

News Sharing, Activation and Perceived Polarization on Social Media

Tiago Ventura

University of Maryland, College Park

Job Talk McCourt School - 11/05/2021

About me



Tiago Ventura

Researcher Civic Integrity at Twitter

PhD Candidate at University of Maryland, College Park

 @TiagoVentura_  TiagoVentura  venturat@umd.edu  <https://tiagoventura.rbind.io/>

Plans for the talk

Plans for the talk

CSS & Political Communication

- News Sharing, activation and perceived polarization on social media

Plans for the talk

CSS & Political Communication

- News Sharing, activation and perceived polarization on social media

Overview of my research agenda

- New Methods in CSS
- Political Communication
- Comparative Politics

News Sharing, Activation and Perceived Polarization on Social Media

Social media increases sorting



Echo Chambers



Reduce cross-cutting exposure



Polarization

Are social media echo-chambers real?

- Online Sorting is similar to offline levels (Gentzkow and Shapiro, 2011)
- Users' friendship networks are heterogeneous outside of politics (Bakshy et al. 2012; Barbéra et al., 2015; Wojcieszak and Mutz, 2009)
- Users' digital media diets are balanced, and strongly influenced by big reputable outlets (Guess 2021)
- Exposure to counter-attitudinal and uncivil content increases affective polarization (Bail et al., 2018, Banks et al., 2021; Suhay et al., 2018)

?

Most users are embedded in heterogenous online networks, yet widespread perceptions of polarization on social media are still strong among users, experts, and policymakers

Main contributions

Main contributions

Activation x Sorting: *Content on social media depends fundamentally on the users' decision to propagate messages.*

Main contributions

Activation x Sorting: *Content on social media depends fundamentally on the users' decision to propagate messages.*

Social media bubbles emerge from activation, even if the users' friendship networks are heterogenous

Modeling Sharing Behavior on Social Media Data

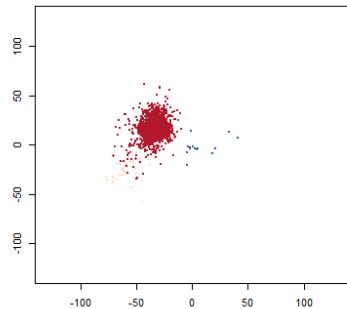
Activation Bubbles in Social Media

Brazil

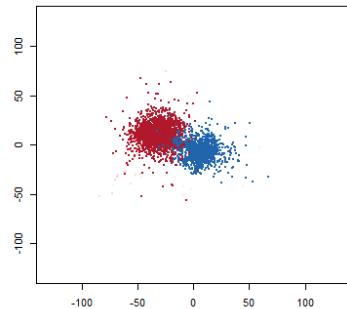
Argentina

United States

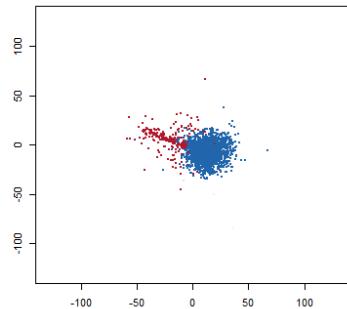
www.brasil247.com



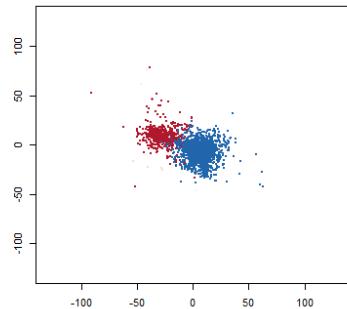
politica.estadao.com.br



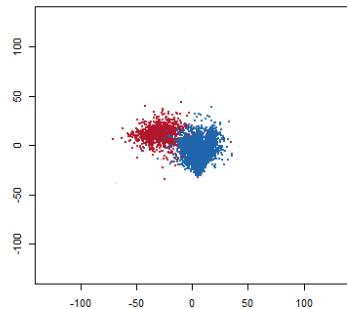
www.gazetadopovo.com.br



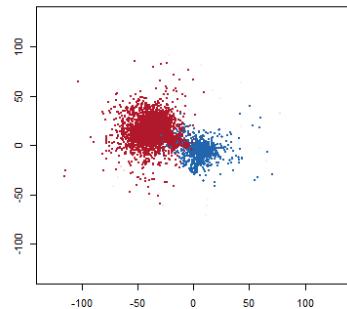
istoe.com.br



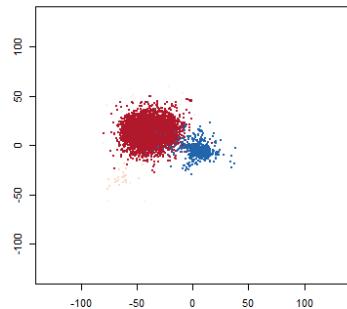
abr.ai



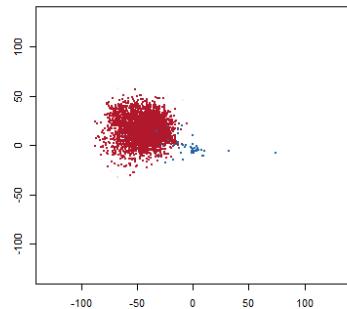
noticias.uol.com.br



aovivo.folha.uol.com.br



www.sensacionalista.com.br



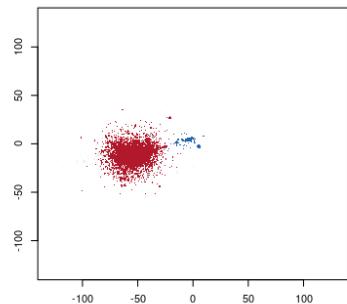
Activation Bubbles in Social Media

Brazil

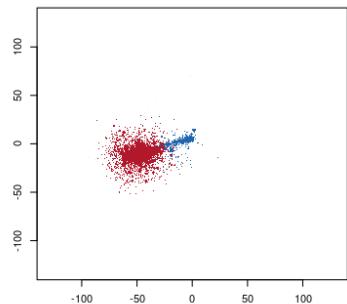
Argentina

United States

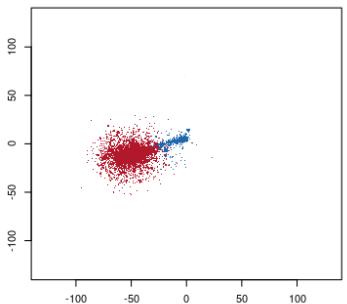
www.laizquierdadiario.com



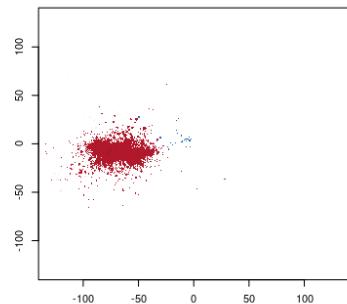
www.minutouno.com



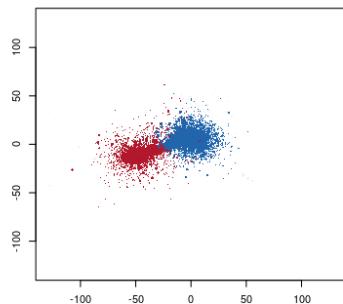
www.enorsai.com.ar



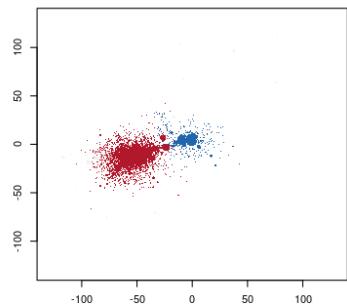
www.telesurtv.net



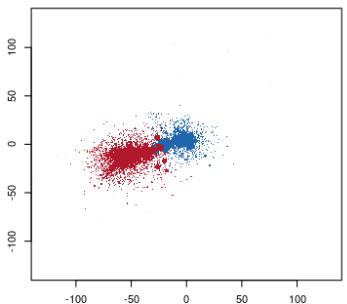
www.infonews.com



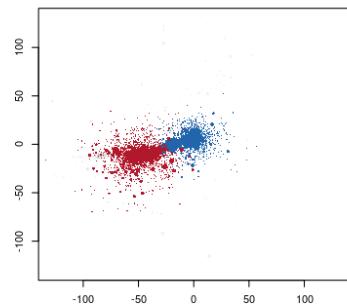
www.facebook.com



www.perfil.com



youtu.be



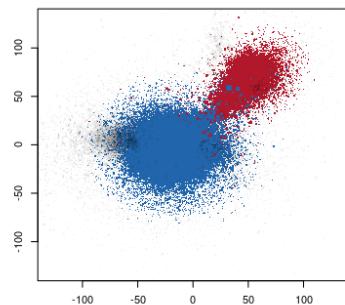
Activation Bubbles in Social Media

Brazil

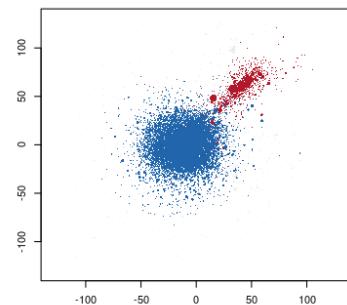
Argentina

United States

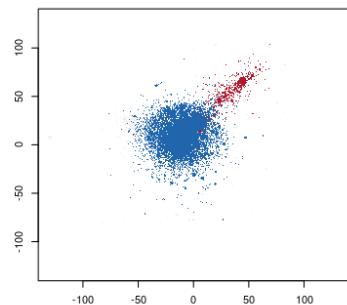
twitter.com



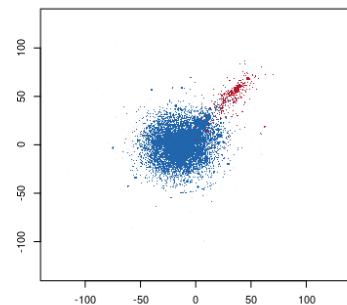
www.nytimes.com



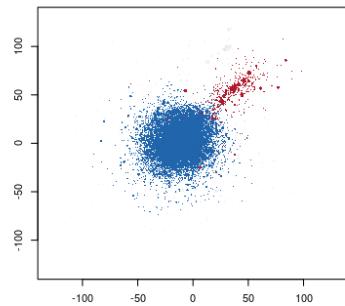
cnn.it



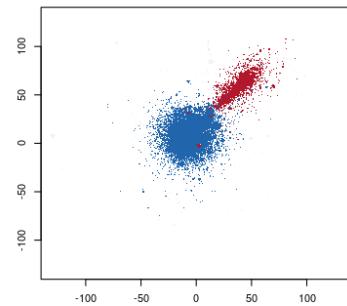
apnews.com



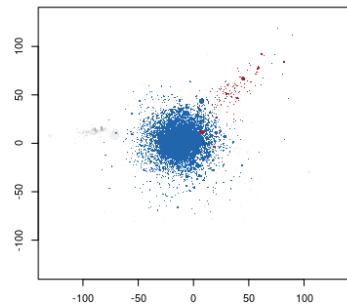
www.washingtonpost.com



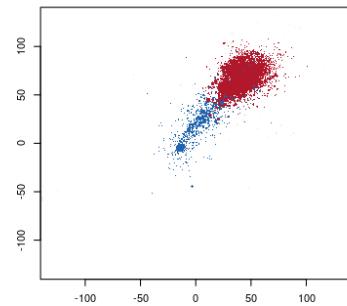
thehill.com



www.huffingtonpost.com

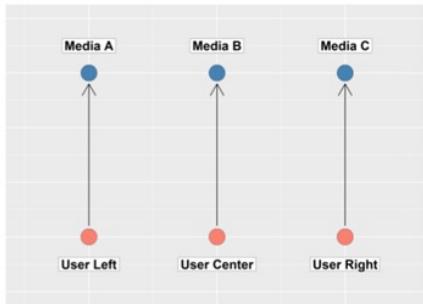


www.foxnews.com



News Sharing Model

Why do Users Share News on Social Media?



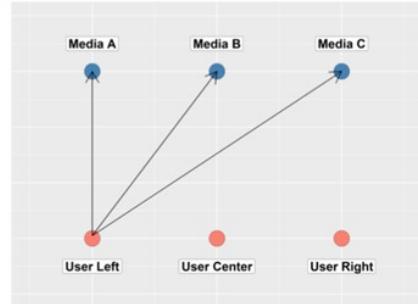
Ideology

$$U_{(ij)}^k = -\alpha_{q(i)}^k \left(x_i^k - \beta_j^k \right)^2 + R_{q(i)}^k N_{ij}^k + A_i^k + G_j^k$$

I R A

Table 1: Observational data with Users (row) and Media (columns)

$$\begin{array}{c} \leftarrow \\ \leftarrow \\ \leftarrow \end{array} \left(\begin{array}{ccccc} M_1 & M_2 & M_3 & \cdot & M_J \\ 0 & 2 & 0 & 4 & 1 \\ 5 & 8 & 2 & 6 & 4 \\ 10 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 9 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 5 & 0 & 1 & 7 \end{array} \right) \begin{array}{l} U_1 \\ U_2 \\ U_3 \\ U_4 \\ \cdot \\ U_I \end{array} \xrightarrow{\hspace{1cm}} \begin{array}{c} \downarrow \\ \downarrow \\ \downarrow \end{array}$$



Attention

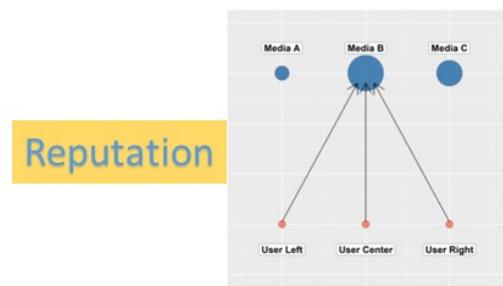


Figure 1: Effect of Reputation in a social media embeds. Users on the left, center, and right of the political spectrum embed content from Media B

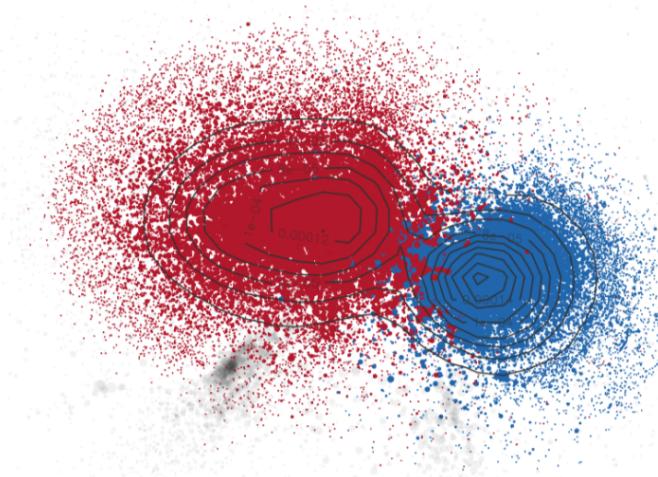
News Sharing Model

Table 1: Observational data with
Users u_i (row) and Media m_j (columns)

m_1	m_2	m_3	.	m_J	
0	2	0	4	1	u_1
5	8	2	6	4	u_2
10	5	0	0	0	u_3
0	0	0	5	9	u_4
.
0	5	0	1	7	u_I

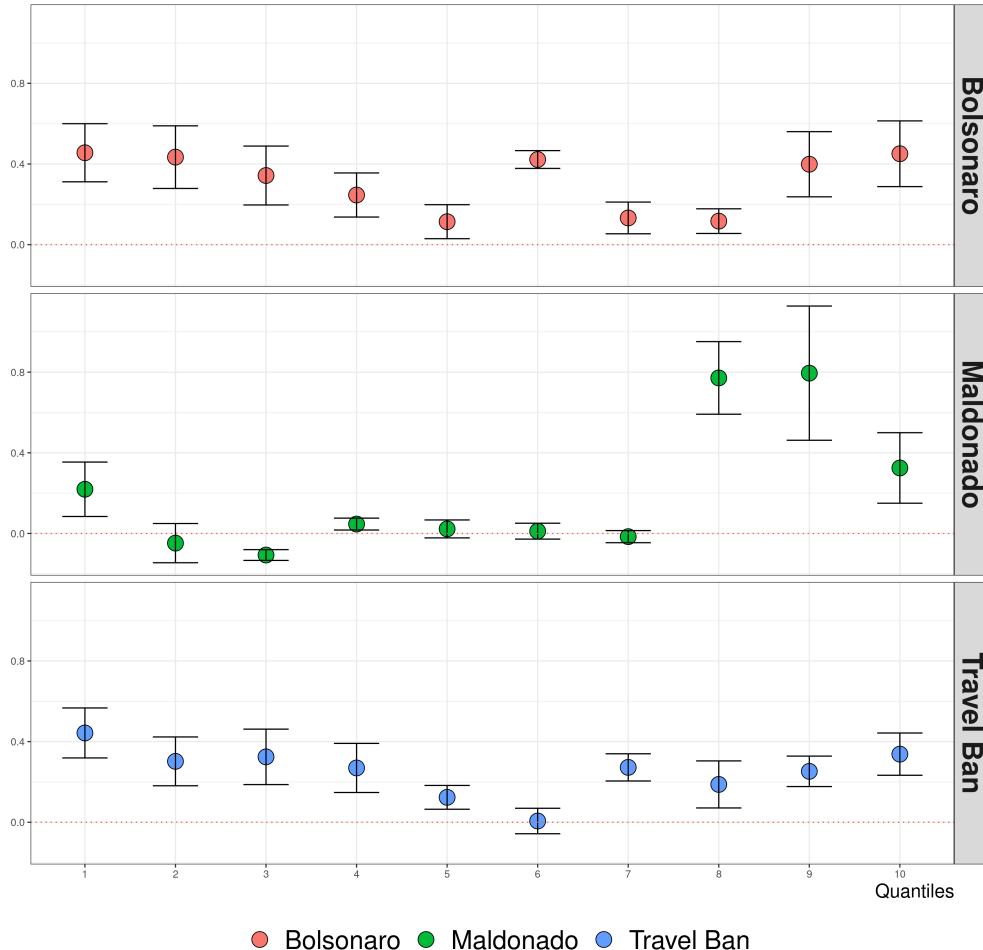
Bolsonaro Twitter Network

ENCG: 2.14



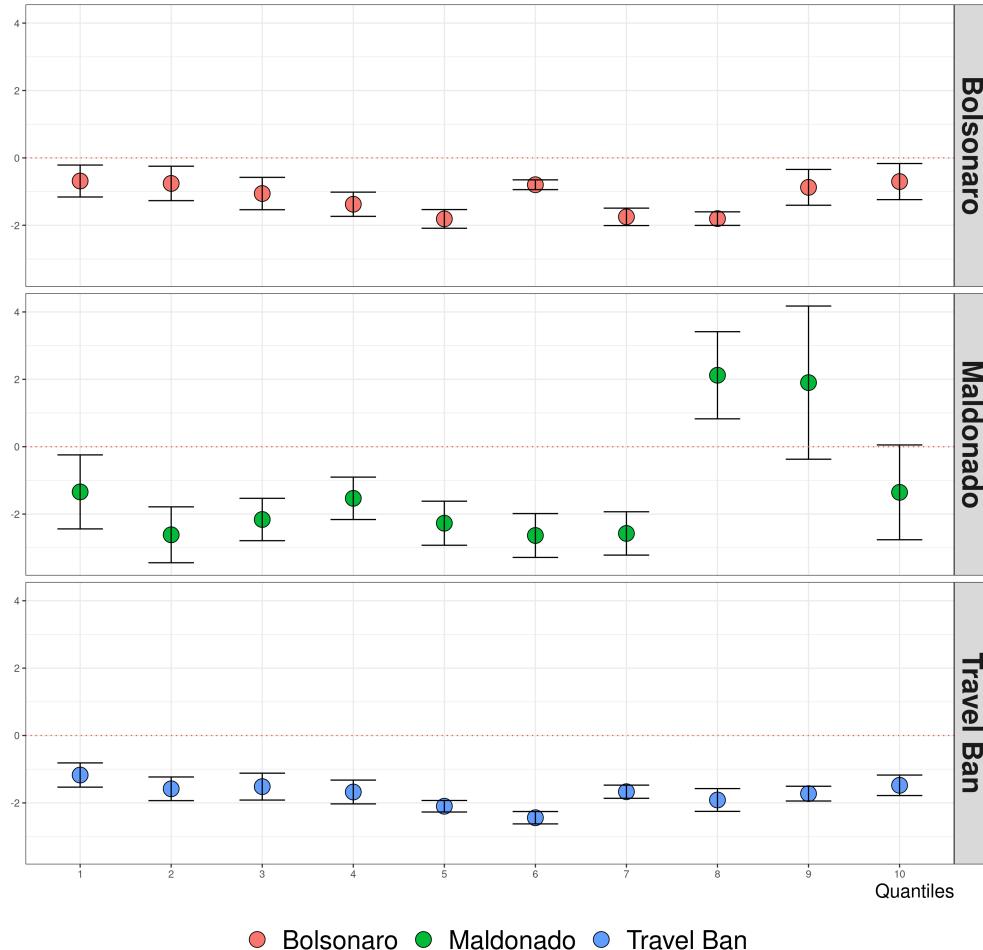
Results: Ideology

Ideology by quantile



Results: Attention

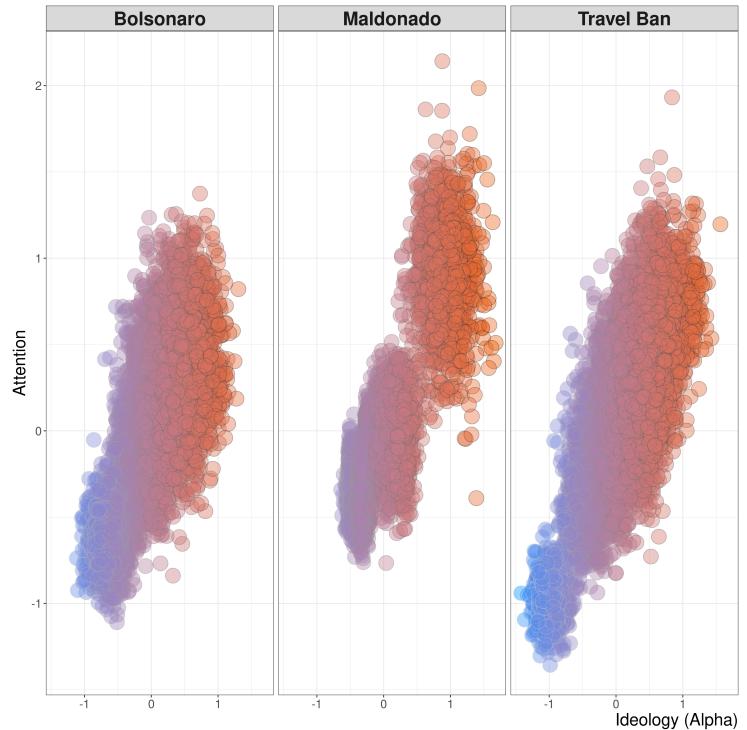
Attention by quantile



Conclusion: Attention and Ideology

Strong correlation between ideology and attention:

- More partisan users share more content
- Partisan content will be over-represented.
- Social media bubbles will emerge from activation, even if networks are heterogenous.



Modeling Social Media Activation with Image-Based Conjoint Experiments

Observational data cannot fully separate activation from sorting.

Solution: Image-Based Conjoint Experiments

Advantages:

- Flexible
- Equal probability for all the frames
- High ecological validity

Outcomes: Which message would you share? Which message are your friends more likely to share?

Image-Based Conjoint

Brazil

Argentina

Mexico

Endorsement →

Text →

Images →

Support →

Leia Atentamente os seguintes tuítes

Tweet 1

Folha de S.Paulo

"Assassino e Bandido. Esse tipo de gente que a turma dos direitos humanos protegem para incriminar policial e cidadão de bem. Não dá pra reduzir crime sendo bomzinho com a bandidagem"

Confira aqui a opinião exclusiva do Deputado Carlos Jordy.

Tweet 2

Folha de S.Paulo

O Deputado Marcelo Freixo destacou em entrevista: "Não podemos mais perder tempo com políticas ineficazes para reduzir o crime. É preciso ir na raiz do problema. Reduzir as desigualdades e criar oportunidades para os jovens pobres que estão morrendo no crime."

8:31 AM - 9 Aug 2019 5012 Retweets 15981 Likes

8:31 PM - 9 Aug 2019 4018 Retweets 10881 Likes

Qual dos dois tuítes é mais provável que VOCÊ COMPARTILHE na sua página?

Tuite 1

Tuite 2

Ambos

Nenhum

Image-Based Conjoint

Brazil

Argentina

Mexico

Endorsements



Pagina12
@Pagina12

Follow



LA NACIÓN
@LANACION

Follow

Images



Text

Alberto Fernández: "Es hora de terminar con la grieta y ponerse a trabajar para cuidar la salud de los Argentinos".

"Es hora de terminar con la grieta y ponerse a trabajar para cuidar la salud de los Argentinos", dice Alberto Fernández.

"Estamos así porque Macri prefirió financiar la salida de dólares en lugar de la inversión en hospitales", dice Alberto Fernández.

Alberto Fernández: "Estamos así porque Macri prefirió financiar la salida de dólares en lugar de la inversión en hospitales".

Support

2:45 PM - 18 Mar 2020

3 Retweets 2 Likes

1 3 z

2:45 PM - 18 Mar 2020

211 Retweets 312 Likes

1 211 312

Image-Based Conjoint

Brazil

Argentina

Mexico

Endorsements



La Jornada
@lajornadaonline



Reforma
@Reforma

Follow

v

Images



0:33 3.3K views La gente pueda regresar a las actividades que antes hacia cotidianamente.

Text

"Los niños soldados de Guerrero son una señal que en México la estrategia de seguridad de AMLO ha fallado", afirma Denise Dresser.

Afirmó Denise Dresser que "Los niños soldados en Guerrero son una señal que en México la estrategia de seguridad de AMLO ha fallado"

Jugar a la guerra: Los niños que perdieron a sus padres, ejecutados, quemados, masacrados, hoy se enlistan como soldados para defenderse.

¿Juegos peligrosos? Como entrena los niños que se enlistan como soldados para defenderse y defender a sus familias de los ataques narcos.

Support

2:45 PM - 18 Mar 2020

3 Retweets 2 Likes

1 3 2 1

2:45 PM - 18 Mar 2020

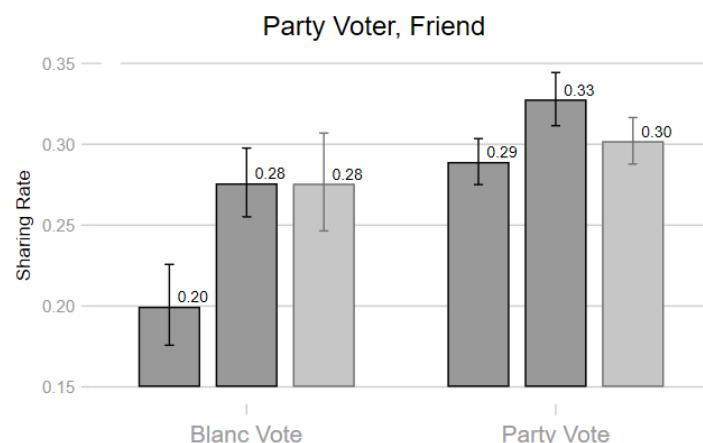
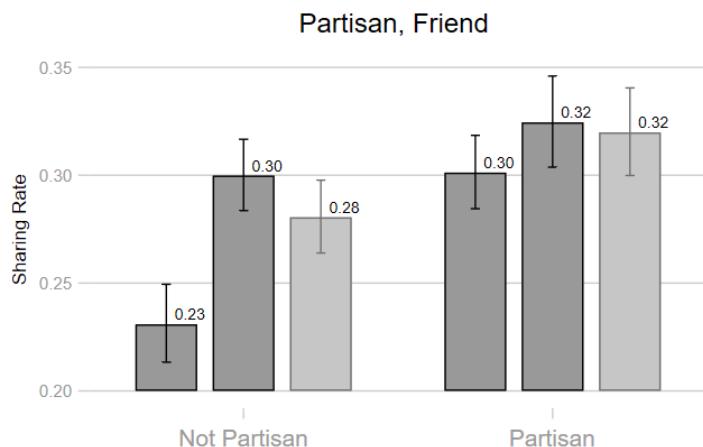
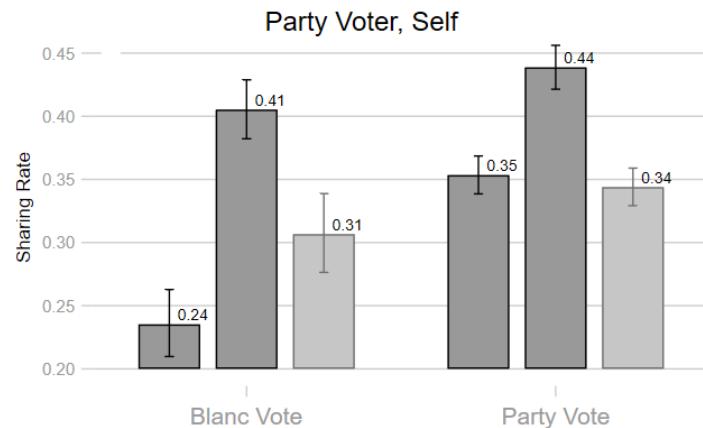
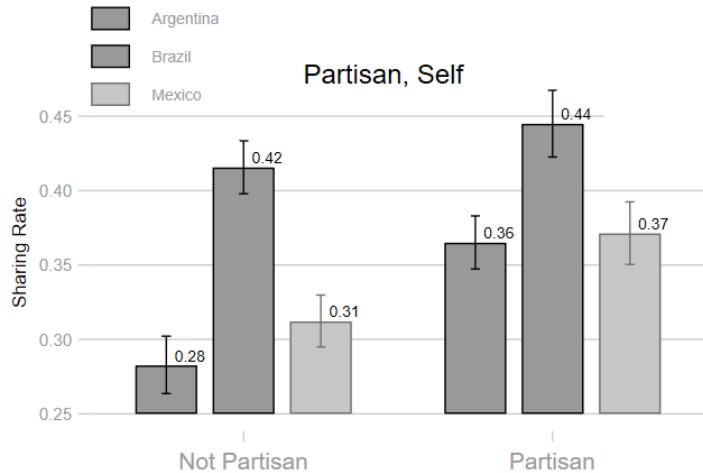
211 Retweets 312 Likes

1 211 312 1

Hypothesis

- H1: Partisan users will be unconditionally more issue motivated than non-partisan voters to share political content.
- H2: Users will share congruent content that aligns politically with the preferences of their co-partisans (in-group cognitive congruence)
- H3: Partisan users expect their friends to share more partisan frames

H1: Sharing-Rate



H2: Framing Activation

Brazil

Argentina

Mexico

Header:

Liberal Media
Conservative Media

Content:

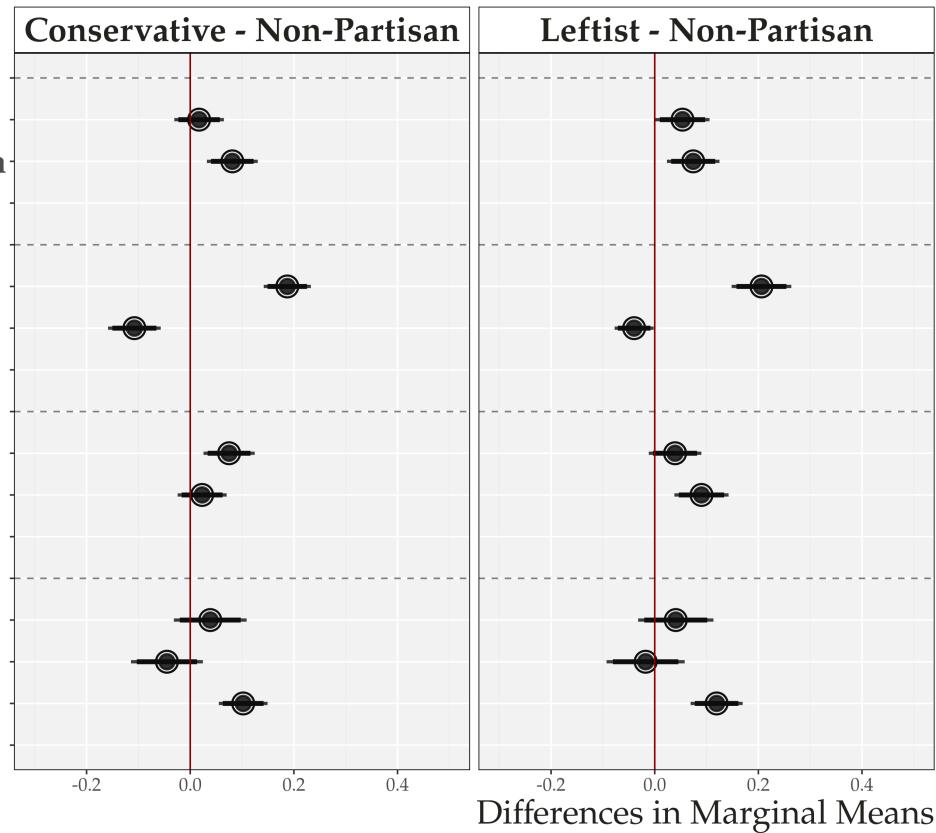
In-group
Out-group

Endorsement:

Law And Order
Civilian

Picture:

Military
School
Neutral



H2: Framing Activation

Brazil

Argentina

Mexico

Header:

Liberal Media
Conservative Media

Content:

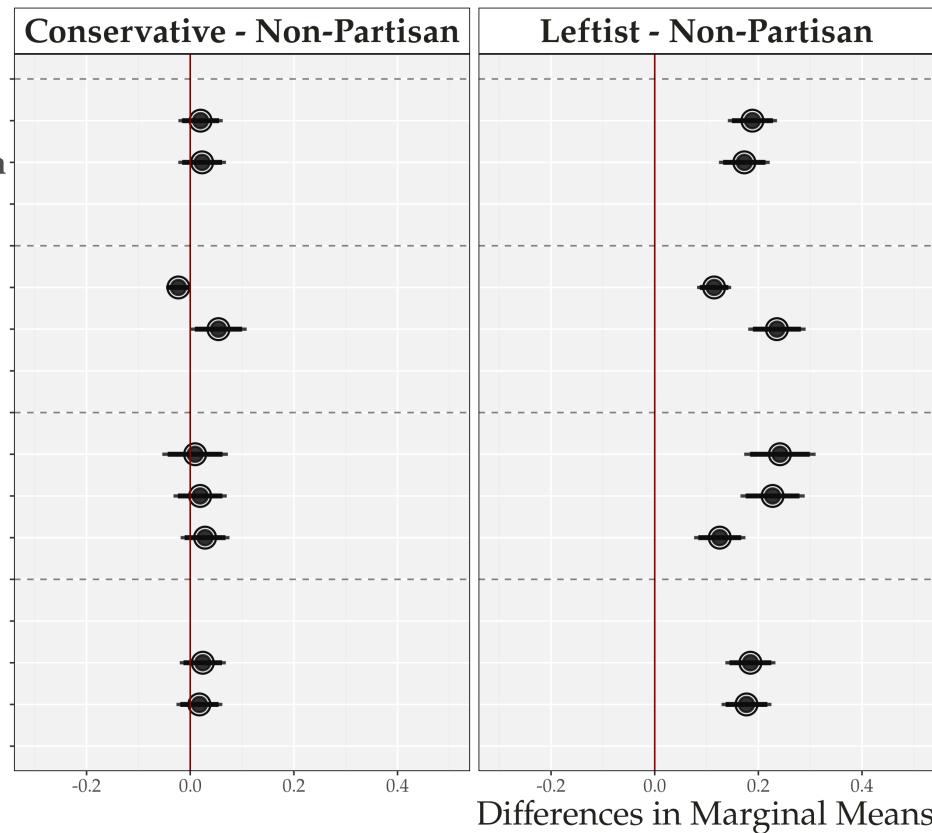
Politics-As-Normal
Cross-The-Isle

Image:

A. Fernandez
Cf. Kirchner
Neutral Image

Foot:

High Support
Low Support



H2: Framing Activation

Brazil

Argentina

Mexico

Header:

Liberal Media
Conservative Media

Content:

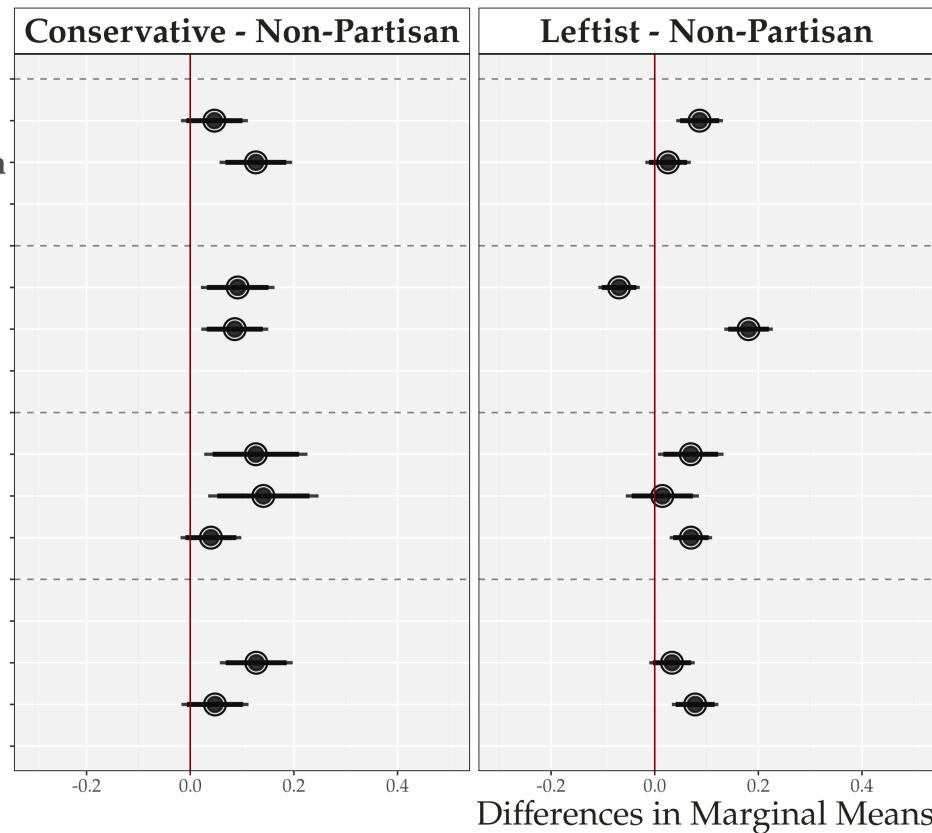
Anti Government
General Framing

Image:

Militarization
Kids
Neutral

Foot:

High Support
Low Support



H3: Would your friend share?

Brazil

Argentina

Mexico

Header:

Liberal Media
Conservative Media

Content:

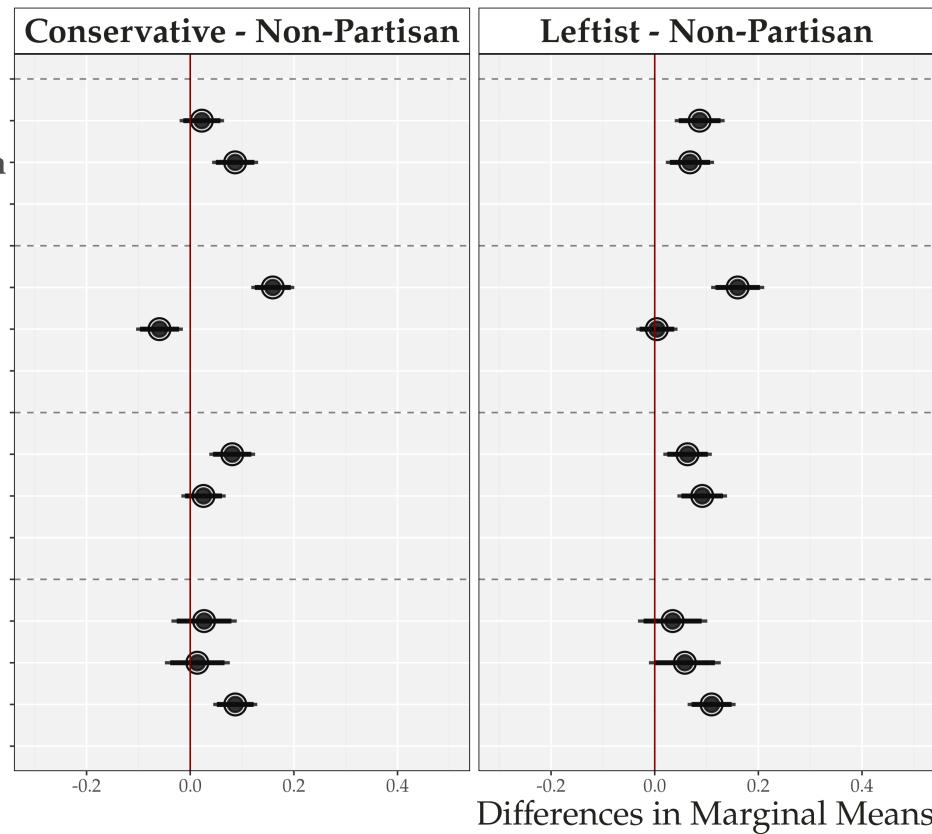
In-group
Out-group

Endorsement:

Law And Order
Civilian

Picture:

Military
School
Neutral



H3: Would your friend share?

Brazil

Argentina

Mexico

Header:

Liberal Media
Conservative Media

Content:

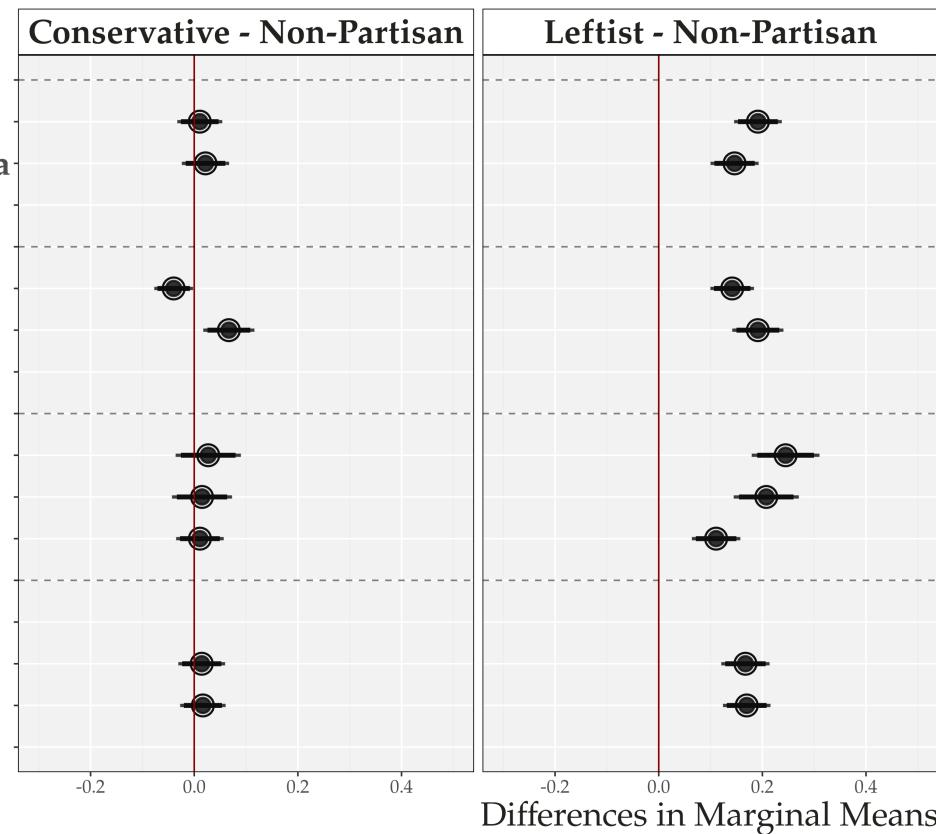
Politics-As-Normal
Cross-The-Isle

Image:

A. Fernandez
Cf. Kirchner
Neutral Image

Foot:

High Support
Low Support



H3: Would your friend share?

Brazil Argentina

Mexico

Header:

Liberal Media

Conservative Media

Content:

Anti Government
General Framing

Image:

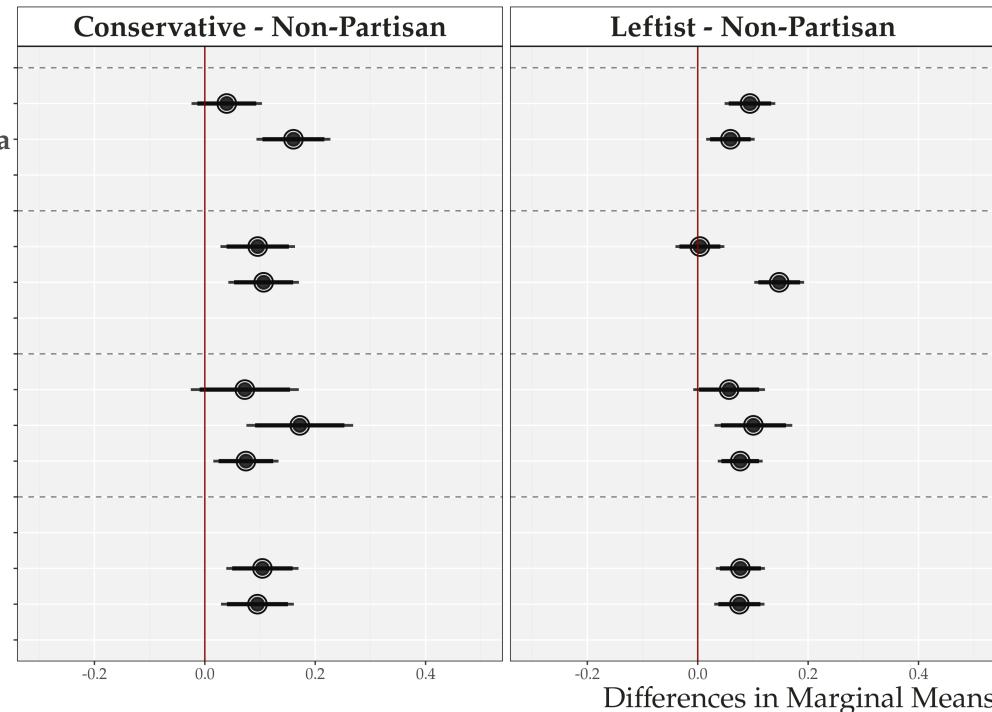
Militarization

Kids

Neutral

Foot:

High Support
Low Support



Main Take Aways

Main Take Aways

*Clear experimental evidence that a social network with input frame elements with **random uniform probabilities** will output local frames that will **over-represent** the preferences of partisan respondents*

Main Take Aways

*Clear experimental evidence that a social network with input frame elements with **random uniform probabilities** will output local frames that will **over-represent** the preferences of partisan respondents*

*As in observational data, bubbles emerge from **propagation** formed as a consequence of statistical correlation between **attention and ideology***

Main Take Aways

Clear experimental evidence that a social network with input frame elements with random uniform probabilities will output local frames that will over-represent the preferences of partisan respondents

As in observational data, bubbles emerge from propagation formed as a consequence of statistical correlation between attention and ideology

Affordances from social media activation creates an ecosystem where even non-sorted users are exposed to more ideological content on social media

Policy Implications

Slowing Sharing Behavior: Social media circuit breakers

Double click retweet policy during the US Presidential Election

Overall Research Agenda

And its CSS applications

Overall Research Agenda

Political
Communication

CSS Methods

Comparative
Politics

Streaming Chats:

- Political Effects of Streaming Chats: Field Digital Experiment
- Toxicity and Streaming Chats on Facebook: Develop scrappers to collect more 100k chatboxes' comments from Facebook

Overall Research Agenda

Political
Communication

CSS Methods

Comparative
Politics

- Strategies to work with Big Social Media Network Data: **The Path-Weighted Regression Model**
- Image-Based Behavioral Experiments: **voting, preferences for police allocation, social media activation**
- Natural Language Processing: Word Embeddings for Ideal-Points Estimation.

Overall Research Agenda

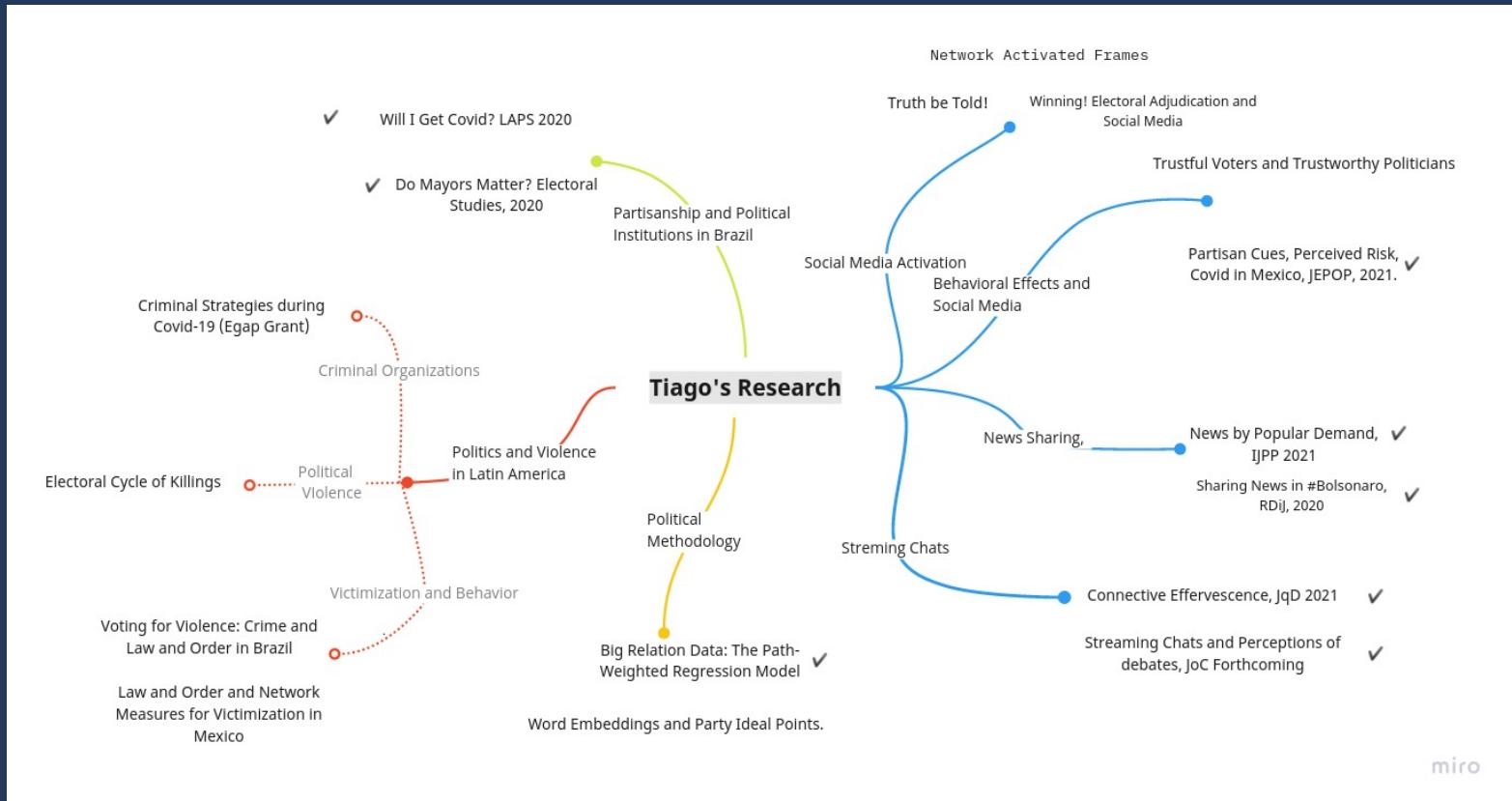
Political
Communication

CSS Methods

Comparative
Politics

As a comparativist, most of my work focuses on issues of **criminal violence, security policies, and preferences for law and order candidates** in Latin America.

- **A Network Approach to Crime Victimization:** Novel network strategies to model overdispersion of crime.
- **Who owns the issue of security in fragmented democracies?:** 140k Legislative Sppeches (Audio + Text) - Topic Models + NLP.
- **Criminal Organizations and provision of public goods during Covid-19 pandemic:** Facebook Ads to target hard-to-reach population



Thank you

Path-Weighted Regression Model (With J. Timoneda and E. Calvo)

Problem

Results

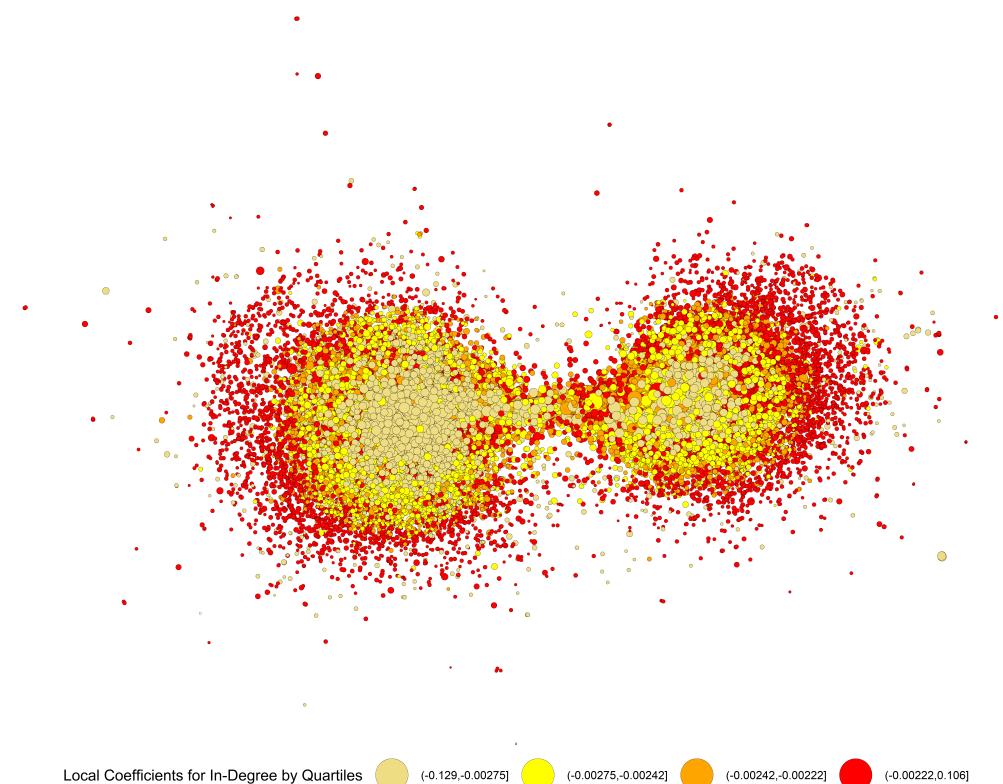
How to model network dependency with Big Data?

- Most network models are too computationally intensive for large networks.
- Solution: Model Locally + Weighted by Closest Paths
- Extension of Geographically Weighted Regression Model
- Use Cross-Validation to Determine Optimal Number of Paths
- Allows to model heterogeneity on networks and easy to parallelize.

Path-Weighted Regression Model (With J. Timoneda and E. Calvo)

Problem

Results

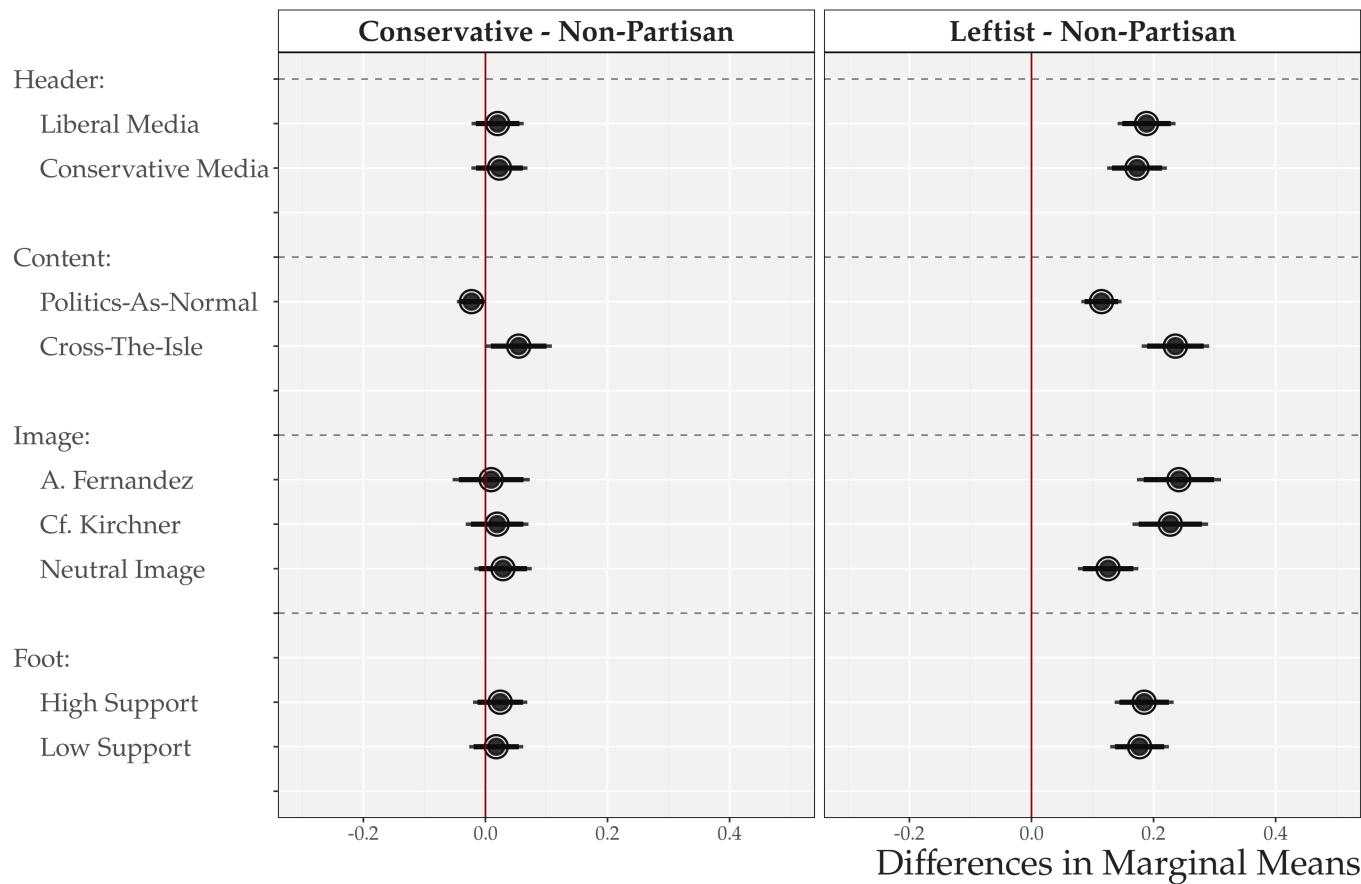


Local Coefficients for In-Degree by Quartiles

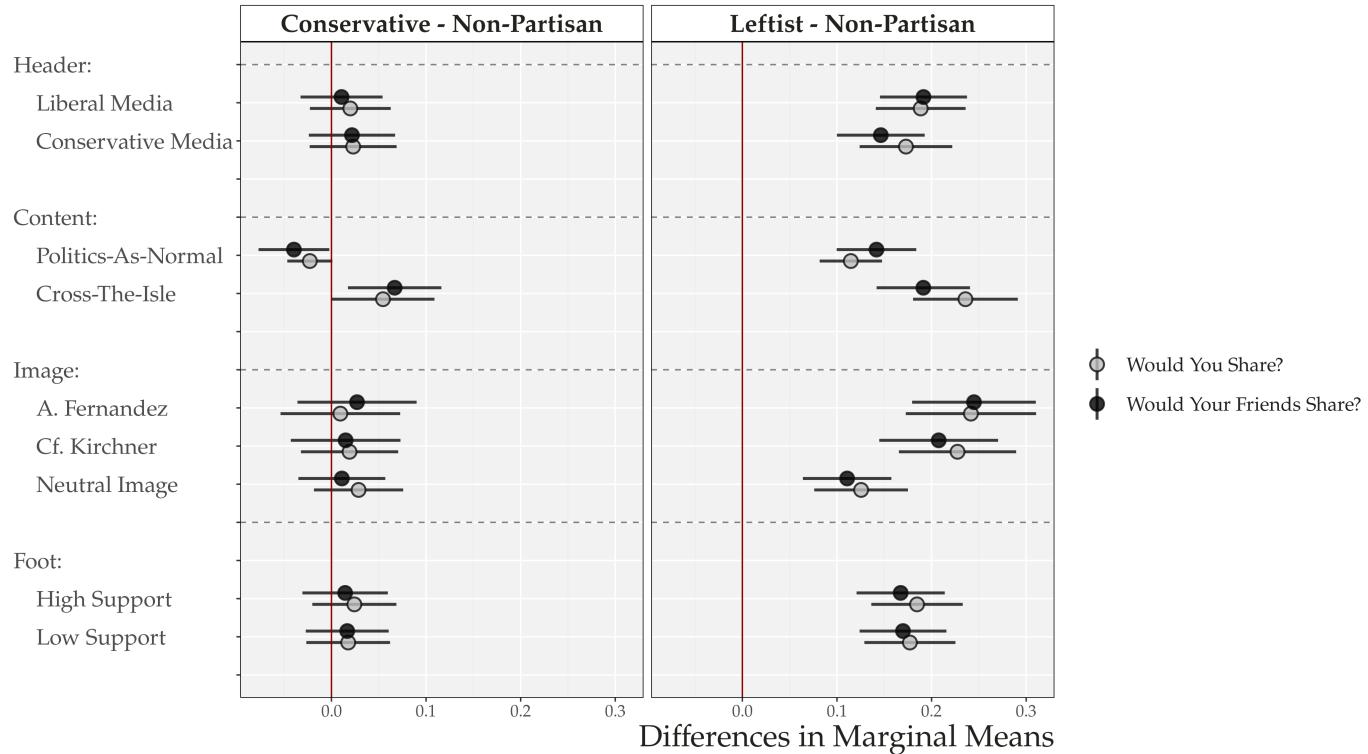
CB Ford's testimony. time to retweet ~ In-degree

NAF Extra: Argentina Conjoint

Framing Activation (Partisans vs Non-Partisans)

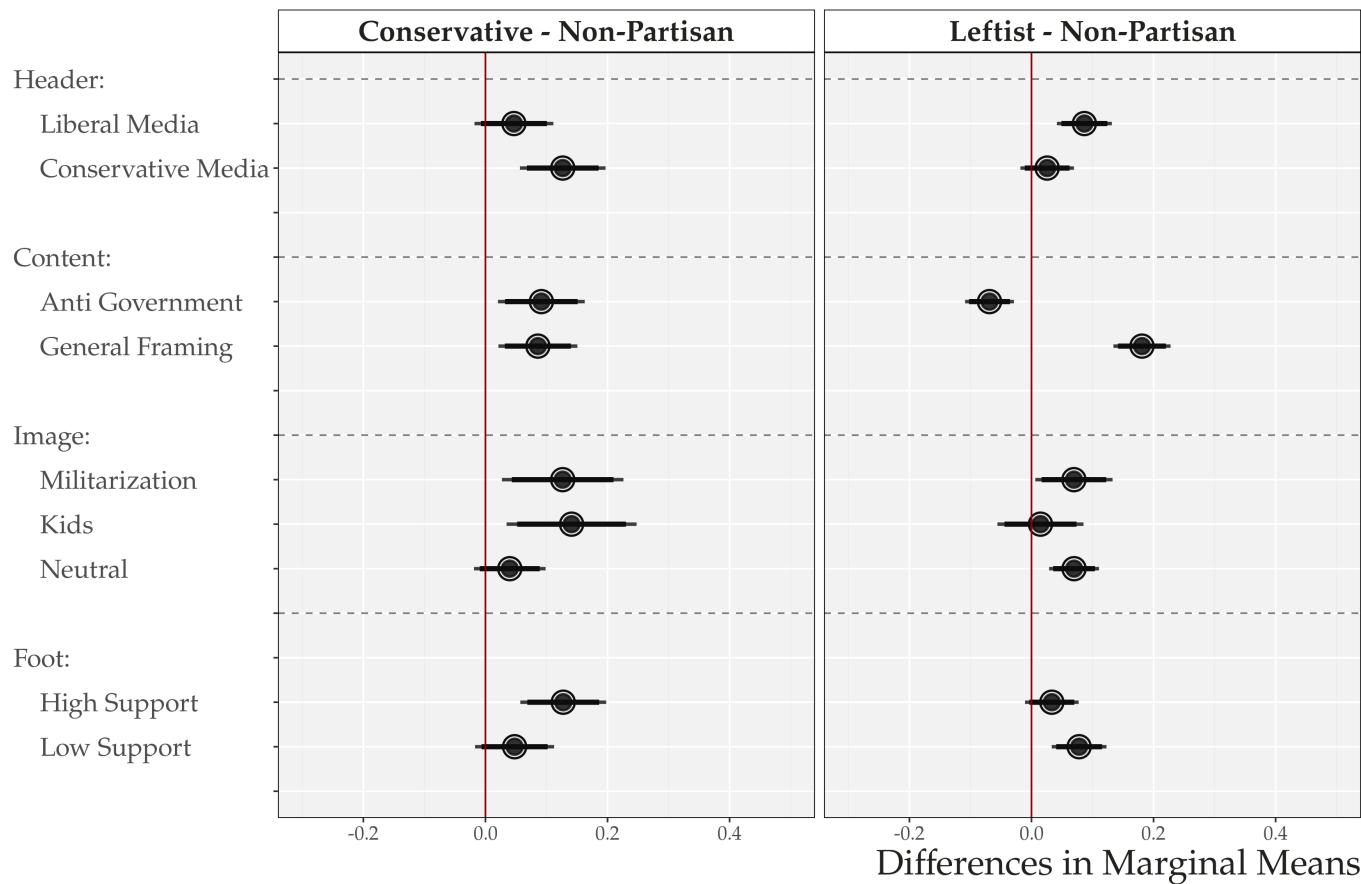


Expectations in the Network

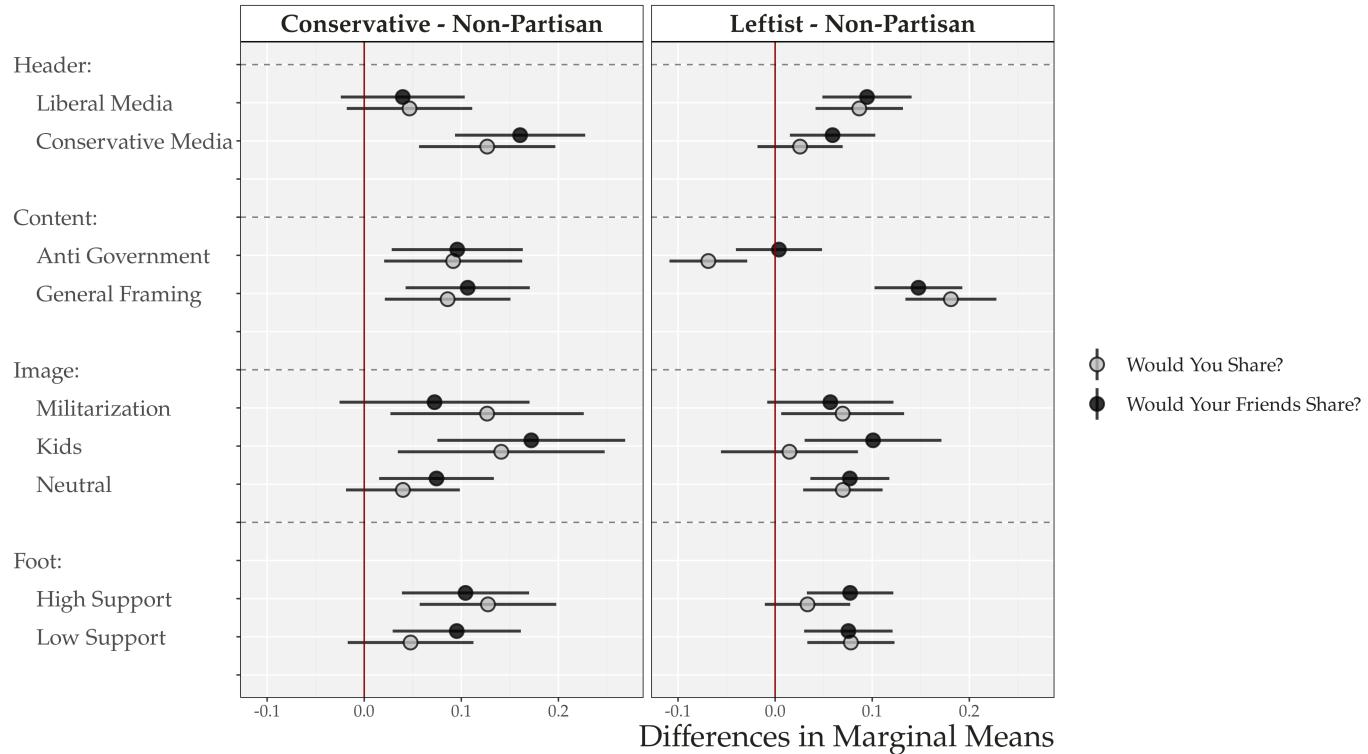


Mexico Conjoint

Framing Activation (Partisans vs Non-Partisans)



Expectations in the Network



Streaming Chats: Extra

New Technologies: Streaming Chats

Streaming Chats

Field Experiment

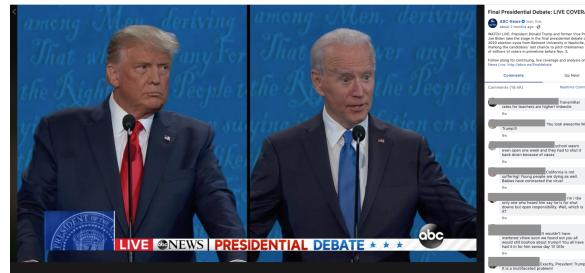
Observational Setup

Next Steps

Video Feed + Social Chat: All in one screen.

Popular among the younger generations.

- Twitch (Amazon)
- YouTube/YouTube Gaming (Google)
- Mixer (Microsoft)
- Facebook Lives



New Technologies: Streaming Chats

Streaming Chats

Field Experiment

Observational Setup

Next Steps

We conducted a large scale "field" experiment that assigns would-be debate viewers to watch on different platforms the October 2019 Democratic Debate.

Two-Wave On-line Survey in September 2019 through MTurk (following Gross, Porter and Wood, 2019).

Three main experimental conditions

- Control (standard NBC broadcast)
- Expert chat (538 website)
- Streaming chat (Facebook)

We then surveyed the respondents after the debate and measured the effects of streaming chats.

New Technologies: Streaming Chats

Streaming Chats

Field Experiment

Observational Setup

Next Steps

Text Analysis

To complement our analysis, we built automatic scrapper (Selenium) to collect comments on facebook pages

- One Hundred Thousand Comments from Facebook pages.
- Several Expert Chats
- Performed Dictionary Methods Sentiment Analysis.
- Deep Learning Models to identify toxicity.

New Technologies: Streaming Chats

Streaming Chats

Field Experiment

Observational Setup

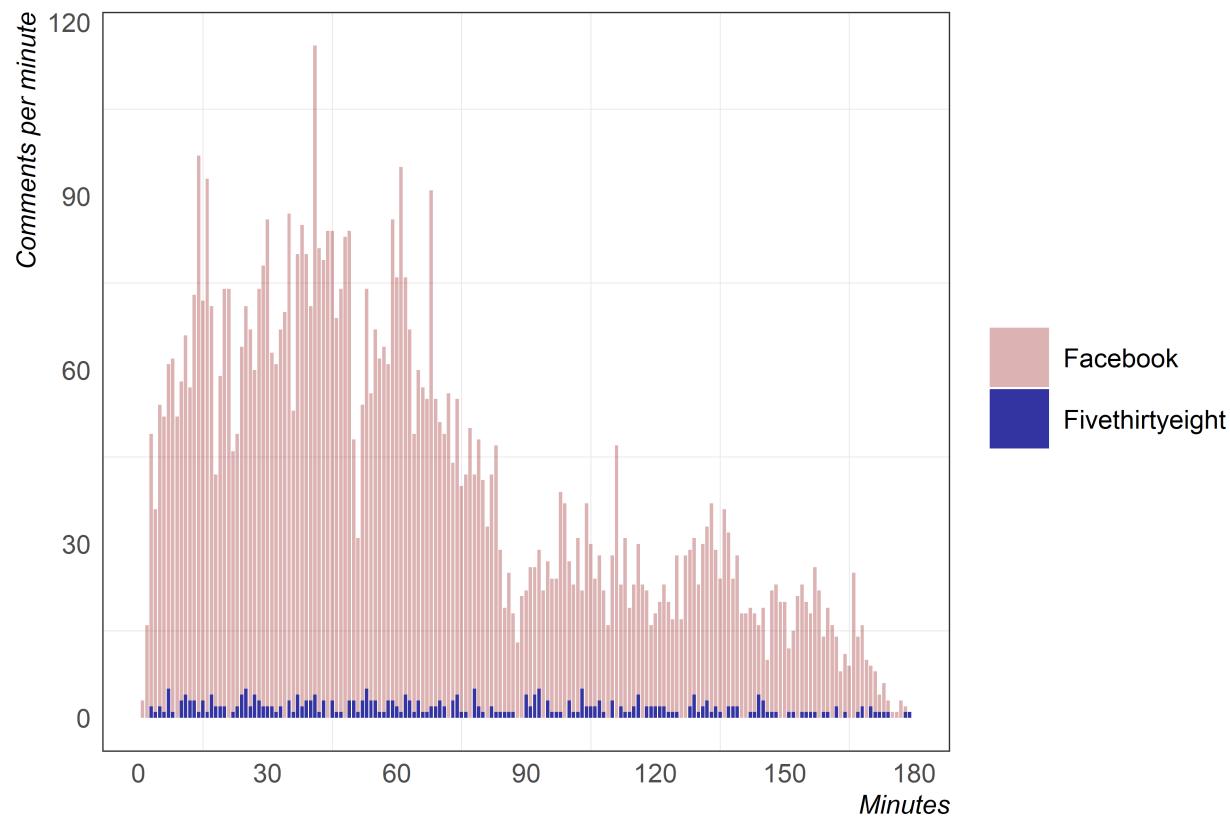
Next Steps

Grant under review to:

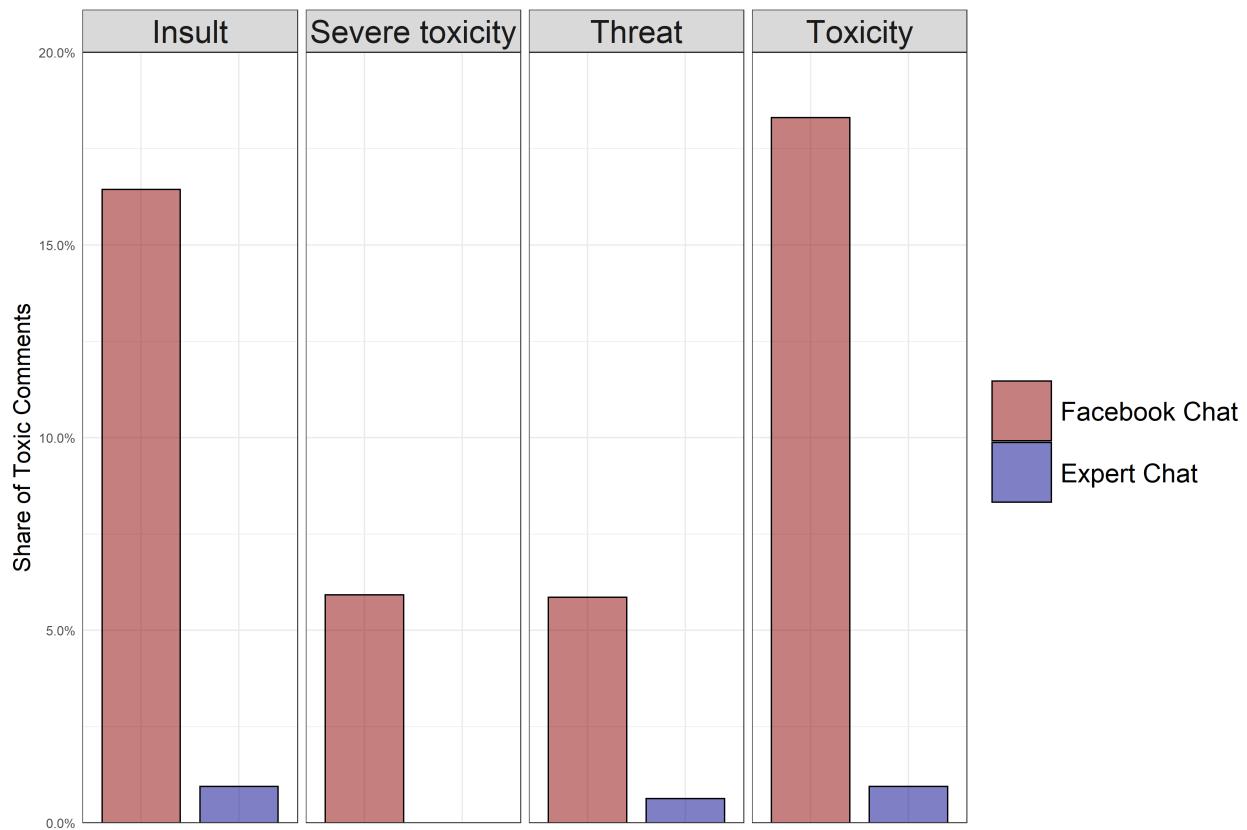
- Deploy a similar experiment in Brazil, during the Presidential Election
- Combine the field experiments with collection of web-browsing data
- Measure behavioral mechanisms with a laboratory experiment



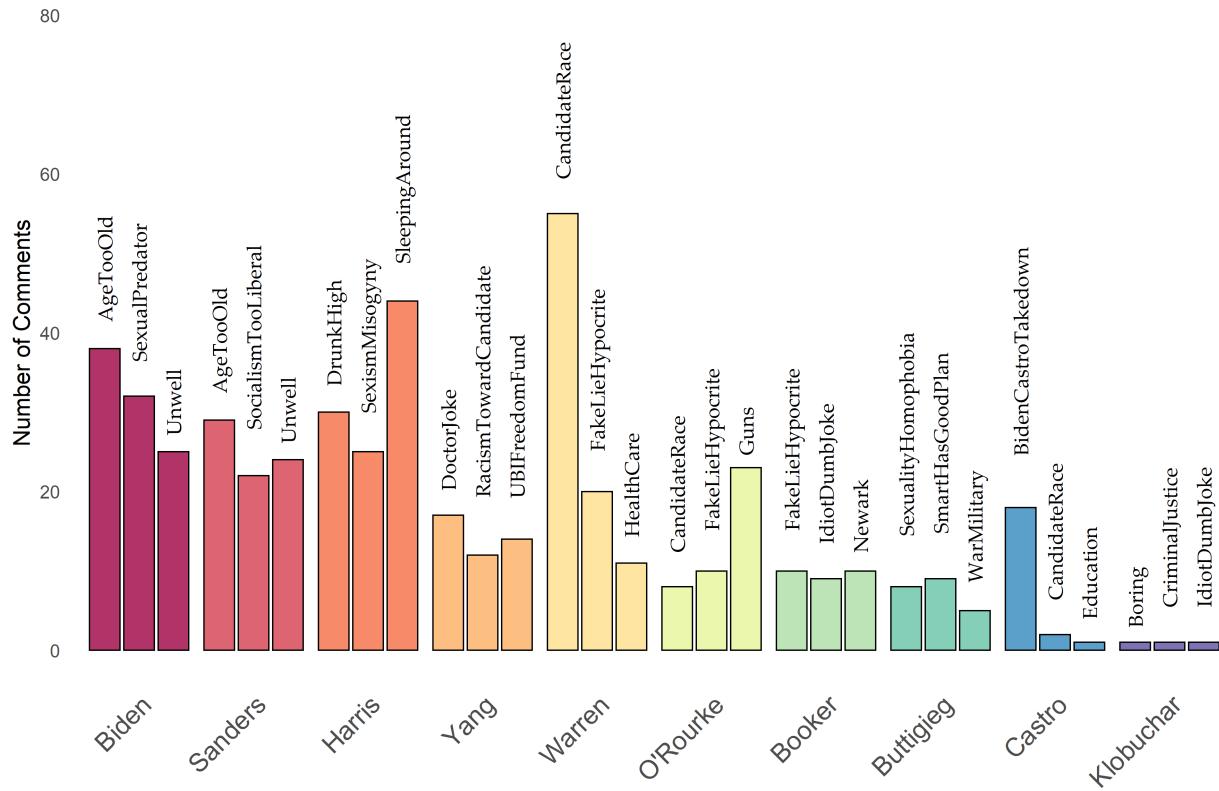
High Frequency



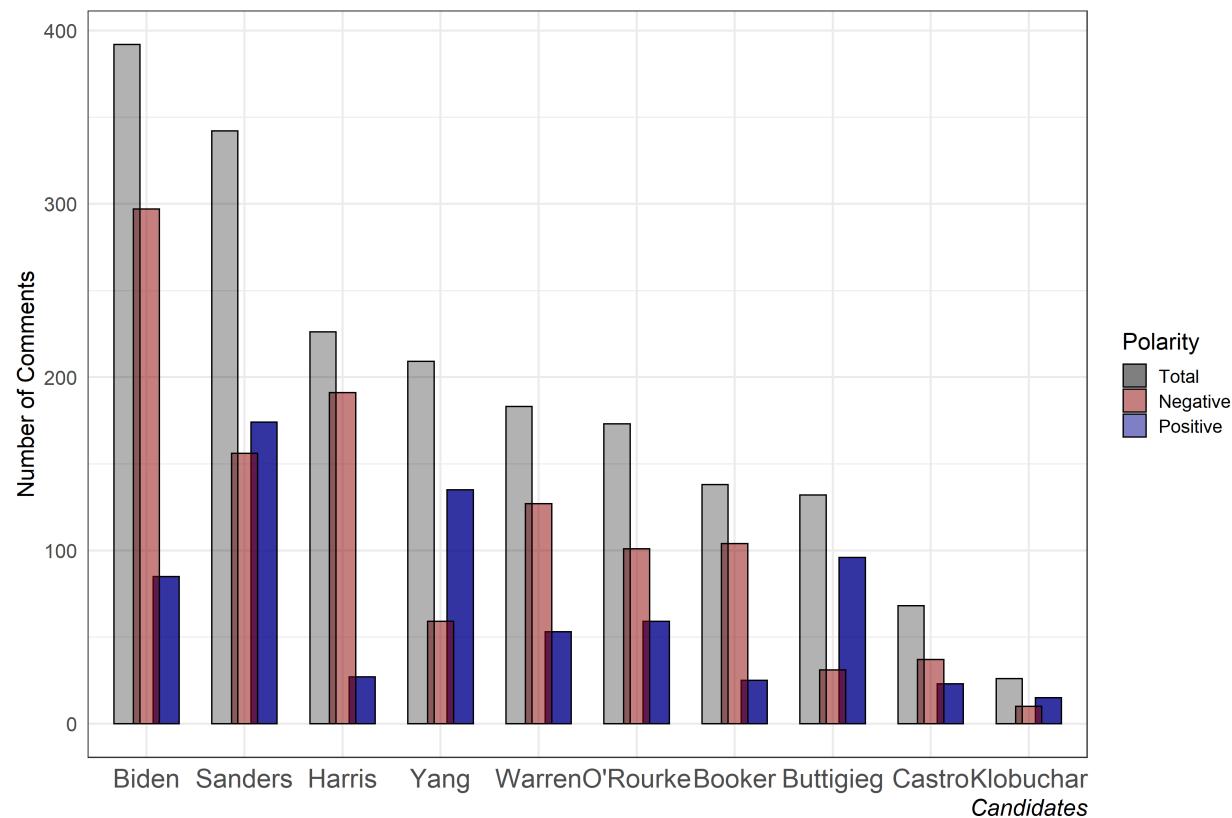
Much more toxic:



Contains Mostly Negative Primes



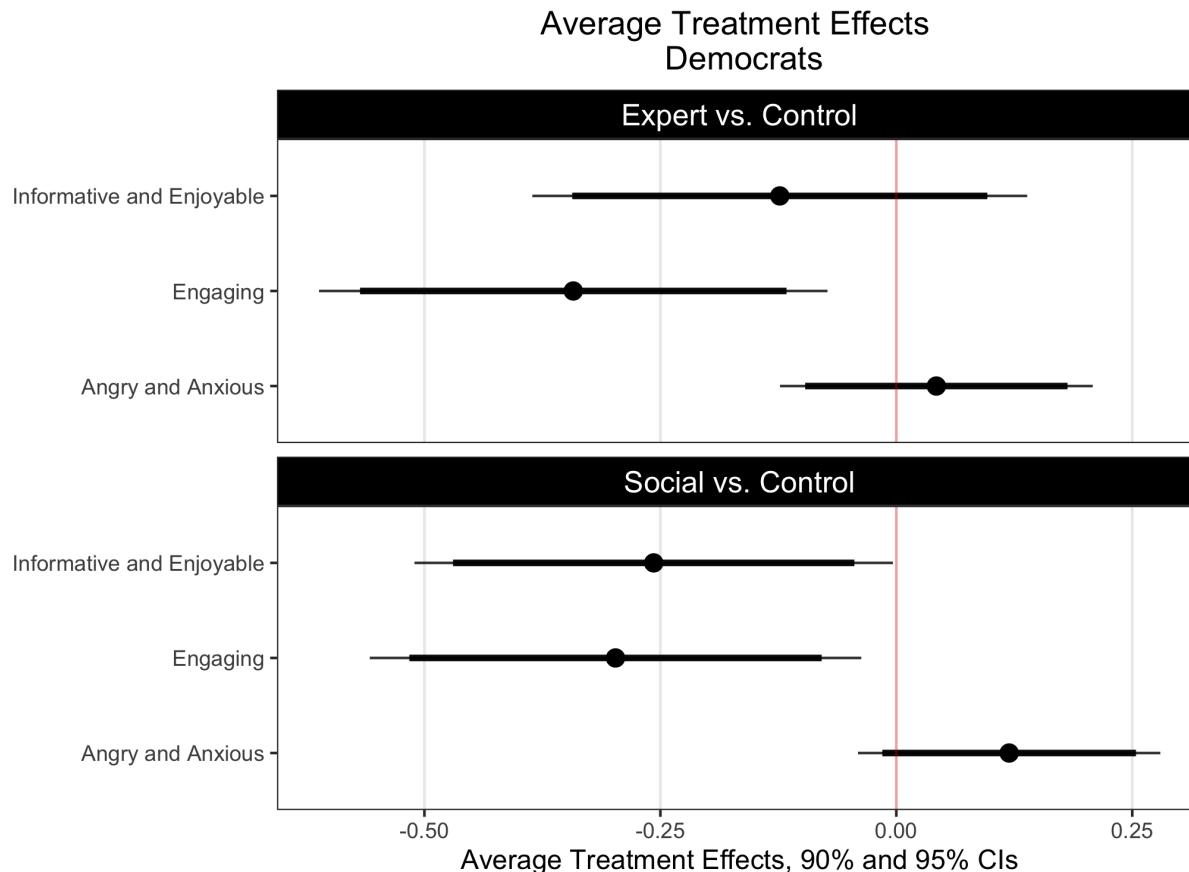
Mostly Negative Polarity about the Candidates



Experimental Results

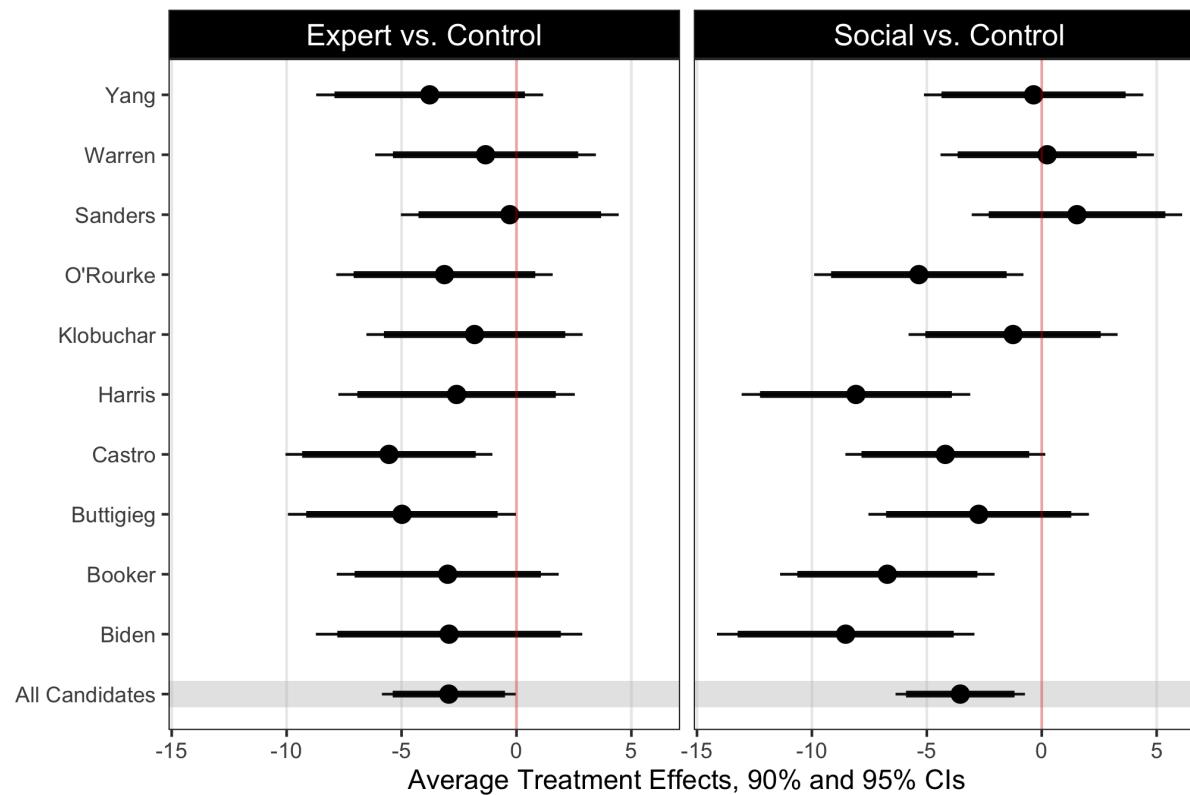
Frequency and Toxicity Hypotheses

Facebook chat somewhat less informative, enjoyable, and engaging

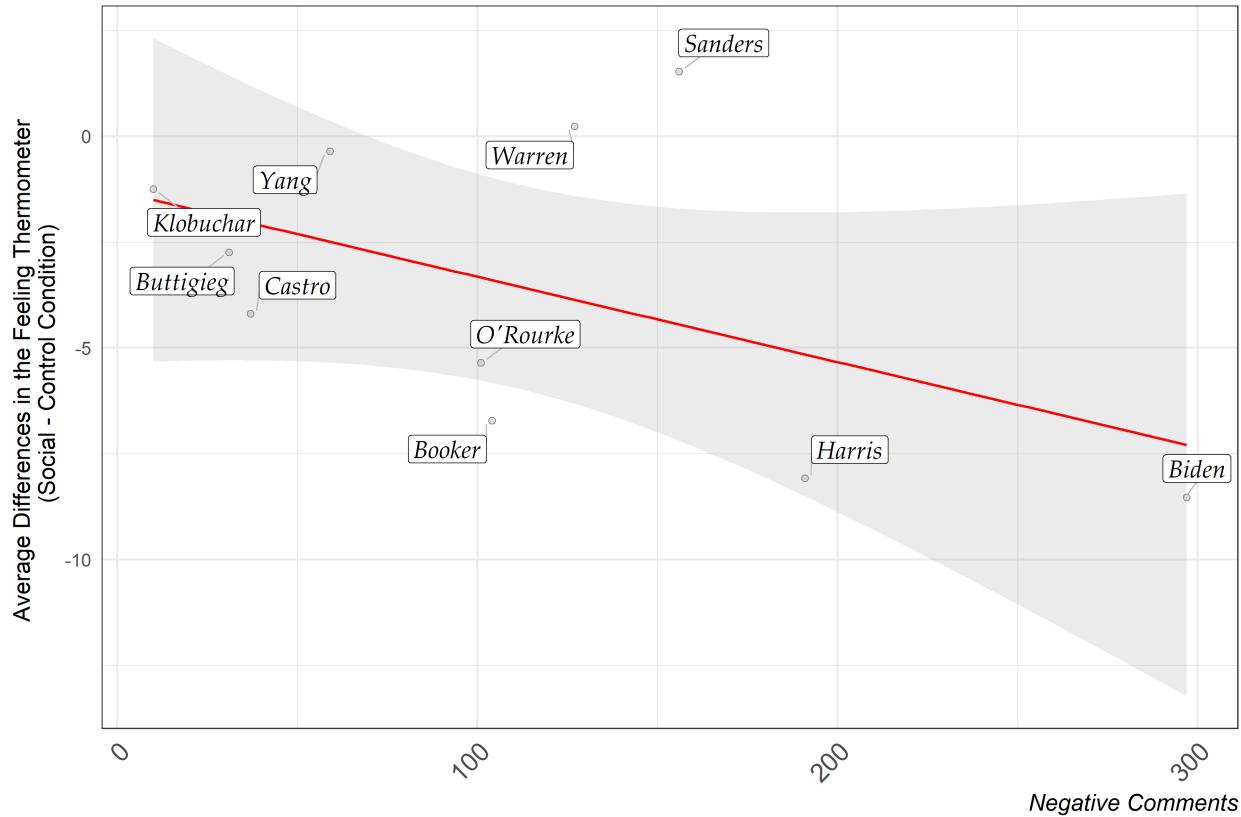


Feeling Thermometers about the Candidates

Average Treatment Effects for Feeling Thermometers
Democrats

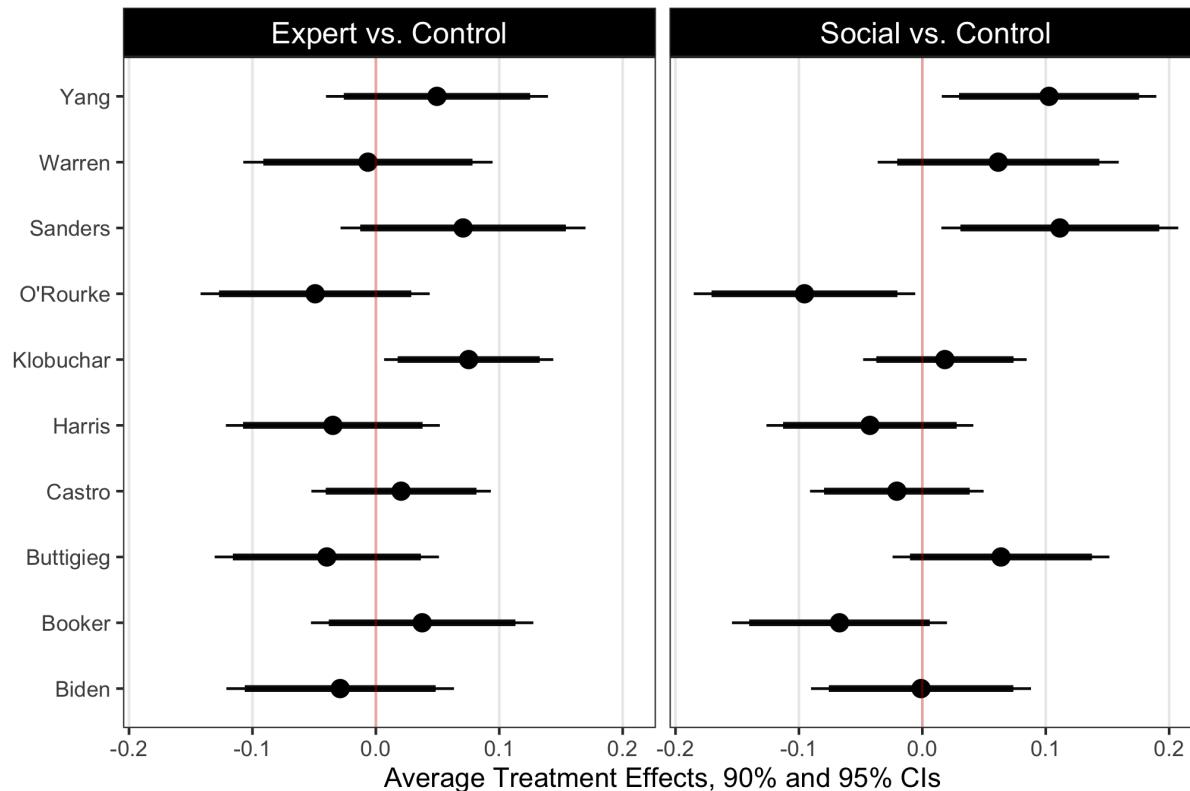


Content Effect

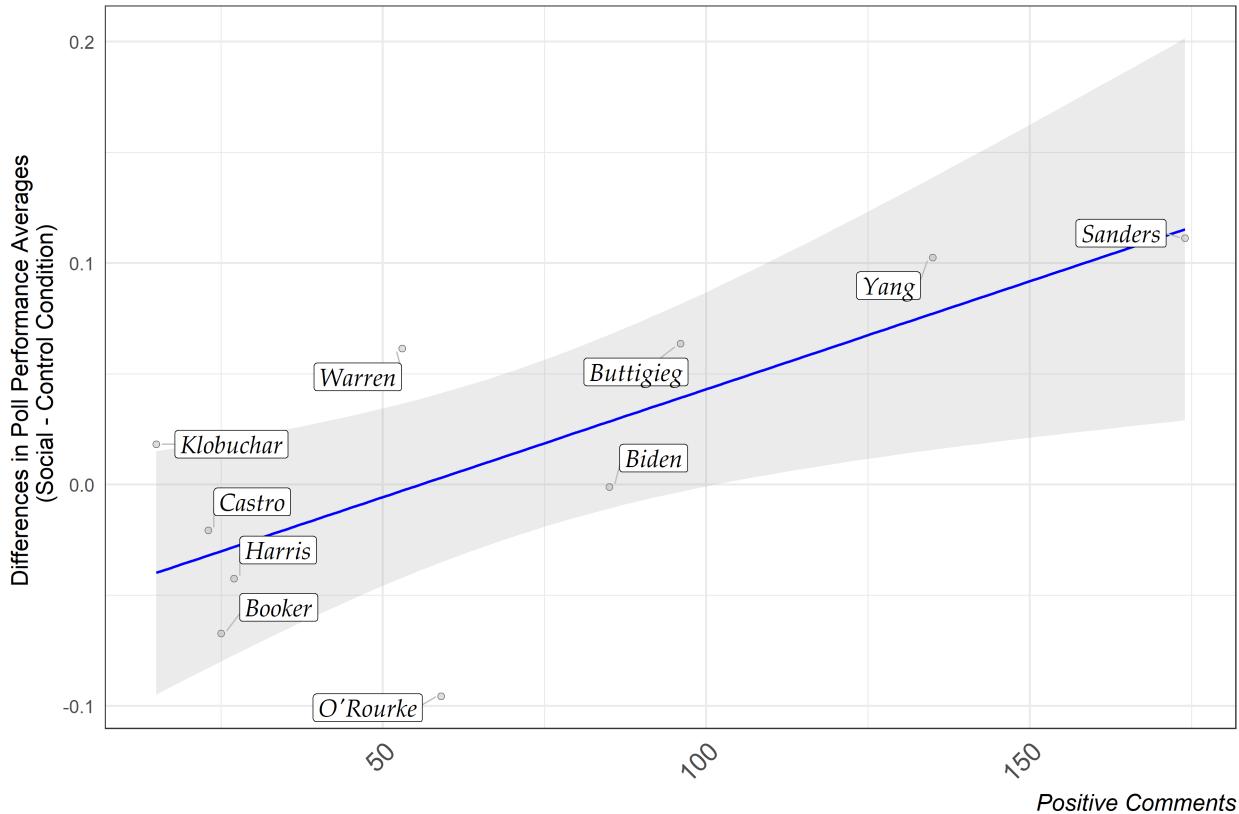


Poll Performance

Average Treatment Effects for Poll Performance
Democrats



Context Effect



Summary

Main Findings:

- Creates worse viewing experience.
- May disproportionately negatively affect certain candidates subject to toxic, negative comments.
- May distort inferences about candidate viability.

Next Steps:

- More research on the mechanism (laboratory experiments).
- More comparative evidence (Does it replicate to other platforms or other countries?).
- More descriptive evidence about these news technologies.

Comparative Politics, Violence and CSS

CSS and Political Violence

Legislating For Violence

Social Networks of Victimization

Facebook Surveys

140 thousand legislative speeches in Brazil (text + audio)

- Topic modeling to detect speeches about violence + ML Models to detect Issue Ownership.

Next Steps: Modern NLP Techniques (Word Embeddings)

- Emotions in context (word2vec + cossine similarity + dictionary)
- Scaling of policy preferences (doc2vec)

Table 2 Violence and Security on Congressional Speeches in the Brazilian House (2002-2019)

Topics	Most Likely Words	FREX Words
Topic 9: Police and Military	milit,seguranc,polic,polic,forç,policia,larmad,public,exérçit,civil	polic,milit,armad,bombeir,policial,seguranc,exérçit,civ,forç,polic
Topic 11 : Gender and Violence	mulh,violênc,homens,contr,lut,tod,feminin,direit,aind,gêner	mulh,homens,violênc,feminin,gêner,igualdad,lut,comemor,internacional,contr
Topic 25: Children and Violence	crianc,jovens,adolescent,anos,idad,menin,sexual,infantil,explor,iov	crianc,adolescent,jovens,menin,sexual,idad,infantil,infanc,iov,adult
Topic 37: Crime	crim,violênc,pre,seguranc,crimin,penal,organiz,armas,combat,public	crim,crimin,armas,pre,penal,criminal,homicidi,assassin,violênc,träfic
Topic 45: Race and Violence	pobr,neigr,popul,forç,pobrez,desigualdad,social,viv,ric,misér	negr,pobr,desigualdad,pobrez,misér,forç,ric,branc,igualdad,rac

Note: Results are estimated using a Structural Topic Model with 60 topics, in a corpus of 133,485 speeches from Representative in the Brazilian Lower Chamber. The table presents only the five topics addressing issues of violence, crime, and public security. For each topic, I present the words with i) highest probability to be part of the topic, and ii) highest FREX (Frequency and Exclusivity) (Roberts et al., 2014a))

CSS and Political Violence

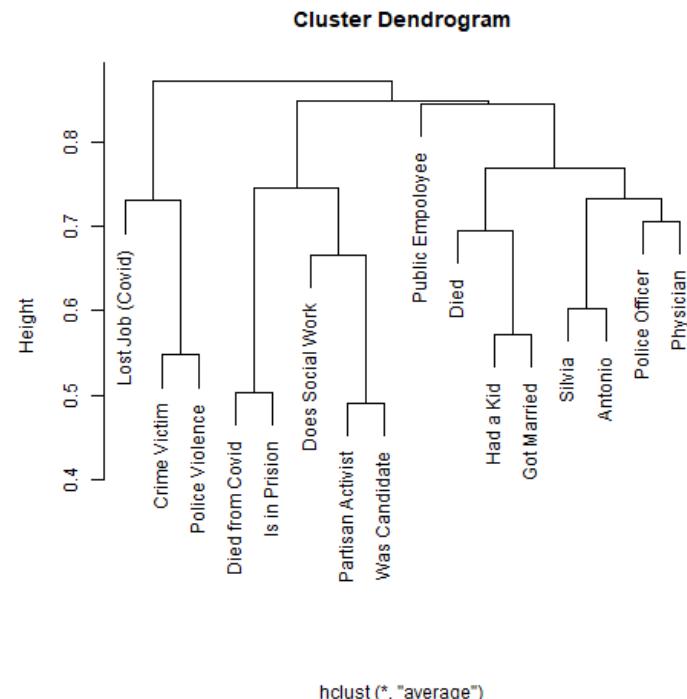
Legislating For Violence

Social Networks of Victimization

Facebook Surveys

Novel network models to estimate
contextual victimization

- How many friends do you know that had a kid last year?
- How many Silvia do you know?
- How many friend do you know that work as teachers?
- **How many friends do you know that were victims of crime?**



CSS and Political Violence

Legislating For Violence

Social Networks of Victimization

Facebook Surveys

BORDER REPORT

Cartels use COVID-19 as excuse to raise drug prices, distribute food to poor in Mexico, report says

Far from suffering losses, transnational criminal organizations adapted to challenge and are now in expansion mode, Congressional Research Service says



A man poses with a box with donations of basic goods handed to him by employees of the foundation of Alejandrina Guzman, daughter of Mexican drug lord Joaquin "El Chapo" Guzman, amid the new coronavirus pandemic in Guadalajara, Mexico, on April 17, 2020. (Photo by Ulises Ruiz / AFP) (Photo by ULISES RUIZ/AFP via Getty Images)

EGAP COVID-19 grant

A screenshot of a Facebook post. The post has a green profile picture with a white letter 'E' and the word 'Encuesta'. Below it says 'Sponsored · *'. The text reads: 'Conteste esta encuesta y participe por un iPhone'. It shows a hand holding a smartphone displaying a survey interface with three checked boxes. At the bottom, it says 'UMDSURVEY.UMD.EDU' and 'Conteste esta encuesta y participe por un iPhone' with a 'LEARN MORE' button. Below the post are standard social media interaction buttons: 'Like', 'Comment', and 'Share'.

Statistical Model

The estimation of the utility function for sharing uses a multilevel overdispersed poisson model.

$$y_{ij} \sim \text{Po}(\mu_i)$$

$$\mu_i = \exp(\alpha_{i[q]}(x_i - L_j)^2 + A_{[q]} + R_{[j]} + \gamma_{[i]})$$

Where:

- y_i = Number of links embedded by user x media
- α_q = Random Slope by quantile
- A_q = Random intercept by quantile
- R_j = Random intercept by media outlet
- $\gamma_{[i]}$ = Overdispersion parameter