GEORGETOWN UNIVERSITY

# Comparing Online and Offline Partisan Segregation Using a Novel Panel of Twitter Users

Tiago Ventura[1] , Megan Brown[2], Jonathan Nagler[3], and Joshua A. Tucker[3]

[1] Georgetown University, [2] University of Michigan, [3] Center for Social Media and Politics, NYU

IC$^2$S$^2$ - 07/18/2024

# Motivation

OPINION > CAMPAIGN

THE VIEWS EXPRESSED BY CONTRIBUTORS ARE THEIR OWN AND NOT THE VIEW OF THE HILL

## How social media fuels U.S. political polarization — what to do about it

BY PAUL BARRETT, JUSTIN HENDRIX AND GRANT SIMS, OPINION CONTRIBUTORS - 09/13/21 4:00 PM ET

SHARE TWEET

---

BOOKS

LISTEN & FOLLOW

## How the polarizing effect of social media is speeding up

September 9, 2022 · 5:01 AM ET

By Ari Shapiro, Michael Levitt, Christopher Intagliata

---

KEYWORDS: CHRISTOPHER MIMS

## Why Social Media Is So Good at Polarizing Us

Mathematicians are teaming up with political scientists to create models of how social media divides us, and results suggest at least one popular solution might actually make the problem worse

**Social media**

⬇

**Echo Chambers**

⬇

**Reduce cross-cutting exposure**

⬇

**Polarization**

# Most of the literature:

Social media

⬇

**Echo Chambers**

⬇

Reduce cross-cutting exposure

⬇

Polarization

# Our contributtion

➜ **To solve the causal chain**, we need to measure online segregation relative to other channels through which voters **consume information or interact** with ingroup and outgroup voters

➜ **Previous related studies:**

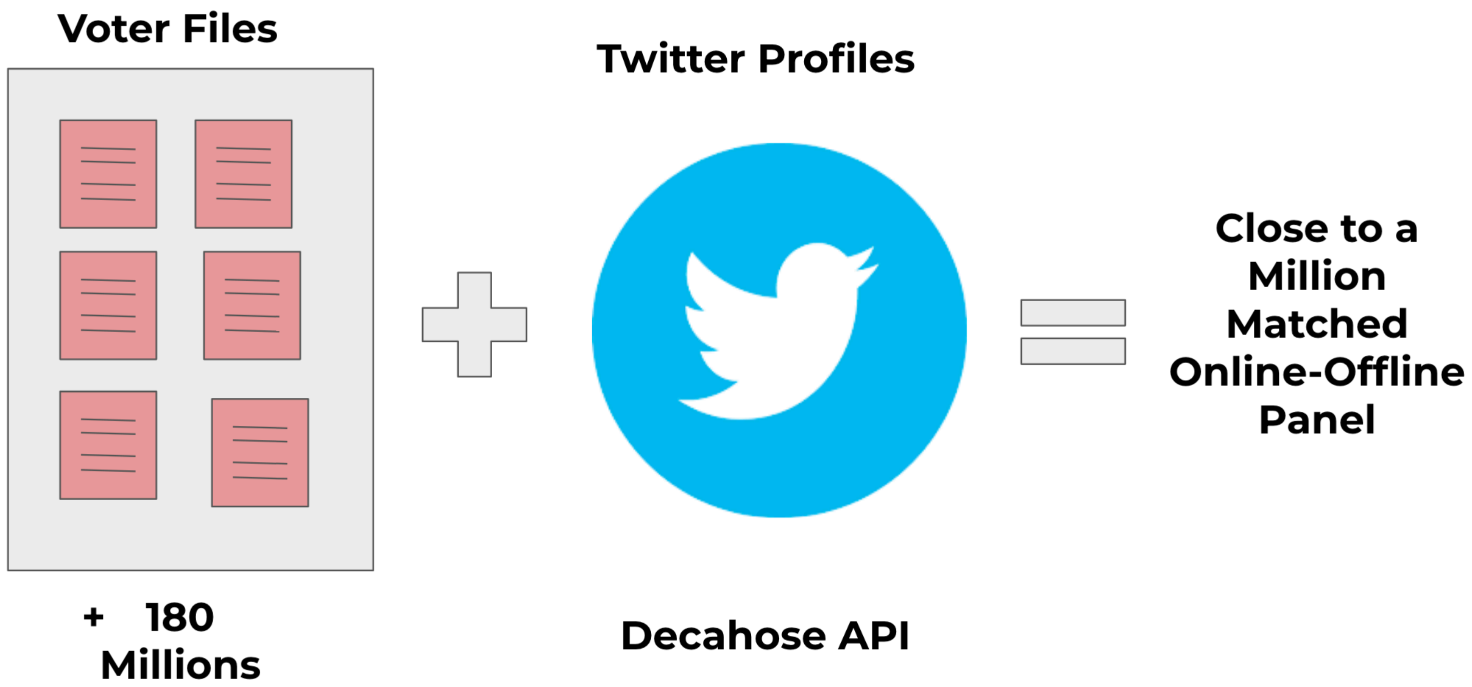
- TV news Consumption (Muise et. al. 2022)
- Self-reported online vs offline networks (Gentzkow and Shapiro 2011)

➜ We provide meaningful comparisons for the same social media users between levels of **offline and online segregation**

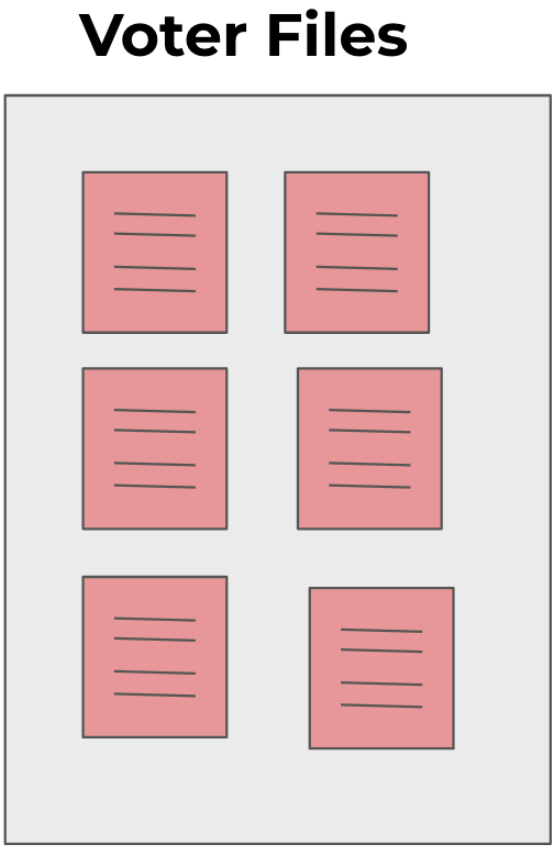
# Research Question

What is the relationship between offline partisan sorting and online echo-chambers?

# Data Infrastructure



# Offline Information: Voter Files




**Data Collection for every matched voters:**

- Voter file demographics (gender, race, partisanship, religion)
- Residential location (9 digits lat and long)
- Closest 1.000 neighboors + their partisanship.

# Online Information: Twitter Data

**Data Collection for every matched voters:**

- Collect their full network (people they follow and follow them) ~ 57M
- Collect their most recent timelines (3200 tweets) + 900k * 3,2k
- Parse their timelines.

# Offline Partisan Segregation

$$\text{Offline Outgroup Proximity} = \frac{\sum_{k=1}^{1000} \frac{1}{d+1} (p_k = q_i)}{\sum_{k=1}^{1000} \frac{1}{d+1}}$$

**Where:**

- $i$ is a matched voters
- $k$ is a given neighbor
- $d$ is the distance in meters between the neighbor and the individual
- $p_k$ is the partisanship of the neighbor
- $q_i$ is the opposite party of the individual whose exposure is being measured.

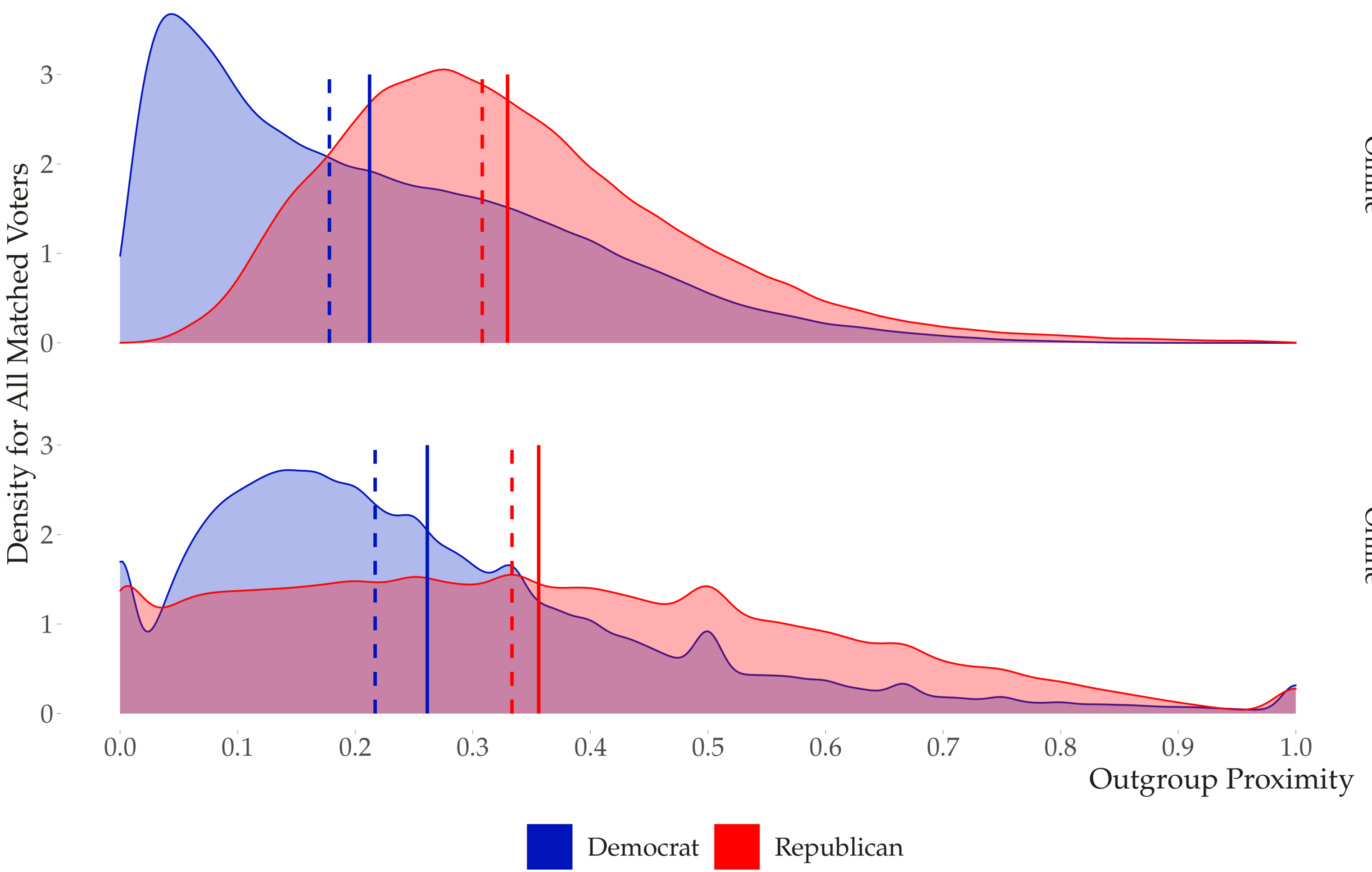
# Online Partisan Segregation

$$\text{Online Outgroup Proximity} = \frac{\sum_{k=1}^{n} \log(a+1)(p_k = q_i)}{\sum_{k=1}^{n} \log(a+1)}$$

**Where:**

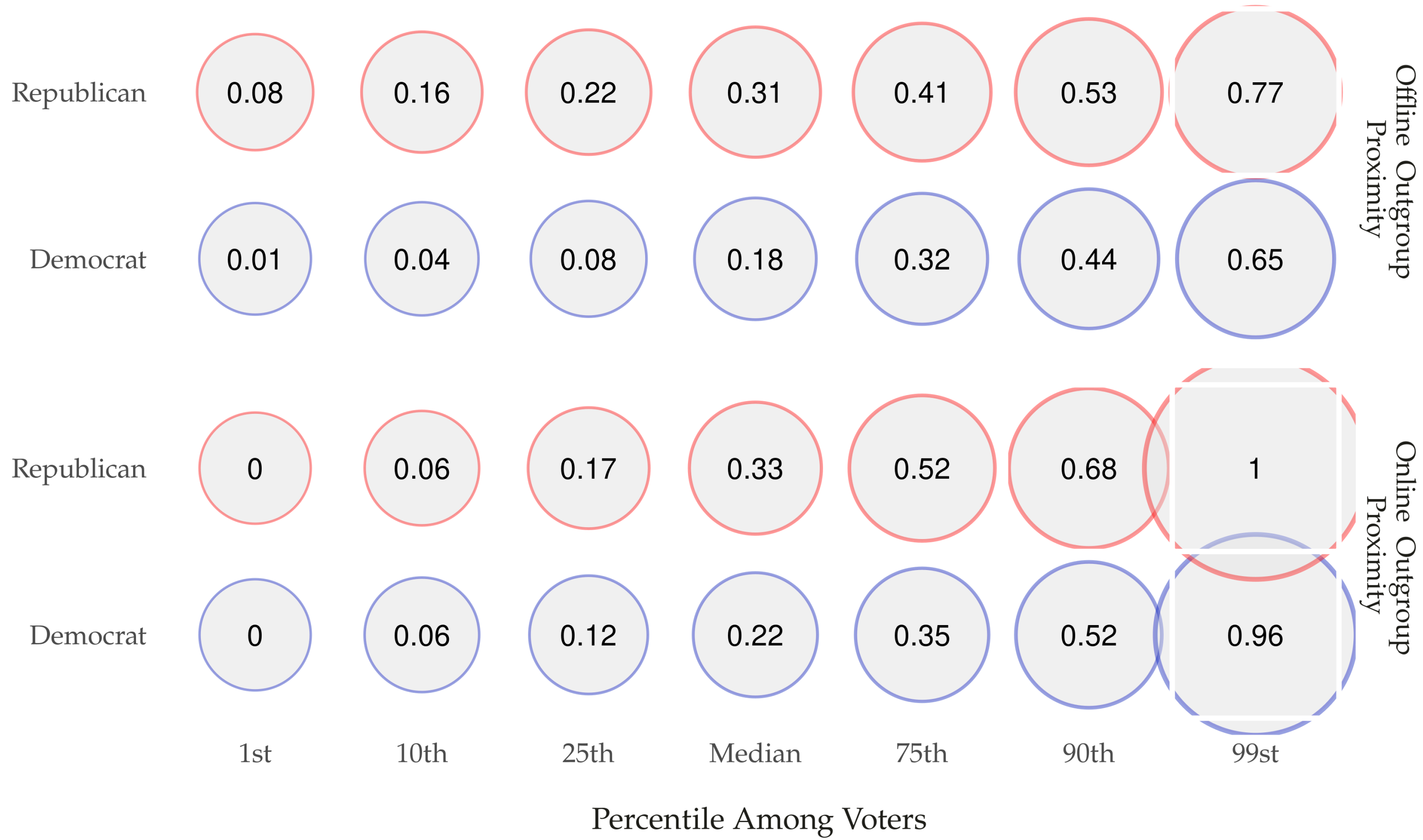
- $i$ is a matched voter
- $k$ is a friend (followed by) the matched voter
- $a$ is the number of interactions between the friend and a user $i$
- $p_k$ is the partisanship of the friend
- $q_i$ is the opposite party of the individual whose exposure is being measured. ]

# Online vs Offline Exposure




*Straight Lines Represents the Median and Dashed Lines the Mean of the Distributions*

# Online vs Offline Exposure by Quantiles



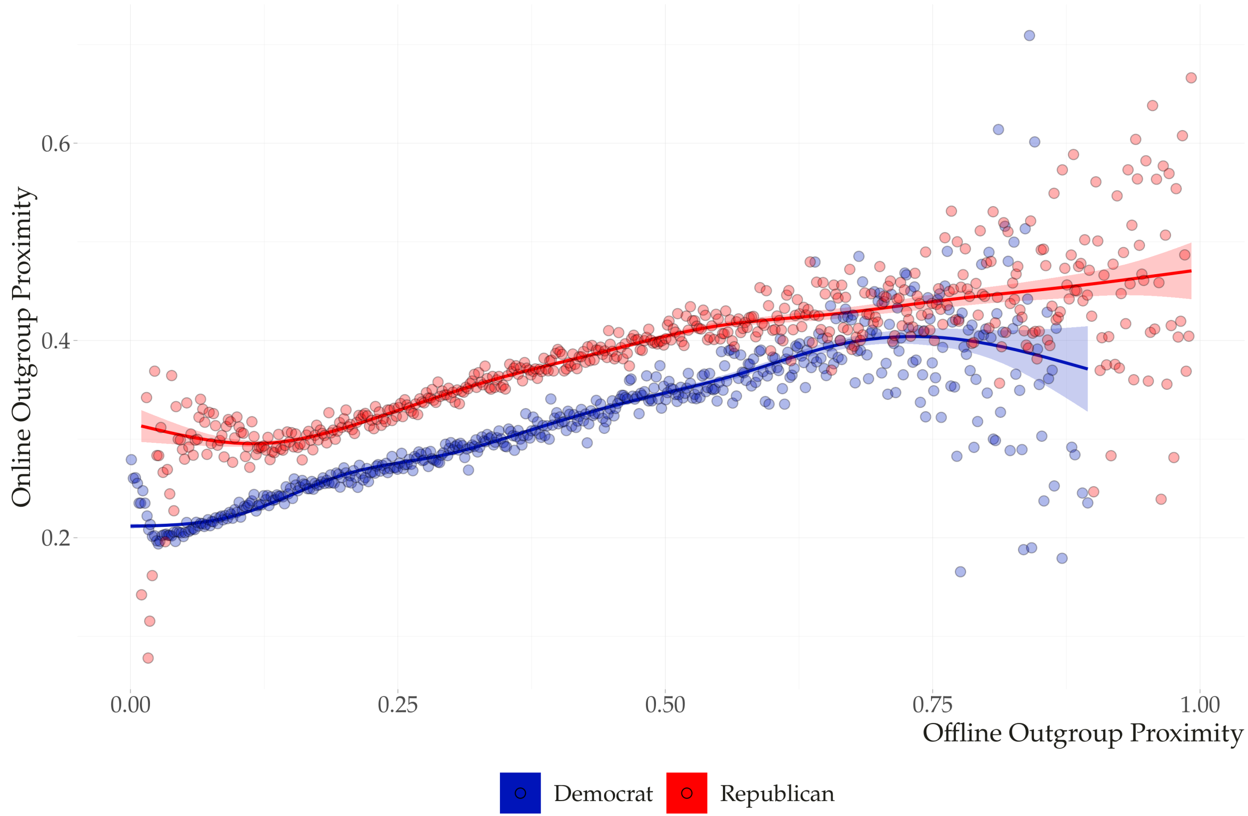
# Comparing Offline and Online Exposure Across Subgroups

**Table 2** Comparing Online and offline outgroup proximity Across Subgroups

| Variable | Online Outgroup Proximity | Offline Outgroup Proximity | Paired-difference | Z-Score(p-Value) |
|---|---|---|---|---|
| **Sociodemographics** | | | | |
| Male | 0.306 | 0.259 | 0.047 | 151.094 ⋆⋆⋆ |
| Female | 0.280 | 0.249 | 0.030 | 79.614 ⋆⋆⋆ |
| White | 0.300 | 0.278 | 0.022 | 76.691 ⋆⋆⋆ |
| Non-White | 0.287 | 0.212 | 0.074 | 178.788 ⋆⋆⋆ |
| Age: <35 | 0.297 | 0.242 | 0.055 | 141.319 ⋆⋆⋆ |
| Age: 35 - 60 | 0.296 | 0.259 | 0.037 | 112.312⋆⋆⋆ |
| Age: +60 | 0.289 | 0.274 | 0.013 | 18.051 ⋆⋆⋆ |
| **Political Variables** | | | | |
| Democrat | 0.261 | 0.212 | 0.050 | 144.023 ⋆⋆⋆ |
| Republican | 0.356 | 0.329 | 0.023 | 44.661 ⋆⋆⋆ |
| Blue States | 0.256 | 0.205 | 0.049 | 147.366 ⋆⋆⋆ |
| Swing States | 0.304 | 0.270 | 0.033 | 66.905 ⋆⋆⋆ |
| Red States | 0.331 | 0.293 | 0.036 | 83.729 ⋆⋆⋆ |

*Note: Significance of differences between the datasets were tested with a two-sided mean difference paired z-test(⋆⋆⋆ means $p - value < 0.001$)*

# Correlation between online and offline exposure

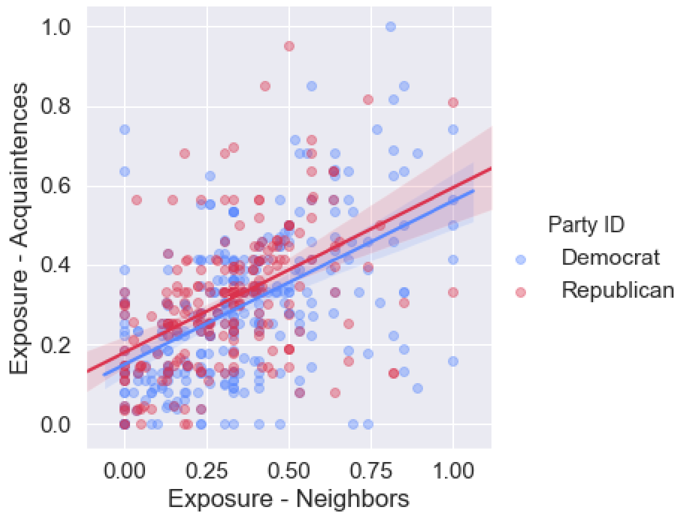


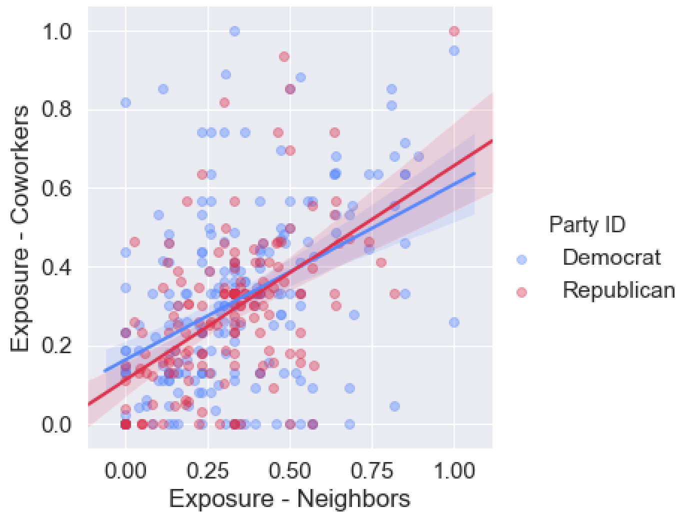
# Modeling Online Echo Chambers

| Dependent Variable: | Online Proximity | | |
|---|---|---|---|
| Model: | (1) | (2) | (3) |
| *Variables* | | | |
| Age: 35-60 | -0.0126** | -0.0122*** | -0.0103*** |
| | (0.0059) | (0.0023) | (0.0012) |
| Age: +60 | 0.0094 | 0.0114* | 0.0132*** |
| | (0.0170) | (0.0058) | (0.0019) |
| Male | -0.0260*** | -0.0247*** | -0.0243*** |
| | (0.0082) | (0.0028) | (0.0011) |
| White | 0.0193*** | 0.0161*** | 0.0141*** |
| | (0.0043) | (0.0020) | (0.0012) |
| Republican | 0.0113 | 0.0143*** | 0.0165*** |
| | (0.0133) | (0.0033) | (0.0013) |
| Log of Friends on Twitter | -0.0406*** | -0.0427*** | -0.0431*** |
| | (0.0022) | (0.0008) | (0.0005) |
| Offline Proximity | -0.3205*** | -0.2593*** | -0.2361*** |
| | (0.0523) | (0.0129) | (0.0041) |
| *Fixed-effects* | | | |
| State-Level (50) | Yes | | |
| Congressional Districts (500) | | Yes | |
| census_tract (47,680) | | | Yes |
| *Fit statistics* | | | |
| Observations | 599,707 | 599,707 | 599,707 |
| $R^2$ | 0.05349 | 0.06785 | 0.15011 |
| Within $R^2$ | 0.03496 | 0.03166 | 0.02992 |

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

# Robustness Check : Yougov Survey



# Discussion and Next Steps

- We provide initial evidence that:
  - Twitter users indeed live in segregated online communities with overall lower levels of exposure to outgroup users in their networks;
  - But these levels of segregation are not particularly distinct from their geographical offline levels of proximity to outgroup voters;
  - These differences hold across distinct political, racial, ethnical and age groups.
- Future work: Focus on the effects of partisan segregation on online behavior
  - Toxicity
  - Outgroup Hostility
  - Sharing of low-quality content

# Thank you!