Pew Research Center

Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump

Democratic legislators' opposition on Facebook spiked after Trump's election, while angry reactions increased among all congressional Facebook followers.

BY *Patrick van Kessel, Adam Hughes, and Solomon Messing*

About Pew Research Center

Pew Research Center is a nonpartisan fact tank that informs the public about the issues, attitudes and trends shaping America and the world. It does not take policy positions. The Center conducts public opinion polling, demographic research, content analysis and other data-driven social science research. It studies U.S. politics and policy; journalism and media; internet, science and technology; religion and public life; Hispanic trends; global attitudes and trends; and U.S. social and demographic trends. All of the Center's reports are available at www.pewresearch.org. Pew Research Center is a subsidiary of The Pew Charitable Trusts, its primary funder.

Terminology

Political opposition and political support refer to statements made by members of Congress on Facebook that oppose or support the actions, decisions or positions of political actors and groups. This report examines support or opposition directed toward Donald Trump, Barack Obama, Hillary Clinton, Democrats or liberals, and Republicans or conservatives. Researchers determined whether or not individual posts expressed opposition or support by classifying a set of posts manually and then training a machine learning model to classify the rest. Researchers classified whether posts expressed both opposition to some figures or groups *and* support for others, but did not classify both support and opposition directed at the same figures or groups.

Local issues refer to references within congressional statements to any place, group, individual(s) or event in the politician's state or district. Researchers determined whether or not individual posts discussed local issues by classifying a set of posts manually and then training a machine learning model to classify the rest.

DW-NOMINATE is a measure of political ideology that places members of the U.S. House and Senate on a liberal-to-conservative ideology scale according to their roll-call voting history in each legislative session of Congress.[1] The scale ranges from -1 (very liberal) to 1 (very conservative) across all Congresses. For the time period studied here, it ranges between -0.77 and 0.99. Very liberal and very conservative legislators are defined as those with DW-NOMINATE scores in the furthest left 10% and furthest right 10% of all DW-NOMINATE scores. Moderates are defined as legislators with DW-NOMINATE scores in the middle 20% of all scores.

Engagement on Facebook refers to "likes," "comments" and "shares," as well as specific reactions, such as "angry." On the Facebook platform, these digital actions provide users a means for interacting with posts created by members of Congress.

---

1. See Lewis, Jeffrey B., Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet. 2017. "Voteview: Congressional Roll-Call Votes Database." https://voteview.com/
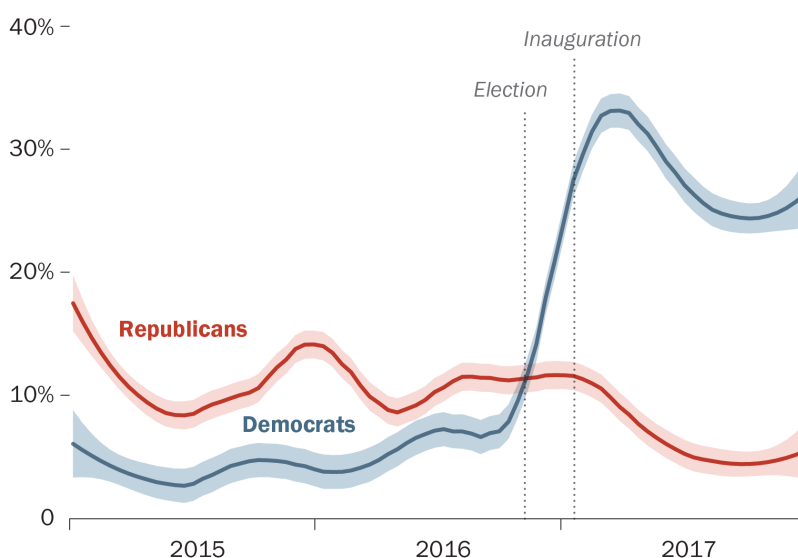
Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump

Democratic legislators' opposition on Facebook spiked after Trump's election, while angry reactions increased among all congressional Facebook followers.

The 2016 presidential election coincided with substantial shifts in the ways that members of Congress communicated with their constituents online. A new Pew Research Center analysis examines congressional Facebook posts from Jan. 1, 2015, through Dec. 31, 2017, a three-year timespan that includes the entire 114th session of Congress, the 2016 primary and general elections, the first year of the 115th Congress, and Republican President Donald Trump's first year in office. The analysis finds that Democrats expressed political opposition nearly five times as much under Trump as they did during the last two years of Barack Obama's presidency. Much of this opposition was directed at President Trump, though Democrats also increasingly opposed Republican members of Congress.

Following Trump's election, Facebook posts from Democrats in Congress included more oppositional language

Average % of posts expressing political opposition

Note: Political opposition includes statements that oppose President Trump or Republicans and conservatives (for Democrats) and statements that express opposition to President Obama, Hillary Clinton or Democrats and liberals (for Republicans). Lines are based on LOWESS estimates. The shaded regions are the 95% confidence bands for the estimated trends.
Source: Pew Research analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Meanwhile, congressional Republicans posted in *support* of Trump more than twice as much as congressional Democrats posted in support of President Obama during his final two years in office (researchers did not have access to posts from Obama's first year in office).[2]
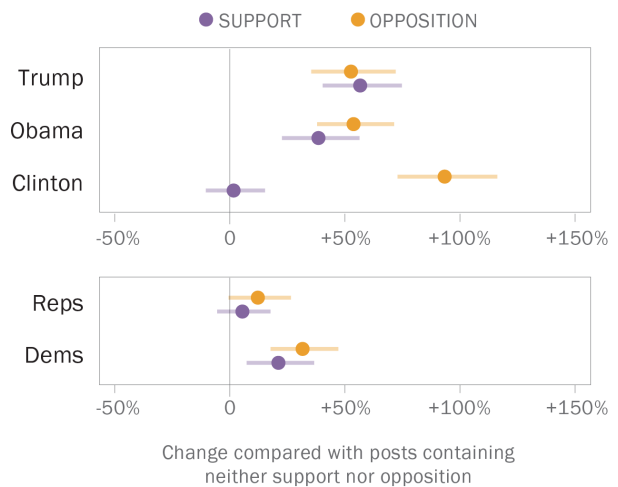
Members of Congress who expressed political opposition most often were also the most liberal or conservative.[3] This pattern is in line with the Center's previous research on how members of Congress express political disagreement. But the new analysis also shows that the most ideological members were also the most likely to express support for others in their party. In other words, the most liberal and conservative members of Congress both attacked those on the other side more often and were more likely to affirm their own side's decisions and positions. Moderates, meanwhile, tended to focus most of their posts on local issues.

Changes occurred not only in what members posted, but also in how their online audiences responded. The Facebook audience increasingly expressed angry reactions when responding to congressional posts. In early 2016, Facebook introduced alternatives to the traditional "like" reaction – and between late February 2016 and Election Day, just 2% of all reactions to posts used the "angry" option. But by the end of 2017, 9% of all reactions to posts by Democrats and 13% of reactions to posts by Republicans were angry.

Nearly universally, both supportive and oppositional posts about Trump or Obama drew more engagement – including likes, comments

Legislators' Facebook audiences 'liked' posts that opposed political figures more than posts that didn't take sides

Estimated percentage increase in Facebook likes for posts containing …

Change compared with posts containing neither support nor opposition

Note: Points are estimated likes for posts that express opposition and support directed at labeled political figures and groups. Estimates are predictions from a multiple regression OLS model that includes random intercepts for each member of Congress and week in the data, alongside indicator variables for each presidential administration, legislator party and account type. The lines in the figure represent the 95% confidence interval for each estimate. See methodology section for additional details.
Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

---

2. Complete data could not be obtained for legislators' accounts prior to the 114th Congress, as some members were not active on Facebook in earlier sessions, or have deleted their accounts since leaving office.
3. Researchers used the DW-NOMINATE ideology measure, which is based on members' roll-call voting records, to classify members as liberal or conservative.

and shares – than posts about other topics. But the pattern was somewhat different for Hillary Clinton. Congressional posts that supported her drew the same number of likes as posts that did not take sides either way, while posts opposing her received 93% more likes on average – the largest increase in likes across all the kinds of posts examined here.

The new analysis used a combination of human coders and machine learning techniques to examine both the changing nature of congressional Facebook outreach and the way Facebook audiences responded to varying kinds of messages. To create the dataset, researchers studied more than 700,000 posts from 599 members of Congress during a three-year period surrounding the November 2016 election and Trump's first year in office, beginning Jan. 1, 2015, and ending Dec. 31, 2017. Among the key findings:

After Trump's election

- **Change of tone for Democrats after Trump took office:** Following Trump's inauguration, the share of Democratic legislators' Facebook posts that included oppositional language – defined here as posts that oppose or disagree with the actions, decisions or positions of Trump and his administration *or* Republicans and conservatives – peaked in March 2017 at an average of 33% of all of their posts before ramping down to 24% toward the end of the year. That compares with an average of 12% of Republican lawmakers' posts expressing opposition to Democrats and liberals *or* Obama during the last two years of his presidency. Democratic opposition during Obama's presidency – at that point mostly aimed at congressional Republicans – appeared in just 6% of their Facebook posts.

- **Republicans expressed more support for Trump in his first year than Democrats did for Obama in the previous Congress:** Just 4% of the average congressional Democrat's Facebook posts from January 2015 through December 2016 expressed support for Obama. In contrast, the average Republican member expressed support for Trump in 9% of their Facebook posts in 2017.

- **After the election, the Facebook audience was far more likely to use the "angry" reaction to respond to outreach:** Between Feb. 24, 2016 (when the reactions were first made available to Facebook users) and Election Day, 2% of all Facebook reactions to congressional posts were angry. But after the election through the end of 2017,

that share tripled to 6%.[4] By December 2017, the average was 9% for posts by Democrats and 13% for posts by Republicans.

During the 2016 campaign

- **Members of both parties focused more on Clinton than Trump during the 2016 campaign:** Both presidential candidates drew modest attention from members of Congress on Facebook, and then-candidate Trump received less support from members of his party than Clinton did from hers. Between each party's convention and Election Day, Democrats in Congress posted in support of Clinton substantially more often (a total of 1,614 posts) than Republicans posted in support of Trump (a total of 690 posts). However, Republicans opposed Clinton in 2,041 posts, far more than Democrats expressed opposition to Trump (1,383 posts).

Consistent patterns

- **Moderates went local, while very liberal and very conservative members took sides:** Moderates in Congress were less likely to express political support or opposition than were very liberal or conservative members. The majority of moderates' outreach focused on local issues (54%, compared with 38% for the most liberal or conservative members). Those in the middle of the ideological spectrum issued statements of political support and opposition about half as often as those on either end of the ideological spectrum.

- **More online followers engaged when elected officials took sides, especially when opposing individuals on the other side:** Across the full time frame of the study, congressional posts that opposed Obama, Trump or Clinton earned more likes, comments and shares than posts that didn't take sides either way. Posts that expressed support for politicians also received more engagement some of the time, but the pattern was not always consistent.

This analysis is based on 737,598 Facebook posts issued by 599 members of Congress between Jan. 1, 2015, and Dec. 31, 2017. The total number of legislators is greater than 535 (the current number of voting officials in the U.S. House and Senate) because members who were newly elected in the 115th Congress or in special elections are included in the study, as long as they produced at least 10 posts within a given Congress.

---

4. For the average member's average post. Other Facebook reactions include "like," "love," "haha," "thankful," "wow," and "sad."

Researchers included both official Facebook accounts (those managed by congressional staff) and unofficial accounts (those used in a personal or campaign capacity) for members of Congress in this analysis. They did so in order to capture a more complete range of outreach on social media than would be possible with official accounts alone. As a result, the study includes a total of 1,129 accounts belonging to the 599 individual legislators.

Official accounts are used to communicate information as part of the member's representational or legislative capacity, and U.S. Senate and House members may draw upon official staff resources appropriated by Congress when releasing content via these accounts. Unofficial accounts – often used in a personal and campaign capacity – may not draw on these government resources under official House and Senate guidelines. Members posted more often on official accounts across the study period: 76% of the average member's posts came from their official account (for Democrats, the share was 78%; for Republicans, it was 75%).

How Pew Research Center studied congressional Facebook posts

Congressional outreach

Researchers identified 737,598 Facebook posts created by 599 members of Congress between Jan. 1, 2015 and Dec. 31, 2017. The posts were published on 1,129 official and unofficial Facebook pages managed by members of Congress and their staff.

Audience engagement

In total, congressional Facebook posts received over 481 million likes, 45 million comments, and 141 million shares from the Facebook audience. Researchers wanted to see if particular kinds of posts were especially likely to receive likes, comments, and shares.

Content classification

To answer those questions, a team of human coders classified a sample of 11,000 posts as to whether each expressed support or opposition directed toward President Trump, President Obama, Hillary Clinton, Republicans, and/or Democrats. They also classified whether or not the posts discussed organizations, events, or individuals in each member's state or district. Then, using machine learning models, the entire set of 737,598 congressional Facebook posts was classified.

Source: "Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

To classify the posts, Pew Research Center manually analyzed a subset (11,000 total) of all the posts, classifying each post's contents for the events, topics and issues raised or discussed in each one. Specifically, the analysts coded each post based on whether it expressed disagreement with presidents, candidates or parties; expressed support for the same; or mentioned local events, places or people. Next, researchers trained machine learning algorithms to make predictions – based on what the human coders reported – in order to classify the content of the entire set of posts.

How researchers classified more than 700,000 posts

Researchers asked human coders recruited from Amazon's Mechanical Turk to read a sample of 11,000 posts and then note whether or not the post mentioned a list of political figures and groups. That list included Trump, Obama and Clinton, as well as two political groups: Democrats and/or liberals and Republicans and/or conservatives. The coders determined whether each post that mentioned one of those individuals or groups expressed support, opposition or neither toward each figure or group. Human coders also determined whether or not each post discussed local issues, defined as relating to "a place, a group, individual(s), or an event in the politician's state or district."

Each set of particular kinds of post – those labeled as opposing Obama, or supporting Republicans, for instance – was used as a "training dataset" for separate machine learning models. Those models used the human decisions about what words convey political support or opposition as a guide for determining how to classify the remaining posts that human coders did not read and classify. This approach builds upon previous Pew Research Center work that examined similar kinds of rhetoric in congressional statements. Performance statistics and additional details are provided in the methodology.

While Facebook is one important part of members' media outreach efforts, members also communicate with their constituents through press releases, town hall meetings, media appearances and on other social media outlets. Although this report does not examine communication across all these channels, Facebook posts constitute a useful way to compare members' communication, as they can be systematically captured and analyzed. Previous research suggests that statements that members of Congress express on Facebook are similar in many ways to those they make in press releases. Focusing on Facebook posts also makes it possible to measure how much a member's audience interacts with their posts via likes, comments and shares. Facebook is the most widely used social media website (excluding YouTube) and the social media site from which most Americans get news.

1. Democrats posted more and expressed more opposition after Trump took office

Shortly after Election Day 2016, Democratic members of Congress became more active on Facebook, posting more frequently than Republican members for the first time since at least January 2015. And in those posts, they expressed more than twice as much political opposition – directed at both President Trump and Republicans – than Republican legislators expressed on Facebook toward Obama and Democrats during the last two years of his term. Researchers tracked congressional rhetoric on Facebook beginning in 2015.
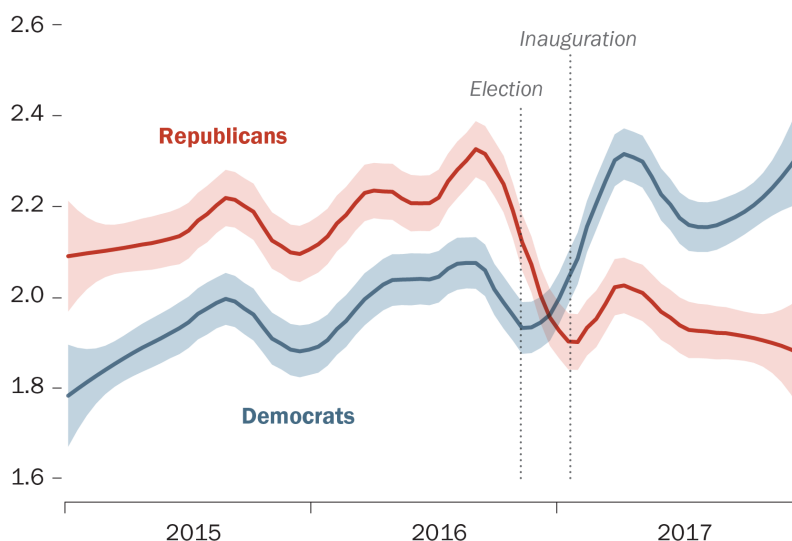
Democrats posted more under Trump than under Obama; Republicans posted less

In 2017, Democrats in Congress started posting more often, while Republicans posted less. The average Democrat posted 33% more often during the 115th Congress than in the 114th, going from 34 posts per month to 45 posts per month. Republicans posted less often, from an average of 42 to 37 posts per month.

Cumulatively, these changes have had a substantial impact on the total volume of Facebook posts being produced across all members of Congress in each party. Democratic members in the 115th Congress produced over 34,000 more Facebook posts across 2017 than they averaged in each year of the 114th Congress. In contrast,

After 2016 election, Democrats in Congress posted on Facebook more often than Republicans

Average daily number of Facebook posts created by individual ...

Note: Each line is a smoothed estimate of the daily number of Facebook posts created by the average Republican (red line) and the average Democratic (blue line) member of Congress. Lines are based on LOWESS estimates. The shaded regions are the 95% confidence band for the estimated trends.
Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Republicans produced over 25,000 fewer posts in 2017 than they averaged during the previous two years.

How members of Congress expressed support and opposition

Overall, 19% of all Facebook posts included in the study contained statements of political support or opposition directed at one of several political figures or groups. Researchers identified the members of Congress who were most likely to express support for or opposition to the key political figures and groups analyzed here: Donald Trump, Barack Obama, Hillary Clinton, Democrats and/or liberals, and Republicans and/or conservatives, in order to provide examples of the kinds of language they used when taking sides on Facebook.

Among Republicans, Sen. Luther Strange, R-Miss., was most likely to voice support for Trump, doing so in 18% of all his posts, while Rep. Jim Jordan, R-Ohio, expressed support for Republicans in 16% of his posts, more than any other member.

One post on Strange's Facebook page in support of Trump, created Aug. 21, 2017, said: "Tonight, President Donald J. Trump put forward a bold plan to win in Afghanistan. This is the speech our troops deserved to hear for years under President Obama and never did." On May 4, 2017, Jordan created a post that said, "The legislation that passed the House today is better because of the intense involvement of conservatives over the past few weeks."

When it came to Republican opposition, Sen. Johnny Isakson, R-Ga., expressed opposition to Obama in 35% of all posts from 2015 through 2017. On April 4, 2016, he wrote: "Yet again President Obama has undermined the trust of our military leaders and jeopardized the well-being of our nation." Rep. Jeff Duncan, R-S.C., opposed Democrats most often, in 9% of all posts, while Rep. Trey Gowdy, R-S.C., expressed the most opposition to Hillary Clinton, doing so in 8% of his posts.

On Sept. 25, 2017, Duncan wrote: "When will Democrats admit their great social engineering experiment is a failure and that people are suffering because of it?" Almost all of Gowdy's criticism focused on Clinton's tenure as secretary of state and controversy over her email use. On July 7, 2016, he wrote: "if you are a private citizen in the Army and you email yourself classified information, you will be kicked out. But if you are Hillary Clinton and you seek a promotion to commander-in-chief, you will not be."

On the Democratic side, former Sen. Harry Reid, D-Nev., was both most likely to support Obama (18% of his posts) and most likely to oppose Republicans (44%).[5] On Jan. 21, 2015, he wrote a post that did both: "President Obama laid out a clear vision for strengthening America's middle class that draws a clear contrast with the Republican Congress' partisan political agenda of special interest giveaways and tax breaks for the wealthiest Americans." Minority Leader Nancy Pelosi, D-Calif., was most likely to support fellow Democrats and did so in 24% of her posts. One such post, from April 25, 2016, stated that "Democrats have introduced an emergency supplemental bill that will provide the swift and decisive action needed to help protect American communities from the Zika virus this summer."

Former Sen. Barbara Boxer, D-Calif., had the largest share of posts (12%) that supported Hillary Clinton.[6] For example, a post created shortly after Clinton's nomination on July 29, 2016, stated: "I've known Hillary for more than 20 years, and I couldn't be more proud of her."

When it came to conflict with Trump, Rep. Jerrold Nadler, D-N.Y., was the most likely to express opposition and did so in 40% of his Facebook posts. On Aug. 15, 2017, in the wake of a violent white supremacist rally in Charlottesville, Virginia, Nadler wrote, "When the man occupying the White House calls racists and neo-Nazis 'very fine people,' just saying 'hate is bad,' doesn't remotely cut it."

---

5. Reid left office at the end of the 114th Congress.
6. Boxer left office at the end of the 114th Congress.

Democratic focus of opposition shifted from Trump to Republicans in late 2017
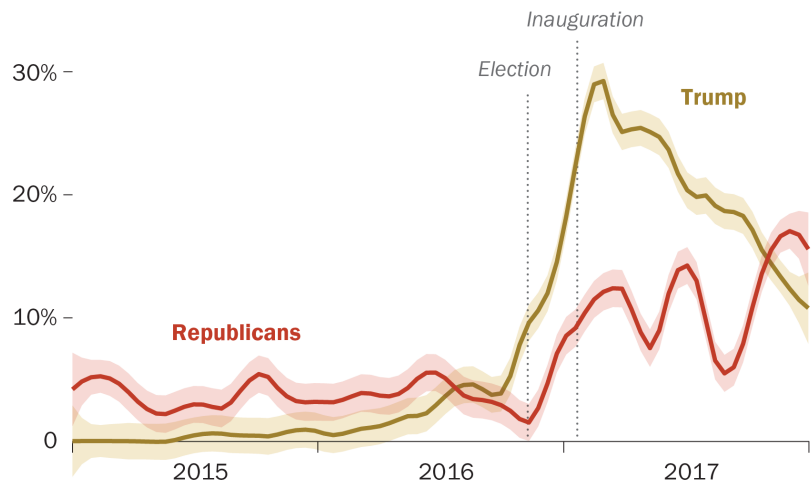
Overall, 30% of the average Democrat's posts in 2017 contained some form of opposition toward Trump, Republicans or both. At the outset of the new administration, the majority of this opposition was directed at the president. Out of all the oppositional posts that the average Democrat produced in 2017, 71% of those posts targeted Trump, while 41% targeted Republicans more generally (some targeted both).

However, throughout Trump's first year in office, Democrats shifted their focus away from the president, choosing to target their opposition toward Republicans with increasing frequency. By the end of the year, Democratic Facebook posts were more likely to express opposition toward Republicans than Trump. In December 2017, 70% of Democratic posts that expressed opposition were directed at Republicans, while only 43% of oppositional posts targeted the president.
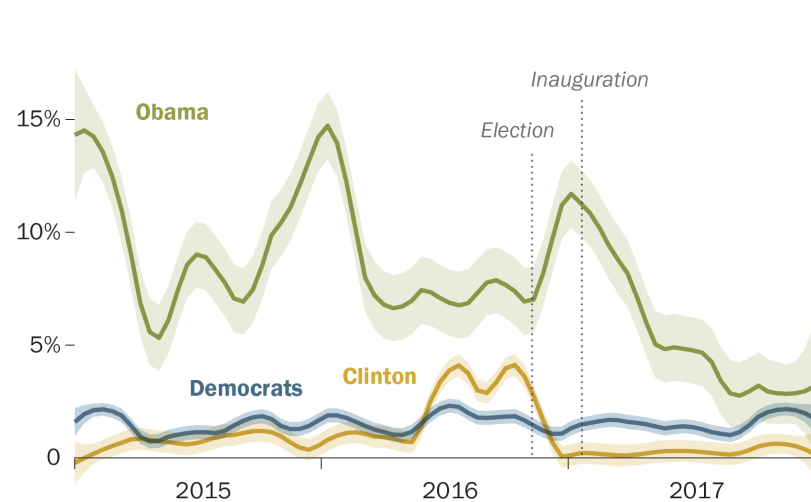
These patterns parallel an increased appetite for political conflict among legislators' Democratic constituents. From

Democratic opposition in Facebook posts peaked after Trump's inauguration; Republicans focused opposition on former President Obama

Average % of Democrats' posts expressing opposition to ...

Average % of Republicans' posts expressing opposition to ...

Note: Each line is a smoothed estimate of the percent of all posts created by the average Democratic or Republican member of Congress that expressed opposition to Republicans and conservatives or Trump (top panel) or Obama, Clinton or Democrats and liberals (lower panel). The shaded regions are the 95% confidence band for the estimated trends.
Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

2017 to 2018, the percentage of Democrats in the U.S. public who said they like elected officials who make compromises with those they disagree with dropped from 69% to 46%.

Across the study period, Republicans focused on President Obama in 82% of oppositional posts on average (8% of all posts by the average lawmaker). A much smaller share of Republicans' posts expressed opposition toward Democrats or Hillary Clinton (2% and 1% of all posts by the average lawmaker, respectively). This focus on Obama persisted after the election. Across 2017, statements opposing Obama and Obama-era policies like the Affordable Care Act were more common than those opposing Democrats. At the same time, throughout 2017, fewer than 1% of the average Republican's posts expressed opposition to Clinton, despite some anecdotal reports to the contrary.
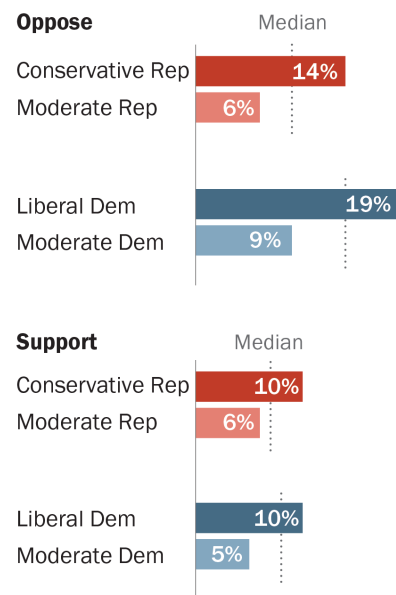
Very liberal and conservative legislators took sides; moderates went local

More liberal or conservative legislators – based on the DW-NOMINATE estimate of ideology – were about twice as likely to express either support or opposition toward other political figures and groups as compared with more moderate members.

Moderate members – defined as those that fell in the middle 20% of the roll-call-based ideology estimate – publicly opposed others in about 7% of posts, compared with a rate of about 16% for the most conservative or liberal members, defined as those in the most liberal tenth or most conservative tenth of the ideology measure. Moderates expressed support for the president or their party in roughly 5% of their Facebook posts, while the most liberal or conservative members did so in 10% of all their Facebook outreach.

Very liberal and conservative members expressed more political opposition and support

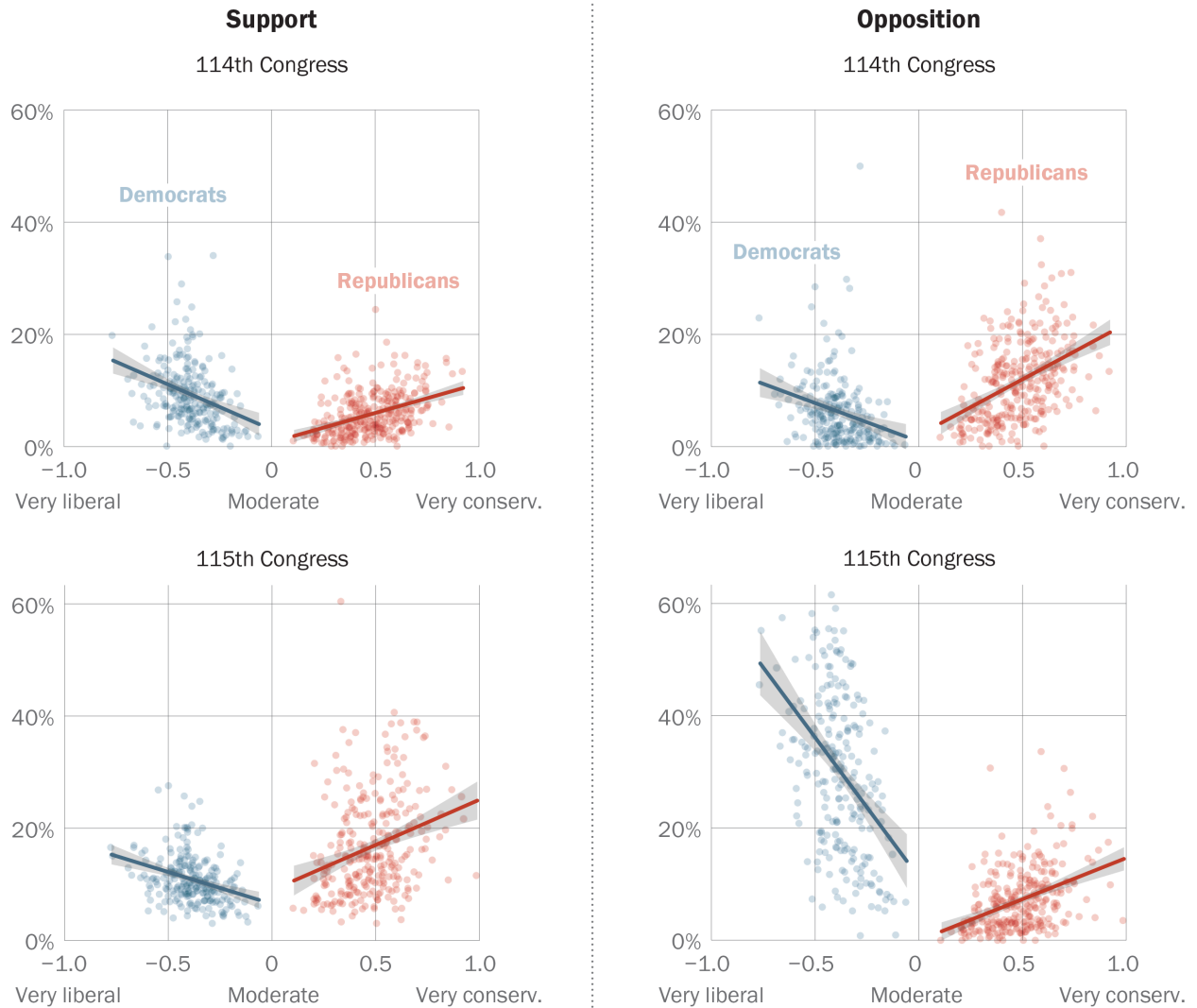*Average % of Facebook posts that expressed opposition or support*

Note: "Ideological" members of Congress refers to the 10% most liberal and 10% most conservative members, based on their DW-NOMINATE ideology scores. "Moderate" members include those whose DW-NOMINATE ideology scores fell in the middle 20% of the distribution.
Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Ideological patterns in support and opposition consistent over time, but Democratic opposition increased in 115th Congress overall

% of legislators' posts expressing opposition and support, by estimated legislator ideology

Support

114th Congress

Opposition

114th Congress

115th Congress

115th Congress

Note: Each point represents a member of Congress. Estimated ideology is based on roll call votes (DW-NOMINATE). The fitted line shows the relationship between each legislator's proportion of Facebook posts containing support or opposition and their ideology, for each party and across each Congress. The shaded region is that estimate's 95% confidence band, an attempt to quantify the uncertainty in this relationship.

Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.

"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

During the 115th Congress, starting in January 2017 and including the inauguration of President Trump, the link between ideology and political opposition became even more pronounced among Democrats. For that Congress, the most liberal Democrats expressed opposition in 35% of posts on average, compared with 10% for moderates.
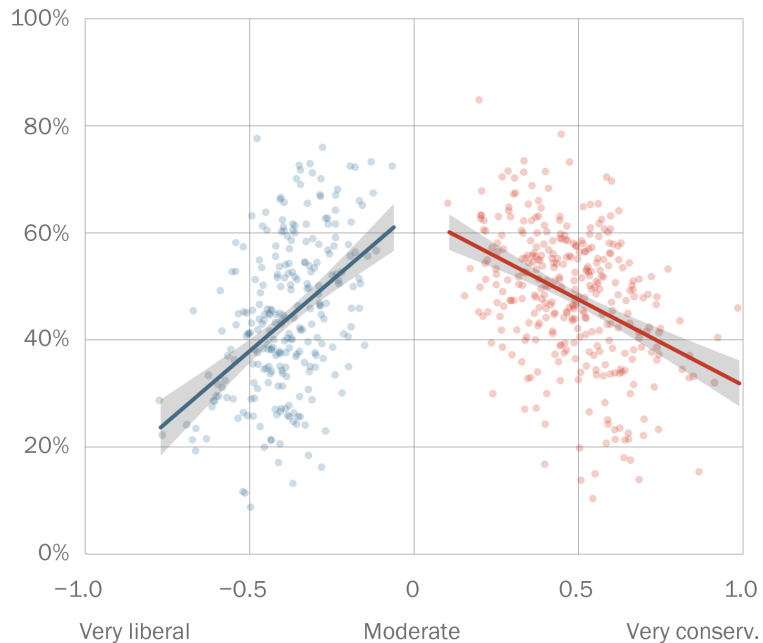
In contrast, while the most liberal and conservative members of Congress focused their rhetoric on expressions of political support and opposition, moderates disproportionately talked about local issues in their Facebook outreach. These posts, which draw attention to individuals, groups, and organizations in the state or district the member represents, made up more than half of all posts from the most moderate members of Congress, compared with about one-third of posts for those on each end of the ideological spectrum.

In general, members were more likely to cover these local issues than engage in other forms of outreach described here, including opposition to political opponents. Over the course of the full three-year study period, nearly half of congressional outreach on Facebook focused on local issues (45% of posts), compared with the 13% of all posts that expressed oppositional views. The average Republican focused on local topics in 48% of Facebook posts, compared with a rate of 44% for the average Democrat. Meanwhile, only 16% of posts from the average Democrat and 10%

Moderates in Congress were more likely to discuss local issues

% of legislators' posts focused on local issues, by estimated legislator ideology

Note: Each point in the figure represents a member of Congress. Estimated ideology is based on roll call votes (DW-NOMINATE). The fitted line shows the relationship between each legislator's proportion of Facebook posts discussing local issues and their ideology, for each party. The shaded region is that estimate's 95% confidence band, an attempt to quantify the range of uncertainty in this relationship.
Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

from the average Republican expressed opposition to the other side. Across both parties, the average legislator expressed support for others in their own party in 8% of posts.

During the 2016 campaign, members of both parties focused on Clinton

During the run-up to the 2016 election, Trump and Clinton drew modest attention from members of Congress on Facebook, and then-candidate Trump received less support from members of his party than Clinton did from hers. Between each party's convention and Election Day, Democrats in Congress posted in support of Clinton substantially more often (a total of 1,614 posts) than Republicans posted in support of Trump (a total of 690 posts). However, Republicans opposed Clinton in 2,041 posts, far more than Democrats expressed opposition to Trump (1,383 posts).

Legislators also weighed in on the two candidates using both their unofficial and official accounts. In total, 154 (11%) of the 1,383 pre-election posts in which Democrats opposed then-candidate Trump came from legislators' official accounts. Republicans leaned even more heavily on their official accounts when it came to candidate-focused election outreach, using those accounts to express opposition to Clinton in about one out of every four posts (26%) that did so across both account types (522 of 2,041 total posts). Congressional rules prohibit legislators from posting campaign-related content on their official accounts, but these restrictions may not apply to posts that mention political candidates outside the context of elections. Clinton was often discussed in the context of her role as secretary of state under President Obama, and Republican opposition to Clinton frequently focused on issues such as the attack on the U.S. Embassy in Benghazi, Libya, and the FBI investigation into her use of a personal email account as secretary of state. In contrast, posts by Democrats that opposed Trump focused on his role as a presidential candidate and public figure.

2. How the Facebook audience engaged with congressional posts

When members of Congress posted in support of or opposition to political candidates, their Facebook audiences engaged more compared with posts that did not. At the same time, an

increasingly large proportion of "reactions" to congressional posts – a set of emoji introduced by Facebook in February 2016 to capture user's responses – featured the "anger" reaction. Over the full range of the study, legislators' Facebook audiences liked congressional posts over 481 million times, created more than 45 million comments in response to posts, and shared the posts over 141 million times. Those audiences also used the "angry" reaction more than 24 million times.
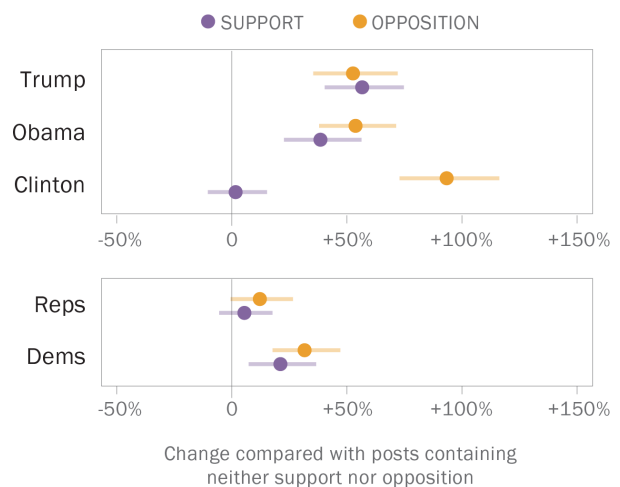
This analysis uses statistical models to examine how the presence of support or opposition in a Facebook post can affect the number of likes, comments and shares it receives. The models help account for other factors that could impact that post's engagement, such as the number of followers a given post's author has and when it was posted. See methodology section for additional details. All of the reported results are based on the entire three-year time frame.

The Facebook audience engaged with oppositional posts more than with ones that didn't take sides

Posts opposing Presidents Trump and Obama and former Secretary Clinton drew more likes from Facebook audiences compared with posts that didn't express political support or

Posts opposing political actors received more likes, especially when opposing Clinton

*Estimated percentage change in Facebook **likes** for posts containing support or opposition to ...*

Change compared with posts containing neither support nor opposition

Note: Points are estimated likes for posts that express opposition and support directed at labeled political figures and groups. Estimates are predictions from a multiple regression OLS model that includes random intercepts for each member of Congress and week in the data, alongside indicator variable for each presidential administration, legislator party and account type. The lines in the figure represent the 95% confidence interval for each estimate. See methodology section for additional details.
Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

opposition. On average, posts opposing Trump received 53% more likes, posts opposing Obama received 54% more and posts opposing Clinton received 93% more likes. However, this pattern
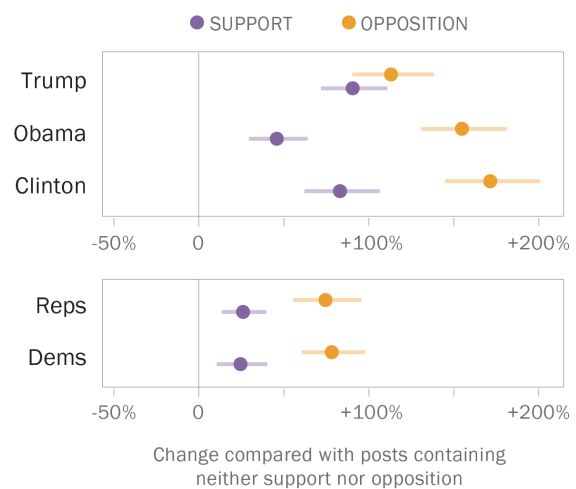
was more muted when posts opposed Democrats or Republicans more generally: Those posts received 32% and 12% more likes than posts that didn't take sides. These findings parallel earlier research that examined the relationship between political disagreement and Facebook engagement.

Among posts expressing political support, the results were less consistent. Posts supporting Trump, Obama and Democrats received an estimated boost in likes of 56%, 38% and 21%, respectively, relative to a post that did not contain any support or opposition. But the boost in likes for posts that supported Clinton or Republicans was much smaller: 2% and 5%, respectively.[7]

The Facebook audience was also more likely to leave comments on posts that expressed either support or opposition than on posts that did neither. When it came to comments, oppositional posts were consistently more likely to result in comments than posts that expressed support. Posts that opposed Clinton received the largest boost in comments, garnering 171% as many comments as the average post that did not express either support or opposition. Similarly, posts opposing Obama and Trump received 155% and 113% more comments. Posts expressing support for Trump, Obama and Clinton received boosts in comments of 90%, 46% and 83%, respectively. A similar but less pronounced pattern emerged for posts opposing and supporting the two political parties.

Post opposing or supporting political actors received *more* comments than posts that didn't take sides

*Estimated percentage change in Facebook **comments** for posts containing support or opposition to …*

Change compared with posts containing neither support nor opposition

Note: Points are estimated comments for posts that express opposition and support directed at labeled political figures and groups. Estimates are predictions from a multiple regression OLS model that includes random intercepts for each member of Congress and week in the data, alongside indicator variable for each presidential administration, legislator party and account type. The lines in the figure represent the 95% confidence interval for each estimate. See methodology section for additional details. Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017. "Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

---

7. The reported results are based on of an analysis of the entire three-year time frame, but similar effects are also observed for the period between July 26 and Nov. 8, 2016, prior to Trump's election, when both Clinton and Trump were competing as candidates.

When it came to Facebook shares, which refer to users reposting congressional posts for their own Facebook audiences, the difference between posts expressing opposition and posts expressing support was most pronounced. Posts that opposed Trump, Obama and Clinton received an estimated boost in shares of 141%, 155% and 225%, respectively. However, posts that expressed support for Trump or Obama received smaller increases, of 41% and 35%. And posts that supported Clinton actually received fewer shares than posts that did not take sides; a decrease of 33%. Posts supporting Democrats and Republicans also received fewer shares than posts that didn't take sides, decreasing 26% and 9% of shares, respectively.

Posts opposing political actors much more likely to be shared than supportive posts or posts that didn't take sides

Estimated percentage change in Facebook **shares** *for posts containing support or opposition to ...*

Change compared with posts containing neither support nor opposition

Note: Points are estimated shares for posts that express opposition and support directed at labeled political figures and groups. Estimates are predictions from a multiple regression OLS model that includes random intercepts for each member of Congress and week in the data, alongside indicator variable for each presidential administration, legislator party, and account type. The lines in the figure represent the 95% confidence interval for each estimate. See methodology section for additional details.
Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.
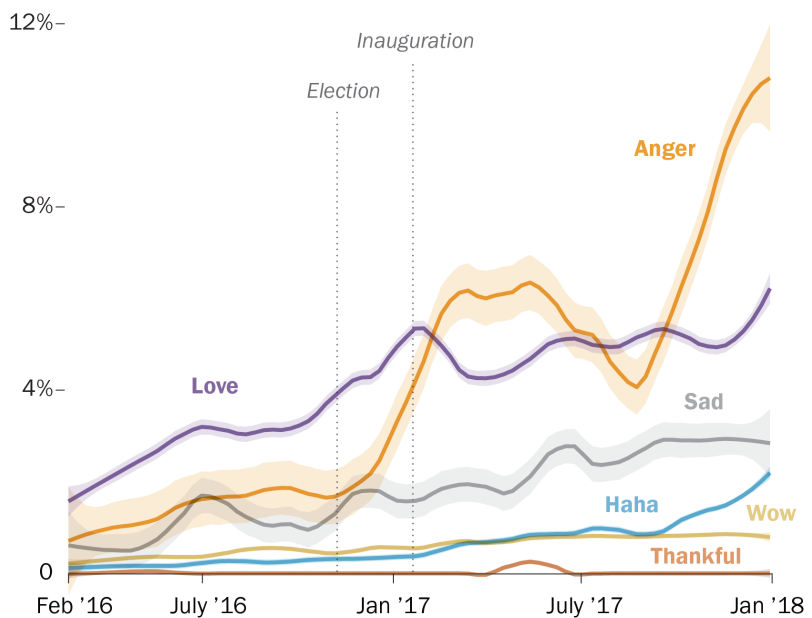"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Angry reactions from the Facebook audience increased; posts expressing opposition received most anger

Legislators' Facebook audiences became much more likely to react to posts with Facebook's "angry" button in the wake of the 2016 election. Prior to the election (but after the "angry" feature was released), just 1% of all reactions to posts by Democrats were angry. After the election, that share increased to 5%, on average. Among Republicans, the share of angry reactions increased from 2% before the election to 6% after. While "likes" remain the most common reaction, "angry" was the most frequently used of the six alternatives (such as "haha," "wow," and "love"). This has not always been the case. Prior to Trump's inauguration, the "love" reaction was the most commonly used alternative to "likes," but it has since been largely eclipsed by "angry." The use of angry reactions to congressional Facebook posts rose throughout 2017, reaching its highest observed rates at the end of the year, comprising 9% of all reactions to the average Democrat's posts in December 2017, and 13% of the average Republican's.

After 2016 election, angry reactions overtook love reactions from the Facebook audience

% of reactions to posts that were ...

Note: The chart shows the composition of reactions to the average member's post over time, as a proportion of the total, including likes. Likes, the most common reaction, are excluded from the graph, but included in the calculation of each reaction's share of the total. The shaded regions are the 95% confidence band for the estimated trends.
Source: Pew Research Center analysis of Facebook posts created by members of Congress between Feb. 23, 2016 and Dec. 31, 2017.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Angry reactions were especially likely to ensue when posts expressed political opposition. Posts that expressed opposition to Trump received an estimated five times as many angry reactions as posts that did not express support or opposition toward any figure or group. When Democrats

expressed opposition to Republicans, they earned six times as many angry reactions, on average. Because the emotional reactions were not available across the entire timeframe, this analysis is based upon posts created between Feb. 23, 2016 (the day before the reactions were released) and Dec. 31, 2017.
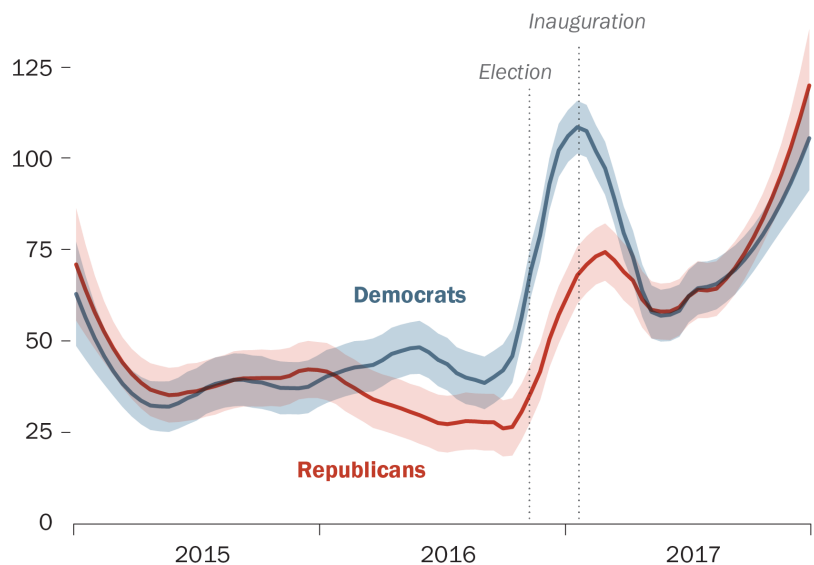
Comments increased after Trump won 2016 election

In the wake of the 2016 election, legislators' Facebook audiences became much more likely to engage online by posting comments. Across both parties, the average number of comments per post increased more than 200%, comparing the time period before the 2016 election with all posts created afterward.

Just as the use of angry reactions peaked in December 2017, the Facebook audience became more likely to post comments at the end of that year, averaging 125 comments per post. By contrast, in December 2015, the average post received just 65 comments. This pattern is even more pronounced when examining the median number of comments per post. That number more than tripled, from seven comments per post at the end of 2015 to 22 comments by December 2017.

Comments on congressional posts increased after 2016 election and in final months of 2017

Average number of comments per post

Note: The chart shows the smoothed average number of comments per post for both posts by Democrats and Republicans in Congress. The shaded regions are the 95% confidence band for the estimated trends.
Source: Pew Research Center analysis of Facebook posts created by members of Congress between Jan. 1, 2015 and Dec. 31, 2017.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Acknowledgments

This report is a collaborative effort based on the input and analysis of the following individuals. Find related reports online at pewresearch.org.

Methodology

To analyze the content of congressional Facebook posts, researchers studied a complete set of posts created by members of the U.S. Senate and House of Representatives and displayed on their official and unofficial public pages between Jan. 1, 2015, and Dec. 31, 2017. Data from previous Congresses were not available. Researchers used the Facebook Graph API to download the posts from members' public pages. Next, they classified the content of the posts using a combination of expert coding, crowd-sourcing and machine learning. Once the posts were classified, researchers used regression analysis to evaluate the relationship between particular posts and the level of engagement Facebook users had with each type of post.

Data for this report came from three sources:

1. Facebook posts collected directly from legislators' public pages through Facebook Graph Application Programming Interface (API)

2. Human coding of Facebook posts completed by Amazon Mechanical Turk workers and researchers within Pew Research Center.

3. Existing measures of congressional ideology based on roll-call voting.

Terminology

Random effects regression is a type of statistical model used here to estimate the relationship between particular kinds of posts and the amount of engagement those posts received. This model is appropriate for the structure of the collected Facebook data – a large number of posts that are associated with a smaller number of legislators – and adjusts for the fact that some legislators have much larger online audiences than others.

DW-NOMINATE is a measure of political ideology that places members of the U.S. House and Senate on a liberal-to-conservative ideology scale according to their roll-call voting history in each legislative session of Congress.[8] The scale ranges from -1 (very liberal) to 1 (very conservative) across all Congresses. For the time period studied here it ranges between -0.77 and 0.99. Very liberal and very conservative legislators are defined as those with DW-NOMINATE scores in the

---

8 See Lewis, Jeffrey B., Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet. 2017. "Voteview: Congressional Roll-Call Votes Database." https://voteview.com/

furthest left 10% and furthest right 10% of all DW-NOMINATE scores. Moderates are defined as legislators with DW-NOMINATE scores in the middle 20% of all scores.

Data collection

The first step in the analysis was to identify each member's Facebook pages. Many members of Congress maintain multiple social media accounts, consisting of one or more "official," campaign or personal accounts. Official accounts are used to communicate information as part of the member's representational or legislative capacity, and U.S. Senate and House members may draw upon official staff resources appropriated by Congress when releasing content via these accounts. Personal and campaign accounts may not draw on these government resources under official House and Senate guidelines.

Researchers started with an existing dataset of official and unofficial accounts for members of the 114th Congress and expanded it with data on members of the 115th Congress from the open-source @unitedstates project. Official accounts are managed by congressional staff and used for official legislative business, while unofficial Facebook accounts are used in a personal or campaign capacity. Researchers also manually checked for additional accounts by reviewing the House and Senate pages of members who were not found in the initial dataset. Every account was then manually reviewed and verified.

The research team first examined each account's Facebook page and confirmed that it was associated with the correct politician. All misattributions were manually corrected by Center experts, resulting in a list of 1,129 total Facebook accounts. Accounts were then classified as official or unofficial based on the links to and from their official ".gov" pages. Accounts were considered official if they were referenced by a member's official house.gov or senate.gov homepage. Congressional rules prohibit linking between official (.gov) and campaign websites or accounts, as well as linking from an official site or account to a personal site or account.

In cases where it was not clear that a Facebook page had ever been used in an official capacity (particularly for members that are no longer in Congress with active webpages), the most recent historical copy of the member's official webpage was manually reviewed using the Library of Congress online archive to determine if a link to the account had been present when the webpage was active. The resulting list of all official accounts for members of the 114th and 115th Congresses was then used to collect the Facebook posts published by each page between Jan. 1, 2015, and Dec. 31, 2017.

Using the Facebook Graph API, researchers obtained Facebook posts for members of the 114th Congress (2015-2016) between Dec. 30 and 31, 2016, so that members who left office before the 115th Congress began would be included in the sample. Between Jan. 2 and Jan. 6, 2018, researchers obtained posts for members of the 115th Congress (2017-2018).

After obtaining posts, researchers checked the combined dataset and identified a small number of duplicate posts from members of Congress who served in both the 114th and 115th Congresses. The duplicates had been introduced due to changes in the public pages' unique Facebook API identifiers, resulting in mismatches between the latest copy of certain posts and older copies that had previously been collected. These duplicates frequently occurred on posts that had been edited or modified slightly – often with nearly identical time stamps and only single character variations (e.g., deleting a space). The unique identifiers of these duplicates were also very similar themselves, differing by only a few digits in specific locations of the identifier string. In all of these cases, the posts' time stamps were rarely separated by more than a few minutes, and were always within 24 hours of each other.

An additional set of duplicates was also found among posts that were produced by pages that had changed names at some point during the time frame. These posts most frequently occurred after the end of election season, when a number of politicians change the titles of their Facebook pages – removing suffixes such as "for Congress" or adding honorifics like "Senator" to their name. In these cases, the time stamps and content of the posts were perfectly identical but the prefixes of the posts' unique identifiers were different.

There were several patterns across multiple post fields that appeared to distinguish duplicates from unique posts. However, no clear set of rules could be identified that comprehensively explained these patterns, so researchers employed a machine learning approach to isolate and remove the duplicate posts.

First, researchers scanned the entire set of posts for each account using a sliding window of two days and identified all pairs of potential duplicates within each window that matched either of the following criteria:

1. Identical time stamps

2. TF-IDF cosine similarity of 0.6 or above, and a Levenshtein difference ratio of 60% or higher on the text of the post

From these "candidate duplicates," a random sample of 1,000 pairs was extracted and manually reviewed. Researchers identified whether or not the two posts in each candidate pair were in fact duplicates. Only 24% were determined to be true duplicates. These results were then used to train a machine learning algorithm using 750 of the pairs to train the model and 250 to evaluate its performance. Researchers trained a random forest machine learning model using a variety of features – or model inputs – representing the similarity of the two posts across different fields and interactions between these features. The features that best identified duplicates included whether the two posts shared an identical time stamp, the number of digits that overlapped between the posts' ID numbers and the difference between the posts' time stamps in seconds. The resulting model achieved high performance, with an average precision and recall of 98%. Of the 250 potential duplicate pairs used to evaluate the model, it missed only four duplicates and correctly classified the remaining 246.

The model was then applied to the entire collection of potential duplicates, removing duplicates when detected. In total, 30,508 posts (4% of the original dataset) were identified as duplicates and excluded before the analysis began.

The final dataset included only those posts that were produced by a member's primary official and/or unofficial Facebook accounts during the time in which they were serving a term as a representative or senator in Congress. The resulting dataset contains 737,598 Facebook posts from 599 different members of Congress who used a total of 1,129 official and unofficial accounts. Photo and video posts were included in this analysis. The findings presented in this report exclude posts by nonvoting representatives, and only posts produced by members that were active online in a given Congress, defined as members that produced at least 10 Facebook posts during each Congress. Members that meet this threshold for just one of the Congresses are only included for that specific Congress.

Content coding

Researchers created an online content classification interface to label the Facebook posts. This interface included the time the post was created, its author, the party of its author, an indication of who was president when the post was created, and all links and captions associated with the post. Content coders were instructed to indicate whether the post mentioned Donald Trump, Barack Obama, Hillary Clinton, federal agencies, Democrats and/or Republicans. If the post did mention any of these entities, coders could select whether the post expressed support/agreement with them, opposition/disagreement with them, or if it was angry/insulting toward them. Finally, coders were instructed to note whether the post asked the reader to engage in political actions,

whether the post was election related, or whether the post concerned a specific local event, institution, organization or individual.

Detailed instructions were provided for each of these items (note that not all items are used in this analysis); these instructions appear in Appendix A. An example of the content classification window appears below:

Extracting samples

Researchers extracted two samples of the Facebook posts for coding: one large sample of 11,000 posts to be classified by workers on Amazon Mechanical Turk, and a smaller sample of 1,100 posts for internal evaluation. In order to ensure that researchers examined enough posts that contained expressions of opposition or support, researchers drew samples of Facebook posts using keyword oversampling. Building on previous research that also used this approach, a series of regular expressions were used to disproportionately select posts that contained keywords likely to co-occur with expressions of political opposition or support. The patterns used appear in Appendix B.

Subsequent analysis adjusted for this process by using probability weights computed by comparing all of the keyword combination strata against the full population. The final proportion

of posts containing a keyword in each set are reported below (some posts match multiple keyword sets; proportions are not mutually exclusive).

The samples were extracted in two phases. Initially, an evaluation sample of 650 posts was extracted from a sampling frame of 447,675 Facebook posts created by members' official Facebook accounts between Jan. 1, 2015 and July 20, 2017. These were coded by multiple in-house experts to establish a baseline of interrater reliability, and also coded each by five different Mechanical Turk users, to be compared against the in-house baseline. A larger sample of 6,499 different posts was extracted from the frame and coded only by Mechanical Turk users to serve as training data for the

Keyword oversamples

% of posts that contain keyword(s) for ...

Keywords	Sampling frame (Population)	Training sample (N=11,000)	Evaluation sample (N=1,100)
No match	24	6	3
Clinton	1	16	11
Trump	7	21	21
Support	50	60	59
Republicans	9	32	29
Agencies	40	49	49
Democrats	6	32	25
Anger	19	41	55
President	6	17	21
Obama	7	20	26

"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

machine learning classifiers. Mechanical Turk workers completed these assignments in the fall of 2017.

After researchers determined that they had reliably identified all of the unofficial Facebook accounts for members of Congress and determined that nearly every member had both an official and unofficial account, the scope of the sampling frame was expanded to include unofficial posts, increasing its size by 40%. The sampling frame was then also extended to encompass a wider time frame, including three full years (from the beginning of 2015 through the end of 2017). Expanding the time frame increased the post count by an additional 17%. The final population consisted of 737,598 posts – 65% larger than the original dataset. As a result, the training and evaluation samples were expanded in a proportional fashion. To do so, new rows were added to the existing samples by drawing exclusively from posts from the new accounts that were created during the original time frame, increasing the sample sizes up by 40%. These new rows were drawn using post-stratification weights based on the party of the member, whether the account was official or not, and the month during which the post was made. These weights ensured that the resulting sample preserved representative proportions across those dimensions. The newly added posts were nearly all unofficial posts, although a handful of the additional posts were drawn from several

new official accounts that had been created since the initial sample was drawn, due to special elections.

The process of adding additional posts was divided into 10 different iterative steps in order to accurately represent the characteristics of the population data. Prior to each step, researchers computed post-stratification weights using the aforementioned post characteristics and then multiplied by weights based on the original sample's keyword oversampling proportions in order to preserve the oversampling rates at which new posts would be added. One-tenth of the desired number of new posts was then sampled using these weights, after which the weights were recomputed. Through this iterative process, the weights were adjusted to compensate for random variation, allowing the sample to smoothly converge to the correct population proportions in each stratum as it approached its target size.

Finally, the newly expanded samples were expanded once more using this same process, this time increasing their size 17% to proportionally match the expansion of the time frame. These new rows were drawn exclusively from both official *and* unofficial posts produced *after* the time frame of the initial sample and were again extracted using post-stratification weights on month, official account status and political party. Finally, to round the sample sizes up to an even number, a small number of additional rows were drawn from the entire final sampling frame using the post-stratification weights, bringing the evaluation sample to 1,100 posts, up from 650, and the training sample to 11,000 posts, up from 6,499. The new posts were then coded by in-house experts and Mechanical Turk users in early 2018.

Every post in the training and evaluation samples was coded by five different Mechanical Turk users, and every post in the evaluation sample was also coded by in-house experts. A total of five in-house experts divided up the coding tasks and each post in the evaluation sample was coded by two different coders.

Validating crowd-sourced data

To assess how closely the expert coders agreed on these labeling tasks, researchers computed the average pairwise Cohen's Kappa, weighted by the number of overlapping tasks each pair of coders completed. To evaluate the crowd-sourced results, the Mechanical Turk results were first collapsed using a 2-out-of-5 threshold that was determined to be effective at maximizing agreement in prior research. Researchers then reviewed all coding disagreements between the expert coders and resolved the disputes, resulting in a single label for each post. These single ground-truth expert labels were then compared against the collapsed crowd-sourced results, using Cohen's Kappa to compute reliability. The results indicate that the agreement between the collapsed Mechanical

Turk labels and those from in-house coders are comparable to the agreement between individual in-house coders.

Item agreement

Average pairwise Cohen's Kappa

Classifier	Expert coders	Expert coders vs. Mechanical Turk
Local topic	0.77	0.80
Oppose Clinton	0.92	1.00
Oppose Democrats	0.65	0.78
Oppose Obama	0.84	0.91
Oppose Republicans	0.93	0.94
Oppose Trump	0.89	0.87
Support Clinton	0.90	0.74
Support Democrats	0.64	0.65
Support Obama	0.78	0.82
Support Republicans	0.75	0.74
Support Trump	0.81	0.80

Note: Kappa describes inter-coder reliability adjusted for chance agreement. The Kappa's reported here are weighted to account for keyword oversampling in post selection.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Content classification

Cleaning the text

To produce a dataset useful for machine learning, the text of each post was converted into a set of features, representing words and phrases. To accomplish this, each post was passed through a series of pre-processing functions. First, to avoid including words that could bias the machine learning models toward particular politicians or districts, a set of custom stopword lists were used to filter out names and other proper nouns, comprised of the following:

- A list of 318 common English stopwords, taken from the Glasgow Information Retrieval Group

- A list of 9,938 first and last names, taken from a Pew Research Center database of 14,289 current and historical politicians and filtered using WordNet

- A list of 896 state names and state identifiers (e.g. "West Virginian", "Texan")

- A list of 18,128 city and county names, taken from a Pew Research Center database of geocoded campaign contributions, and filtered using WordNet

- A list of 24 month names and abbreviations

- 377 additional stopwords, manually identified through a process of iterative review by Center researchers

Some people and locations have names that are also common English words, some of which are used far more frequently as the latter. To avoid unnecessarily excluding these words from our training data, potential stopwords were assessed using WordNet, which provides information on a word's alternative definitions and where they fall on a spectrum of generality to specificity (using a hyponymy taxonomy). Words were flagged as being common and/or versatile enough to be included in the training data and removed from the stopword list if they met two or more of the following criteria:

- The word has more than two different definitions (synsets)

- One or more of the word's definitions (synsets) had a variation (lemma) with a depth of less than five (indicating generality)

- One or more of the word's definitions (synsets) had at least two variations (lemmas)

After combining all of these lists into a single set of stopwords, an additional 77 words were removed, based on a list compiled by Center researchers during a process of manual iterative review. This ultimately resulted in a list of 27,413 stopwords that were removed from the text of all posts. After removing stopwords, the text of each post was lowercased and URLs and links were removed using a regular expression.

Common contractions were expanded into their constituent words, punctuation was removed and each sentence was tokenized using the resulting whitespace. Finally, words were lemmatized (reduced to their semantic root form) and filtered to those containing three or more characters.

Extracting features

Machine learning classifiers were trained using a variety of features:

1. Term-frequency inverse-document frequency matrices with l2 normalization, containing 1 to 4 grams with a minimum document frequency of 10 and maximum document proportion of 90%

2. Term frequency matrices with no normalization, containing 1 to 4 grams with a minimum document frequency of 10 and maximum document proportion of 90%

3. Features based on 300 Word2Vec dimensions extracted using a pre-trained model (Google News); the average, maximum, minimum, standard deviation, and median of each vector, aggregated from the words in each post that matched to any of the 500,000 most frequent words in the pre-trained model

For target-specific support and opposition models, an additional set of features were extracted using the relevant target-specific regular expression from the keyword oversampling patterns described earlier, as well as the more general "anger" and "support" regular expressions:

1. TF and TF-IDF matrices with the same parameters as 1 and 2 above, extracted from the subset of sentences in each post that contained keywords relevant to the classification variable, identified using the oversampling regular expression relevant to the target being classified

2. A set of features based on the "anger" keyword oversampling regular expression: a boolean flag indicating whether a post contained a match, the total number of matches in the post, and the squared and logged values of the total number of matches

3. The same set of features in 2, except for the "support" keyword oversampling regular expression

4. The same set of features in 2 and 3, except for the "president" keyword oversampling regular expression (Trump/Obama/Clinton models only)

Model training

Researchers used machine learning algorithms to classify the entire set of Facebook posts used in this report based on the training results from human coders. Training data was weighted and evaluated using traditional sample weighting based on the population proportions of oversampled keywords. Additional weights were used only during the model training process (not during evaluation), weighting cases using the inverse proportion of their class. The XGBoost classification algorithm was used for each of the classifiers used in this report.

Researchers trained the algorithms separately for posts authored by Republican members of Congress and those authored by Democratic members. As a result, the models predicted opposition to Trump and Republicans and/or conservatives within the set of posts created by Democrats in Congress, but not for posts created by Republicans. For posts authored by Republicans, the models predicted opposition to Obama, Democrats and/or liberals, and Clinton. Likewise, the models predicted support for Trump and Republicans and/or conservatives within the set of Republican-authored posts, but not posts authored by Democrats, and vice versa. Overall, 46% of the training data and 47% of the test data were used for Democratic post classification models, leaving 54% and 53% of the training and test data for Republican models, respectively. Posts by both Democrats and Republicans in Congress were used together when training and evaluating the local topic mention classification algorithm.

Rather than use the discrete predictions produced by the classifiers, a custom threshold was identified for each model and applied to the raw probabilities of its predictions. To find this threshold, the models were trained on the full training dataset and applied to the evaluation posts. These predictions were compared against the in-house expert codes across full range of possible thresholds (0-1), evaluated based on the precision and recall of the model at each possible cutoff. Next, the model was trained using a five-fold split of the training data, and evaluated on each of

the five folds. For each fold, the classifier was again evaluated across the full range of possible thresholds.

This time, the cutoffs were evaluated in terms of the minimum of the precision and recall for that fold *unless* the precision or recall at that threshold had been lower in the evaluation set, in which case the latter was taken as the score. This resulted in a score at each potential threshold, representing the model's worst performance at that cutoff across both the held-out fold and the expert evaluation set. The threshold that maximized this score was selected for each fold, and these five thresholds were averaged together to produce a final threshold for the classifier. In effect, this resulted in classifiers that had optimal performance across multiple hold-out folds and the expert evaluation set, and that achieved a balance between precision and recall, ensuring that false positive and false negative rates were as equal as possible.

The following table shows the performance for each model, including the five-fold precision and recall averages and the precision and recall when the hand-coded expert and Mechanical Turk data to the machine learning predictions. In most cases, the machine-learning classifiers agree more closely with in-house experts than they do with the aggregated Mechanical Turk labels.

Model performance

Classifier	Five-fold cross validation		Experts		Mechanical Turk	
	Precision	Recall	Precision	Recall	Precision	Recall
Local topic	0.76	0.80	0.78	0.77	0.81	0.81
Oppose Clinton	0.78	0.85	0.92	0.79	0.92	0.79
Oppose Democrats	0.62	0.67	0.80	0.84	0.64	0.68
Oppose Obama	0.81	0.86	0.83	0.74	0.81	0.73
Oppose Republicans	0.79	0.81	0.84	0.85	0.82	0.84
Oppose Trump	0.83	0.84	0.88	0.96	0.84	0.86
Support Clinton	0.81	0.78	0.91	0.92	0.83	0.57
Support Democrats	0.57	0.62	0.73	0.65	0.46	0.66
Support Obama	0.61	0.63	0.74	0.79	0.75	0.94
Support Republicans	0.54	0.60	0.74	0.76	0.55	0.79
Support Trump	0.68	0.74	0.71	0.63	0.55	0.49

Note: Precision is a measure of how well a model avoids false positives, and recall measures how well the model identifies true positives/avoids false negatives. All performance statistics reported here are weighted to account for keyword oversampling in post selection.
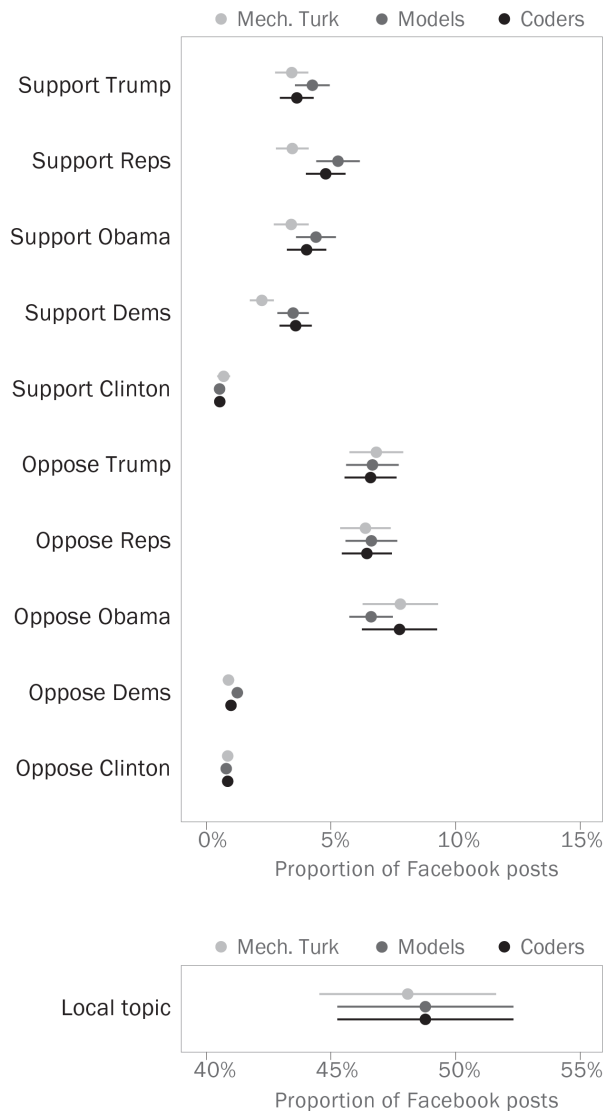"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Evaluating potential machine learning bias

After researchers classified individual posts, they aggregated those classification decisions across members, parties and time periods to arrive at global estimates of the proportion of posts that contained support or opposition directed at the individuals and groups described above. However, this procedure risks producing biased estimates of overall rates if the models used to classify posts systematically over- or under-state the prevalence of particular kinds of posts. To assess this risk, researchers calculated the global proportion of posts containing a positive value for each classification, across the expert coder sample (1,100 posts, all human coded), the Mechanical Turk sample (11,000 posts, all human coded) and the sample of predicted classifications (all remaining posts, machine coded). These estimates are shown above. In all cases, the estimated proportion of posts containing each kind of content is very similar across the three estimation strategies. None of the differences across estimates within each classification type are statistically significant, indicating that the model predictions are unlikely to be biased for the full sample of posts.

Estimates from machine learning models closely align with those from human-coded data

% of Facebook posts containing ...

Note: The lines in the figure represent the standard error for each estimate.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Engagement analysis

Researchers used multiple regression models in order to estimate the relationship between certain kinds of posts and the rate at which the Facebook audience engaged with those posts. These models are useful because they include terms (random intercepts) that act as baselines for drawing comparisons across different groups (in this case, groups of posts all made by the same member of Congress or groups of posts created during the same week).

Because some members of Congress are particularly likely to both have large numbers of online followers and to post particular kinds of content, researchers used models to help account for the fact that some posts were seen by many more Facebook users than other posts. The models included interaction terms between each "opinion" variable (meaning that the post either supported or opposed the political figure or group attached to that opinion) and the partisanship of the post's author. For Republicans, the predicted number of likes, comments and shares for each political target was based upon the sum of the predictions from component terms in the interaction model (Republican, opinion of some target and the interaction term that captures both), with all other variables held at their mean. For

Facebook engagement analysis

Coefficients and standard errors...

Variable	Log(Likes)	Log(Comments)	Log(Shares)
Republican	-0.125 (0.033)	0.111 (0.029)	-0.104 (0.030)
Trump Opinion	0.181 (0.003)	0.296 (0.003)	0.348 (0.004)
Clinton Opinion	0.007 (0.009)	0.235 (0.009)	-0.145 (0.010)
Obama Opinion	0.139 (0.005)	0.145 (0.005)	0.115 (0.006)
Dem. Opinion	0.082 (0.004)	0.083 (0.004)	-0.111 (0.005)
Rep. Opinion	0.049 (0.003)	0.216 (0.003)	0.294 (0.004)
Trump Opinion × Republican	0.010 (0.005)	-0.039 (0.005)	-0.219 (0.006)
Clinton Opinion × Republican	0.276 (0.011)	0.169 (0.011)	0.603 (0.013)
Obama Opinion × Republican	0.044 (0.006)	0.233 (0.006)	0.245 (0.007)
Dem. Opinion × Republican	0.036 (0.007)	0.147 (0.007)	0.362 (0.009)
Rep. Opinion × Republican	-0.027 (0.005)	-0.125 (0.005)	-0.327 (0.006)
Before Election	0.020 (0.022)	-0.184 (0.021)	-0.020 (0.015)
After 2017 Inauguration	0.170 (0.017)	0.142 (0.016)	0.081 (0.015)
Senator	0.560 (0.043)	0.646 (0.038)	0.507 (0.039)
Official Account	0.060 (0.002)	0.178 (0.002)	0.085 (0.002)
Intercept	1.596 (0.032)	0.670 (0.029)	0.695 (0.028)
Observations	737,494	737,494	731,493
Random effect for legislator	Yes	Yes	Yes
Random effect for week	Yes	Yes	Yes

Note: Dependent variables are logged (base 10) to account for a small number of posts with very high levels of engagement. "Opinion" variables are a 0 or 1 depending on whether the creator of the post expressed support or opposition to the target.
"Taking Sides on Facebook: How Congressional Outreach Changed Under President Trump"

PEW RESEARCH CENTER

Democrats in Congress, the Republican term and the interaction terms always had a coefficient of 0, and the predicted number of likes, comments or shares for each political target was based on the coefficients for the opinion variables, alongside the coefficients for all other control variables set to their means.

Researchers estimated percentage increases in likes, comments and shares relative to a post by the average member of Congress that did not express an opinion (supportive or oppositional) about any of the political targets included here. For several thousand posts, the number of shares were not available. These posts are excluded from the shares regression model.

Appendix A: Content classification instructions

Human coders read the following instructions when classifying Facebook posts. They were presented with the post, when the post was created, the author of the post, information about who was president when the post was made, and each of the numbered items as possible selections. If the coder clicked on any of the items, the interface displayed additional information about what the item should or should not include. Only some of the coding items are included in this analysis.

1. POLITICAL ACTION: Does the post invite the reader to vote, volunteer, call or send messages, sign a petition, attend an event/rally/protest, or make a donation? Explicitly asks for readers to take political action, such as attending an event, volunteering, voting, donating, etc.

    Includes invitations to:
    Vote, sign a petition, call elected officials
    Join a protest, volunteer, or attend a town hall
    Make a donation
    Attend an upcoming event

    Does not include:
    Announcements for past events
    Descriptions of other individuals taking a political action
    Informational posts

2. LOCAL REFERENCE: Does the post mention a place, group, individual(s), or event in the politician's state or district? Has a specific reference to a place (town/city/county/state), group, individual(s), or event in a politician's district or state (or mentions the district or state itself)

    Includes:
    Discussion of government programs, facilities, or commemorative events in a politician's district or state
    Town hall meetings or other community events
    References to policies that primarily apply to a particular state, district, city, or community
    Mention of local or state places, organizations, or events

    Does not include:
    National policies

Regional issues or events
Other topics unrelated to localities

3. DONALD TRUMP: Does the post mention "Donald Trump, his administration, or his campaign"?
President Donald Trump himself, his administration or campaign, or one of his own decisions and actions. Only includes VP or other government, party, or campaign officials if they are explicitly tied to Trump or his administration. INCLUDES references to "Trumpcare" but NOT the AHCA.

Includes:
The President, The/This Administration, The White House, or other use of the word 'executive' when referring to President Trump (you can see who was in office at the time by looking below the post date)
Mentions of Donald Trump as a candidate or president-elect
Mentions of the Trump campaign
'Trumpcare' (but NOT the AHCA)

Does not include:
Vice President Pence, unless referred to as 'Trump's VP' or otherwise linked to Trump directly
Mentions of a member of Trump's administration, cabinet, or campaign, that are not explicitly linked to Trump in the text

4. BARACK OBAMA: Does the post mention "Barack Obama or his administration"?
President Barack Obama himself, his administration or campaign, or one of his own decisions and actions. Only includes VP or other government, party, or campaign officials if they are explicitly tied to Obama or his administration. INCLUDES references to "Obamacare" but NOT the Affordable Care Act or ACA.

Includes:
The President, The/This Administration, The White House, or other use of the word "executive" when referring to President Obama (you can see who was in office at the time by looking below the post date)
Mentions of Barack Obama as a candidate or president-elect
Mentions of the Obama campaign
"Obamacare" but NOT the Affordable Care Act or ACA.

Does not include:
Vice President Biden, unless referred to as "Obama's VP" or otherwise linked to Obama directly
Mentions of a member of Obama's administration, cabinet, or campaign, that are not explicitly linked to Obama in the text

5. HILLARY CLINTON: Does the post mention "Hillary Clinton or her campaign"?
   Hillary Clinton in any of her roles - as Secretary of State, Senator, or presidential candidate - or mentions of her campaign

   Includes:
   Hillary Clinton in any of her roles - as Secretary of State, Senator, or presidential candidate
   Mentions of her campaign, or "the campaign" when it clearly refers to Clinton's 2016 presidential bid

   Does not include:
   Mentions of Vice Presidential candidate Tim Kaine, unless he's explicitly associated with Clinton and her campaign

6. REPUBLICANS: Does the post mention "Republicans," "conservatives," or "conservative values"?
   Republican politician(s) (EXCEPT Trump) if their party or ideology is mentioned. Also includes the party itself, and the "conservative" ideology more generally. Does NOT include specific politicians UNLESS the text associates them with the Republican party or conservative ideology.

   Includes:
   Any Republican politician ONLY IF their party affiliation or conservative ideology is specifically mentioned
   The Republican Party, GOP, RNC, House Liberty Caucus
   Republican candidates for office

   Does not include:
   Mentions of specific politicians if their party affiliation or political ideology is unclear
   Mentions of party leaders like Paul Ryan or prominent candidates like Ted Cruz and Marco Rubio - if the post doesn't mention their political affiliation

7. DEMOCRATS: Does the post mention "Democrats," "liberals," or "liberal values?" Democratic politician(s) (EXCEPT Obama and Clinton) if their party or ideology is mentioned. Also includes the party itself, and the "liberal" or "progressive" ideology more generally. Does NOT include specific politicians UNLESS the text associates them with the Democratic party or liberal ideology.

    Includes:
    Any Democratic politician ONLY IF their party affiliation or liberal ideology is specifically mentioned
    The Democratic Party, DNC
    Progressives or the Progressive Caucus
    Democratic candidates for office

    Does not include:
    Mentions of specific politicians if their party affiliation or political ideology is unclear
    Mentions of party leaders like Chuck Schumer or prominent candidates like Hillary Clinton and Bernie Sanders - if the post doesn't mention their political affiliation

For all political entities listed in items 3-7, coders were asked to evaluate whether or not the post expressed:

8. SUPPORT: Support / agrees

9. OPPOSE: Opposes / disagrees

10. ANGER: Angry or insulting?

Appendix B: Regular expressions for keyword oversampling

Researchers used the following regular expression patterns to oversample posts in the training data.

Clinton	Hillary|Clinton
Trump	[Tt]rump
Support	\b[Ss]trong|\s[Gg]rateful|\b[Dd]edicated|\b[Pp]romise|\b[Ss]upport*|\b[Cc]onfident*|\b[Hh]onor*|\b[Aa]dvocate|\b[Pp]romot*|\b[Vv]ision|\b[Ss]trengthen*|\b[Aa]dvance*|\b[Oo]pportunit*|\b[Ss]pirit|\b[Hh]allmark|\b[Jj]oin*|\b[Aa]pprove|\b[Cc]ommitt*|\b[Aa]gree|\b[Ee]ncourag*|\b[Hh]onest*|\b[Cc]ollabora*|\b[Pp]rivilege*|\b[Pp]roud|\b[Hh]istoric|\b[Ff]riends|\b[Ii]mprov*|\b[Tt]ogether|\b[Ss]marter|\b[Cc]onsensus|\b[Mm]omentum|\b[Ee]ncouraged|\b[Ss]tand with|\b[Pp]leased|\b[Aa]dvanc*|\b[Pp]rovid*|\b[Ss]olidarity|\b[Gg]lad|\b[Ff]orward|\b[Ee]xcellent|\b[Ww]onderful|\b[Gg]ratif*|\b[Gg]ood [Nn]ews|\b[Ss]ervice|\b[Ll]eadership
Republicans	[Cc]onservative*|[Rr]epublican*|\bGOP\b|\bgop\b|[Tt]ea [Pp]art*|Freedom Caucus
Agencies	[Aa]genc*|[Bb]ureau*|[Dd]epartment*|[A-Z]{3,5}|regulat*|VA|[Dd]ept
Democrats	[Ll]iberal*|[Dd]emocrat*|[Pp]rogressiv*
Anger	\b[Cc]rooked|\birresponsib*|\bunconstitution*|\boutrag*|\bcorrupt*|\breckless|\bdangerous*|\bbigot*|\bthreat*|\bmysogy*|\bnasty|\bcondemn|\bworst|\bundermin*|\bintoleran*|\bdestest*|\banti*|\bracis*|\bhorrif*|\bdivisiv*|\binexcusab*|\bbetray*|\bweak*|\bshame*|\breprehensib*|\bviolat*|\bdisturb*|\bagenda|\bfalsehood|\bcredibility|\bunfit|\bresign|\blies|\blied|\blying|\bendanger*|\bnegative|\bpartisan*|\burge|\bunconscionab*|\boverzealous*|\babysmal*|\bmistake*|\bthreaten*|\bnaive*|\bdestroy*|\bharmful|\bunnecessar*|\bfailure*|\bextreme|\bstruggle|\breject*|\bshock*|\belite|\belistist|\bcrony*|\bsquander*|\bradical|\bcapitulat*|\bneglect*|\bhostage|\bpuppet*|\bburdensome|\bunprecedented|\bobstruct*|\bideolo*|\bdiscriminat*|\binexcus*|\bwrong*|\bworse*|\bdodg*|\bendanger*|\bunconscionabl*|\bdevastat*|\brisky|\bcostly|\bblatant*|\bfailed|\bdisgrace*|\bsurrender*|\bunfit|\bmishandl*|\bcareless*|\bignoran*|\baccountable|\binsult*|\babsurd|\bhorribl*|\bimpeach*|\bresign*|\bscandal*|\bpower grab|\bbroken
President	([Tt]he|[Tt]his|[Hh]is)([a-z\s\W]{0,20})[Aa]dministration|White House|President\'s desk|([Tt]he|[Oo]ur|[Tt]his) [Pp]resident*
Obama	[Oo]bama