

Appendices 1-2

Appendix 1: Far-Right Dictionary Examples and Sourcing

Source	Term Examples
Daily Stormer	globalist cuckservative racial holy war (rahowa)
AltRight Open Intelligence	goolag deus vult cultural marxism
Manosphere	red/black/bluepill femoid wojak

Note: Sample of ideologically significant and group-relevant terms tapping central concepts were sampled from various sources, including explicitly alt-right communities like *The Daily Stormer*.

Appendix 2: Far-Right and Precursor Community Examples

Community Type	Subreddit Examples
Far-right	The_Donald alternativeright thenewright physicalremoval
Conspiracy	greatawakening pizzagate conspiracy
Racist	greatapes wewuzkangs whitebeauty
Manosphere	redpill incels men going their own way (mgtow)
Violent	gore watchpeopledie
Gamer	kotakuinaction crusaderkings
Offensive Humor	darkhumorandmemes ingoingtohellforthis darkjokecentral

Note: Examples listed here have been stripped of case. All permutations of the above communities (e.g. both r/GreatAwakening and r/the_greatawakening) were included in analysis.

Appendices 3a-c: Robust Regression Models

Robust regression (RR) models for H2-4 have been provided below. Appendices 3a, 3b, and 3c correspond to ordinary least squares (OLS) models displayed in Tables 4, 5, and 6 respectively. I note the implications of the robust version for each hypothesis test and note takeaways below.

- H3 (Appendix 3b): R-squared values in the RR model are higher than those in the OLS model displayed in Table 5 (0.13 RR versus 0.014 OLS). The relationship between FREP (IV) and term rate within far-right communities (DV) reaches statistical significance ($p < 0.01$) in RR but not in OLS. The direction of the effect is unchanged.
- H4 (Appendix 3c): R-squared values in the RR model are higher than those in the OLS model displayed in Table 6 (0.07 RR versus 0.02 OLS). The relationship between FREP (IV) and term rate outside of far-right communities (DV) is reversed in direction and statistically significant ($p < 0.01$).

After running robust versions of each model and comparing results to OLS models displayed in Tables 4-6, I conclude that OLS is the most theoretically sound approach. Switching from OLS to RR results in higher R-squared values but changes some results by discarding large percentages of the sample without good theoretical reason (12 percent of the sample, in the RR model of H4). This would certainly be appropriate if I had reason to believe that the outliers were generated by extreme or unusual cases (e.g. by bots that escaped detection during the data cleaning process). Instead, I find that outlier observations appear to be generated by authentic users who do not use language in an unusual way. This study relies on experimental measures without well-defined preexisting expectations, making the cutoff between outlier and typical observations difficult to determine.

Appendix 3a: Robust Regression - Effect of Socialization on Change in Global Term Rate (H2)

	<i>Dependent variable:</i>
	Change in Term Rate, Global
FREP	0.043*** (0.009)
Conspiracy Precursor	-0.460 (0.416)
Racist Precursor	0.718 (1.705)
Violent Precursor	-0.416 (0.504)
Sexist Precursor	-0.291 (0.483)
Offensive Precursor	-0.561 (0.494)
Gamer Precursor	-0.351 (0.497)
Constant	1.196*** (0.295)
Observations	484
R ²	0.115
Adjusted R ²	0.102
Residual Std. Error	3.463 (df = 476)

Note: *p<0.1; **p<0.05; ***p<0.01. Table entries are coefficients from a robust regression model, with standard errors in parentheses. Term rate is denominated by incidence per 1000 words across all posts made in the relevant time period, and the change in this value is determined by the difference in term rate before and after three months of far-right engagement across all user activity.

Appendix 3b: Robust Regression - Effect of Socialization on Term Rate in Far-Right Spaces (H3)

<i>Dependent variable:</i>	
Term Rate, Far Right Spaces	
FREP	0.069*** (0.010)
Conspiracy Precursor	0.108 (0.610)
Racist Precursor	1.924 (3.990)
Violent Precursor	-0.196 (1.119)
Sexist Precursor	-0.258 (0.718)
Offensive Precursor	0.393 (0.875)
Gamer Precursor	0.667 (0.780)
Constant	1.807*** (0.525)
Observations	
	496
R ²	
	0.126
Adjusted R ²	
	0.113
Residual Std. Error	
	4.912 (df = 488)

Note: *p<0.1; **p<0.05; ***p<0.01. Table entries are coefficients from a robust regression, with standard errors in parentheses. The dependent variable, term rate in far-right spaces, is denominated by incidence per 1000 words across posts made in far-right communities in the relevant time period.

Appendix 3c: Robust Regression - Effect of Socialization on Term Rate in Other Spaces (H4)

	<i>Dependent variable:</i>
	Term Rate, Other Spaces
FREP	−0.009*** (0.002)
Conspiracy Precursor	0.002 (0.156)
Racist Precursor	−0.022 (0.329)
Violent Precursor	−0.043 (0.272)
Sexist Precursor	−0.084 (0.165)
Offensive Precursor	0.370* (0.217)
Gamer Precursor	0.262 (0.192)
Constant	0.889*** (0.236)
Observations	531
R ²	0.066
Adjusted R ²	0.054
Residual Std. Error	1.026 (df = 523)

Note: *p<0.1; **p<0.05; ***p<0.01. Table entries are coefficients from a robust regression model, with standard errors in parentheses. The dependent variable, term rate in other spaces, is denominated by incidence per 1000 words across posts made in other (non-far-right-affiliated) communities in the relevant time period.