POLITICAL SCIENCE

# Exposure to ideologically diverse news and opinion on Facebook

Eytan Bakshy,[1*†] Solomon Messing,[1†] Lada A. Adamic[1,2]

**Exposure to news, opinion, and civic information increasingly occurs through social media. How do these online networks influence exposure to perspectives that cut across ideological lines? Using deidentified data, we examined how 10.1 million U.S. Facebook users interact with socially shared news. We directly measured ideological homophily in friend networks and examined the extent to which heterogeneous friends could potentially expose individuals to cross-cutting content. We then quantified the extent to which individuals encounter comparatively more or less diverse content while interacting via Facebook's algorithmically ranked News Feed and further studied users' choices to click through to ideologically discordant content. Compared with algorithmic ranking, individuals' choices played a stronger role in limiting exposure to cross-cutting content.**

Exposure to news and civic information is increasingly mediated through online social networks and personalization (*1*). Information abundance provides individuals with an unprecedented number of options, shifting the function of curating content from newsroom editorial boards to individuals, their social networks, and manual or algorithmic information sorting (*2*–*4*). Although these technologies have the potential to expose individuals to more diverse viewpoints (*4*, *5*), they also have the potential to limit exposure to attitude-challenging information (*2*, *3*, *6*), which is associated with the adoption of more extreme attitudes over time (*7*) and misperception of facts about current events (*8*). This changing environment has led to speculation around the creation of "echo chambers" (in which individuals are exposed only to information from like-minded individuals) and "filter bubbles" (in which content is selected by algorithms according to a viewer's previous behaviors), which are devoid of attitude-challenging content (*3*, *9*). Empirical attempts to examine these questions have been limited by difficulties in measuring news stories' ideological leanings (*10*) and measuring exposure—relying on either error-laden, retrospective self-reports or behavioral data with limited generalizability—and have yielded mixed results (*4*, *9*, *11*–*15*).

We used a large, comprehensive data set from Facebook that allows us to (i) compare the ideological diversity of the broad set of news and opinion shared on Facebook with that shared by individuals' friend networks, (ii) compare this with the subset of stories that appear in individuals' algorithmically ranked News Feeds, and (iii) observe what information individuals choose to consume, given exposure on News Feed. We constructed a deidentified data set that includes 10.1 million active U.S. users who self-report their ideological affiliation and 7 million distinct Web links (URLs) shared by U.S. users over a 6-month period between 7 July 2014 and 7 January 2015. We classified stories as either "hard" (such as national news, politics, or world affairs) or "soft" content (such as sports, entertainment, or travel) by training a support vector machine on unigram, bigram, and trigram text features (details are available in the supplementary materials, section S1.4.1). Approximately 13% of these URLs were classified as hard content. We further limited the set of hard news URLs to the 226,000 distinct hard-content URLs shared by at least 20 users who volunteered their ideological affiliation in their profile, so that we could accurately measure ideological alignment. This data set included ~3.8 billion potential exposures (cases in which an individual's friend shared hard content, regardless of whether it appeared in her News Feed), 903 million exposures (cases in which a link to the content appears on screen in an individual's News Feed), and 59 million clicks, among users in our study.

We then obtained a measure of content alignment (*A*) for each hard story by averaging the ideological affiliation of each user who shared the article. Alignment is not a measure of media slant; rather, it captures differences in the kind of content shared among a set of partisans, which can include topic matter, framing, and slant. These scores, averaged over websites, capture key differences in well-known ideologically aligned media sources: FoxNews.com is aligned with conservatives ($A_s = +.80$), whereas the HuffingtonPost.com is aligned with liberals ($A_s = -0.65$) (additional detail and validation are provided in the supplementary materials, section S1.4.2). We observed substantial polarization among hard content shared by users, with the most frequently shared links clearly aligned with largely liberal or conservative populations (Fig. 1).

The flow of information on Facebook is structured by how individuals are connected in the network. The interpersonal networks on Facebook are different from the segregated structure of political blogs (*16*); although there is clustering according to political affiliation on Facebook, there are also many friendships that cut across ideological affiliations. Among friendships with individuals who report their ideological affiliation in their profile, the median proportion of friendships that liberals maintain with conservatives is 0.20, interquartile range (IQR) [0.09, 0.36]. Similarly, the median proportion of friendships that conservatives maintain with liberals is 0.18, IQR [0.09, 0.30] (Fig. 2).

How much cross-cutting content individuals encounter depends on who their friends are and what information those friends share. If individuals acquired information from random others, ~45% of the hard content that liberals would be exposed to would be cross-cutting, compared with 40% for conservatives (Fig. 3B). Of course, individuals do not encounter information at random in offline environments (*14*) nor on the Internet (*9*). Despite the slightly higher volume of conservatively aligned articles shared (Fig. 1), liberals tend to be connected to fewer friends who share information from the other side, compared with their conservative counterparts: Of the hard news stories shared by liberals' friends, 24% are cross-cutting, compared with 35% for conservatives (Fig. 3B).

The media that individuals consume on Facebook depends not only on what their friends share but also on how the News Feed ranking

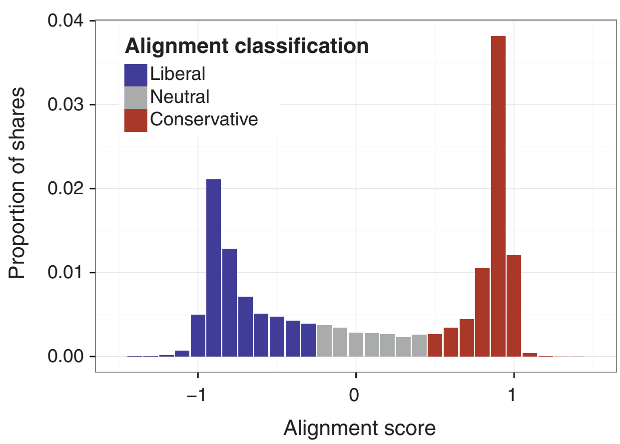



**Fig. 1. Distribution of ideological alignment of content shared on Facebook measured as the average affiliation of sharers weighted by the total number of shares.** Content was delineated as liberal, conservative, or neutral on the basis of the distribution of alignment scores (details are available in the supplementary materials).

[1]Facebook, Menlo Park, CA 94025, USA. [2]School of Information, University of Michigan, Ann Arbor, MI, USA.
*Corresponding author. E-mail: ebakshy@fb.com †These authors contributed equally to this work.

algorithm sorts these articles and what individuals choose to read (Fig. 3A). The order in which users see stories in the News Feed depends on many factors, including how often the viewer visits Facebook, how much they interact with certain friends, and how often users have clicked on links to certain websites in News Feed in the past. We found that after ranking, there is on average slightly less cross-cutting content: The risk ratio comparing the probability of seeing cross-cutting content relative to ideologically consistent content is 5% for conservatives and 8% for liberals (supplementary materials, section S1.7).

Individual choice futher limits exposure to ideologically cross-cutting content. After adjusting for the effect of position [the click rate on a link is negatively correlated with its position in the News Feed (fig. S5)], we estimated the risk ratio comparing the likelihood that an individual clicks on a cross-cutting content relative to a consistent content to be 17% for conservatives and 6% for liberals, a pattern that is consistent with prior research (*4*, *17*). Despite these tendencies, there is substantial room for individuals to consume more media from the other side; on average, viewers clicked on 7% of hard content available in their feeds.

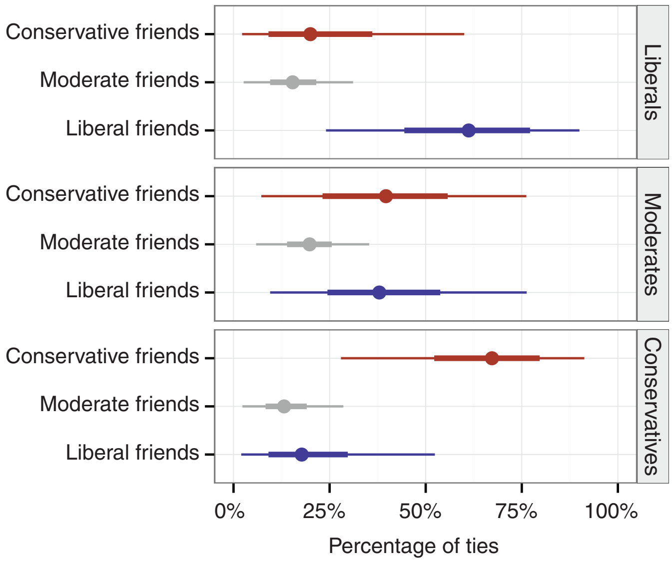



**Fig. 2. Homophily in self-reported ideological affiliation.** Proportion of links to friends of different ideological affiliations for liberal, moderate, and conservative users. Points indicate medians, thick lines indicate interquartile ranges, and thin lines represent 10th to 90th percentile ranges.

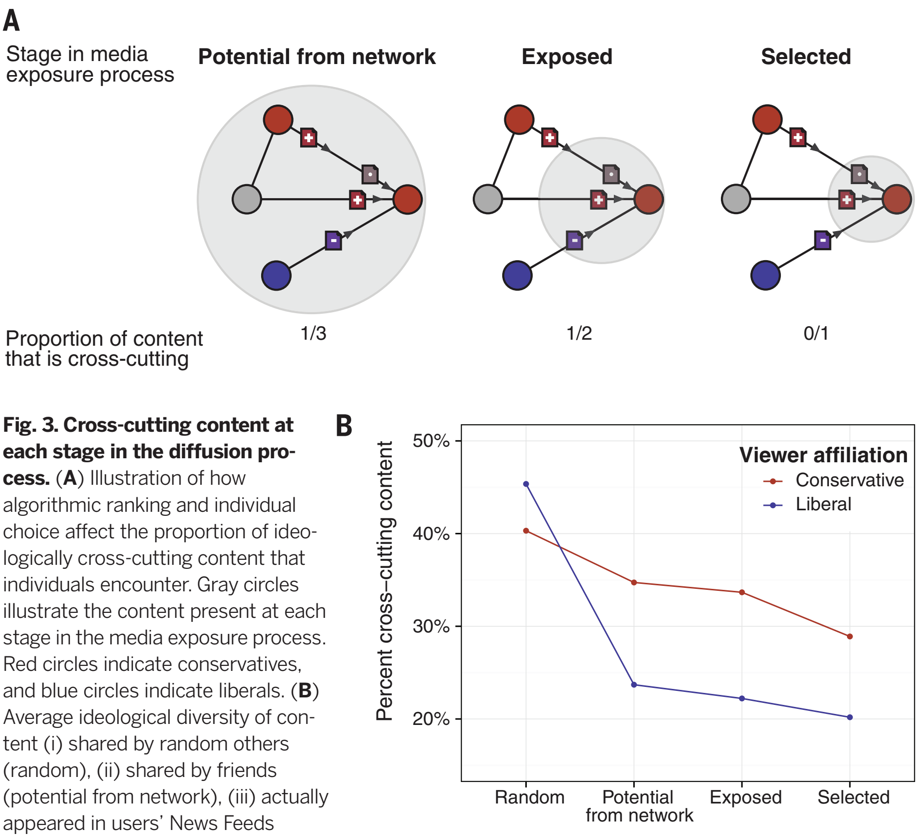



**Fig. 3. Cross-cutting content at each stage in the diffusion process.** (**A**) Illustration of how algorithmic ranking and individual choice affect the proportion of ideologically cross-cutting content that individuals encounter. Gray circles illustrate the content present at each stage in the media exposure process. Red circles indicate conservatives, and blue circles indicate liberals. (**B**) Average ideological diversity of content (i) shared by random others (random), (ii) shared by friends (potential from network), (iii) actually appeared in users' News Feeds (exposed), and (iv) users clicked on (selected).

Our analysis has limitations. Although the vast majority of U.S. social media users are on Facebook (*18*), our study is limited to active users who volunteer an ideological affiliation on this social media platform. Facebook's users tend to be younger, more educated, and more often female as compared with the U.S. population as a whole (*18*). Other forms of social media, such as blogs or Twitter, have been shown to exhibit different patterns of homophily among politically interested users, largely because ties tend primarily to form based on common topical interests and/or specific content (*16*, *19*), whereas Facebook ties primarily reflect many different offline social contexts: school, family, social activities, and work, which have been found to be fertile ground for fostering cross-cutting social ties (*20*). In addition, our distinction between exposure and consumption is imperfect; individuals may read the summaries of articles that appear in the News Feed and therefore be exposed to some of the articles' content without clicking through.

This work informs long-standing questions about how media exposure is shaped by our social networks. Although partisans tend to maintain relationships with like-minded contacts [which is consistent with (*21*)], on average more than 20% of an individual's Facebook friends who report an ideological affiliation are from the opposing party, leaving substantial room for exposure to opposing viewpoints (*22*, *23*). Furthermore, in contrast to concerns that people might "listen and speak only to the like-minded" while online (*6*), we found exposure to cross-cutting content (Fig. 3B) along a hypothesized route: traditional media shared in social media (*4*, *24*). Perhaps unsurprisingly, we show that the composition of our friend networks is the most important factor limiting the mix of content encountered in social media. The way that sharing occurs within these networks is not symmetric: Liberals tend to be connected to fewer friends who share conservative content than are conservatives (who tend to be linked to more friends who share liberal content).

Within the population under study here, individual choices (*2*, *13*, *15*, *17*) more than algorithms (*3*, *9*) limit exposure to attitude-challenging content in the context of Facebook. Despite the differences in what individuals consume across ideological lines, our work suggests that individuals are exposed to more cross-cutting discourse in social media than they would be under the digital reality envisioned by some (*2*, *6*). Rather than people browsing only ideologically aligned news sources or opting out of hard news altogether, our work shows that social media expose individuals to at least some ideologically cross-cutting viewpoints (*4*). Of course, we do not pass judgment on the normative value of cross-cutting exposure. Although normative scholars often argue that exposure to a diverse "marketplace of ideas" is key to a healthy democracy (*25*), a number of studies have found that exposure to cross-cutting viewpoints is associated with lower levels of political participation (*22*, *26*, *27*). Regardless, our work suggests that the power to expose oneself to perspectives from the other side in social media lies first and foremost with individuals.

## REFERENCES AND NOTES


1. K. Olmstead, A. Mitchell, T. Rosenstiel, *Navigating news online*. Pew Research Center (2011); available at www.journalism.org/analysis_report/navigating_news_online.
2. W. L. Bennett, S. Iyengar, *J. Commun.* **58**, 707–731 (2008).
3. E. Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (Penguin Press, London, 2011).
4. S. Messing, S. J. Westwood, *Communic. Res.* **41**, 1042–1063 (2012).
5. E. Bakshy, I. Rosenn, C. Marlow, L. Adamic, *Proc. 21st Int. Conf. World Wide Web Pages* **1201.4145** (2012).
6. C. R. Sunstein, *Republic.com 2.0* (Princeton Univ. Press, Princeton, NJ, 2007).
7. N. Stroud, *Polit. Behav.* **30**, 341–366 (2008).
8. S. Kull, C. Ramsay, E. Lewis, *Polit. Sci. Q.* **118**, 569–598 (2003).
9. S. Flaxman, S. Goel, J. M. Rao, "Ideological segregation and the effects of social media on news consumption," SSRN Scholarly Paper ID 2363701, Social Science Research Network, Rochester, NY (2013).
10. T. Groeling, *Annu. Rev. Polit. Sci.* **16**, 129–151 (2013).
11. M. Gentzkow, J. M. Shapiro, *Q. J. Econ.* **126**, 1799–1839 (2011).
12. M. J. LaCour, "A balanced information diet, not echo chambers: Evidence from a direct measure of media exposure," SSRN Scholarly Paper ID 2303138, Social Science Research Network, Rochester, NY (2013).
13. E. Lawrence, J. Sides, H. Farrell, *Perspect. Polit.* **8**, 141 (2010).
14. D. O. Sears, J. L. Freedman, *Public Opin. Q.* **31**, 194 (1967).
15. N. A. Valentino, A. J. Banks, V. L. Hutchings, A. K. Davis, *Polit. Psychol.* **30**, 591–613 (2009).
16. L. A. Adamic, N. Glance, in *Proceedings of the 3rd International Workshop on Link Discovery* (ACM, New York, 2005), pp. 36–43.
17. S. Iyengar, K. S. Hahn, *J. Commun.* **59**, 19–39 (2009).
18. M. Duggan, A. Smith, "Social media update 2013," Pew Research Center (2013); available at www.pewinternet.org/2013/12/30/social-media-update-2013.
19. M. D. Conover, J. Ratkiewicz, M. Francisco, B. Gonçalves, A. Flammini, F. Menczer, Political polarization on Twitter. *Fifth International AAAI Conference on Weblogs and Social Media* (2011).
20. D. C. Mutz, J. J. Mondak, *J. Polit.* **68**, 140 (2006).
21. S. Goel, W. Mason, D. J. Watts, *J. Pers. Soc. Psychol.* **99**, 611–621 (2010).
22. D. C. Mutz, *Am. J. Polit. Sci.* **46**, 838–855 (2002).
23. B. Bishop, *The Big Sort: Why the Clustering of Like-Minded America Is Tearing Us Apart* (Houghton Mifflin Harcourt, New York, 2008).
24. D. C. Mutz, P. S. Martin, *Am. Polit. Sci. Rev.* **95**, 97 (2001).
25. T. Mendelberg, *Deliber. Particip.* **6**, 151–193 (2002).
26. R. Huckfeldt, J. M. Mendez, T. Osborn, *Polit. Psychol.* **25**, 65–95 (2004).
27. R. Bond, S. Messing, *Am. Polit. Sci. Rev.* **109**, 62–78 (2015).


## ACKNOWLEDGMENTS


We thank J. Bailenson, D. Eckles, A. Franco, K. Garrett, J. Grimmer, S. Iyengar, B. Karrer, C. Nass, A. Peysakhovich, S. Taylor, R. Weiss, S. Westwood, J. M. White, and anonymous reviewers for their valuable feedback. The following code and data are archived in the Harvard Dataverse Network, http://dx.doi.org/10.7910/DVN/LDJ7MS: "Replication Data for: Exposure to Ideologically Diverse News and Opinion on Facebook"; R analysis code and aggregate data for deriving the main results (tables S5 and S6); Python code and dictionaries for training and testing the hard-soft news classifier; aggregate summary statistics of the distribution of ideological homophily in networks; and aggregate summary statistics of the distribution of ideological alignment for hard content shared by the top 500 most shared websites. The authors of this work are employed and funded by Facebook. Facebook did not place any restrictions on the design and publication of this observational study, beyond the requirement that this work was to be done in compliance with Facebook's Data Policy and research ethics review process (www.facebook.com/policy.php).


### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/348/6239/1130/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S6
References (*28–35*)


20 October 2014; accepted 27 April 2015
Published online 7 May 2015;
10.1126/science.aaa1160


ECOPHYSIOLOGY

# Climate change tightens a metabolic constraint on marine habitats


**Curtis Deutsch,[1]* Aaron Ferrel,[2]† Brad Seibel,[3] Hans-Otto Pörtner,[4] Raymond B. Huey[5]**



**Warming of the oceans and consequent loss of dissolved oxygen ($O_2$) will alter marine ecosystems, but a mechanistic framework to predict the impact of multiple stressors on viable habitat is lacking. Here, we integrate physiological, climatic, and biogeographic data to calibrate and then map a key metabolic index—the ratio of $O_2$ supply to resting metabolic $O_2$ demand—across geographic ranges of several marine ectotherms. These species differ in thermal and hypoxic tolerances, but their contemporary distributions are all bounded at the equatorward edge by a minimum metabolic index of ~2 to 5, indicative of a critical energetic requirement for organismal activity. The combined effects of warming and $O_2$ loss this century are projected to reduce the upper ocean's metabolic index by ~20% globally and by ~50% in northern high-latitude regions, forcing poleward and vertical contraction of metabolically viable habitats and species ranges.**


Climate change is altering ecosystems by shifting distributions, phenologies, and interactions among species, but understanding how these changes are caused by climatic influences on physiology and fitness remains a challenge (*1*). In the ocean, increased metabolic rates due to rising temperatures will be accompanied by declines in dissolved $O_2$, potentially restricting organismal aerobic capacities (*2–4*). The physiology of hypoxic and thermal tolerance of marine species is well understood (*3*, *5–7*). Lacking, however, is a general mechanistic model that quantifies how $O_2$ and temperature jointly restrict large-scale biogeographic distributions now and in the future. Here, we combine laboratory and field data to demonstrate that temperature and $O_2$ together limit the contemporary ranges of marine ectotherms and to derive empirically based estimates of habitat loss in the warmer and less oxygenated oceans projected by this century's end.

For marine habitats to be metabolically viable, the environmental $O_2$ supply rate ($S$) must exceed an animal's resting metabolic demand ($D$). The rate of $O_2$ supply increases with ambient $O_2$ pressure ($PO_2$) and with respiratory efficacy (*8*). Thus, $S = \alpha_S B^\delta PO_2$, where respiratory efficacy is the product of $\alpha_s$, a per-mass rate of gas transfer between water and animal and its scaling with body mass, $B^\delta$. Resting metabolic demand also scales with $B$ and with absolute temperature ($T$), according to $D = \alpha_D B^\varepsilon \exp(-E_o/k_B T)$, where $\alpha_D$ is a taxon-specific baseline metabolic rate, $\varepsilon$ is its allometric scaling, $E_o$ is its temperature dependence, and $k_B$ is Boltzmann's constant (*9*).

We define a metabolic index, denoted $\Phi$, as the ratio of $O_2$ supply to an organism's resting $O_2$ demand

$$\Phi = A_o B^n \frac{PO_2}{\exp(-E_o/k_B T)} \quad (1)$$

where $A_o = \alpha_s/\alpha_d$ is the ratio of rate coefficients for $O_2$ supply and metabolic rate, and $n$ is the difference between the respective allometric scalings ($n = \delta - \varepsilon$). If $\Phi$ falls below a critical threshold value of 1, organisms must either suppress aerobic activity (*5*) or initiate anaerobic metabolism, conditions that are physiologically unsustainable. Conversely, values above 1 enable organismal metabolic rates to increase by a factor of $\Phi$ above resting levels, permitting critical activities such as feeding, defense, growth, and reproduction. Thus, for a given environment, $\Phi$ estimates the ratio of maximum sustainable metabolic rate to the minimum rate necessary for maintenance for a given species.

We analyzed data from published studies in which hypoxia tolerance was determined at


[1]School of Oceanography, University of Washington, Seattle, WA 98195, USA. [2]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA 90095, USA. [3]Biological Sciences Department, University of Rhode Island, Kingston, RI 02881, USA. [4]Alfred Wegener Institute, D-27570 Bremerhaven, Germany. [5]Department of Biology, University of Washington, Seattle, WA 98195, USA.
***Corresponding author. E-mail: cdeutsch@uw.edu** †Present address: Los Angeles Unified School District, Los Angeles, CA 90085, USA.

Correction: 28 May 2015

www.sciencemag.org/cgi/content/full/science.aaa1160/DC1

# Supplementary Material for

## Exposure to ideologically diverse news and opinion on Facebook

Eytan Bakshy,* Solomon Messing, Lada Adamic

*Corresponding author. E-mail: ebakshy@gmail.com

Published 7 May 2015 on *Science* Express
DOI: 10.1126/science.aaa1160

**This PDF file includes:**

Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S6
Full Reference List

**Correction:** Several edits to text were made for clarification purposes in sections 1.1, 1.4, and 1.7 and in the Table S6 caption.

# Supporting Materials for Exposure to Ideologically Diverse News and Opinion on Facebook

Eytan Bakshy* Solomon Messing
Lada Adamic
*Corresponding author. E-mail: eytan@fb.com

**This PDF file includes:**

Materials and Methods

Supplementary Text

Fig. S1-S10

Tables S1 to S7

Full Reference List

**Other Supplementary Material for this manuscript includes the following:**

The following code and data are archived in the Harvard Dataverse Network,

http://dx.doi.org/10.7910/DVN/LDJ7MS,

"Replication Data for: Exposure to Ideologically Diverse News and Opinion on Facebook".

R analysis code and aggregate data for deriving the main results (e.g., Table S5, S6)

Python code and dictionaries for training and testing the hard-soft news classifier

Aggregate summary statistics of the distribution of ideological homophily in networks

Aggregate summary statistics of the distribution of ideological alignment for hard content shared by the top 500 most shared websites.

# 1 Materials and Methods

## 1.1 Population

To construct our population, we consider active U.S. adults on Facebook who report their political affiliation. This includes U.S. Facebook users who are 18 or older, log in at least 4/7 days per week (i.e., 105/185 days during study period, July 7, 2014-January 7, 2014; this removes approximately 30% of users). We limit this population further to those who self-report their ideological affiliation, which comprises 25% of active U.S. adults as defined above. We also limit the population to those who have clicked on at least one link shared on Facebook that we classified as hard news/opinion over the course of the study. The final population includes 10.1 million users. See Table S1.

## 1.2 Political designations

All Facebook users can self-report their political affiliation; 9% of U.S. users over 18 do. We mapped the top 500 political designations on a five-point, -2 (Very Liberal) to +2 (Very Conservative) ideological scale; those with no response or with responses such as "other" or "I don't care" were not included. 46% of those who entered their political affiliation on their profiles had a response that could be mapped to this scale. We validated a sample of these labels against a survey of 79 thousand U.S. users in which we asked for a 5-point very-liberal to very-conservative ideological affiliation; the Spearman rank correlation between the survey responses and our labels was 0.78.

## 1.3 Ideological homophily

Individuals vary with respect to how many friends they have of different affiliations. Figure S1 shows the proportion of ties to friends of different ideological affiliations for each user. Each panel indicates a focal user of a particular affiliation, and the distributions show the kernel density estimate of the percentage of ties to friends of different affiliations. Vertical lines indicate medians. Both conservatives and liberals exhibit substantial homophily in friend networks, while moderates maintain a similar distribution of ties to both liberal and conservative friends. See also Table S2.

## 1.4 Hard news and opinion data

The final dataset used in the analysis of the main text includes 226,310 hard news and opinion stories (which we refer to simply as "hard content" below). Because we wish to classify each article according to its ideological alignment, we only consider stories that had been shared by at least 20 U.S. users who self-report a mappable ideological affiliation. This set of links includes over 90% of all URLs labeled as hard content seen by individuals in our study. We describe the procedure used to classify hard content below.

### 1.4.1 Hard-soft classification

We build our hard-soft classifier using an approach often referred to in the Natural Language Processing literature as "bootstrapping" [28, 29, 30, 31] which entails using regular expressions to build a set of training labels (and should not be confused with bootstrapping in statistics).

We begin with URL content shared by at least 100 U.S. users. To extract features from the documents in question (text summaries of the articles sent to Facebook when a user shares content from an external website), we apply English stopwords; tokenize using unigrams, bigrams, and trigrams; and use tokens that have occurred in at least 2 and no more than half of all documents.

To construct positive training labels, we use stories from 81 of the most shared news sites on Facebook in 2012, among them *nytimes.com, foxnews.com, cnn.com*, and *latimes.com*, that contain indicators of explicit hard news/opinion topics in the URL (see `FitSVM.py` in replication materials for a complete list). These include the following strings: “politi”, “usnews”, “world”, “national”, “state”, “elect”, “vote”, “govern”, “campaign”, “war”, “polic”, “econ”, “unemploy”, “racis”, “energy”, “abortion”, “educa”, “healthcare”, “immigration.” We construct negative training cases by matching the URL to the following strings “sports”, “entertainment”, “arts”, “fashion”, “style”, “lifestyle”, “leisure”, “celeb”, “movie”, “music”, “gossip”, “food”, “travel”, “horoscope”, “weather”, “gadget.” Our training data consisted of 147,958 stories; of which 114,121 were labeled soft news and 33,837 were labeled hard news/opinion.

We then trained our classifier using a linear SVM with the standard L2 penalty and hinge loss using SciKitLearn. We classified 694,989 URLs as hard content and 6,929,907 as soft content in the complete set. Our classifier achieves ten-fold cross-validated accuracy of 97.1 percent. Limiting this set to URLs shared by at least 20 affiliated users yields 226,310 hard news and opinion stories, which we use in our analysis.

#### 1.4.2 Measuring alignment

We measure the ideological *alignment* of these items, which constitutes a behavioral indicator of the extent to which partisan identifiers actively share content. We derive the alignment score $A_u$ of an individual URL by averaging the political alignment of the set of people $P_u$ who share the URL.

$$A_u = \frac{1}{|P_u|} \sum_{i \in P_u} a_i \tag{1}$$

As shown in Figure S2, the measure indicates substantial polarization among the content shared by individuals who provide an ideological affiliation, with the most frequently shared URLs coming from sources at the ends of the distribution of alignment. We note that while the alignment scores of URLs classified as hard content exhibit strong alignment with either the right or left, soft content generally does not show strong patterns of ideological separation (see

Figure S2a). We take quintiles of the measure and color the left- and rightmost quantiles blue and red, respectively.

Alignment scores of URLs, averaged over their respective websites, have substantial face validity. We note a few highly shared and well-known media sources in Table S3.

#### 1.4.3 Validation

We validate this metric against other efforts to quantify ideology in media content. To produce these estimates, we matched domains from our study to those listed in three other studies: [32], [33], and [34].

We first compare our alignment measure to recent work by Budak et al. (2014) that utilizes crowd-sourced content analysis to quantify media bias in commonly visited online news sources. By employing humans to manually annotate articles for media bias, then using machine learning to infer ideology based on the vocabulary used in unseen news articles, Budak et al. generate measures of political slant. We matched all of the 15 sources in Budak et al. to domain-level alignment scores in our data and find the two scores to be highly correlated (Pearson's $\rho = 0.91$, 95% CI = $[0.750.97]$, see also Figure S3a).

Next, we examine the correlation between our measure and that of Groseclose and Milyo (2005), which leverages Americans for Democratic Action (ADA) ratings of the ideology of organizations such as Washington think tanks that news organizations cite as news sources [33]. We match 17 of the 20 sources in Groseclose and Milyo (matching hard news broadcasts to the same organization's website when the news source constituted a television show) to domains in our data; the correlation is -0.47 (Pearson's $\rho$, 95% CI = $[-0.78, 0.01]$), see also Figure S3c).

Finally, we validate our measure of alignment against Gentzkow and Shapiro's (2010) measure of media *slant*, which models the similarity between the language used by Democrats and Republicans in congressional proceedings and local newspapers in each congressional district

[34]. Of the 435 local papers scored by Gentzkow and Shapiro, we consider the 20 sources most shared on Facebook (this includes 4.6 percent of hard news/opinion shares from U.S. Facebook users; the remaining domains comprise just one percent). The correlation between alignment and Gentzkow and Shapiro's slant is $0.56$ (Pearson's $\rho$, 95% CI = $[0.15, 0.80]$), see also Figure S3b).

## 1.5 Quantifying cross-cutting content

To measure whether content is cross-cutting, we then take the quintiles of the URL alignment score create five alignment categories: content shared by audiences that are on balance primarily liberal (-2), somewhat liberal (-1), bipartisan (0), somewhat conservative (1), and primarily conservative (2) (see also Figure S2). Finally, we show the proportion of stories published on websites of interest broken down by each alignment categories (Table S4).

## 1.6 Alignment and unaffiliated sharers

Our measure of content alignment uses data from individuals who self-report their political affiliation. It is important to note that many link shares are from users who do not report their political affiliation (Figure S4). Furthermore, both moderates and those who do not self-report their affiliation tend to share more liberally aligned content than conservative. Because liberals are more connected on Facebook (see Table S1) and because more individuals across the graph share liberal content, there is a higher likelihood that a given individual's friends have shared liberal content.

## 1.7 Relative risk in exposure probabilities

The main text demonstrates that individuals have the potential to be and are exposed to ideologically cross-cutting content. Although we do not identify the causal effects of a URL being cross-cutting on whether a user will see that URL in the New Feed, or select that URL if presented in the News Feed, it is still informative to look at the relative differences in the probability of exposure for ideologically consistent versus cross-cutting content. We summarize these relative differences in terms of risk ratios, e.g.,

$$1 - \frac{\text{Pr(click }|\text{exposed, cross-cutting)}}{\text{Pr(click }|\text{exposed, not cross-cutting)}}.$$

Such expressions can also be written in terms of the proportion of URLs that are cross-cutting at each stage in the consumption process. For example, the above equation is mathematically equivalent to $\frac{\pi_e(1-\pi_n)}{\pi_n(1-\pi_e)} - 1$, where $\pi_e$ and $\pi_n$ are the proportion of actual News Feed exposures and potential exposures (articles shared by friends), respectively, that are ideologically cross-cutting. We present these raw proportions in Table S5.

Table S6 presents these quantities at each stage: the relative likelihood of friends sharing cross-cutting content compared to the probability of sharing ideologically consistent content (given it was shared on Facebook at all), the relative likelihood an individual encounters cross-cutting content compared to consistent content in the News Feed given that friends' shared it, and the relatively likelihood an individual clicks on cross-cutting content compared to consistent content given exposure in the News Feed. In addition, we include a position-adjusted change for the selection probabilities (described in S1.8, below).

## 1.8 Position effects

In this section, we describe how the order in which the News Feed displays items could affect what information individuals are exposed to and select. In particular, we show the click rate on a link is negatively correlated with its position in the News Feed. Ignoring this variation could lead one to attribute differences in individual choice to differences induced by the feed ranking algorithm via ordering. We derive an adjusted estimate of the difference in the probability of selecting content, and show the sensitivity of the ratio reported in the main text to potential position effects.

If the News Feed ranking algorithm differentially placed ideologically cross-cutting content in lower positions (or higher positions), and being placed in a different position had a causal effect on click rates, then part of the observed difference in click-through rates for ideologically similar vs cross-cutting content could be explained by position effects alone.

We investigate this possibility by examining how exposure to ideologically diverse content and click through rates vary by position. Exposures to friends' link sharing behaviors are logged when viewers load the Facebook News Feed and the story renders in the visible portion of their Web browser or mobile device (i.e., a "validated viewport view").

Let $y_{ij} = 1$ when an individual $i$ selects (clicks) a story $j$, and 0 otherwise. Let $z_{ij}$ indicate whether a story appears on the individual's screen for more than 250 milliseconds, such that $z_{ij} = 1$ when $i$ has a viewport view for story $j$, and $0$ otherwise. Finally, we define $C(i, j)$ to be $1$ when $j$ is ideologically cross-cutting for user $i$, and $0$ otherwise. We denote the total number of items seen by $i$, conditional on whether or not the content is crosscutting ($c$) as $n_i(c) = \sum_{j=1}^{M} z_{ij} C(i, j)$.

We then define $\bar{y}(c)$ to be the average probability that an individual clicks on an article, conditional on whether or not it is cross-cutting:

$$\bar{y}(c) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_i(c)} \sum_{j=1}^{M} y_{ij} z_{ij} C(i, j),$$

where $N$ is the total number of individuals and $M$ is the total number of distinct URLs that viewers could potentially be exposed to. Let the difference in probability of clicking on a link for ideologically similar and cross-cutting content be $\delta = \bar{y}(0) - \bar{y}(1)$.

Most users visit Facebook multiple times per day, so the same story may appear in different positions at different points throughout the day. We map each viewer-URL pair to the highest (i.e., minimum) position that a URL-share story rendered in each viewers' News Feed. We then examine the relationship between this position, ideological diversity, and click-through rates in Figure S5.

From Figure S5a, one can see that there is a strong correlation between position and click-through rates, and that furthermore, a slightly lower proportion of cross-cutting stories appear in positions toward the top of users' News Feeds for liberals, while conservatives encounter a slightly higher proportion of cross-cutting stories toward the top of their feeds (Figure S5b). Some positions—particularly the second position of the News Feed—are often allocated to sponsored content, which may include links to articles shared by friends which are associated with websites associated with a particular advertiser. Since we aim to characterize interactions with all hard content shared by friends, such links are included in our analyses. These links appear to be more ideologically consistent with the viewers; however further investigation is beyond the scope of this work. The decreasing relationship between click-through rate and position, combined with differences in ideological alignment that vary by position, could obscure the relative contribution of algorithmic ranking and selective exposure to decreased levels of ideologically cross-cutting news consumption.

One approach to adjust for the potential imbalance in story position is to stratify the difference in selection probabilities for ideologically similar vs cross-cutting content based on position. Using $n_{ip}$ to denote the total number of stories for user $i$ that appear in position $p$, $n_{i\cdot}$ to denote the total number of stories seen by a user, and $n_{\cdot\cdot}$ as the total number of stories seen by all users, we express the difference in click rates conditional on position as,

$$\tilde{\delta}_p = \bar{y}(0) - \bar{y}(1) \mid P = p,$$

And then write the stratified estimator of the $\delta$ as:

$$\tilde{\delta} = \frac{1}{n_{..}} \sum_{p=1}^{\mathrm{P}} n_{.p} \tilde{\delta}_p$$

As stated above, we provide these stratified estimates in Table **??**.

# 2 Supplementary Text

## 2.1 Facebook usage and exposure to cross-cutting content

Because the Facebook News Feed orders content shared by friends, the potential effects of ranking on exposure to cross-cutting content might be lower for highly active users who view many stories because these individuals scroll further down and see more of their friends' content. We investigate this possibility by examining the raw proportion of cross-cutting content that individuals could have potentially been exposed to, based on what their friend networks share, what they are exposed to via the Facebook News Feed, and what they select as a function of the number of stories viewed (Figure S6). We transform the number of stories viewed into deciles, using the distribution of stories viewed over all users, for ease of interpretation.

For conservatives, there is very little change in the diversity of content when moving from potential to exposed at each activity level. However, despite the lower proportion of cross-cutting content active conservatives encounter in News Feed, they select it at similar rates. It is possible that more active conservatives may be more open to consuming content from the other side, being younger and more female (both of which are associated with higher levels of openness [35]). There are other possible causes for this pattern, but a full analysis is beyond the scope of this work. Liberals, on the other hand, see less content from the other side, without much consistent variation related to activity.

## 2.2 Proportion of Individuals Exposed to Cross-Cutting Content

We also present the fraction of individuals in our sample who encounter at least one cross-cutting and aligned item at each stage, among those whose friends have shared said content (Figure S7). This provides an indication of the effects of algorithmic ranking and selectivity among those who are on the margins of being exposed to no cross-cutting content at all and allows comparisons between each stage with a common denominator (the total number of individuals in the sample). The figure shows that among those at the margin, selection choices play a greater role in determining whether individuals encounter ideologically cross-cutting content, compared to algorithmic ranking.

## 2.3 Sample Data

## 2.4 Classifications of political and nonpolitical URLs

A random sample of the text utilized by the hard-soft URL classifier. 20 URLs classified as hard-, and 20 classified as soft content, drawn from completely out of sample cases (e.g., URLs that were not in the training or test sets), are provided. When a user shares a URL from a news website, blog, or other site, these text summaries are sent to Facebook from the external website. They have been truncated and stripped of some types of punctuation.

### 2.4.1 Hard news and opinion content

- The government keeps backups of every federal record ever But it won't turn over Lois Lerner's missing emails ...

- On Sept 25th Governor Jerry Brown signed into law SB 1135 Prison Anti Sterilization bill authored by Senator Hannah Beth Jackson, sponsored by legal and human rights organization Justice Now and included bi partisan co authorship The bill went before the Governor after passing unanimously out ...

- COURT HALTS EXECUTION OF MENTALLY ILL TEXAS INMATE USA TODAY 1 32 P M WEDNESDAY, DECEMBER 3RD, 2014 AUSTIN A federal appeals court in New Orleans on Wednesday halted the execution of Texas killer Scott Panetti, whose case has spark...
- Foreigners love oil, gold, diamonds, and cheap labor of Islamic world They like the quarrels of the Middle East Believe me, they don t like us ...
- So far, the debate over the proposed Islamic center near Ground Zero has unfolded along predictable lines, with the man at the center of the project, Imam Feisal Abdul Rauf, drawing attacks from the right painting him as a terrorist sympathizer with ties to Hamas and the Muslim Brotherhood ...
- A Philadelphia police officer was caught on video cursing and threatening a teenager The video was posted on Facebook October 17, showing the unidentified officer following a teen boy as he walked home with ...
- The attack sent terrorized villagers fleeing into the bush in search of safety ...
- Philadelphia Police are looking for a woman who was seen on surveillance video stealing Halloween decorations from a home in South Philadelphia ...
- Malaysian Airlines fight MH17 carrying 295 passenger shot down by missile in Ukraine's east, Interfax reports ...
- As Congress and the White House mull a federal response to the surge of illegal immigration on the southwest border, here s a list of four Democratic governors who support letting illegal i...

- Pennsylvania Governor Tom Corbett on Tuesday signed into law a prisoner gag order that rights groups say is an affront to the First Amendment and a denial of all citizens' right to understand "an area of U S life physically removed from public scrutiny " More than 50 House Republicans have signed on to the strategy ...

- Australian Prime Minister Tony Abbott says authorities don t know the hostage taker s motivation yet The situation at the Lindt Chocolat Cafe has been going on for hours ...

- Aboard the papal plane, Nov 26, 2014 12 08 am ( CNA ) Pope Francis has said he aims to express the social doctrine of the Church, not the views of partisan political philosophies, suggesting it is reductionistic to say otherwise ...

- On a trip to Afghanistan during President Barack Obama's first term, Defense Secretary Robert Gates was stunned to find a telephone line at the military's special operations headquarters that linked directly back to a top White House national security official ...

- Dinesh D'Souza, a reliable producer of worthless garbage opinions, has another one The protesters in Ferguson, Missouri aren't so different from ISIS, the ruthlessly violent terrorist group currently wreaking havoc in Iraq ...

- United States will provide 47 million (35 million euros) in humanitarian aid to help Palestinians hit by Israel's campaign in the Gaza Strip, Secretary of State John Kerry pledged Monday ...

- Sen John McCain, the Republican nominee for president in 2008, joked on the Colbert Report that he might once again seek the nation s highest office The Comedy Central program s host, Stephen Colbert, asked the Arizona senator whether he would be interested in succeeding Chuck Hagel as secre...

- it is very likely that sometime next year, the Supreme Court will take up yet another major Texas redistricting case In 1991, the Democrats redrew the state s congressional map to create what the Almanac of American Politics called the shrewdest gerrymander of the 1990s with incre...

- A Florida woman claims she was beaten and choked by her husband of 4 years after a heated argument about not having enough fried chicken leftovers, according to an arrest affidavit obtained by The Smoking Gun ...

### 2.4.2 Soft content

- "The Power" un hit pop elettronico dal tedesco gruppo musicale Snap! dal loro album potenza mondiale It was released in January 1990 and reached number o ...

- Downtown Detroit is home to Quicken Loans, a company that sells mortgages The founder of Quicken Loans, Dan Gilbert is a one ...

- Garth Brooks LIVE in Missouri Posted by Clear99 on October 15, 2014 Experience the electrifying return of one of country music s most influential icons! Garth Brooks is returning to Missouri his first stop marks his first St Louis performance in over 18 years! He ll do two shows at the Scot...

- Official video of Blind Melon performing Tones of Home from the album Blind Melon Buy It Here ...

- Written by Bill Payne and Richie Hayward From the classic 1977 live album "Waiting For Columbus" Produced by Lowell George ...

- Special savings on Cyber Monday only ...

- 100+ positions available Licensed Insurance Agents, Sales Reps, and Customer Service Reps ...
- By now, there are very few Americans who haven t heard of the ...
- A quick clip from our new single “It’s A Revolution” The single drops with a new music video on September 15! Stay Tuned!
- Purchase Fam Jam (Fe Sum Immigrins) on iTunes ...
- Citando este pasaje del Antiguo Testamento, Jes s se alaba la diferencia entre la conducta externa y la vida interior del ser humano A qu se refer a el Se or? A la creencia de los fariseos de que cumpliendo las obras externas de la Ley de Mois s complac an a Dios, tales como lavarse las manos...
- In this post, we bring to you some of the most inspirational pictures quotes on Saying Images These quotes are about life, love, happiness & motivation We believe that you ll find some inspiration through this post & pictures quotes & they can change your life positively Inspirational Life qu...
- Janice and James Raffle and their three sons Angus, Barney and Joshua rushed to historic landmark when they heard the American president was in the area ...
- NEW YORK, United States For many, the market for wearables often called the next major technology battleground refers to gadgets worn on ...
- Fifth track of ”IN RAINBOWS”, disc 2, by Radiohead The band had worked on In Rainbows for more than two years, beginning in early 2005 In between recording ...
- Sandude Brewing Co of Turlock is trying to make a name for itself in the craft beer industry ...

- Protesters are expected to gather in downtown Greenville Sunday afternoon to stage a Die In along Main Street ...
- Help us reach 1,000,000 signatures today, telling LEGO to ditch Shell and their dirty Arctic oil!
- Join us for the expert webinar hosted by Zo Kessler on Wednesday, July 16, 2014 at 1 PM Eastern Time ...
- We live in a culture that says that the only way to get what you want is to DO DO DO and then DO some more Push Struggle Make it happen Put your nose to the grindstone Work your ass off Try And never ever give up I once believed this was true Surrender was ...

### 2.5 Ego networks

The figures in the main text are chosen as illustrative examples of liberal, moderate, and conservative networks; the proportion of links to liberals and conservatives fall within the interquartile ranges for each affiliation, shown in the main text. To illustrate the variety of ego network compositions in which Facebook users find themselves, we show separate random samples for conservatives, moderates, and liberals (Figures S8 - S10).

## 3 Figures and Tables

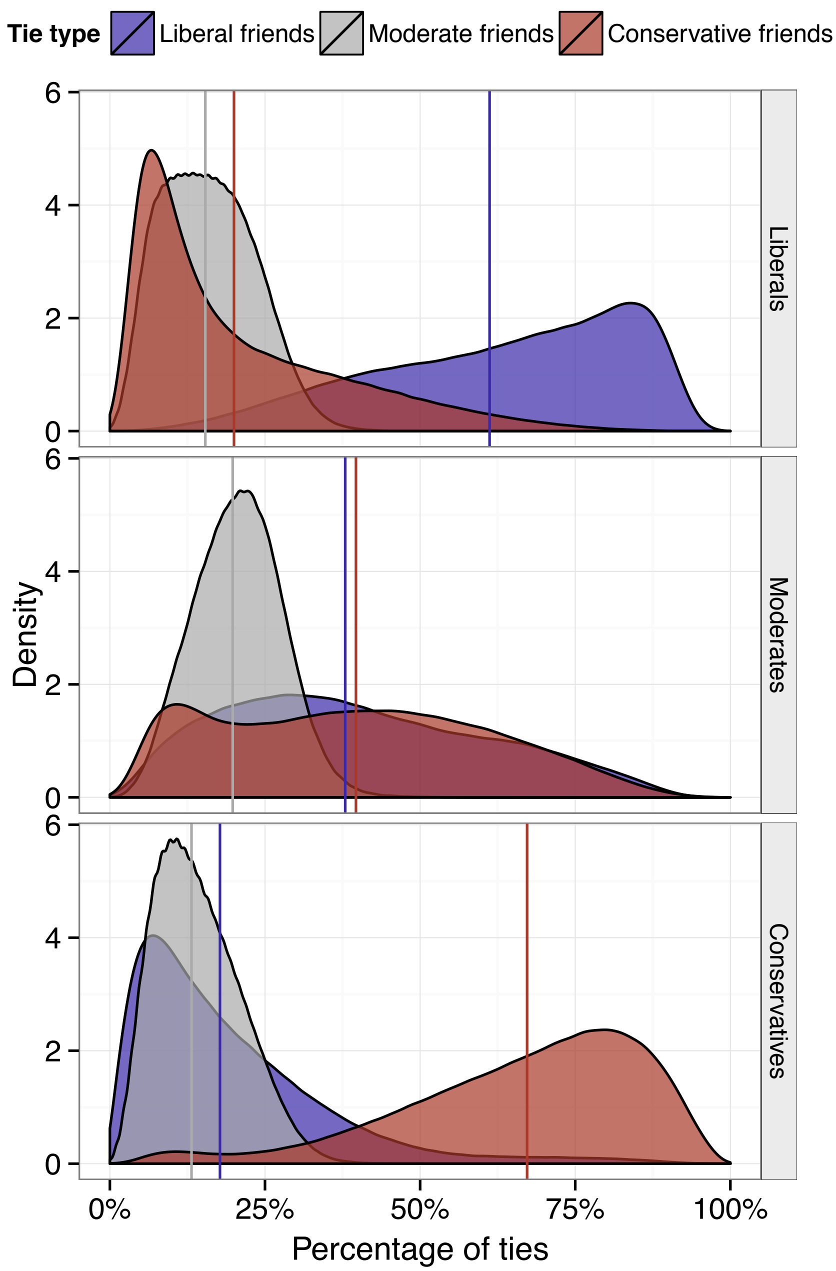
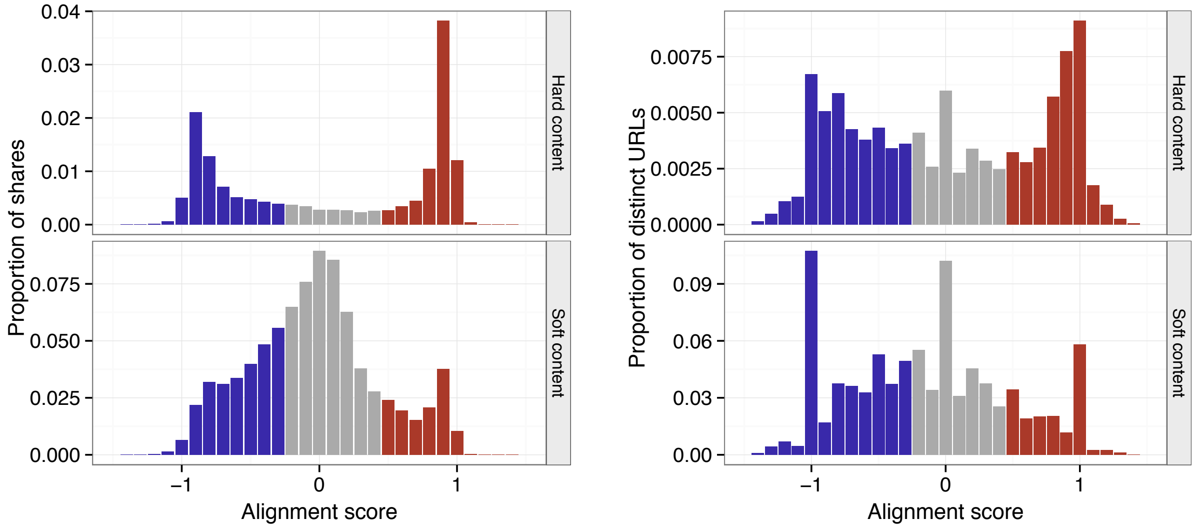
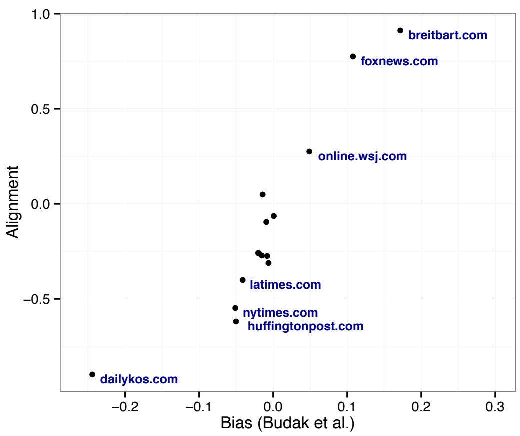



Figure S1: Kernel density plot of ties to friends of different affiliations, for liberals, moderates, and conservatives. Vertical lines indicate medians.

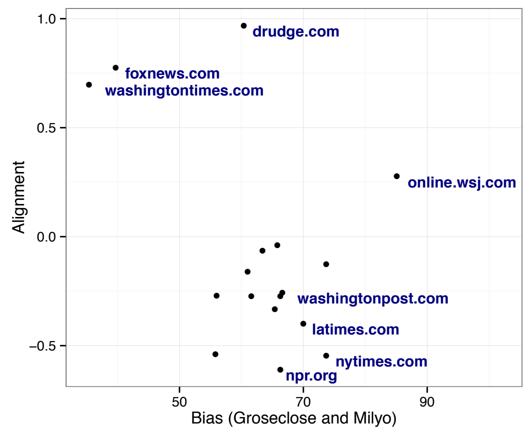



(a) (b)

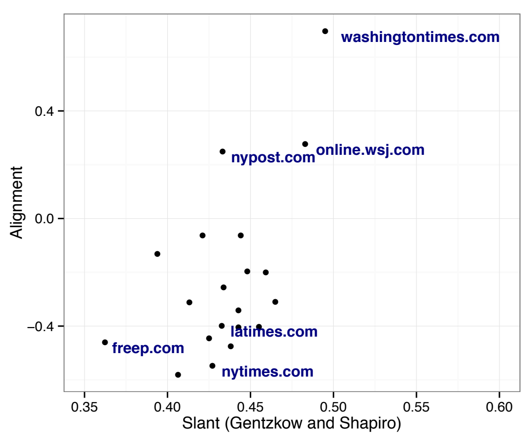
Figure S2: Distribution of alignment scores for hard news and opinion compared to soft content, (a) weighted by the total number of shares (b) weighted by the number of distinct URLs.




(a)

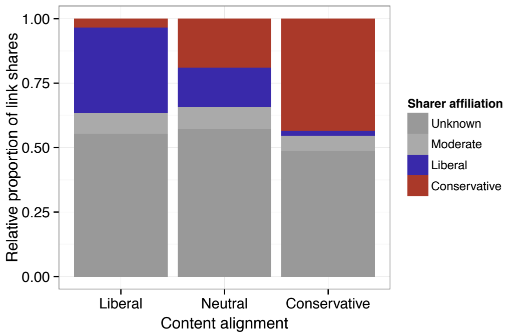



(b)

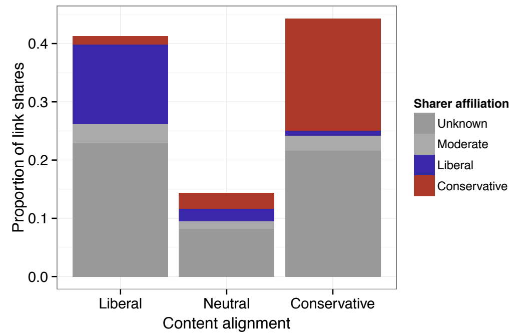



(c)

Figure S3: Average alignment by domain compared to (a) Budak et al.'s measure of media bias based on crowd-sourced annotations of partisan leanings; (b) compared to Groseclose and Milyo's measure of media bias, based on Americans for Democratic Action (ADA) ratings of the ideology of organizations that news organizations cite as sources; and (c) compared to Gentzkow and Shapiro's measure of media slant based on similarity to congressional records for 20 most shared domains.

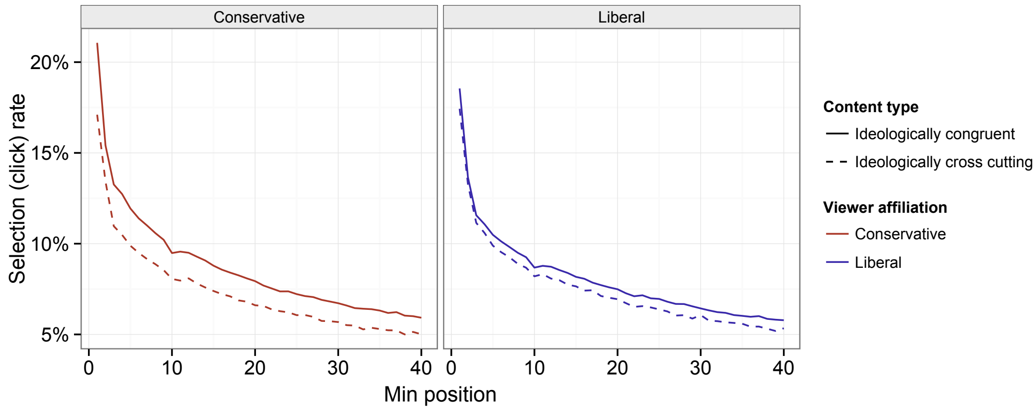



(a)

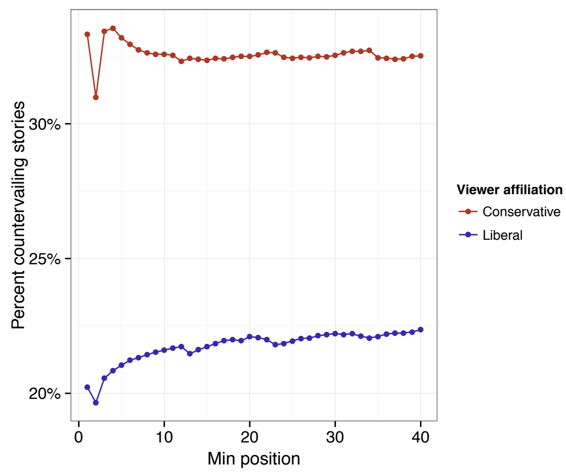



(b)

Figure S4: Ideological affiliations of sharers of liberal, neutral, and conservative hard content, including those who do not self-report their affiliation, as (a) the relative proportion of shares, from each alignment category (b) total number of link shares, as a proportion of all link shares




(a)




(b)

Figure S5: Relationship between story position and (a) click rate for ideologically congruent and cross-cutting content (b) percent of cross-cutting content shown in News Feed, for liberals and conservatives. Note that the relationship between click-through rate and position is both caused by relevance (including selective exposure) and individuals' tendencies to engage with content that is positioned toward the top of the News Feed.

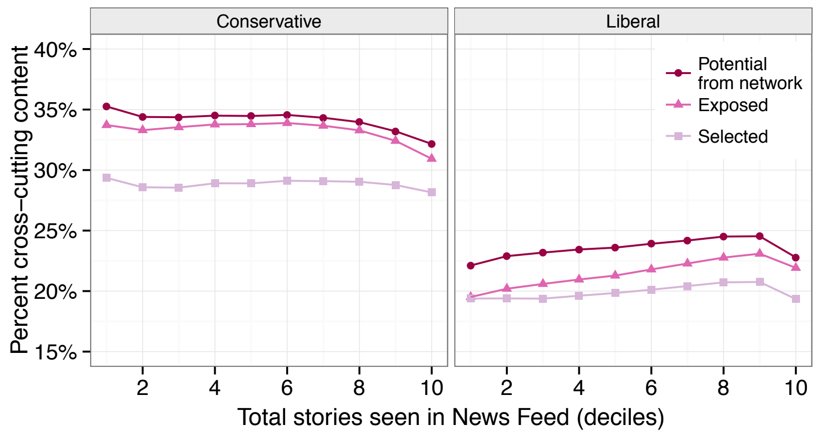



Figure S6: Proportion of ideologically cross-cutting content individuals could potentially be exposed to via their network, are exposed to via Facebook News Feed, and select (click), conditional on the decile-transformed total number of stories they are exposed to on Facebook.

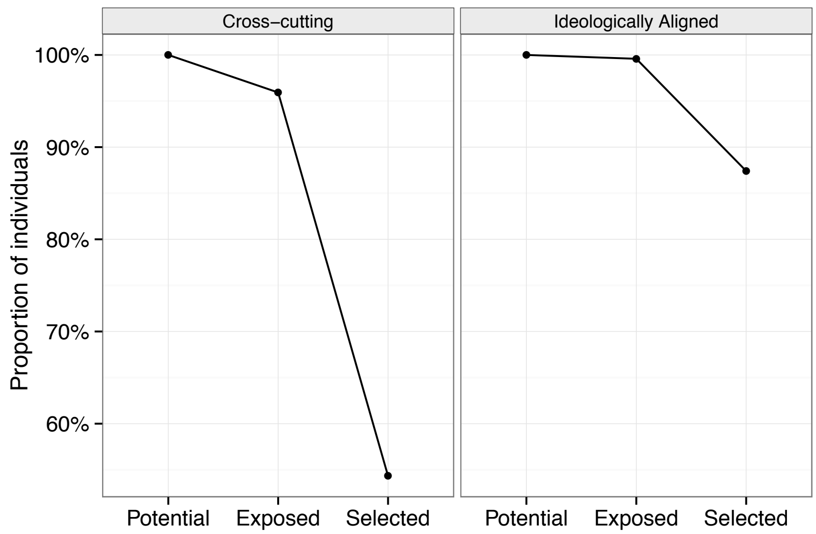



Figure S7: Proportion of individuals with at least one cross-cutting story (1) shared by friends (potential), (2) actually appearing in peoples' News Feeds (exposed), (3) clicked on (selected). The proportion of individuals with at least one ideologically aligned item is provided for comparison. Among individuals at the margins, choices about what to consume are more important in determining exposure to cross-cutting content than algorithmic ranking.

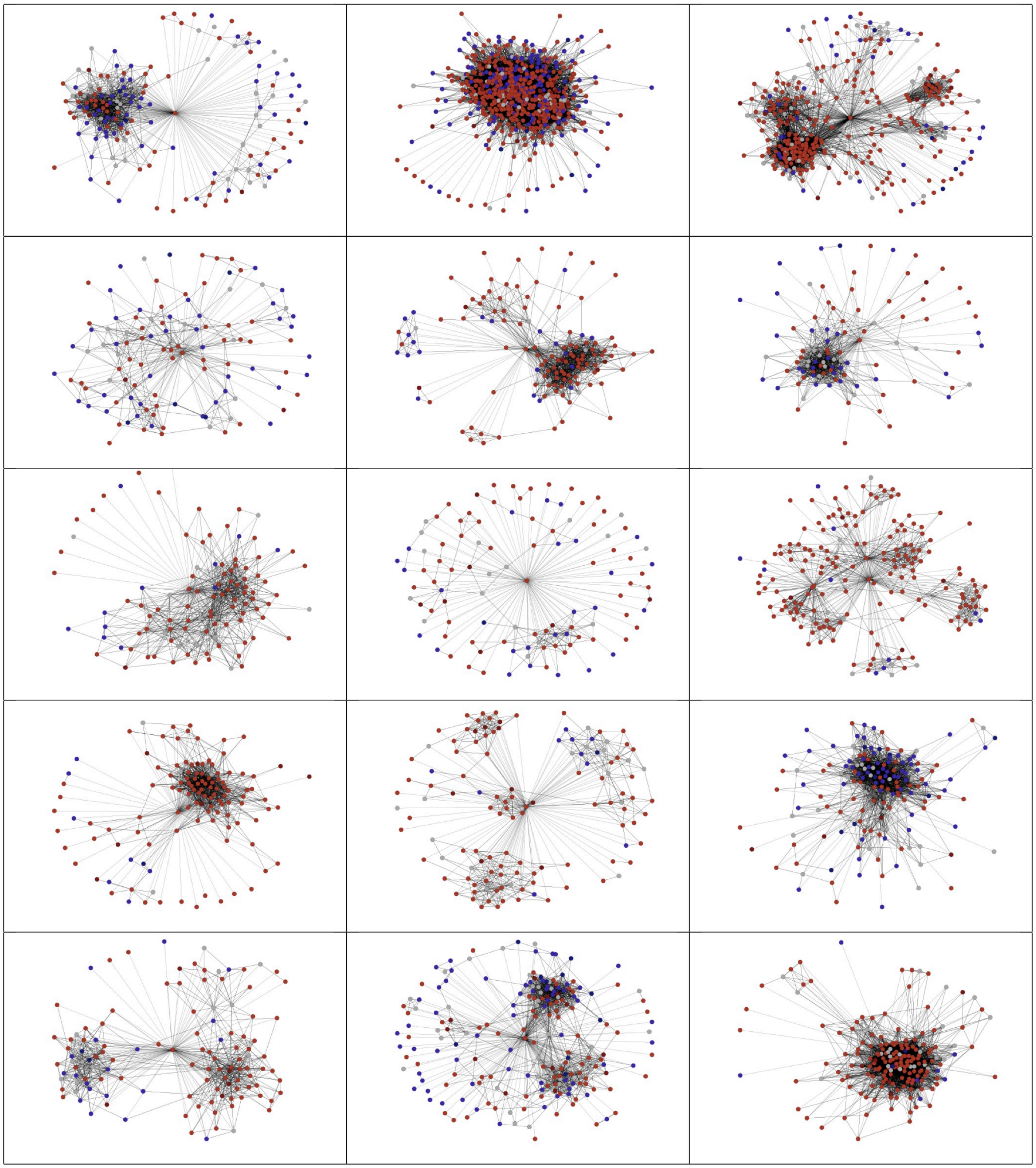


Figure S8: Ego networks of a random sample of conservatives, each having at least 100 friends with a declared political affiliation that was either conservative (red), moderate (gray) or liberal (blue). Only friends with political affiliation are shown.

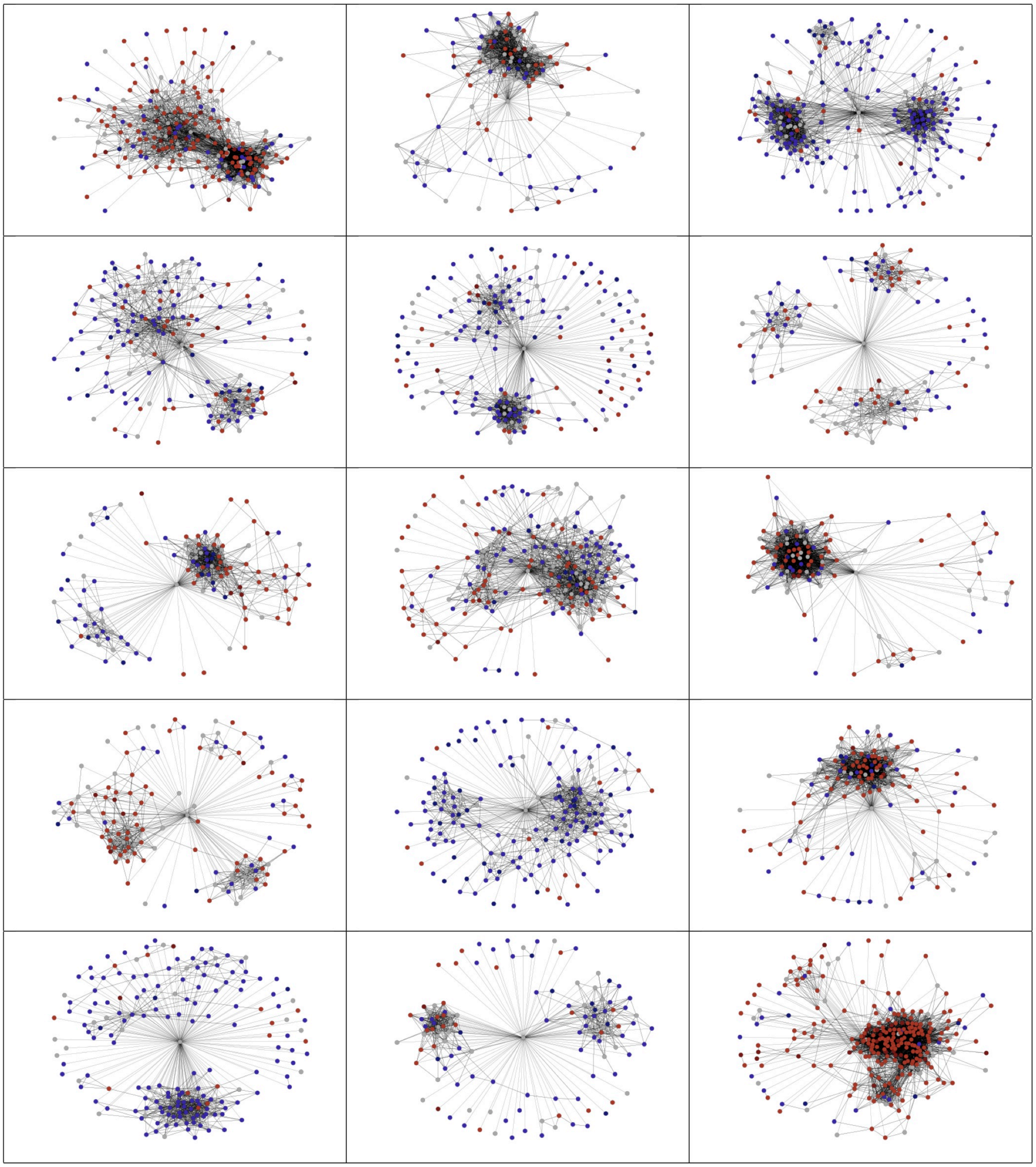


Figure S9: Ego networks of a random sample of moderates, each having at least 100 friends with a declared political affiliation that was either conservative (red), moderate (gray) or liberal (blue). Only friends with political affiliation are shown.

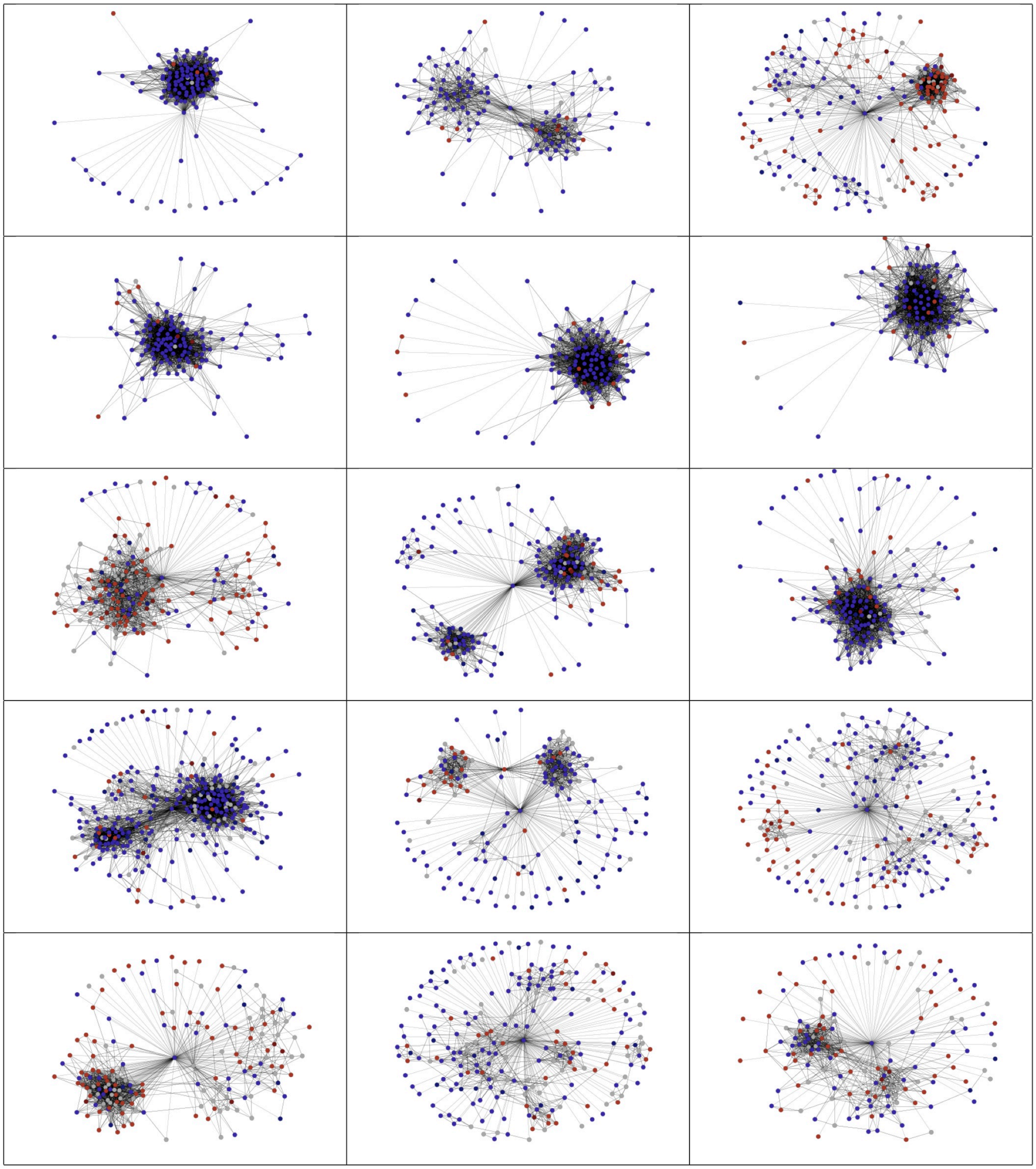


Figure S10: Ego networks of a random sample of liberals, each having at least 100 friends with a declared political affiliation that was either conservative (red), moderate (gray) or liberal (blue). Only friends with political affiliation are shown.

| Variable | Viewer affiliation | Mean | 25th perc. | Median | 75th perc. |
|---|---|---|---|---|---|
| Age | Liberal | 36.21 | 26 | 32 | 44 |
| | Moderate | 36.42 | 26 | 31 | 44 |
| | Conservative | 39.13 | 26 | 36 | 50 |
| Female | Liberal | 0.61 | 0 | 1 | 1 |
| | Moderate | 0.48 | 0 | 0 | 1 |
| | Conservative | 0.53 | 0 | 1 | 1 |
| Login days | Liberal | 171.82 | 167 | 182 | 185 |
| | Moderate | 172.64 | 169 | 182 | 185 |
| | Conservative | 172.83 | 169 | 182 | 185 |
| Num. friends | Liberal | 551.00 | 198 | 369 | 683 |
| | Moderate | 487.45 | 192 | 355 | 620 |
| | Conservative | 475.46 | 194 | 350 | 615 |

Table S1: Summary statistics for the population in our study. N = 4,063,793 (liberal), 1,602,164 (moderate), 4,469,394 (conservative).

| User affiliation | Friend affiliation | Mean | 5% | 25% | 50% | 75% | 95% |
|---|---|---|---|---|---|---|---|
| Liberal | Liberal | 0.60 | 0.24 | 0.44 | 0.61 | 0.77 | 0.90 |
| | Moderate | 0.16 | 0.03 | 0.10 | 0.15 | 0.22 | 0.31 |
| | Conservative | 0.24 | 0.02 | 0.09 | 0.20 | 0.36 | 0.60 |
| Moderate | Liberal | 0.40 | 0.10 | 0.24 | 0.38 | 0.54 | 0.76 |
| | Moderate | 0.20 | 0.06 | 0.14 | 0.20 | 0.26 | 0.35 |
| | Conservative | 0.40 | 0.07 | 0.23 | 0.40 | 0.56 | 0.76 |
| Conservatives | Liberal | 0.21 | 0.02 | 0.09 | 0.18 | 0.30 | 0.52 |
| | Moderate | 0.14 | 0.02 | 0.08 | 0.13 | 0.19 | 0.29 |
| | Conservative | 0.65 | 0.28 | 0.52 | 0.67 | 0.80 | 0.91 |

Table S2: Summary statistics for the distribution of the proportion of ties to friends of different affiliations, for liberals, moderates, and conservatives.

| Domain | Avg. alignment |
|---|---|
| www.dailykos.com | -0.90 |
| www.huffingtonpost.com | -0.62 |
| www.nytimes.com | -0.55 |
| www.cnn.com | -0.27 |
| www.washingtonpost.com | -0.26 |
| www.foxnews.com | 0.78 |
| www.theblaze.com | 0.89 |
| www.tpnn.com | 0.93 |

Table S3: Domain-level alignment for a sample of well-known media sources. Site alignment scores are obtained by averaging the alignment of URLs from a particular domain.

| Domain | -2 | -1 | 0 | 1 | 2 |
|---|---|---|---|---|---|
| www.dailykos.com | 0.967 | 0.029 | 0.000 | 0.002 | 0.000 |
| www.huffingtonpost.com | 0.361 | 0.330 | 0.291 | 0.014 | 0.001 |
| www.nytimes.com | 0.428 | 0.368 | 0.125 | 0.074 | 0.002 |
| www.cnn.com | 0.037 | 0.513 | 0.394 | 0.045 | 0.009 |
| www.washingtonpost.com | 0.208 | 0.361 | 0.213 | 0.159 | 0.056 |
| www.foxnews.com | 0.001 | 0.002 | 0.008 | 0.532 | 0.454 |
| www.theblaze.com | 0.000 | 0.000 | 0.004 | 0.309 | 0.686 |
| www.tpnn.com | 0.000 | 0.000 | 0.000 | 0.019 | 0.980 |

Table S4: Proportion of links from popular news outlets that are shared by primarily liberal (-2), somewhat liberal (-1), bipartisan (0), somewhat conservative (1), and primarily conservative (2) audiences.

| Viewer affiliation | $\pi_r$ | $\pi_n$ | $\pi_e$ | $\pi_s$ |
|---|---|---|---|---|
| Liberal | 0.454 | 0.237 | 0.222 | 0.211 |
| Conservative | 0.403 | 0.347 | 0.337 | 0.296 |

Table S5: Proportion of content that is ideologically cross-cutting for content that is shared by random others ($\pi_r$), within individuals' networks ($\pi_n$), was displayed in the News Feed ($\pi_e$), and got clicked on ($\pi_s$).

| Viewer affiliation | Random → Potential | Potential → Exposed | Exposed → Selected | Exposed → Selected* |
|---|---|---|---|---|
| Liberal | -0.626 | -0.080 | -0.063 | -0.065* |
| Conservative | -0.212 | -0.046 | -0.172 | -0.165* |

Table S6: Relative risk in probability of encountering cross-cutting versus consistent content at each transition (minus 1, see S1.7 and S1.8). * indicates position-adjusted estimate.

**References and Notes**


1. K. Olmstead, A. Mitchell, T. Rosenstiel, *Navigating news online*. Pew Research Center. Online at: http://www. journalism. org/analysis_report/navigating_news_online (2011).

2. W. L. Bennett, S. Iyengar, A New Era of Minimal Effects? The Changing Foundations of Political Communication. *J. Commun.* **58**, 707–731 (2008). doi:10.1111/j.1460-2466.2008.00410.x

3. E. Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (Penguin Press, London, 2011).

4. S. Messing, S. J. Westwood, *Communic. Res.* (2012).

5. E. Bakshy, I. Rosenn, C. Marlow, L. Adamic, *Proceedings of the 21st international conference on World Wide Web Pages* **1201.4145** (2012).

6. C. R. Sunstein, *Republic.com 2.0* (Princeton University Press, 2007).

7. N. Stroud, Media Use and Political Predispositions: Revisiting the Concept of Selective Exposure. *Polit. Behav.* **30**, 341–366 (2008). doi:10.1007/s11109-007-9050-9

8. S. Kull, C. Ramsay, E. Lewis, Misperceptions, the Media, and the Iraq War. *Polit. Sci. Q.* **118**, 569–598 (2003). doi:10.1002/j.1538-165X.2003.tb00406.x

9. S. Flaxman, S. Goel, J. M. Rao, Ideological segregation and the effects of social media on news consumption, *SSRN Scholarly Paper ID 2363701*, Social Science Research Network, Rochester, NY (2013).

10. T. Groeling, Media Bias by the Numbers: Challenges and Opportunities in the Empirical Study of Partisan News. *Annu. Rev. Polit. Sci.* **16**, 129–151 (2013). doi:10.1146/annurev-polisci-040811-115123

11. M. Gentzkow, J. M. Shapiro, Ideological Segregation Online and Offline. *Q. J. Econ.* **126**, 1799–1839 (2011). doi:10.1093/qje/qjr044

12. M. J. LaCour, A balanced information diet, not echo chambers: Evidence from a direct measure of media exposure, *SSRN Scholarly Paper ID 2303138*, Social Science Research Network, Rochester, NY (2013).

13. E. Lawrence, J. Sides, H. Farrell, Self-Segregation or Deliberation? Blog Readership, Participation, and Polarization in American Politics. *Perspect. Polit.* **8**, 141 (2010). doi:10.1017/S1537592709992714

14. D. O. Sears, J. L. Freedman, Selective Exposure to Information: A Critical Review. *Public Opin. Q.* **31**, 194 (1967). doi:10.1086/267513

15. N. A. Valentino, A. J. Banks, V. L. Hutchings, A. K. Davis, Selective Exposure in the Internet Age: The Interaction between Anxiety and Information Utility. *Polit. Psychol.* **30**, 591–613 (2009). doi:10.1111/j.1467-9221.2009.00716.x

16. L. A. Adamic, N. Glance, *Proceedings of the 3rd International Workshop on Link Discovery* (ACM, New York, 2005), pp. 36–43.

17. S. Iyengar, K. S. Hahn, Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use. *J. Commun.* **59**, 19–39 (2009). doi:10.1111/j.1460-2466.2008.01402.x

18. M. Duggan, A. Smith, Social media update 2013. Pew Research Center. Online at: http://www.pewinternet.org/2013/12/30/social-media-update-2013/ (2013). (2013).

19. M. D. Conover *et al*., *Fifth International AAAI Conference on Weblogs and Social Media* (2011).

20. D. C. Mutz, J. J. Mondak, The Workplace as a Context for Cross-Cutting Political Discourse. *J. Polit.* **68**, 140 (2006). doi:10.1111/j.1468-2508.2006.00376.x

21. S. Goel, W. Mason, D. J. Watts, Real and perceived attitude agreement in social networks. *J. Pers. Soc. Psychol.* **99**, 611–621 (2010). Medline doi:10.1037/a0020697

22. D. C. Mutz, Cross-cutting Social Networks: Testing Democratic Theory in Practice. *Am. Polit. Sci. Rev.* **96**, 112 (2002). doi:10.1017/S0003055402004264

23. B. Bishop, *The Big Sort: Why the Clustering of Like-Minded America is Tearing Us Apart* (Houghton Mifflin Harcourt, New York, 2008).

24. D. C. Mutz, P. S. Martin, *Am. Polit. Sci. Rev.* **95**, 97 (2001).

25. T. Mendelberg, Political Decision Making. *Deliber. Particip.* **6**, 151–193 (2002).

26. R. Huckfeldt, J. M. Mendez, T. Osborn, Disagreement, Ambivalence, and Engagement: The Political Consequences of Heterogeneous Networks. *Polit. Psychol.* **25**, 65–95 (2004). doi:10.1111/j.1467-9221.2004.00357.x

27. R. Bond, S. Messing, Quantifying Social Media's Political Space: Estimating Ideology from Publicly Revealed Preferences on Facebook. *Am. Polit. Sci. Rev.* **109**, 62–78 (2015). doi:10.1017/S0003055414000525

28. E. Riloff, *Proceedings of the Thirteenth National Conference on Artificial Intelligence* (AAAI Press, Portland, Oregon, 1996), vol. 2 of AAAI'96, pp. 1044–1049.

29. E. Riloff, *Case-Based Reasoning Research and Development*, K. D. Ashley, D. G. Bridge, eds., no. 2689 in Lecture Notes in Computer Science (Springer Berlin Heidelberg, 2003), pp. 4–4.

30. S. Gupta, C. D. Manning, *Proceedings of the SIGNLL Conference on Computational Natural Language Learning* (SIGNLL, Baltimore, MD, 2014), vol. 98.

31. S. Gupta, C. D. Manning, *Proceedings of the ACL 2014 Workshop on Interactive Language Learning, Visualization, and Interfaces* (ACL-ILLVI) (Baltimore, MD, 2014), p. 38.

32. C. Budak, S. Goel, J. M. Rao, Fair and Balanced? Quantifying Media Bias through Crowdsourced Content Analysis (2014); available at http://ssrn.com/abstract=2526461.

33. T. Groseclose, J. Milyo, A Measure of Media Bias. *Q. J. Econ.* **120**, 1191–1237 (2005). doi:10.1162/003355305775097542

34. M. Gentzkow, J. M. Shapiro, What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica* **78**, 35–71 (2010). doi:10.3982/ECTA7195

35. S. K. Whitbourne, Openness to experience, identity flexibility, and life change in adults. *J. Pers. Soc. Psychol.* **50**, 163–168 (1986). Medline doi:10.1037/0022-3514.50.1.163