Anger makes fake news viral online

Yuwei Chuai\textsuperscript{1}, Jichang Zhao\textsuperscript{1, 2, *}

\textsuperscript{1}School of Economics and Management, Beihang University, China
\textsuperscript{2}Beijing Advanced Innovation Center for Big Data and Brain Computing, China
*Correspondence to: jichang@buaa.edu.cn

Fake news that manipulates political elections, strikes financial systems and even incites riots is more viral than real news online, resulting in unstable societies and buffeted democracy. The easier contagion of fake news online can be causally explained by the greater anger it carries. Offline questionnaires reveal that anger leads to more incentivized audiences in terms of anxiety management and information sharing and accordingly makes fake news more contagious than real news online. Our results suggest that the digital contagion of emotions, in particular anger, should be comprehensively considered in profiling the spread of online information. Cures such as tagging anger in social media could be implemented to slow or prevent the contagion of fake news at the source.

Fake news refers to information that is fabricated, misleading, and verifiably false (1, 2). Most people broadly accept information instead of critically questioning its authenticity (1). In particular, with the boom of social media, on which individuals can be simultaneously producers and consumers of information, ordinary people can easily participate in circulation and gain influence through posting (e.g., tweeting) and reposting (e.g., retweeting). Consequently, the
impact of fake news on social media could be disproportionate (3) and profound (4), especially in the political (2, 4–7) and economic fields (8). In the first few months of the 2016 U.S. presidential election, on average, each adult was exposed to more than one fake news item that was not only widely spread but also deliberately biased (6). Furthermore, fake news is more likely to appear in the highly uncertain conditions of emergencies, such as disease epidemics and outbreaks (9, 10), accidents and conflicts (11), which makes the spread of fake news a byproduct of the natural response that people have to disastrous events, and social media can be fertile ground for this spread (43) online.

Fake news is more viral than real (true) news online (2). The mechanism underlying its fast spread, though critical, remains unresolved. Unique structural features in the circulation of fake news, such as long diameters of penetration, have been revealed and have been found to be platform independent (13–16). However, fake news is generally verified to be false after explosive circulation (17); thus, in the early spread, it is essentially not thought to be fake, so the structural uniqueness is the manifestation of its fast spread, rather than a cause that can fundamentally explain its viral proliferation. Individuals, either human or bots (18), posting and reposting fake news on social media are an alternative cause, in particular, the human that occupies the dominant partition (19). The spread of news is associated with the friends and followers of the author. Nevertheless, user characteristics fail to sufficiently explain the easy contagion of fake news due to their greater effects on the dissemination of real news (2). The content of fake news, which was also found to be entangled with spread (2, 20), could offer promising directions in probing the mechanism of its fast spread. More importantly, instead of examining spreading structures (2, 13) and reposter demographics (21) after the circulation was ignited, revealing the mechanism at the source that independent to user demographics would be powerful in inspiring new cures with the minimum invasion of privacy. Hence, we would rather differentiate fake news from real news at the very beginning of their spread through scrutiny of
content to figure out new treatments against fake news that can be implemented without delay.

Online news content not only delivers factual information but also carries sophisticated emotional signals. Being embedded in information spread, the digital contagion of emotions, in which individuals experience the same feelings on social media, is similar to the face-to-face emotion exchange offline (22, 23). Emotions further impact the spread of information, e.g., promoting the sharing (24) or shaping the paths (25). When the relevance between content quality and popularity is not strong (26), the emotions involved and their influence on psychological arousal may be key (2, 27–29). Moreover, the spread of different emotions can inherently be distinguished (29), implying that emotions conveyed by both fake and real news could offer comparative proxy measurements to examine the mechanisms underlying their circulation. In fact, fake news is preferentially injected with emotions such as anger for political attacks (30). However, differentiating fake news from real news is rarely based on emotions delivered in the content and incentives beyond reposting in extant efforts. The discrepancy in user perceptions between fake news and real news were unraveled through emotions of replies (2), while the emotions that inherently carried by news themselves are not considered in explaining the circulation. In fact, it has been found that negative emotions in content might cause positive responses (e.g., sympathy) (31), meaning the emotions in particular the negative parts should be directly examined in the spread of fake news. At the same time, although content on social media can be short, simplifying the emotions it carries into a single emotion might miss the emotional richness (2, 32, 33) and lead to failure of emotion recognition and inconsistent results (23, 28, 29, 34).

In this study, by successfully combining digital traces on social media and offline questionnaires, we aim to unravel the mechanism underlying the fast spread of fake news by answering three key questions: What are the differences in the emotional distributions of real and fake news? Can these differences explain why fake news is more infectious than real news? How do
they affect the incentives behind news reposting?

We collected a large dataset of both fake news and real news from Weibo, the most popular Twitter-like service in China, that includes 10,000 true news items posted by credibly verified users and 22,479 fake news items endorsed by an official committee of Weibo after wide dissemination (see SM S1 for more details). On the basis of the number of followers on behalf of the broadcasting potential of authors and the number of retweets on behalf of the spreading capability of news (16), we assemble both categories of news into treatment and control groups. For example, taking fake news with low numbers of followers and high volumes of retweets (LHF news) as the treatment group, the controlled counterparts consist of either fake news with high volumes of followers and low numbers of retweets (HLF news) or true news with high volumes of followers and low numbers of retweets (HLT news) (see SM S2 for details). By intentionally selecting news that is weakly retweeted but posted by highly followed authors, the possible effects from users can be controlled to amplify the spread promotion resulting from the particular emotion content it carries. Moreover, although fake news is statistically more contagious (longer path, faster speed, lasts longer, and gets more retweets) than real news (see SM S2.3 and 3), not every fake news item is necessarily more viral than any real news item. For instance, the diffusion capability of highly retweeted true news is definitely more powerful than that of lowly circulated fake news. Therefore, we would compare LHF news with HLF news and HLT news first and then extend the comparison to the full spectrum of discrepancies between true (T) news and fake (F) news in terms of emotions.

Emotional signals carried in either fake or real news can be sophisticated, i.e., a combination of elementary compounds rather than a single one (33). The distribution of five emotions that represent basic human feelings (2, 35, 36), namely, anger, disgust, joy, sadness and fear, is inferred for each news item in our data through a lexicon that is manually labeled to cover 87.1% of news items with the remaining considered neutral (see SM S4). Emotions with strong
presence in the distribution are the feelings that the sender of the news wishes the receivers to experience (37). The proportion of anger (Fig. 1A) in LHF news is expected to be significantly higher than that in both HLF and HLT news, while joy is expected to be lower (Fig. 1E). The comparison is then extended to a full spectrum between all fake news and real news, and consistent results, though with shrinking gaps for anger and joy, as expected, are obtained (Figs. 1B, F). Furthermore, the dominance of anger in fake news (especially highly retweeted news) and joy in real news (even lowly retweeted news) is further confirmed with better resolution in the distribution of emotions of keywords that precisely separate the treatment groups from control groups (see SM S6 and Fig. S10). These observations persistently suggest that fake news carries more anger yet less joy than real news and imply the possibility that anger might promote the fast spread of fake news online. The divergence in anger and joy between fake news and real news is robust and independent of emotion inference models and emotion distribution measures (see SM S7). Even in specific events like COVID-19, the dominance of anger to joy in highly retweeted fakes news conformably suggests the promotion of anger in fast spread of fake news (see SM S7). By contrast the near overlap in disgust between different types of news (Fig. 1C, D), the less occupation of fake news when sadness more than 0.5 (Fig. 1H), and the more dominant position of fear in HLF news (Fig. 1I) indicate their less positive roles in the virality of fake news (24, 34, 38). Therefore, significant gaps across news groups could also be independent of circulation, and well-controlled causal inference is accordingly necessary for anger and joy.

To causally infer and qualify the promotion of anger and the prevention of joy in the spread of fake news, internal factors related to content (39), user (2) and external shocks such as disaster events (9) should be comprehensively controlled. Specifically, internal factors, including mention, hashtag, location, date, URL, length, topic, other emotions, follower (number of followers), friend (number of reciprocal followers) and external shocks including emergency (a
disaster event) constitute control variables in the logit and linear inference models (see SM S8). The results of the logit model (see SM S9) for lowly retweeted true (LT) news (control group) and highly retweeted fake (HF) news (treatment group) show that the coefficient of anger is significantly positive and the coefficient of joy is negative, implying that anger causally promotes the fast spread of fake news online. Other emotions are omitted (Table 1(1)) due to multicollinearity and their trivial impact on circulation. Moreover, for the logit model used to estimate all true and fake news, anger is positively associated with fake news, though with a smaller coefficient and narrower deviation, as anticipated (Table 1(2)). Recalling the gaps observed in emotion distributions across groups of news, all the results consistently suggest the positive promotion effect of anger in the circulation of fake news, particularly for news that is highly retweeted. The causally negative relationship between joy and fake news contrarily indicates its prevention in dissemination. To further qualify the influence of both anger and joy in the spread of fake news, a linear regression model with the number of retweets as the dependent variable is established (see SM S9). It is congruously found for fake news and all news that the coefficients of anger are significantly positive while the coefficients of joy are negative (Table 1 (3) and (4)), suggesting that anger can promote circulation and joy can prevent the spread. Specifically, supposing that other factors are fixed, increasing the occupation of anger by 0.1 and reducing that of joy by 0.1 in fake news leads to 5.8 more retweets, and 2.2 more retweets occur if anger is increased by 0.1 in place of other negative emotions but joy is fixed. The above causal relationships between emotions and circulation are robust to alternative emotion detection approaches such as competent machine learning models (see Table S17). For other significant factors, although mention can promote the spread of news (Table 1(3) and (4)), the coefficient is not significant for LHF news (Table 1(1)) and even prevents the spread of fake news (Table 1(2)); emergency is significantly positive in the logit models (Table 1(1) and (2)) but inconsistently negative in the linear models (Table 1(3) and (4)) (see SM S8 for more de-
Therefore, carrying more anger and less joy is the mechanism behind the fast spread of fake news that makes it more viral than real news online. More importantly, additional evidence from extensive datasets of English news on both Twitter and mainstream media further confirms the independence to cultures and platforms of this mechanism (see SM 10).

Negative stimuli such as anger elicit stronger and quicker emotional reactions and even behavioral responses than positive stimuli such as joy (40, 41). The odds of being forwarded through e-mails are also causally impacted by the physiological arousal caused by emotional articles, and those evoking high-arousal positive or negative emotions could be more viral (34). In the spread of fake news, the incentives behind the action of reposting that reignites circulation are therefore hypothetically associated with the anger and joy the news carries. Taking LHF news as the treatment group and HLF news and HLT news as the control groups, the possible associations between reposting incentives and emotions are examined through offline questionnaires. By selecting 15 typical news items with keywords from these groups (see SM S11), questionnaires are implemented to investigate four motivations for news reposting in social media (42), including anxiety management, information sharing, relationship management and self-enhancement. The subjects of the surveys are Weibo users, and the overlapping between offline subjects and online users is ensured (see SM S12). Preliminary results indicate that the motivation of anxiety management in LHF news is significantly higher than that in the control groups (Fig. 2A). Moreover, compared to HLT news, subjects are more intensively incentivized to share information when reposting HLF news and LHF news (Fig. 2B). Thus, fake news can stimulate strong motivation for information sharing; in particular, news that is widely disseminated can also strengthen the motivation for anxiety management. There is no significant variation in the motivation for relationship management across news groups (Fig. 2C), and the motivation for self-enhancement in HLT news is stronger than that in fake news (Fig. 2D). What is more interesting is that in questionnaires with keywords highlighted with marks, the unique
stimuli of widely circulated fake news for anxiety management is strengthened (see Fig. S23A). The incentive of information sharing is similarly enhanced for fake news (see Fig. S23B). All these results imply that the responses to the anger carried by fake news are sharing information and even managing anxiety. To validate this finding, the news in questionnaires is further split into anger-dominated news and joy-dominated news (see SM S13.2) to directly probe the impact of emotions. Compared to the retweeting motivations of joy-dominated news, anger-dominated news stimulates stronger incentives for anxiety management (Fig. 2E) and information sharing (Fig. 2F). Joy-dominated news ultimately excites stronger self-enhancement (Fig. 2H) than anger-dominated news. Meanwhile, no significant difference is observed between anger and joy in terms of relationship management motivation (Fig. 2G). Shuffling emotions randomly further testifies to the significance of these observations (see SM S13.2). Therefore, the greater anger delivered in fake news leads to more incentivized audiences with respect to anxiety management and information sharing, resulting in a greater likelihood of retweets and, thus, more viral contagion.

Our findings emphasize the necessity of considering emotions, particularly anger, in understanding the spread of information online. On social media, the associations between information diffusion and embedded emotions have been noted for a long time; however, the profiles of the roles of both positive and negative emotions are inconsistent and even contradictory across diverse contexts (23). Considering the heterogeneous influence on spreading from negative emotions such as anger and sadness (24, 34, 38), the causal impact on information diffusion should be examined with respect to well-resolved negative emotions. Instead of simplifying emotions binarily into positive and negative emotions, more elementary emotions are considered in this study, and the distribution of five emotions is inferred to reflect the complete emotional spectrum of news online. This more detailed spectrum of emotions identifies anger’s unique role in provoking strong incentives of anxiety management and information shar-
ing, which results in the virality of fake news online. From this perspective, emotions could be genes of fake news circulation, and similar to small mutations, they could make the virus go viral. Mutations that increase anger or reduce joy in fake news enhance its likelihood of being retweeted. Distinguishing structures in the circulation of fake news, which have been pervasively revealed in both Twitter (14) and Weibo (14), could also be deciphered based on the anger such news predominantly carries since anger prefers weak ties in social networks (2) and may inherently forge the diffusion structure of fake news (see SM 2.3). Meanwhile, the role of joy in preventing spread, especially in fake news, underlines the fundamentality of considering negative emotions of fine granularity to control and deepen future explorations. Therefore, it is anticipated that insights from emotions will improve the extant understanding of online information spread.

The vigorous promotion in circulation from anger implies new weapons against fake news. Although structural signals can be sensed at an early stage to target fake news (14), fake news spreads rapidly and reaches the peak of new retweets in less than one hour (see Fig. S7), so the negative impact has been exposed to a large population of audiences before identification. Moreover it can take more than three days for a post to be rated as false by outside fact-checkers on Facebook (44). What is worse, like a cat-and-mouse game between manipulation and detection, features derived from content or users that were found to be helpful in machine learning on targeting fake news (45) can be easily converted to inspire future countermeasures for fabricating more sophisticated false news (11). In particular, fake news related to emergencies is widely disseminated because of its clever combination with anger, which may explain why efforts to counter misperceptions about diseases during epidemics and outbreaks are not always effective (10). Inefficient or ineffective efforts to detect fake news and debunk misinformation by correcting both calls for new treatments and preventing the spread of anger could be a profound and promising direction. The early deviation in dissemination paths between fake news and real
news suggests the rapid effect of anger in shaping retweeting (25). For example, platforms such as Facebook, Twitter and Weibo should warn and discourage users as they try to retweet news that delivers too much anger and persuade them to assess the credibility of the information more critically. The trade-off between free speech and fake news prevention is the prime principle; however, a better balance would be achieved by tagging angry news (e.g., with an occupation of anger of more than 20%, see SM S14 for more details) at the very beginning to make audiences and potential spreaders less emotional and more rational (46).
References

1. D. M. J. Lazer, et al., The science of fake news. Science 359, 1094-1096 (2018).

2. S. Vosoughi, D. Roy, S. Aral, The spread of true and false news online. Science 359, 1146-1151 (2018).

3. J. Allen, B. Howland, M. Mobius, D. Rothschild, D. J. Watts, Evaluating the fake news problem at the scale of the information ecosystem. Sci. Adv. 6, eaay3539 (2020).

4. S. Aral, D. Eckles, Protecting elections from social media manipulation. Science 365, 858-861 (2019).

5. A. Bovet, H. A. Makse, Influence of fake news in Twitter during the 2016 US presidential election. Nat. Commun. 10, 1-14 (2019).

6. H. Allcott, M. Gentzkow, Social media and fake news in the 2016 election. J. Econ. Perspect 31, 211-236 (2017).

7. N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, D. Lazer, Fake news on Twitter during the 2016 U.S. presidential election. Science 363, 374-378 (2019).

8. K. Rapoza, “Can ‘Fake News’ Impact The Stock Market?” Forbes, 26 February 2017; www.forbes.com/sites/kenrapoza/2017/02/26/can-fake-news-impact-the-stock-market/.

9. L. Spinney, In Congo, fighting a virus and a groundswell of fake news. Science 363, 213-214 (2019).
10. J. M. Carey, V. Chi, D. J. Flynn, B. Nyhan, T. Zeitzoff, The effects of corrective information about disease epidemics and outbreaks: evidence from Zika and yellow fever in Brazil. *Sci. Adv.* **6**, eaaw7449 (2020).

11. I. Khaldarova, M. Pantti, Fake news: the narrative battle over the Ukrainian conflict. *Journal. Pract.* **10**, 891-901 (2016).

12. G. Miller, “Researchers are tracking another pandemic, tooof coronavirus misinformation,” *Science*, 24 March 2020; https://www.sciencemag.org/news/2020/03/researchers-are-tracking-another-epidemic-too-misinformation.

13. M. Del Vicario, *et al.*, The spreading of misinformation online. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 554-559 (2016).

14. Z. Zhao, *et al.*, Fake news propagates differently from real news even at early stages of spreading. *EPJ Data Sci.* **9**, 7 (2020).

15. N. F. Johnson, *et al.*, Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature** **573**, 261-265 (2019).

16. X. Wang, Y. Lan, J. Xiao, Anomalous structure and dynamics in news diffusion among heterogeneous individuals. *Nat. Hum. Behav.* **3**, 709-718 (2019).

17. S. Iyengar, D. S. Massey, Scientific communication in a post-truth society. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7656-7661 (2019).

18. C. Shao, *et al.*, The spread of low-credibility content by social bots. *Nat. Commun.* **9**, 1-9 (2018).
19. K. Langin, “Fake news spreads faster than true news on Twitter thanks to people, not bots,” *Science*, 8 March 2018; [https://www.sciencemag.org/news/2018/03/fake-news-spreads-faster-true-news-twitter-thanks-people-not-bots](https://www.sciencemag.org/news/2018/03/fake-news-spreads-faster-true-news-twitter-thanks-people-not-bots).

20. D. A. Scheufele, N. M. Krause, Science audiences, misinformation, and fake news. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 7662-7669 (2019).

21. A. Guess, J. Nagler, J. Tucker, Less than you think: prevalence and predictors of fake news dissemination on Facebook. *Sci. Adv.* **5**, eaau4586 (2019).

22. A. D. Kramer, J. E. Guillory, J. T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8788-8790 (2014).

23. A. Goldenberg, J. J. Gross, Digital emotion contagion. *Trends Cogn. Sci.* **24**, 316-328 (2020).

24. S. Stieglitz, L. Dang-Xuan, Emotions and information diffusion in social media: sentiment of microblogs and sharing behavior. *J. Manag. Inf. Syst.* **29**, 217-248 (2014).

25. W. J. Brady, J. A. Wills, J. T. Jost, J. A. Tucker, J. J. Van Bavel, Emotion shapes the diffusion of moralized content in social networks. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7313-7318 (2017).

26. A. Acerbi, Cognitive attraction and online misinformation. *Palgrave Commun.* **5**, 1-7 (2019).

27. L. Nummenmaa, *et al.*, Emotions promote social interaction by synchronizing brain activity across individuals. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 9599-9604 (2012).
28. K. Peters, Y. Kashima, A. Clark, Talking about others: emotionality and the dissemination of social information. *Eur. J. Soc. Psychol.* **39**, 207-222 (2009).

29. B. E. Weeks, Emotions, partisanship, and misperceptions: how anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *J. Commun.* **65**, 699-719 (2015).

30. M. Higgins, Mediated populism, culture and media form. *Palgrave Commun.* **3**, 1-5 (2017).

31. Q. Wang, Y. Jin, T. Yang, S. Cheng, An emotion-based independent cascade model for sentiment spreading. *Knowl.-Based Syst.* **116**, 86-93 (2017).

32. S. Du, Y. Tao, A. M. Martinez, Compound facial expressions of emotion. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E1454-1462 (2014).

33. E. Penz, M. K. Hogg, The role of mixed emotions in consumer behaviour. *Eur. J. Mark.* **45**, 104-132 (2011).

34. J. Berger, K. L. Milkman, What makes online content viral? *J. Mark. Res.* **49**, 192-205 (2012).

35. D. A. Sauter, F. Eisner, P. Ekman, S. K. Scott, Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 2408-2412 (2010).

36. P. Ekman, An argument for basic emotions. *Cogn. Emot.* **6**, 169-200 (1992).

37. J. Bollen, H. Mao, A. Pepe, “Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena,” in *Fifth International AAAI Conference on Weblogs and Social Media*. (AAAI, 2011).
38. R. Fan, J. Zhao, Y. Chen, K. Xu, Anger is more influential than joy: sentiment correlation in Weibo. *PLOS ONE* **9**, 1-8 (2014).

39. B. Suh, L. Hong, P. Pirolli, E. H. Chi, “Want to be retweeted? large scale analytics on factors impacting retweet in twitter network,” in *2010 IEEE Second International Conference on Social Computing*. (IEEE, 2010), pp. 177–184.

40. T. A. Ito, J. T. Larsen, N. K. Smith, J. T. Cacioppo, Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *J. Pers. Soc. Psychol.* **75**, 887-900 (1998).

41. R. F. Baumeister, E. Bratslavsky, C. Finkenauer, K. D. Vohs, Bad is stronger than good. *Rev. Gen. Psychol.* **5**, 323-370 (2001).

42. S. Sudhir, A. B. Unnithan, Measuring consumer motivations to share rumors: scale development. *Int. J. Online Mark.* **4**, 51-67 (2014).

43. R. Fan, K. Xu, J. Zhao, [http://arxiv.org/abs/2005.01924](http://arxiv.org/abs/2005.01924) (2020).

44. “Fact-checking fake news on Facebook works - just too slowly,” *Physorg*, 16 October 2017; [https://phys.org/news/2017-10-fact-checking-fake-news-facebook-.html#jCp](https://phys.org/news/2017-10-fact-checking-fake-news-facebook-.html#jCp).

45. K. Shu, A. Sliva, S. Wang, J. Tang, H. Liu, Fake news detection on social media: a data mining perspective. *ACM SIGKDD Explor. Newsl.* **19**, 22-36 (2017).

46. J. G. Bullock, A. S. Gerber, S. J. Hill, G. A. Huber, Partisan bias in factual beliefs about politics. *Q. J. Political Sci.* **10**, 519-578 (2015).

47. B. Shi, K. Xu, J. Zhao. [https://arxiv.org/abs/2004.05591](https://arxiv.org/abs/2004.05591) (2020).
48. S. Goel, A. Anderson, J. Hofman, D. J. Watts, The structural virality of online diffusion. Manage. Sci. 62, 180-196 (2016).

49. V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. J. Stat. Mech.: Theory Exp. 2008, P10008 (2008).

50. R. Lambiotte, J. Delvenne, M. Barahona, Random walks, markov processes and the multiscale modular organization of complex networks. IEEE Trans. Netw. Sci. Eng. 1, 76-90 (2014).

51. O. Luminet IV, P. Bouts, F. Delie, A. S. Manstead, B. Rim, Social sharing of emotion following exposure to a negatively valenced situation. Cogn. Emot. 14, 661-688 (2000).

52. J. A. Suykens, J. Vandewalle, Least squares support vector machine classifiers. Neural Process. Lett. 9, 293-300 (1999).

53. J. Ramos, “Using tf-idf to determine word relevance in document queries,” in Proceedings of the First Instructional Conference on Machine Learning (Piscataway, NJ, 2003), pp. 133–142.

54. I. Guyon, A. Elisseeff, An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157-1182 (2003).

55. J. Zhao, L. Dong, J. Wu, K. Xu, “Moodlens: an emoticon-based sentiment analysis system for chinese tweets,” in Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM, 2012), pp. 1528–1531.

56. B. Shi, J. Zhao, K. Xu, “A word2vec model for sentiment analysis of Weibo,” in 2019 16th International Conference on Service Systems and Service Management (ICSSSM) (IEEE, 2019), pp. 1–6.
57. R. Fan, J. Zhao, K. Xu, Topic dynamics in Weibo: a comprehensive study. *Soc. Netw. Anal. Min.* **5**, 41 (2015).

58. R. P. Bagozzi, N. Wong, Y. Yi, The role of culture and gender in the relationship between positive and negative affect. *Cogn. Emot.* **13**, 641-672 (1999).

59. S. M. Mohammad, P. D. Turney, “Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon,” in *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text* (ACL, 2010), p. 2634.

60. S. M. Mohammad, P. D. Turney, Crowdsourcing a wordemotion association lexicon. *Comput. Intell.* **29**, 436-465 (2013).
Acknowledgments

Funding: This work was supported by NSFC (Grant No. 71871006). Author contributions: YC conducted the analysis and wrote the manuscript. JZ conceived of the study, conducted the analysis, wrote the manuscript and oversaw the work. Competing interests: The authors declare no competing interests. Data and materials availability: All data and code used in this study are publicly available through the permanent link https://doi.org/10.6084/m9.figshare.12163569.v2.
Supplementary materials

Supplementary Text

Figs. S1 to S25

Tables S1 to S26
Fig. 1. Complementary cumulative distribution functions (CCDFs) of emotions. (A and B) The proportion of anger. The proportion of anger greater than 0.5 in LHF news is nearly 3 times that in HLT news (A). (C and D) The proportion of disgust. (E and F) The proportion of joy. The proportion of joy in HLT news is more than 2 times that of LHF news (E). (G and H) The proportion of sadness. (I and J) The proportion of fear. The results of K-S tests are shown in SM S5, and consistent results from other methods can be seen in SM S7.
Fig. 2. The CCDFs of motivations. (A and E) Anxiety management (M1-avg). (B and F) Information sharing (M2-avg). (C and G) Relationship management (M3-avg). (D and F) Self-enhancement (M4-avg). (A to D) The CCDFs of four motivations in HLT news, HLF news and LHF news. (E to H) The CCDFs of four motivations in anger-dominated news and joy-dominated news. The results of the K-S tests can be seen in SM S13.
| Variables   | Fake (1) | Retweet (2) | Retweet (3) | Retweet (4) |
|-------------|----------|-------------|-------------|-------------|
| Anger       | 0.889*** | 0.385       | 23.959***   | 22.278      |
|             | (0.097)  | (0.077)     | (6.752)     | (5.628)     |
| Joy         | -1.503***| -1.279**    | -29.555***  | -38.978***  |
|             | (0.074)  | (0.055)     | (5.452)     | (3.936)     |
| Other Emotions | (omitted) | (omitted)   | (omitted)   | (omitted)   |
| Follower    | -6.10e-08** | -3.14e-07*** | 0.00002***  | 0.00001***  |
|             | (1.04e-08) | (1.70e-08)  | (3.39e-06)  | (7.08e-07)  |
| Friend      | 0.001***  | -3.57e-06   | 0.048**     | 0.040**     |
|             | (0.00004) | (0.00003)   | (0.003)     | (0.003)     |
| Mention     | 0.104***  | -0.201***   | 23.998***   | 17.067***   |
|             | (0.067)   | (0.050)     | (4.294)     | (3.521)     |
| Hashtag     | -1.264*** | -1.631***   | 2.851       | -3.350      |
|             | (0.072)   | (0.052)     | (6.268)     | (4.018)     |
| Location    | -0.066    | -0.198***   | -5.034      | -4.438*     |
|             | (0.069)   | (0.048)     | (3.011)     | (2.572)     |
| Date        | -0.542*** | -1.217***   | 14.641***   | 0.424       |
|             | (0.056)   | (0.040)     | (4.270)     | (2.982)     |
| URL         | -2.205*** | -1.592***   | -20.438***  | -24.866**   |
|             | (0.062)   | (0.040)     | (2.664)     | (2.263)     |
| Length      | -0.005*** | 0.009**     | -0.281***   | -0.197*     |
|             | (0.0007)  | (0.0005)    | (0.054)     | (0.036)     |
| Emergency   | 5.576***  | 4.915**     | 33.522***   | 23.012**    |
|             | (0.722)   | (0.585)     | (7.911)     | (6.545)     |
| Finance     | -0.361*** | 0.153**     | -18.488**   | -19.635***  |
|             | (0.093)   | (0.062)     | (8.130)     | (5.065)     |
| International| -0.379   | -0.547***   | 53.856      | 12.386      |
|             | (0.153)   | (0.118)     | (22.359)    | (12.479)    |
| Military    | 0.928***  | 0.879***    | 11.864      | 13.618      |
|             | (0.154)   | (0.122)     | (14.884)    | (11.159)    |
| Society     | 0.942***  | 1.513***    | -21.502***  | -15.074***  |
|             | (0.071)   | (0.053)     | (6.915)     | (4.401)     |
| Sports      | -0.742*** | -1.393***   | 110.648***  | 63.475***   |
|             | (0.137)   | (0.110)     | (29.290)    | (11.564)    |
| Technology  | 0.253**   | -0.143*     | -1.712      | -6.322***   |
|             | (0.104)   | (0.080)     | (11.131)    | (6.508)     |
| Cons        | 0.205**   | 1.470***    | 81.871***   | 73.831***   |
|             | (0.098)   | (0.077)     | (10.733)    | (6.806)     |
| N           | 10486     | 26831       | 20323       | 26831       |

Table 1. Results of logit and linear models in different groups. (1) The results of the logit model in LT news and HF news. (2) The results of the logit model in all true news and fake news. (3) The results of the linear model in LF news and HF news. (4) The results of the linear model in lowly retweeted (L) news and highly retweeted (H) news (see SM S9 for more details). The values in brackets are the standard errors. *P < 0.1, **P < 0.05, ***P < 0.01.
Supplementary Materials for
Anger makes fake news viral online

Yuwei Chuai¹, Jichang Zhao¹,²,∗
¹School of Economics and Management, Beihang University, China
²Beijing Advanced Innovation Center for Big Data and Brain Computing, China

∗Correspondence to: jichang@buaa.edu.cn

This PDF file includes:
Supplementary Text
Figs. S1 to S25
Tables S1 to S26
S1 Fake news and Real news

The fake news and real news in this study were collected from Weibo, the most popular Twitter-like service in China, which had 200 million daily active users and generated over 100 million daily tweets (news) at the end of 2018. Here news refers to tweets including news-related content on Weibo. The users of Weibo are dominated by young people, and those aged between 18 and 30 years old account for 75% of all users. There is also a distinctive verification mechanism in Weibo that ensures the reliability of the user demographics. Specifically, all users have to provide their IDs during registration because of the real-name certification regulation in China. Besides, influential users, including elites with a certain reputation and influence in specific domains, well-known enterprises and their executives, the mainstream media, and government agencies such as public authorities, can be further manually verified through documentary evidence (47). Weibo even presents red or blue badges on their online profiles. Weibo officially organizes a committee composed of professional fact-checkers outside Weibo to tag fake news authoritatively and publicly.

Through the open API of Weibo, we collected fake news rated and exposed by the official committee. Considering that fake news always draws attention from the committee after being widely disseminated, the digital traces of the spread of such news on Weibo can be completely traversed. Further probes on the timelines of all news items confirm this fact in S3. Real news, also termed true news in this study, refers to information that was not tagged as false by the committee and was posted by verified users, such as mainstream media, elites, or public authorities, with credibility. In total, we collected 22,479 fake news items (with 1,189,186 users) and 10,000 real news items (with 409,865 users) from 2011 to 2016. For each news item on Weibo, we also collected its attributes, namely, text, posting time, author profile (number of followers, number of reciprocal followers, etc.), retweets, and reposting time. A subset of

1https://data.weibo.com/report/reportDetail?id=433
the fake news and real news used in this study was employed in the previous study (14) on the structural uniqueness of fake news, in which equivalent results are derived from both Weibo and Twitter, implying the reliability and universality of our data. Also, authentic tweets from credible nonverified authors of Weibo further testified the representativeness of our real news data (14).

We have made the data publicly available at https://doi.org/10.6084/m9.figshare.12163569.v2.


S2 News groups

S2.1 Partition strategy

The number of followers intuitively represents the influence of users on social media, i.e., more followers mean the news will be broadcast to a larger audience and accordingly result in more retweets. Additionally, the number of retweets can represent the spreading capability of a given news item. Fake news might be widely retweeted because of the influence of its author; however, the broadcasting potential of authors does not sufficiently explain the fast spread of fake news (2), e.g., fake news posted by lowly followed authors might be massively retweeted. To examine the causal impact of emotions on the circulation of fake news, treatment groups and control groups are established to control for variables and infer the significant roles of emotions underlying the spread. Considering that the role of emotions in information spreading might be subtle and easily interfered with by other variables, such as the influence of authors, we aim to split news, either fake or real, into a treatment group (e.g., highly retweeted news posted by authors with a low volume of followers) and a control group (e.g., lowly retweeted news posted by authors with a high volume of followers), through which the possible influence of authors can be controlled and the effects of emotions are amplified. Intuitively, for highly retweeted news posted by authors with a low volume of followers, promotion from the content, in particular, the emotions carried, would be more powerful and thus easier to detect. Therefore, we group the news according to the number of its authors followers \(x\) and the number of retweets \(y\) (16). For example, based on real news with a high number of followers but a low number of retweets and fake news with a low number of followers but a high number of retweets, a division model of maximizing the difference between true and fake news is defined to determine the splitting interface, which is specified as

\[
D = \left( \frac{N_{\text{HFF}}}{N_F} - \frac{N_{\text{HT}}}{N_T} \right) + \left( \frac{N_{\text{HLF}}}{N_F} - \frac{N_{\text{H}}}{N_T} \right) - \left| \frac{N_{\text{LFF}}}{N_F} - \frac{N_{\text{LT}}}{N_T} \right| - \left| \frac{N_{\text{HF}}}{N_F} - \frac{N_{\text{HT}}}{N_T} \right|
\]
where

- \( N_T \) is the number of true (T) news items.
- \( N_F \) is the number of fake (F) news items.
- \( N_{LLT} \) is the number of true news items with a low number of followers \((< x)\) and a low number of retweets \((< y)\).
- \( N_{LHT} \) is the number of true news items with a low number of followers \((< x)\) and a high number of retweets \((\geq y)\).
- \( N_{HHT} \) is the number of true news items with a high number of followers \((\geq x)\) and a high number of retweets \((\geq y)\).
- \( N_{HLT} \) is the number of true news items with a high number of followers \((\geq x)\) and a low number of retweets \((< y)\).
- \( N_{LLF} \) is the number of fake news items with a low number of followers \((< x)\) and a low number of retweets \((< y)\).
- \( N_{LHF} \) is the number of fake news items with a low number of followers \((< x)\) and a high number of retweets \((\geq y)\).
- \( N_{HHF} \) is the number of fake news items with a high number of followers \((\geq x)\) and a high number of retweets \((\geq y)\).
- \( N_{HLF} \) is the number of fake news items with a high number of followers \((\geq x)\) and a low number of retweets \((< y)\).
We let the number of followers (from 10 to $10^4$) and the number of retweets (from 10 to $10^8$) grow exponentially with a step size of 1 to maximize the value of $D$ and find the optimal partition line. As shown in Fig. S1, the best tuple is $(x^*, y^*) = (10, 1000)$.

According to the tuple $(10, 1000)$, we divide the news into low volume of followers and lowly retweeted true (LLT) news, low volume of followers and highly retweeted true (LHT) news, high volume of followers and highly retweeted true (HHT) news, high volume of followers and lowly retweeted true (HLT) news, low volume of followers and lowly retweeted fake (LLF) news, low volume of followers and highly retweeted fake (LHF) news, high volume of followers and highly retweeted fake (HHF) news and high volume of followers and lowly retweeted fake (HLF) news (Fig. S2). Lowly retweeted true (LT) news includes LLT news and HLT news, highly retweeted true (HT) news includes LHT news and HHT news, lowly retweeted fake (LF) news includes LLF news and HLF news and highly retweeted fake (HF) news includes LHF news and HHF news. Additionally, ignoring the label of fake or true, lowly retweeted news is categorized as L news, and highly retweeted news is categorized as H news.

By pairing various groups, diverse assemblies of treatments and controls can be established to examine the causal impact of emotions on circulation. Specifically, HLT news accounts for the largest proportion of true news, and LLF news accounts for the largest proportion of fake news (Table S1).

| True (T) news | Fake (F) news |
|---------------|---------------|
| LT news       | LF news       |
| HLT news      | HLF news      |
| LHT news      | HLH news      |
| HHT news      | HHF news      |
| LLT           | LLF           |
| (388)         | (12805)       |
| (3.88%)       | (56.96%)      |
| HLT           | HLF           |
| (7867)        | (3513)        |
| (78.67%)      | (15.63%)      |
| LHT           | LHF           |
| (36)          | (1397)        |
| (0.36%)       | (6.21%)       |
| HHT           | HHF           |
| (1709)        | (4764)        |
| (17.09%)      | (21.19%)      |

Table S1: Numbers and proportions of all groups of both fake and real news.
Fig. S1: The difference ($D$) varies with the tuple $(x, y)$, where $x = 10^i$ ($i = 1, 2, 3, 4$) and $y = 10^j$ ($j = 1, 2, \cdots, 8$).

Fig. S2: (A) Scatter plot of true news. (B) Scatter plot of fake news.
S2.2 Information dominance

To verify the rationality of the partition strategy in S2.1, we first examine the information dominance between different author groups. Here information dominance measures to which extent the authors of news items could dominate the spread in other spreader groups. According to their numbers of followers \((x)\), all users are divided into eight groups, including \(G_0\) (users whose follower counts fall in the interval \([0, 10)\)), \(G_1\) (users whose follower counts fall in the interval \([10, 10^2)\)), \(G_2\) (users whose follower counts fall in the interval \([10^2, 10^3)\)), \(G_3\) (users whose follower counts fall in the interval \([10^3, 10^4)\)), \(G_4\) (users whose follower counts fall in the interval \([10^4, 10^5)\)), \(G_5\) (users whose follower counts fall in the interval \([10^5, 10^6)\)), \(G_6\) (users whose follower counts fall in the interval \([10^6, 10^7)\)), and \(G_7\) (users whose follower counts fall in the interval \([10^7, \infty)\)). The information transmitted from the news item \(m\) in \(G_i\) (if the author of \(m\) belongs to \(G_i\), \(m\) is accordingly split to \(G_i\)) to \(G_j\) is defined as

\[
T_{i,m,j} = \frac{N_{i,m,j}}{\sum_{g=1}^{G} N_{i,m,g}},
\]

where \(N_{i,m,j}\) is the number of spreaders belonging to \(G_j\) in the retweets of \(m\) in \(G_i\) and \(G\) is the number of groups. Meanwhile, the coverage of \(m\) to \(G_j\) is defined as

\[
C_{i,m,j} = \frac{N_{i,m,j}}{N_j},
\]

where \(N_j\) is the number of users belonging to \(G_j\). According to \(T_{i,m,j}\) and \(C_{i,m,j}\), the transmission coverage of \(G_i\) to \(G_j\) is defined as

\[
TC_{i,j} = \frac{1}{M_i} \sum_{m=1}^{M_i} T_{i,m,j} C_{i,m,j},
\]

where \(M_i\) is the number of news items in \(G_i\). Then, the information dominance of \(G_i\) to \(G_j\) is
ID\left(G_i, G_j\right) = \frac{TC_{i,j} - TC_{j,i}}{TC_{i,j} + TC_{j,i}}.

When the information dominance of $G_i$ ($G_{out}$) to $G_j$ ($G_{in}$) is positive, i.e., $ID\left(G_i, G_j\right) > 0$, it is defined that $G_i$ has more information influence as compared to $G_j$. As shown in Fig. S3, since $G_2$, the information dominance of $G_{out}$ to $G_{in}$ is constantly larger than 0.5, implying authors with numbers of followers higher than $10^3$ indeed possess more information influence. Hence, it is reasonable to divide L users (with low influence) and H users (with high influence) by $10^3$ according to our partition strategy.
S2.3 Structural virality

The spreading capability of news may not be comprehensively represented by the number of retweets, and the diffusion structure can also reflect the very viral nature of news. Therefore, we further examine the rationality of the partition strategy according to retweeting number ($y$) in S2.1 from the perspective of circulation structure. The structural virality is the average distance between all pairs of nodes in a diffusion (48), which can measure the diversity of diffusion structure. It is defined as

$$v = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} d_{i,j},$$

where $d_{i,j}$ denotes the length of the shortest path between nodes $i$ and $j$. When $v \sim 2$, it can be thought an approximately pure broadcast (48). The average structural virality of news diffusion with the number of retweets is shown in Fig. S4. For all true and fake news, approximately 97% of the structural virality is lower than 2 when the number of retweets is less than 10, which is exactly same to the cutting point previously obtained, verifies the reliability of the division in S2.1 and again consolidates our partition strategy of news groups for treatment and control. Meanwhile, fake news is more viral (longer average path) than true news (K-S test $\sim 0.159$, $P \sim 0$) in terms of structural virality, which is consistent with previous results on Twitter (2), implying the universality of our dataset from Weibo. Six typical diffusion networks of both fake and real news are also shown in Fig. S5 to further illustrate this.
Fig. S4: The average structural virality with the growing retweets.
Fig. S5: Typical examples of diffusion networks for true and fake news items. (A) A true news diffusion network with 630 nodes and $v \sim 2$ (pure broadcast). (B) A fake news diffusion network with 600 nodes and $v \sim 2$ (pure broadcast). (A) and (B) are both pure broadcast structures. (C) A true news diffusion network with 102 nodes and $v \sim 7.142$. It talks about the descendant of Confucius and includes 9 communities (49, 50). (D) A fake news diffusion network with 207 nodes and $v \sim 9.895$. It talks about Red Cross Society of China and includes 17 communities. (E) A true news diffusion network with 800 nodes and $v \sim 5.763$. It talks about a North Korean diplomat who joined South Korea and includes 21 communities. (F) A fake news diffusion network with 997 nodes and $v \sim 7.748$. It talks about some people using babies to make soup and includes 63 communities. Different colors represent different communities in the spread.
S3 News timelines

As mentioned in S1, both fake news and real news were collected before 2017 (our commercial access to Weibo API expired in 2017), and the news in our data set was posted from 2011 to 2016 (Fig. S6). The lifecycle of a news item starts from the posting time and ends with the final retweet in the sampling period. The timeline of each true or fake news item is analyzed by calculating the proportion of the number of new retweets within each hour of its lifecycle. For both true and fake news, retweets reach their peak within one hour after posting (Fig. S7A and S7B), illustrating the quick circulation on social media and, in particular, the explosive spread in the very early stage. Furthermore, we count the number of retweets every ten minutes and calculate the cumulative distribution functions (CCDFs) for different types of news. Fake (F) news demonstrates stronger vitality than true (T) news (K-S test $\sim 0.140$, $P \sim 0.0$) (Fig. S7C). Specifically, fake news still obtains 26% of its retweets after 48 hours, while that proportion for true news is 20%. More importantly, the stronger vitality of fake news than true news is consistently observed in groups of LT news vs. LF news (K-S test $\sim 0.114$, $P \sim 0.0$) (Fig. S7D) and HT news vs. HF news (K-S test $\sim 0.138$, $P \sim 0.0$) (Fig. S7E). Besides, we compared the distributions of the number of retweets within 48 hours of the posting and found that the propagation speed of fake news is significantly higher than that of true news (K-S test $\sim 0.195$, $P \sim 0.0$) (Fig. S7F). All this evidence suggests findings similar to those for Twitter (2), that is, fake news is more viral than real news online. Compared to that of real news, its circulation lasts longer, has higher speed, and ultimately produces more retweets.
Fig. S6: Yearly counts of news.

Fig. S7: (A) The proportion of new retweets in each hour for both HT and LT news. (B) The proportion of new retweets in each hour for fake news. (C) CCDFs for retweeting time for true news and fake news. (D) CCDFs for retweeting time for LT news and LF news. (E) CCDFs for retweeting time for HT news and HF news. (F) CCDFs for the number of retweets within 48 hours for true news and fake news.
S4 Emotion lexicon

In this study, the emotional texts of news in social media, both fake and true, are assumed to carry sophisticated signals that cannot be fully represented by binary values such as positive or negative. In contrast, emotions, in particular, negative emotions, are split into elementary compounds, including anger, disgust, sadness, and fear (35, 51). Then, together with joy, which is used to reflect positive emotion, the distributions of the five emotions are derived to fully represent the emotional spectrum of each news item. An emotion lexicon must be established to obtain the emotional distribution of the text in both fake and true news intuitively and accurately; then, the occupation of a certain emotion can be calculated as the fraction of terms with this emotion in all emotional terms of the news text. We first segment all the texts into terms, filter by parts of speech, and keep nouns, verbs, adverbs, gerunds, adjectives, adjectives directly used as adverbials and adjectives with noun function to compose a candidate set. As a result, 34227 preselected terms are obtained. Note that there might also be terms of nonemotion in the candidate set. We then hire human coders to manually label the terms: those without emotions are marked as neutral. A WeChat applet, named Word Emotion (Fig. S8), is built to make the labeling convenient. The whole labeling task was completed by nine well-instructed coders who are active users of Weibo with ages between 18 and 30 years old, and each term is labeled three times by randomly selected coders. Finally, terms with more than two identical emotional labels are screened out to build the lexicon. Ultimately, there are 6155 emotional terms in total, including 1323 anger terms, 710 disgust terms, 2066 joy terms, 1243 sad terms, and 813 fear terms. The emotion lexicon covers 87.1% of the text of all fake and true news, and the remaining news items are labeled neutral, suggesting that the news in social media is indeed emotional.

The emotion lexicon is publicly available and can be downloaded freely at https://doi.org/10.6084/m9.figshare.12163569.v2.
Fig. S8: Main page of the WeChat applet Word Emotion. The Chinese word on the left describes a very angry state. The Chinese word on the right describes rejoice with wild excitement.
S5 Emotion distributions

The emotion distributions of news in the different groups are derived utilizing the established emotion lexicon. After the inference of emotion distributions, possible differences between treatment and control groups of news are comprehensively examined. These differences are expected to help reveal the mechanism underlying the circulation of fake news. In particular, more insights might be derived by splitting negative emotion into more elementary emotions.

In the main text, we discussed that the amount of anger in fake news is significantly higher than that in true news, and the amount of joy in true news is significantly higher than that in fake news. This phenomenon is more obvious in HLT news and LHF news after excluding the influence of the author. Moreover, to further examine the difference between anger and joy and its possible association with the fast spread of fake news, we compare the emotional differences between HLF news and LHF news. The results show that the amount of anger in LHF news is significantly higher than that in HLF news (Fig. 1A in the main text), and the amount of joy is significantly lower than that in HLF news (Fig. 1E in the main text), which is consistent with the comparison between L news and H news (Fig. S9A, S9C). That is, the amount of anger in widely circulated news is significantly higher than that in less widely circulated news. The statistics of the emotional distributions and the results of K-S tests are shown in Table S2-5. All these observations consistently suggest an association between anger and the virality of fake news and inspire later causal inference through regression models.
Fig. S9: CCDFs of emotions in L news and H news. (A) Anger, (B) Disgust, (C) Joy, (D) Sadness, (E) Fear. The results of the K-S tests can be seen in Table S5.

Table S2: Statistics and K-S tests for HLT news and LHF news.

|       | Mean     | Std      | K-S test         |
|-------|----------|----------|------------------|
|       | HLT | LHF | HLT | LHF | D   | p-value |
| Anger | 0.110781 | 0.266855 | 0.256383 | 0.343774 | ~0.275 | ~0 |
| Disgust | 0.065549 | 0.052674 | 0.196524 | 0.154399 | ~0.039 | 1.0 |
| Joy   | 0.610843 | 0.328504 | 0.42096 | 0.346331 | ~0.366 | ~0 |
| Sadness | 0.119657 | 0.157584 | 0.260941 | 0.240621 | ~0.157 | ~0 |
| Fear  | 0.09317  | 0.194382 | 0.23423 | 0.280941 | ~0.264 | ~0 |
|             | Mean          | Std           | K-S test         |
|-------------|---------------|---------------|------------------|
|             | HLF (3132)    | LHF (1238)    |                  |
| **Anger**   | 0.183563      | 0.266855      | 0.305268 0.343774 | D ~ 0.135, p-value ~ 0 |
| **Disgust** | 0.059838      | 0.052674      | 0.167523 0.154399 | D ~ 0.033, p-value ~ 0.34 |
| **Joy**     | 0.391998      | 0.328504      | 0.36497 0.346331  | D ~ 0.105, p-value ~ 0    |
| **Sadness** | 0.133024      | 0.157584      | 0.244897 0.240621 | D ~ 0.086, p-value ~ 0    |
| **Fear**    | 0.231577      | 0.194382      | 0.309433 0.280941 | D ~ 0.058, p-value ~ 0    |

Table S3: Statistics and K-S tests for HLF news and LHF news.

|             | Mean          | Std           | K-S test         |
|-------------|---------------|---------------|------------------|
|             | T (6550)      | F (20352)     |                  |
| **Anger**   | 0.112438      | 0.165286      | 0.255438 0.290279 | D ~ 0.101, p-value ~ 0   |
| **Disgust** | 0.066563      | 0.047572      | 0.197113 0.149817 | D ~ 0.031, p-value ~ 0   |
| **Joy**     | 0.609413      | 0.442912      | 0.418222 0.354057 | D ~ 0.349, p-value ~ 0   |
| **Sadness** | 0.120355      | 0.122562      | 0.258137 0.233947 | D ~ 0.045, p-value ~ 0   |
| **Fear**    | 0.09123       | 0.221667      | 0.229021 0.282974 | D ~ 0.357, p-value ~ 0   |

Table S4: Statistics and K-S tests for true news and fake news.
|       | Mean     | Std      | K-S test        |
|-------|----------|----------|-----------------|
|       | L (20066)| H (6836) | L               | H               |
| Anger | 0.122546 | 0.240105 | 0.259731        | 0.327235        |
|       |          |          | D ~ 0.210, p-value ~ 0 |
| Disgust | 0.043108 | 0.078873 | 0.148363        | 0.196844        |
|       |          |          | D ~ 0.089, p-value ~ 0 |
| Joy   | 0.524593 | 0.362686 | 0.368586        | 0.377564        |
|       |          |          | D ~ 0.249, p-value ~ 0 |
| Sadness | 0.113074 | 0.148299 | 0.234553        | 0.253735        |
|       |          |          | D ~ 0.094, p-value ~ 0 |
| Fear  | 0.196679 | 0.170037 | 0.276485        | 0.275805        |
|       |          |          | D ~ 0.108, p-value ~ 0 |

Table S5: Statistics and K-S tests for L news and H news.
S6 Keywords in separating news groups

The existence of highly retweeted tweets posted by authors with a low volume of followers in both fake news and real news implies the potential influence of content on circulation. Besides, emotions are carried by words in the text. The distinguishing distributions of emotions, in particular, anger and joy, between fake news and real news inspire us to pinpoint keywords that could split news groups. Additionally, these keywords could help in later offline questionnaires to strengthen the stimuli of anger and joy on the reposting incentives of the audience (see S13).

Specifically, for LHF news, HLT news, and HLF news, we train an SVM (52) and a logistic regression model, which are commonly employed to weigh words in text mining tasks, to evaluate the separability of the text and extract keywords that influence the separation. These groups of news are further split into two corpora to learn binary classification models, i.e., one corpus is composed of LHF news (positive class) and HLT news (negative class) and the other corpus is composed of LHF news (positive class) and HLF news (negative class). Words are used as text features to calculate the TF-IDF matrix (53) for classification. After 5-fold cross-validation, the average accuracies are 0.94 (SVM) and 0.98 (logistic regression) in the corpus of LHF-HLT and 0.75 (SVM) and 0.81 (logistic regression) in the corpus of LHF-HLF, implying that using words as features results in good separation of LHF news from HLT news and HLF news. Moreover, content carrying emotions such as anger and joy could be an influential driver of news circulation. In particular, the better separability between LHF news and HLT news suggests the feasibility of keywords in strengthening the divergence of different news items in reposting stimuli. On this basis, we combine the chi-square test, mutual information, AdaBoost, and extra-trees for feature selection (54), and 150 influential keywords with the greatest weights in the classification are selected from each group of news items (Fig. S10A, C, and E). By analyzing the emotional distributions of keywords in each type of news, we found that

---

2 These methods are implemented with the scikit-learn package in Python.
the emotional keywords in HLT news are all joyful (Fig. S10B), and those in HLF news are mainly joyful (Fig. S10F), followed by fearful. However, negative emotions, especially anger, dominate the keywords in LHF news (Fig. S10D). These observations support the initial assumption that emotions carried by news, in particular, the dominant emotions of anger and joy, can be reflected by keywords that effectively separate different groups of news; therefore, these keywords will affect the incentives underlying retweets. Meanwhile, the exact same difference in the emotion distribution at the keyword level further confirms the consistency and robustness of the emotional divergence between fake news and true news revealed at the collective level (see S5).
Fig. S10: (A) Word cloud of keywords in HLT news. (B) Emotional distribution of keywords in HLT news. (C) Word cloud of keywords in LHF news. (D) Emotional distribution of keywords in LHF news. (E) Word cloud of keywords in HLF news. (F) Emotional distribution of keywords in HLF news. All the keywords in the word cloud are translated into English and can be found in the publicly available data at https://doi.org/10.6084/m9.figshare.12163569.v2.
S7 Additional tests for emotion inference and divergence

S7.1 Alternative approaches of emotion inference

In addition to the emotion lexicon, which offers an intuitive measure of emotion occupation, machine learning models, in particular, state-of-the-art solutions such as deep neural networks, are alternative models to infer the emotion distributions of both fake and true news. In this study, to ensure the consistency and accuracy of our results on emotion distributions, we also considered classic machine learning and deep learning models. Specifically, two classifiers built for emotion detection in Chinese tweets from Weibo are employed to perform the additional tests, namely, a naïve Bayesian classifier (termed Bayes, with an accuracy of 0.64) (55) and a back-propagation neural network based on an emotional dictionary (termed BP1, with an accuracy of 0.69)\(^3\), to calculate the emotion distributions of the texts in terms of probabilities of belonging to certain emotions. Then, the occupations of different emotions are further compared across groups, and the results are shown in Table S6-11. All the results support our conclusions obtained from the emotion lexicon, in particular, the difference in emotion distributions between anger and joy, suggesting the robustness of our understanding of emotion divergence between fake news and real news.

S7.2 Alternative measure of emotion distribution

In the previous analysis and the additional test on emotion divergence, the emotion distribution of each news item is inferred exclusively by one method, i.e., lexicon-based, Bayes or BP1, and is simply represented as the occupations of the emotions in each text. However, it is possible that different methods could result in different inferences on the same text, which might undermine the consistency of emotion divergence we previously observed at the text level. To further assess the robustness of our conclusions about the different occupations of anger and joy in fake news

\(^3\)BP1 was built with Keras.
### Table S6: Statistics and K-S tests for HLT news and LHF news based on Bayes.

|       | Mean   | Std   | K-S test                       |
|-------|--------|-------|--------------------------------|
|       | HLT    | LHF   |                               |
| Anger |        |       |                                |
|       | (6797) | (1326)|                                |
|       | 0.260271 | 0.321154 | 0.125017 | 0.10956 | D ~ 0.294, p-value ~ 0 |
| Disgust | 0.208931 | 0.150072 | 0.094848 | 0.086206 | D ~ 0.355, p-value ~ 0 |
| Joy   | 0.253518 | 0.149253 | 0.137519 | 0.110704 | D ~ 0.403, p-value ~ 0 |
| Sadness | 0.216766 | 0.315336 | 0.122535 | 0.127497 | D ~ 0.367, p-value ~ 0 |
| Fear  | 0.060514 | 0.064185 | 0.126408 | 0.114596 | D ~ 0.053, p-value ~ 0.004 |

### Table S7: Statistics and K-S tests for T news and F news based on Bayes.

|       | Mean   | Std   | K-S test                       |
|-------|--------|-------|--------------------------------|
|       | T      | F     |                               |
|       | (8836) | (22065)|                               |
| Anger |        |       |                                |
|       | 0.257017 | 0.32336 | 0.124972 | 0.101632 | D ~ 0.334, p-value ~ 0 |
| Disgust | 0.206955 | 0.16109 | 0.095054 | 0.081381 | D ~ 0.315, p-value ~ 0 |
| Joy   | 0.25383 | 0.163993 | 0.13615 | 0.103545 | D ~ 0.389, p-value ~ 0 |
| Sadness | 0.222754 | 0.304678 | 0.122226 | 0.128804 | D ~ 0.368, p-value ~ 0 |
| Fear  | 0.059445 | 0.046879 | 0.125032 | 0.103895 | D ~ 0.034, p-value ~ 0 |
### Table S8: Statistics and K-S tests for L news and H news based on Bayes.

|       | Mean       | Std       | K-S test       |
|-------|------------|-----------|----------------|
|       | L (23215)  | H (7686)  | L              | H              |
| Anger | 0.302948   | 0.308743  | 0.107285       | 0.12818        | $D \sim 0.095$, p-value $\sim 0$ |
|       | 0.174278   | 0.173983  | 0.085508       | 0.095092       | $D \sim 0.058$, p-value $\sim 0$ |
| Disgust | 0.194394  | 0.175448  | 0.118335       | 0.127085       | $D \sim 0.112$, p-value $\sim 0$ |
| Sadness | 0.283035  | 0.275866  | 0.131581       | 0.134087       | $D \sim 0.111$, p-value $\sim 0$ |
| Fear | 0.045345   | 0.065959  | 0.10502        | 0.124328       | $D \sim 0.093$, p-value $\sim 0$ |

### Table S9: Statistics and K-S tests for HLT news and LHF news based on BP1.

|       | Mean       | Std       | K-S test       |
|-------|------------|-----------|----------------|
|       | HLT (2607) | LHF (893) | HLT            | LHF            |
| Anger | 0.061185   | 0.296197  | 0.154826       | 0.350258       | $D \sim 0.436$, p-value $\sim 0$ |
|        | 0.095284   | 0.104351  | 0.164124       | 0.10821        | $D \sim 0.288$, p-value $\sim 0$ |
| Joy   | 0.552737   | 0.209983  | 0.414815       | 0.301277       | $D \sim 0.399$, p-value $\sim 0$ |
| Sadness | 0.178144  | 0.226401  | 0.299973       | 0.312545       | $D \sim 0.131$, p-value $\sim 0$ |
| Fear  | 0.112649   | 0.163068  | 0.256806       | 0.266463       | $D \sim 0.216$, p-value $\sim 0$ |
|        | Mean | Std | K-S test |        | Mean | Std | K-S test |
|--------|------|-----|----------|--------|------|-----|----------|
|        | T    | F   |          |        | T    | F   |          |
|        | (3692)| (15000) |          |        | (14098)| (4594) |          |
| Anger  | 0.060797 | 0.142411 | 0.153548 | 0.262294 | D ~ 0.262, p-value ~ 0 | Anger  | 0.098229 | 0.212406 | 0.217867 | 0.303962 | D ~ 0.260, p-value ~ 0 |
| Disgust| 0.092621 | 0.079041 | 0.161604 | 0.102752 | D ~ 0.189, p-value ~ 0 | Disgust| 0.073238 | 0.107763 | 0.107646 | 0.138275 | D ~ 0.190, p-value ~ 0 |
| Joy    | 0.559246 | 0.343842 | 0.413631 | 0.339062 | D ~ 0.367, p-value ~ 0 | Joy    | 0.417254 | 0.291668 | 0.360852 | 0.36234  | D ~ 0.238, p-value ~ 0 |
| Sadness| 0.176255 | 0.214104 | 0.299774 | 0.303651 | D ~ 0.124, p-value ~ 0 | Sadness| 0.196246 | 0.238491 | 0.293552 | 0.329259 | D ~ 0.069, p-value ~ 0 |
| Fear   | 0.111082 | 0.220601 | 0.25587  | 0.292651 | D ~ 0.359, p-value ~ 0 | Fear   | 0.215033 | 0.149672 | 0.29317  | 0.27021  | D ~ 0.225, p-value ~ 0 |

Table S10: Statistics and K-S tests for T news and F news based on BP1.

Table S11: Statistics and K-S tests for L news and H news based on BP1.
and true news, a new text-level measure is presented to represent the emotion distribution by
ranks. Specifically, for each news item text, a batch of models is employed separately to infer
the probability of belonging to the five emotions, which are then ranked according to these
probabilities: lower-ranking values represent higher probabilities of the texts belonging to the
corresponding emotions. Note that emotions with the same probability are ranked randomly.
By aggregating the ranks of a certain emotion over all models, a distribution of rank can be
obtained for the emotion in each text. Then, for each group of news items, the distributions
of the five emotions can be obtained by averaging the rank distributions of the corresponding
emotions in all texts.

First, employing a word2vec (56) model that inferred over 560 million tweets of Weibo, each
term is embedded into a vector of 200 dimensions. Then, the text of a news item is converted
into a vector of 200 dimensions by averaging the embeddings of all terms in the text. To increase
the number of inference models of emotions, six additional emotion classifiers are constructed
on the emotion lexicon: including AdaBoost, decision tree, logistic regression, ridge classifier,
SVM, and backpropagation neural network (BP2). 4 Specifically, terms with emotional labels in
the emotion lexicon are first embedded to train these models; then, the emotions of news item
text in the same embedding space are inferred. The accuracies of these models in 5-fold cross-
validations are 0.67, 0.73, 0.79, 0.76, 0.75 and 0.86. From the results of the rank distributions,
ranks of anger in LHF news, F news and H news are significantly lower than those in HLT
news, T news, and L news (Fig. S11A, B, C), while the ranks of joy show the opposite trends
(Fig. S11G, H, I). Note that a lower rank represents a higher probability of belonging to the
corresponding emotion. This result is consistent with all previous results, indicating that the
divergence in anger and joy between fake news and real news is robust and independent of
emotion inference model and emotion distribution measure. However, the differences in other

4The classic machine learning models are built with scikit-learn and BP2 is built with PyTorch.
negative emotions across news groups, though significant, are inconsistent and varying. The ranks of sadness in LHF news, F news, and H news are significantly higher than those in HLT news, T news, and L news (Fig. S11J, K, L), which is inconsistent with the previous results (see Fig. 1 in the main text). The ranks of disgust fluctuate inconsistently across different assemblies of news groups. Although the rank of fear in LHF news is significantly lower than that in HLT news, as the rank is smaller than 4, it becomes higher than that of HLT, as the rank is 5. (Fig. S11M). Therefore, in the following causal inference on the impact of emotions on circulation, negative emotions other than anger are not considered separately.

### S7.3 A case study of fake news in COVID-19

Emergent events in particular those disastrous ones always spur fake news items and social media can be fertile ground for their fast spread. With the sudden outbreak of COVID-19, the epidemic-related fake news floods the Internet, disseminating false information and resulting in collective panic. Here, we further collect 324 fake news (including 31,284 retweets) related to the epidemic from January 22 to March 1, 2020,\(^5\) and examine the divergence between anger

\(^5\)https://covid19.thunlp.org/archives/5/
Fig. S11: CCDFs of emotional ranks in HLT news and LHF news, T news and F news, L news and H news. (A, B, C) Anger, (D, E, F) Disgust, (G, H, I) Joy, (J, K, L) Sadness, (M, N, O) Fear. The results of the K-S tests are shown in Table S12-14.
|       | Mean  | Std   |       | K-S test                         |
|-------|-------|-------|-------|----------------------------------|
|       | T     | F     | T     | F                               |
| Anger | 3.442262 | 2.82039 | 0.685257 | 0.746122 | D ~ 0.429, p-value ~ 0 |
| Disgust | 3.883473 | 4.087925 | 0.547386 | 0.630079 | D ~ 0.260, p-value ~ 0 |
| Joy   | 2.567956 | 3.296035 | 1.12176 | 0.87241 | D ~ 0.442, p-value ~ 0 |
| Sadness | 2.765095 | 2.893998 | 0.520543 | 0.579022 | D ~ 0.142, p-value ~ 0 |
| Fear  | 3.747426 | 3.357023 | 0.911518 | 0.874507 | D ~ 0.285, p-value ~ 0 |

Table S13: Statistics and K-S tests for the rank distributions of T news and F news.

|       | Mean  | Std   |       | K-S test                         |
|-------|-------|-------|-------|----------------------------------|
|       | L     | H     | L     | H                               |
| Anger | 3.00335 | 2.747687 | 0.724269 | 0.883641 | D ~ 0.181, p-value ~ 0 |
| Disgust | 4.124329 | 3.818081 | 0.593661 | 0.642307 | D ~ 0.246, p-value ~ 0 |
| Joy   | 3.072798 | 3.408962 | 0.908621 | 1.095298 | D ~ 0.237, p-value ~ 0 |
| Sadness | 2.853962 | 2.915554 | 0.542974 | 0.644931 | D ~ 0.122, p-value ~ 0 |
| Fear  | 3.396054 | 3.54197 | 0.858397 | 0.990407 | D ~ 0.157, p-value ~ 0 |

Table S14: Statistics and K-S tests for the rank distributions of L news and H news.
Table S15: Statistics for LF news and HF news in COVID-19.

|       | Anger       | Disgust     | Joy         | Sadness     | Fear        | Anger-Joy   |
|-------|-------------|-------------|-------------|-------------|-------------|-------------|
| LF (146) Mean | 0.108992   | 0.050385    | 0.370184    | 0.126049    | 0.344389    | -0.261192   |
|       | Std         | 0.25049     | 0.176962    | 0.398416    | 0.25443     | 0.402195    | 0.519583    |
| HF (54) Mean  | 0.145137   | 0.154795    | 0.268133    | 0.151615    | 0.28032     | -0.122996   |
|       | Std         | 0.306189    | 0.310037    | 0.330733    | 0.258752    | 0.342644    | 0.500691    |

and joy in their emotional distributions to testify our findings in the circumstance of specific emergence events. Using the emotion lexicon built in this paper, the emotional distributions of 200 fake news items are inferred. It is consistently found that HF news carries more anger and less joy than LF news. The dominance of anger to joy (the occupation of anger minus that of joy) is significantly larger in the group of HF news (\( T \sim 2.851, P \sim 0.006 \)) (Fig. S12 and Table S15). However, it should be noted that here we only support a case study on the fake news that caused by specific events like COVID-19. Due to the very small size of samples (e.g., 200) and no control groups of real news, further explorations like regression models (see S9) will not be performed on this data set.
Fig. S12: Probability density function (PDF) of Anger-Joy in LF news and HF news.
S8 Control variables

Carrying more anger but less joy is significantly associated with the fast spread of fake news. To further examine the causal impact of anger and joy on the circulation of news online, variables that might be correlated with the spread should be comprehensively considered and controlled. In addition to emotions inferred from texts, other factors such as content (39), user profiles (2), and external shocks such as disaster events (9) that could be obtained from the content are considered and controlled. Note that considering the fast spread of fake news (see S3) and, in particular, that most people do not critically question its credibility (1), only variables that can be derived at the very beginning of the posting are considered, while those related to spreading structures that are usually employed in the detection of fake news (14) are not considered due to the ex post facto inference. In addition to variables derived from content at the source, we introduce the number of followers and the number of friends, i.e., reciprocal followers, in Weibo as control variables to further consider the possible impact from user profiles. Notably, the ages of the authors are missing from the user profiles returned by Weibos open API. However, evidence from previous efforts of the impact of age on spread is inconsistent (2, 21). In the meantime, according to the annual report⁶, most Weibo user ages are concentrated in a narrow range between 18 and 30 years old, so the impact of age could be trivial because of context dependence. Also, according to recent results in (21), ages of users are associated with topics of the content, e.g., the one with ages more than 60 are more likely to post/repost tweets in politics, hence here in our model, the factor of user age could be indirectly controlled through topics that comprehensively considered. Thus, age can be omitted without significant disturbance to the results.

In total, the following variables will be derived and controlled:

⁶https://data.weibo.com/report/index
• Mention: Whether the text contains @.

• Hashtag: Whether the text contains a hashtag.

• Location: Whether the text contains location information.

• Date: Whether the text contains date information.

• URL: Whether the text contains a URL.

• Length: The length of the text.

• Emergency: Whether the text content is related to a disaster event. The emergency event in this study refers to the explosion accident in the Tianjin Binhai New Area on August 12, 2015, which occurred within the sampling period.

• Topic: The topic discussed in the text.

• Follower: The number of followers of the author.

• Friend: The number of friends of the author.

S8.1 Analysis of binary factors

Table S16 shows the statistics of binary factors including mention, hashtag, location, date, URL, and emergency. From the perspective of the proportions of all binary factors, mention, and emergency have high proportions in LHF news, followed by H news, suggesting that both promote the spread of fake news. Hashtag, date, and URL have higher proportions in true news than in fake news, implying that they contribute little to the spread of fake news. Meanwhile, although the proportion of location is relatively high in fake news, it is concentrated mainly in L news, so its impact on spread might be trivial. These preliminary analyses offer directions for examining the causal impact of these factors on the spread of news.
|                  | HLT | HLF | LHF | T   | F   | L   | H   | All |
|------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| **Mention**      |     |     |     |     |     |     |     |     |
| Yes              | 941 | 510 | 293 | 1388| 3346| 3276| 1458| 4734|
| No               | 5856| 2942| 1033| 7448| 18719| 19939| 6228| 26167|
| **P (%)**        | **13.84** | **14.77** | **22.10** | **15.71** | **15.16** | **14.11** | **18.97** | **15.32** |
| **Hashtag**      |     |     |     |     |     |     |     |     |
| Yes              | 1675| 353 | 252 | 2369| 1833| 2827| 1375| 4202|
| No               | 5122| 3099| 1074| 6467| 20232| 20388| 6311| 26699|
| **P (%)**        | **24.64** | **10.23** | **19.00** | **26.81** | **8.31** | **12.18** | **17.89** | **13.60** |
| **Location**     |     |     |     |     |     |     |     |     |
| Yes              | 1249| 794 | 270 | 1524| 4614| 4917| 1221| 6138|
| No               | 5548| 2658| 1056| 7312| 17451| 18298| 6465| 24763|
| **P (%)**        | **18.38** | **23.00** | **20.36** | **17.25** | **20.91** | **21.18** | **15.89** | **19.86** |
| **Date**         |     |     |     |     |     |     |     |     |
| Yes              | 3670| 1085| 504 | 4661| 5791| 7215| 3237| 10452|
| No               | 3127| 2367| 822 | 4175| 16274| 16000| 4449| 20449|
| **P (%)**        | **53.99** | **31.43** | **38.01** | **52.75** | **26.25** | **31.08** | **42.12** | **33.82** |
| **URL**          |     |     |     |     |     |     |     |     |
| Yes              | 3744| 853 | 212 | 4693| 5364| 8353| 1704| 10057|
| No               | 3053| 2599| 1114| 4143| 16701| 14862| 5982| 20844|
| **P (%)**        | **55.08** | **24.71** | **15.99** | **53.11** | **24.31** | **35.98** | **22.17** | **32.55** |
| **Emergency**    |     |     |     |     |     |     |     |     |
| Yes              | 2   | 82  | 180 | 3   | 663 | 404 | 262 | 666 |
| No               | 6795| 3370| 1146| 8833| 21402| 22811| 7424| 30235|
| **P (%)**        | **0.03** | **2.38** | **13.57** | **0.03** | **3.00** | **1.74** | **3.41** | **2.16** |

Total 6797 3452 1326 8836 22065 23215 7686 30901

Table S16: Statistics of binary factors.
S8.2 Analysis of Length

We calculated the length distribution of the text as the number of characters and letters. The length of LHF news has a more concentrated distributed than that of HLT news (K-S test $\sim 0.145$, $P \sim 0$) (Fig. S13A), and the difference is also significant in fake news and true news (K-S test $\sim 0.134$, $P \sim 0$) (Fig. S13B). Therefore, fake news may be more deliberate and planned in terms of linguistic organization, while real news is more casually narrated. However, the text length is more concentrated in HLF news (compared with LHF news, K-S test $\sim 0.073$, $P \sim 0$) (Fig. S13A) and L news (compared with H news, K-S test $\sim 0.095$, $P \sim 0$) (Fig. S13C), indicating that this factor might have little effect on promoting the spread of false news.

S8.3 Analysis of Topic

The topics discussed in the news are also important features of the text. We used a naïve Bayesian topic classifier (57) to analyze the topic distributions of different types of news. The classifier was trained on more than 410,000 Weibo tweets, which were grouped into seven
categories that fit the news taxonomy of Weibo: entertainment, finance, international, military, society, sports, and technology. The accuracy and F-measure are greater than 0.84, indicating good performance in topic classification. Besides, incremental training in this classifier can help solve the problem of new words. News that cannot be classified into the above seven categories is omitted in the analysis. As shown in Fig. S14, significant differences are observed in the distribution of topics among different groups of news. Specifically, the topic of society accounts for the largest proportion in HLF news, LHF news, and F news, suggesting that fake news focuses on social issues that are closely related to people's daily lives. Hot social topics would make fake news more likely to spread but do not necessarily make fake news widely spread because H news's proportion of society topic is lower than that of L news.

Through the analysis of the above eight variables derived from content, the differences between true and fake news are examined, but many do not promote the spread of fake news. Two factors, mention and emergency, may play promoting roles in the spread of fake news; however, they only occupy small proportions of all news items, which might undermine their effect on fast circulation.

**S8.4 Analysis of variables from authors**

We also examine the variables from the author profiles. Interestingly, whether true or fake, news with more retweets was posted by authors with more followers (Fig. S15) and friends (Fig. S16). However, the greater numbers of followers and friends associated with true news (as compared to fake news, and is consistent with the finding in Twitter (2)) suggest that these factors might not be the key factors making fake news more viral than true news online. By controlling all these variables, we establish both logit and linear models to examine the causal impact of anger and joy on the spread of fake news.
Fig. S14: Topic distributions of different groups of news.

Fig. S15: The boxplots of followers in true (LT and HT) news and fake (LF and HF) news.
Fig. S16: The boxplots of friends in true (LT and HT) news and fake (LF and HF) news.
S9 Logit and linear regression models

Logit and linear regression models are established to causally examine the impact of anger and joy on the spread of fake news. Note that for emotion variables, we focus primarily on anger and joy and combine the other emotions into other emotions. Note that there is a linear relationship between emotion-related variables because the ratios of the five emotions sum to 1. All the control variables from content, user profiles, and the external shock, as presented above, are comprehensively introduced into both models. The logit model is defined as

\[ \logit(p_{\text{fake}}) = \beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + \beta_4 v_4 + \beta_5 v_5 + \beta_6 v_6 + \beta_7 v_7 + \beta_8 v_8 + \beta_9 v_9 + \beta_{10} v_{10} + \beta_{11} v_{11} + \beta_{12} v_{12} + \beta_{13} v_{13}, \]

where

- \( p_{\text{fake}} \) is the probability of fake news.
- \( \beta_0 \) is the intercept.
- \( \beta_1, \beta_2, ..., \beta_{13} \) are the coefficients of variables.
- \( v_1, v_2, ..., v_{13} \) represent anger, joy, other emotions, follower, friend, mention, hashtag, location, date, URL, length, emergency, and topic.
- Mention, hashtag, location, date, URL, emergency, and topic are virtual variables.

Emotion variables derived from emotion distributions in the logit model are calculated for all methods, namely, emotion lexicon, Bayes, and BP1. The results of the model based on the emotion lexicon are shown in Table 1 of the main text. We hereby supplement the estimation results for the remaining two methods (Table S17). In all the results, the coefficients of anger are uniformly and significantly positive after controlling for all other variables, indicating that anger is causally associated with fake news, particularly news that is highly retweeted. By contrast, the
coefficients of joy are significantly negative in all results, especially for HF news and H news, indicating its prevention on the spread or news, particularly fake news. The coefficients of emergency and military and the topic of society are significantly positive, while the coefficients of mention are positive but nonsignificant (Table 1 in the main text and Table S17), which is consistent with our analysis in S8.

Then, a linear regression model is established to further qualify the influence of anger and joy on the spread of fake news. The model is defined as

\[
\text{reg (Num}_{\text{retweet}}) = \beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + \beta_4 v_4 + \beta_5 v_5 + \beta_6 v_6 + \beta_7 v_7 + \beta_8 v_8 + \beta_9 v_9 + \beta_{10} v_{10} + \beta_{11} v_{11} + \beta_{12} v_{12} + \beta_{13} v_{13},
\]

where

- The dependent variable \(Num_{\text{retweet}}\) is the number of retweets within 48 hours of news release. Note that over 70% of retweets of fake news and 80% of retweets of real news occurred within 48 hours after posting (see S3). Other settings, e.g., longer than 48 hours, do not influence the results.

- The independent variables are consistent with the explanatory variables of the logit model.

We first estimate the linear model on fake news and then for all news, neglecting the labels of true or fake; the results can be found in Table 1 (3, 4) of the main text, in which the emotion distributions are inferred through the method based on the emotion lexicon. We also apply the linear model on emotion distributions from the other two methods, and consistent results are obtained, as shown in Table S17 (3, 6). Specifically, the positive coefficient of anger indicates its causal promotion on the spread, while the negative coefficient of joy indicates its preventive effect on the circulation of fake news. Furthermore, the coefficients of emergency, military topic, and social topic are significantly positive, implying their roles in enhancing the spread of information.
| Variables          | Bayes Fake | Bayes Retweet | BP1 Fake | BP1 Retweet |
|--------------------|------------|---------------|----------|-------------|
|                    | (1)        | (2)           | (3)      | (4)         |
| Anger              | 2.809***   | 2.834***      | 36.203***| 2.450***    |
|                    | (0.230)    | (0.176)       | (13.766) | (0.175)     |
|                    |            |               | (0.144)  | (7.364)     |
| Joy                | -4.678***  | -4.266***     | -103.948**| -1.306***   |
|                    | (0.235)    | (0.174)       | (12.597) | (0.098)     |
|                    |            |               | (0.071)  | (3.906)     |
| Others             | 0          | 0             | 0        | 0           |
|                    | (omitted)  | (omitted)     | (omitted)| (omitted)   |
| Follower           | -5.84e-08  | -3.07e-07     | 0.00001* | -8.63e-08   |
|                    | (0.25e-09) | (1.60e-08)    | (6.57e-07)| (1.55e-08) |
|                    |            |               | (2.31e-08)| (8.66e-07) |
| Friend             | 0.0008***  | 0.00005***    | 0.041*** | 0.0007***   |
|                    | (0.00003)  | (0.00003)     | (0.002)  | (0.00005)   |
|                    |            |               | (0.00003)| (0.003)    |
| Mention            | 0.069      | -0.161***     | 18.374***| -0.023      |
|                    | (0.061)    | (0.046)       | (3.900)  | (0.091)     |
|                    |            |               | (0.069)  | (4.690)     |
| Hashtag            | -1.240***  | -1.560***     | -2.510   | -1.161***   |
|                    | (0.067)    | (0.048)       | (3.610)  | (0.096)     |
|                    |            |               | (0.069)  | (5.086)     |
| Location           | -0.397     | -0.433***     | -8.549** | -0.312      |
|                    | (0.063)    | (0.044)       | (2.504)  | (0.093)     |
|                    |            |               | (0.065)  | (3.550)     |
| Date               | -0.661     | -1.324***     | -2.310   | -0.647***   |
|                    | (0.051)    | (0.036)       | (2.611)  | (0.074)     |
|                    |            |               | (0.054)  | (3.741)     |
| URL                | -2.272***  | -1.580***     | -26.224**| -2.035***   |
|                    | (0.057)    | (0.035)       | (2.076)  | (0.084)     |
|                    |            |               | (0.055)  | (2.787)     |
| Length             | -0.007***  | 0.006***      | -0.195** | 0.006***    |
|                    | (0.005)    | (0.0004)      | (0.029)  | (0.001)     |
|                    |            |               | (0.007)  | (0.049)     |
| Emergency          | 5.418***   | 4.793***      | -25.962  | 5.169**     |
|                    | (0.721)    | (0.584)       | (6.470)  | (0.733)     |
|                    |            |               | (0.592)  | (7.287)     |
| Finance            | -1.129***  | -0.475***     | -27.532  | -0.209***   |
|                    | (0.094)    | (0.056)       | (4.395)  | (0.132)     |
|                    |            |               | (0.085)  | (5.840)     |
| International      | -6.952***  | -1.012***     | -9.402   | -0.265      |
|                    | (0.142)    | (0.110)       | (10.466) | (0.209)     |
|                    |            |               | (0.164)  | (17.156)    |
| Military           | 0.252      | 0.263***      | 13.958   | 1.550***    |
|                    | (0.134)    | (0.104)       | (10.264) | (0.203)     |
|                    |            |               | (0.163)  | (14.566)    |
| Society            | 0.242***   | 0.726***      | -26.640**| 1.307***    |
|                    | (0.069)    | (0.051)       | (4.160)  | (0.097)     |
|                    |            |               | (0.071)  | (5.237)     |
| Sports             | -0.855***  | -1.422***     | 52.292** | -0.926**    |
|                    | (0.121)    | (0.099)       | (10.198)| (0.190)     |
|                    |            |               | (0.150)  | (13.741)    |
| Technology         | 0.108      | -0.198***     | -7.469   | 0.568       |
|                    | (0.091)    | (0.070)       | (5.745)  | (0.139)     |
|                    |            |               | (0.109)  | (7.304)     |
| Cons               | 0.065      | 1.278***      | 75.999** | -0.294**    |
|                    | (0.120)    | (0.091)       | (7.639)  | (0.139)     |
|                    |            |               | (0.106)  | (8.469)     |
| R²                 | 0.353      | 0.360         | 0.123    | 0.395       |
|                    | 0.361      | 0.308         | 0.632    | 0.865       |
| N                  | 13063      | 30816         | 30816    | 6382        |

Table S17: (1,4) The logit model for LT news and HF news. (2,5) The logit model for T news and F news. (3,6) The linear model for L news and H news. *P < 0.1, **P < 0.05, ***P < 0.01.
S10 Additional validations on English news

It was stated that emotion expression is culture dependent (58). Though previous results on diffusion networks (see S2) and timeline analysis (see S3) demonstrate the consistency with English tweets in Twitter and suggest the universality of our data from Weibo, more evidence on the roles of anger and joy in the circulation through regression models of causal inference is still necessary. Here six datasets publicly available online are accordingly utilized to ensure that our results could apply to English news (tweets) that from Twitter and even other mainstream news media like WASHINGTON (Reuters). These datasets include:

(1) Dataset S1: 12,247,065 coronavirus (COVID-19) tweets posted from 4 March 2020 to 28 March 2020 in Twitter.\(^7\)

(2) Dataset S2: 8,642,360 coronavirus (COVID-19) tweets posted from 29 March 2020 to 15 April 2020 in Twitter.\(^8\)

(3) Dataset S3: 3,835,546 coronavirus (COVID-19) tweets posted from 16 April 2020 to 24 April 2020 in Twitter.\(^9\)

(4) Dataset S4: 397,629 election day tweets scraped on the day of 2016 United States election in Twitter.\(^10\)

(5) Dataset S5: 23,481 fake news and 21,417 real news posted from 31 March 2015 to 19 February 2018 and miss retweets.\(^11\)

\(^7\)https://www.kaggle.com/smid80/coronavirus-covid19-tweets#2020-03-00%20Coronavirus%20Tweets%20(pre%202020-03-12).CSV
\(^8\)https://www.kaggle.com/smid80/coronavirus-covid19-tweets-early-april
\(^9\)https://www.kaggle.com/smid80/coronavirus-covid19-tweets-late-april
\(^10\)https://www.kaggle.com/kinguistics/election-day-tweet#selection_day_tweets.csv
\(^11\)https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset
(6) Dataset S6: 478 fake news (tweets) posted during breaking news related to the events including Prince Toronto, Charlie Hebdo, Germanwings-crash, Sydney siege and etc. in Twitter.\textsuperscript{12}

Each English tweet contain text, retweet counts, follower counts, friend counts and etc. in Datasets S1-4. Though there are no labels of whether fake or real on these tweets, the promoting effect of anger on retweeting can still be verified. We randomly extract 2,000 news items from each file (one file per day) in Datasets S1-3 and obtain 90,000 news items (57,508 news with retweets) related to COVID-19 totally. Besides, there are 72,182 politically related news with retweets in the Dataset S4. News with retweets extracted in Dataset S1-3 (COVID-19) and Dataset S4 (Politics) are thus combined to examine the effects of emotions on the spread. We divide news into L news and H news according to the number of retweets and build the logit model as

$$\text{logit}(p_h_{-\text{news}}) = \beta_0 + \beta_1 w_1 + \beta_2 w_2 + \beta_3 w_3 + \beta_4 w_4 + \beta_5 w_5 + \beta_6 w_6 + \beta_7 w_7 + \beta_8 w_8 + \beta_9 w + \beta_{10} w_{10} + \beta_{11} w_{11} + \beta_{12} w_{12} + \beta_{13} w_{13} + \beta_{14} w_{14} + \beta_{15} w_{15} + \beta_{16} w_{16} + \beta_{17} w_{17},$$

where

- $p_h_{-\text{news}}$ is the probability of H news (tweets with more than 10 retweets in the dataset).
- $\beta_0$ is the intercept.
- $\beta_1, \beta_2, ..., \beta_{17}$ are the coefficients of variables.
- $w_1, w_2, ..., w_{17}$ represent variables of anger, disgust, joy, sadness, fear, surprise, anticipation, trust, follower, friend, mention, hashtag, location, date, URL, length, and topic.
- Topic indicates politics or COVID-19.

\textsuperscript{12}https://figshare.com/articles/PHEME_dataset_for_Rumour_Detection_and_Veracity_Classification/6392078
The emotion lexicon from the National Research Council of Canada (NRC) is employed to infer the emotional distributions of all English news. It contains 14,182 words with eight emotions: anger, disgust, joy, sadness, fear, surprise, anticipation, and trust \((59, 60)\). The coverage of this emotion lexicon is 73.3\% on the dataset used in the logit model. Though emotions carried by English news here is expanded to eight emotions, the promoting effect of anger is still significant and joy is opposite as expected (Table S18). These results suggest that the promoting effect of anger in spread is independent to cultural differences and our results can be confidently extended to English news. Other emotions such as disgust and anticipation are also found to be significant but with negative coefficients, implying their prevention on the spread. It should also be noted that here the linear model is not examined due to the missing of retweeting time in these datasets, i.e., whether the spread of news is sufficiently sampled cannot be assured, and consequently it would be problematic to treat the number of retweets directly as dependable variables.

Since whether the news in the Dataset S1-4 is true or fake is not labeled, Dataset S5 of containing 21,417 true news (with 11,264 political news and 10,133 world news) and 23,481 fake news (with 9,050 news, 6,718 political news, 1,548 government news, 4,415 left-news, 781 U.S. news, and 776 Middle-east news) is further utilized to verify the divergence of anger between true and fake news. Note that true news may from sources such as WASHINGTON (Reuters) and Twitter, hence here the texts of title and body of the news are jointed together to perform the emotion inference (the coverage of emotion lexicon is nearly 100\%). As expected, anger occupation in fake news is higher than that in true news (true news \(\sim 0.110\), fake news \(\sim 0.123\), K-S test \(\sim 0.108\), \(P \sim 0\)). There is also a very small dataset (Dataset S6) of fake news containing 117 LF news (tweets) with emotions and 361 HF news (tweets) with emotions. It is consistent to results from Weibo (see Table 1) that HF news in Twitter carries more anger than the counterpart of LF news (LF news \(\sim 0.020\), HF news \(\sim 0.142\), K-S test \(\sim 0.416\), \(P \sim 0\)).
To sum up, results from these supplementary datasets of English news consolidate our conclusions derived from Weibo and support that independent to cultures and platforms, fake news carries more anger than real news and anger promotes the circulation of news online.
Table S18: The logit model for the English news about COVID-19 and politics. \( *P < 0.1, **P < 0.05, ***P < 0.01. \)
S11 Selecting typical news for questionnaires

Emotions of high arousal, such as anger and joy, are associated with information diffusion, particularly information sharing (24). To further investigate how anger and joy carried in news influence incentives underlying retweeting, which reignites the circulation of news on social media, offline questionnaires are conducted to bind the emotion divergence between fake news and real news with retweeting incentives. Due to the time consuming and intensive labor costs, it is challenging for questionnaires to cover all the fake news and true news in our data. Therefore, five typical news items respectively from groups of HLT news, LHF news, and HLF news are selected to perform the surveys. Similarly, in terms of news in these groups, the possible stimuli from emotions such as anger and joy to the retweeting incentives are hoped to be amplified to ease the following detection. To guarantee that the selection of news samples from each group is representative, each group of news is clustered before sampling. First, we use the word2vec model to convert the words in each news item into vectors of 200 dimensions and take the mean of these word vectors to represent the news item, i.e., the news item is similarly embedded in the space of 200 dimensions. Then, K-means clustering is employed to cluster each group of news items into five clusters. Next, based on including keywords with high importance in each news item (see S6) and intrinsic factors such as mention and hashtag in each group (see S8), representative texts are sampled from those near the cluster centers. Note that we do not deliberately consider emotion distributions in the selection to avoid the impact of subjective bias on subsequent incentive stimuli and to ensure the objectivity of the results. Finally, we select 15 typical news items (Table S19-22), and their positions in the group can be found in Fig. S17. The sampled texts and the keywords in these texts are distributed evenly in the embedding space of different groups of news, suggesting that they are indeed typical and representative. Notably, the selected keywords that help separate the groups of news in sampling the texts are anticipated
to help strengthen the stimuli of reposting incentives, which would further enhance the impact of anger and joy.
### News1

| CN | 西部资源重组媒体说明会【阙文彬回答媒体提问: 继续加大稀贵金属投资】西部资源实际控制人阙文彬说, 从新能源到文化再到稀贵金属, 我个人认为新的董事会或者新的经营班子接上后, 应该在2亿-5亿的利润差不多, 在这个基础上将现有的一些企业通过一种合法的途径出售。... 全文: [http://m.weibo.cn/1315875979/4010238174942685](http://m.weibo.cn/1315875979/4010238174942685) |
| EN | Western Resources Reorganization Media Briefing # [Wenbin Que Answers Media Questions: Continue to Increase Investment in Rare and Precious Metals] In response to media questions about the company's main business, Wenbin Que, the actual controller of Western Resources, said that from new energy to culture to rare and precious metals, I personally think that after the new board or new management team is connected, it should have a profit of about 200 million to 500 million. On this basis, some existing enterprises will be sold through a legal way. ... Full text: [http://m.weibo.cn/1315875979/4010238174942685](http://m.weibo.cn/1315875979/4010238174942685) |

### News2

| CN | 聚焦赣州【心】爸妈在哪里? 崇义文昌塔旁发现的小男孩至今还在福利院】8月17日, 一则“崇义县横水派出所民警在文昌塔附近一脐橙园树下, 发现一名哭泣的小男孩至今无人认领”的消息, 在微信朋友圈广泛转发。文章中还附有几张小男孩的照片。当日下午, 记者了解到, 目前小男孩在医院检查无碍后已被送往... 全文: [http://m.weibo.cn/1970239225/4009774025014136](http://m.weibo.cn/1970239225/4009774025014136) |
| EN | # Focus on Ganzhou # [heart] Where are the parents? The little boy found next to Wenchang Pagoda in Chongyi is still in the welfare home] On August 17th, a policeman from the Hengshui Police Station in Chongyi County found a crying little boy under a navel orange tree near Wenchang Pagoda. The "claim" message was widely reposted in WeChat Moments, and there are several pictures of the little boy in the article. In the afternoon of the same day, the reporter learned that the little boy was sent to... after being checked by the hospital. ... Full text: [http://m.weibo.cn/1970239225/4009774025014136](http://m.weibo.cn/1970239225/4009774025014136) |

### News3

| CN | 鹤壁身边事【淇滨区兰苑社区刘振强: 带爸妈旅行, 收获满满的幸福】“我父亲一直想出门走走，特别想去北京看一看。我以前没有在意，感觉父母还年轻，以后有的是机会。直到父亲生了一场大病，需要借助轮椅出行，我才感到了后悔，幸好还来得及补救。”8月16日，淇滨区兰苑社区的刘振强告诉 记者, 最近他... 全文: [http://m.weibo.cn/2514256341/4009491428875467](http://m.weibo.cn/2514256341/4009491428875467) |
| EN | # Things around Hebi # [Zhengqiang Liu, Lanyuan Community, Qibin District] Take my parents to travel and reap the full happiness] "My father always wanted to go out for a walk, especially to go to Beijing to take a look. I didn’t care before, I felt my parents were still young. There would be opportunities in the future. It was not until my father had a serious illness and needed to use a wheelchair to travel. I regretted it. Fortunately, I had time to remedy it." On August 16, Zhengqiang Liu of Lanyuan Community in Qibin District told reporters that he recently ... Full text: [http://m.weibo.cn/2514256341/4009491428875467](http://m.weibo.cn/2514256341/4009491428875467) |

### Table S19: HLT-News1-3 selected in HLT news. Keywords are highlighted in red.
"Fat man who doesn't know the ball" is hard [powerful] The fat man behind the Chinese table tennis team [sneers], yes! Guoliang Liu is definitely an all-rounder. The devil trains the team members, provides shouting, cheering, wake-up services, water and towels, kiss the team members [kiss] and have to cook the noodles to reward the three troops ... So China won all the gold medals in table tennis for the third consecutive Olympic Games [Olympic gold medal]. Some netizens said: "Being a father and being a mother ... full text:

http://m.weibo.cn/1891503444/4009944795388322

On July 12, the second edition of the newspaper presented to you: China Cultural Relics Conservation Foundation held a special fund work symposium, Tongling in Anhui made an emergency rescue of the ancient mining site of Jinniu Cave at Fenghuangshan Copper Mine, and Guobo held the "Four Medical Books". Ben Thangka Art Inheritance Achievement Exhibition, Hubei implemented a "three-level joint review" model, accelerated the promotion of cultural relics census data review, Xinjiang held the first national mobile cultural relics census training class, "South China Sea Geography Strategy" ...

Full text: http://m.weibo.cn/1250227403/3997198805156773
News1: Don’t go to the cinema on May 12. Please don’t enter the theater. Let’s work hard for Sadako’s box office zero. The box office of “The Flowers of War” filmed by Chinese is in zero in Japan. Sadako 3D filmed by Japanese will be released in the mainland of China on May 12. However, May 12 is also the anniversary of the Nanjing Massacre and the national disaster day. Don’t forget the national shame! As a Chinese, dare to make Sadako 3D’s box office zero on May 12. Friends, you must repost, repost.

News2: # Tianjin Tanguo Big Bang # I am not sure whether the text is true. I only know that I am very moved. I only know that a few batches of firefighters did not survive when they went to the scene of the explosion for people. They are so fearless and great. I can’t do anything except give them the most sincere thanks [prayer]. I am just a student who has just grown up. The views are natural immature. I only hope that people can do what they should do and not be blinded by their interests.

News3: A friend picked up an admission ticket and let her know if you know her: name: Yaqian Bai, examination center: No. 1 middle school, examination room: 013, seat number: 11, admission ticket number: 204101311. Contact number: 15935078941. Don’t delay the child's college entrance examination. Help others and make your hands fragrant! Thank you! @ Happy Xiao-Xiao-Le @ Happy Zhang-Jiang @ Shanghai Pudong Chuansha Police Station @ Interactive Chuansha

News4: Love Relay: Yuming Hu, female, 4 and a half years old, from Yuncheng. Save her, she suffers from a rare "Bugat's syndrome" that has produced antibodies to hormones, and her weight continues to rise. She repeats a sentence every day: Mom, hurt! I hope everyone can help her, one more person forwards more power. @ Han Hong Caring Charity Foundation @ 365 Child Rescue Caring Fund

News5: Starting at 6 o'clock this afternoon, the city's high-definition probes will be all activated. The co-pilot does not wear a seat belt, which is subject to the same penalty. It will be fined 50 yuan for talking on the phone while driving, 200 yuan for breaking yellow flash lamp, and 100 yuan for parking over the line. The 60-day national traffic police centrally investigates and deals with drunk driving. Once seized, they will be detained for six months in total, and no driver’s certificate is allowed within five years. Please interactively tell the car owners, friends and relatives to avoid fines.
| News1 | Urgent notice: The Maternal and Child Health Hospital notice: There are more and more children with leukemia. The Maternal and Child Health Hospital reminds you, please don't give your baby Shuangwai milk and milk drinks with additives. Wangzai milk, Coca-Cola, Shuangwaiwai, Wahaha AD Calcium Milk, Futurestar, Qstar, Mei zhiyuan Fruit Milk. Both contain botulinum. Now an emergency recall. Anyone with a child reposts! ! ! People without baby repost! ! ! |
|-------|--------------------------------------------------------------------------------------------------|
| News2 | Look, Amway boss is dead! Just 56 years old, eating Nutrilite for 27 years, so ironic. Look again! Amway Chengguan, the founder of the 3S system, Guantian Chen died of liver cancer at the age of 56, worked for 27 years in Amway, and 27 years in Nutrilite, making 27 years of money for Americans. Please look again, the author of 'Away from Poverty', the founder of Amway Master Ciguan Wang, died in Fuzhou. He eats Amway health care products every day at the age of 61. |
| News3 | The first Ebola in China has been discovered in Ningbo, with a basic mortality rate of 90%. The time of inflow into China is ten days earlier than the time estimated by experts. Everyone must remind children and their families to wash their hands with soap at any time, do not eat street stalls and open-air food, and buy the finished food home to boil and eat, precautions! Remember this time Ebola is likely to develop into a more terrible plague than SARS. Pay attention to hygiene and take care! [blush] |
| News4 | The little girl died after using up the uncharged mobile phone charger. She put one end of the charger into her mouth and was electrocuted. The girl’s parents regretted it and stood up to warn everyone! Please don’t let the tragedy repeat! |
| News5 | 【可恶！骆驼被砍四肢当街行乞】骆驼一般只在动物园才能见到，但近日，人们却在福州街头看到一只乞讨的骆驼。骆驼身旁有两位衣衫褴褛的人跪在地上磕头乞讨。民警发现，骆驼的四肢均有不同程度的损伤，四肢均无蹄子。据伤口观测，很大可能是人为造成。警方已协调相关部门处理。 |

Table S22: HLF-News1-5 selected news from HLF news. Keywords are highlighted in red.
S12 Questionnaires

We employ a carefully designed questionnaire that is commonly used for rumor sharing motivation surveys on social media (42), which comprehensively measures four motivations of the subjects: anxiety management, information sharing, relationship management and self-enhancement. There are six items for anxiety management (Fig. S18), six items for information sharing (Fig. S19), five items for relationship management (Fig. S20) and four items for self-enhancement (Fig. S21). Each item is measured on a four-point scale (1-strongly disagree, 2-disagree, 3-agree, 4-strongly agree). There are six questionnaires in total. For each group of news items, we implement two online questionnaires, one showing the original text and one showing the text with keywords marked in red squares (Fig. S22). Meanwhile, five news items from each group appear in each questionnaire randomly. Except for the news presented, all other circumstances in the questionnaires, e.g., author profile, posting time, and posting source, are carefully controlled to be consistent. Specifically, the difference in stimuli to the incentives of subjects is only the news itself. For the presentation of the text, we attempted to simulate the real Weibo interface by adding the background of the mobile version of the Weibo App to each news item (Fig. S22). For subjects who completed the questionnaires, we required them to be Weibo users aged between 18 and 30 years old (according to the 2018 Weibo user development report, this age group accounts for 75% of all users), matching users in online data as much as possible.\(^\text{13}\) Note that subjects are not specifically targeted based on occupation or income level because we want to probe the general effect of emotion divergence on the retweeting incentives for the majority of Weibo users. More importantly, considering the widespread global impact of fake news online, revealing a mechanism that is independent of user demographics would be powerful in inspiring new cures.

\(^{13}\)https://data.weibo.com/report/index
Fig. S18: Anxiety Management Motivation (M1).

I will feel relaxed after sharing this message
分享这条消息后我会感到轻松

I am worried about others and sharing this message will help keep them safe
我很担心别人的情况分享这条消息有助于保护他们的安全

Sharing this message will make me feel in control of the situation
分享这条信息会让我感觉自己能掌控局面

Sharing this message will create a pleasant mood in me
分享这条信息会给我带来愉快的心情

Sharing this message will make me feel confident
分享这条信息会让我感到自信

I am motivated to share this message and reduce my anxiety regarding the event
我有动机分享这条信息以减少我对事件的焦虑。
Fig. S19: Information Sharing Motivation (M2).
Fig. S20: Relationship Management Motivation (M3).
Fig. S21: Self Enhancement Motivation (M4).

I will share this message to pass time
我会分享这条信息来打发时间

I will share this message to let others know about my activities
我会分享这条信息让其他人知道我的动态

Sharing this message will help others know about my interests
分享这条信息会帮助其他人了解我的兴趣

I will share this message because it’s enjoyable to me
我会分享这条信息，因为它对我来说很有趣
Fig. S22: Questionnaire examples of original text (left) and text with marked keywords (right).
S13 Questionnaire results

We hired a well-reputed online survey company\(^{14}\) and collected a total of 1291 valid responses from 1316 subjects within China. Specifically, we obtained 224 responses to the unmarked HLT news questionnaire (HLT-Q1), 214 responses to the marked HLT news questionnaire (HLT-Q2), 210 responses to the unmarked LHF news questionnaire (LHF-Q1), 212 responses to the marked LHF news questionnaire (LHF-Q2), 211 responses to the unmarked HLF news questionnaire (HLF-Q1) and 210 responses to the marked HLF news questionnaire (HLF-Q2). All the responses are carefully validated, and the values of Cronbach’s alpha are provided in Table S23. The collected responses are also publicly available at https://doi.org/10.6084/m9.figshare.12163569.v2. Since subjective bias may exist, that is, the response degree might vary across different subjects, the following method is adopted to eliminate the subjective bias:

\[
Mi - avg = m_i - \frac{m_1 + m_2 + m_3 + m_4}{4}, \quad i = 1, 2, 3, 4
\]

where \(m_i\) is the average score of all the items in motivation \(Mi\) and \(Mi - avg\) is the debiased average score for \(Mi\).

S13.1 Differences in motivations between different groups of news

The main text showed that the motivation of information sharing of false news is stronger than that of real news, and the motivation of anxiety management of LHF news is significantly stronger than that of news in both HLF and HLT. For responses with keywords outlined, these differences are significant and even augmented, and interestingly, the differences between LHF news and the other two groups of news are more significant in M1 (Fig. S23A), implying

\[^{14}\text{https://www.wjx.cn/}\]
Table S23: The values of Cronbach’s alpha in different questionnaires.

|        | M1   | M2   | M3   | M4   | N  |
|--------|------|------|------|------|----|
| LHF-Q1 | 0.775| 0.718| 0.771| 0.768| 210|
| LHF-Q2 | 0.774| 0.682| 0.794| 0.732| 212|
| HLF-Q1 | 0.787| 0.706| 0.799| 0.721| 211|
| HLF-Q2 | 0.759| 0.714| 0.773| 0.768| 210|
| HLT-Q1 | 0.702| 0.562| 0.714| 0.695| 224|
| HLT-Q2 | 0.744| 0.642| 0.777| 0.724| 214|

Fig. S23: (A to D) CCDFs for the motivations in different groups of news with marked keywords. (E to F) CCDFs for the motivations in the two groups separated randomly.

audiences of highly retweeted fake news are more incentivized in terms of anxiety management.

The statistics and K-S tests are shown in Table S24 and Table S25.

**S13.2 Differences in motivations between anger and joy**

Next, we divide the news in the questionnaires according to the emotions it carries with the largest occupation. News1 and News5 in LHF news are dominated by anger. Joy dominates News2 in LHF news and News1, 3, 4, 5 in HLT news. The rest of the news is dominated by other emotions. In the analysis in S13.1, we found that the marked keywords play a role in
|          | M1-avg Mean | Std     | M2-avg Mean | Std     | M3-avg Mean | Std     | M4-avg Mean | Std     |
|----------|-------------|---------|-------------|---------|-------------|---------|-------------|---------|
| LHF-Q1   | 0.051052    | 0.321949| 0.073487    | 0.380414| -0.115521   | 0.378259| -0.05692    | 0.345747|
| LHF-Q2   | 0.073487    | 0.380414| 0.498801    | 0.40541 | -0.264092   | 0.384499| -0.308196   | 0.409732|
| HLF-Q1   | -0.115521   | 0.378259| 0.58906     | 0.43338 | -0.228476   | 0.382048| -0.245063   | 0.42301 |
| HLF-Q2   | -0.097599   | 0.347347| 0.565893    | 0.421701| -0.188552   | 0.333062| -0.279742   | 0.447777|
| HLF-Q1   | -0.05692    | 0.345747| 0.390253    | 0.414137| -0.317336   | 0.382803| -0.015997   | 0.416897|
| HLF-Q2   | -0.037578   | 0.337936| 0.353388    | 0.395185| -0.264213   | 0.360444| -0.051597   | 0.425809|

Table S24: The statistics of each motivation in each questionnaire.

|          | LHF-HLF-Q1 D ~ 0.235, p-value ~ 0 | LHF-HLT-Q1 D ~ 0.144, p-value ~ 0.019 | HLF-HLT-Q1 D ~ 0.117, p-value ~ 0.091 | LHF-HLF-Q2 D ~ 0.242, p-value ~ 0 | LHF-HLT-Q2 D ~ 0.153, p-value ~ 0.012 | HLF-HLT-Q2 D ~ 0.107, p-value ~ 0.158 |
|----------|-----------------------------------|---------------------------------------|--------------------------------------|-----------------------------------|---------------------------------------|--------------------------------------|
| M1-avg   |                                   |                                       |                                      |                                   |                                       |                                      |
| M2-avg   | D ~ 0.127, p-value ~ 0.056        | D ~ 0.167, p-value ~ 0.004            | D ~ 0.212, p-value ~ 0               | D ~ 0.094, p-value ~ 0.282        | D ~ 0.187, p-value ~ 0.001            | D ~ 0.266, p-value ~ 0               |
| M3-avg   | D ~ 0.103, p-value ~ 0.193        | D ~ 0.100, p-value ~ 0.207            | D ~ 0.130, p-value ~ 0               | D ~ 0.136, p-value ~ 0.058        | D ~ 0.136, p-value ~ 0.004            | D ~ 0.106, p-value ~ 0               |
| M4-avg   | D ~ 0.087, p-value ~ 0.359        | D ~ 0.284, p-value ~ 0.230            | D ~ 0.096, p-value ~ 0               | D ~ 0.274, p-value ~ 0.261        | D ~ 0.274, p-value ~ 0.214            | D ~ 0.214, p-value ~ 0               |

Table S25: The results of K-S tests
widening differences. Hence, we directly combine the responses without keywords and those with keywords according to their dominant emotions to further examine the emotions stimuli with respect to retweeting motivation. The results are analyzed in the main text, and the K-S tests results are shown in Table S26. Furthermore, in terms of neglecting emotion dominance, all the data of the questionnaires are divided into two groups randomly to analyze the difference in motivations. Surprisingly, no significant differences were observed in the four motivations (Fig. S23E-H) (anxiety management: K-S test $\sim 0.040$, $P \sim 0.673$; information sharing: K-S test $\sim 0.062$, $P \sim 0.168$; relationship management: K-S test $\sim 0.053$, $P \sim 0.317$; self-enhancement: K-S test $\sim 0.059$, $P \sim 0.200$), suggesting the significance of the different incentives provoked by anger and joy.

Table S26: The statistics and K-S tests for anger and joy.
Carrying more anger makes fake news more viral than real online news. According to this conclusion, instead of determining new features in fake news detection, developing new cues of tagging anger on social media is a promising approach to restrain the spread of fake news at the source. Because the intervention can be implemented immediately after posting, there will be no lag in the fight against fake news. More importantly, the principle of guaranteeing the freedom of speech will be respected, and an acceptable trade-off between free sharing and fake news prevention can be achieved. By alerting users of angry tweets, audiences can be persuaded to assess them more critically before emotionally retweeting, consequently leading to less emotional and more rational retweeters. Specifically, for tweets (news) that deliver too much anger, e.g., the occupation of anger surpasses a predetermined threshold ($\theta$), a retweeting warning could be provided on platforms such as Twitter, Facebook, and Weibo. According to a report of Facebook in battling misinformation of COVID-19, warning labels can effectively prevent 95% users from further access.\(^{15}\) In accordance with this, it is very optimistically assumed here that no angry tweets with warning tags from the platform will be retweeted. To determine the value of $\theta$, we focus on news with high volumes of retweets (HT news and HF news in our data) and define a measure to optimize $\theta$, i.e., preventing fake news that will be highly retweeted but not real news that will be popular. The measure is denoted as $\beta$ and is defined as

$$\beta = \frac{N_{HF(\geq \theta)}}{N_{HF}} - \frac{N_{HT(\geq \theta)}}{N_{HT}},$$

where

- $N_{HF}$ is the number of HF news items.
- $N_{HF(\geq \theta)}$ is the number of HF news items with an occupation of anger greater than $\theta$.

\(^{15}\)https://about.fb.com/news/2020/04/covid-19-misinfo-update/
• $N_{HT}$ is the number of HT news items.

• $N_{HT(\geq \theta)}$ is the number of HT news items with an occupation of anger greater than $\theta$.

The values of $\beta$ for $\theta$ values increasing with a step size 0.1 and 0.05 are shown in Fig. S24 and Fig. S25, and the values peak when $\theta = 0.2$. In our dataset from Weibo, warning about news in which anger occupies more than 20% will efficiently and effectively prevent 46% of highly retweeted fake news and only influence the circulation of 22% of popular real news. And for all highly retweeted news in our dataset (i.e., HF+HT), in those with occupation of anger higher than 0.2, HF news accounts for 89%, implying further that our treatment can predominantly target highly retweeted fake news. Though the fraction of prevented fake news that otherwise will be widely circulated is not as high as expected, considering the intrinsic characteristics of very low cost and timely intervention, the newly presented treatment should be weighted with high priority in the toolbox of cures against fake news. Hence, it is worth trying on social media platforms such as Weibo, Twitter or Facebook to prevent the spread of fake news online at the source through this new approach.
Fig. S24: The value of $\beta$ with $\theta$ growing by 0.1.
Fig. S25: The value of $\beta$ with $\theta$ growing by 0.05.