Identity Representation in News and Social Media

Elliott Ash (ETH Zürich)

Ruben Durante (ICREA–UPF)

Carlo Schwarz (Bocconi)

Introduction

Motivation:

- Large and growing literature on media bias/slant (Strömberg 2015)

- Focus on whether content favors a certain political agenda (e.g., left vs. right) or the interests of particular actors (e.g., owners, advertisers, lenders)

- Scant evidence on whether and how news media and politicians promote or perpetuate certain cultural stereotypes (e.g., gender, ethnicity, religion) or engage in identity politics

Methodology:

- Previous measures of media bias use hand–coding or text analysis

- Recent work starting to bring in computer vision (e.g., Boxell 2021; Adukia et al. 2021)

- We provide a combined perspective that compares content of images *and* text

Data

- 600K web articles from New York Times (NYT) and Fox News Channel with text, images, and metadata (2000–2020)

- 1.8M Facebook posts by U.S. Congress Members, also with text, images, and metadata (2010–2020).

Machine Learning

- gender and ethnicity assigned using computer–vision tools applied to image

- topics assigned using text–analysis tools applied to texts

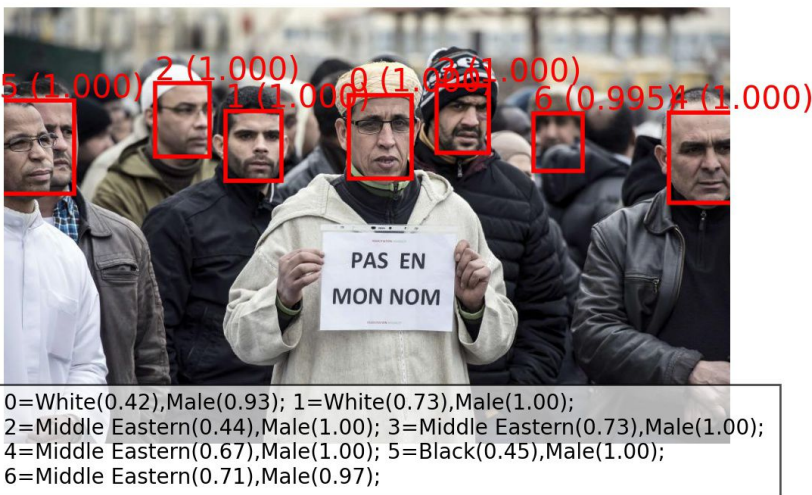- validation and comparison of ML outputs

Goals

- measure visual stereotypes in NYT and Fox News imagery, relating article text and associated image.

- analyze identity politics in politician's Facebook posts.

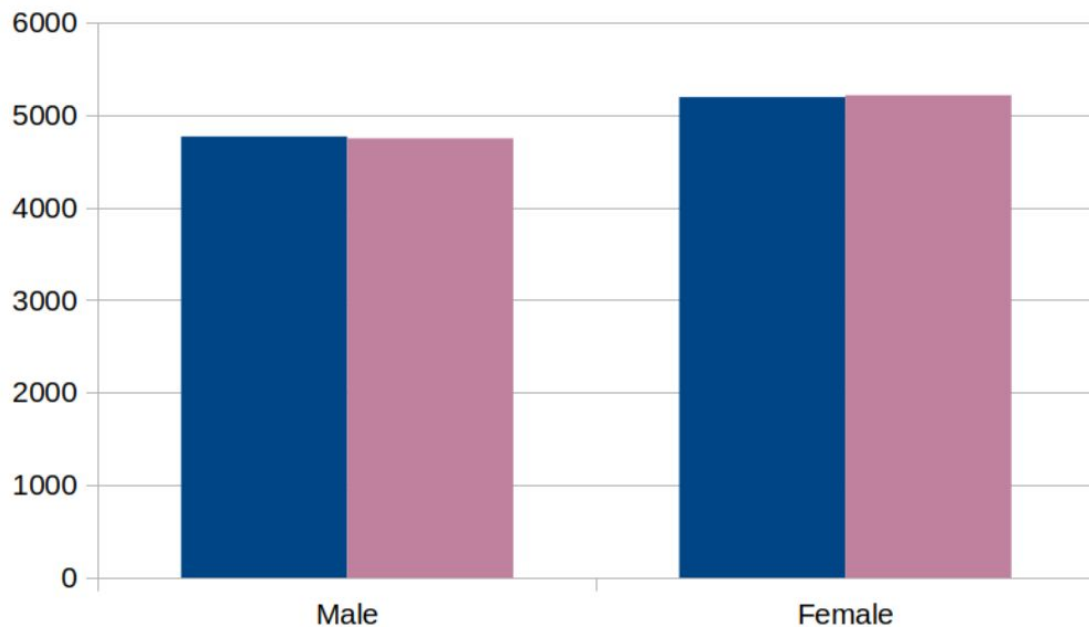# Methodology

Classifier for Gender and Ethnicity

- Recognize all faces in picture, return associated faces as pixels

  - python face_recognition package (github.com/ageitgey/face_recognition)

- Each face's pixel data converted to predictive feature vector (512 dims) using FaceNet (Schroff et al 2015)

- Input to feed-forward neural net classifiers (one for gender, one for ethnicity)

  - 3 dense hidden layers with 300, 300, and 120 neurons

  - hidden layers followed by normalization, ReLU activation, then dropout

  - SoftMax output layer

- Output → predicted probabilities by gender and race/ethnicity
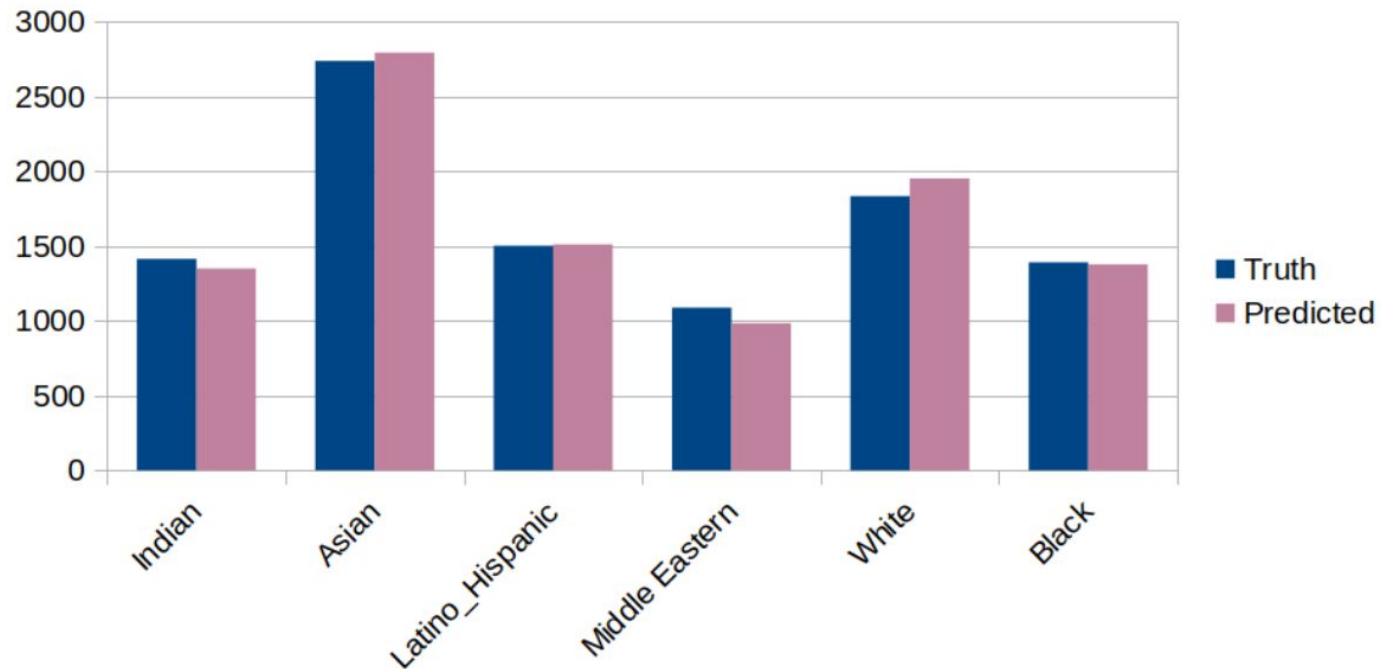
Classifier for Gender and Ethnicity

0=Asian(0.98),Male(1.00); 1=White(0.97),Male(1.00);
2=Black(0.99),Female(0.99); 3=Latino_Hispanic(0.55),Female(1.00);

0=Black(0.99),Female(0.99);

0=White(0.67),Female(1.00); 1=Middle Eastern(0.40),Female(1.00);
2=White(0.72),Female(1.00); 3=White(0.93),Male(1.00);
4=White(0.79),Female(0.71); 5=White(0.98),Female(0.81);

0=White(0.42),Male(0.93); 1=White(0.73),Male(1.00);
2=Middle Eastern(0.44),Male(1.00); 3=Middle Eastern(0.73),Male(1.00);
4=Middle Eastern(0.67),Male(1.00); 5=Black(0.45),Male(1.00);
6=Middle Eastern(0.71),Male(0.97);

Classifier Reproduces Correct Class Shares

True/Predicted Counts in Test Set, by Gender

True/Predicted Counts in Test Set, by Race/Ethnicity

Analysis of News Web Sites

Ash, Durante, Grebenshikova, and Schwarz

The New York Times (left)

PLAY THE CROSSWORD

Account ⌄

The New York Times

Thursday, May 20, 2021
Today's Paper

☀ 17°C 23° 11°

Your Thursday Briefing
Here's what you need to know.

Listen to 'The Ezra Klein Show'
Nicole Hemmer on minoritarian politics and the G.O.P.

Listen to 'The Modern Love Podcast'
After 25 years, returning to the place she had left him.

New Pressures Push U.S. and Europe to Stop Israel-Gaza Conflict

- President Biden faces a leftward shift in his party. In Europe, Muslim migration, terrorism fears and populist politics make diplomacy more urgent.
- Israeli officials said a cease-fire agreement with Hamas would most likely be reached within the next two days.

Displaced Palestinian families take refuge in a school in Gaza City. Hosam Salem for The New York Times

The Violence Is Stoking an 'Identity Crisis' for Some Young American Jews

A new generation is confronting the region's longstanding conflict with very different pressures from their parents' and grandparents' generations.

'I Was Surrounded by Death': Gaza Father Is Rescued, Emerging to Grief

Riad Ishkontana said that when rescuers found him and his 7-year-old daughter, he awoke to a new life — one without his wife and four other children.

Sharply Divided House Backs Jan. 6 Commission, but Senate Path Dims

The vote was a victory for Democrats, who were joined by 35 Republicans in pushing for a full accounting of the deadly riot.

Speaker Nancy Pelosi told reporters on Wednesday that Republicans seemed to be "afraid of the truth." Stefani Reynolds for The New York Times

Fed officials tiptoed toward a conversation about dialing back some support for the economy.
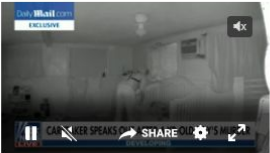
Opinion

Nicholas Kristof

The 'Unshakable' Bonds of Friendship With Israel Are Shaking

Israel might 'win' in Gaza, and lose America.

Charles M. Blow

History Can Be Erased. It Often Has Been.

I see Republicans trying to alter our perception of the insurrection at the U.S. Capitol. I do not take that lightly.

Senators introduced a bipartisan bill to significantly overhaul the U.S. Postal Service.

Fox News (right)

FOX NEWS
4:04 AM

Hot Topics　LIVE UPDATES: More Israeli Gaza airstrikes as Hamas predicts ceasefire

Markets　S&P 500　-0.09%

DailyMail.com EXCLUSIVE

ON NOW
Fox & Friends First

Exclusive Clips

GUTFELD
Gutfeld reacts to NYC Pride parade banning police

THE INGRAHAM ANGLE
Ingraham: Democrats believe 'America isn't exceptional, it's systemically racist'

HANNITY
Hannity sounds off on Fauci's latest COVID mixed messaging

TUCKER CARLSON TONIGHT
Tucker: Equity is racism

NO CONSPIRACY THEORY

House Intel Republicans tout proof of COVID's origins in Wuhan lab leak
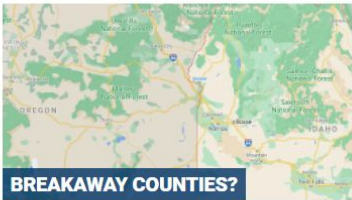
Pompeo calls on China to release evidence disproving Wuhan lab leak theory: 'Come on, bring it'

Ex-New York Times health reporter: 'Lab leak' coronavirus theory looking stronger
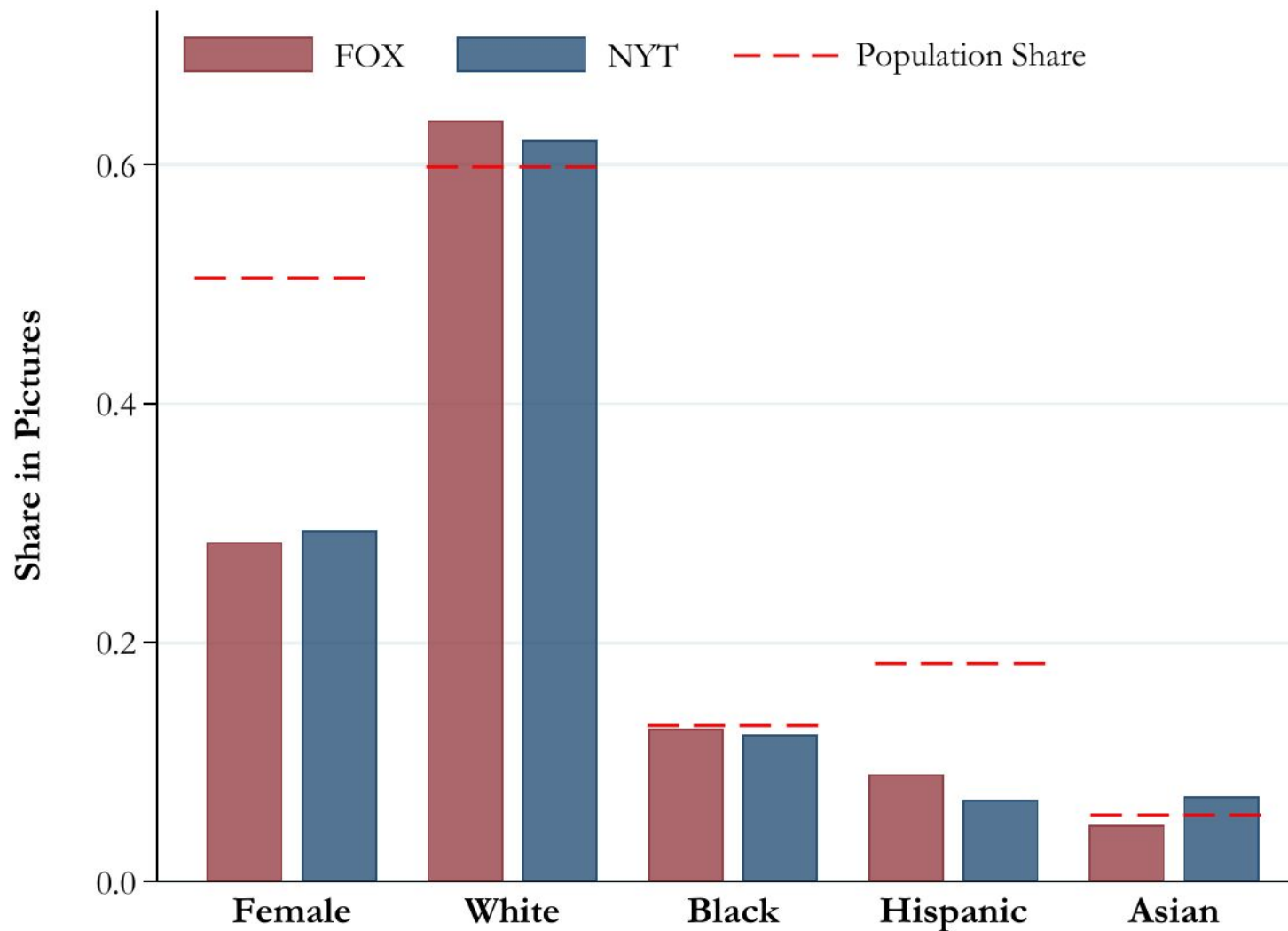
WHERE'S WHITMER?

GOP looks to track Dem's movements after recent Fla. trip

BREAKAWAY COUNTIES?

Portion of blue state votes in hopes of joining red state neighbor

Representation in News Web Sites

Notes: This figure shows the shares of different gender and ethnic groups in the pictures of Fox and NYT. The red lines indicate the share of each group in the U.S. population.

Exploring Specific Stereotypes

- Based on previous literature (e.g. Gilens 2009), we explored several specific, policy-relevant topics, and their association with visual identity representation:
  - crime
  - immigration
  - poverty
  - science
- for each of these, let w_i be the log number of mentions in article i
  - use lexicon of closely related words

- We estimate

$$y_i = \alpha_i + \beta_1 w_i \times NYT_i + \beta_2 w_i \times Fox_i + \varepsilon_i$$

  - y_i = visual identity share
  - α_i includes fixed effects (outlet-section and year)
  - NYT_i, Fox_i = indicators for NYT or Fox News.
  - standard errors clustered by outlet-section
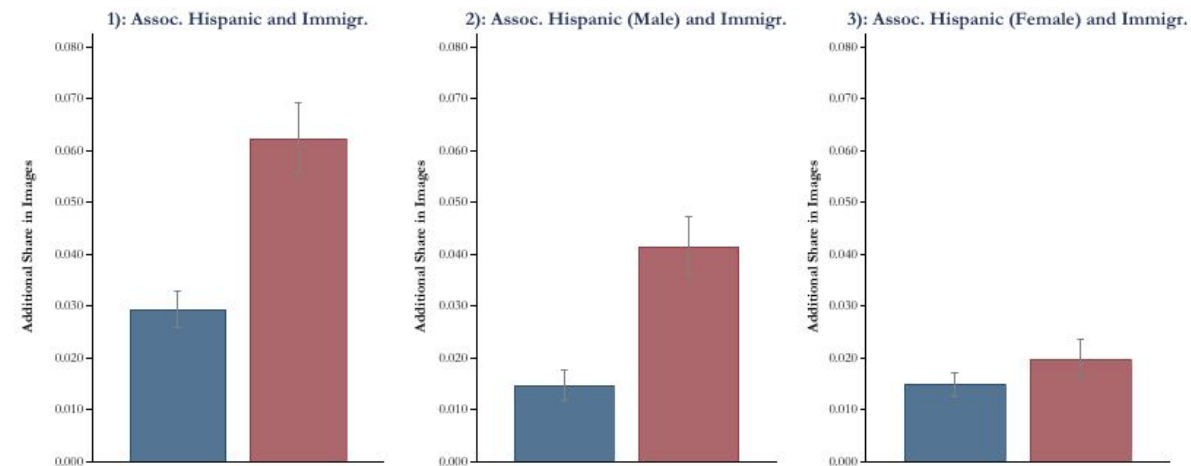
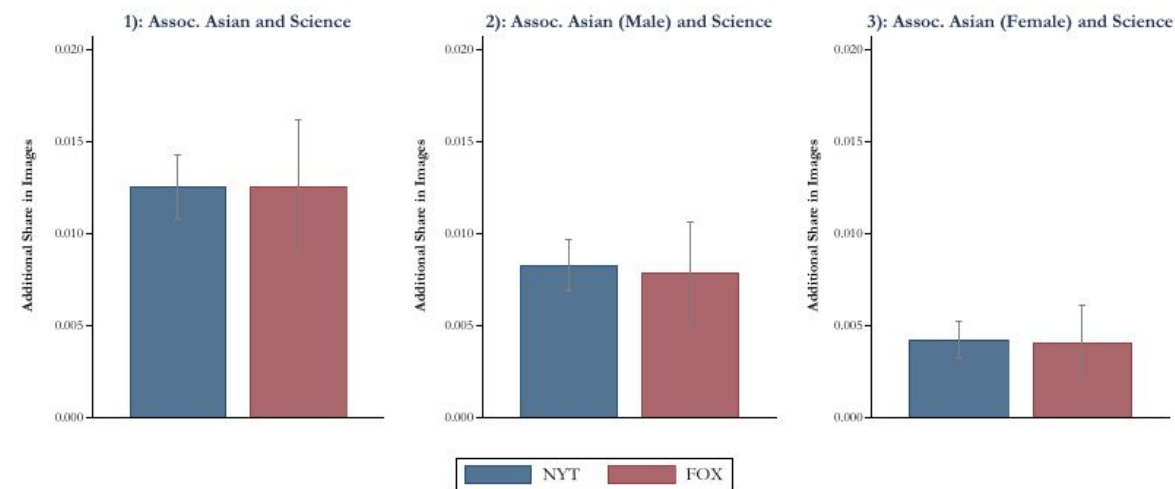# Topics and Visual Characteristics

(a) BLACKS AND CRIME

1): Assoc. Black and Crime

2): Assoc. Black (Male) and Crime

3): Assoc. Black (Female) and Crime

(b) BLACKS AND POVERTY

1): Assoc. Black and Poverty

2): Assoc. Black (Male) and Poverty
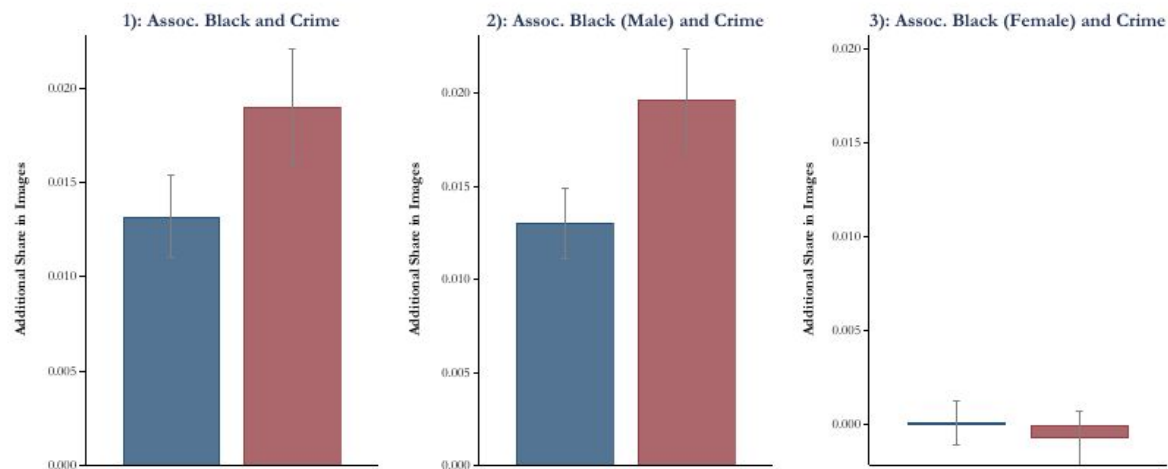
3): Assoc. Black (Female) and Poverty
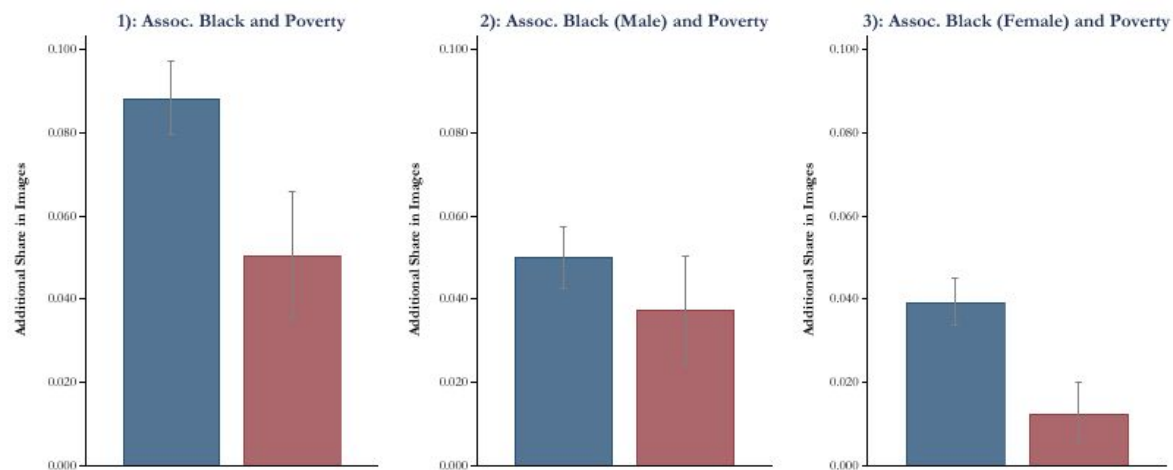
(c) HISPANICS AND IMMIGRATION

1): Assoc. Hispanic and Immigr.

2): Assoc. Hispanic (Male) and Immigr.

3): Assoc. Hispanic (Female) and Immigr.

(d) ASIANS AND SCIENCE

1): Assoc. Asian and Science

2): Assoc. Asian (Male) and Science

3): Assoc. Asian (Female) and Science

NYT   FOX

Effect of Editor/Author Identity on Visual Representation

- Hand-code gender and ethnicity of NYT editors.
- Machine-code author gender and ethnicity from byline, based on distinctive first names.
  (doesn't work for distinctive black names (e.g. "Jamal"), as they are very rare among NYT authors.)

- Estimate

$$y_i = \alpha_i + \beta D_i + \varepsilon_i$$

  - y_i = visual identity share
  - α_i includes fixed effects (section and month)
  - D_i =
    - share of editors in a section with identity characteristic.
    - indicator for for article author identity characteristic.
  - standard errors clustered by outlet-section

New Producer Representation and Visual Representation

Panel A: Identity of Author — Panel B: Share of Editors
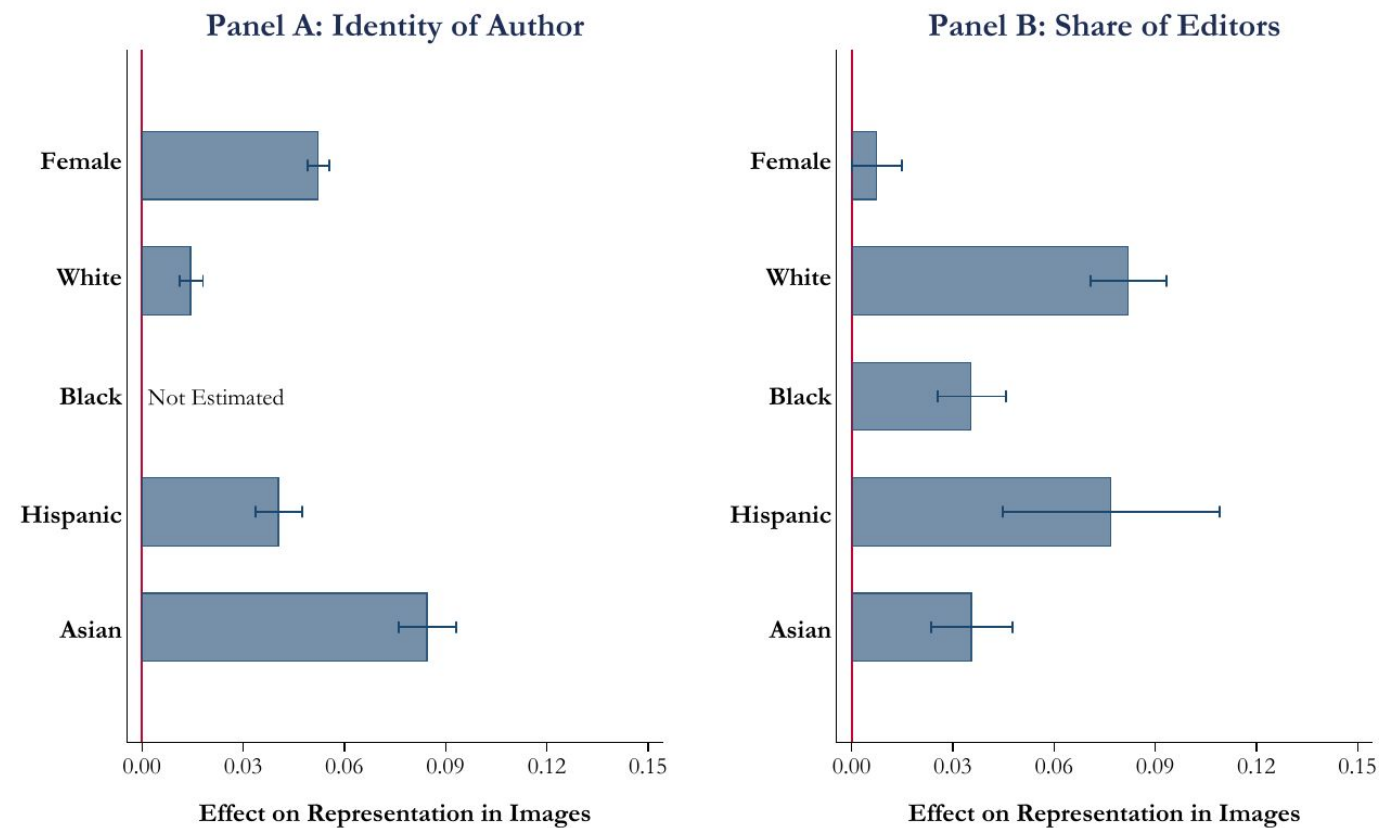
Notes: This table reports the estimates of Equation (3) for NYT, in which the dependent variable is the image share of each identity group in an article. The independent variable in Panel A is an indicator variable for the identity group of the article's author. In Panel B the independent variable is the share of editors in a newspaper section of the respective identity group. All regressions control for section and month fixed effects. Author identity (Panel A) is coded from the name, which does not provide a reliable signal of Black names; hence that estimate is not included. The whiskers indicate 95% confidence intervals.

Analysis of Politician Facebook Posts

Ash, Durante, and Schwarz
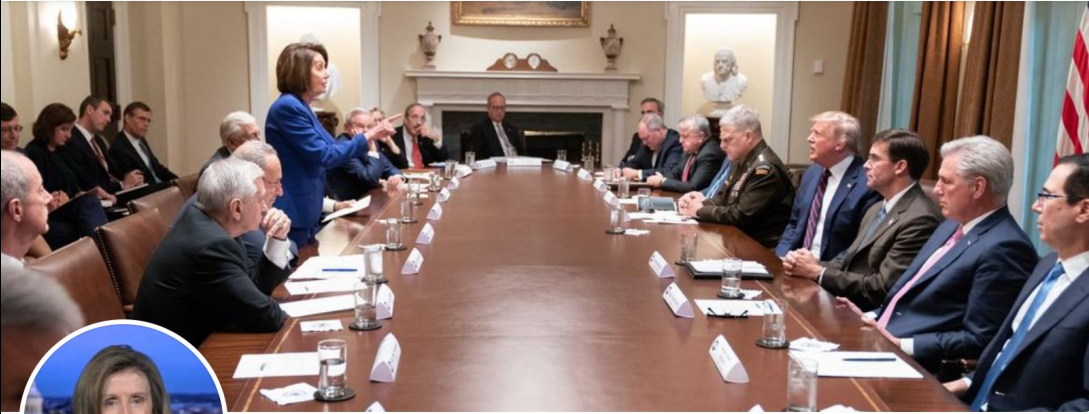
Nancy Pelosi ✓

533K followers • 19 following

Posts   About   Photos   Videos   ...

Intro

This account is managed by Speaker Pelosi's political team. Posts from Nancy are signed -NP.

ⓘ Page · Politician

🔒 NANCY PELOSI FOR CONGRESS is responsible for this Page

📍 San Francisco, CA, United States, California

🌐 store.pelosiforcongress.org

Photos   See all photos

Nancy Pelosi ✓
June 17 at 7:09 AM · 🌐

Congratulations Warriors — CHAMPIONS on and off the court!
#DubNation

👍❤️ 4.6K   470 Comments 337 Shares

Senator Mitch McConnell ✓

145K followers • 1 following

Posts   About   Photos   Videos   ...

Intro

Mitch McConnell is the senior Senator of Kentucky. He also serves as the U.S. Senate Republican Lead

ⓘ Page · Government Official

📞 +1 202-224-2541

🌐 mcconnell.senate.gov

Photos   See all photos

Senator Mitch McConnell ✓
May 30 at 10:32 PM · Instagram · 🌐

This Memorial Day we remember the brave men and women in uniform who laid down their lives in defense for our nation. Let us carry their memory close today and honor their sacrifice.

👍❤️🤗 194   559 Comments 11 Shares

👍 Like   💬 Comment   ↗ Share

Most relevant ▾

Edward DeMarco
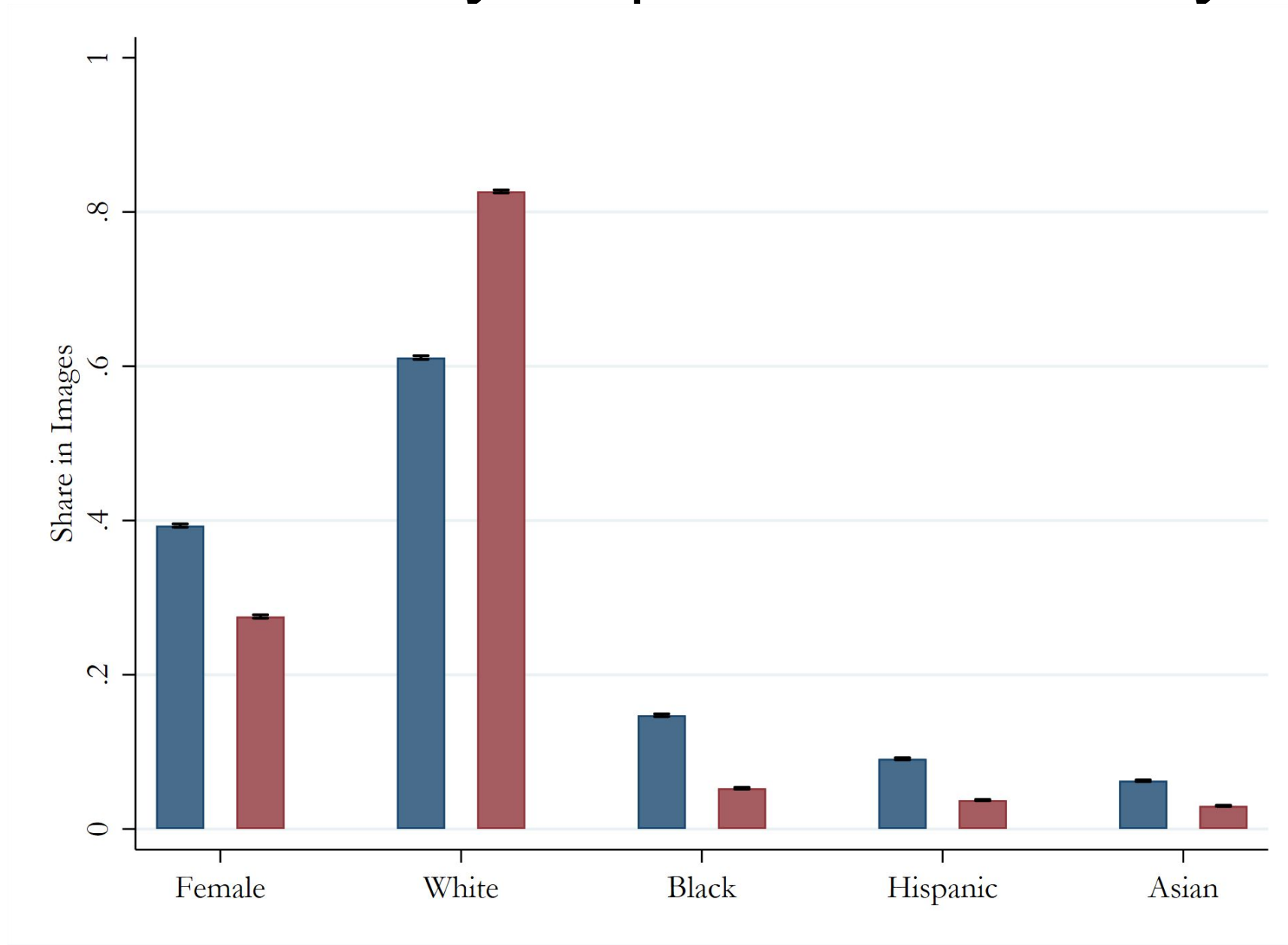Let's also remember the brave parents who tried to rescue their children at the Robb school
2w Edited   😡❤️ 57
↳ 1 Reply
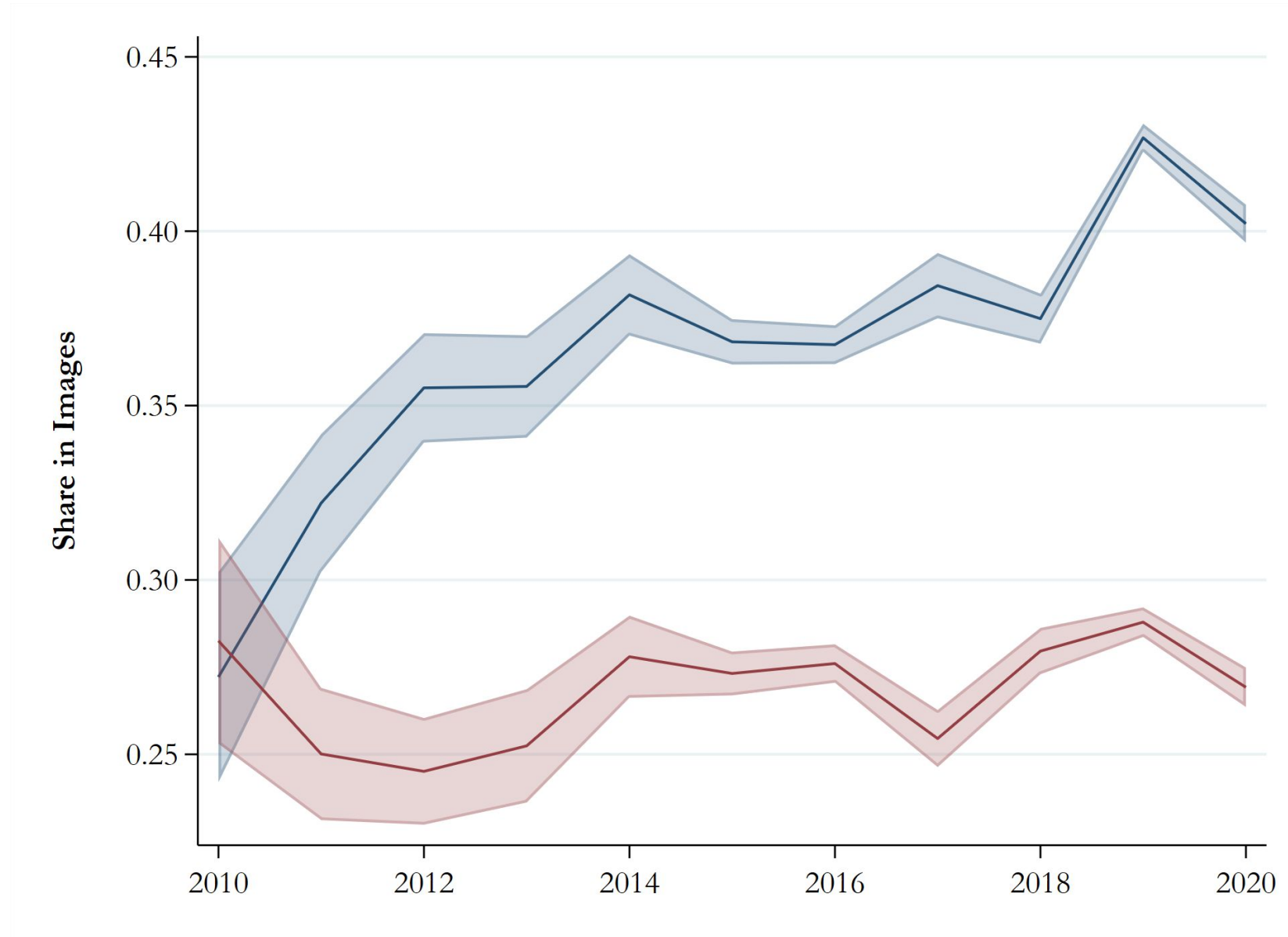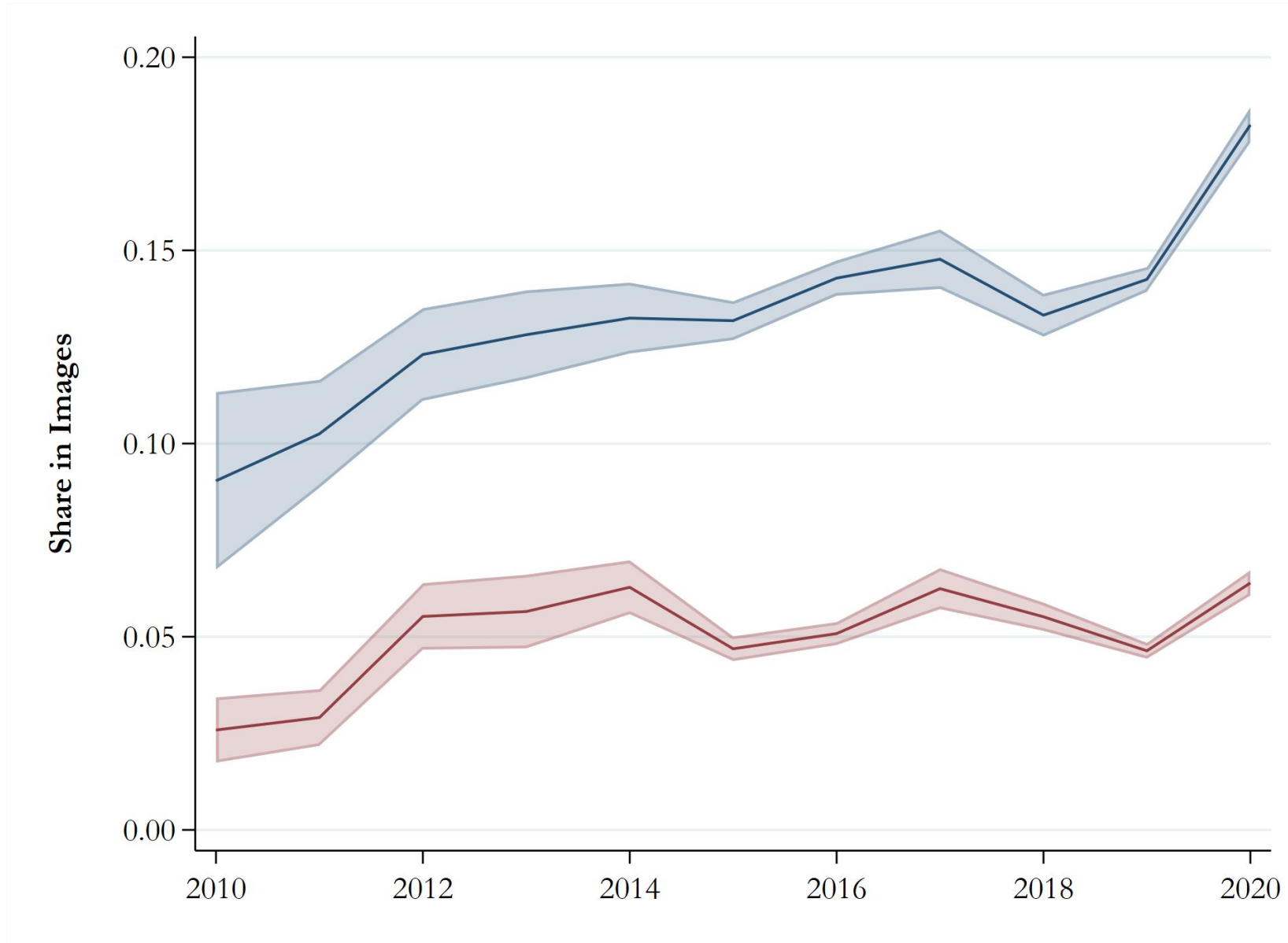
View more comments   1 of 214
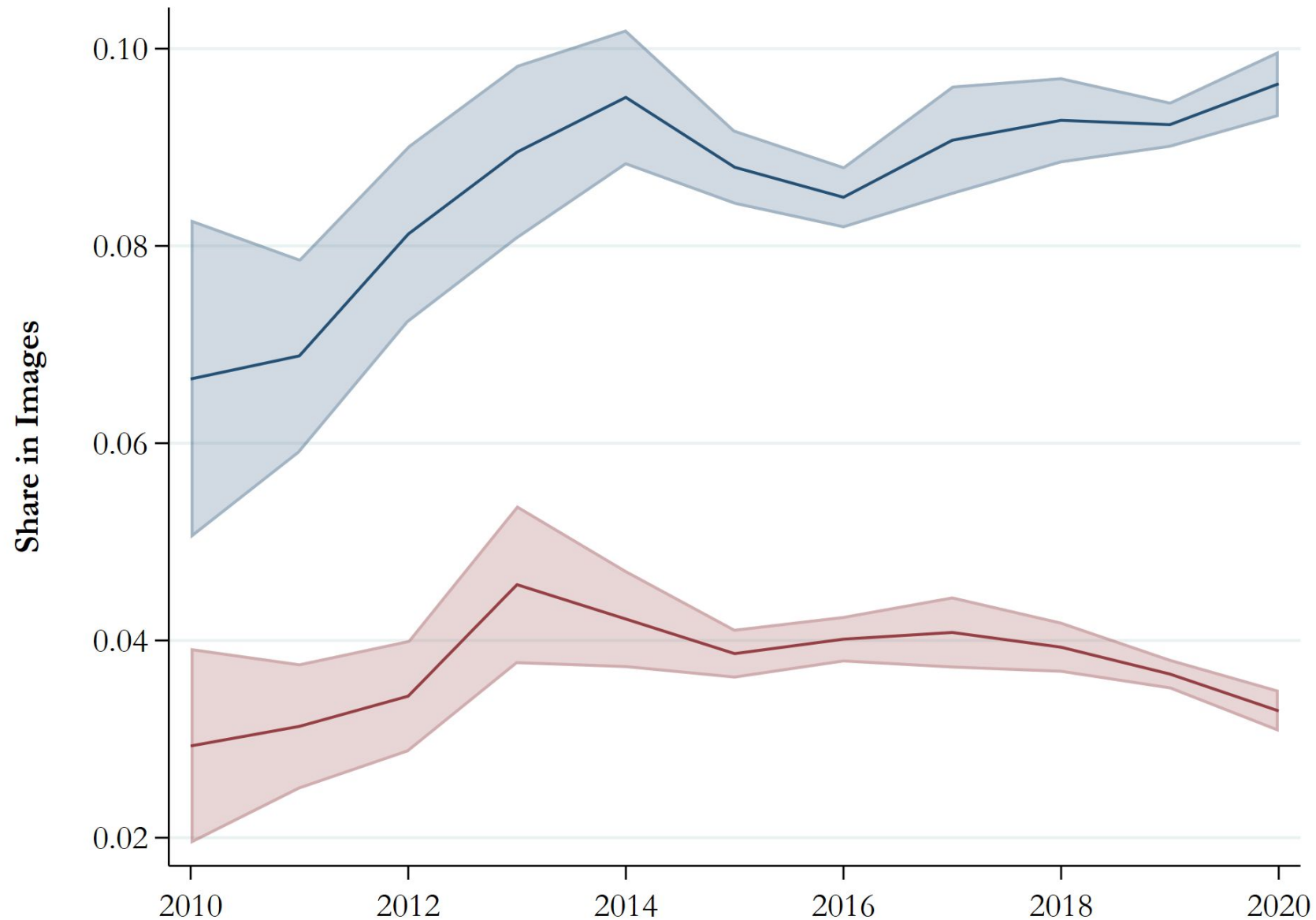
Facebook Identity Representation, By Party

Female Representation Over Time, By Party
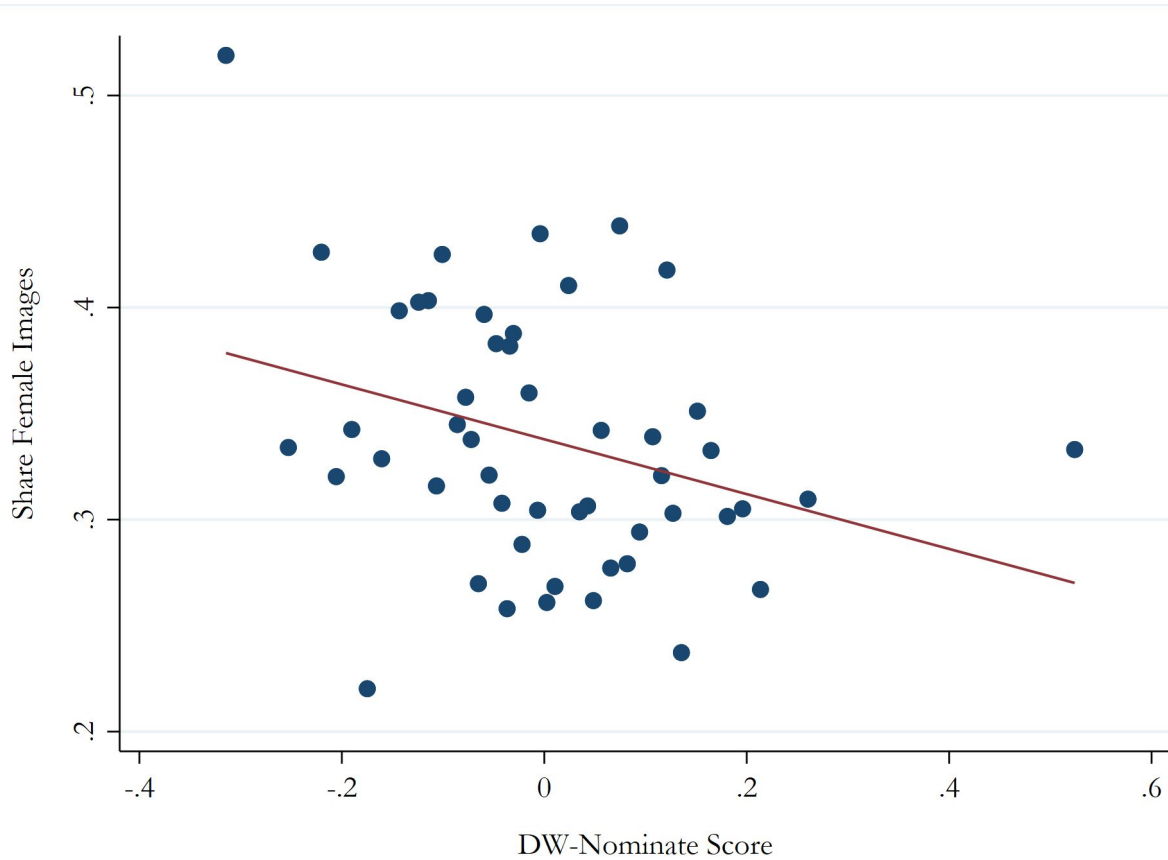
Black Representation Over Time, By Party
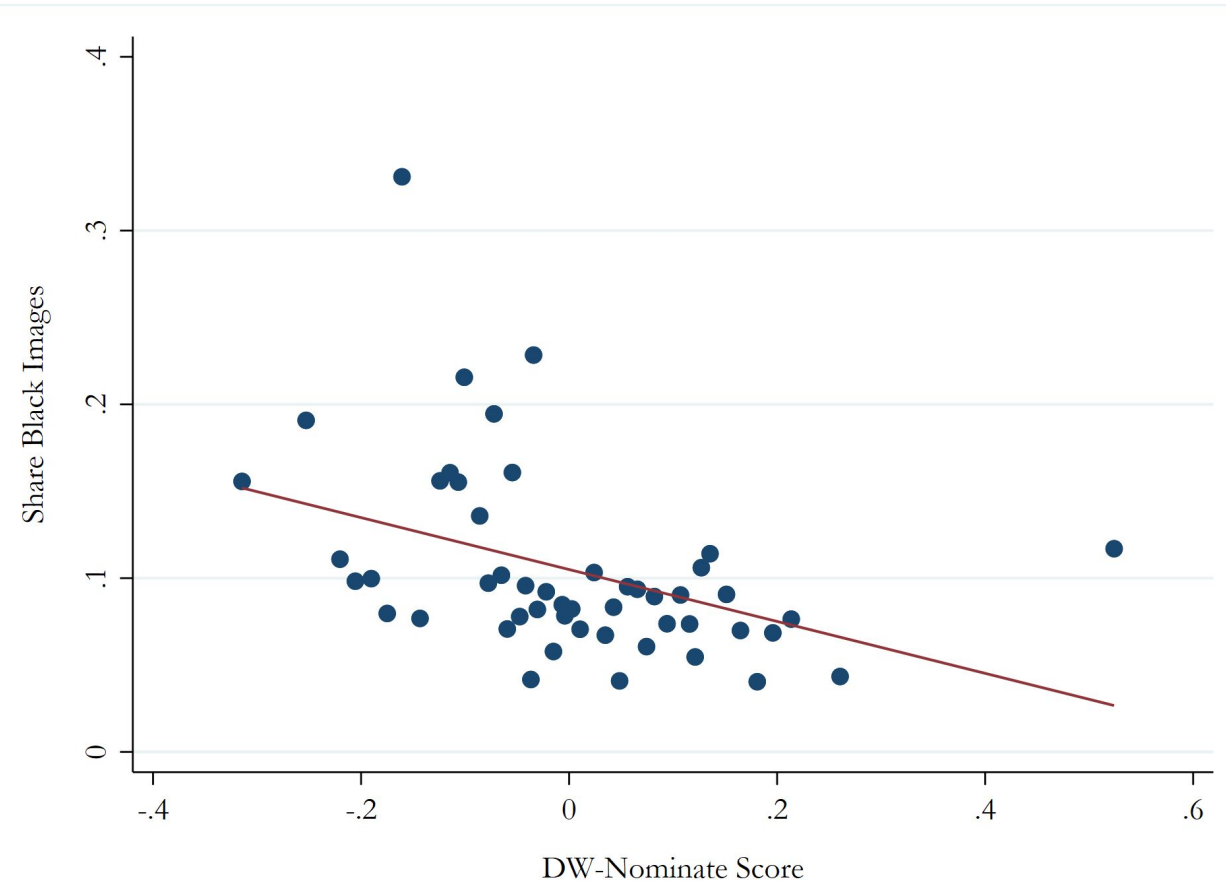
Hispanic Representation Over Time, By Party

Representation and Ideology (DW–NOMINATE)

Controlling for Party FE

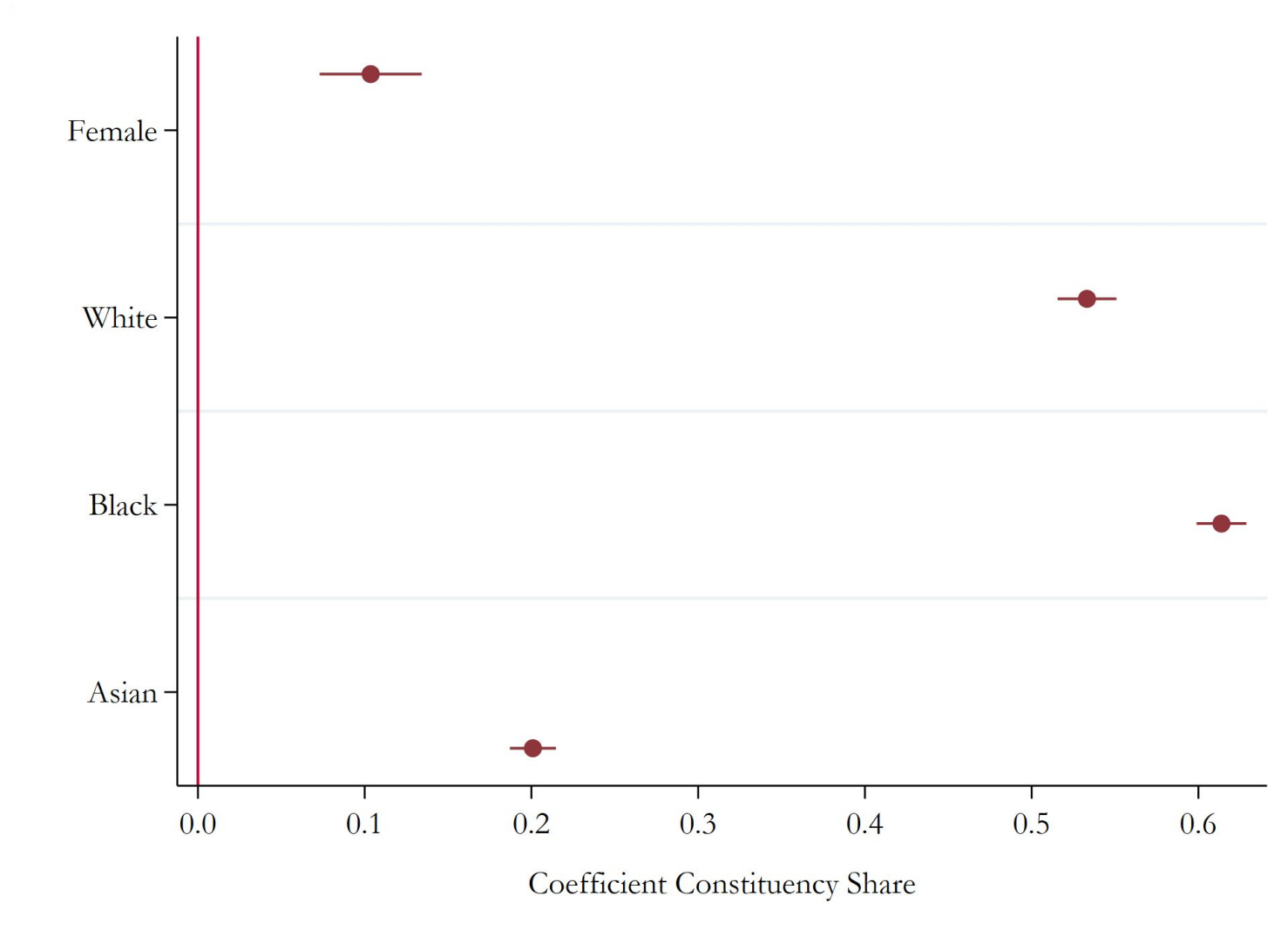Female Representation vs. Conservatism of Roll Call Votes

Black Representation vs. Conservatism of Roll Call Votes

Constituency Representation and Visual Representation

Controlling for Party and Time FE

Conclusion

Summary

First analysis of gender and ethnicity stereotypes in news media and identity politics on social media.

1. **Measurement**
   - New large-scale databases of news articles and social media posts with images and metadata
   - Validated image-based measures of gender/ethnicity representation
2. **Representation:**
   - men and Whites are over-represented in images at the expense of women and Hispanics, who are under-represented
   - differences across news web sites and political parties
3. **Stereotypes:**
   - news images associate Blacks with crime, larger effect for Fox News
4. **News production:**
   - visual representation reflects author and editor identity.
5. **Identity Politics**
   - Democrats increasingly feature more women and minorities
   - Additional intensive-margin effect with roll-call ideology.
   - Constituency demographics related to representation

Next Steps on Facebook Posts

- Stereotypes analysis on Facebook posts
- Exogenous variation in district demographics due to redistricting
- Effect of electoral cycle / close elections
- Effect of Trumpism / China Shock
- Analyze user responses – eg likes, comments

Thanks!
Elliott Ash, ETH Zurich
ashe@ethz.ch