The Wisdom of Partisan Crowds

Joshua Becker[a,b,c,1], **Ethan Porter**[d], **and Damon Centola**[a,f]

[a]Annenberg School for Communication, University of Pennsylvania; [b]Kellogg School of Management, Northwestern University; [c]Northwestern Institute on Complex Systems, Northwestern University; [d]School of Media and Public Affairs, George Washington University; [e]School of Engineering, University of Pennsylvania

This manuscript was compiled on April 7, 2019

Theories in favor of deliberative democracy are based on the premise that social information processing can improve group beliefs. While research on the "wisdom of crowds" has found that information exchange can increase belief accuracy on non-controversial factual matters, theories of political polarization imply that groups will become more extreme—and less accurate—when beliefs are motivated by partisan political bias. A primary concern is that partisan biases are associated not only with more extreme beliefs, but also a diminished response to social information. While bipartisan networks containing both Democrats and Republicans are expected to promote accurate belief formation, politically homogeneous networks are expected to amplify partisan bias and reduce belief accuracy. To test whether the wisdom of crowds is robust to partisan bias, we conducted two web-based experiments in which individuals answered factual questions known to elicit partisan bias before and after observing the estimates of peers in a politically homogeneous social network. In contrast to polarization theories, we found that social information exchange in homogeneous networks not only increased accuracy but also reduced polarization. Our results help generalize collective intelligence research to political domains.

collective intelligence | polarization | networks | social influence

A major concern for democratic theorists is that citizens are simply too ignorant of basic political facts to benefit from deliberation (1), yet research on the "wisdom of crowds" (2–4) has found the aggregated beliefs of large groups can be "wise"—i.e., factually accurate—even when group members are individually inaccurate. While these statistical theories offer optimistic support for democratic principles (5, 6), normative theories of deliberative democracy remain challenged by the argument that social influence processes—in contrast with the aggregation of independent survey responses—amplify group biases (7–9).

One argument against deliberative democracy derives from a common premise in wisdom of crowds theory, which states that in order for groups to produce accurate beliefs, individuals within those groups must be statistically independent, such that their errors are uncorrelated and cancel out in aggregate (3, 10, 11). When individuals can influence each other, the dynamics of herding and groupthink are expected to undermine belief accuracy (10, 11), an argument that has raised concerns about the value of deliberative democracy (12). However, experimental research has shown that when individuals in a group can observe the beliefs of other members, information exchange can improve group accuracy even as individuals become more similar (13, 14). This effect can be explained by the observation that individuals who are more accurate revise their answers less in response to social information, thus pulling the mean belief toward the true answer (13, 15).

While such results are promising, political beliefs are shaped by cognitive biases that are not present in the nonpartisan estimation tasks (e.g., distance estimates) that have frequently been employed in experimental studies of the wisdom of crowds (11, 13, 14). A key finding of political attitude research is that partisan bias can shape not only value statements but also beliefs about facts (16–19). Such biases persist even when survey respondents are offered a financial incentive for their accuracy (17, 20). One explanation for the emergence of partisan bias in factual beliefs is motivated reasoning (21). Motivated reasoning results from the psychological preference for cognitive consistency, which means that people will adjust their beliefs to be consistent with each other (22). This preference can affect political attitudes, such that people will adjust their beliefs about the world to support their preferences for different parties or politicians (18).

Even when inaccurate beliefs are shaped by motivated reasoning and when corrected beliefs would be less supportive of party loyalties, experimental evidence suggests that accuracy can be improved by information exposure (23). In politically heterogeneous networks containing both Democrats and Republicans, social influence has been found to improve belief accuracy and reduce partisan biases (20, 24). However, theories of political polarization suggest that homogeneous networks—containing members of only one political party—will reverse the expected learning effects of social information processing and instead amplify partisan biases (9, 25, 26).

The risk of homogeneous networks derives from the expectation that response to social information on partisan topics is correlated with belief extremity, rather than belief accuracy (25, 26). However, previous research on political polarization (9, 16, 26) has been concerned primarily with attitude differences, and has not directly examined the effect of social

Significance Statement

Normative theories of deliberative democracy are based on the premise that social information processing can improve group beliefs. Research on the "wisdom of crowds" has found that information exchange can increase belief accuracy in many cases, but theories of political polarization imply that groups will become more extreme—and less accurate—when beliefs are motivated by partisan political bias. While this risk is not expected to emerge in politically heterogeneous networks, homogeneous social networks are expected to amplify partisan bias when people communicate only with members of their own political party. However, we find that the wisdom of crowds is robust to partisan bias. Social influence not only increases accuracy but decreases polarization without between-group network ties.

J.B. and E.P. designed the experiment, analyzed the data, and wrote the paper. J.B. collected the data. J.B. and D.C. constructed the data collection tool for Experiment 1. J.B. constructed the data collection tool for Experiment 2. All authors commented on and approved the final manuscript.

The authors declare no conflict of interest

[1]To whom correspondence should be addressed. E-mail: joshua.aaron.becker@gmail.com

influence on belief accuracy. To understand the potential effects of partisan bias on the wisdom of crowds, we first study a formal model of belief formation to generate hypotheses relating polarization theories to political belief accuracy. This model is formally identical to that used in previous research on the wisdom of crowds (13, 27), but parameterized to account for a possible correlation between belief extremity and adjustment to social information. Echoing previous experimental findings (20), this model shows that opposing biases cancel out in in politically diverse bipartisan networks, leaving the average belief unchanged even when bias is correlated with response to social information. However, in politically homogeneous "echo chamber" networks, a correlation between bias and adjustment causes group beliefs to become more extreme and less accurate (Fig. S4), consistent with political theories of polarization (26) (see *SI Appendix* for detailed model results).

To test whether the wisdom of crowds is robust to partisan bias, we conducted two web-based experiments examining social influence in homogeneous social networks. Contrary to predictions based on the "law of group polarization" (26) we find that homogeneous social networks are not sufficient to amplify partisan biases. Instead, we find that beliefs become more accurate and less polarized. These results suggest that prior models of the wisdom of crowds generalize to factual belief formation on partisan political topics in politically homogeneous networks.

1. Experimental Design

Following a pre-registered experimental design, our first experiment asked subjects recruited from Amazon Mechanical Turk to answer four fact-based questions (e.g., "What was the unemployment rate in the last month of Barack Obama's presidential administration?"). Subjects were compensated for their participation according to the accuracy of their final responses. The four questions used in this experiment (*Materials and Methods*) were selected because they showed the greatest levels of partisan bias among 25 pre-tested questions.

Subjects were randomly assigned to either a social condition or a control condition. For each question, subjects first provided an independent answer ("Round 1"). In the social condition, subjects were then shown the average belief of four other subjects connected to them in a social network, and were prompted to provide a second, revised answer ("Round 2"). Subjects in the social condition were then shown the average revised answer of their network neighbors and were prompted to provide a third and final answer ("Round 3"). In the control condition, subjects were prompted to provide their answer three times, but with no social information. Besides the absence or presence of social information, subject experience was identical in both social and control conditions. Subjects in both conditions were provided 60 seconds to provide their answer each round, for a total of 3 minutes per question. As soon as subjects provided their response, they were advanced to the next round, even if there was time remaining.

Each trial contained 35 subjects. For each trial in the social condition, all subjects participated simultaneously. Subjects in the social condition were connected to each other in random networks in which each subject observed the average response of four other subjects and was observed by those same four subjects, forming a single connected network of 35 subjects. To test whether the wisdom of crowds is robust to partisan bias in politically homogeneous networks, each trial in each condition consisted of either only Republicans or only Democrats. Subjects in the social condition interacted anonymously and were not informed that they were observing the responses by people who shared their partisan preferences.

We controlled for question order effects by using four question sets, each of which were identical except for the order in which questions were presented (see *SI Appendix*). For each question set, we collected data for 3 networked groups and 1 control group for each political party (i.e., 4 independent groups for each party). In total, we collected data for 12 networks and 4 control groups for each party (1,120 subjects in total). Figure S1 (*SI Appendix*) illustrates our experimental design.

The experimental questions have true answers with values ranging from 4.9 to 224,600,000. In order to compare across questions, we follow similar studies (11) and log-transform all responses and true values prior to analysis using the natural logarithm. This allows for comparison across conditions because log(A)-log(B) approximates percent difference, and thus calculated errors for each response are approximately equal to percent error. This also accounts for the observation that estimates of this type are frequently distributed log-normally (11, 28). We find that alternative normalization procedures produce comparable results (*SI Appendix*).

Because responses by individuals within a social network are not independent, we measure all outcomes at the trial level. To produce this metric, we first calculate the mean (logged) belief of the 35 responses given for a single round of a single question in a single trial. We then measure group error for each round of each question as the absolute value of the arithmetic difference between the mean (logged) belief and the (logged) true value. We then measure the change in error for each question of each trial as the arithmetic difference between the error of the mean at Round 1 and the error of the mean at Round 3. This method produces four measurements of change in error for each trial, i.e. one for each question. We then calculate the average of this value over all four questions completed by each trial to measure average change in error for each trial. We thus produce 24 independent observations of the effect of social influence on group accuracy when beliefs are motivated by partisan bias, including 12 independent observations of Republican networks and 12 independent observations of Democrat networks. In addition, we produce 8 independent control observations, including 4 independent observations of Republican control groups and 4 independent observations of Democrat control groups.

We replicated this entire design in a second experiment, with modifications intended to increase the effect of partisan bias on responses to social information. We describe this replication below.

Results (Experiment 1)

We find no evidence that social influence in homogeneous networks either reduces accuracy or increases polarization on factual beliefs. Instead, we find that social influence increased accuracy for both Republicans and Democrats and also decreased polarization despite the absence of between-group ties. We begin our analysis by confirming that in Experiment 1, subjects' independent beliefs demonstrated partisan bias, as expected based on previous research (5, 17, 20). In Round

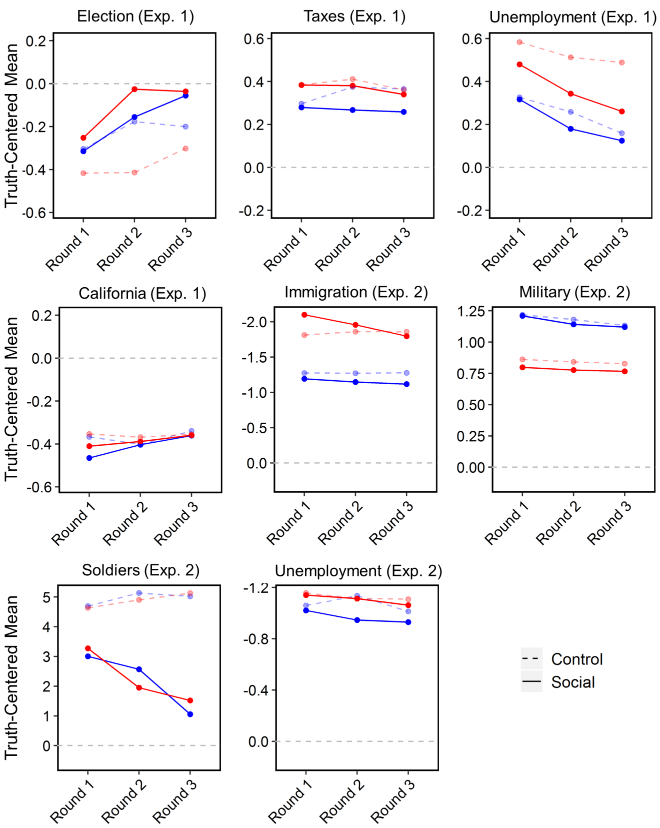

Fig. 1. Normalized, truth-centered mean at each round, averaged across 12 social trials (solid line) and 4 control trials (dashed line). Control groups show more random variation than social groups due to the smaller sample size. Each panel shows one question. Red indicates responses by Republicans, and blue indicates responses by Democrats. For questions with a negative true answer (Immigration, Unemployment) the normalization process in Experiment 2 reverses the sign, and the y-axis is inverted to show relative under- and over-estimates (e.g., subjects overestimated immigration.)

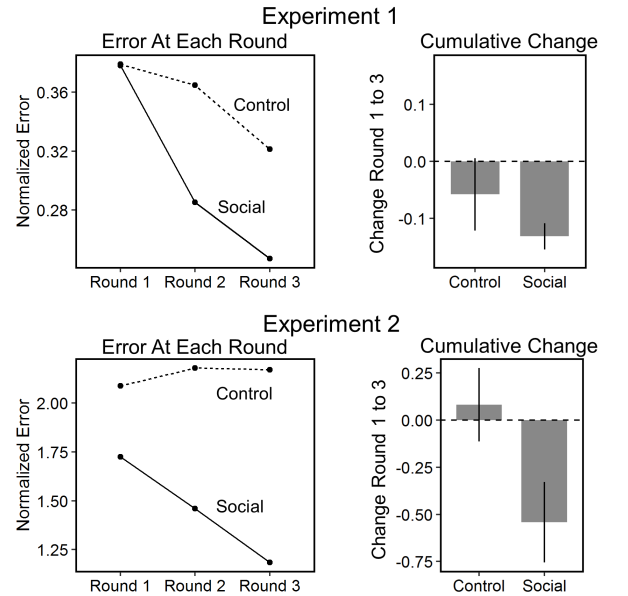

Fig. 2. LEFT: Normalized error of the mean, averaged across 24 social conditions (solid line) and 8 control conditions (dashed line) at each round of the experiment. RIGHT: Cumulative change in error from Round 1 to Round 3. Error bars display standard 95% confidence interval around the mean.

1 (before social influence), responses provided by Democrats were significantly different from responses provided by Republicans for all questions (See Fig. 1; $P<0.001$ for all questions except race in California, for which $P<0.05$) (see *SI Appendix*).

Effect of Social Influence on Belief Accuracy. To illustrate the change in beliefs for each question, Figure 1 shows the truth-centered mean of normalized beliefs (so that a negative value indicates an underestimate, and a positive value indicates an overestimate) in social conditions at each round of both experiments. The value for each data point is obtained by calculating the arithmetic difference between the mean belief and the true value at each round for each question, and then averaging this value across all 12 social network trials for each political party. In every case, the average estimate became closer to the true value after social influence.

To test whether this change could be explained by random fluctuation, we calculate the error for each round of each question as the absolute value of the truth-centered mean (i.e., the absolute distance from truth). We then calculate the change in error from Round 1 to Round 3, and average this value across all 4 questions to measure average change in absolute error within each trial. This analysis determines whether, on average, the group mean became closer to the true value after social influence. For those in the social condition, we find that error at Round 3 was significantly lower than error at Round 1 for every one of the 12 Republican trials ($P<0.001$) as well as every one of the 12 Democrat trials ($P<0.001$) in Experiment 1. Across both Republicans and Democrats, we find that the average error of the mean decreased by 35% from Round 1 to Round 3.

One possibility is that improvement in the social condition is due to the opportunity for subjects to revise their answers. To test whether this is the case, we compared improvement in the social condition with improvement in the control condition. Following the procedure described above, we calculate the average change in error for the 24 social network trials and the 8 control trials, shown in Figure 2. We find that error did decrease slightly in the control condition ($P<0.15$), but that the change in the social condition was significantly greater than the control condition ($P<0.03$), indicating that the reduction in error in homogeneous social networks cannot be explained by individual learning effects. The error of the mean in control groups decreased by only 15%, a substantially smaller change than the 35% decrease in social networks. Thus while providing individuals the opportunity to revise their answer may improve belief accuracy, these results suggest that social information processing—even in homogeneous partisan groups—can help counteract the effects of partisan bias.

Another possibility is that individuals became less accurate even as the group mean became more accurate, which would occur if individual beliefs become more widely dispersed—e.g., if moderates and extremists moved in opposite directions. To investigate this possibility, we first measure the standard deviation of responses by each of the 24 networked groups in Experiment 1 before and after information exchange, averaging across all four questions. We find that standard deviation decreased significantly from Round 1 to Round 3 in social networks ($P<0.001$) but did not significantly change for control groups ($P=0.25$). We find that the change in networks was significantly greater than change in control groups (P<0.001), suggesting that information exchange in homogeneous social networks leads to increased similarity among group members.

We also directly test the effect of social influence on average individual error (as opposed to the error of the average). This quantity is measured by first averaging error across all individuals within a group for a given question, then averaging across all questions in a trial, and then averaging across all 24 social network trials. For Experiment 1, we find that average individual error decreased in social networks (P<0.001). While individual error also decreased slightly in control groups (P<0.11), the improvement was significantly smaller in control groups than social networks (P<0.001), with a 7% decrease in the average error of isolated individuals as compared with a 33% decrease in error by individuals in social networks.

Robustness to Partisan Priming (Experiment 2)

One possibility is that Experiment 1 did not fully capture the effects of partisan bias. A notable observation is that estimation bias—the tendency to under- or overestimate—was in the same direction for both Republicans and Democrats. However, nearly all of the 25 pilot questions generated bias in the same direction. We also find this pattern in previous research on partisan factual beliefs (17), suggesting that same-direction bias is a common feature of partisan beliefs. While this same-direction bias runs counter to intuitive expectations about partisan polarization, it is consistent with previous research on estimation bias, which shows that people have a general tendency to under- or overestimate for any given question (28). The belief differences between Democrats and Republicans can be understood as an additional partisan bias added on top of a general estimation bias.

Nonetheless, a limitation of Experiment 1 is that questions were chosen based on the numeric magnitude of bias in pre-testing, and not the controversial nature of the questions. Moreover, the experimental interface was politically neutral and did not communicate to subjects in the social condition that they were in homogeneous partisan networks, factors which may have prevented subjects from perceiving the questions as partisan in nature. We therefore replicated our initial experiment with several changes designed to increase the effect of partisan bias on response to social information.

Replication Methods. Instead of choosing questions based on numeric polarization in pre-testing, we selected questions based on their connection to controversial policy topics. For example, we asked participants about the number of illegal immigrants in the U.S. at a time when illegal immigration was at the center of national debate (when disagreement over "the wall" with Mexico led to a U.S. government shutdown in January 2019). We also framed questions to emphasize change (i.e. we requested numeric estimates for the magnitude and direction of change) to allow for more partisan expressiveness. We re-used one question from Experiment 1, asking about unemployment, because that question taps into a strong policy controversy (the economy) and showed the greatest partisan bias in the first experiment. By re-using this question with an emphasis on directional change, we expected to observe demonstration of a split-direction partisan bias. Exact wording of all four questions is provided in *Materials and Methods*.

In addition to selecting more controversial questions, we also modified the experimental interface to include partisan primes that have been shown in prior research (20) to enhance the effects of partisan bias on social information processing. First, we required all subjects to confirm their political party prior to entering the experimental interface, to prime them to the political nature of the study. Second, we included an image of an elephant and a donkey (i.e., symbols for the Democratic and Republican parties) on the experimental interface (see Fig. S3 in the *SI Appendix*). Third, for subjects in the social condition, we indicated the party membership of other subjects in the study when providing social information. Finally, subjects upon recruitment were invited to participate in the "Politics Challenge," and the URL to the web platform included the phrase "Politics Challenge."

Questions in this second experiment allowed negative answers, for which the logarithm is not defined, and so we normalize results by dividing by the true answer, which also represents percent difference. However, this method leaves our analysis extremely sensitive to large values as might occur through typographic error. While these extreme values do not change our statistical analysis, the inclusion of all responses yields implausible effect sizes. (For example, we find that error in the social condition decreased by 3.6×10^{7}% while error in the control groups increased by 5.3×10^{4}%.) We therefore present results in the main text and figures after manually removing extremely large values, a process which impacts fewer than 1% of responses. An analysis that includes all submitted responses is provided in the *SI Appendix*.

Replication Results. As with Experiment 1, we begin our replication analysis by ensuring that subjects showed partisan bias, finding significant differences between Republicans and Democrats for all four questions (P<0.001). For the question on unemployment, which was re-used from Experiment 1 and reframed to emphasize change, we now observe a meaningful split between the two parties: a majority (54%) of Democrats stated that unemployment decreased under Obama, while a majority (67%) of Republicans stated the opposite. Nonetheless, the overall numeric bias was still in the same direction: the mean answer for both parties was an overestimate. As this example shows, divergent beliefs between Democrats and Republicans can nonetheless generate numeric estimation bias in the same direction.

Figures 1, 2 and 3 show outcomes of the replication. We again find that social influence increased the accuracy of mean beliefs for both Democrats (P<0.03) and Republicans (P<0.001). Across all trials, we found that the error of the mean decreased by 31% for subjects in the social condition, approximately the same effect size observed in Experiment 1. In contrast, we saw a 4% increase in error for the control condition, though this change was not statistically significant (P>0.46). The two conditions were significantly different (P<0.002), indicating that the benefits of social information cannot be explained by individual learning effects.

Similar to Experiment 1, we found that standard deviation decreased significantly in the social condition (P<0.001), but increased slightly in the control condition (P>0.19) and the two conditions were significantly different (P<0.001). This result shows that subjects became more similar over time as a result of social information, indicating that social learning effects are robust to explicit partisan primes. In addition to learning at the group level, we found a 34% decrease in individual error for

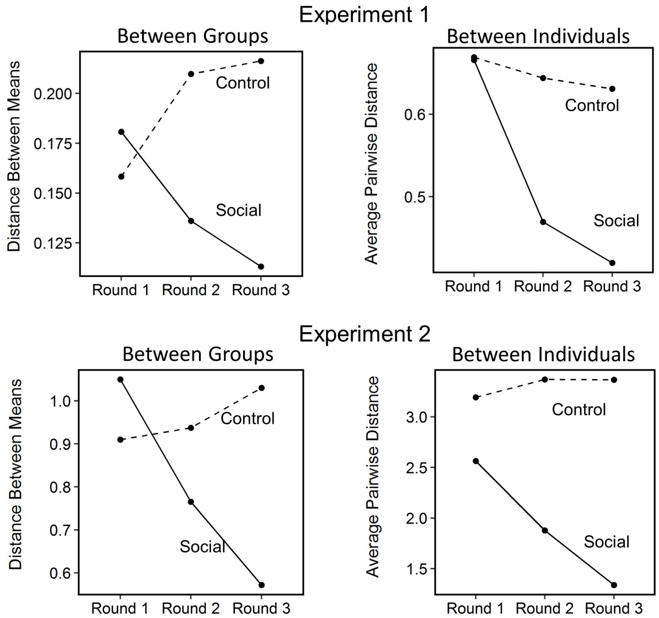

Fig. 3. Points indicate polarization at each round of the experiment for both social networks (solid line) and control groups (dashed line). LEFT: Difference in the normalized mean belief of Democrats and the normalized mean belief of Republicans. RIGHT: Average pairwise distance of normalized responses, which measures the expected difference between a randomly selected Democrat and Republican.

subjects in the social conditions (P<0.001) and a nominal 3% increase in individual error for control subjects (P>0.74). The two conditions were significantly different (P<0.001), showing that social learning is robust to partisan priming for both group-level improvement and individual improvement.

Polarization and the Wisdom of Crowds

Results from both experiments show that the wisdom of crowds in networks is robust to political partisan bias. We find that an increase of in-group belief similarity generates improvements at both the group level and the individual level. One risk, however, is that this increase of in-group similarity is accompanied by a decrease in between-group similarity, generating increased belief polarization even as groups become more accurate. To measure belief polarization, we conduct a paired analysis for each experiment matching the 12 Republican networks with the 12 Democrat networks (based on trial number, as per our pre-registered analysis) and calculating their similarity at each Round (see *SI Appendix*).

We measured polarization using two outcomes. Figure 3 (left) shows the average distance (absolute value of the arithmetic difference, see *SI Appendix*) between the mean normalized belief for Republicans and the mean normalized belief for Democrats at each round of the experiment. Among subjects in the social condition, the average distance between the mean belief of Democrats and the mean belief of Republicans decreased by 37% for Experiment 1 (P<0.01) and 46% for Experiment 2 (P<0.02). In contrast, the distance between the mean Republican and Democrat belief nominally increased for the control condition in both experiments, though the effects were not statistically significant (P<0.13 for Exp. 1, and P>0.87 for Exp. 2). Overall, the change in polarization was significantly different between the control and social conditions (P<0.01 for Exp. 1, P<0.08 for Exp. 2).

As a second measure of polarization, Figure 3 (right) shows the average pairwise distance between individual Republicans and Democrats. This metric measures the average distance between every possible 2-person cross-party pairing, and reflects the expected distance between the belief of a randomly selected Democrat and a randomly selected Republican. This outcome can be understood as reflecting the expected distance in belief between a Democrat and a Republican who could meet by chance in a public forum. For this metric, we found that Democrats and Republicans embedded in homogeneous social networks became more similar in all 24 trials across both experiments, with a 37% decrease in average pairwise distance for Experiment 1 (P<0.001) and a 48% decrease for Experiment 2. Outcomes for control groups show that this value did not change reliably in the absence of social information, showing a nominal decrease in Experiment 1 (6% change, P>0.12) but a nominal increase in Experiment 2 (5% change, P=0.25). Overall, decrease in average pairwise distance was significantly greater in social networks than in control groups (P<0.01 for each experiment).

Discussion

We observed that the mean response to objective, fact-based questions became more accurate as a result of social influence, despite the fact that beliefs were shaped by partisan bias and individuals were embedded in politically homogeneous social networks. In contrast to theories of polarization (26), our results are consistent with the explanation that accurate individuals exert the greatest influence on factual political beliefs as predicted by prior research on the wisdom of crowds (13). In the context of growing concerns about the effects of partisan echo chambers, our results suggest that deliberative democracy may be possible even in politically segregated social networks. Homogeneous social networks, such as those we study, are not on their own sufficient to increase partisan political polarization.

This finding, however, presents a tension: information exchange can mitigate partisan bias, yet public opinion remains polarized. Although we observe decreased polarization and increased accuracy, some error remains as well as some differences between political parties. Polarization can exist despite the potential for social learning. The co-existence of polarization and social learning may be due to structural factors such as network centralization (i.e., the presence of disproportionately central individuals), which can generate and sustain belief polarization in social networks. Network centralization in general has been found to undermine the wisdom of crowds (13); and the ability to obtain central positions in social networks (e.g., through broadcast media or web-based platforms) could allow extremists to exert disproportional influence on group beliefs. In simulation (*SI Appendix*) we find that a correlation between belief extremity and social network centrality can cause the wisdom of crowds to fail, such that social influence simply enhances existing partisan bias, as predicted by the law of group polarization.

In considering the limitations of our study, it is important to address the generalizability of our research. One concern is that our subject population is not a nationally representative sample; Amazon Mechanical Turk (MTurk) attracts subjects who are younger and more digitally sophisticated than the general population (29). Subjects in our experiment may thus have relied more effectively on web search, placing less weight on social information, and so our results may be weaker than would be expected in the general population. MTurkers also tend to skew liberal, and so our sample may have underestimated initial polarization. Generally, however, analyses of political research find that research on non-representative samples such as MTurk typically replicate well on nationally representative samples (30), suggesting our experimental results are likely to replicate. A second concern about generalizability is ecological validity, i.e. whether our experiment reflects the dynamics of political belief formation more broadly. We paid subjects for accuracy, which was necessary to discourage subjects from entering nonsense answers, but political attitudes are typically formed without financial incentive. However, prior research on political beliefs has found that subjects can become more accurate even when they are not compensated for accuracy (23), suggesting that financial incentives could impact the effect sizes (17) but not the direction of belief change. Nonetheless, some empirical contexts may produce perverse incentives that drive people away from accuracy, if, for example, people are motivated to be provocative instead of accurate.

Because accuracy incentives appear necessary for the wisdom of crowds to emerge, an important direction for future work is to examine how individual motivations toward accuracy can vary across empirical settings. A single person motivated by controversy would not be likely to disrupt the wisdom of crowds (unless they hold a central network position), but an entire population motivated by controversy might meet the conditions required for the law of group polarization to hold. Under the assumption that some people are not generally motivated toward accuracy, the robustness of our findings to different empirical settings would depend on the proportion of individuals who are motivated to hold accurate beliefs and the proportion of individuals who are motivated to advance controversial views.

The primary goal of this research was to test whether the wisdom of crowds is robust to partisan bias by studying belief formation about controversial topics in politically homogeneous networks. Based on our experimental results, we reject the hypothesis that social information in politically homogeneous networks will always amplify existing biases. Rather, we find that in the networks studied here, information exchange increases belief accuracy and reduces polarization. While the wisdom of crowds may not hold in all possible empirical settings, our results open the question of when—if ever, and in what circumstances—the wisdom of partisan crowds will fail.

Materials and Methods

Subjects provided informed consent prior to entering the experimental interface. Experiment 1 was run on a custom platform and approved by University of Pennsylvania IRB, Experiment 2 was run on the Empirica.ly platform and approved by Northwestern University IRB. See *SI Appendix* for replication data and code.

Questions for Experiment 1: (1) In the 2004 election, individuals gave $269.8 million to Republican candidate George W. Bush. How much did they give to Democratic candidate John Kerry? (Answer in millions of dollars - e.g., 1 for $1 million.) (2) According to 2010 estimates, what percentage of people in the state of California identify as Black/African-American, Hispanic, or Asian? (Give a number 0-100) What was the U.S. unemployment rate at the end of Barack Obama's presidential administration – i.e., what percent of people were unemployed in December 2016? (Give a number 0-100) (4) In 1980, tax revenue was 18.5% of the economy (as a proportion of GDP). What was tax revenue as a percent of the economy in 2010? (Give a number 0 to 100).

Questions for Experiment 2: (1) For every dollar the federal government spent in fiscal year 2016, about how much went to the Department of Defense (US Military)? Answer with a number between 0 and 100. (2) In 2007, it was estimated that 6.9 million unauthorized immigrants from Mexico lived in the United States. How much did this number change by 2016, before President Trump was elected? Enter a positive number if you think it increased, and a negative number if you think it decreased. Express your answer as a percent change. (3) How much did the unemployment rate in the United States change from the beginning to the end of Democratic President Barack Obama's term in office? Enter a positive number if you think it increased, and a negative number if you think it decreased. Express your answer as a percent change. (4) About how many U.S. soldiers were killed in Iraq between the invasion in 2003 and the withdrawal of troops in December 2011?

ACKNOWLEDGMENTS. This project was made possible by the Summer Institute in Computational Social Science (SICSS). We thank Matt Salganik, Chris Bail, participants at SICSS, Michael X. Delli-Carpini, and Steven Klein for helpful comments. We are grateful to Joseph "Mac" Abruzzo for help with recruitment and to Abdullah Almaatouq for support with the Empirica.ly platform.

1. Somin I (2010) Deliberative democracy and political ignorance. *Crit Rev* 22(2-3):253–279.
2. Pennycook G, Rand DG (2019) Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proc Natl Acad Sci* 116(7):2521–2526.
3. Nofer M, Hinz O (2014) Are crowds on the internet wiser than experts? The case of a stock prediction community. *Journal of Business Economics* 84(3):303–338.
4. Wolf M, Krause J, Carney PA, Bogart A, Kurvers RHJM (2015) Collective intelligence meets medical decision-making: the collective outperforms the best radiologist. *PLoS ONE* 10(8):e0134269.
5. Bartels LM (1996) Uninformed votes: Information effects in presidential elections. *American Journal of Political Science* pp. 194–230.
6. Galton F (1907) Vox populi (The wisdom of crowds). *Nature* 75:450–51.
7. Chambers S (2003) Deliberative democratic theory. *Annu Rev Polit Sci* 6(1):307–326.
8. Moscovici S, Zavalloni M (1969) The group as a polarizer of attitudes. *Journal of Personality and Social Psychology* 12(2):125–135.
9. Sunstein CR (2002) The law of group polarization. *J Polit Philos* 10(2):175–195.
10. Surowiecki J (2004) *The wisdom of crowds.* (Anchor).
11. Lorenz J, Rauhut H, Schweitzer F, Helbing D (2011) How social influence can undermine the wisdom of crowd effect. *Proc Natl Acad Sci* 108(22):9020–9025.
12. Solomon M (2006) Groupthink versus the wisdom of crowds: The social epistemology of deliberation and dissent. *The Southern Journal of Philosophy* 44(S1):28–42.
13. Becker J, Brackbill D, Centola D (2017) Network Dynamics of Social Influence in the Wisdom of Crowds. *Proceedings of the National Academy of Sciences* 114(26):E5070–E5076.
14. Jayles B, et al. (2017) How social information can improve estimation accuracy in human groups. *Proceedings of the National Academy of Sciences* p. 201703695.
15. Madirolas G, de Polavieja GG (2015) Improving collective estimations using resistance to social influence. *PLoS computational biology* 11(11):e1004594.
16. Bartels LM (2002) Beyond the running tally: Partisan bias in political perceptions. *Political Behavior* 24(2):117–150.
17. Bullock JG, Gerber AS, Hill SJ, Huber GA (2015) Partisan Bias in Factual Beliefs about Politics. *Quarterly Journal of Political Science* 10(4):519–578.
18. Lodge M, Taber CS (2013) *The rationalizing voter.* (Cambridge University Press).
19. Taber CS, Lodge M (2006) Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science* 50(3):755–769.
20. Guilbeault D, Becker J, Centola D (2018) Social learning and partisan bias in the interpretation of climate trends. *Proceedings of the National Academy of Sciences* p. 201722664.
21. Kunda Z (1990) The case for motivated reasoning. *Psychological bulletin* 108(3):480.
22. Gawronski B (2012) Back to the future of dissonance theory: Cognitive consistency as a core motive. *Social Cognition* 30(6):652–668.
23. Wood T, Porter E (2018) The elusive backfire effect: Mass attitudes' steadfast factual adherence. *Political Behavior* pp. 1–29.
24. Shi F, Teplitskiy M, Duede E, Evans JA (2019) The wisdom of polarized crowds. *Nature Human Behaviour* p. 1.
25. Schkade D, Sunstein CR, Kahneman D (2000) Deliberating about dollars: The severity shift. *Columbia Law Review* pp. 1139–1175.
26. Sunstein CR (2009) *Going to extremes: How like minds unite and divide.* (Oxford University Press).
27. DeGroot MH (1974) Reaching a consensus. *Journal of the American Statistical Association* 69(345):118–121.
28. Kao AB, et al. (2018) Counteracting estimation bias and social influence to improve the wisdom of crowds. *J R Soc Interface* 15(141).
29. Munger K, Luca M, Nagler J, Tucker J (2019) Age Matters: Sampling Strategies for Studying Digital Media Effects.
30. Coppock A, Leeper TJ, Mullinix KJ (2018) The Generalizability of Heterogeneous Treatment Effect Estimates Across Samples.