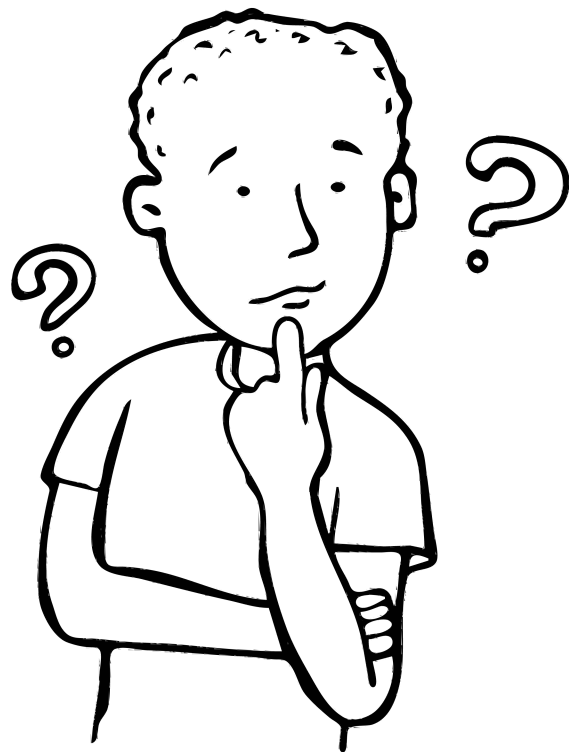
How do different types of offline events affect crowd behavior in online social movements?

Why would events matter?

Offline events $\longrightarrow$ Crowd behavior online

police violence events heightened protests

Sustained participation and Communication

Dataset

Online

- 53 *BlackLivesMatter* related hashtags;
- 36M tweets (each contains at least one #);
- 27M retweets;
- Time range: 01/02/2014 - 05/10/2015;

Offline

- 5 most prominent police violence events;
  (Eric Garner, Michael Brown, Tamir Rice, Walter Scott, Freddie Gray)
- Protest events: elephrame.com;

Event categorization of the online dataset

Defining a time window w around each event e:

1. *violence* period: $[t_e, t_e + w]$;

2. *protest* period: $[t_e - w, t_e + w]$;

3. combine adjacent periods with the same event type;

4. *violence* overrides *protest* in case of overlap;

5. let w = one week (Gallagher et al., 2018)

- activities often follow violence events (unexpected);

Event categorization of the online dataset

- the background bins: created using offline events;
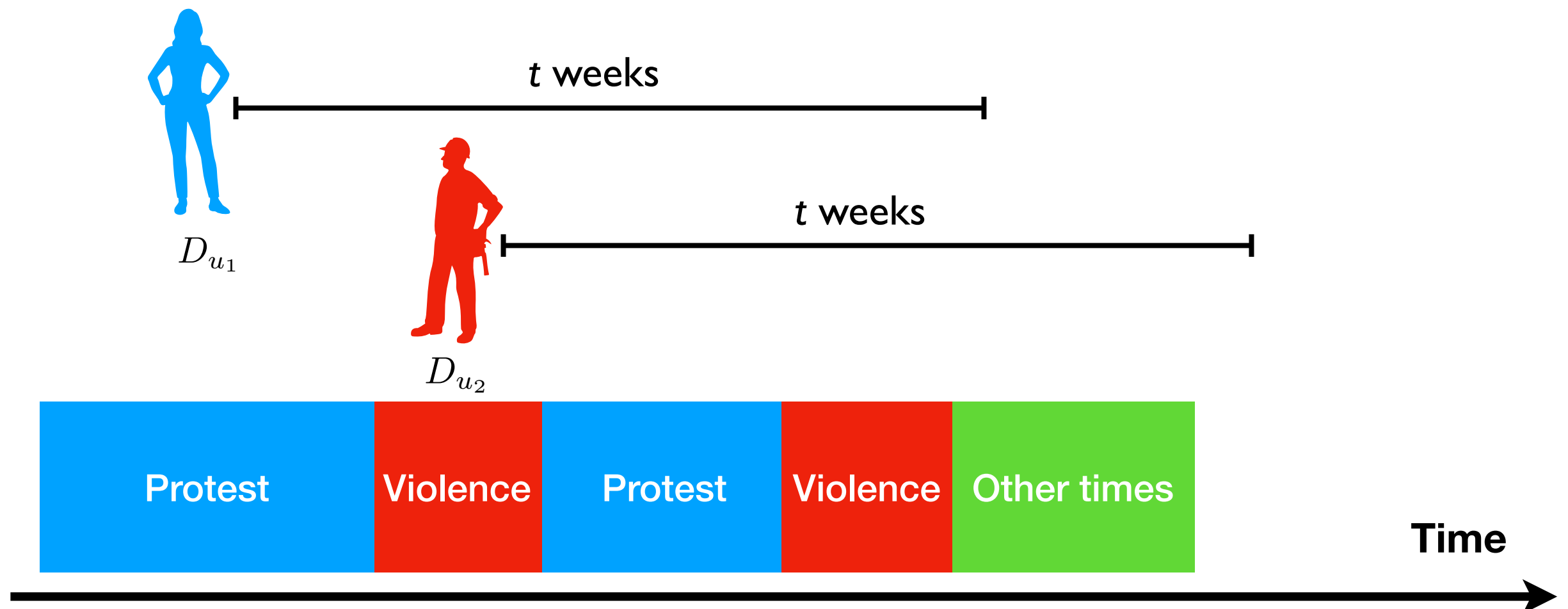- two line plots: calculated based on the online dataset;
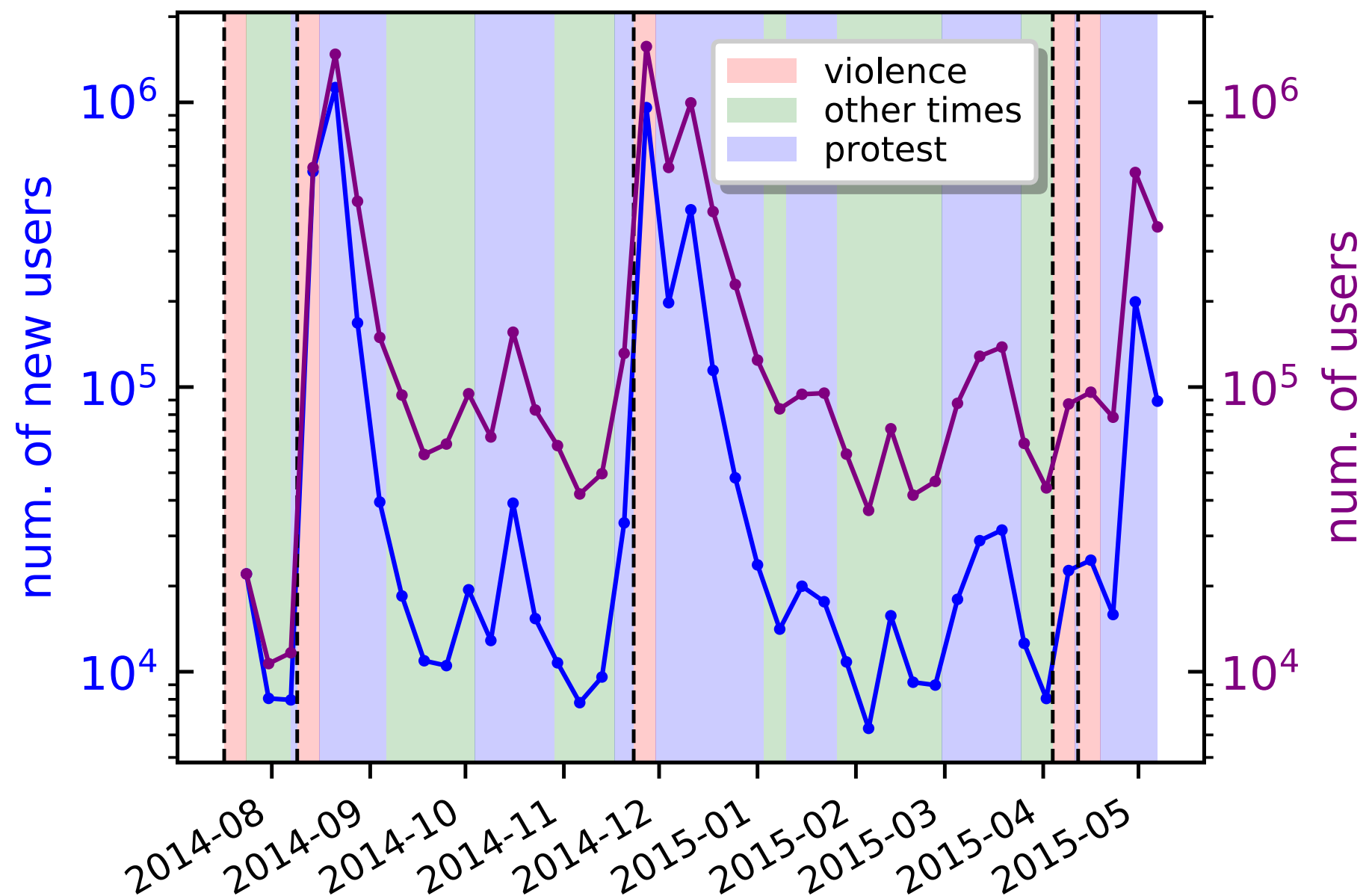
Research Questions

How arriving at the movement during a type of event is related to continued user participation in the BLM OSM?
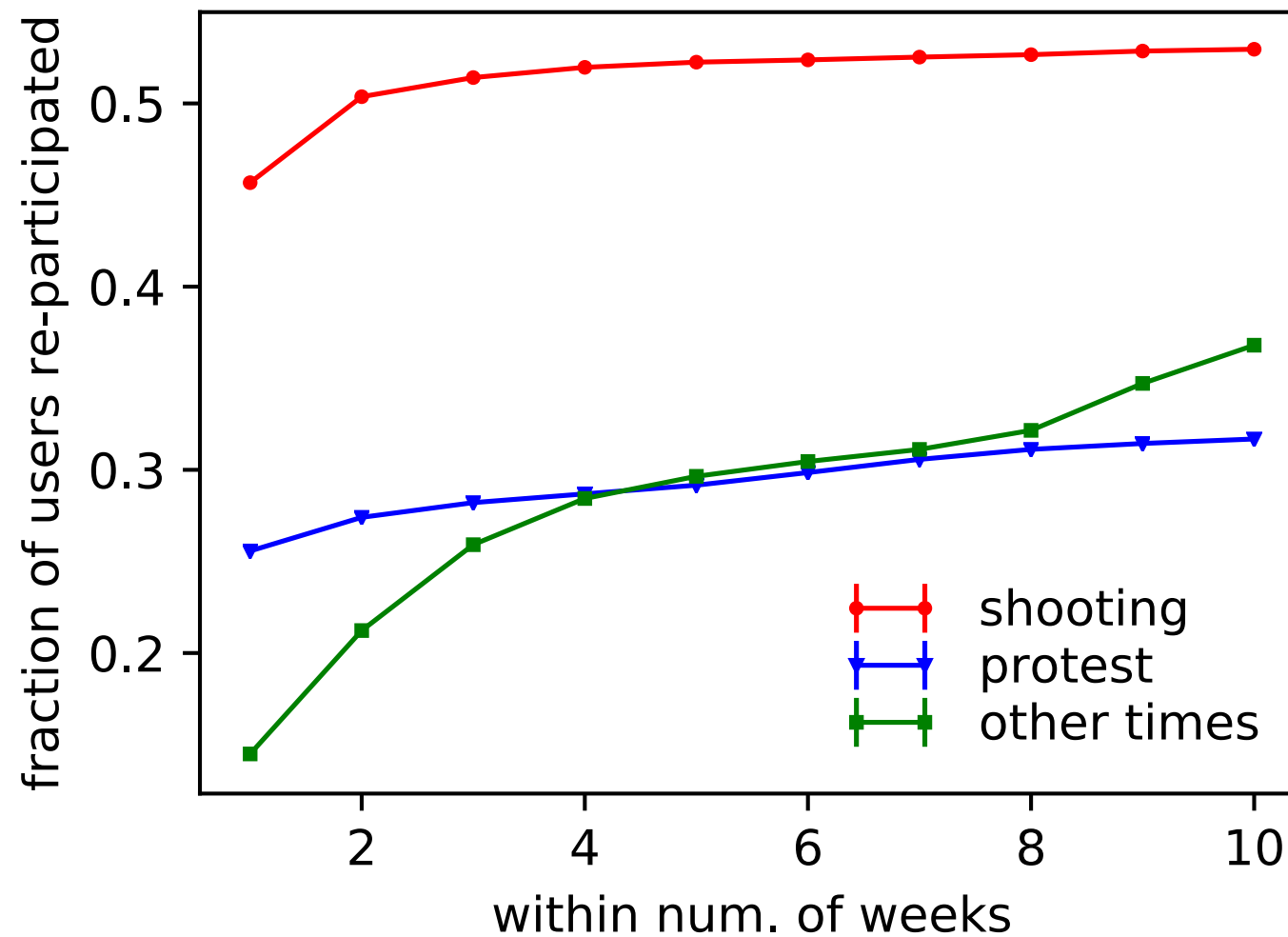
How likely are users combing back?

need controls!

Regression framework

- each user is a data point (3.8 million);
- focus on the event type of their first tweets;
- control for temporal factors and user attributes;

for t in [1, 10]:

    for each measure y:

- use y as the dependent variable;

$$logit(p(y = 1)) = \beta_1 I + \sum_{2}^{k} \beta_k x_k + \epsilon \qquad (1)$$

$$\log(y + 1) = \beta_1 I + \sum_{2}^{k} \beta_k x_k + \epsilon \qquad (2)$$

- the same set of control variables;
- independent variables: event type + controls;

Regression framework

Independent Variables

(focus) Event type →

1	the event type of D_u (categorical)
2	the week number of D_u
3	the fraction of days labeled as *shooting* in t weeks starting from D_u
4	the fraction of days labeled as *protest* in t weeks starting from D_u
5	the tweet count of u at D_u
6	the friend count of u at D_u
7	the follower count of u at D_u
8	the account age (in days) of u at D_u

Note: Variables $5 - 8$ are user meta data.

Temporal factors { (rows 2–4)

User attributes { (rows 5–8)

Focus is not the prediction accuracy, but the coefficients of Events;
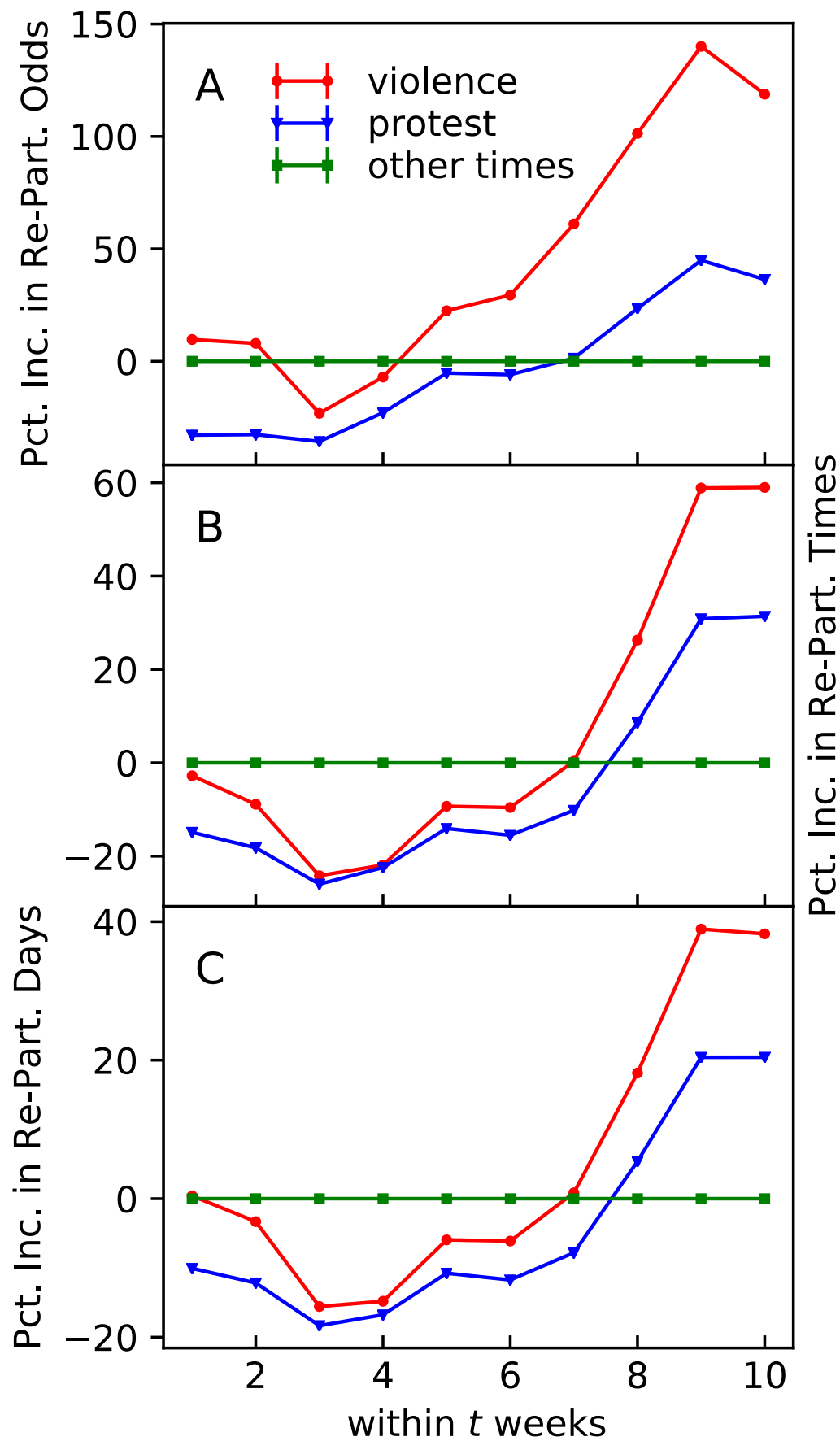
- Use other times as the baseline group;
- Report the perc. increase for violence and protest;

← Re-participate or not?

← Re-participation frequency

← Re-participation persistence

↑ the tracking window

Research Questions

How starting a tie during a type of event is related to its future interactions?
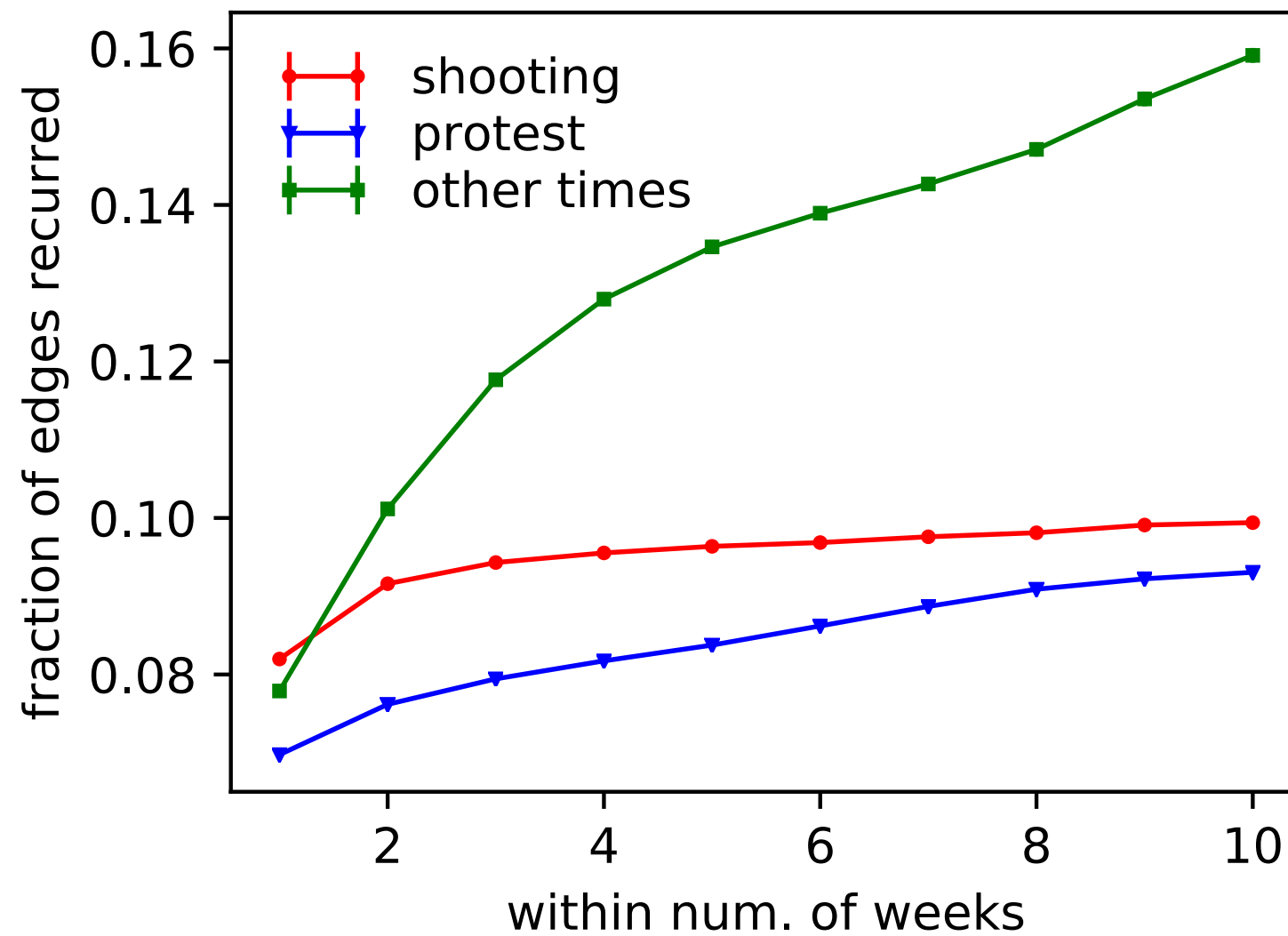
Social Interactions

Measure tie strength along three dimensions:
1. retweet again or not? (binary);
2. num. of retweets (frequency);
3. num. of days retweeted (persistence);

t weeks

t weeks

$D_{e_{i,j}}$

| Protest | Violence | Protest | Violence | Other times |
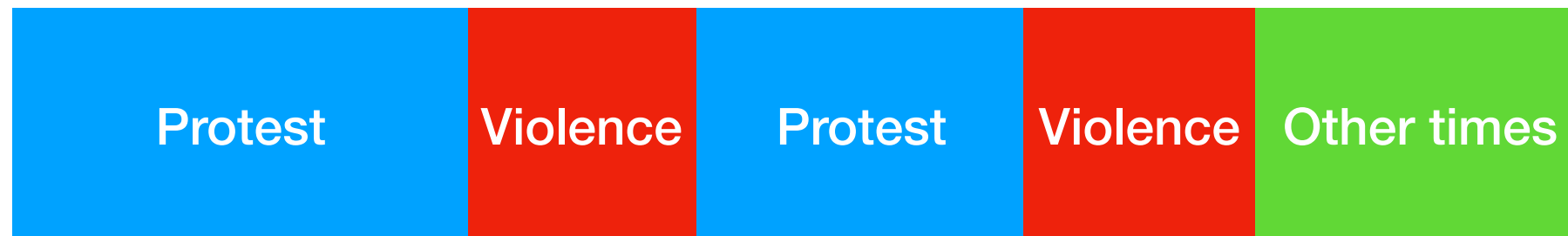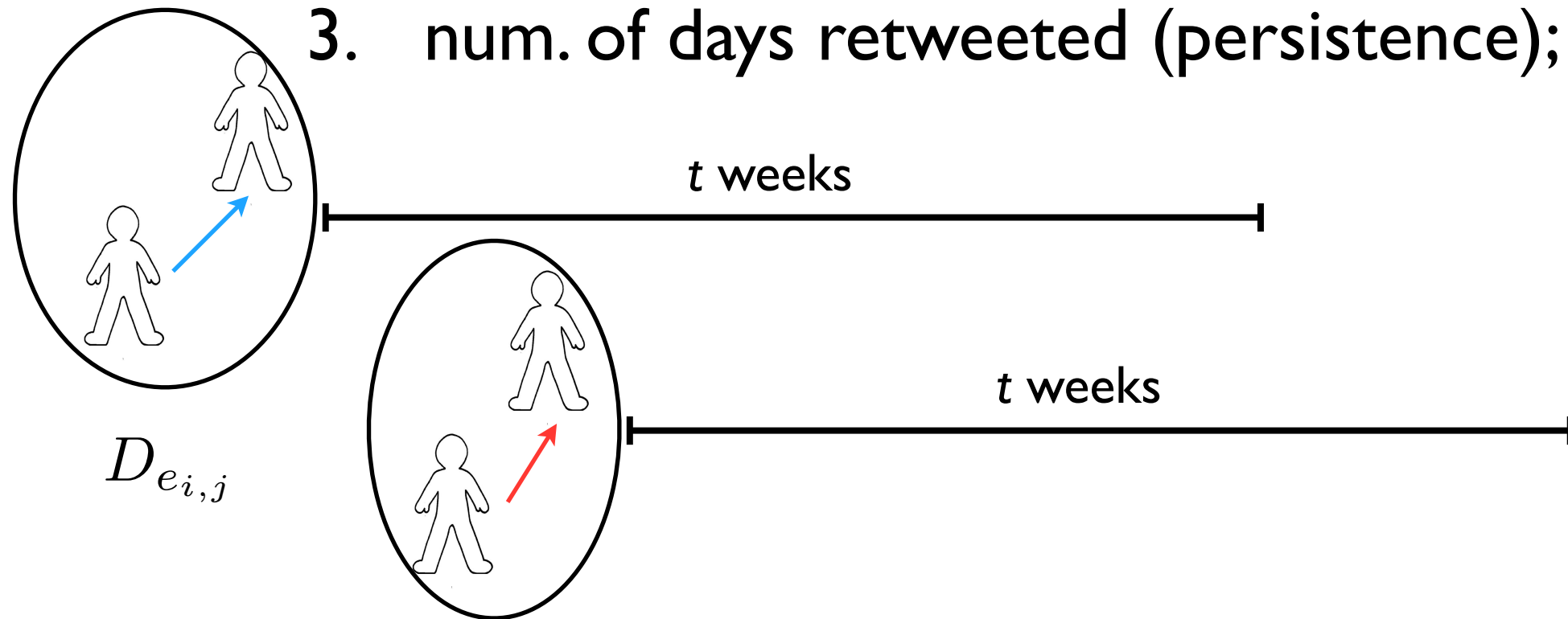
Time

How likely do interactions recur?

need controls!

Regression framework

- each pair of users is a data point (17 million);
- focus on the event type of their first retweets;
- control for temporal factors and user attributes;

```
for t in [1, 10]:
    for each measure y:
```

- use y as the dependent variable;
- the same set of control variables;
- independent variables: event type + controls;

Regression framework

(focus) Event type ➜

Temporal factors

User activities

prior interactions

1	the event type of $D_{e_{i,j}}$
2	the event type of D_{u_i}
3	the event type of D_{u_j}
4	the week number of $D_{e_{i,j}}$
5	the # protest tweets of u_i in $[D_0, D_{e_{i,j}}]$
6	the fraction of (5) in *shooting*
7	the fraction of (5) in *protest*
8	the # protest tweets of u_j in $[D_0, D_{e_{i,j}}]$
9	the fraction of (8) in *shooting*
10	the fraction of (8) in *protest*
11	the times u_i has retweeted others in $[D_0, D_{e_{i,j}}]$
12	the times u_i has been retweeted by others in $[D_0, D_{e_{i,j}}]$
13	the times u_j has retweeted others in $[D_0, D_{e_{i,j}}]$
14	the times u_j has been retweeted by others in $[D_0, D_{e_{i,j}}]$
15	the times u_j has retweeted u_i in $[D_0, D_{e_{i,j}}]$
16	the # users u_i and u_j have retweeted in common in $[D_0, D_{e_{i,j}}]$
17	the # protest tweets of u_i in t weeks from $D_{e_{i,j}}$
18	the # protest tweets of u_j in t weeks from $D_{e_{i,j}}$
19	the fraction of days labeled as *shooting* in t weeks from $D_{e_{i,j}}$
20	the fraction of days labeled as *protest* in t weeks from $D_{e_{i,j}}$

Note: Variables $5 - 20$ are calculated based on our dataset.

retweet or not? →

num. of retweets →

num. of days retweeted →

- other times: baseline;
- violence: short term and long term effects;
- protest: ties are less likely to be sustained
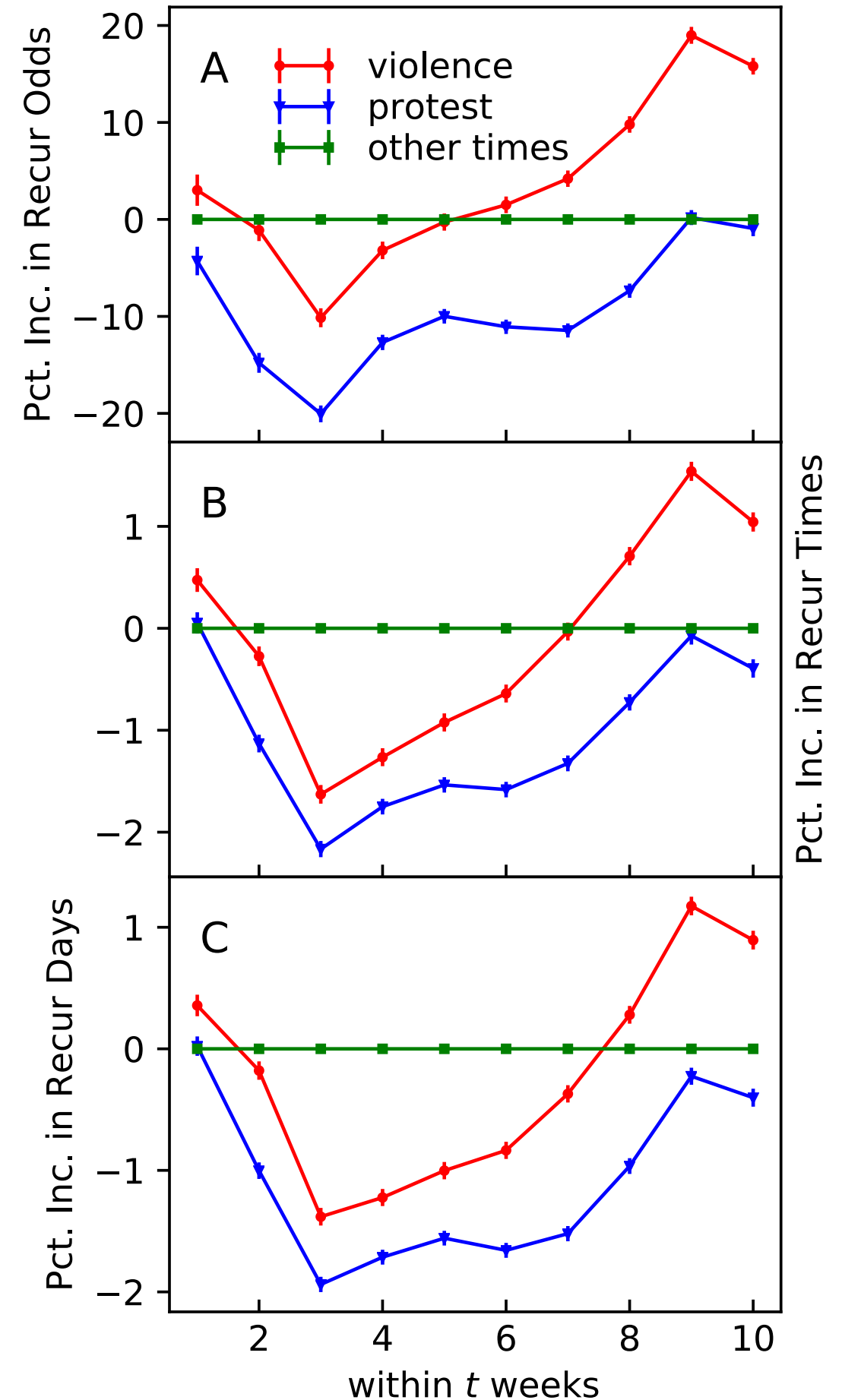
Conclusion

- police violence and protest events have both long term and short term effects on user commitment in the *BLM* OSM.

- interactions formed during violence events are more likely to be sustained than those formed during other times, with the latter expressing more engagement than the protests group.

- **Implication** for policymakers, movement organizers, and online social movement observers. e.g. encourage newcomers during other times.

Collaborators

Ceren Budak

Daniel M. Romero

Questions?