

Tweetkeeping

Why Some Politically-Interested Iranian Users Get Suspended on Twitter?



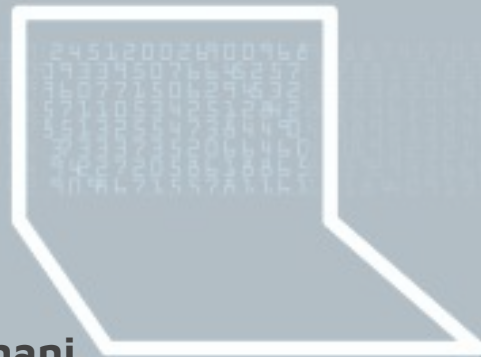
Andreu Casas

Vrije Universiteit Amsterdam

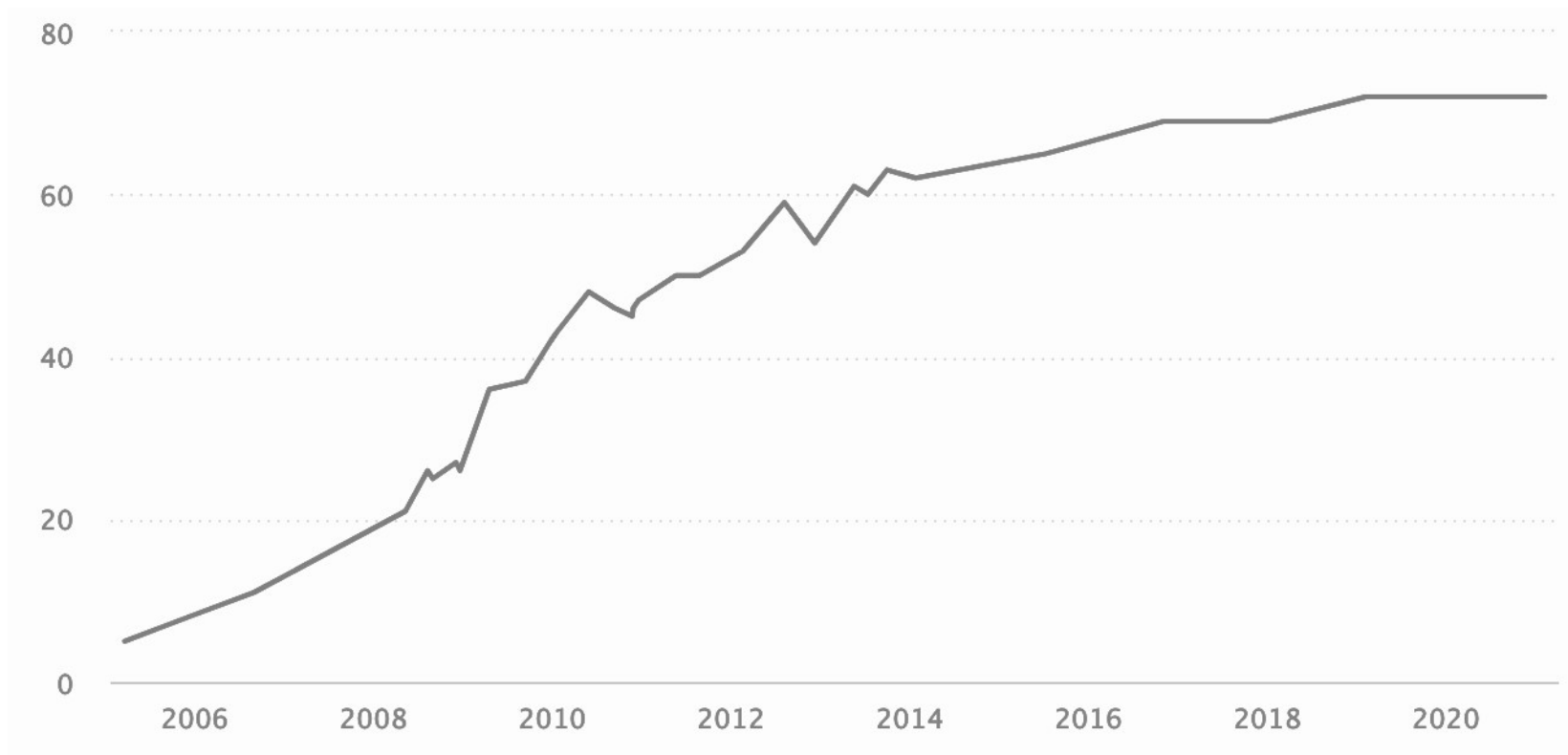


Mehdi Zamani

University of Amsterdam

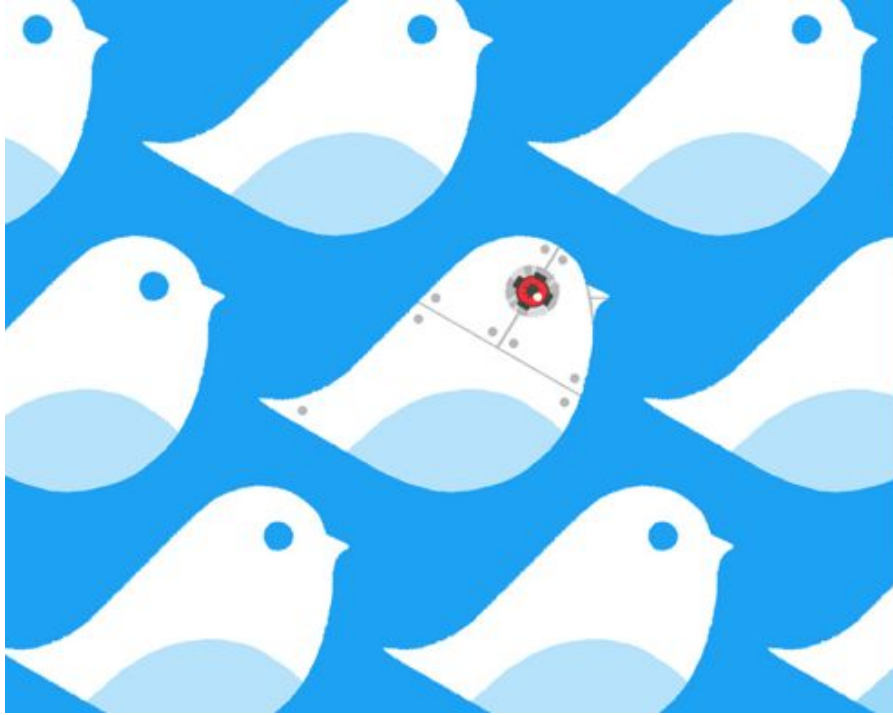


People increasingly use social media to interact with friends and family, consume information, engage and participate in politics, etc.

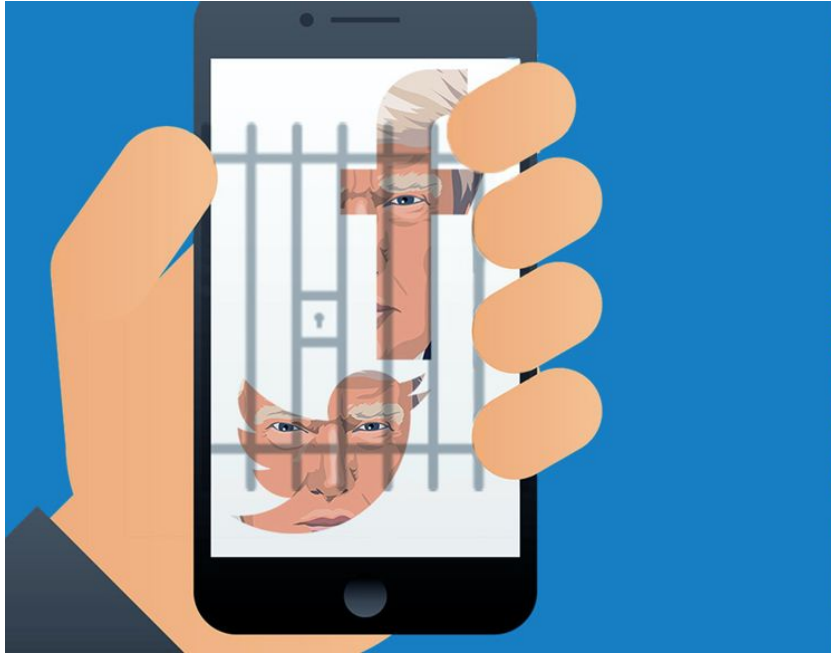


PEW (2021): % of U.S. adults who say they use at least one social media platform

Concerns about **bots**, **trolls**, **misinformation**, and **uncivil behavior** brings social media to **suspend accounts**.



Some argue that social media companies also **cancel content and voices based on their political views.**



The companies in particular report to suspend accounts who support the position of governments in a **geopolitical rivalry with the United States**

Politics

Twitter Suspends Several Accounts Used by Maduro's Government

By [Alex Vasquez](#)

January 8, 2020, 6:35 PM GMT+1

Facebook and Twitter 'dismantle Russian network'

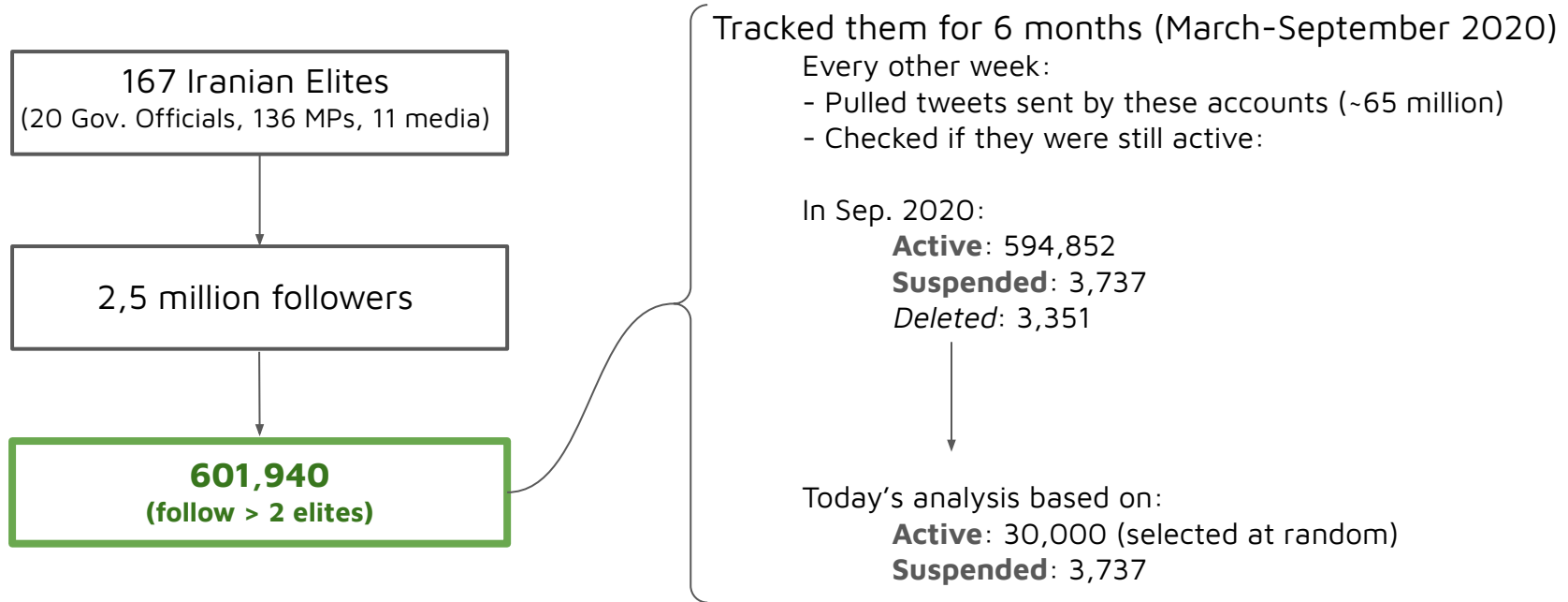
🕒 2 September 2020

Twitter suspends account apparently linked to Iranian supreme leader, citing rules about 'fake' accounts

Adam Taylor - 1/22/2021



We lack independent analysis of which accounts are suspended and for what reasons. An analysis of **suspensions in the Iranian Twittersphere**.



We had some **baseline expectations**

- Accounts using **uncivil/hateful** language will be *more* likely to be suspended
- **Malicious accounts** will be *more* likely to be suspended
Accounts that are likely to be **bots** and/or be **coordinated**
- **Verified accounts** will be *less* likely to be suspended
- Accounts spreading **misinformation** will be *more* likely to be suspended

We had some **baseline expectations**

- Accounts using **uncivil/hateful** language will be *more* likely to be suspended
- **Malicious accounts** will be *more* likely to be suspended
Accounts that are likely to be **bots** and/or be **coordinated**
- **Verified accounts** will be *less* likely to be suspended
- Accounts spreading **misinformation** will be *more* likely to be suspended

Expectations/RQ related to their **political views**

- Accounts **posting political messages** will be *more* likely to be suspended
- **Conservative (Principlist)** accounts will be *more* likely to be suspended
- Accounts **messaging in favor of the Iranian government** will be more likely to be suspended

Measuring **I**deology

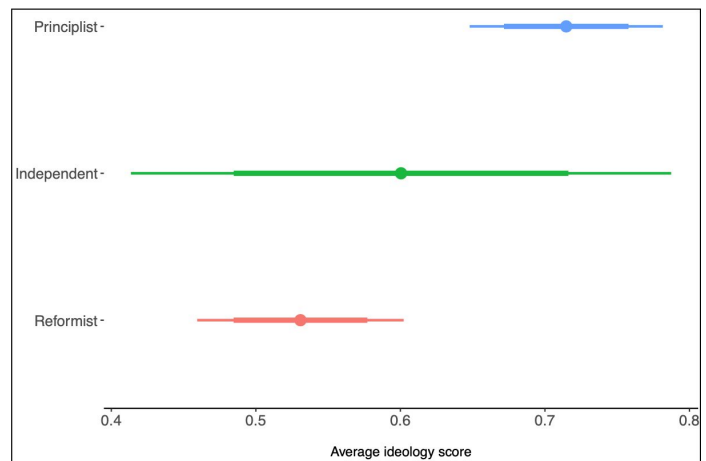
152 elites
(out of 167)

Sampled 5,000
followers
(out of ~600,000)

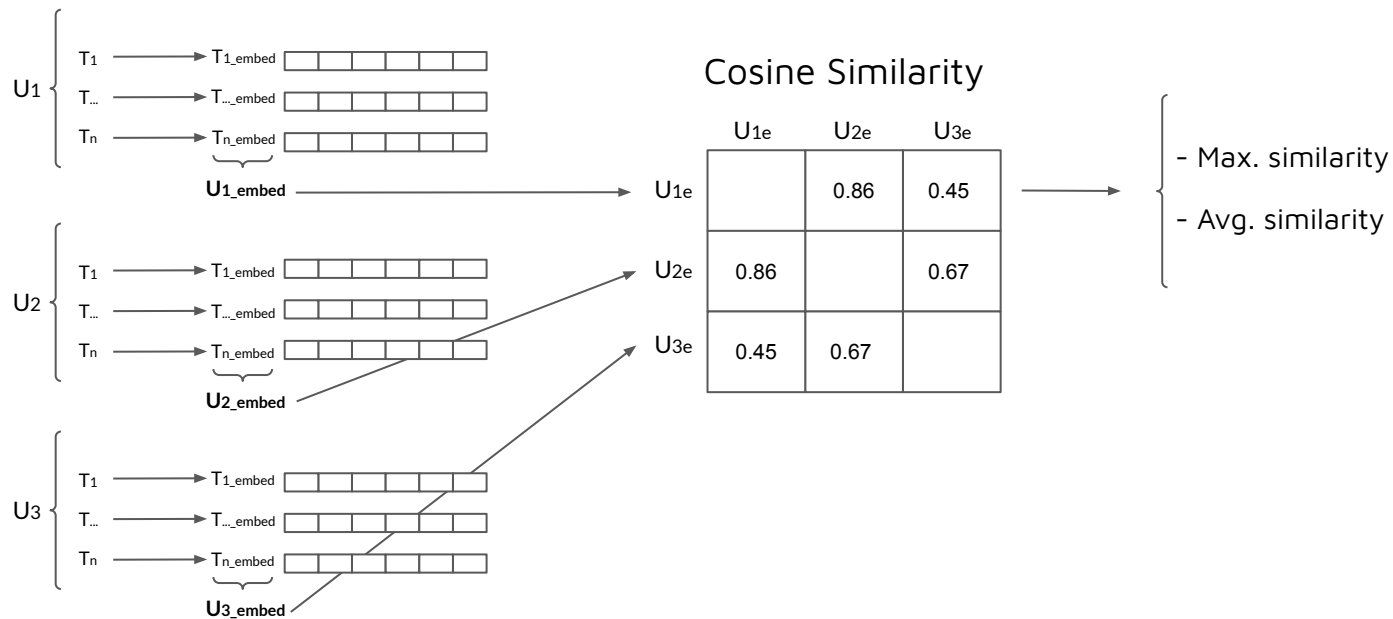
	E1	E2	E3	E4	...	E152
U1	1	0	0	0	0	1
U2	1	0	0	0	0	0
U3	1	1	0	1	0	0
U4	0	0	0	1	0	0
U5	0	0	0	1	1	1
U6	0	0	0	0	0	1
U7	0	0	0	0	1	0
...						
U5000	1	1	1	0	0	0

Item Response
Theory model

Reformist ← → Principlist



Measuring **Coordination**

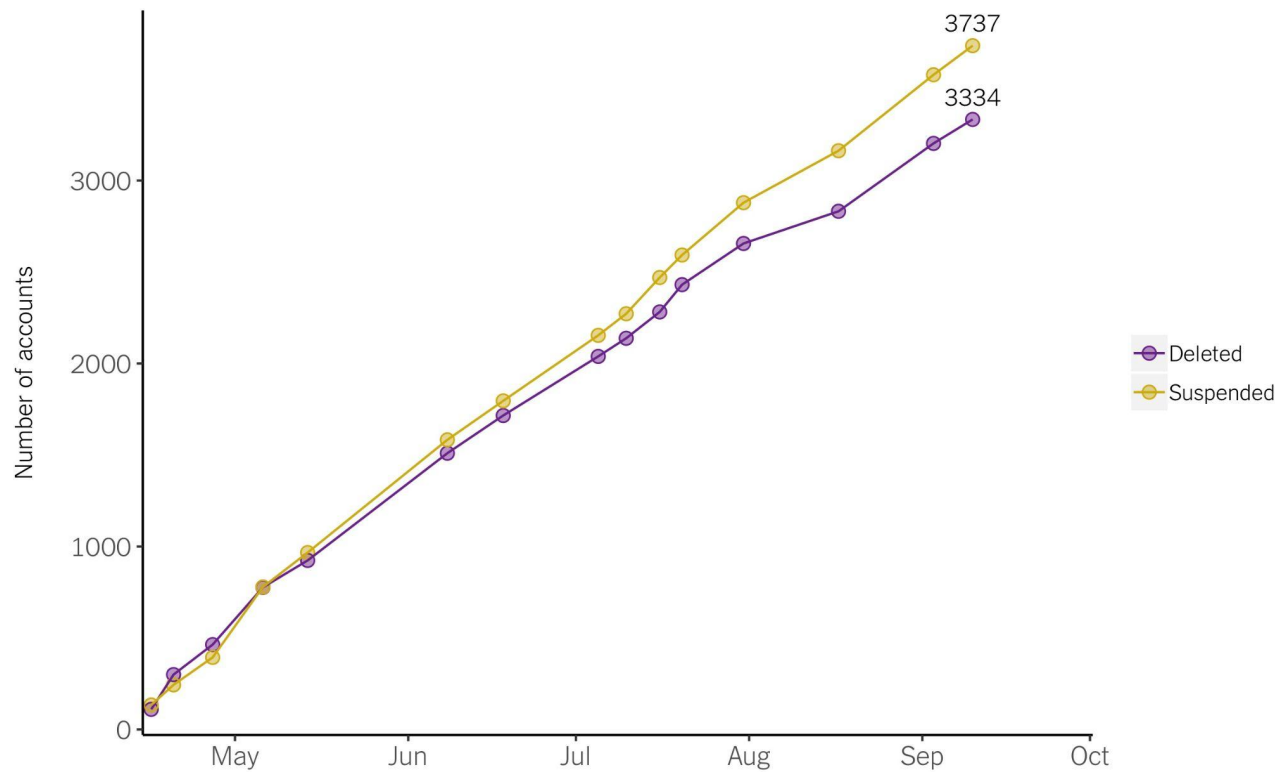


Measuring **Political**, **Uncivil/Hateful**, and **Pro-Iran** tweets

- Annotated tweets in Farsi, Arabic, and English
- Fine-tuned 3 BERT multilanguage model to predict binary outcomes
- Generated predictions for unlabeled tweets: political and hateful (binary), pro-iran (Pr)
- Aggregated predictions at the user level: political-hateful (counts), pro-iran (max, avg)

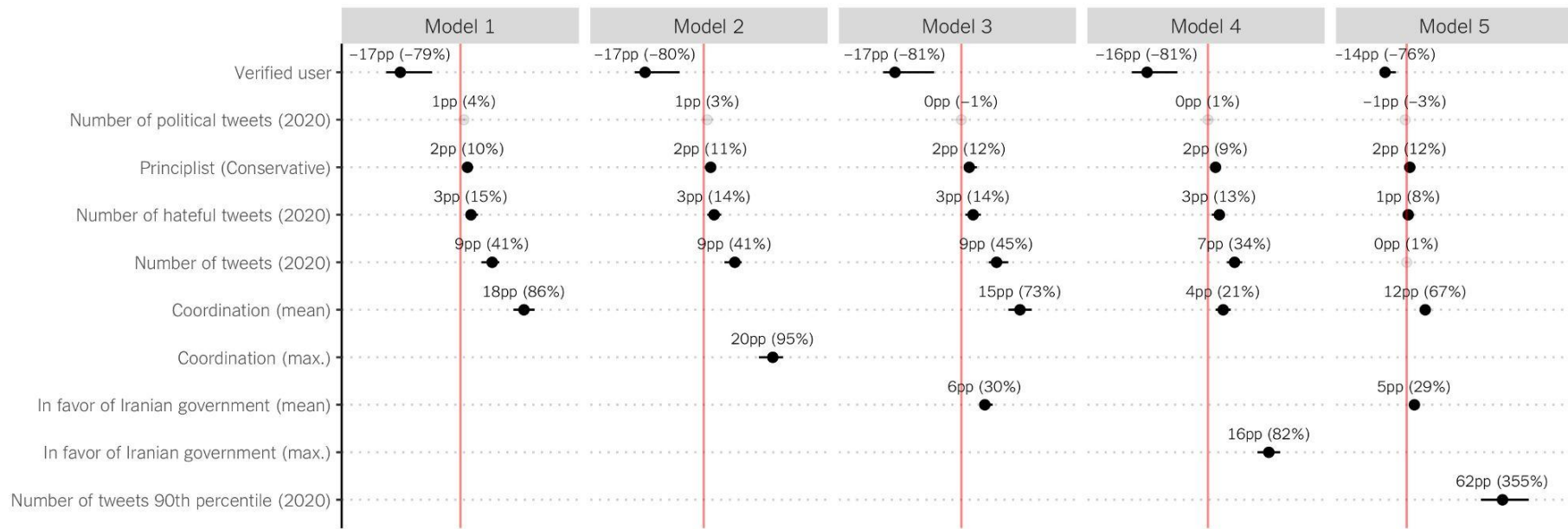
	Labeled (N)	Negative (%)	Positive (%)	Epochs	Accuracy	Precision	Recall	F-Score
Political	2,893	64%	44%	6	83%	81%	83%	82%
Hateful	1,228	90%	10%	10	93%	66%	46%	53%
Pro-Iran	607	61%	39%	8	72%	63%	69%	66%

A stable **suspension trend (2020).**



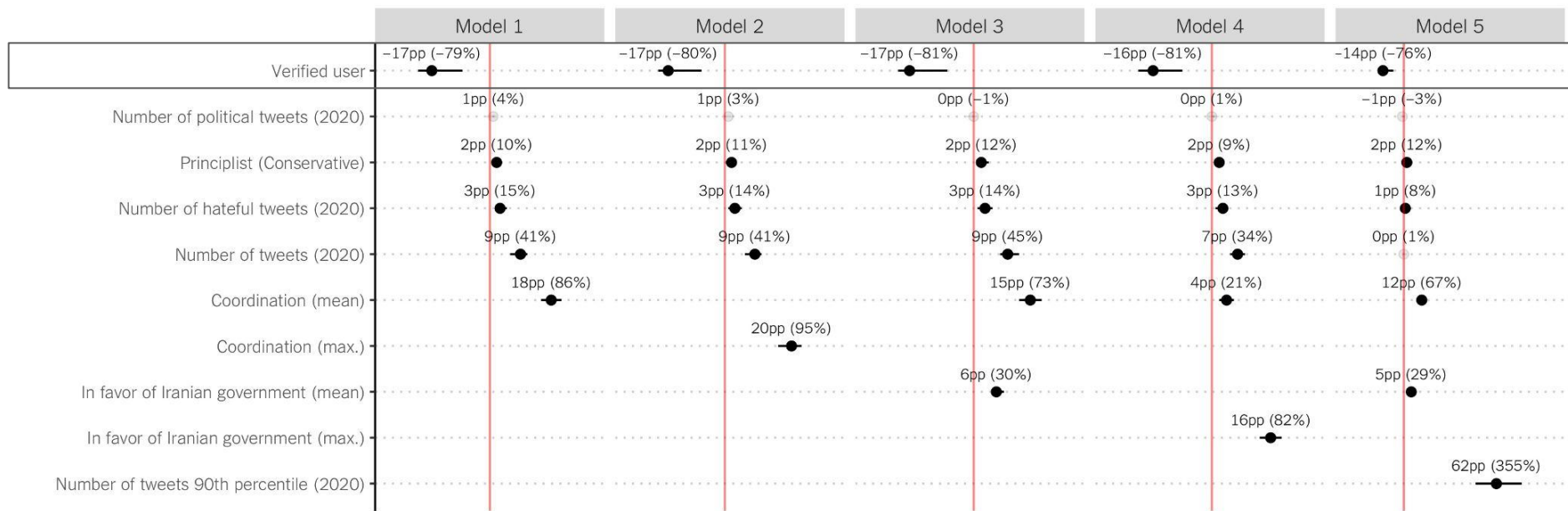
Logistic regressions **predicting suspension**.

Showing the marginal effect of a 1sd change (continuous), or of being **Verified** and in the **90th perc**
20 → 22 Pr of suspension (2pp or 10%)



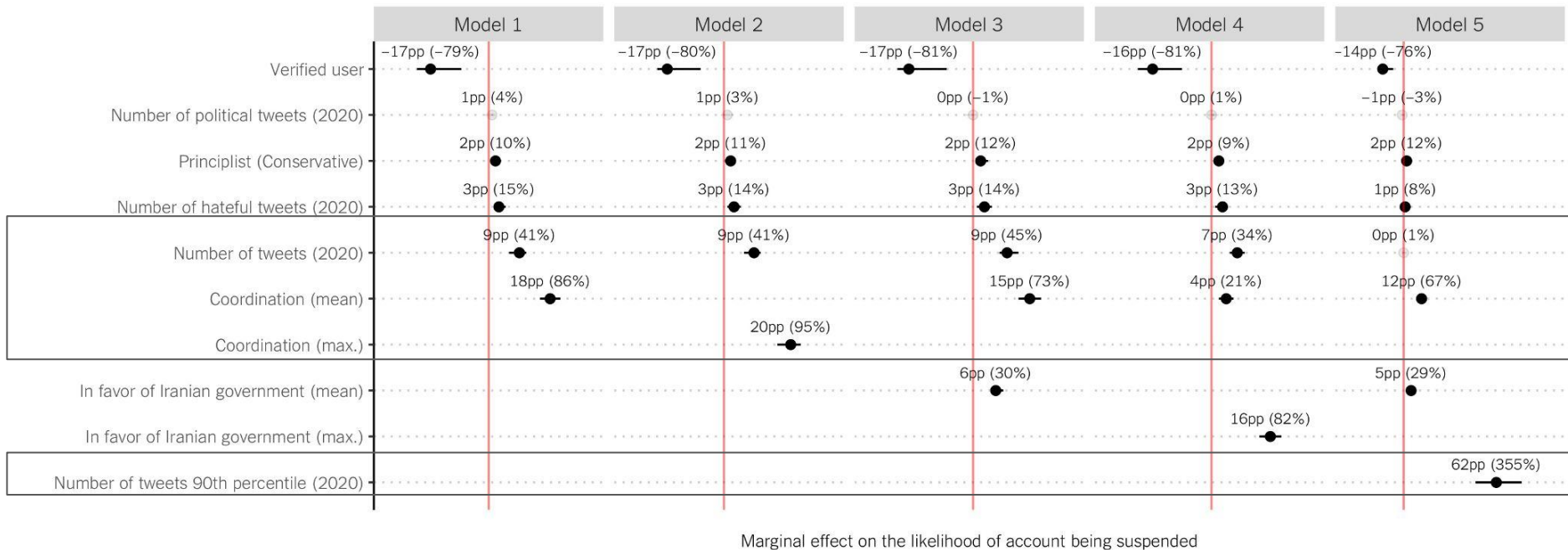
Marginal effect on the likelihood of account being suspended

Verified users substantially *less* likely to be shutdown.

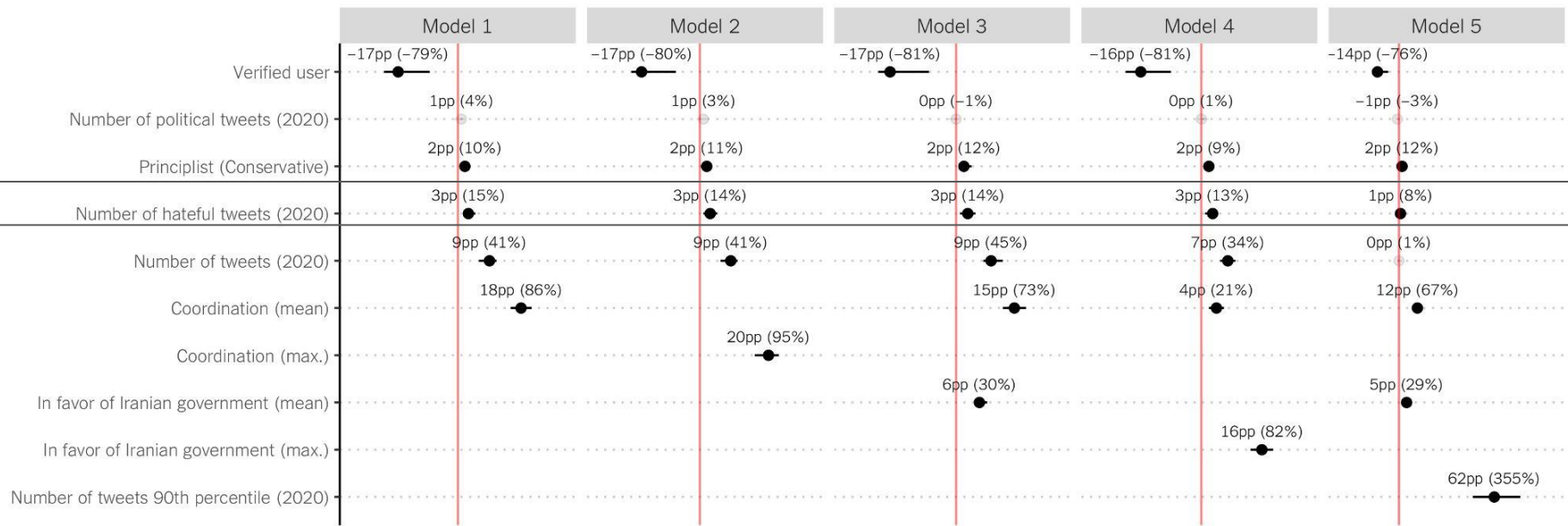


Marginal effect on the likelihood of account being suspended

Malicious accounts (likely bots & coordinated accounts) substantially *more* likely to be shutdown.

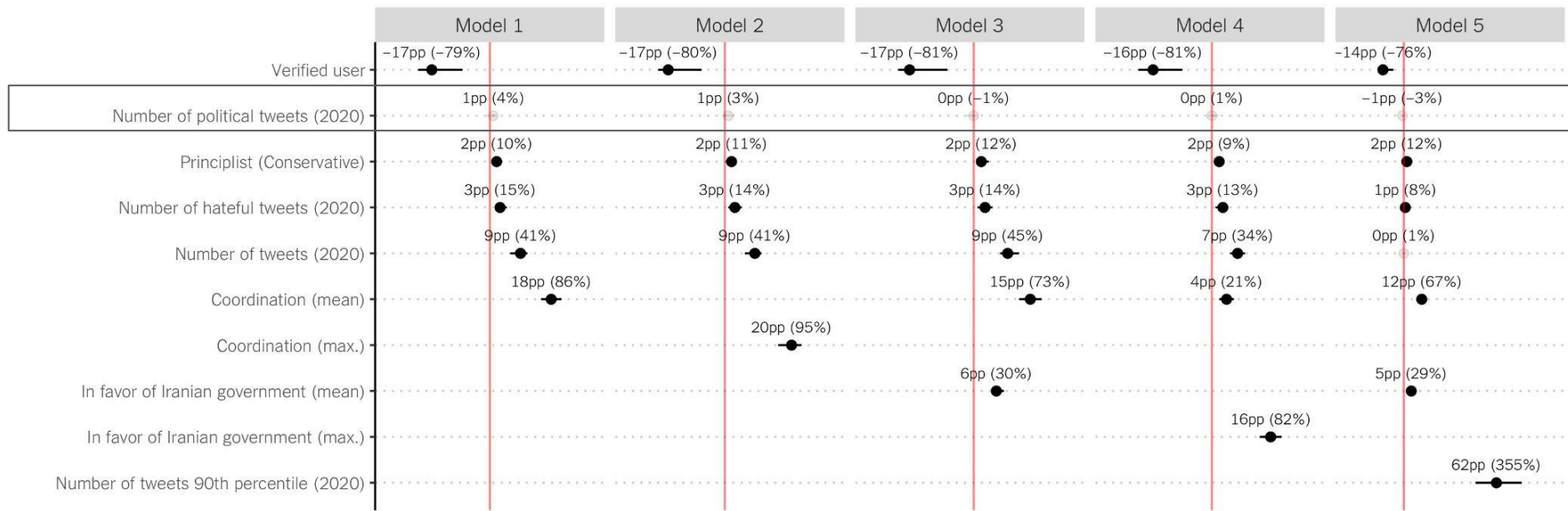


Accounts using **hateful language** are *more* likely to be shutdown.



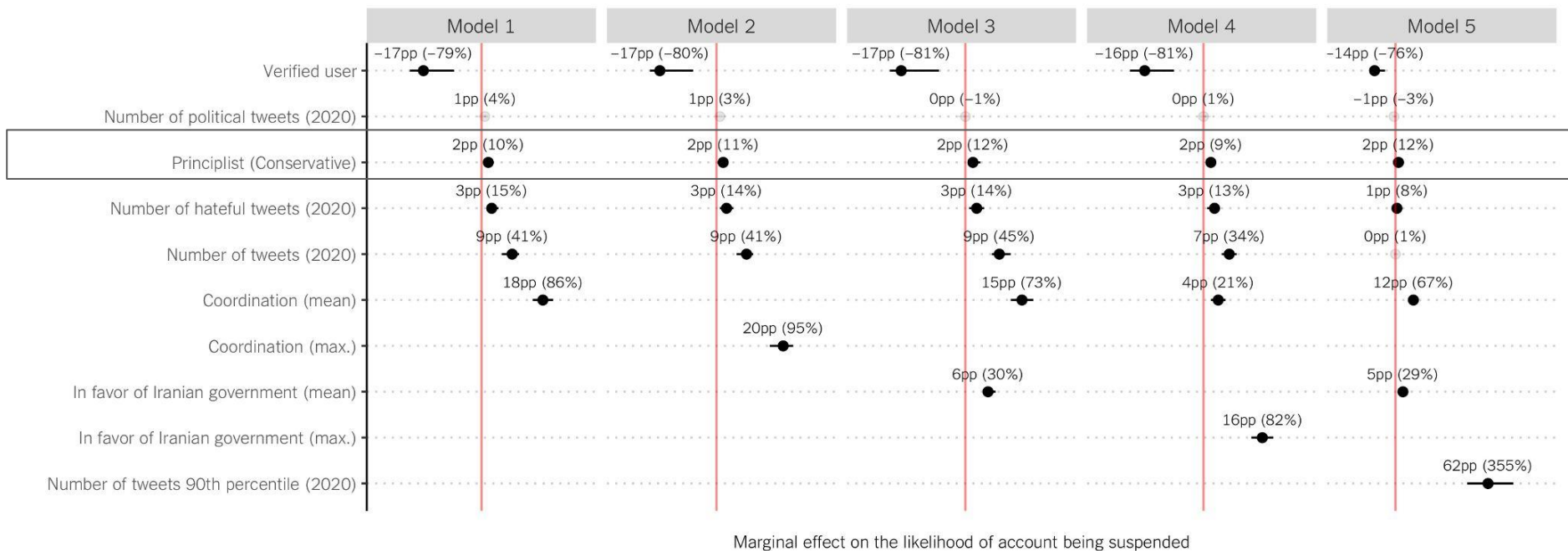
Marginal effect on the likelihood of account being suspended

Accounts **messaging about politics** are NOT more likely to be suspended.

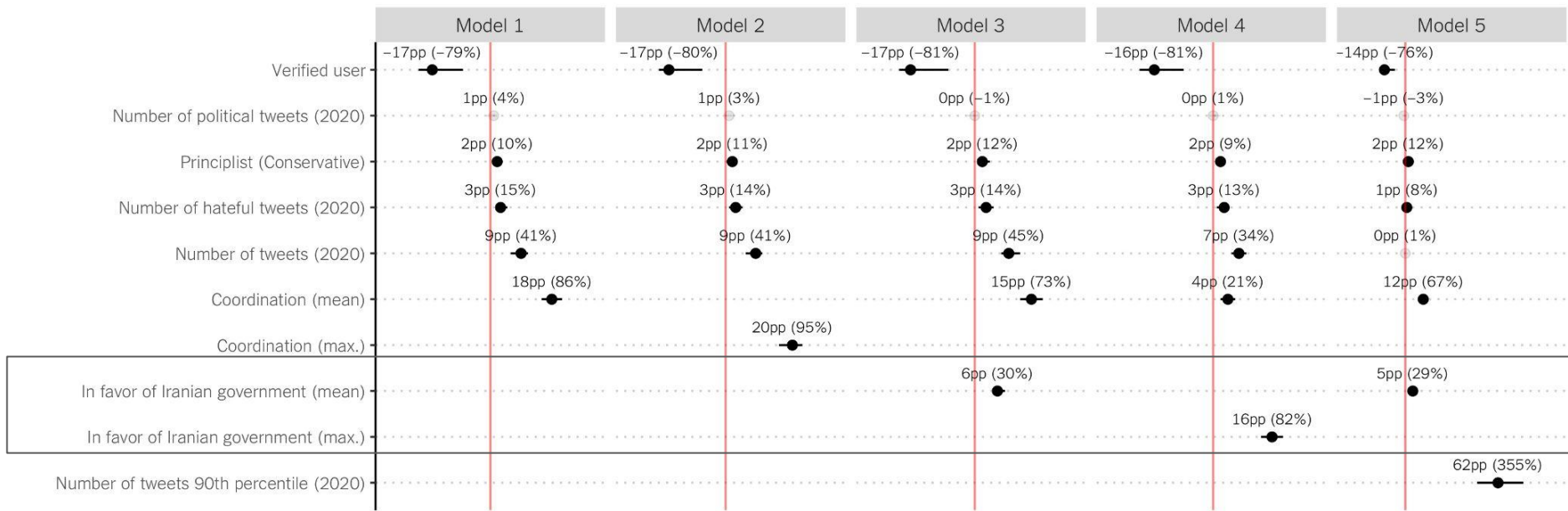


Marginal effect on the likelihood of account being suspended

Principlists (Conservatives) are more likely to be suspended.



Accounts **supportive of the Iranian gov.** are more likely to be suspended.



Marginal effect on the likelihood of account being suspended

Conclusions

- Non-trivial amount of suspensions (~ 4,000 in 6 months)
- A rather stable suspension rate
- Accounts suspended because of toxic behavior (likely bots, coordination, hateful language)
- Accounts also **suspended for geopolitical reasons**
 - Principlists & those supportive of the Iranian government
- **Crucial democratic consequences**
 - Regulation of *Freedom of Speech* increasingly in the hands of private companies
 - Clear lack of transparency despite ongoing efforts from the companies

Next Steps

- Controlling for **misinformation**
- Run analysis on the **full dataset**
- Leveraging the **panel nature** of the data
- **Across** (Non-Western & Western) **country** analysis
- **Across platform** analysis