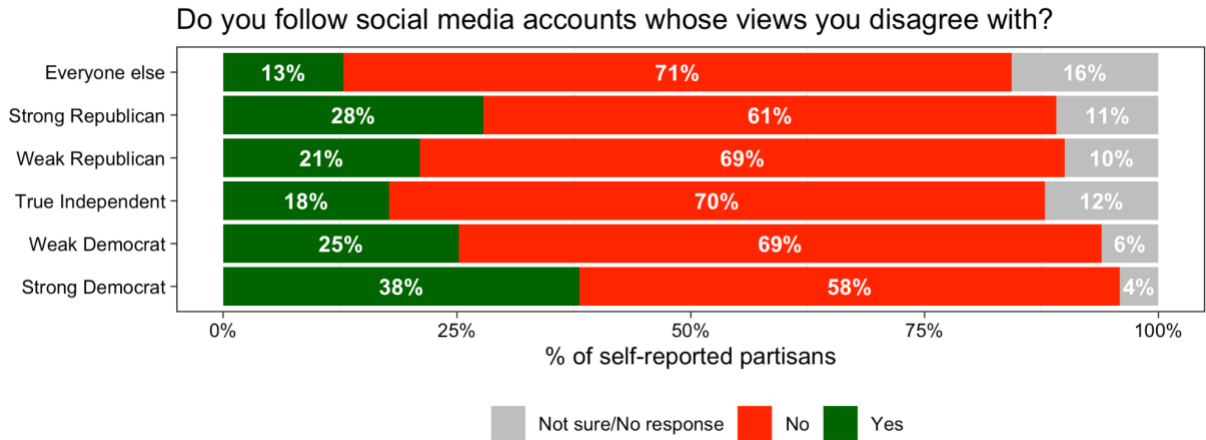
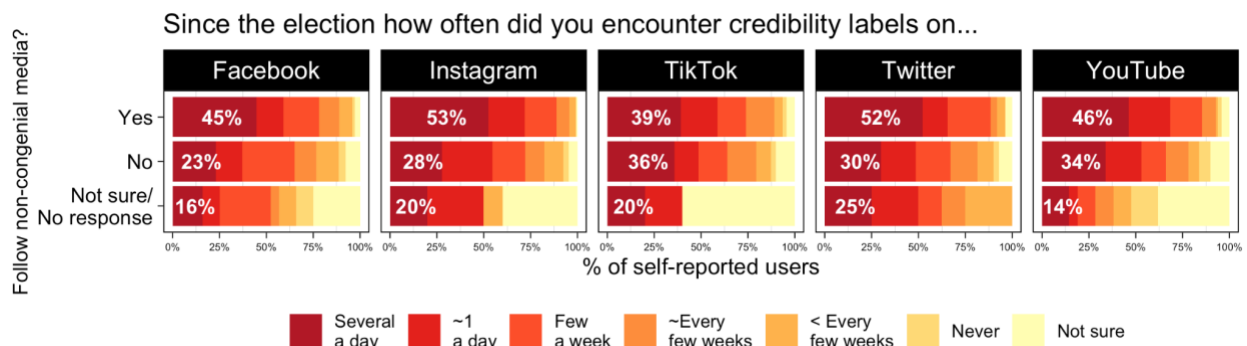


## Appendix C: Supplementary exploratory results

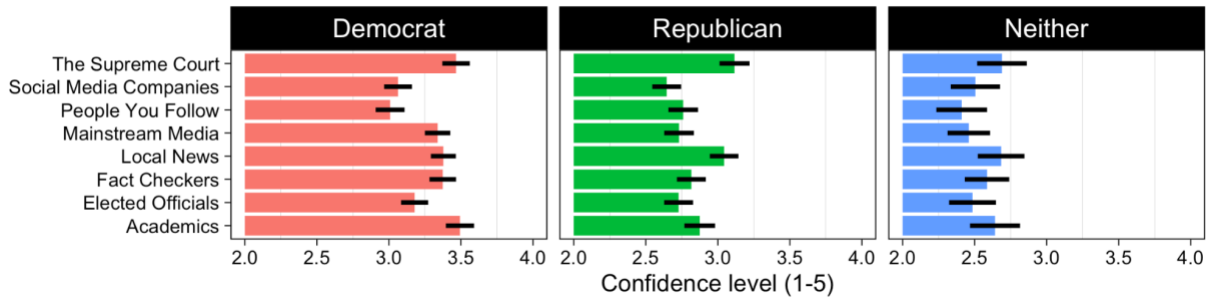


**Figure 4. Non-congenial news consumption by party.** This plot shows the baseline levels of following non-congenial news by different partisans. It shows that Democrats do so slightly more than Republicans (31% vs. 22%) and nearly twice as often as non-partisans or independents (17%). Overall, 27% of respondents answered that they follow social media accounts whose views they disagree with.



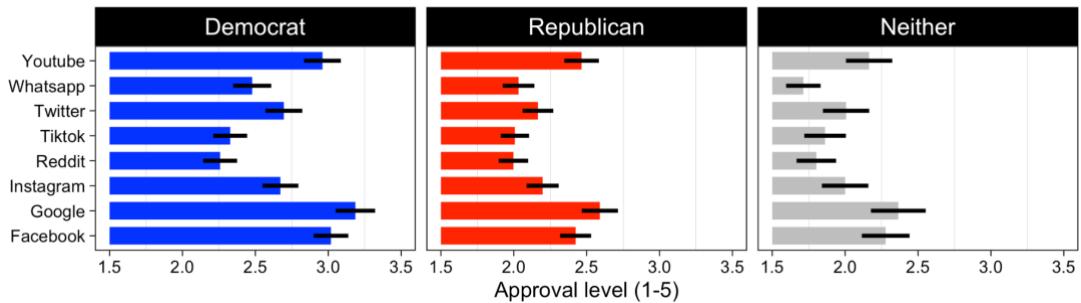
**Figure 5. Frequency of exposure to credibility labels by non-congenial news consumption.** This grid shows percentage (x-axis) of responses in each exposure frequency category (y-axis) for each respondent group (rows) and social media platform (columns).

Confidence in groups to make decisions in public interest

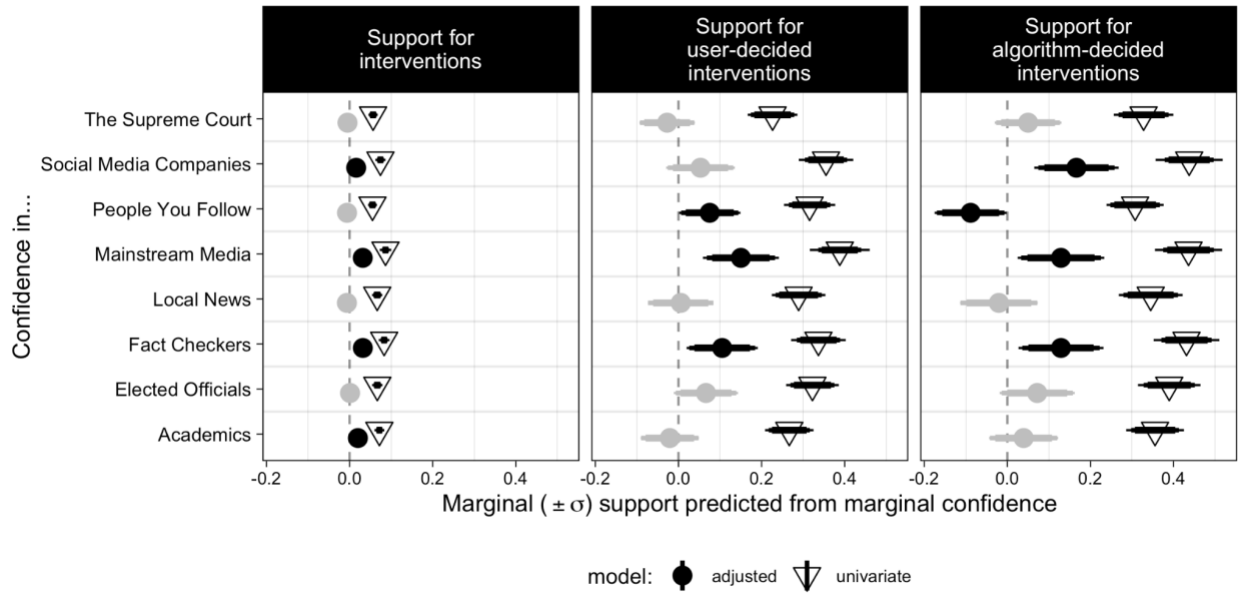


**Figure 6. Trust in specific institutions by party.** This grid shows mean confidence – 1.96 standard errors measured on a 6-point scale (x-axis) in specific American “institutions” (y-axis), originally asked in a series of Pew Research Center surveys (with the addition of People You Follow), for each respondent group (columns), used to create the institutional trust measure used in the paper.

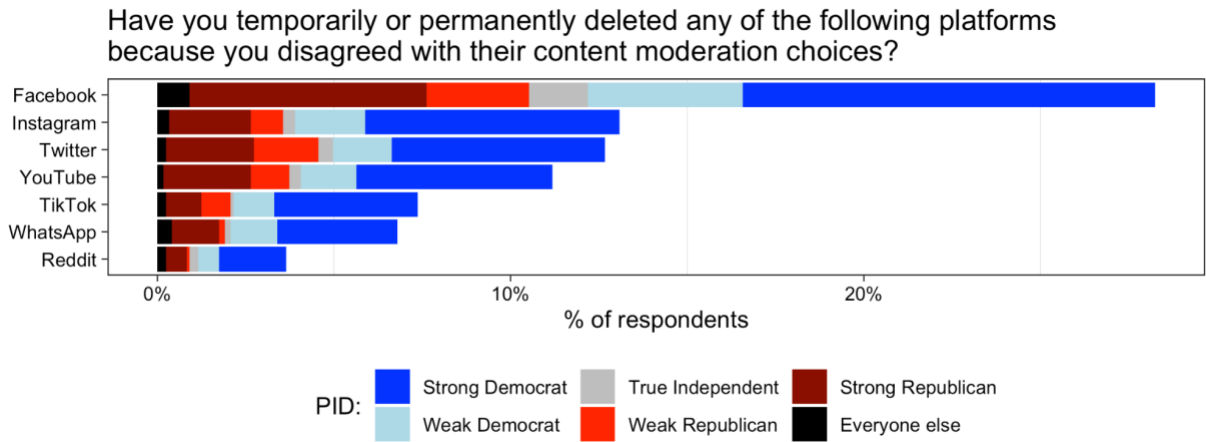
Confidence in social media companies to make decisions in public interest



**Figure 7. Trust in specific social media platforms by party.** This grid shows mean confidence – 1.96 standard errors measured on a 6-point scale (x-axis) in specific social media companies (y-axis), for each respondent group (columns). Across all groups, Google is marginally most trustworthy, while WhatsApp and TikToc are least so, though there are small differences.



**Figure 8. Association between trust in specific social media platforms and support for specific interventions.** This grid shows standardized coefficients predicting the increase in support for specific types of interventions (columns) from increases in support for specific institutions (rows), both measured originally on 5-point scales. Each adjusted model controls for the usual covariates as well as confidence in all other institutions besides the one shown. We see that trust effects are highly correlated across the board, though trust in Mainstream Media, Fact Checkers, and Social Media Companies most strongly and robustly predicts increased support.



**Figure 9. Self-reported rates of temporary or permanent platform deletion by party.**

*Predicting intervention attitudes*

Finally, to place our findings about partisanship (RQ1) and institutional trust (RQ2) in context, we sought to discover the strongest predictors of support for each intervention (downranking, label, removal) out of *all* possible variables including the demographic characteristics adjusted for in our previous models, additional demographic traits, all other disaggregated attitudinal responses (e.g., trust for specific institutions, companies) and self-reported social media experiences (e.g., frequency of usage, exposures to good/bad interventions).

To discover these, we fit an Elastic Net regression model of each support outcome on this predictor

set.<sup>1</sup> Following standard machine learning procedures, we estimated the hyper-parameters of this model via 10-fold cross-validation (or repeating re-fitting and evaluation).

Figure 9 visualizes the strength (coefficient) of the top ten non-zero predictors on support for each specific intervention as well as the aggregated intervention support index that includes other support-related questions. We see that strong Republican partisanship consistently surfaces as the strongest (negative) predictor of support, however institutional trust (particularly in the Mainstream Media) and frequency of social media usage are typically and consistently the next strongest predictors of increased support.

---

<sup>1</sup> The Elastic Net regression model is a regularized modification of linear regression, useful for identifying a selection of high-signal predictors of the outcome variable within a much larger set of predictors. See [“Regularization and variable selection via the elastic net” \(Hastie & Zou 2005\)](#) for details.

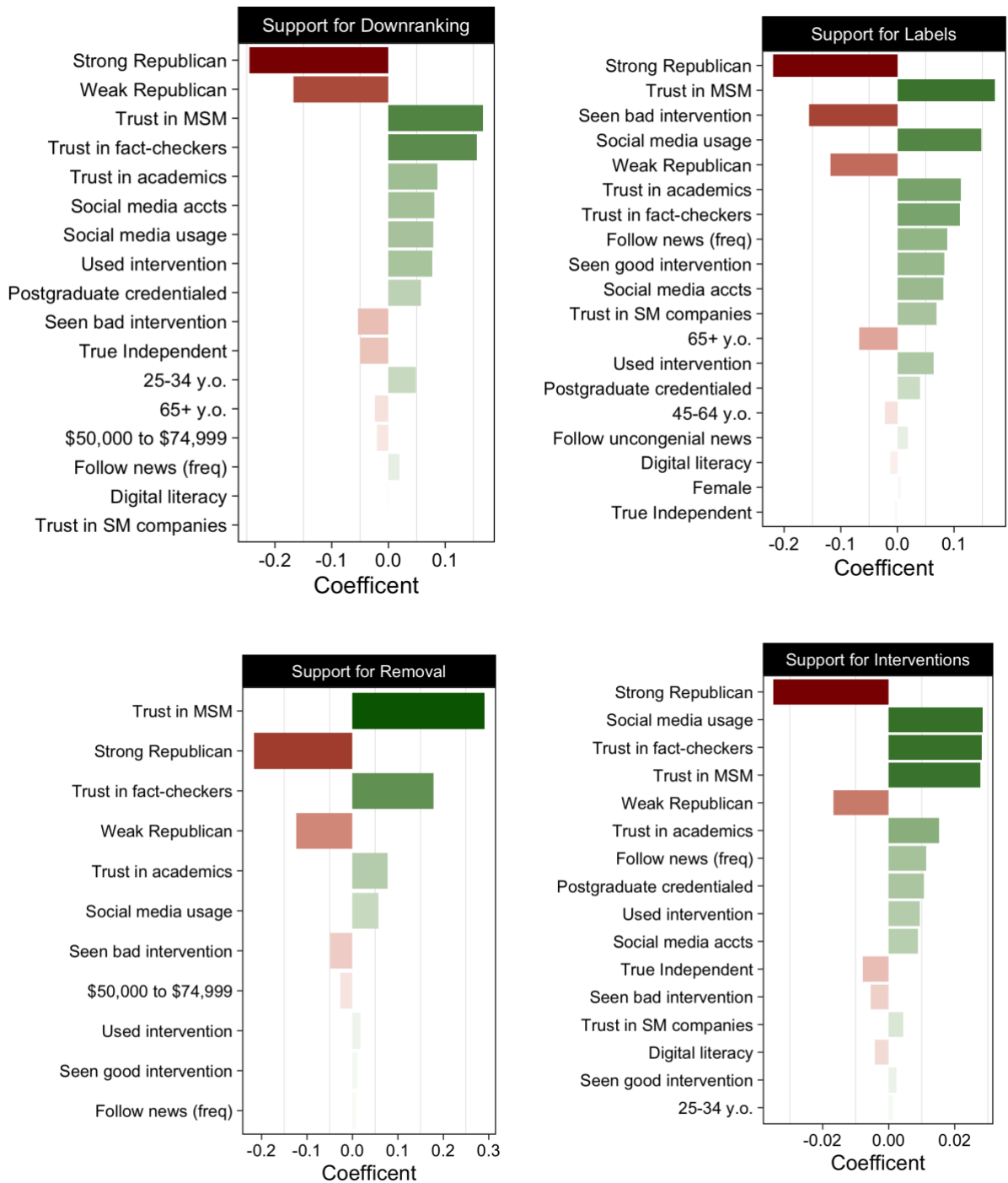


Figure 10. Top non-zero predictors of intervention support in elastic net model.