

Self-Censorship on Facebook

Sauvik Das and Adam Kramer
sauvik@cmu.edu , akramer@fb.com

Self-censorship?



- Self-censorship is the act of preventing oneself from sharing a thought.
- We've all done it. (We're doing it right now).

On Social Media?

- Social media adds an additional phase of filtering: *after* a thought has been expressed.
- **Last-minute self-censorship**

Why study it?

- Last-minute self-censorship can be both helpful and hurtful.
- Ripe opportunity to explore design implications *and* understand user behavior.

WHAT WE DO KNOW

What we know

Scarcely studied in its own right: it's hard to measure what's not there!

Boundary Regulation Strategy

- People present themselves differently to different social circles (Goffman '59)
- People have trouble maintaining consistency of presentation across social-contexts on social media (Fredric & Woodrow '12, Wisniewski et al. '12)

Tied to Audience

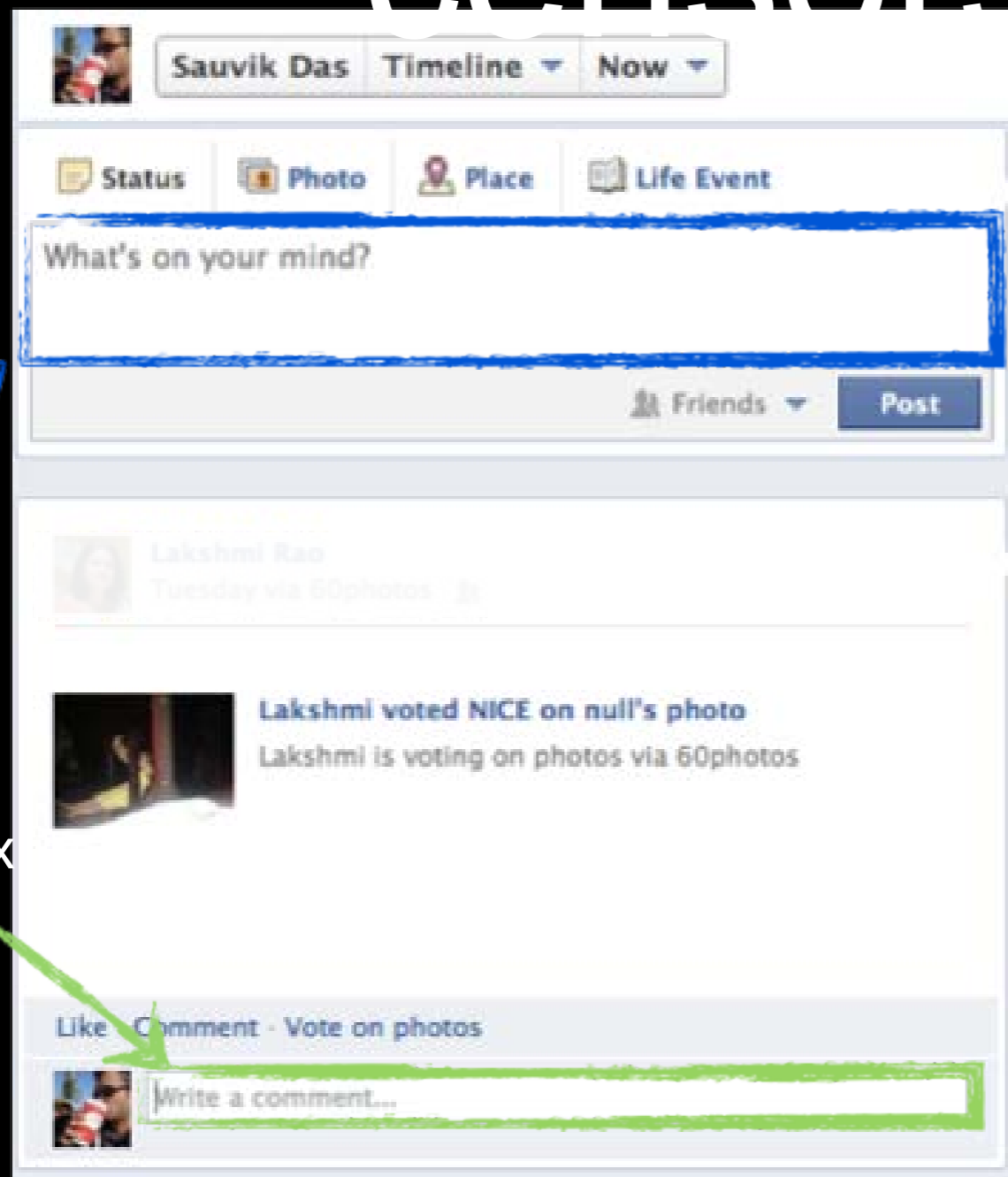
- People have an “imagined audience” when they share content on social media (Marwick & boyd ‘10)
- But are bad at estimating their audience (Bernstein et al. ‘13)
- When given the right tools, people selectively share and exclude content from different audiences (Kairam et al. ‘12)
- People said they would self-censor half as much content if given the right audience selection and exclusion tools (Sleeper et al. ‘13)

What we don't know

- How often do people self-censor?
- What sorts of content gets self-censored?
- What factors are associated with being a more frequent self-censor?

Methodology

Measuring censorship



Composer

Comment Box

Registered that input occurred after 5 characters entered.

10 minute reset time, or after submission.

Compared “possible” posts with “shared” posts. Difference was considered **self-censored**.

Measuring censorship

- Censorship measured as a per-user count.
- 3.9 million randomly selected U.S./U.K. Facebook users.
- Logged for 17 days (July 6th-22nd, 2012).

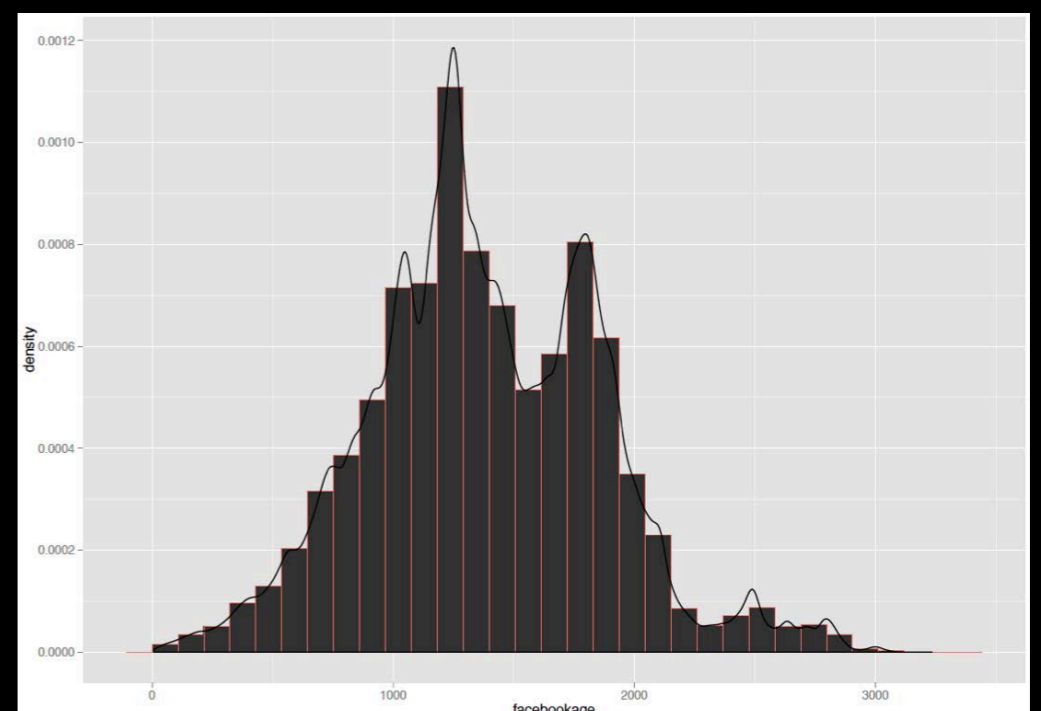
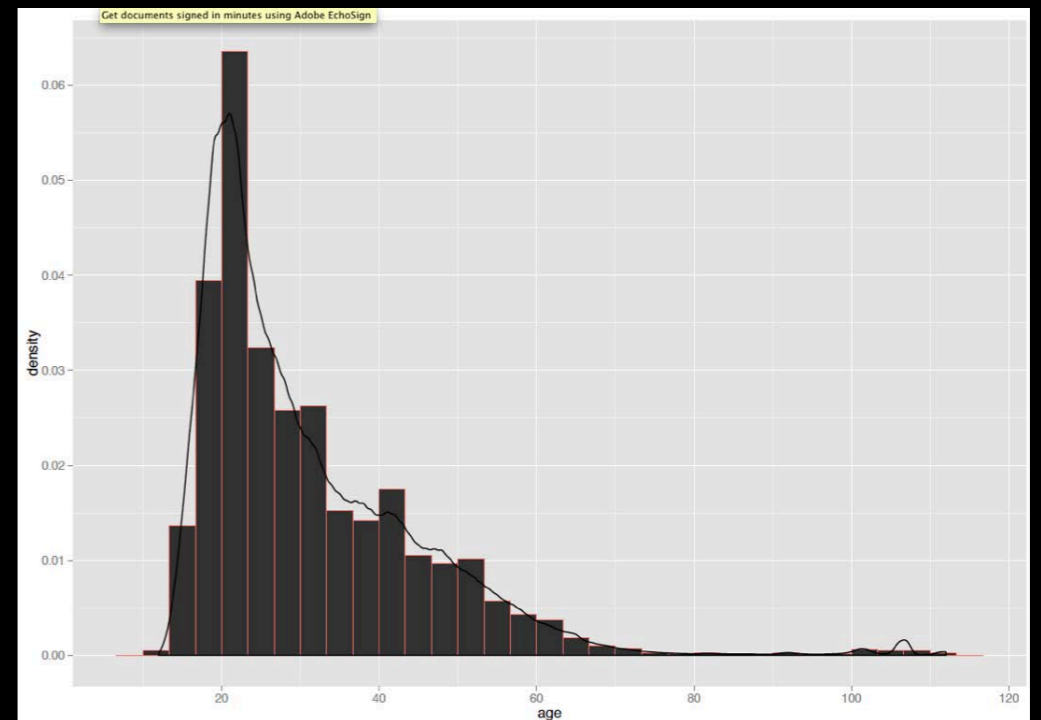
Descriptive Stats

n=3,941,161

mean age: 30.9 years (s.d. = 14.1)

mean experience: 1386 days (s.d. = 401)

57% female



FINDINGS

How often do people self-censor?

Scale

71% of our sample self-censored at least once.

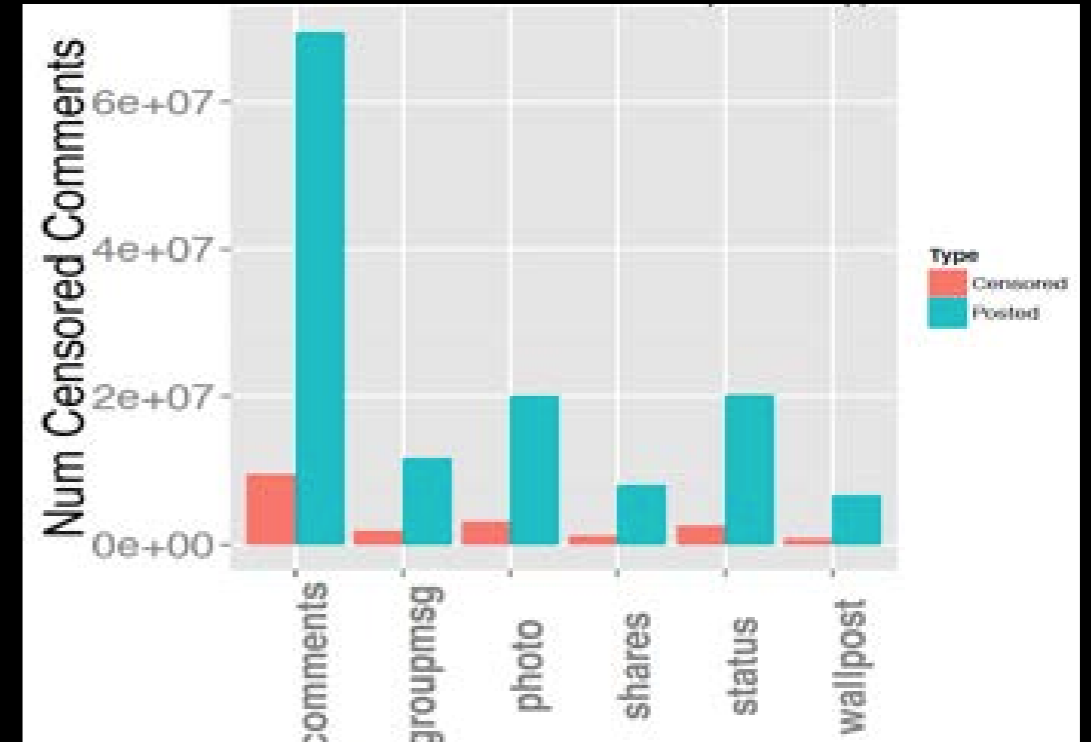
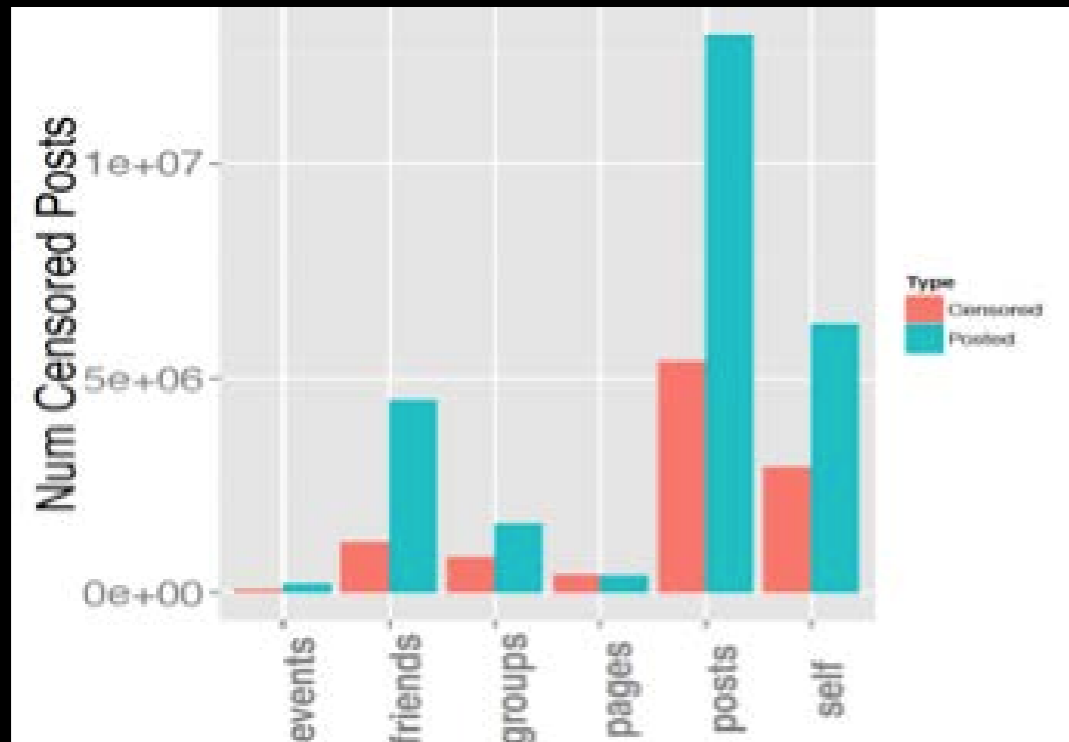
- 51% censored at least one post.
- 44% censored at least one comment.

Scale

- 33% of all potential posts censored.
- 13% of all potential comments censored.

What sort of content gets
self-censored?

What gets censored?



Content Type	Censorship Rate
Groups	38.2%
Status Updates	34.5%
Events	25.3%
Friend's Timeline	24.8%

Content Type	Censorship Rate
Photos	14.7%
Group Msg	14.5%
Shares	12.7%
Status Update	12.2%
Wall Post	10.8%

Audience Uncertainty

Low

High

Comments Events Friend Timeline Status Updates

- Low censorship
- High censorship

Audience Uncertainty




Low

High

Comments Events Friend Timeline Status Updates

Groups



-  Low censorship
-  High censorship
-  High censorship

Audience Uncertainty




Low

High

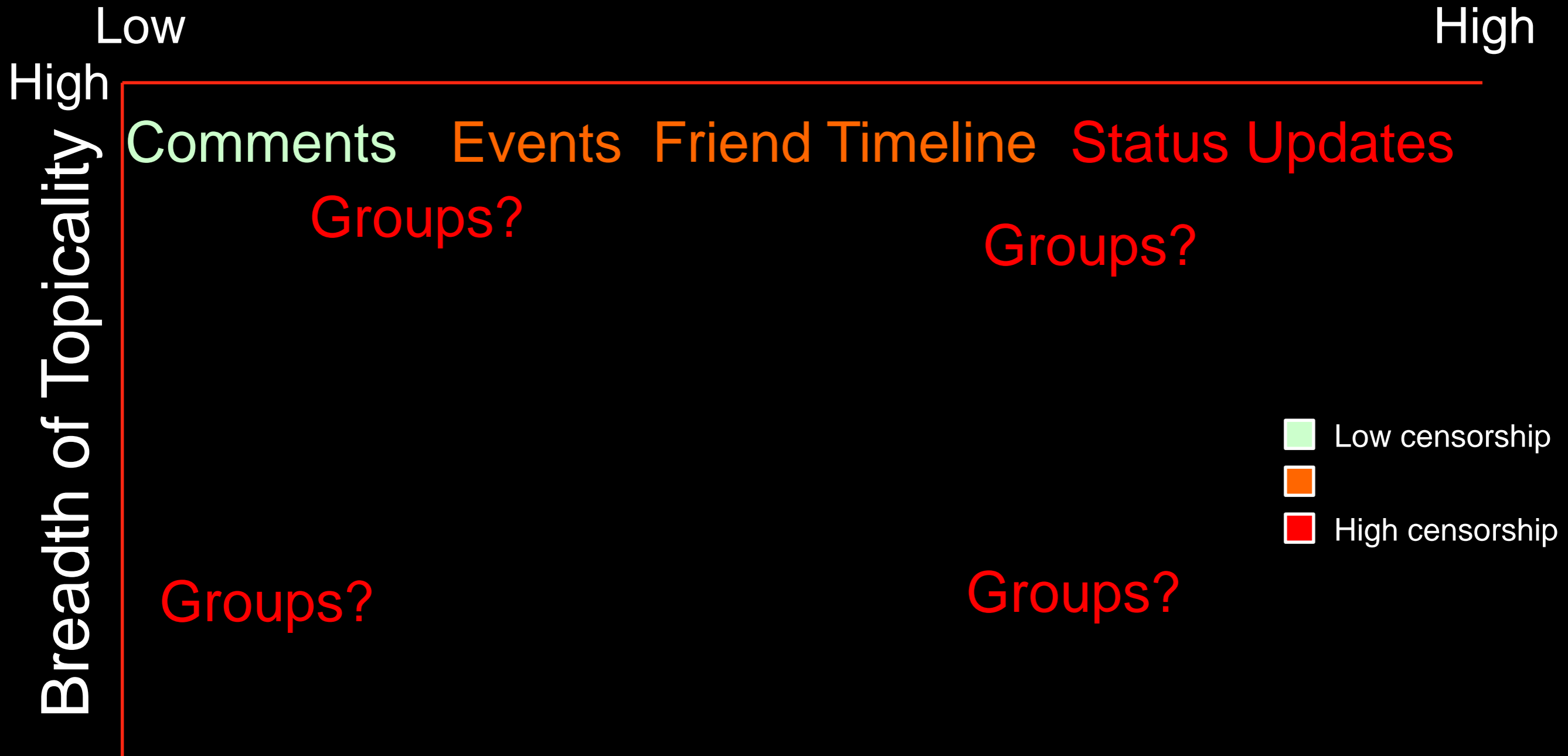
Comments Events Friend Timeline Status Updates

Groups + Seen State



-  Low censorship
-  High censorship
-  High censorship

Audience Uncertainty



Summary

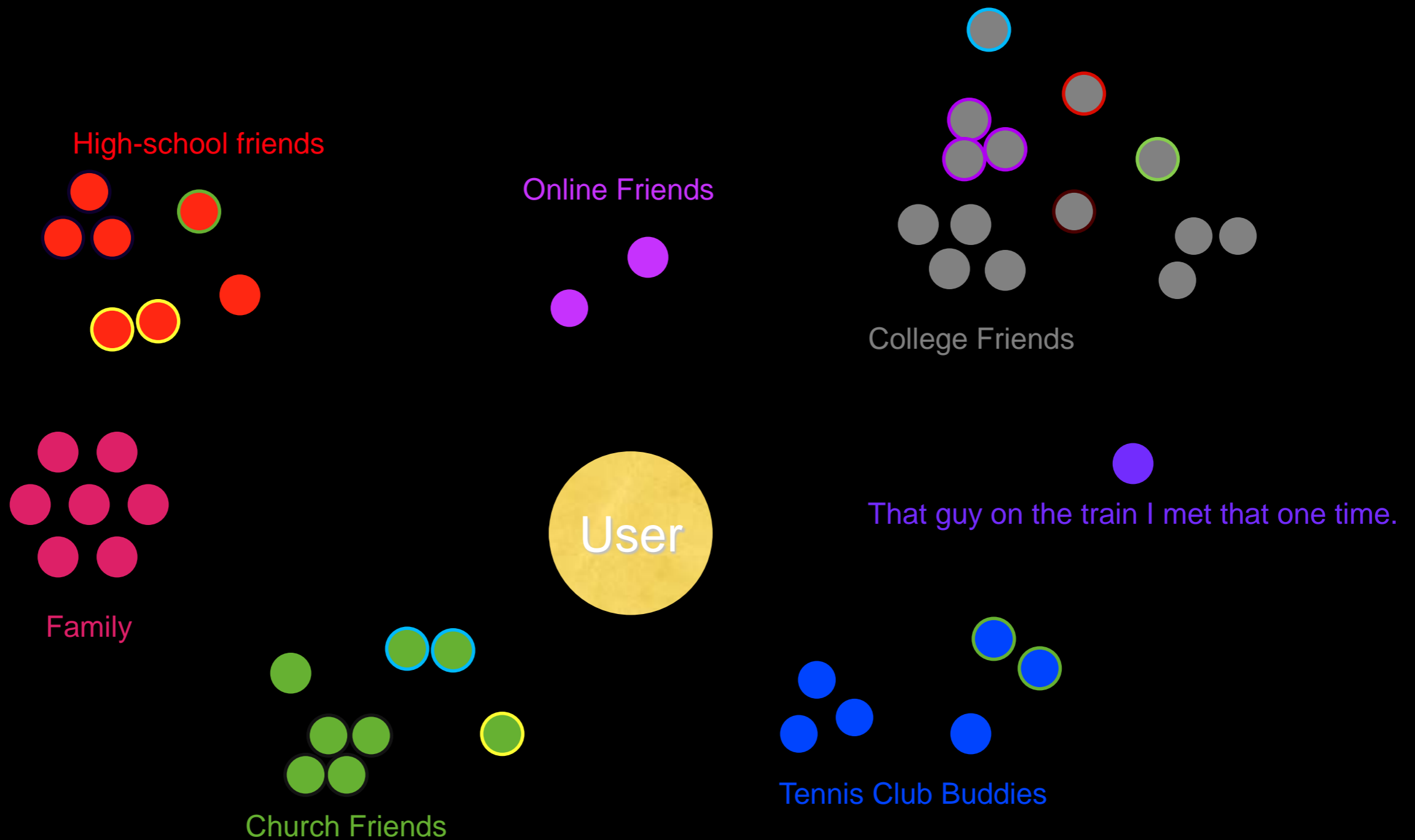
- High audience uncertainty correlates with higher self-censorship.
- Broader topicality correlates with higher self-censorship.
- Groups can help us understand how these dimensions intermix.

What factors are associated
with being a more frequent
self-censorer?

Factors of interest... Also controlled for...

- Social Graph Diversity
- Audience Selection Tools
- Activity on Site
- Age
- Experience with the Site
- Privacy Settings
- Gender

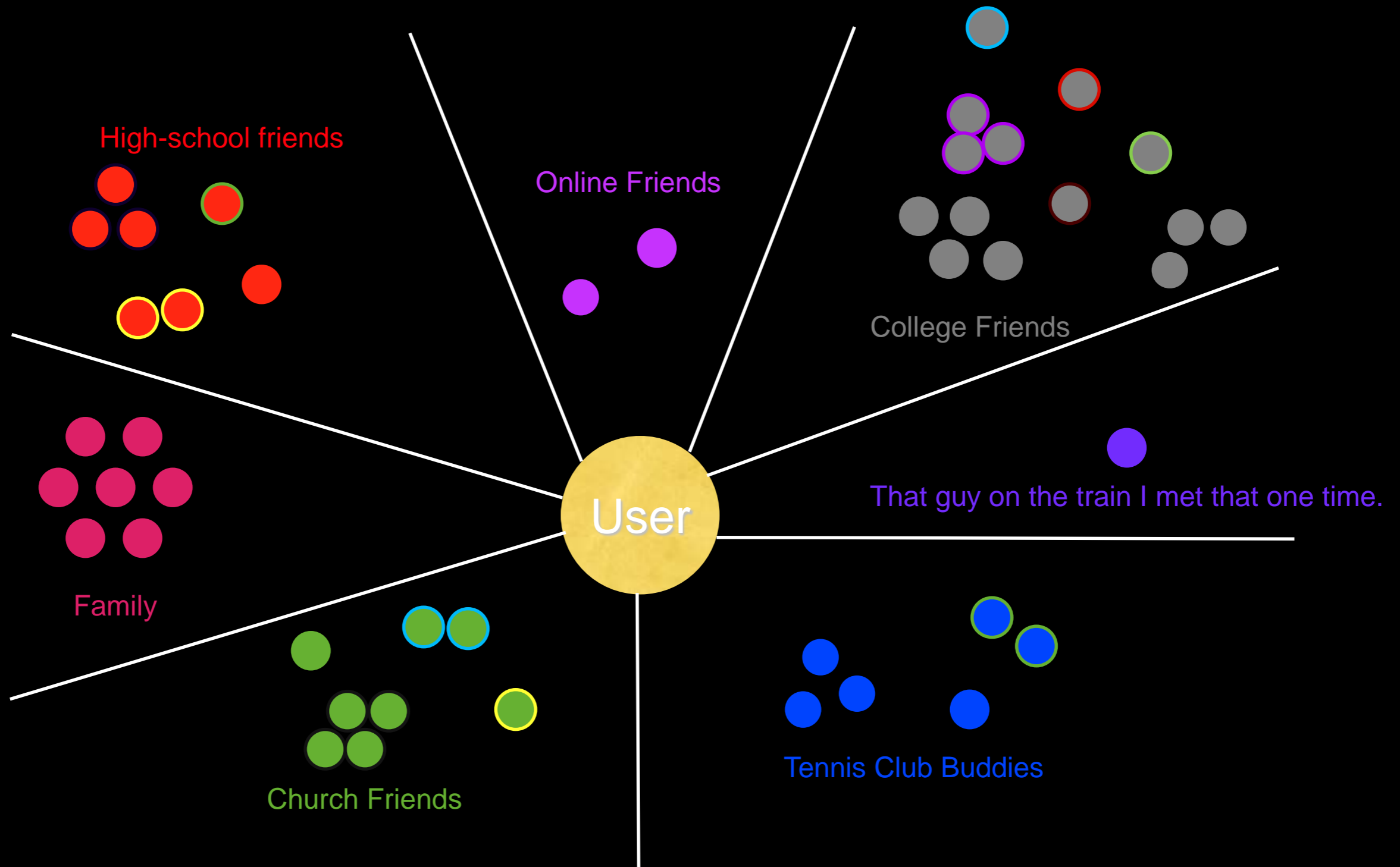
Hypotheses



People with more diverse social graphs will censor more.

Social Graph Diversity Features

Feature	Expected Effect
Average number friends of friends	+
Biconnected components	+
Friendship density	-
Friend age entropy	+
Friend political entropy	+



People who use audience selection tools will censor less.

Audience Selection Tool Features

Feature	Expected Effect
Group member count	-
Friendlist created	-
Private messages sent	-

Modeling self-censorship

- Employed a Negative Binomial Regression.
- Negative Binomial was favored over Poisson because of the presence of overdispersion.
- Response was the count of censored count, offset with amount of created content.
- Coefficients estimated separately for posts and comments. Only posts reported in this presentation.

Posts Model Negative Binomial Regression Coefficients

Category	Feature	Coefficient	Exp.
Social Graph Diversity	Average number friends of friends	1.32	+
	Biconnected components	1.12	+
	Friendship density	0.97	-
	Friend age entropy	0.96	+
	Friend political entropy	0.92	+
Audience Selection Tools	Group member count	1.29	-
	Buddylists created	1.13	-

All significant at $p = 0.01$

Posts Model Negative Binomial Regression Coefficients			
Category	Feature	Coefficient	Expectation
Social Graph Diversity	Average number friends of friends	1.32	+
	Biconnected components	1.12	+
	Friendship density	0.97	-
	Friend age entropy	0.96	+
	Friend political entropy	0.92	+

- Diversity appears to have two components:
 - People with larger 2nd degree networks and more distinct social circles censor more
 - People whose friends are more diverse censor less.

Posts Model Negative Binomial Regression Coefficients			
Category	Feature	Coefficient	Expectation
Audience Selection Tools	Group member count	1.29	-
	Friendlists created	1.13	-

- Audience selection tools had the opposite effect that we expected!
- People who were part of more groups and people who created more friend lists actually self-censor more posts.

Conclusion

Magnitude

- Self-censorship occurs frequently in social media, as expected.
- 71% censored at least once.
- Frequency varies by the nature of the content (post vs. comment) and its context (group post vs status update).

Not Just Audience

- People censor more as audience uncertainty increases.
- People censor more as the breadth of the topicality increases.
- Groups are weird, and possibly the key to understanding the interaction.

Diversity

- People with more diverse friends censor less.
- But, people with more disparate social contexts censor more.

Boundary Regulation

- Self-censorship does appear to be a boundary regulation strategy.
- Users who have more distinct social circles self-censor more.
- Even controlling for the use of audience selection tools!

Boundary Regulation

Present audience selection tools are insufficient or untrusted by users who need to balance many different social contexts.

Limitations

- Our metric is only a correlate.
- Groups are still wildcards.
- We don't actually know what gets self-censored.

Questions?

Boundary Regulation

References

- Farnham, S.D. & Churchill, E.F. Faceted identity, faceted lives. Proc. CSCW 2011, 359.
- Frederic, S. & Woodrow, H. Boundary regulation in social media. Proc. CSCW 2012, 769.
- Kairam, S., Brzozowski, M., Huffaker, D., & Chi, E. Talking in circles. Proc. CHI 2012, 1065
- Marwick, A.E. & boyd, d. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. New Media & Society, 13 (2010), 114.
- Wisniewski, P., Lipford, H., & Wilson, D. Fighting for my space. Proc. CHI 2012, 609.

Audience References

- Sleeper, M., Balebako, R., Das, S., McConohy, A., Wiese, J., & Cranor, L.F. The Post that Wasn't: Exploring Self-Censorship on Facebook. Proc. CSCW 2013 under review.
- Acquisti, A. & Gross, R. Imagined Communities: Awareness, Information Sharing, & Privacy on the Facebook. Priv. Enhancing Tech. 4258, (2006), 36–58.
- Tufekci, Z. Can You See Me Now? Audience and Disclosure Regulation in Online Social Network Sites. Bull. Sci., Tech, & Soc, 28 (2007), 20.
- Ardichvili, A., Page, V., and Wentling, T. Motivation and barriers to participation in virtual knowledge-sharing communities of practice. Journ. of Knldg. Mgmt. 7, 1 (2003), 64–77.

Politics References

- Hayes, A.F., Glynn, C.J., and Shanahan, J. Validating the Willingness to Self-Censor Scale: Individual Differences in the Effect of the Climate of Opinion on Opinion Expression. *International Journal of Public Opinion Research* 17, 4 (2005), 443–455.
- Hayes, A.F., Scheufele, D.A., and Huges, M.E. Nonparticipation as Self-Censorship: Publicly Observable Political Activity in a Polarized Opinion Climate. *Political Behavior* 28, 3 (2006), 259–283.
- Huckfeldt, R., Mendez, J.M., and Osborn, T. Disagreement, Ambivalence, and Engagement: The Political Consequences of Heterogeneous Networks. *Political Psychology* 25, 1 (2004), 65–95.
- Mutz, D. The Consequences of Cross-Cutting Networks for Political Participation. *American Journal of Political Science* 46, 4 (2002), 838–855.
- Noelle-Neumann, E. The spiral of silence: a theory of public opinion. *Journal of Communication* 24, (1974), 43–51.

Gender References

- Bowman, N.D., Westerman, D.K., & Claus, C.J. How demanding is social media: Understanding social media diets as a function of perceived costs and benefits – A rational actor perspective. *Comp. in Hum. Behav.*, 28 (2012), 2298.
- Cho, S.H. Effects of motivations and gender on adolescents' self-disclosure in online chatting. *Cyberpsychology & Behavior* 10, 3 (2007), 339–45.
- Dindia, K. & Allen, M. Sex differences in self-disclosure: A meta-analysis. *Psych. Bulletin* 112 (1992), 106.
- Seamon, C.M. Self-Esteem, Sex Differences, and Self-Disclosure: A Study of the Closeness of Relationships. *Osprey J. Ideas & Inquiry*, (2003).
- Walther, J.B. Selective self-presentation in computer-mediated communication: Hyperpersonal dimensions of technology, language, and cognition. *Comp. in Hum. Behav.* 23 (2007), 2538.

ZINB References

- Lambert, D. Zero-Inflated Poisson With an Application to Defects in Manufacturing. *Technometrics* 34, 1 (1992), 1–14.
- Mwalili, S.M., Lesaffre, E., and Declerck, D. The zero-inflated negative binomial regression model with correction for misclassification: an example in caries research. *Statistical methods in medical research* 17, 2 (2008), 123–39.

Full List of Features

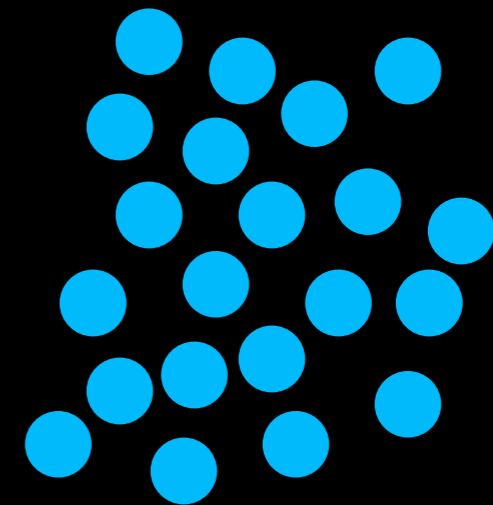
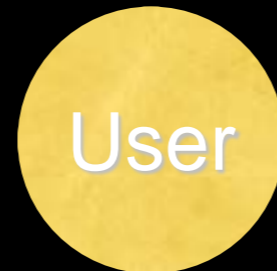
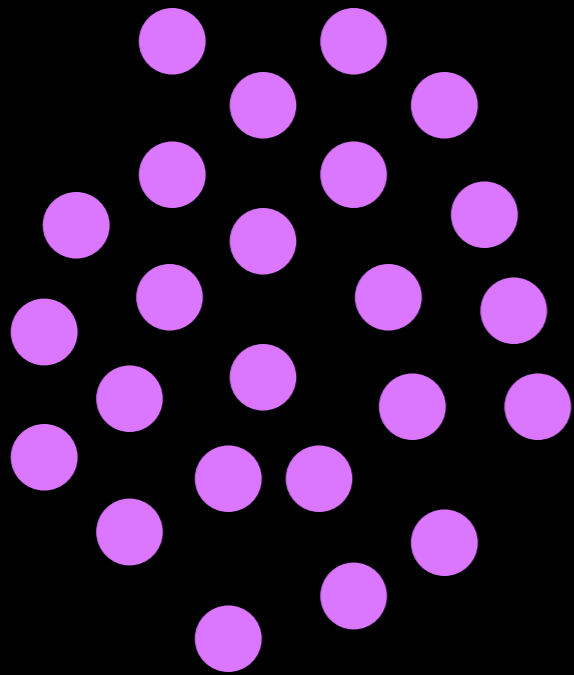
Demographic	Behavioral	Social Graph
Gender	Messages sent	Number of friends
Age	Photos added	Connected components
Political affiliation	Friendships initiated	Biconnected components
Media privacy	Deleted posts	Average age of friends
Wall privacy	Deleted comments	Friend age entropy
Group member count	Buddylists created	Mostly (C/L/M)* Friends
Days since joining Facebook	Checkins	Percent male friends
	Checkins deleted	Percent friends (C/L/M)
	Created posts	Friend political entropy
	Created comments	Density of social graph

* Conservative/Liberal/Moderate

Measuring Censorship

- **Gold standard:** honest users reporting instances of self-censorship
- Practical constraints:
 - **speed:** slower site speeds would present a confound.
 - **invisibility:** had to run behind the scenes-- manipulating the UI would require extensive user testing.
 - **privacy:** ethical consideration that prevents us from logging content that users do not want to share.

Female Friends

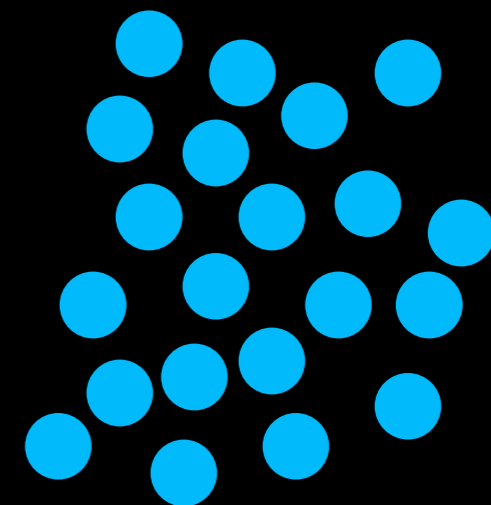
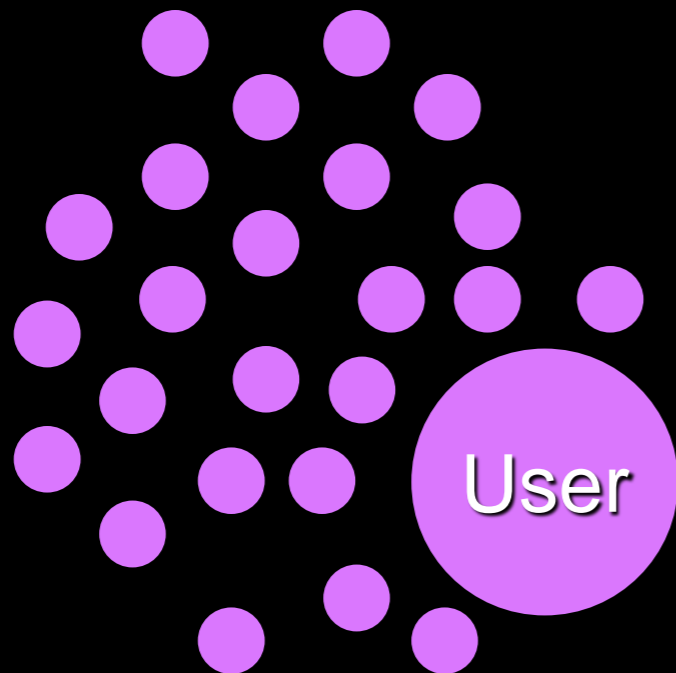


Male Friends

Males will censor more than females.

Percentage Male Friends
Own Gender X Percentage Male Friends

Female Friends

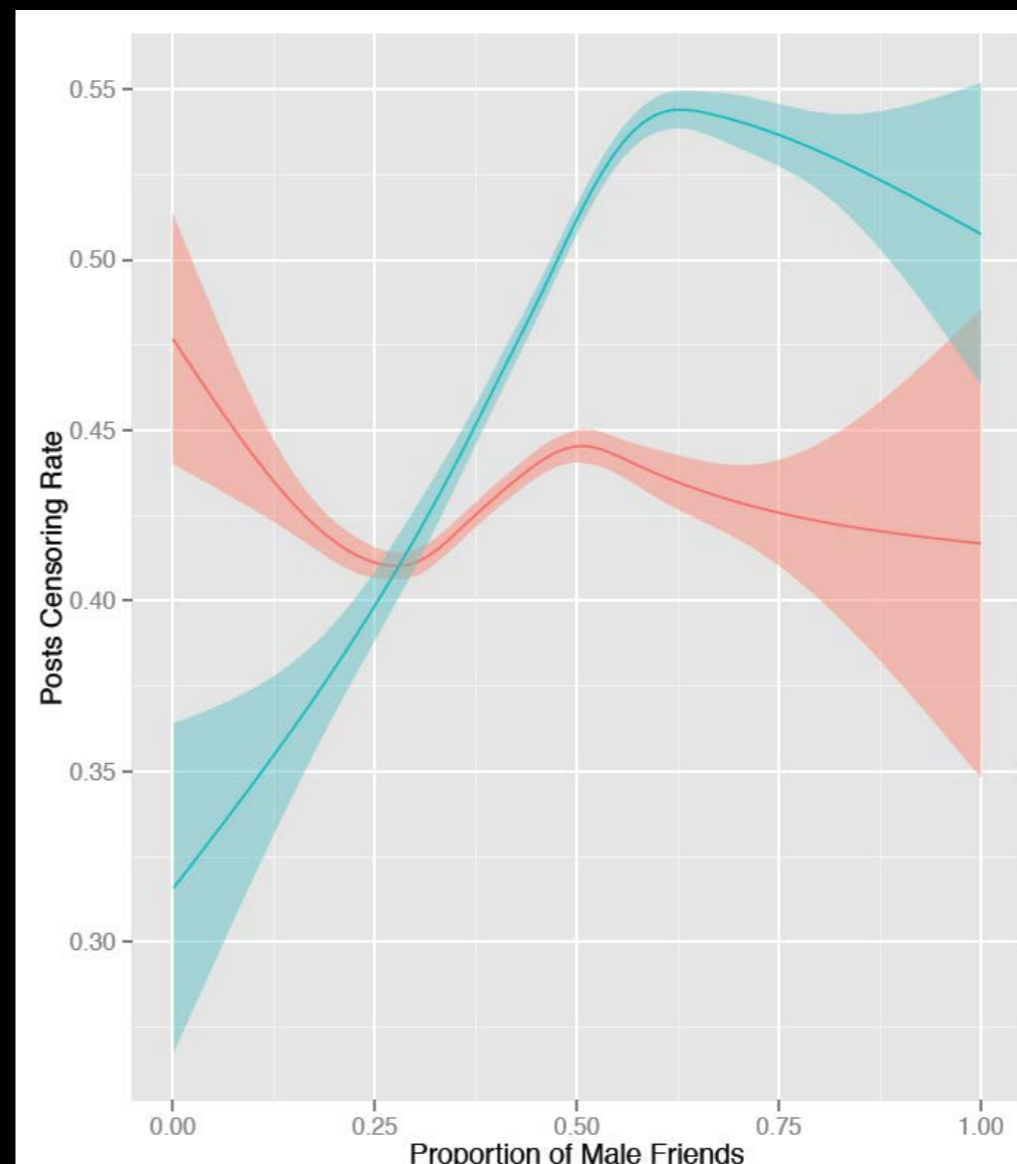


Male Friends

People with more opposite sex friends will censor more.

Posts Model Negative Binomial Regression Coefficients			
Category	Feature	Coefficient	Baseline
Gender	Gender: Male	1.26	Female
	Gender: Male X Percentage Male Friends	1.11	Female X Percentage Male Friends

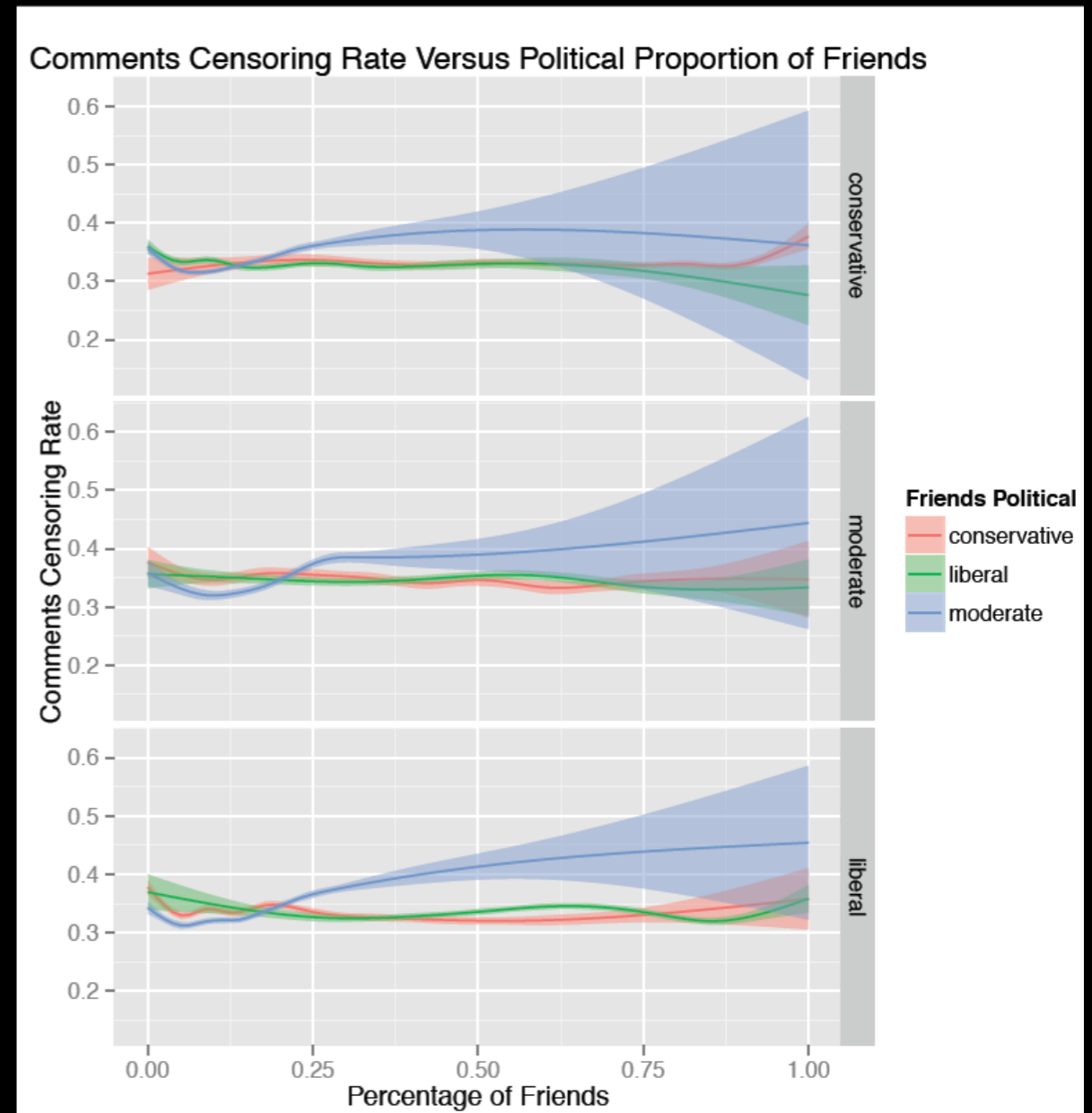
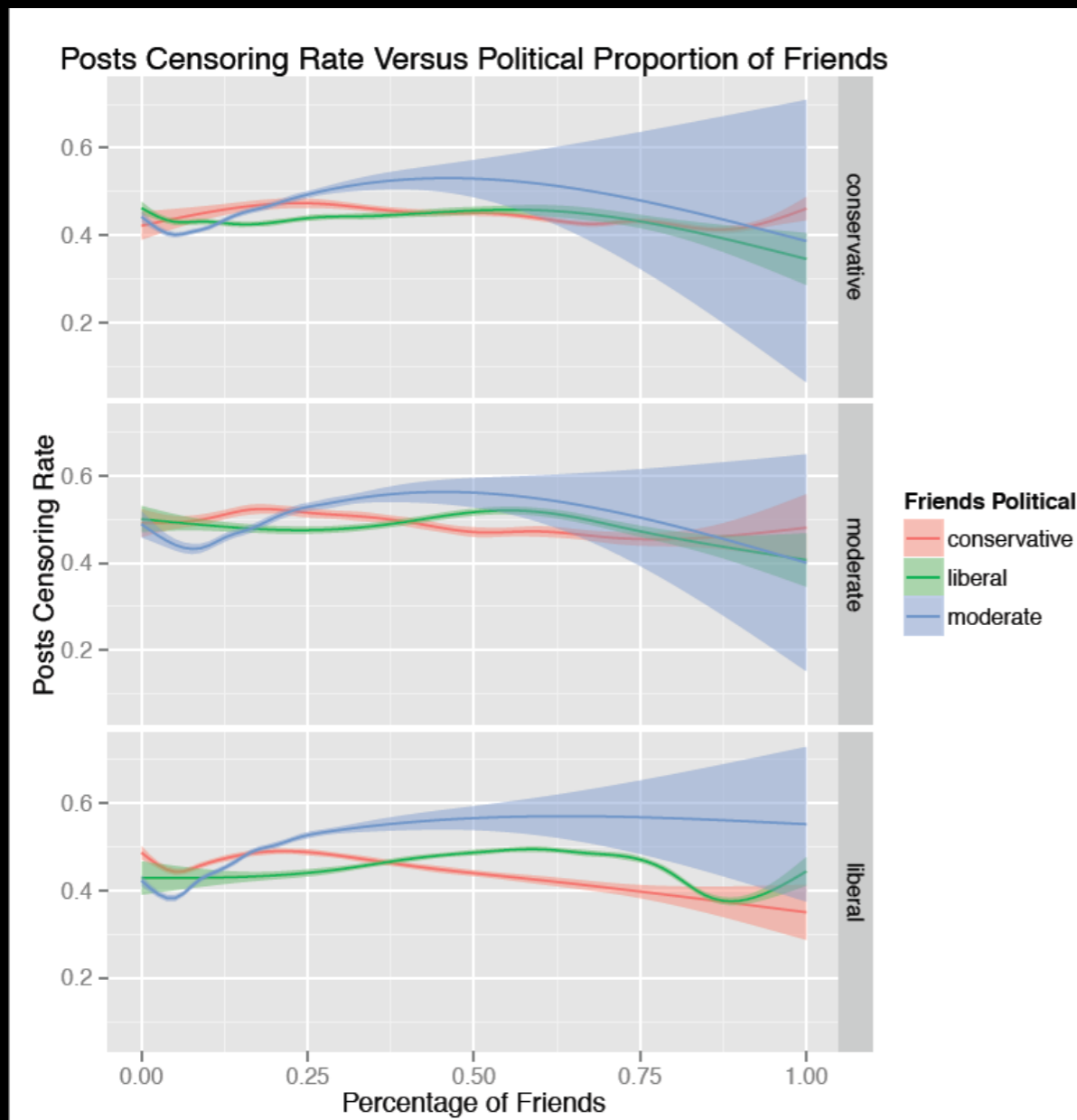
- Male censor substantially more posts than females.
- Interestingly, this is even true as more males are part of their social graph.

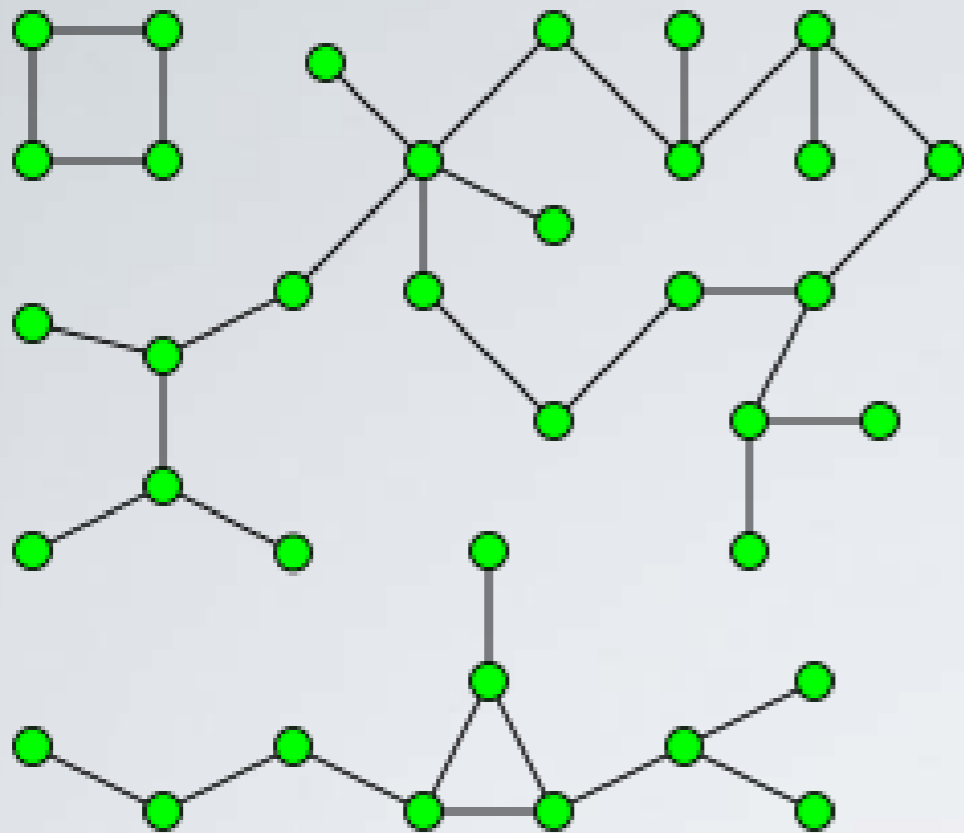


Take-aways

- User-specific factors do seem to be associated with self-censorship.
- Males censor much more than females.
- Further research will be needed to discern why.

Modeling Self-censorship





Connected Component

Biconnected Components

