

Social Media and Fake News in the 2016 Election

Hunt Allcott & Matthew Gentzkow

Journal of Economic Perspectives, 2017

美国民主一直受到传媒技术更替的影响，19 世纪廉价的出版印刷显著扩张了党派报纸的传播，很多人认为媒体的权力监督作用因此被削弱 (Kaplan 2002)；20 世纪，广播电视成为主流，人们开始担心严肃的政治辩论变成剪短的原声摘要 (sound bites)、真正有领导能力的候选人不及更上镜的有魅力者、权力被集中于寡头 (Lang and Lang 2002; Bagdikian 1983)；21 世纪初，网络新闻的兴起带来新的顾虑，纷繁的观点使公民只相信自己愿意相信的而回避反面视角 (Sunstein 2001a, b, 2007; Pariser 2011)；最近，担忧集中于社交媒体，它没有第三方筛选审核，个人也可以获得极大的知名度。有证据表明：

- 62% 的美国成年人在社交网络上获取新闻 (Gottfried and Shearer 2016)
- 最流行的假新闻在 Facebook 上比主流新闻流传得更广 (Silverman 2016)
- 很多看了假新闻的人表示相信 (Silverman and Singer-Vine 2016)
- 被议论最多的假新闻更倾向于 Donald Trump 而不是 Hillary Clinton (Silverman 2016)

经济学模型中，媒体公司收集和出售关于真实情况的信号，消费者受益于真实情况，而假新闻将信号扭曲至与真实情况无关。假新闻因为比准确信号成本低而兴起，消费者不能无成本地判断真伪，且他们也喜欢政治新闻。假新闻提高了一些消费者的效用，但了解真实情况变得更难了，这也增加了私人和社会成本。

而假新闻的相关话题在经济、心理、政治、计算机领域都有研究，Flynn, Nyhan, and Reifler (2017) 近期对政治中的错误认知做过综述：许多研究关注新的信息怎样影响政治信念 (Berinsky 2017; DiFonzo and Bordia 2007; Taber and Lodge 2006; Nyhan, Reifler, and Ubel 2013; Nyhan, Reifler, Richey, and Freed 2014), 流言怎样传播 (Friggeri, Adamic, Eckles, and Cheng 2014), 媒体曝光的作用 (Bartels 1993, DellaVigna and Kaplan 2007, Enikolopov, Petrova, and Zhuravskaya 2011, Gerber and Green 2000; Gerber, Gimpel, Green, and Shaw 2011; Huber and Arceneaux 2007; Martin and Yurukoglu 2014; Spenkuch and Toniatti 2016; and for overviews, DellaVigna and Gentzkow 2010 and Napoli 2014), 新闻消费中的意识形态分离 (Bakshy, Messing, and Adamic 2015; Gentzkow and Shapiro 2011; Flaxman, Goel, and Rao 2016).

1 Background: The Market for Fake News

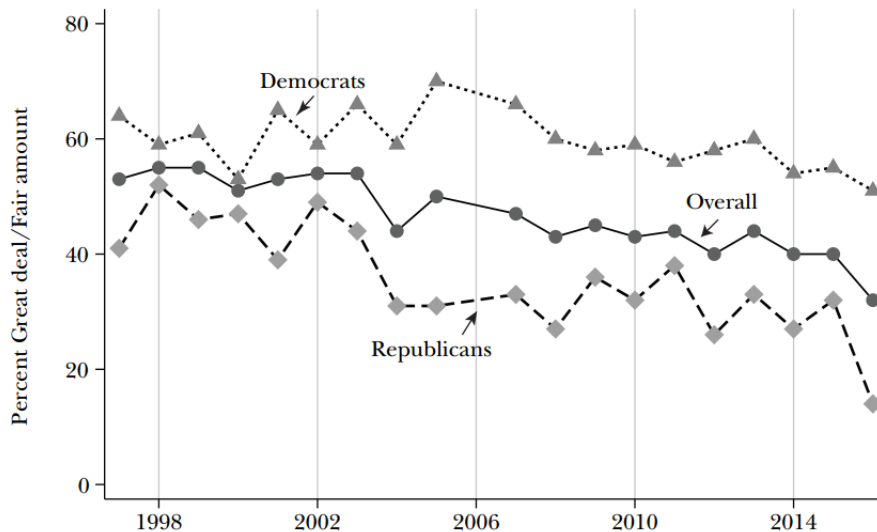
- 假新闻的定义：有意地且实际上虚假的，能误导读者的新闻文章。
- 网站易于搭建、广告平台使网络内容易于盈利，造成传媒产业的进入壁垒下降，传媒巨头顾及名誉，没有报道假新闻的激励，而社交网络则非常适合假新闻的传播，且社交网络在快速增长。同时，大众对主流媒体的信任也在下降 (Figure 2A)，政治变得两极化，双

方的相互评价都在走低 (Figure 2B), 这使得大家都更容易相信关于对方的负面假新闻。

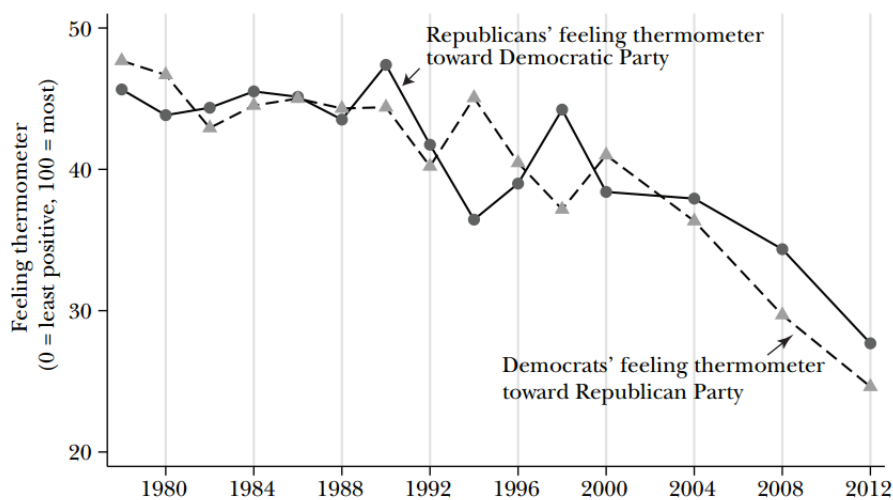
Figure 2

Trends Related to Fake News

A: Trust in Mainstream Media



B: Feeling Thermometer toward Other Political Party



- 假新闻主要来源于网站, 且这些网站的存在时间都不长。产生假新闻有两个动机, 一是大量点击有广告盈利, 二是出于意识形态。
- Gentzkow, Shapiro, and Stone (2016) 提出了关于假新闻的经济模型:
 - 存在两个未观测到的事实, 即左倾或者右倾候选人的执政表现是否更好。媒体公司得到了有关这两个事实的信号, 但这些信号的准确性存在差异。公司可以增加成本提高信号的准确性, 也可以根据公布信号的收益改变报道策略, 他们可以选择是如实报道信号, 还是给信号添加偏差。
 - 消费者对于世界有不同的信念, 自由派认为左倾候选人会表现更好, 而保守派认为右倾候选人会表现更好。消费者从两个途径获得效用, 1. 他们想知道真相, 2. 他们对与自己观点一致的报道产生心理效用。以上两者存在权衡取舍, 消费者选择能给

他带来最大期望效用的公司。当消费者作出决策后，并知道了实际结果，他们更新对于公司的信念，并决定未来选择哪家公司。

- 公司的利润取决于选择消费它的人数（带来广告收益），且公司有保持名誉、提高消费者效用的激励。
- 消费者选了高质量的候选人会产生正外部性。

在这个模型中，有两个不同的激励会使公司扭曲他们的报道，1. 报道真相的回报是有限的，理性的消费者会认为给出接近他们信念的报道的公司质量更高 (Gentzkow and Shapiro 2006)，2. 由于心理效用，消费者可能更喜欢印证他们信念的报道 (Mullainathan and Shleifer 2005)。这些激励都会导致均衡中的报道有偏。假新闻的产生者有两个显著的特征：1. 他们对准确报道没有投资，所以他们产生的信号与真相无关。2. 他们不追求长期声誉，只追求短期利润最大化。

新闻市场中加入假新闻生产者增加了潜在社会成本：1. 消费者因为有不准确的信念而受损，2. 不准确的信念降低了社会正外部性，难以识别高质量的候选人，3. 消费者更加怀疑真实的新闻生产者，更难识别新闻造假者，4. 这些因素在供给端被强化，对于高准确性的需求下降，低偏差的报道提高准确性的激励减少。与这些负面影响对应，乐于阅读印证自己信念的假新闻的读者效用增加，两者存在权衡取舍。

2 Real Data on Fake News

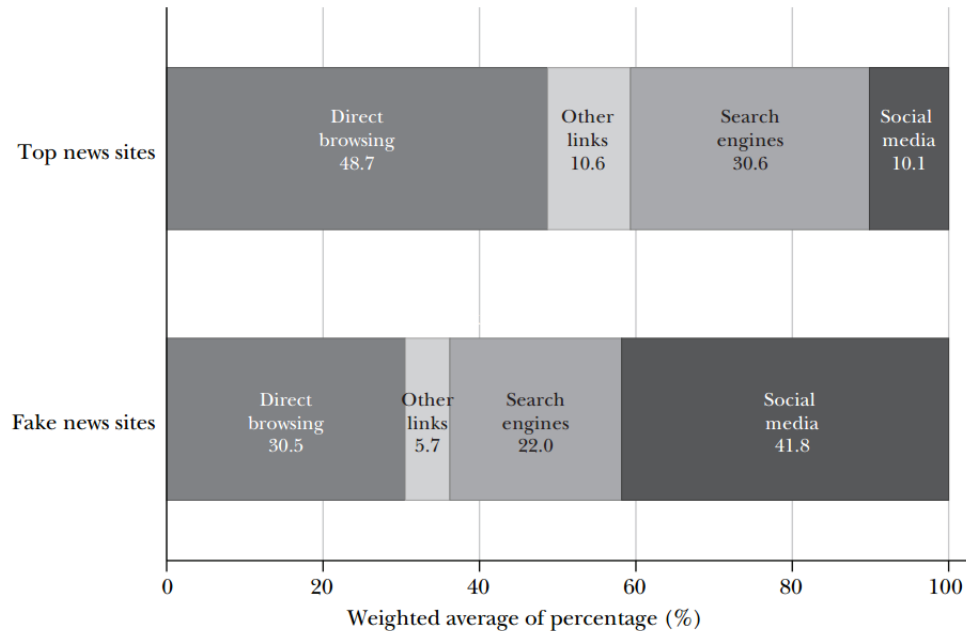
- 数据来源于 snopes.com、plotifact.com 上爬取的带有 Donald Trump 和 Hillary Clinton 标签的文章，以及 21 个在 Facebook 上备受点击的假新闻，这三个来源一共包含 156 篇假新闻，分为支持 Clinton(反对 Trump) 和支持 Trump(反对 Clinton) 两类
- 另一组数据来自在线调查问卷，包括政治立场、种族、受教育程度、在大选上的花费（阅读大选新闻花费的时间以及在社交网络上阅读花费的时间）、获取大选信息的最主要来源。最后对于每个人，给 15 个大选新闻标题，问其是否记得有这个新闻，以及当时是否相信这条新闻。当然，数据集也包括年龄、收入水平、性别等背景信息。

3 Social Media as a Source of Political Information

社交网络上进入市场和生产内容的固定成本非常低，这提高了小范围、短期策略（这也是假新闻制造者常常采用的）的相对利润，降低了打造长期声誉提高质量的利润；另外，社交媒体的形式——手机上的窄条信息和新闻推送窗——难以辨认文章的准确性；Bakshy, Messing, and Adamic (2015) 发现 Facebook 上有意识形态分割，意识形态相近的好友圈很少有对立意识形态的内容，人们更喜欢阅读和分享与他们意识形态相近的新闻。这意味着社交网络上的用户获得真实信息的可能性更低。

从网络流量来源可以看出社交网络对假新闻生产者的重要性 (Figure 3):

Figure 3
Share of Visits to US News Websites by Source



4 Partisanship of Fake News

在上文的数据集中，有 41 条支持 Clinton 的假新闻，115 条支持 Trump 的假新闻，在 Facebook 上分别被分享了 760 万次和 3030 万次。也有研究证实了这一点，假新闻生产者发现“支持 Trump(反对 Clinton)”假新闻的需求更大，便相应地提供更多这样的假新闻 (Sydell 2016)。这种现象可能是因为：对主流媒体共和党新闻的信任下降 (Figure 2) 提高了对非主流渠道新闻的需求；支持 Trump 的新闻更吸引人；也可能是共和党人更喜欢相信假新闻（理论上有这个可能:）。最后一点在之前的研究中不成立，McClosky and Chong(1985)、Uscinski, Klostad, and Atkinson (2016) 发现无论是偏左还是偏右的人群，对于阴谋论的倾向是一致的。Bakshy, Messing, and Adamic (2015) 发现保守派比自由派接触了更多跨领域的新闻，这使得他们更能识别假政治新闻。

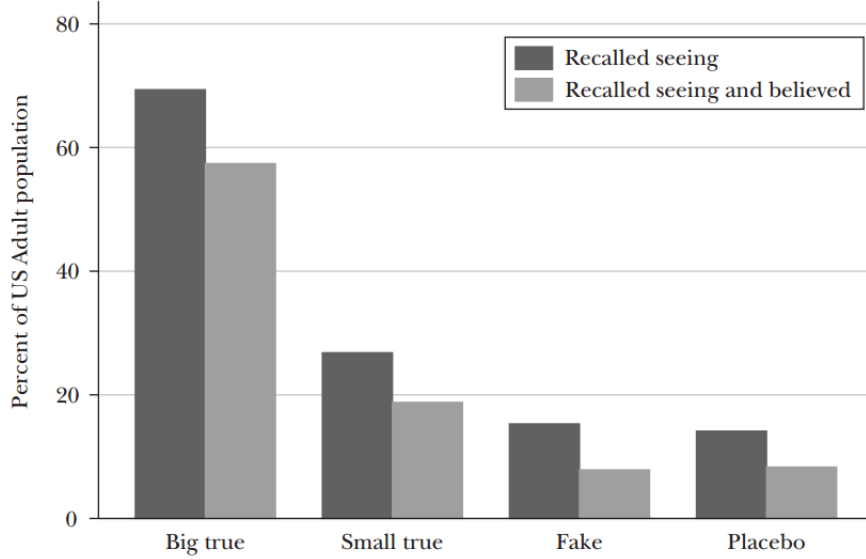
5 Exposure to Fake News

选民在 2016 大选中看了多少假新闻？主要有三种估计方法：

- 用分享次数估计阅读量。Eventbrite (2012) 和 Jessica Novak 估计 Facebook 的阅读转发比大致为 14, Zhao, Wang, Tao, Ma, and Guan (2013) 发现人人网视频的观看转发比在 3 至 8 之间，综合估计大选假新闻的阅读转发比不超过 20，这意味着 3800 万的假新闻分享量大致意味着 7.6 亿阅读量。
- 用网站浏览数据衡量假新闻网站的影响力。
- 用大选后调查问卷估计受众看到并记得的文章。结果如 Figure 5, Fake 和 Placebo 的差值 (1.2%) 代表真正被受试者记住的假新闻：

Figure 5

Percent of US Adult Population that Recall Seeing or that Believed Election News



Notes: In our post-election survey, we presented 15 headlines. For each headline, the survey asked whether respondents recall seeing the headline (“Do you recall seeing this reported or discussed before the election?”) and whether they believed it (“At the time of the election, would your best guess have been that this statement was true?”). The left bars present the share of respondents who recall seeing the headlines in each category, and the right bars present the share of respondents who recall seeing and believed the headlines. “Big True” headlines are major headlines leading up to the election; “Small True” headlines are the minor fact-checked headlines that we gathered from Snopes and PolitiFact. The Placebo fake news headlines were made-up for the research and never actually circulated. Observations are weighted for national representativeness.

调查中的 Fake 标题被转发了 38.6 万次，如果 1.2% 的美国人看到并记住了这些新闻，那么平均每个人看到并记住了 1.14 条数据集中的假新闻。

6 Who Believes Fake News?

真实消息对个人和社会都有利,那么哪些因素能预测辨别真伪的能力?相关研究包括 Lewandowsky, Oberauer, and Gignac (2013), Malka, Krosnick, and Langer (2009), and Oliver and Wood (2014).

$$C_{ia} = \alpha_1 X_i + \alpha_0 + \epsilon_{ia}$$

如果受试者辨认出新闻标题真伪 $C_{ia} = 1$ ，不确定则 $C_{ia} = 0.5$ ，辨认错则 $C_{ia} = 0$ 。

Table 1

What Predicts Correct Beliefs about News Headlines?

	(1)	(2)	(3)	(4)	(5)
Democrat	0.029 (0.020)	-0.004 (0.023)	0.028 (0.019)	-0.010 (0.021)	0.015 (0.013)
Republican	-0.024 (0.024)	0.040 (0.027)	-0.037* (0.020)	0.021 (0.023)	-0.018 (0.014)
ln(Daily media time)			-0.002 (0.007)	0.042*** (0.008)	0.013*** (0.004)
Social media most important			-0.066*** (0.025)	0.065*** (0.024)	-0.023 (0.016)
Use social media			0.014 (0.030)	-0.023 (0.038)	0.002 (0.019)
Social media ideological segregation			-0.027 (0.036)	0.028 (0.046)	-0.008 (0.024)
Education			0.014*** (0.004)	0.004 (0.004)	0.011*** (0.003)
Undecided			-0.011 (0.017)	0.006 (0.022)	-0.005 (0.013)
Age			0.002*** (0.000)	0.000 (0.001)	0.002*** (0.000)
N	12,080	6,040	12,080	6,040	18,120
p-value (Democrat = Republican)	0.029	0.124	0.004	0.207	0.035
Articles in sample	False	True	False	True	All

Note: This table presents estimates of a regression of a dependent variable measuring correct beliefs about headlines on individual characteristics. Columns 1 and 3 include only false headlines, columns 2 and 4 contain only true headlines, and column 5 contains all headlines. All columns include additional demographic controls: income, race, and gender. “Social media most important” means social media were the respondent’s most important sources of election news. “Social media ideological segregation” is the self-reported share (from 0 to 1) of social media friends that preferred the same presidential candidate. “Undecided” is an indicator variable for whether the respondent decided which candidate to vote for less than three months before the election. Observations are weighted for national representativeness. Standard errors are robust and clustered by survey respondent.

*, **, *** indicate statistically significantly different from zero with 90, 95, and 99 percent confidence, respectively.

相关研究表明：

- 提高受教育程度能减轻政治误解 (Flynn, Nyhan, and Reifler 2017)
- 人们的预设立场能很好地预测其信念 (Oliver and Wood 2014; Uscinski, Kloststad, and Atkinson 2016)，比如共和党人更容易相信奥巴马不是美国人、民主党人更可能相信布什策划了 9.11 事件 (Cassino and Jenkins 2013)

为了验证各因素如何影响“意识形态同盟对信念的作用”，有如下回归：

$$B_{ia} = \beta_D D_i C_a + \beta_R R_i T_a + \gamma_D D_i + \gamma_R R_i + \epsilon_{ia}$$

如果受试者相信新闻标题为真 $B_{ia} = 1$ ，不确定则 $B_{ia} = 0.5$ ，不相信则 $B_{ia} = 0$ 。 D_i, R_i 分别表示民主党和共和党， C_a, T_a 表示标题是支持 Clinton 的还是 Trump 的。

Table 2
Ideological Alignment and Belief of News Headlines

	(1)	(2)	(3)
Democrat × Pro-Clinton	0.172*** (0.021)		
Republican × Pro-Trump	0.147*** (0.023)		
Aligned		0.161*** (0.016)	0.096 (0.140)
Aligned × Republican			0.000 (0.027)
Aligned × ln(Daily media time)			0.024*** (0.009)
Aligned × Social media most important			-0.031 (0.037)
Aligned × Use social media			-0.068 (0.050)
Aligned × Social media ideological segregation			0.147*** (0.046)
Aligned × Education			-0.004 (0.007)
Aligned × Undecided			-0.099*** (0.030)
Aligned × Age			0.001 (0.001)
<i>N</i>	10,785	10,785	10,785

Notes: This table presents estimates of a regression of a variable measuring belief of news headlines on the interaction of political party affiliation indicators and pro-Clinton or pro-Trump headline indicators. The sample includes all news headlines (both true and false) but excludes survey respondents who are Independents. “Social media most important” means social media were the respondent’s most important sources of election news. “Social media ideological segregation” is the self-reported share (from 0 to 1) of social media friends that preferred the same presidential candidate. “Undecided” is an indicator variable for whether the respondent decided which candidate to vote for less than three months before the election. Observations are weighted for national representativeness. Standard errors are robust and clustered by survey respondent. *, **, ***: statistically significantly different from zero with 90, 95, and 99 percent confidence, respectively.