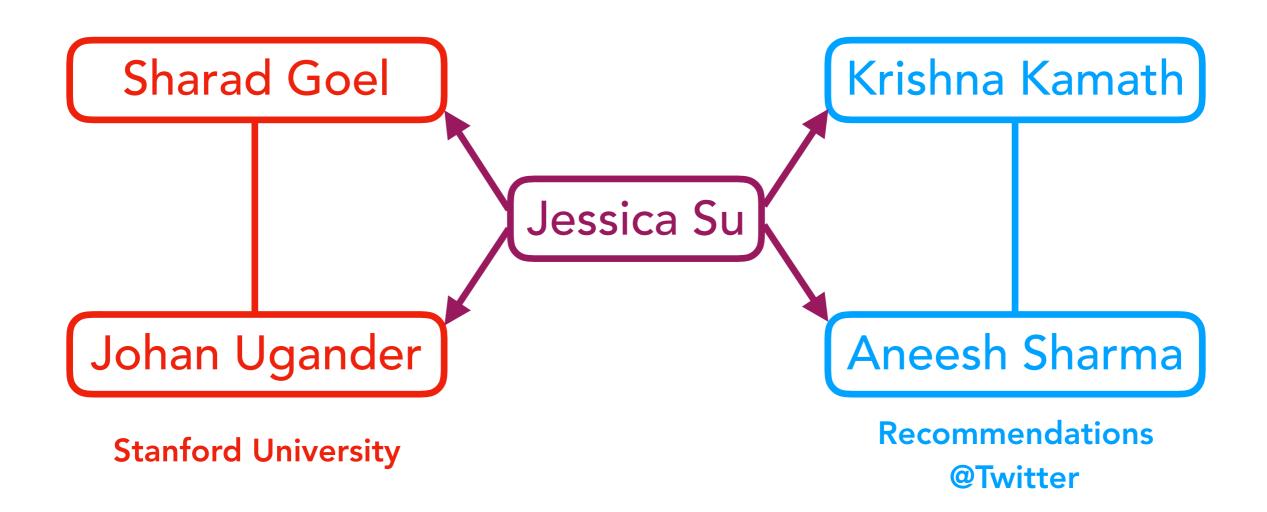
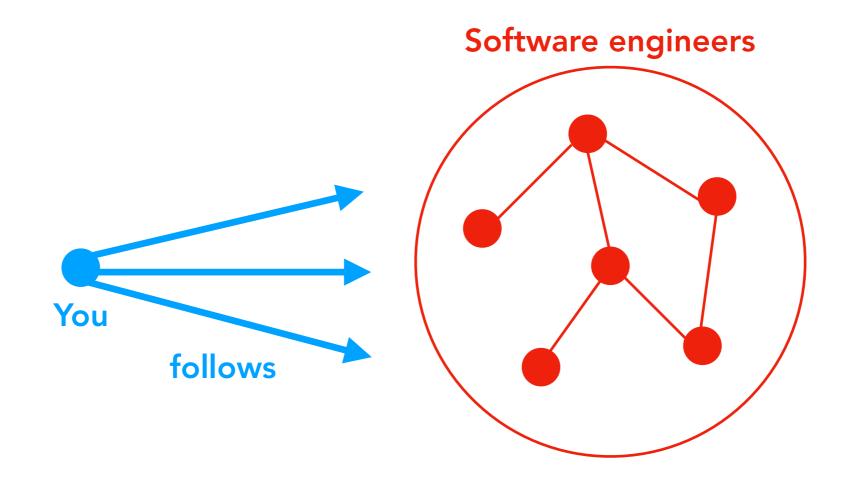
# An Experimental Study of Structural Diversity in Social Networks



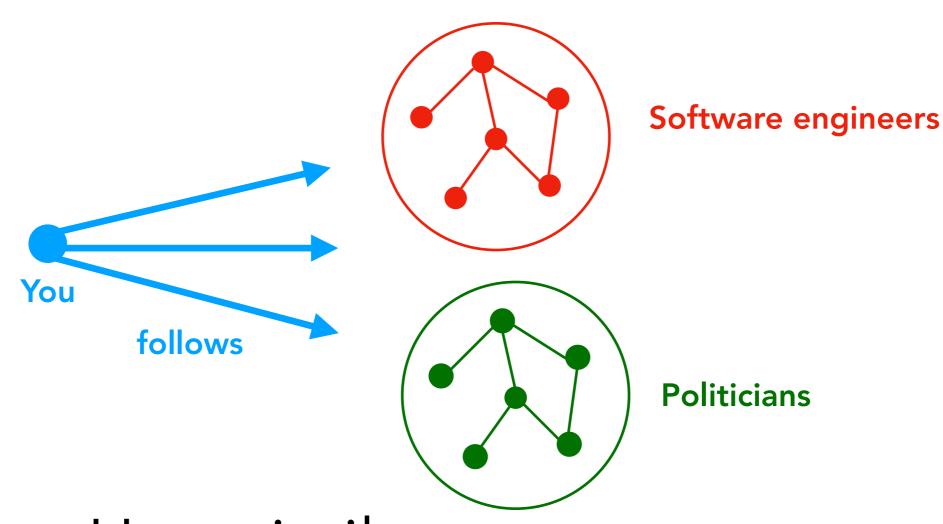
How similar are your social connections?

How similar are your social connections?

(people you follow on Twitter)



How similar are your social connections?



How similar are your social connections?

# Structural diversity is related to engagement

Studies have shown that people with more diverse social networks are more likely to be engaged on the site

(Ugander et al., "Structural Diversity in Social Contagion")

(Aral and Nicolaides, "Exercise Contagion in a Global Social Network")

(Spiliotopoulos and Oakley, "Understanding Motivations for Facebook Use: Usage Metrics, Network Structure, and Privacy")

(Weng et al., "Virality Prediction and Community Structure in Social Networks")

(Backstrom et al., "Characterizing and curating conversation threads: Expansion, focus, volume, re-entry")

# Why is structural diversity related to engagement?

Option #1: Diversity causes engagement

Maybe people like seeing many different types of content

Maybe having a wide variety of content makes it more likely that you see at least one interesting item

## Why is structural diversity related to engagement?

Option #2: Diversity is **correlated** with engagement, but doesn't **cause** it

Maybe young people just happen to have more diverse networks, and also spend more time on the Internet

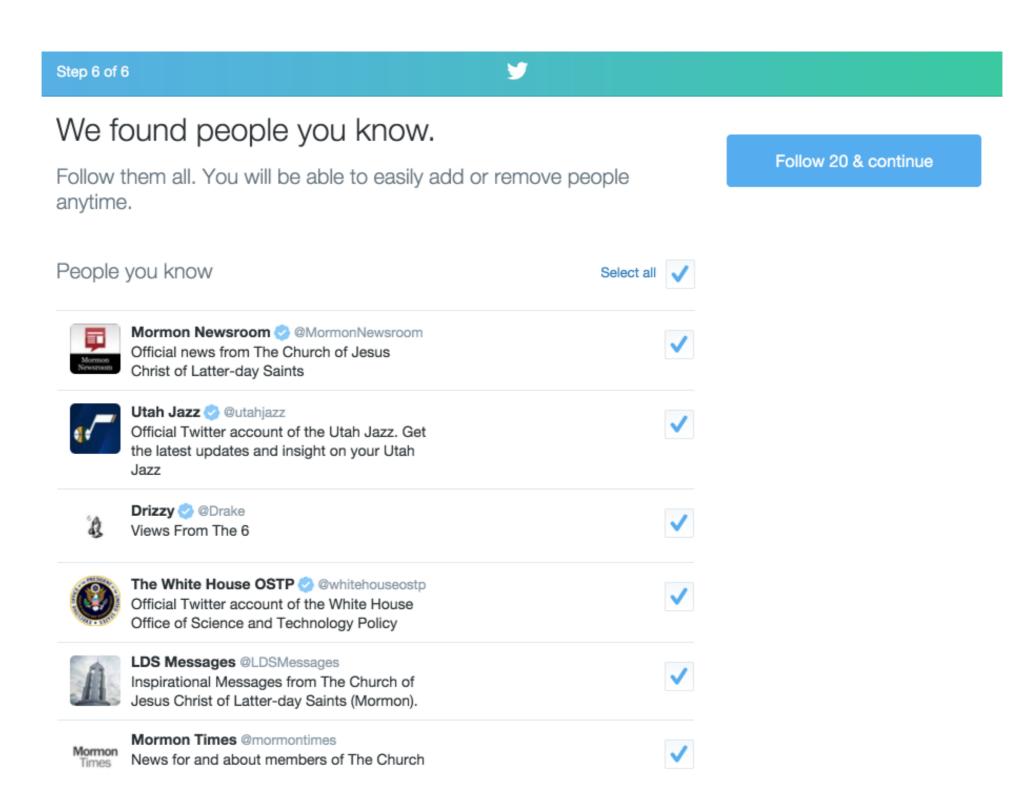
# Let's do an experiment!

If we change diversity, what happens to engagement?

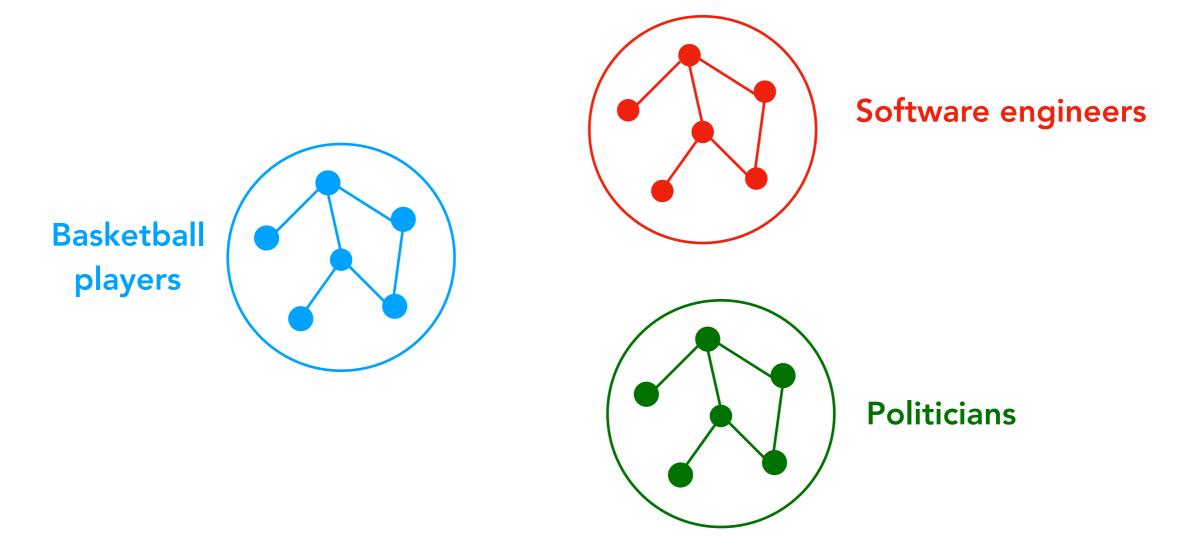
#### Experiments are hard

How do we change the structural diversity of someone's network?

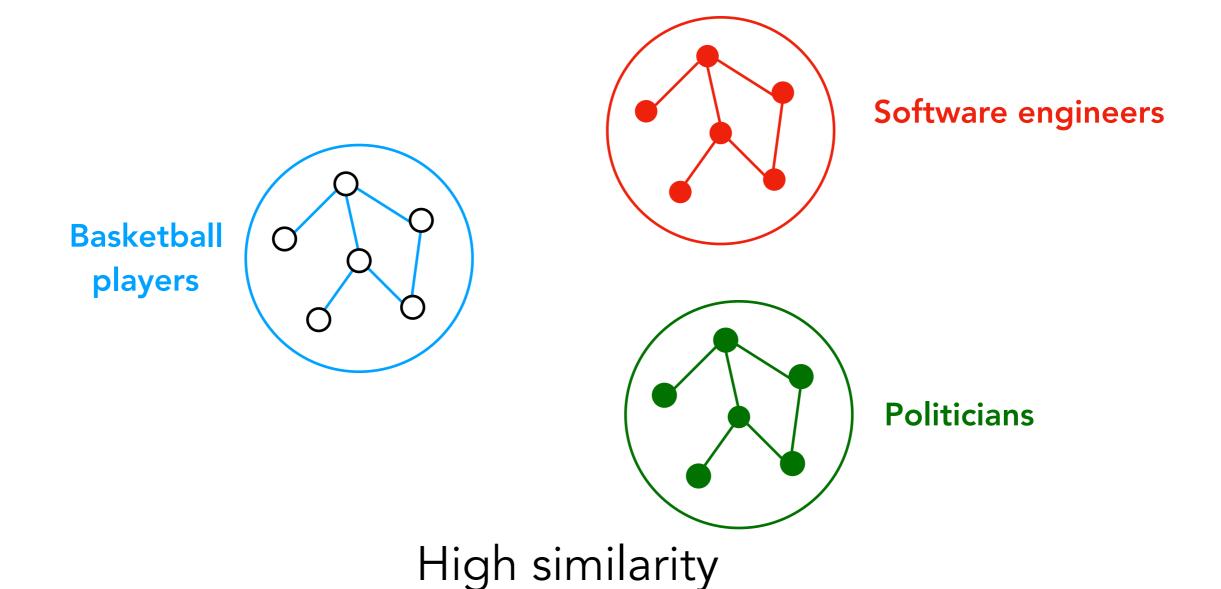
How do we randomly assign people to experimental conditions?



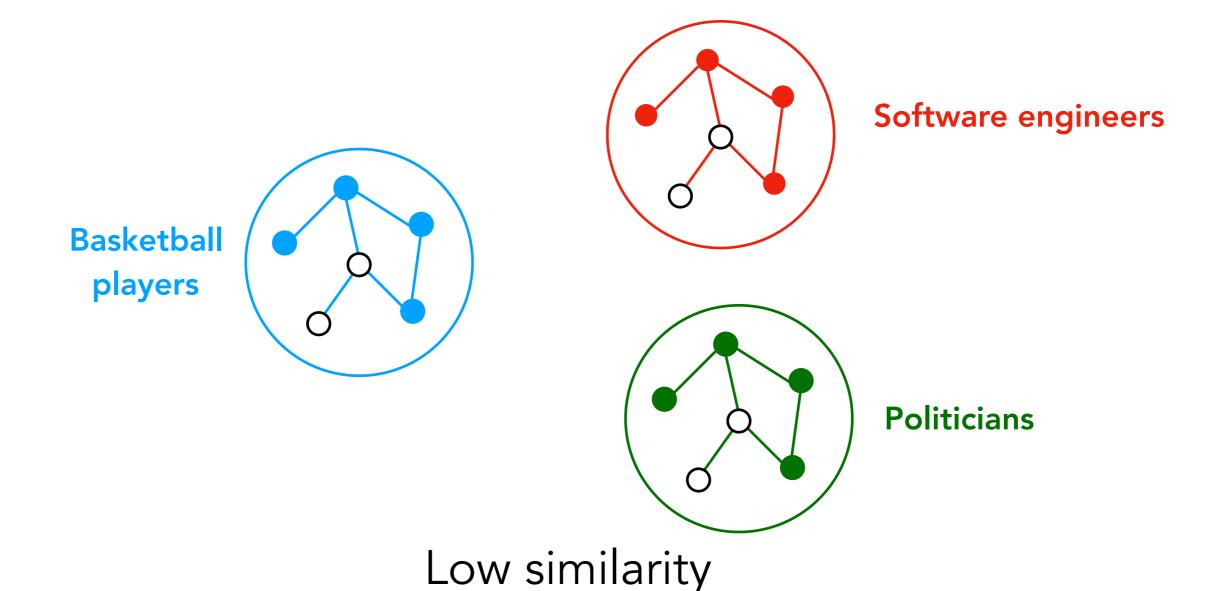
#### Alter structural diversity of Twitter new user recommendations



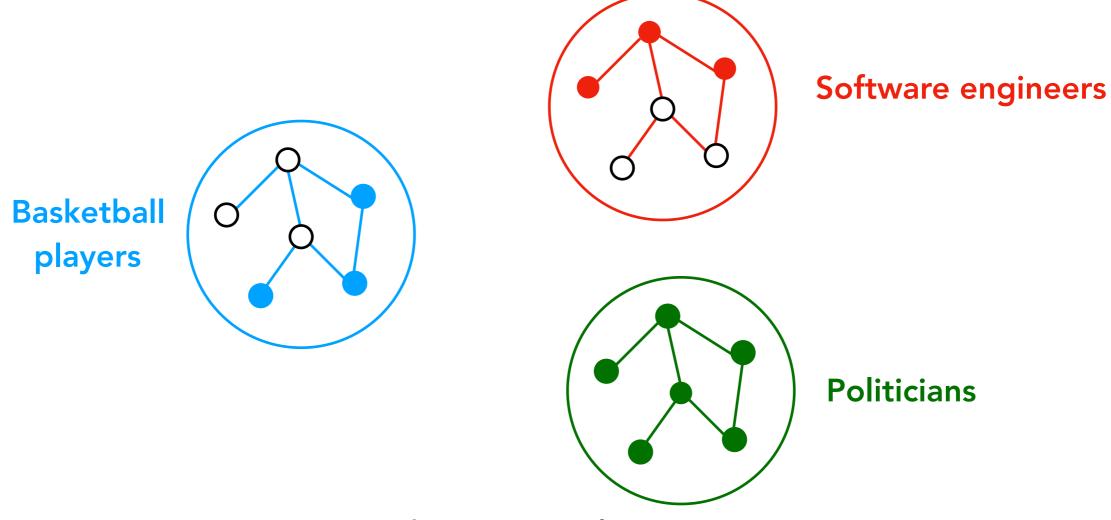
Potential recommendations



(low diversity)

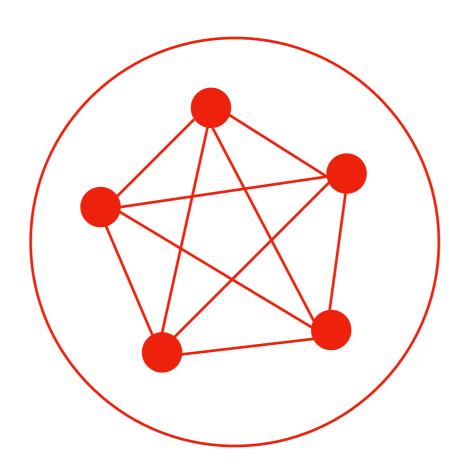


(high diversity)



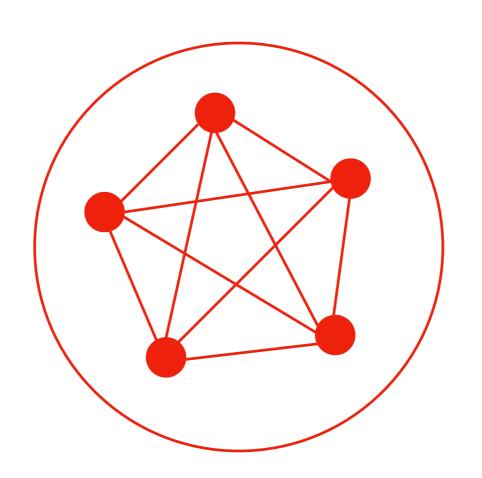
Medium similarity (medium diversity)

#### Experiment eligibility



Ineligible for the experiment (can't choose a low-similarity set)

#### Experiment eligibility

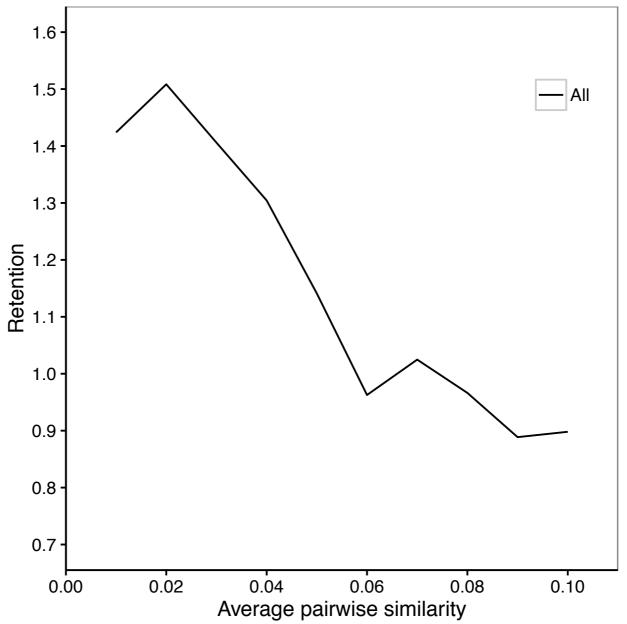


12% of users are eligible for the experiment

Ineligible for the experiment (can't choose a low-similarity set)

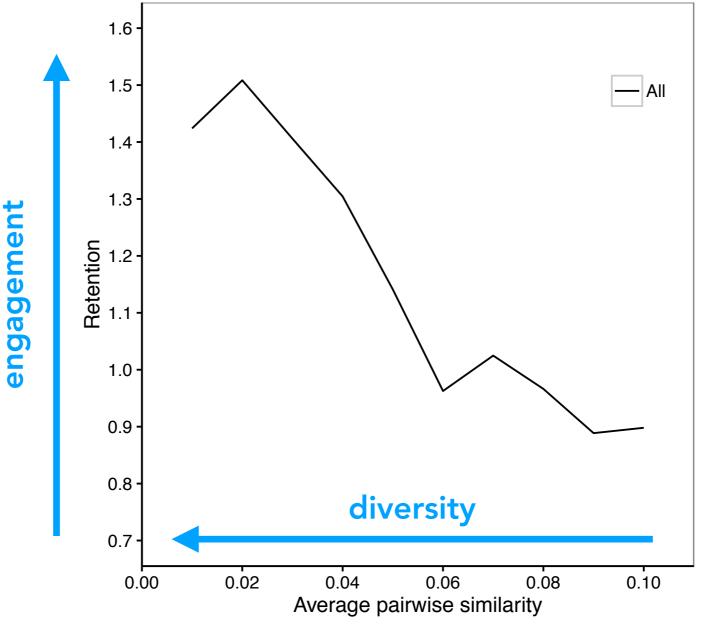
#### Results

(for people who followed exactly 20 users)



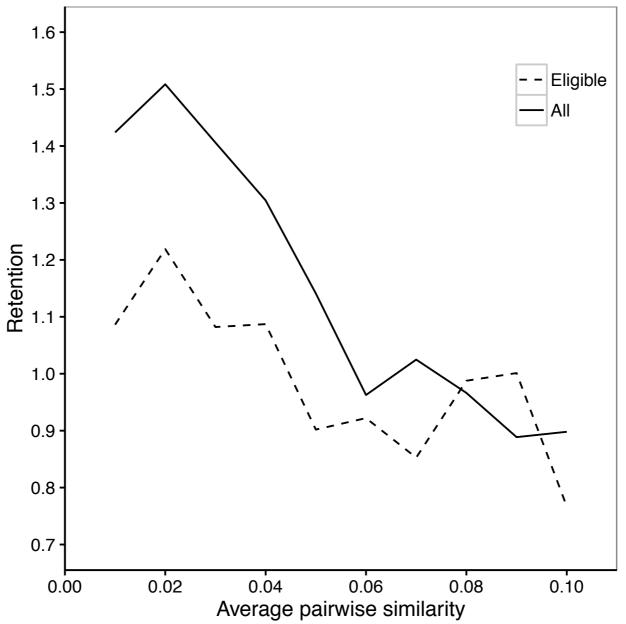
Diversity correlates with retention in the general population of Twitter users

(for people who followed exactly 20 users)



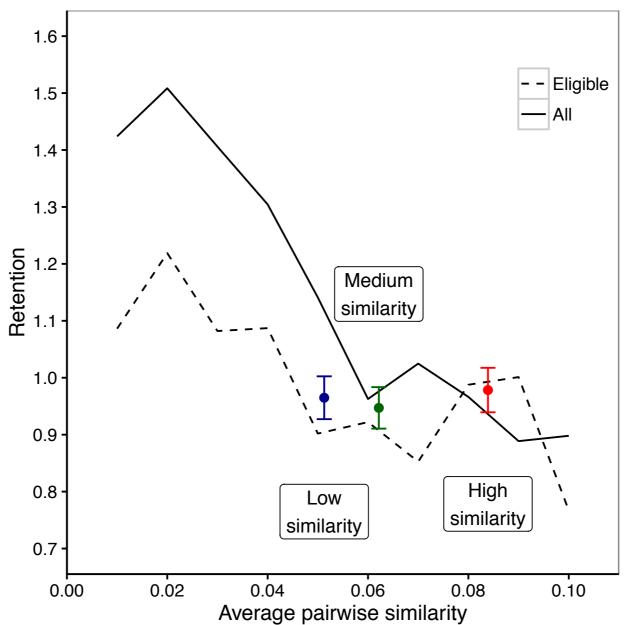
Diversity correlates with retention in the general population of Twitter users

(for people who followed exactly 20 users)



For eligible users, this correlation is attenuated (but still there)

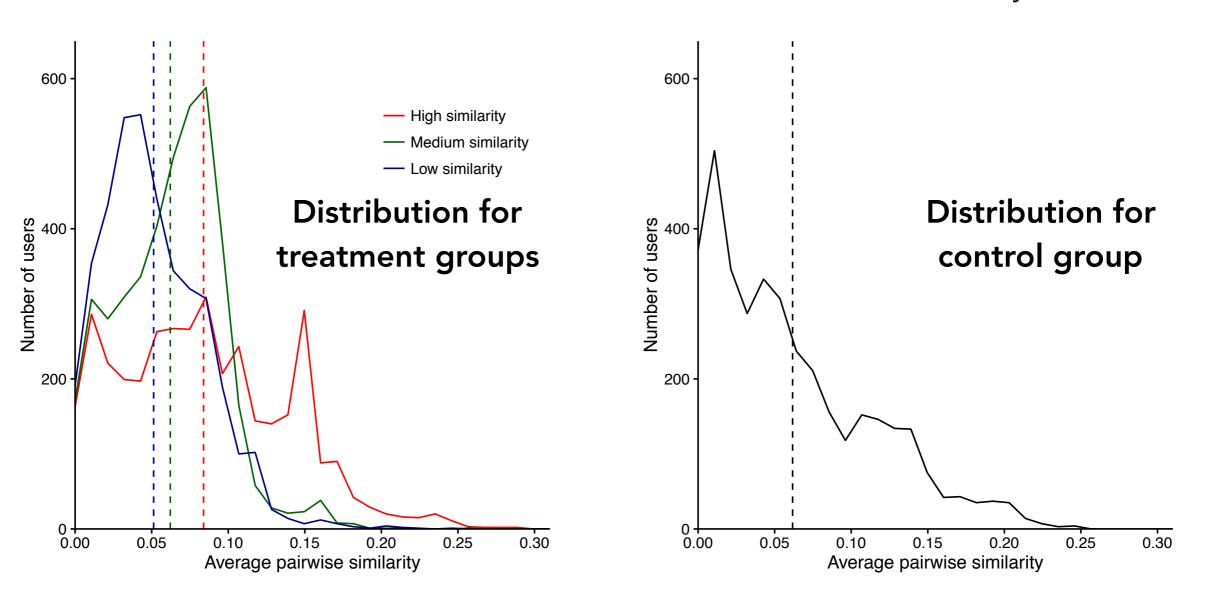
(for people who followed exactly 20 users)



We observe no causal effect of structural diversity on user retention

#### Treatment effect

(for people who followed exactly 20 users)



Different treatment groups have different levels of structural diversity

#### Conclusion

Not everyone accepts our recommendations (our treatment is only valid for people who followed exactly 20 users)

#### Conclusion

Not everyone accepts our recommendations (our treatment is only valid for people who followed exactly 20 users)

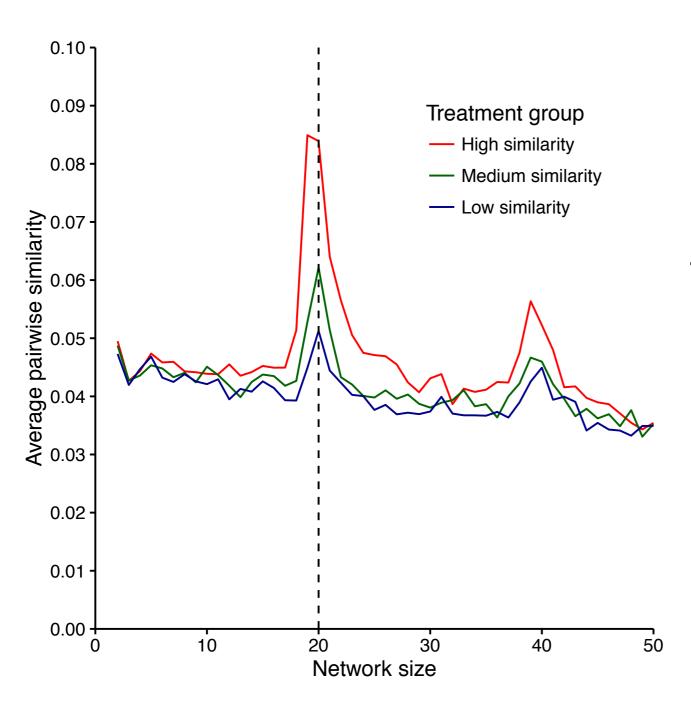
Need to make sure people in certain treatment groups aren't more likely to follow exactly 20 users

#### Conclusion

We're not just modifying structural diversity We're also showing different users...

#### Thanks for listening

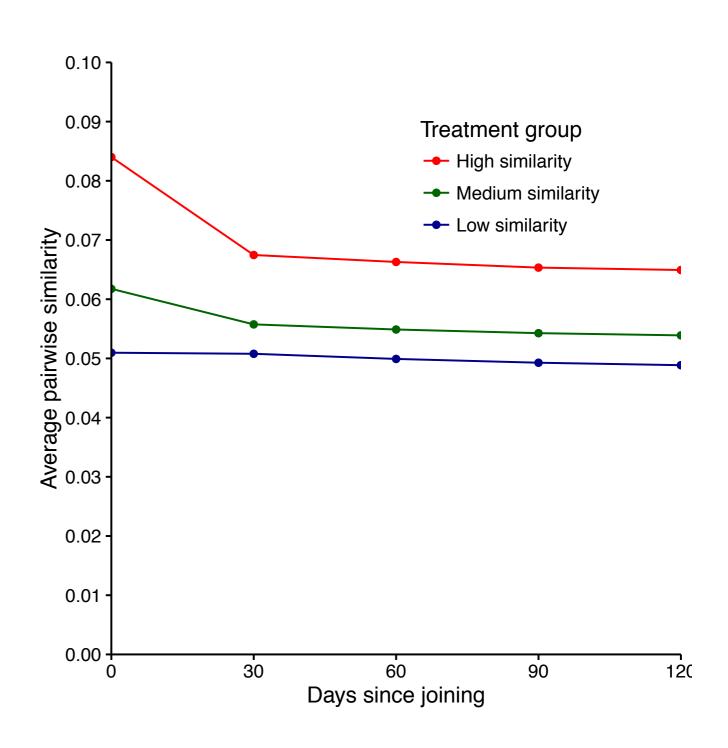
#### FAQ: Treatment validity



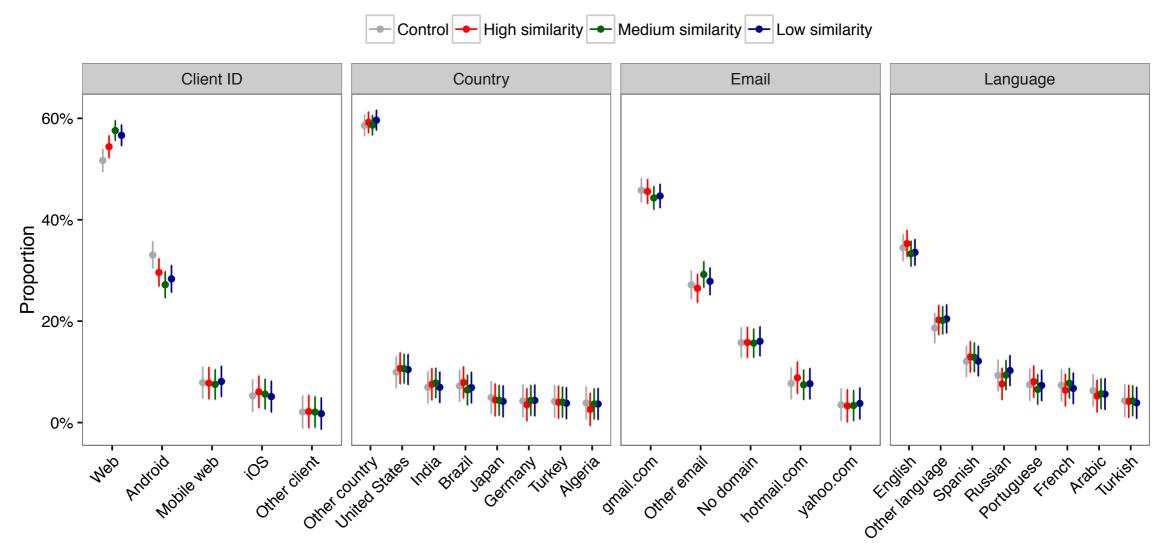
Treatment is most valid on the set of users who followed exactly 20 people (users who clicked "Follow All")

#### FAQ: Treatment validity

For "follow all" users, the similarity gap persists over time

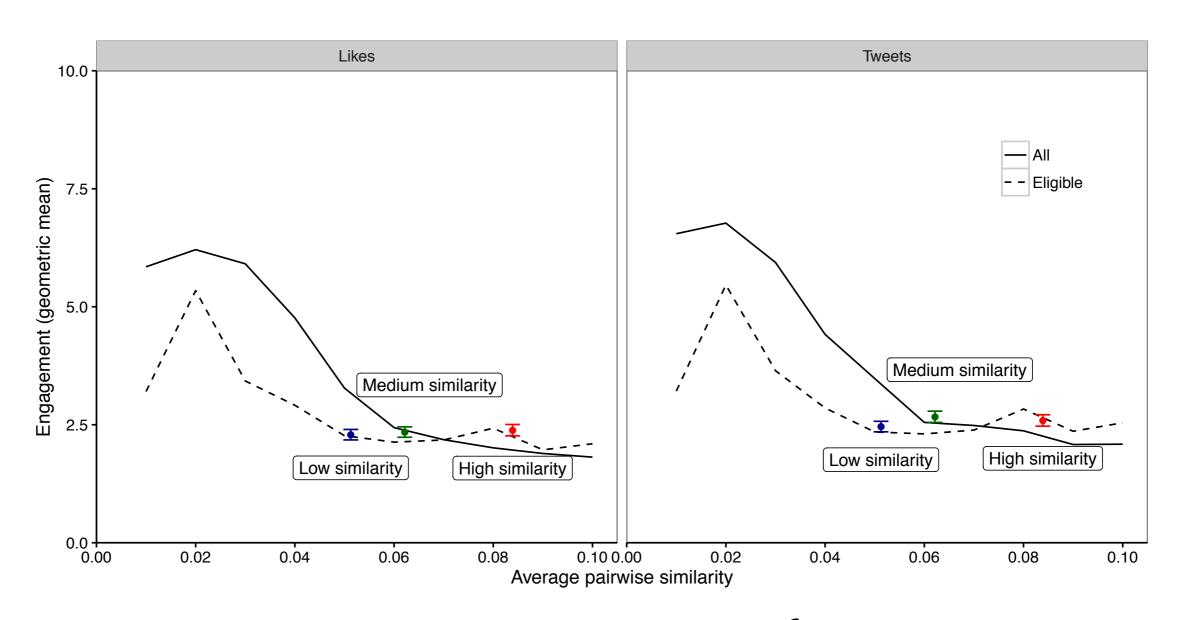


#### FAQ: Post-treatment selection



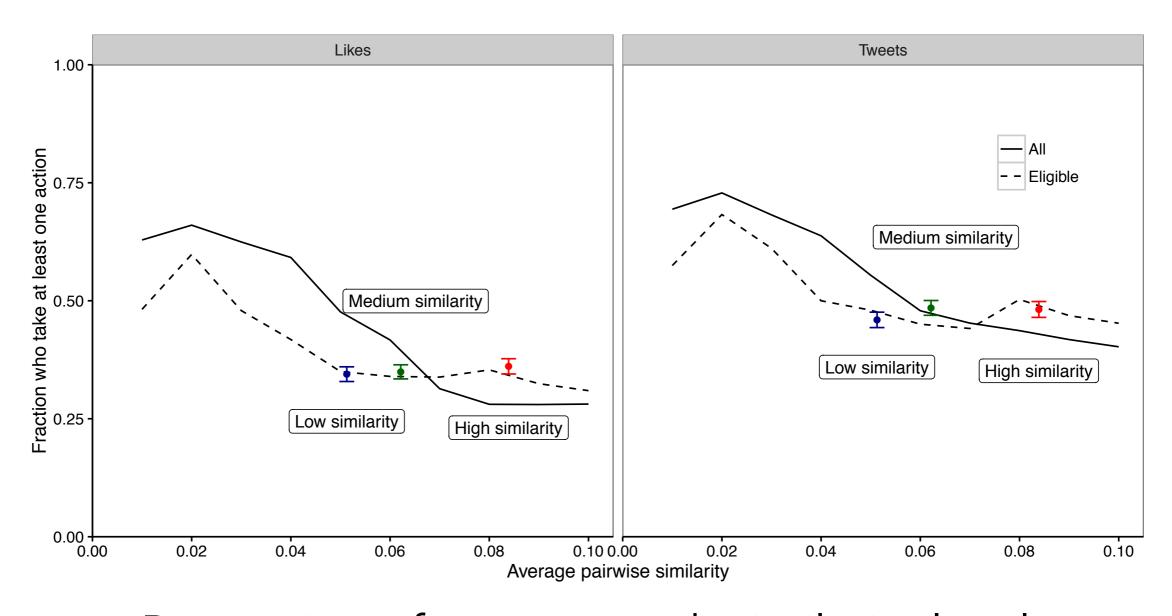
The treatment groups have roughly equal proportions of each "type of user," even after we subset on users who follow exactly 20 people

### FAQ: Does this result hold for engagement measures other than user retention?



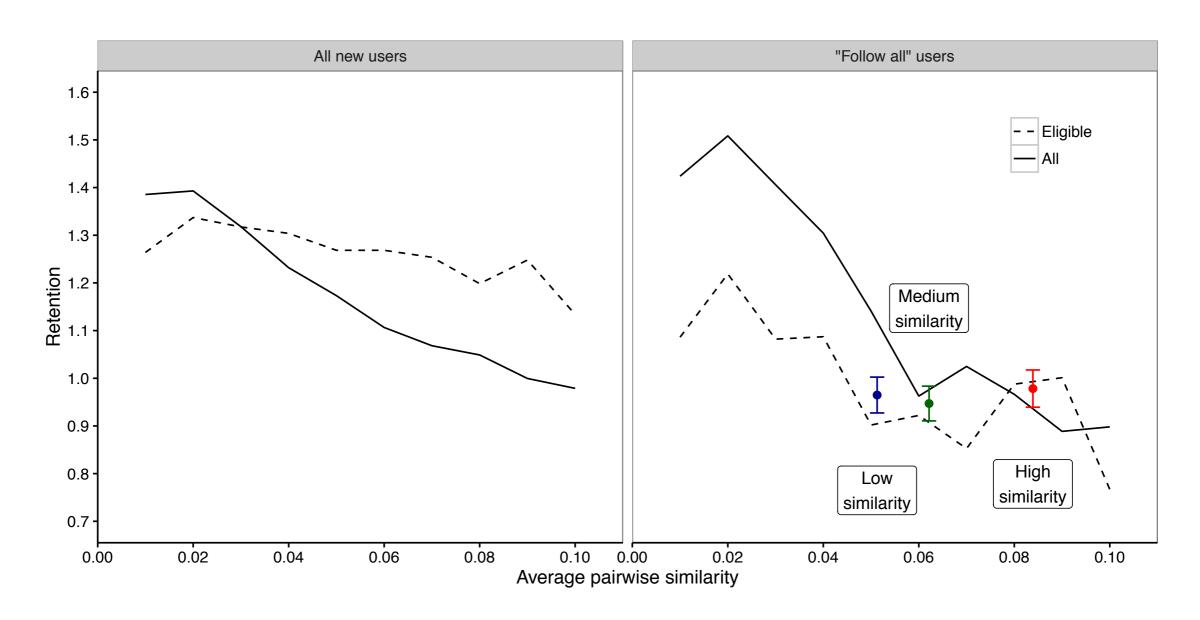
Geometric mean of 1 + (number of likes/tweets sent by the user)

### FAQ: Does this result hold for engagement measures other than user retention?



Proportion of users at each similarity level who send at least 1 like/tweet

### FAQ: What happens when we don't subset on users who followed 20 people?



We can't measure a causal effect, but the correlation still holds