

# Is the internet causing political polarization?

## Evidence from demographics

Levi Boxell, *Stanford University*\*

Matthew Gentzkow, *Stanford University and NBER*

Jesse M. Shapiro, *Brown University and NBER*

March 2017

### Abstract

We combine nine previously proposed measures to construct an index of political polarization among US adults. We find that the growth in polarization in recent years is largest for the demographic groups least likely to use the internet and social media. For example, our overall index and eight of the nine individual measures show greater increases for those older than 75 than for those aged 18–39. These facts argue against the hypothesis that the internet is a primary driver of rising political polarization.

---

\*E-mail: lboxell@stanford.edu, gentzkow@stanford.edu, jesse\_shapiro\_1@brown.edu. We acknowledge funding from the Brown University Population Studies and Training Center and the Stanford Institute for Economic Policy Research (SIEPR). The Pew Research Center, American National Election Study, and the relevant funding agency/agencies bear no responsibility for use of the data or for interpretations or inferences based upon such uses.

# 1 Introduction

By many measures, Americans have become increasingly polarized in recent decades.<sup>1</sup> In 1960, roughly 5 percent of Republicans and Democrats reported that they would “[feel] ‘displeased’ if their son or daughter married outside their political party;” by 2010, nearly 50 percent of Republicans and over 30 percent of Democrats “felt somewhat or very unhappy at the prospect of inter-party marriage” (Iyengar et al. 2012). The relative favorability of party affiliates towards their own party has increased by over 50 percent from 1980 to 2015 (Gentzkow 2016), and the proportion of voters voting for the same party in both presidential and House elections has increased from 71 percent of reported voters in 1972 to 90 percent in 2012 (American National Election Study 2015c).

Many authors attribute this trend, in part, to the rise of social media and the internet in general. Sunstein (2001; 2009; 2017) and Pariser (2011) argue that the internet may create echo chambers in which individuals hear news only from like-minded sources. Gabler (2016) argues that “social media contribute to... more polarization as the like-minded find one another and stoke one another’s prejudices and grievances, no matter what end of the political spectrum.” Haidt (2016) calls social media “one of our biggest problems” and notes that “so long as we are all immersed in a constant stream of unbelievable outrages perpetrated by the other side, I don’t see how we can ever trust each other and work together again.” Discussing the role of social media in the 2016 election, President Barack Obama said “the capacity to disseminate misinformation, wild conspiracy theories, to paint the opposition in wildly negative light without any rebuttal—that has accelerated in ways that much more sharply polarize the electorate and make it very difficult to have a common conversation” (Remnick 2016).

In this paper, we use survey data to study how trends in political polarization relate to respondents’ propensities to obtain news or information online or from social media. Using data from the American National Election Study, we compute nine measures of political polarization that have been proposed in past work and that have increased in recent years. Examples include the affect polarization measure of Iyengar et al. (2012) and the straight-ticket voting measure used in Hetherington (2001). We do not take a stand on whether polarization is increasing (Fiorina et al. 2010; Stein 2016) or on how polarization should be conceptualized (Fiorina and Abrams 2014).

---

<sup>1</sup>Both the appropriate definition of polarization and the extent of its increase are debated in the literature. See Abramowitz and Saunders (2008), Fiorina and Abrams (2008), Prior (2013), Lelkes (2016), and Gentzkow (2016) for reviews.

Instead, we start with the measures others have put forward as evidence of rising polarization and ask whether demographic differences in these measures are consistent with an important role for the internet and social media.

We divide respondents according to demographics that predict internet and social media use. The main predictor we focus on is age. Internet and social media usage rates are far higher among the young than the old—with less than 20 percent of those aged 75 years and older using social media in 2012, as compared to 80 percent of those aged 18–39. We also construct an index of predicted internet use based on a broader set of demographics.

We find that the increase in polarization is largest among the groups least likely to use the internet and social media. A normalized index of our nine polarization measures increases by 0.18 index points overall between 1996 and 2012. Among respondents aged 75 and older, the increase is 0.38 index points, whereas for adults under age 40, the increase is 0.05 index points. Across intermediate age groups, the growth in polarization is consistently higher among older respondents. Polarization increases more for the old than the young in eight of the nine individual measures. A similar pattern emerges for groups of respondents divided by our broader index of predicted internet use.

These findings argue against the hypothesis that the internet in general or social media in particular are the main drivers of increasing polarization. Any such explanation needs to account for the rapid increase in polarization among those with limited internet use and negligible use of social media. However, it is possible to construct such accounts. It may be that social media increases polarization among the young while some other factor increases it among the old. It may be that there are spillovers across demographic groups; young adults polarized through social media might in turn affect the views of older adults or might indirectly influence older adults through channels like the selection of politicians or the endogenous positioning of traditional media. Our evidence, nevertheless, rules out what seem like the most straightforward accounts linking the growth in polarization to the internet.

Much of the empirical evidence on the role of the internet and social media in polarization focuses on segregation of users across information sources or social networks (e.g., Adamic and Glance 2005; Gentzkow and Shapiro 2011; An et al. 2014; Bakshy et al. 2015; Halberstam and Knight 2016). Some work looks directly at online activity and political attitudes. Barbera (2015) finds that, over the course of a year, the average Twitter user who follows at least one political account becomes slightly more moderate with “exposure to political diversity on social

media [having] a positive effect on political moderation” in Germany, Spain, and the US. Liang and Nordin (2012) find that high-speed internet access increases online news consumption but has little to no impact on political attitudes, while Davis and Dunaway (2016) suggest the relationship between internet and polarization is mixed, possibly only impacting those with high levels of engagement with news and politics. Using quasi-random variation in broadband internet access due to state-level, right-of-way legislation, Lelkes et al. (2015) argue “that the new media environment exacerbates already rising tensions.” We contribute to the literature by documenting new evidence on how broad trends in polarization differ across groups with high and low exposure to sources of online information.

The remainder of the paper is organized as follows. Section 2 describes the data and defines the polarization measures, section 3 discusses the trends observed across the polarization measures, and section 4 concludes.

## **2 Data and measures of polarization**

Our primary sources of data are the American National Election Studies’ (ANES) 1948–2012 Time Series Cumulative, 2008 Time Series Study, and 2012 Time Series Study data sets.<sup>2</sup> The ANES is a nationally representative, face-to-face survey of the voting-age population that is conducted in both pre- and post-election rounds and contains numerous demographic variables and political measures. The 2012 ANES survey includes a separate sample of respondents who completed the survey online; we drop these respondents to maintain consistency across years.

### **2.1 Trends in internet and social media usage**

Panel A of figure 1 shows trends in internet access by age group between 1996 and 2012.<sup>3</sup> The internet question was first asked in 1996, when less than 40 percent of 18–39-year-olds had internet. This figure shows that the elderly (75+) have substantially lower levels of internet access/usage across all years, but with similar changes in usage up through 2008. After 2008, the older age

---

<sup>2</sup>See ANES (2015a; 2015b; 2016) respectively. All variables beginning with “VCF” are taken from the 1948–2012 Time Series Cumulative Data File (ANES 2015a). All ANES survey question information is taken from the survey codebooks.

<sup>3</sup>Our measure of internet access for 2012 is responses to “Do you or anyone in this household use the Internet at any location?” (prmedia\_useinet in the 2012 Time Series Study), while our measure of internet access for previous years (1996–2008) is “Do you have access to the Internet or the World Wide Web [exc. 2008: (‘the Web’)]?” (VCF0744 in the 1948–2012 Time Series Cumulative file).

groups experience larger gains in internet usage. Panel B of figure 1 shows that the contrast is even starker when looking at whether respondents had seen campaign information online—with 60 percent of 18–39-year-olds reporting to have seen information about the 2012 presidential campaign online and only 26 percent of 75+-year-olds reporting the same.<sup>4</sup>

Because the ANES does not ask questions regarding social media usage, we supplement it with microdata from the Pew Research Center to plot trends in social media use from 2005 to 2012 in panel C of figure 1.<sup>5</sup> The elderly have substantially lower levels of social media usage across all years.

## 2.2 Measures of political polarization

We collect nine measures of polarization that have been used in prior work and have increased in recent years. In each case, we try to reconstruct the measures exactly as proposed in past work and only intentionally deviate where explicitly stated.

Let  $S_t$  denote the set of all face-to-face survey respondents in year  $t$  who have a valid response for their age.<sup>6</sup> Let  $P_i$  denote a 7-point measure of individual  $i$ 's party from strong Democrat (-3) to Independent (0) to strong Republican (3), and  $B_i$  denote a 7-point measure of individual  $i$ 's ideological affiliation from strong liberal (-3) to moderate (0) to strong conservative (3).<sup>7</sup> Throughout, we

---

<sup>4</sup>Our measure for observing campaign information online comes from VCF0745 (“Have you seen any information about this election campaign on (the Internet/the Web)?”) in the 1948–2012 Time Series Cumulative file for 1996–2004, V085010 and V085021 (“Did you read, watch, or listen to any information about the campaign for President on the Internet?”) in the 2008 Time Series Study for 2008, and mediapo\_net (see 2008 wording) in the 2012 Time Series Study for 2012.

<sup>5</sup>We use four surveys: “September 2005 — Online Dating”, “May 2008 — Cloud computing, politics, and adult social networking”, “July 25–Aug. 26, 2011 — Apps and Social Media Climate”, and “July 16–Aug. 7, 2012 — Civic Engagement” (Pew Research Center 2005; 2008; 2011; 2012). These surveys ask respondents who claim to “use the internet” or “send or receive email” (and, in 2012, those who “access the internet on a cell phone, tablet or other mobile handheld device”) at least occasionally whether they (2012 wording) “Use a social networking site like Facebook, LinkedIn or Google Plus.” A separate question regarding Twitter use was asked in 2011 and 2012, but we exclude it from our analysis. For all surveys, we treat refusals or “Don’t know” responses to the social network or all of its prerequisite questions as missing observations, and we use the included survey weights.

<sup>6</sup>See VCF0101 with 2012 wording: “What is the month, day and year of your birth?” This value is then converted into an age and top-coded by ANES. The top-coding varies across years, but is at least 90 and thus does not impact our age group assignment. A few respondents are 17 years old at the time of the survey, but will be 18 at the time of the election, and are treated as 18-year-olds.

<sup>7</sup>Party and ideology affiliation come from questions VCF0301 (“Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what? [If R/D] Would you call yourself a strong [R/D] or a not very strong [R/D]? [If Independent, Other, or No Preference] Do you think of yourself as closer to the Republican or Democratic party?”) and VCF0803 (“We hear a lot of talk these days about liberals and conservatives. Here is a 7-point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. Where would you place yourself on this scale, or haven’t you thought much about this?”) respectively, and are shifted downwards by four in order to range from -3 to 3. For ideology, people who respond “Don’t know”

let  $R_t := \{i : P_i > 1\}$  and  $R_t^L := \{i : P_i > 0\}$  denote the sets of Republicans excluding and including leaners respectively in time  $t$ . We define  $D_t$  and  $D_t^L$  analogously. Likewise, we let  $C_t = \{i : B_i > 0\}$  and  $L_t = \{i : B_i < 0\}$  denote the set of conservatives and liberals, respectively.

In computing each polarization measure, we restrict the sample to respondents with valid, non-missing responses (as defined by the ANES codebooks) to each of the relevant questions used in constructing the measures. We do not explicitly adjust our set notation to account for this restriction. Throughout, we let  $w_i$  denote the ANES survey weights for respondent  $i$ .<sup>8</sup>

We follow Iyengar et al. (2012) and Gentzkow (2016)<sup>9</sup> in using the ANES thermometer ratings of parties and ideologies to capture how peoples feelings towards those on the other side of the political spectrum have changed over time.<sup>10</sup> These two measures capture the difference in the “warmth” or “coldness” respondents feel toward those of their own and the opposite party or ideology.

**Definition.** *Partisan affect polarization*  $M_t^{AP}$  is

$$M_t^{AP} = \frac{1}{\sum_{i \in D_t^L} w_i} \sum_{i \in D_t^L} w_i (A_i^D - A_i^R) + \frac{1}{\sum_{i \in R_t^L} w_i} \sum_{i \in R_t^L} w_i (A_i^R - A_i^D)$$

where  $A_i^R$  and  $A_i^D$  denote individual  $i$ 's feelings towards Republicans and Democrats, respectively, on a 0–100 scale, with higher values representing more favorable feelings.

**Definition.** *Ideological affect polarization*  $M_t^{AI}$  is

$$M_t^{AI} = \frac{1}{\sum_{i \in L_t} w_i} \sum_{i \in L_t} w_i (A_i^L - A_i^C) + \frac{1}{\sum_{i \in C_t} w_i} \sum_{i \in C_t} w_i (A_i^C - A_i^L)$$

---

or “Haven’t though much about it” are treated as having a missing or non-valid response.

<sup>8</sup>We use the type-0, face-to-face survey weights from variable VCF0009x.

<sup>9</sup>See “Affective Distance” from tables A1 and A2 in the online appendix of Iyengar et al. (2012) and figure 9 in Gentzkow (2016).

<sup>10</sup>We use ANES thermometer questions that ask respondents to rate various groups from 0–100, where higher values indicate more favorable feelings and a value of 50 designates neutral feelings. The exact wording of the questions has changed over time. We use responses for the following groups: liberals (VCF0211), conservatives (VCF0212), Democrat party (VCF0218), and Republican party (VCF0224). Responses between 97 and 100 are top-coded at 97. The 2012 preface for these questions is as follows: “Please look at page [preload: prepg\_c] of the booklet. I’d like to get your feelings toward some of our political leaders and other people who are in the news these days. I’ll read the name of a person and I’d like you to rate that person using something we call the feeling thermometer. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don’t feel favorable toward the person and that you don’t care too much for that person. You would rate the person at the 50 degree mark if you don’t feel particularly warm or cold toward the person. If we come to a person whose name you don’t recognize, you don’t need to rate that person. Just tell me and we’ll move on to the next one.”

where  $A_i^C$  and  $A_i^L$  denote an individual  $i$ 's feelings towards conservatives and liberals, respectively, on a 0–100 scale, with higher values representing more favorable feelings.

We use the measure developed by Mason (2015), and also used by Davis and Dunaway (2016), to capture the degree of partisan sorting.<sup>11</sup> Partisan sorting measures the extent to which partisan identity is correlated with self-reported ideology.

**Definition.** *Partisan sorting*  $M_t^{Sorting}$  is

$$M_t^{Sorting} = \frac{1}{\sum_{i \in S_t} w_i} \sum_{i \in S_t} \frac{w_i}{105} [g(|P_i - I_i| + 1)(|P_i| + 1)(|B_i| + 1) - 7],$$

where  $g(x) = \max_{i \in U_t, S_t} (|P_i - I_i| + 1) + \min_{i \in U_t, S_t} (|P_i - I_i| + 1) - x$ .

We follow Hetherington (2001) in examining the frequency with which individuals split their votes across parties in an election.<sup>12</sup>

**Definition.** *Straight-ticket*  $M_t^{ST}$  is the survey-weighted proportion of respondents who report voting for the same party in both the presidential and House elections.<sup>13</sup>

We use the measure proposed by Abramowitz and Saunders (2008)<sup>14</sup> to examine the similarity of ideological positions across issues by letting  $k_i$  denote a conservative (1), moderate (0), or liberal (-1) response to  $k \in K$ , where  $K$  is a set of seven different questions.<sup>15</sup>

<sup>11</sup>See figure 1 in Mason (2015). Details on the implementation are taken from Mason (2013).

<sup>12</sup>See figure 3 of Hetherington (2001).

<sup>13</sup>Data come from VCF0709, which aggregates responses from previous questions regarding whom the respondent voted for in the recent election. We could also define this measure using straight-ticket voting across Senate and presidential elections (VCF0710), but Senate elections are less frequent.

<sup>14</sup>See Abramowitz and Saunders (2008, 544) where they outline their 7-point measure of ideological polarization. While Abramowitz and Saunders (2008) further collapse this 7-point scale to label individuals as low, moderate, or high and report the percentage of respondents with ideological consistent views in their table 1, we simply report the average of the 7-point measure. Footnote 2 in Abramowitz and Saunders (2008) suggests that “using a simple additive index consisting of the same items produces nearly identical results concerning the trend in polarization between 1984 and 2004.”

<sup>15</sup>The seven questions used are as follows: 7-point ideological position (VCF0803), aid to blacks (VCF0830's 2012 wording: “Some people feel that the government in Washington should make every effort to improve the social and economic position of blacks. Suppose these people are at one end of a scale, at point 1. Others feel that the government should not make any special effort to help blacks because they should help themselves. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, or 6. Where would you place yourself on this scale, or haven't you thought much about this?”), foreign defense spending (VCF0843's 2012 wording: “Some people believe that we should spend much less money for defense.... Others feel that defense spending should be greatly increased....”), government responsibility to guarantee jobs and income (VCF0809's 2012 wording: “Some people feel the government in Washington should see to it that every person has a job and a good standard of living.... Others think the government should just let each person get ahead on their own....”), government health insurance plan (VCF0806's 2012 wording: “There is much

**Definition.** *Issue consistency*  $M_t^{Consistency}$  is

$$M_t^{Consistency} = \frac{1}{\sum_{i \in S_t} w_i} \sum_{i \in S_t} w_i \left| \sum_{k \in K} k_i \right|.$$

In a similar spirit to Abramowitz and Saunders (2008),<sup>16</sup> we define issue divergence as the average correlation between these same seven questions and party affiliation for respondents who are not strictly independents.

**Definition.** *Issue divergence*  $M_t^{Divergence}$  is

$$M_t^{Divergence} = \frac{1}{|K|} \sum_{k \in K} cor_{i \in R_t^L \cup D_t^L}(\mathbf{1}_{P_i > 0}, k),$$

where *cor* calculates Kendall’s tau between the party indicator and the question coding.<sup>17</sup>

Partisan-ideology polarization examines the extent which the self-reported ideological affiliation of Republicans and Democrats differ. For this, we follow a measure outlined in Abramowitz and Saunders (2008, 547).

---

concern about the rapid rise in medical and hospital costs. Some people feel there should be a government insurance plan which would cover all medical and hospital expenses for everyone.... Others feel that all medical expenses should be paid by individuals through private insurance plans like Blue Cross or other company paid plans....”), government services and spending (VCF0839’s 2012 wording: “Some people think the government should provide fewer services even in areas such as health and education in order to reduce spending... Other people feel it is important for the government to provide many more services even if it means an increase in spending...”), and abortion legislation (VCF0838’s 2012 wording: “There has been some discussion about abortion during recent years. Which one of the opinions on this page best agrees with your view? You can just tell me the number of the opinion you choose. 1. By law, abortion should never be permitted. 2. The law should permit abortion only in case of rape, incest, or when the woman’s life is in danger. 3. The law should permit abortion for reasons other than rape, incest, or danger to the woman’s life, but only after the need for the abortion has been clearly established. 4. By law, a woman should always be able to obtain an abortion as a matter of personal choice.”). All questions are answered on a 7-point, liberal-to-conservative scale (except government services and spending whose scale is conservative-to-liberal of the same order and abortion legislation which has four available positions). Liberal positions (1–3) are coded as -1, moderate positions (4) and refusals (“Don’t know” or “Haven’t thought much about it”) are coded as 0, and conservative positions (5–7) are coded as 1. For the abortion questions, we code conservative positions (1 and 2) as 1, moderate (3) and refusals (“Don’t know” or “Haven’t thought much about it”) as 0, and liberal positions (4) as -1.

<sup>16</sup>See table 3 in Abramowitz and Saunders (2008), which uses a similar, but different, set of questions. For simplicity and consistency, we use the same set of questions for our issue consistency and issue divergence polarization measures.

<sup>17</sup>Each correlation is computed for the set of respondents with valid, non-missing responses to both the party identification question  $P_i$  and the given issue question  $k \in K$ .



**Definition.** *Partisan-ideology polarization*  $M_t^{PI}$  is

$$M_t^{PI} = \frac{1}{\sum_{i \in R_t} w_i} \sum_{i \in R_t} B_i w_i - \frac{1}{\sum_{i \in D_t} w_i} \sum_{i \in D_t} B_i w_i.$$

Following the description provided in Davis and Dunaway (2016, 283), we also measure the extent to which individuals *perceive* there to be ideological differences between the Republican and Democrat parties.<sup>18</sup>

**Definition.** *Perceived partisan-ideology polarization*  $M_t^{Perceived}$  is

$$M_t^{Perceived} = \frac{1}{\sum_{i \in S_t} w_i} \sum_{i \in S_t} w_i (\tilde{R}_i - \tilde{D}_i)$$

where  $\tilde{R}_i$  and  $\tilde{D}_i$  denote individual  $i$ 's perception on how conservative the Republican or Democratic party is, respectively, on the 7-point liberal-to-conservative scale.

Finally, we construct a measure of religious polarization similar to Abramowitz and Saunders (2008).<sup>19</sup>

**Definition.** *Religious polarization*  $M_t^{Relig}$  is

$$M_t^{Relig} = \frac{1}{\sum_{i \in CG_t} w_i} \sum_{i \in CG_t} P_i w_i - \frac{1}{\sum_{i \in NC_t} w_i} \sum_{i \in NC_t} P_i w_i,$$

where  $CG_t$  is the set of white respondents who go to church no less than almost every week and  $NC_t$  is the set of white respondents who go to church no more than a few times a year.<sup>20</sup>

Our overall index of polarization is then defined to be

<sup>18</sup>We use questions VCF0503 and VCF0504, and we shift responses down by four. The wording for 2012 is as follows: “We hear a lot of talk these days about liberals and conservatives. Here is a seven-point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative... Where would you place the Republican party on this scale?”

<sup>19</sup>See figure 5 of Abramowitz and Saunders (2008), which uses “observant” and “non-observant,” white respondents.

<sup>20</sup>White individuals are those placed in the “White non-Hispanic” group of ANES summary question VCF0105b, which collapses various responses of an individual's race across questions to four categories: white non-Hispanic, black non-Hispanic, Hispanic, and other or multiple races non-Hispanic. Responses regarding church attendance are taken from question VCF0130 (2012 wording: “Lots of things come up that keep people from attending religious services even if they want to. Thinking about your life these days, do you ever attend religious services, apart from occasional weddings, baptisms or funerals?” and if “Yes”, respondent is asked “Do you go to religious services every week, almost every week, once or twice a month, a few times a year, or never?”). Individuals with no religious preference (relevant for responses between 1970 and 1988 inclusive) are treated as if they never go to church. Individuals that attend church “once or twice a month” are excluded.

$$Index_t = \frac{1}{|M|} \sum_{m \in M} m_t / m_{1996},$$

where  $M$  is the set of all nine polarization measures. We also compute this index for different groups of respondents. In these cases, we normalize the group level values based on the overall 1996 value  $m_{1996}$ .

Figure 2 plots each of these polarization measures, including the index, over time. By design, all of the measures we include show an overall growth in polarization, with the index growing by 0.19 index points between 1996 and 2012. It is interesting to note that the index grew almost as quickly in the decades before the advent of the internet as after it, a pattern also exhibited by many of the individual measures.

### 3 Trends in polarization by demographic group

#### 3.1 By age

The top panel of figure 3 shows trends in our polarization index by age group. Table 1 provides additional quantitative detail and standard errors for inference.<sup>21</sup> Polarization for those aged 75+ grows by 0.38 index points between 1996 and 2012, and polarization for those aged 65+ grows by 0.32 index points over the same period. Polarization among those aged 18–39 increased by 0.05 index points between 1996 and 2012. Bootstrap standard errors show that we can separately reject, at the five percent level, the hypotheses that the increase for those aged 18–39 is equal to the increase for those aged 65+ and that the increase for those aged 18–39 is equal to the increase for those aged 75+.

Table 1 and appendix figure 1 report corresponding values for each of the polarization measures. For every measure, except religious polarization, we see that the oldest age group experiences larger changes in polarization than the youngest age group. In four of the nine measures, young adults actually experience declines in polarization while the other age groups experience large increases. Furthermore, in four of the nine measures, we see a monotonically increasing relationship between age group and the growth in polarization.

In appendix figure 2, we show that these results are robust to using cohorts instead of age groups, restricting to males or females, restricting the sample to those who self-identify with a

<sup>21</sup>We conduct inference using a nonparametric bootstrap with 100 replicates. We draw replicates from the full sample  $\cup_t S_t$  and construct measures of polarization for each replicate following the procedure in section 2.

party, and restricting the sample to those who self-identify as being “very much interested” in the upcoming election.<sup>22</sup> In appendix figure 3, we show that the trends between the 18–39 and 65+ age groups track fairly closely across the entire 1972–2012 time period.

### 3.2 By predicted and actual internet access

The second panel of figure 3 shows trends in polarization according to a broad index of predicted internet use. We suppose that

$$\Pr(\text{internet}_i = 1 | X_i) = X_i' \beta, \quad (1)$$

where  $\beta$  is a vector of parameters and  $X_i$  is a vector of demographic characteristics including indicators for age group, gender, race, education, and whether an individual lives in the political South. We estimate (1) on the sample  $S_{1996}$  using weighted least squares with weights  $w_i$ .<sup>23</sup> We then compute predicted internet use  $\widehat{\text{internet}}_i$  for all respondents. The panel plots the polarization index for respondents in the first and fourth quantiles of  $\widehat{\text{internet}}_i$  for each respective survey year.<sup>24</sup> The group with the greater likelihood of having access to the internet experienced slower growth in polarization between 1996 and 2012.

The third panel of figure 3 shows trends in polarization by actual internet access.<sup>25</sup> The set of respondents with internet access have greater polarization in 2012 than those without internet access. However, the trends are parallel between the two groups between 1996 and 2012.

## 4 Conclusion

Our findings are difficult to square with a straightforward account linking the recent rise of polarization to the internet. This is especially true for accounts in which social media plays a central role. Unless cross-group spillovers are very large or the effects of digital media vary greatly across groups, some other forces must explain the large increase in polarization among the groups least likely to use the internet. None of this is to say that the rise of digital technologies is not important.

---

<sup>22</sup>In fact, among those who self-identify as interested in the election, we see even stronger disparities across age groups, as the 18–39 age group starts out with a value of around 1.5 in 1996 and *drops* to less than 1.2 in 2012, while the 65+ age group grows from a value around 1 to a value of over 1.2.

<sup>23</sup>Coefficient estimates and variable definitions are reported in appendix table 1.

<sup>24</sup>Appendix figure 1 shows analogous plots for the individual polarization measures.

<sup>25</sup>Appendix figure 1 shows analogous plots for the individual polarization measures.

They may well account for some recent polarization, and whatever role they may have played in the past is likely to grow in coming years.

## References

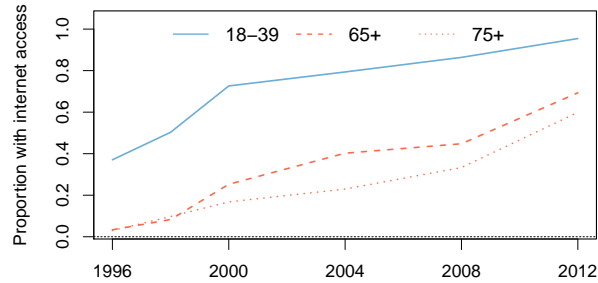
- Abramowitz, Alan I. and Kyle L. Saunders. 2008. Is polarization a myth? *The Journal of Politics*. 70(2): 542–555.
- Adamic, Lada A. and Glance, Natalie. 2005. The political blogosphere and the 2004 U.S. election: Divided they blog. *Proceedings of the 3rd International Workshop on Link Discovery*. 36–43.
- An, Jisun, Daniel Quercia, and Jon Crowcroft. 2014. Partisan sharing: Facebook evidence and societal consequences. *Proceedings of the Second ACM Conference on Online Social Networks*. 13–24.
- American National Election Study. 2016. The ANES 2012 time series study. Stanford University and the University of Michigan. Accessed at [http://www.electionstudies.org/studypages/download/datacenter\\_all\\_datasets.php](http://www.electionstudies.org/studypages/download/datacenter_all_datasets.php) on December 22, 2016.
- . 2015a. The ANES 1948–2012 time series cumulative data file. Stanford University and the University of Michigan. Accessed at [http://www.electionstudies.org/studypages/download/datacenter\\_all\\_datasets.php](http://www.electionstudies.org/studypages/download/datacenter_all_datasets.php) on December 22, 2016.
- . 2015b. The ANES 2008 time series study. Stanford University and the University of Michigan. Accessed at [http://www.electionstudies.org/studypages/download/datacenter\\_all\\_datasets.php](http://www.electionstudies.org/studypages/download/datacenter_all_datasets.php) on February 18, 2017.
- . 2015c. The ANES guide to public opinion and electoral behavior: Split ticket voting presidential/congressional 1952-2012. Accessed at [http://www.electionstudies.org/nesguide/toptable/tab9b\\_2.htm](http://www.electionstudies.org/nesguide/toptable/tab9b_2.htm) on February 22, 2017.
- Bakshy, Eytan, Itamar Rosenn, Cameron Marlow, and Lada Adamic. 2012. The role of social networks in information diffusion. *Proceedings of the 21st International Conference on World Wide Web*. 519–528.
- Barbera, Pablo. 2015. How social media reduces mass political polarization. Evidence from Germany, Spain, and the U.S. *Working paper*. Accessed at [http://pablobarbera.com/static/barbera\\_polarization\\_APSA.pdf](http://pablobarbera.com/static/barbera_polarization_APSA.pdf) on February 19, 2017.
- Davis, Nicholas T. and Johanna L. Dunaway. 2016. Party polarization, media choice, and mass partisan-ideological sorting. *Public Opinion Quarterly*. 80(S1): 272–297.
- Fiorina, Morris P. and Samuel J. Abrams. 2008. Political polarization in the American public. *Annual Review of Political Science*. 11: 563–588.
- Fiorina, Morris P., Samuel J. Abrams, and Jeremy C. Pope. 2010. *Culture war? The myth of polarized America*. London, England: Longman Publishing.

- Fiorina, Morris P. and Samuel Abrams. 2014. Americans aren't polarized, just better sorted. *The Washington Post*. Accessed at <https://www.washingtonpost.com/news/monkey-cage/wp/2014/01/21/americans-arent-polarized-just-better-sorted/> on February 19, 2017.
- Gabler, Neil. 2016. The internet and social media are increasingly divisive and undermining of democracy. *Alternet.org*. Accessed at <http://www.alternet.org/culture/digital-divide-american-politics> on February 19, 2017.
- Gentzkow, Matthew. 2016. Polarization in 2016. *Toulouse Network for Information Technology Whitepaper*. Accessed at <http://web.stanford.edu/~gentzkow/research/PolarizationIn2016.pdf> on February 19, 2017.
- Gentzkow, Matthew and Jesse M. Shapiro. 2011. Ideological segregation online and offline. *Quarterly Journal of Economics*. 126(4): 1799–1839.
- Haidt, Jonathan. 2016. Why social media is terrible for multiethnic democracies. *Vox*. Accessed at <http://www.vox.com/policy-and-politics/2016/11/15/13593670/donald-trump-jonathan-haidt-social-media-polarization-europe-multiculturalism> on February 19, 2017.
- Halberstam, Yosh and Brian Knight. 2016. Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter. *Journal of Public Economics*. 143: 73–88.
- Hetherington, Marc J. 2001. Resurgent mass partisanship: The role of elite polarization. *The American Political Science Review*. 95(3): 619–631.
- Iyengar, Shanto, Gaurav Sood, and Ypthach Lelkes. 2012. Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*. 76(3): 405–431.
- Liang, Che-Yuan and Mattias Nordin. 2012. The internet, news consumption, and political attitudes. *Uppsala Center for Fiscal Studies Working Paper Series*. Accessed at [http://econpapers.repec.org/paper/hhsuufswp/2012\\_5f010.htm](http://econpapers.repec.org/paper/hhsuufswp/2012_5f010.htm) on February 19, 2017.
- Lelkes, Ypthach, Gaurav Sood, and Shanto Iyengar. 2015. The hostile audience: The effect of access to broadband internet on partisan affect. *American Journal of Political Science*. 61(1): 5–20.
- Lelkes, Ypthach. 2016. Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*. 80(S1): 392–410.
- Mason, Lilliana. 2013. Replication data for: “I disrespectfully agree”: The differential effects of partisan sorting on behavioral and issue polarization.” Harvard Dataverse, V3. Accessed at <https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/22346> on January 20, 2017.
- Mason, Lilliana. 2015. “I disrespectfully agree”: The differential effects of partisan sorting on

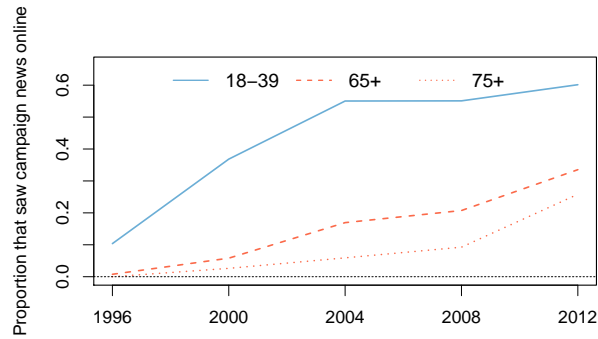
- social and issue polarization. *American Journal of Political Science*. 59(10): 128–145.
- Pariser, Eli. 2011. *The filter bubble: What the internet is hiding from you*. New York, NY: Penguin Press.
- Pew Research Center. 2005. September 2005 — Online dating survey. Accessed at <http://www.pewinternet.org/datasets/september-2005-online-dating/> on February 16, 2017.
- . 2008. May 2008 — Cloud computing, politics, and adult social networking survey. Accessed at <http://www.pewinternet.org/datasets/may-2008-cloud-computing-politics-and-adult-social-networking/> on February 16, 2017.
- . 2011. July 25–Aug. 26, 2011 — Apps and social media climate survey. Accessed at <http://www.pewinternet.org/datasets/august-2011-apps-and-adult-sns-climate/> on February 16, 2017.
- . 2012. July 16–Aug. 7, 2012 — Civic engagement survey. Accessed at <http://www.pewinternet.org/datasets/august-2012-civic-engagement/> on February 16, 2017.
- Prior, Markus. 2013. Media and political polarization. *Annual Review of Political Science*. 16: 101–127.
- Remnick, David. 2016. Obama reckons with a Trump presidency. *The New Yorker*. Accessed at <http://www.newyorker.com/magazine/2016/11/28/obama-reckons-with-a-trump-presidency> on February 20, 2017.
- Stein, Jeff. 2016. Most experts think America is more polarized than ever. This Stanford professor disagrees. *Vox*. Accessed at <http://www.vox.com/policy-and-politics/2016/11/1/13351284/america-polarization> on February 19, 2017.
- Sunstein, Cass R. 2001. *Republic.com*. Princeton, NJ: Princeton University Press.
- . 2009. *Republic.com 2.0*. Princeton, NJ: Princeton University Press.
- . 2017. *#Republic: Divided democracy in the age of social media*. Princeton, NJ: Princeton University Press.

Figure 1: Trends in internet access and social media use by age group

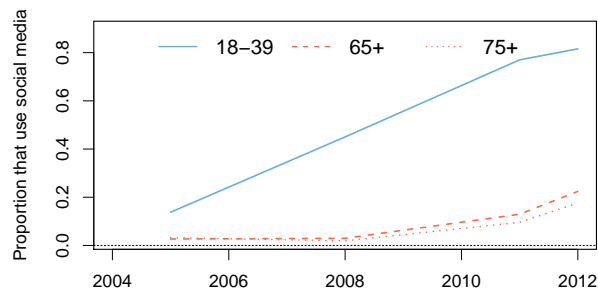
*Panel A: Internet access*



*Panel B: Campaign information online*



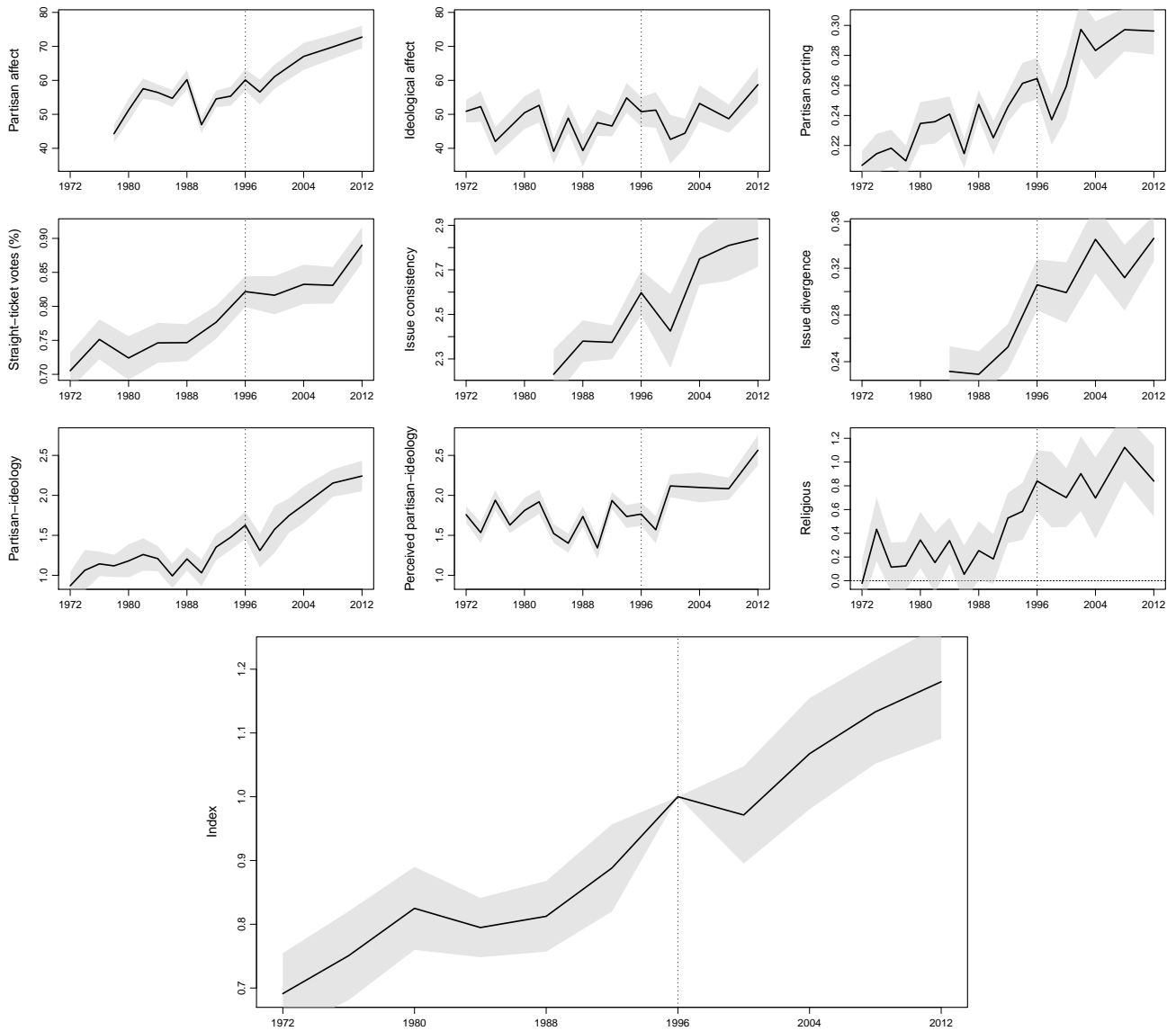
*Panel C: Social media use*



Notes: Panel A shows the weighted proportion of respondents who have internet access by age group. Panel B shows the proportion of respondents that saw campaign information online by age group. Panel C shows the estimated proportion of the adult American population that uses social media by age group according to the Pew Research Center (2005; 2008; 2011; 2012). See section 2.1 for details on each variable.

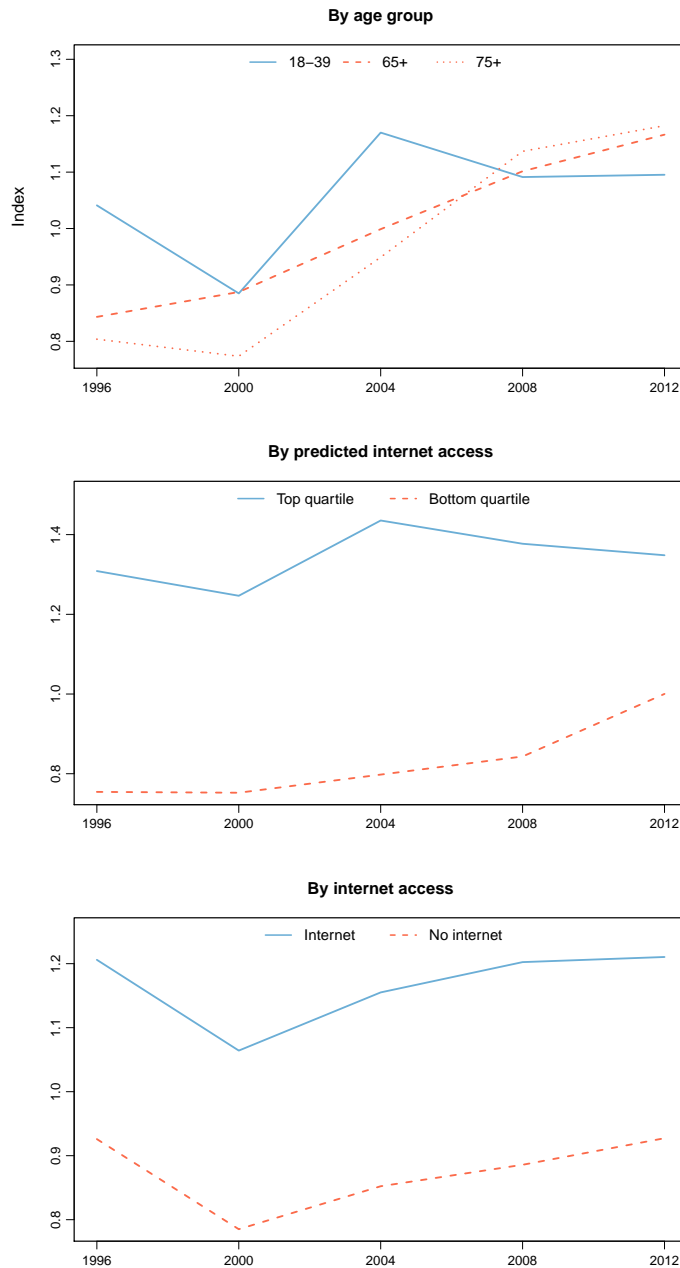


Figure 2: Trends in political polarization



Notes: Each of the nine small plots shows the trend in a given polarization measure across time. The large plot shows the trend in the index, which is computed as the average across all polarization measures available in a given year after normalizing each measure to have a value of one in 1996. The shaded regions are 95 percent confidence intervals constructed using a nonparametric bootstrap with 100 replicates. See section 2 for definitions.

Figure 3: Trends in polarization by demographic group



Notes: Each plot shows the polarization index broken out for different demographic groups. The top plot shows the index by age group. The middle plot shows the index by quartile of predicted internet access. The bottom quartile includes values that are at or below the 25th percentile, while the top quartile includes values greater than the 75th percentile. The bottom plot shows the index by internet access. For a given measure and group, the polarization value is normalized by the 1996 value of the polarization measure in the full sample. The index is then the average of these normalized polarization measures for each group. See section 2 for definitions and section 3 for construction of predicted internet access.

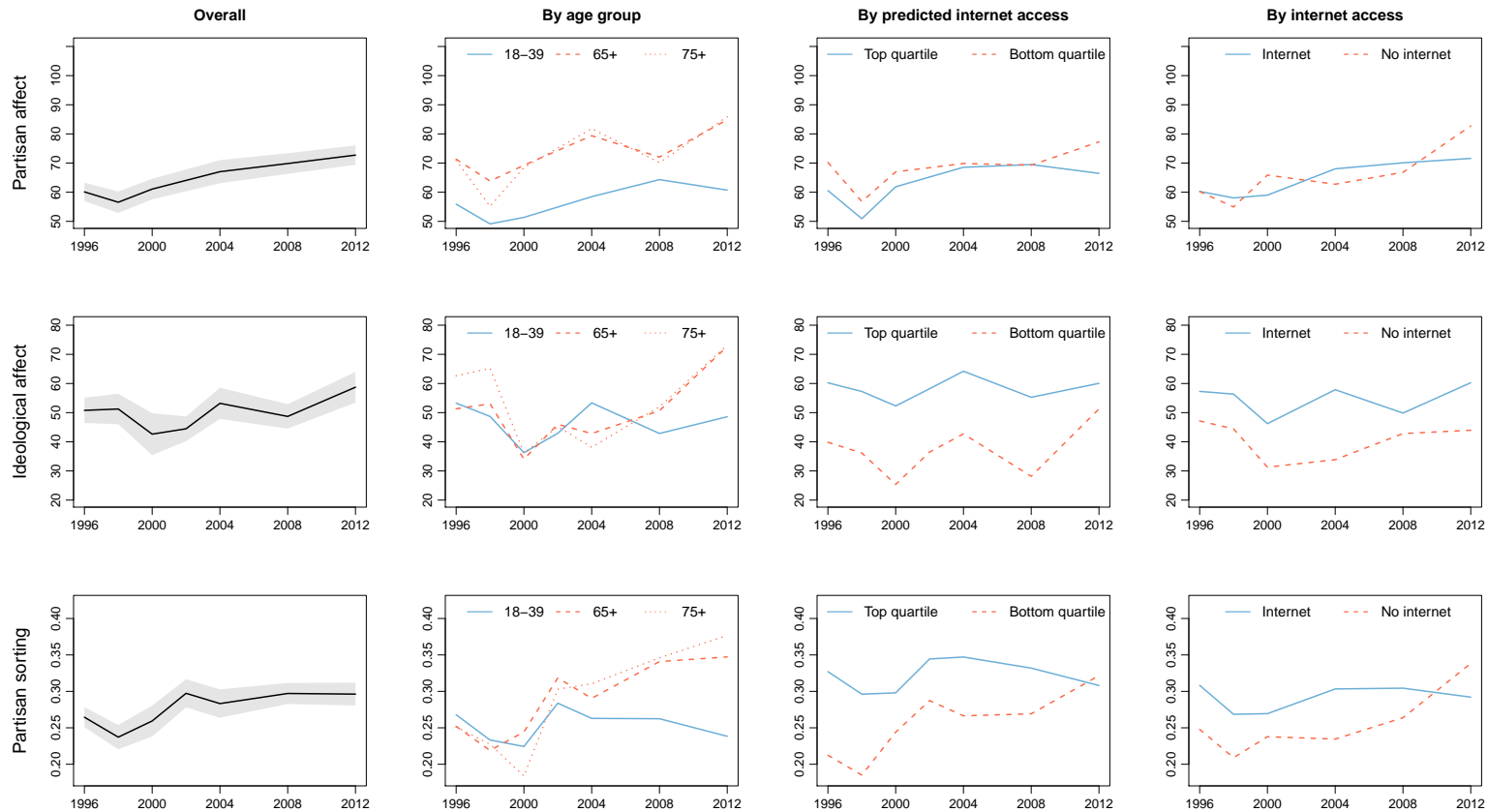
Table 1: Growth in polarization 1996 to 2012

Measure	Overall	Age Groups					
		18–39	40–64	65+	75+	65+ minus 18–39	75+ minus 18–39
Partisan affect	12.6 (2.4)	4.8 (3.7)	16.0 (4.0)	13.5 (5.9)	15.0 (10.0)	8.72 (6.73)	10.16 (10.21)
Ideological affect	7.9 (3.3)	-4.7 (5.6)	8.4 (5.5)	21.4 (9.3)	10.6 (12.9)	26.10 (10.80)	15.25 (13.63)
Partisan sorting	0.03 (0.01)	-0.03 (0.02)	0.05 (0.02)	0.10 (0.03)	0.13 (0.05)	0.12 (0.03)	0.16 (0.05)
Straight-ticket	0.07 (0.02)	0.01 (0.04)	0.11 (0.03)	0.06 (0.04)	0.06 (0.06)	0.05 (0.06)	0.05 (0.07)
Issue consistency	0.24 (0.08)	-0.08 (0.13)	0.17 (0.12)	1.01 (0.16)	1.44 (0.23)	1.09 (0.21)	1.52 (0.27)
Issue divergence	0.04 (0.02)	-0.02 (0.03)	0.03 (0.02)	0.14 (0.04)	0.16 (0.06)	0.16 (0.05)	0.18 (0.07)
Partisan-ideology	0.62 (0.13)	0.17 (0.22)	0.54 (0.20)	1.34 (0.34)	1.63 (0.56)	1.17 (0.40)	1.46 (0.59)
Perceived partisan-ideology	0.80 (0.12)	0.73 (0.19)	0.74 (0.17)	1.01 (0.26)	0.97 (0.35)	0.27 (0.32)	0.24 (0.41)
Religious	-0.00 (0.20)	0.15 (0.29)	0.09 (0.32)	-0.34 (0.42)	-0.18 (0.68)	-0.49 (0.52)	-0.34 (0.76)
Index	0.18 (0.05)	0.05 (0.06)	0.20 (0.08)	0.32 (0.09)	0.38 (0.16)	0.27 (0.11)	0.32 (0.16)

Notes: Table shows the change in each measure, and in the index, from 1996 to 2012. The “Overall” column includes all ages. Columns “18–39,” “40–64,” “65+,” and “75+” re-calculate the measures for each age group. The last two columns show the difference in growth between the two age groups. Standard errors are in parentheses and are constructed using a nonparametric bootstrap with 100 replicates. See section 2 for definitions.

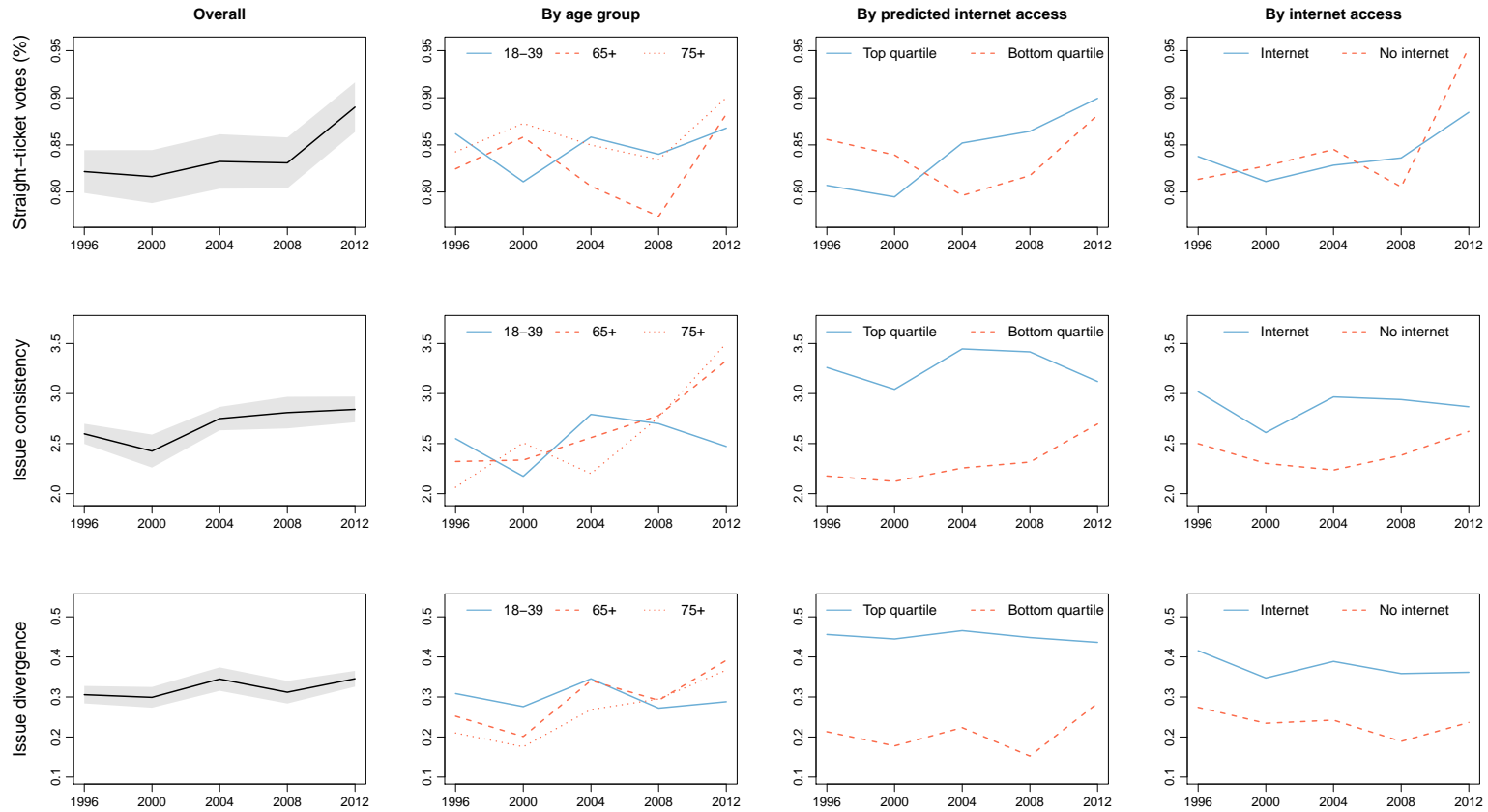
# A Appendix

Appendix figure 1: Trends in polarization by demographic group, individual measures



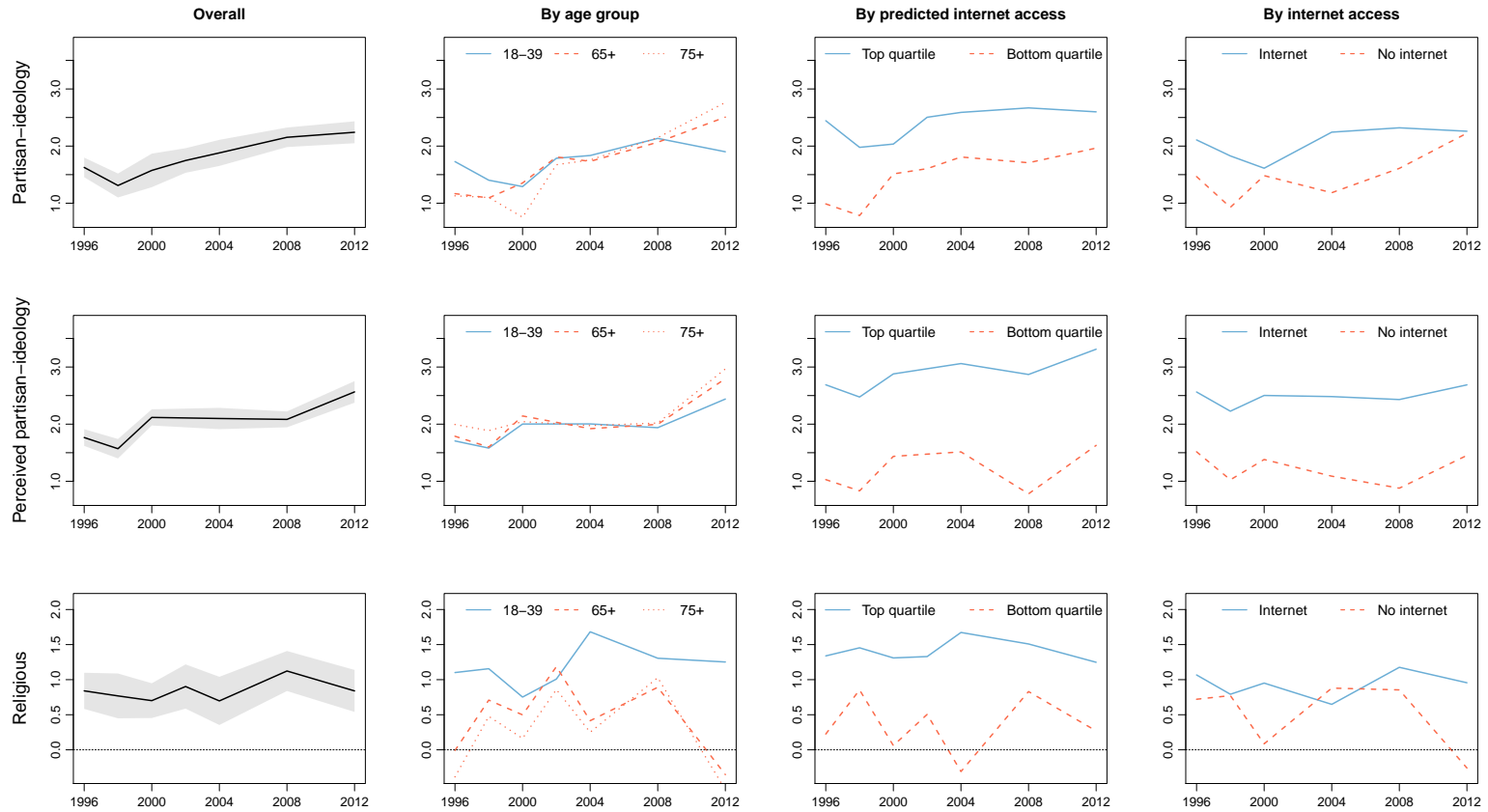
Notes: Each row of plots displays the trends for a given polarization measure for various groups of respondents. The first figure in each row shows the overall measure with a 95 percent confidence interval constructed using a nonparametric bootstrap with 100 replicates. The second figure in each row shows the measure by age group. The third figure in each row shows the measure by predicted internet quantiles where higher values indicate the set of users more likely to have received internet in 1996. The bottom quartile includes values that are at or below the 25th percentile, while the top quartile includes values greater than the 75th percentile. The fourth figure in each row shows the measure by internet access. See section 2 for definitions and section 3 for construction of predicted internet access.

Appendix figure 1: Trends in polarization by demographic group, individual measures (cont.)



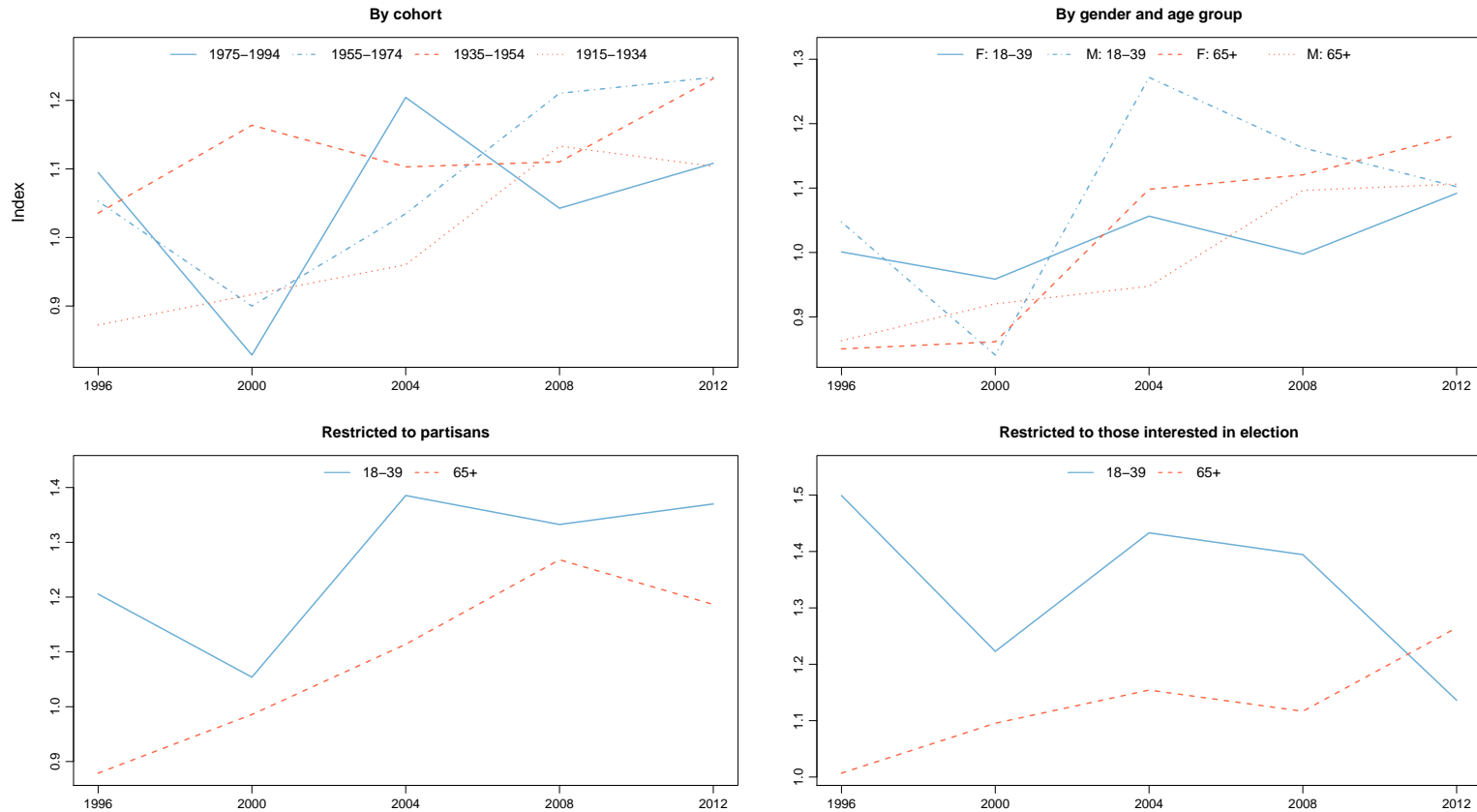
Notes: Each row of plots displays the trends for a given polarization measure for various groups of respondents. The first figure in each row shows the overall measure with a 95 percent confidence interval constructed using a nonparametric bootstrap with 100 replicates. The second figure in each row shows the measure by age group. The third figure in each row shows the measure by predicted internet quantiles where higher values indicate the set of users more likely to have received internet in 1996. The bottom quartile includes values that are at or below the 25th percentile, while the top quartile includes values greater than the 75th percentile. The fourth figure in each row shows the measure by internet access. See section 2 for definitions and section 3 for construction of predicted internet access.

Appendix figure 1: Trends in polarization by demographic group, individual measures (cont.)



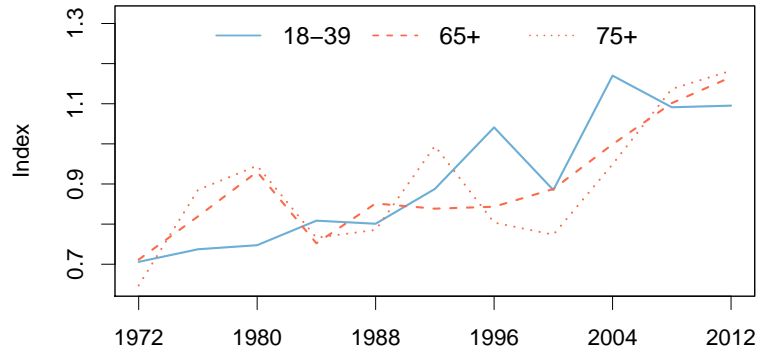
Notes: Each row of plots displays the trends for a given polarization measure for various groups of respondents. The first figure in each row shows the overall measure with a 95 percent confidence interval constructed using a nonparametric bootstrap with 100 replicates. The second figure in each row shows the measure by age group. The third figure in each row shows the measure by predicted internet quantiles where higher values indicate the set of users more likely to have received internet in 1996. The bottom quartile includes values that are at or below the 25th percentile, while the top quartile includes values greater than the 75th percentile. The fourth figure in each row shows the measure by internet access. See section 2 for definitions and section 3 for construction of predicted internet access.

Appendix figure 2: Robustness and heterogeneity



Notes: Each plot displays the polarization index for various alternative specifications. The “by cohort” plot calculates the index for “1975–1994,” “1955–1974,” “1935–1954,” and “1915–1934” cohorts, where birth year is defined as survey year minus age. Note that age is top-coded at a value of 90 or greater (the exact value varies across survey years), which does not impact the assignment of age groups but does impact birth year assignment, and thus, cohort assignments. The “by gender” plot shows the polarization for each age group separately for females (F) and males (M). The “restricted to partisan” plot restricts the sample to those whose party identification ( $P_i$ ) is greater than one in absolute value. Note that some of the nine polarization measures already restrict the sample to partisans. The “restricted to those interested in the election” plot restricts the sample to those who report being “Very much interested” in the political campaigns (VCF0310’s 2012 wording: “Some people don’t pay much attention to political campaigns. How about you, would you say that you have been/were very much interested, somewhat interested, or not much interested in the political campaigns (so far) this year?”). Each polarization measure is divided by the 1996 value of the measure on the full sample and then all measures are averaged to form the index.

Appendix figure 3: By age group, extended series



Notes: Plot shows the polarization index by age group from 1972 to 2012. After normalizing each measure by dividing by the 1996 value of the measure on the full sample, we take the average across all polarization measures available in a given year .



Appendix table 1: Predicted internet in 1996

Estimator: Weighted least squares	
Dependent variable: Has internet	
Intercept	0.428 (0.044)
Age Group: 40-64	-0.124 (0.022)
Age Group: 65+	-0.269 (0.030)
Gender: Male	-0.008 (0.020)
Race: Hispanic	0.046 (0.045)
Race: Other	0.108 (0.069)
Race: White	0.156 (0.034)
Education: Grade School	-0.363 (0.053)
Education: High School	-0.371 (0.026)
Education: Some College	-0.146 (0.029)
Region: South	0.081 (0.022)

Notes: Table shows the coefficients from a weighted least squares regression. Sample is restricted to respondents in 1996 who responded to the questions needed to construct each covariate. Dependent variable is an indicator for whether an individual has internet access. Our measure of internet access for 2012 is responses to “Do you or anyone in this household use the Internet at any location?” (prmedia\_useinet in the 2012 Time Series Study), while our measure of internet access for previous years (1996–2008) is “Do you have access to the Internet or the World Wide Web [exc. 2008: (‘the Web’)]?” (VCF0744 in the 1948–2012 Time Series Cumulative file). All covariates are indicator variables. Age groups are defined using VCF0101, gender is defined using VCF0104, race is defined using VCF0105b, education is defined using VCF0110, and region is defined using VCF0113. Standard errors are in parentheses.