The Dynamics of Peer-Produced Political Information During the 2016 U.S. Presidential Campaign

BRIAN C. KEEGAN, Department of Information Science, University of Colorado Boulder, United States

Wikipedia plays a crucial role for online information seeking and its editors have a remarkable capacity to rapidly revise its content in response to current events. How did the production and consumption of political information on Wikipedia mirror the dynamics of the 2016 U.S. Presidential campaign? Drawing on systems justification theory and methods for measuring the enthusiasm gap among voters, this paper quantitatively analyzes the candidates' biographical and related articles and their editors. Information production and consumption patterns match major events over the course of the campaign, but Trump-related articles show consistently higher levels of engagement than Clinton-related articles. Analysis of the editors' participation and backgrounds show analogous shifts in the composition and durability of the collaborations around each candidate. The implications for using Wikipedia to monitor political engagement are discussed.

CCS Concepts: • **Human-centered computing** → **Collaborative content creation**; *Wikis*; *Empirical studies in collaborative and social computing*; • **Information systems** → *Web log analysis*; • **Applied computing** → *Computing in government*.

Additional Key Words and Phrases: Wikipedia; election; Presidential campaign; enthusiasm gap; systems justification; Donald Trump; Hillary Clinton

ACM Reference Format:
Brian C. Keegan. 2019. The Dynamics of Peer-Produced Political Information During the 2016 U.S. Presidential Campaign. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 33 (November 2019), 20 pages. https://doi.org/10.1145/3359135

INTRODUCTION

How was information about political candidates produced and consumed on Wikipedia during and following the 2016 U.S. presidential campaign? Trump's victory surprised pollsters, pundits, politicians, and the plurality of voters who cast 2.8 million more ballots for Clinton. In the face of nearly unanimous consensus about Clinton's inevitability, what signals of collective political attention might have challenged this consensus? This paper argues Wikipedia's role as an information clearinghouse about current events can reproduce more general biases in media coverage over the course of a high-profile and prolonged political campaign.

Wikipedia is the open encyclopedia that "anyone can edit" and it occupies a central position in the ecology of online information production and consumption. Wikipedia's responses to current events provide important insights into how collective memory processes negotiate and privilege some remembrances into the historical record over others [15, 50]. Its prominence, mutability, and archival records make it a valuable resource for researchers to understand the flows of attention over the course of a political campaign. Wikipedia articles are rapidly revised in response to current

Author's address: Brian C. Keegan, Department of Information Science, University of Colorado Boulder, TLC 129 1045 18th St. Boulder, Colorado, 80309, United States.

2573-0142/2019/11-ART33 $15.00
https://doi.org/10.1145/3359135

events and the dynamics of how this content is produced and consumed can provide important insights into more general flows of collective attention in a complex media ecosystem.

We disentagle the central question of Wikipedia role in generating and disseminating political information around the 2016 campaign into two research questions: (1) how did the production and consumption of political information vary over time and (2) who were the editors involved in this activity? This study examines the dynamics of Trump's and Clinton's biographical and related articles and the users who revised them during the 2016 U.S. presidential campaign. Drawing on systems justification theory from political psychology and the enthusiasm gap in political science, two hypotheses are proposed to explain the variance in political information production and consumption. Data on 375,315 revisions to Clinton's biographical and related articles and 366,268 revisions for Trump were retrieved from the English Wikipedia and analyzed along with the contribution histories of the 2,211 most active editors.

The results find evidence of a early, significant, and sustained gap in enthusiasm favoring Trump's biographical and related articles over Clinton's articles. All articles showed peaks of activity corresponding to major events over the course of the campaign, but Trump's articles consistently had significantly more revisions, page views, editors, and content than Clinton's articles throughout the campaign. The composition of Trump's article collaborations saw a greater influx of users even though there was substantial overlap in the set of editors contributing to both. Engagement with this political content was likewise disruptive: active editors who began contributing during the campaign also had significant changes in behaviors compared to beforehand.

Although coarse aggregations and retrospective analyses of collective attention during an election campaign are tempting to mine for predictive features [16, 45], the correlations between sophisticated forecasts and Wikipedia's activity were negligible. Instead, these results suggest that the observed "enthusiasm gaps" in political information production and consumption are potential artifacts of Wikipedia reproducing the biases in more general media coverage. These findings have implications for political researchers, candidates, and Wikipedia editors for measuring the interactions in a complex media ecosystems and understanding their consequences on political processes.

BACKGROUND

The production and consumption of political information on Wikipedia is a matter of both empirical and general interest given both its prominence and history of as a site of political influence efforts. Additionally, Wikipedia's articles show unique collaborative dynamics around breaking news events. This section will review the literature on peer production and politics, systems justification behavior, and enthusiasm gap to pose two research questions and two hypotheses to explain the variance in activity across politicians' Wikipedia articles during a campaign.

Peer production and politics

Peer production is a mode of content creation involving decentralized task definition and execution, harnessing diverse motivations, and open distribution. Empirical studies of peer production projects like GNU/Linux and Wikipedia have focused on the mechanisms that underpin their organization and governance, the motivations of its volunteer contributors, and the quality of the products they generate [3, 4]. Wikipedia's success as the online encyclopedia that "anyone can edit" can be measured along multiple dimensions, but its prominence as one of the most trafficked sites on the web and increasingly integrated content into other social platforms is of particular concern for political actors. The English Wikipedia was a site of intensive editing and creative vandalism as early as the 2004 U.S. presidential election [7]. Larger institutions began to take note of Wikipedia's political risks and opportunities with China intermittently banning access [10] and revelations about

editing from corporations, intelligence agencies, and congressional staffers [6]. In one particularly memorable episode, in the days preceding Senator John McCain's public announcement that Alaska governor Sarah Palin would be his vice presidential candidate her Wikipedia article was quietly revised to remove negative information [11]. Politically-motivated editing of biographies, institutions, and events remains endemic to this day and across languages [18].

Wikipedia is unique among social media platforms in enforcing a strong editorial identity around verifiability and neutrality [30, 52]. Wikipedia is not a site for agitating or organizing movements in the way that Facebook and Twitter have been alleged [58]. While Wikipedia strongly espouses egalitarian principles, its organization and governance have been characterized as oligarchic through unequal participation and entrenched leadership that would seemingly predispose it to partisan capture [53]. However, Wikipedia's coverage of candidates, elections, and officeholders show a high level of accuracy when a relevant article exists, but there are many gaps and omissions, especially for older and less prominent topics [8]. The content Wikipedia articles about American politics indicates an initially liberal slant [44], but this bias disappears over time as articles employ language used by both Democrats and Republicans [20]. Wikipedia does not generate more politically biased content that encyclopedia articles written by experts [21]. However, negative facts introduced to politicians' Wikipedia biographies are more likely to be quickly removed than positive facts [34]. Given the prevalence and influence of alternative narratives [55] and misinformation in online discourse [1, 45], misinformation remains a persistent issue for Wikipedia editors to address [41, 49].

Article dynamics

Wikipedia's editors and content are extremely responsive to current events [36, 39]. The high-tempo online knowledge collaborations generating content about current and breaking news events are accomplished through a variety of practices: regenerating organizational structures from previous crises [40], adapting specific expertise to fill to emergent social roles [37], developing routines to handle regular but unpredictable events [38], and integrating individual events into larger narratives [59]. Although the supply and demand for quality content on Wikipedia is misaligned [61], political campaigns have realized Wikipedia's influence and proactively edit articles in advance of elections and announcements [11, 18]. The relative quantity and timing of how political information is produced and consumed over the course of a campaign will reveal emergent patterns for subsequent inductive analysis.

> **RQ1:** How does Wikipedia's production and consumption of political information vary during campaigns?

Editor dynamics

Predicting election outcomes with social media activity data is fraught [16, 33], but Wikipedia data can forecast overall turnout and changes in vote share for parties [63]. Wikipedians who self-identify with political ideologies do not exhibit polarized editing behaviors of avoiding collaboration with opponents or preferentially collaborating with allies [46]. Not all editors on Wikipedia are human: automated agents ("bots") play important roles for managing collaborations [17, 47]. But bots can also be abused to simulate popularity, spread misinformation, or amplify attacks in online social systems [14, 51]. The contributions of newcomers, ideologues, or bots to Wikipedia articles about political candidates during campaigns will be instructive for other social web platforms struggling with their own responses to selective exposure and computational propaganda [19].

RQ2: Who are the Wikipedia editors revising information about candidates during campaigns?

Systems justification

The "bandwagon effect" is a well-documented phenomenon in which citizens are influenced to support a candidate by virtue of others' judgments of her viability [54]. System justification theory predicts that actors rationalize future events so that more likely events are more desirable and less likely events are less desirable: attractive outcomes become less desirable as it becomes less likely ("sour grapes") and undesirable outcomes become more desirable as the likelihood increases ("sweet lemons") [32, 35]. Editors' political preferences do not appear to play a strong role in the articles they edit [20, 46], but systems justification theory predicts that editing an article becomes more desirable if the likelihood of the candidate winning increases becomes less desirable if the likelihood of the candidate winning decreases. While Wikipedians are motivated to edit by multiple intrinsic and extrinsic factors [48, 62], editors should allocate more effort toward articles they expect will persist (losing candidates' articles may be deleted after the election for non-notability) or become popularly valued (contributions will be viewed by many other people over the course of the election). Wikipedia editors are more likely to create articles for political challengers who have a better chance of defeating incumbents and timing the creation of these articles during periods of high collective attention [42]. Thus, election-related articles with more activity should reflect editors' judgments of that candidate's likelihood of winning as they volunteer their editing labor or seek information about articles where their effort and interest will not be wasted.

H1: The production and consumption of political information varies with candidates' likelihood of winning.

Enthusiasm gap

In democratic systems, the success of political candidates clearly hinges on voter participation: campaigns must motivate voters to cast ballots rather than staying home. Capturing voters' interest, attention, and enthusiasm can predict these turnout rates, although these are distinctive measures of electoral engagement [5]. But minor deviations in the same voters' turnout across different elections — also known as the *enthusiasm gap* — can produce large changes in first-past-the-post electoral systems [28]. Wikipedia may reproduce these enthusiasm gaps through the relative differences in the ability for each candidate to mobilize user engagement with its political content over the course of a campaign: a candidate motivating their "own" users to engage with their Wikipedia content relatively more than their opponent's editors may reflect differences in enthusiasm among the general electorate. These relative differences in enthusiasm could be attributed to efforts by voters to seek and consume information about candidates' policies to inform their own voting decision and complementary efforts by Wikipedia editors to ensure that the content of the information sought remains up-to-date for these viewer-voters [63]. Under this framework, campaigns should strive to generate news coverage or popular interest that motivates Wikipedia editors to update articles to update articles and views.

H2: Successful candidates mobilize greater cumulative Wikipedia activity than their opponents.

Label	*Description*	*Date*
A	Clinton announces candidacy	12 April 2015
B	Trump announces candidacy	16 June 2015
C	Super Tuesday primary elections	3 March 2016
D	"Acela" primary elections	29 April 2016
E	Trump secures nomination	27 May 2016
F	Clinton secures nomination	8 June 2016
G	Republican National Convention	11 July 2016
H	Democratic National Convention	28 July 2016
I	Access Hollywood controversy	7 October 2016
J	Election day	8 November 2016
K	Inauguration	20 January 2017

Table 1. Annotated events in Figures 1 and 2.

OUR APPROACH

Our approach identified relevant cases by identifying a parent category for each candidate, retrieved the complete revision histories of each of these articles, and used log data analysis and visualization techniques to generate quantitative features for description, statistical analysis, and visualization. While the project uses publicly-available and pre-existing trace data and is therefore exempt from human subjects review, we only report the names of Wikipedia editors in one instance when they received significant press coverage in order to preserve user privacy.

Methods

Wikipedia's MediaWiki platform serves a diverse array of data via its application programming interface (API). Event log data from article and user revision histories, membership from category structure, and markup of page content were retrieved from the English Wikipedia's MediaWiki API with custom Python scripts. Follow-on data cleanup, alignment, and aggregation used `pandas`[43], statistical tests and analyses used `scipy` [31], and data visualization used `matplotlib` [29]. The Jupyter Notebooks for replicating this data collection and analysis workflow are available online at https://github.com/brianckeegan/Wikipedia_2016_Election. Given the skewed distributions of activity and heteroskedastic variance in the samples examined throughout this paper, we employ a non-parametric Kruskall-Wallis H-test to test the null hypothesis that the medians of Trump's and Clinton's activity distributions are identical.

Data

We analyze both the *biographical articles* about the candidates ("Hillary Clinton" and "Donald Trump") as well as the *related articles* that are members or children of the Wikipedia categories for each candidate: 1,336 for Clinton and 949 for Trump. The time range spanning 1 January 2015 through 9 November 2017 was selected for detailed analysis, although the revision data goes back more than a decade beforehand for all articles. Clinton announced her candidacy via web video on 12 April 2015 and Trump announced his candidacy on 16 June 2015. The presidential election was held 17 months later on 8 November 2016. Data was collected for one year afterwards to examine article and editor dynamics through the administration's first year.

Revisions. Wikipedia preserves the user name (or IP address for unregistered editors), timestamp, comment, and content of every revision to an article. The revision history for each English Wikipedia

biographical and related article was retrieved from the Wikipedia's API using the "Revisions" endpoint.[1] This generated a corpus of 375,315 revisions for Clinton's biographical and related articles and 366,268 revisions for Trump. 57,944 (15.4%) of Clinton's revisions and 77,110 (21.1%) of Trump's revisions occurred over the course of the campaign.

Page views. Wikipedia records the daily frequency that every page is accessed by different web agents. Wikipedia introduced a new page view data endpoint in July 2015 and subsequently stopped publishing its page view data from late 2007 through 2015. As a result, we only report the data from the new pageview endpoint covering the campaign. Page view activity is broken down by user type and platform, but not by geographic location for privacy reasons. We report on the aggregated "all-agent" and "all-access" statistics.

Redirects. We acknowledge methodological criticisms of prior Wikipedia research by incorporating data about redirects [26]. For each candidates' biographical article, the list of pages in the main article namespace (0) that redirect to an article (*e.g.*, "Hillary Rodham Clinton" redirects to "Hillary Clinton") were retrieved from the Wikipedia API using the "Linkshere" endpoint.[2] There were a total of 78 redirects linking to the "Hillary Clinton" article and 67 redirects linking to the "Donald Trump" article. The same procedure was done for each of the related articles.

Editor contributions. From the initial set of 4,354 editors who contributed to Clinton's biographical article and 93,012 editors who contributed to Clinton's related articles and the 5,452 editors who contributed to Trump's biographical and 67,643 users who contribute to Trump's related articles, we extracted a sub-sample of *active editors* who made at least five unique revisions, contributed to more than two pages, and were active for more than one day. These thresholds were iteratively developed to exclude editors who had limited activity, narrow topical focus, or brief window of engagement. We prioritized this definition of active editors because of endogenous variance in the activity of inactive editors across candidates' articles could bias subsequent analyses. Of the 3,016 active editors on both candidate's biographical and related articles, 2,467 contributed to Clinton articles and 2,969 to Trump articles. 805 editors making more than 500 revisions in a single month were skipped given their tendency to be "bot" or "cyborg" editors using automated tools for anti-vandalism patrols generating anomalously high levels of activity. For each of the remaining 2,211 editors, their contribution histories for the period from 1 January 2014 through 9 November 2017 were retrieved from the Wikipedia API using the "Usercontribs" endpoint.[3]

FiveThirtyEight probabilities. The political website FiveThirtyEight published daily probabilities of Clinton and Trump winning the election between 8 June 2016 and 8 November 2016. The data for the "chance of winning" under the "Polls-plus forecast" were retrieved as a measure of systems justification behavior.[4] Because polls are sampled over multiple days, the effects of major events take several days to be reflected in the daily probabilities. Consequently, the probabilities are coarsened from days to weeks and compared with corresponding weekly Wikipedia activity.

RESULTS

The results of the analysis to address the two research questions and two hypotheses are presented in the three sub-sections below about the biographical article dynamics, related article dynamics, and editor dynamics.

[1] https://www.mediawiki.org/wiki/API:Revisions
[2] https://www.mediawiki.org/wiki/API:Linkshere
[3] https://www.mediawiki.org/wiki/API:Usercontribs
[4] https://projects.fivethirtyeight.com/2016-election-forecast/

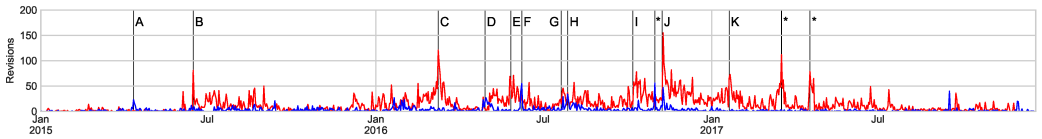

Fig. 1. The daily revisions to the Clinton (blue) and Trump (red) biographical articles.

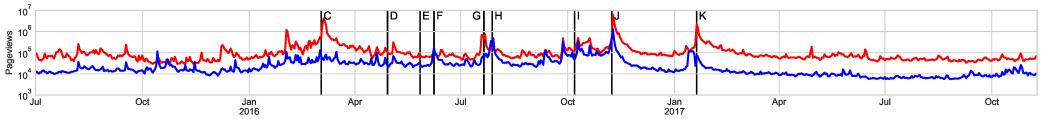

Fig. 2. The daily page views to the Clinton (blue) and Trump (red) biographical articles.

Biographical article dynamics

RQ1 asked "How does Wikipedia's production and consumption of political information vary during campaigns?" Focusing on each candidate's biographical article, we analyzed data about changes in the revisions, page size, page views, and page protections over the course of the campaign.

Revisions. Over the history of their articles (through 9 November 2017), Donald Trump's article received an average of 4.2 ± 9.6 (max. 155) revisions per day while Hillary Clinton's article received an average of 4.2 ± 5.8 (max. 172) revisions per day. The H-test for daily revisions was 105.66, $p < 0.001$). Over the course of the campaign (1 June 2015 through 8 November 2016), Trump's article received an average of 16.5 ± 15.9 (max. 121) revisions per day while Clinton's article received an average of 4.2 ± 5.9 (max. 56) revisions per day ($H = 295.96, p < 0.001$). Figure 1 visualizes the daily revisions to each candidate's biographical article over the course of the campaign. Ten major events are annotated with descriptions given in Table 1. Three additional bursts of activity are annotated with stars ($*$) that do not correspond to major exogenous events, but endogenous bursts of revisions from single users making many sequential changes.

Page size. Over the history of each candidate's biographical articles, Clinton's article had a median size of 167 kB compared to Trump's median size of 52 kB ($H = 510.35, p < 0.001$). While Clinton's article was significantly larger than Trump's preceding the 2016 campaign (241 kB vs. 106 kB), Trump's article more than tripled in size over the course of the campaign, growing to 342 kB on the day of the election compared to Clinton's 280 kB. Clinton's article had a median size of 272 kB, which was significantly smaller than Trump's median size of 286 kB during the campaign ($H = 45.08, p < 0.001$).

Editors. Clinton's article had more cumulative unique editors (3,652) than Trump (3,432) when he announced his campaign (point *B*). Between the start of Trump's campaign and Election day, the cumulative number of unique editors on Trump's article grew to 4,773 editors (39.6% increase) compared to Clinton's article growing to 4,145 editors (13.8% increase). The composition and differences in these editor sets are explored in more detail in a later section.

Page views. Clinton's article (and redirects) received a median of 41,787 page views compared to Trump's 159,283 page views ($H = 460.6, p < 0.001$). Clinton received 19,535,002 page views and Trump received 73,116,431 over the course of the campaign. Figure 2 visualizes the number of daily page views to the Clinton and Trump articles. The same 10 events are annotated with descriptions given in Table 1. There were only 3 date ranges when Clinton's page view activity surpasses Trump's: mid-October 2015, early June 2016 after she clinched the Democratic nomination, and during the Democratic National Convention in July.

FiveThirtyEight probabilities. To test H1 that changes in a candidate's likelihood of winning varies with the production and consumption of political information, the weekly percentage changes in each candidate's page view, revision, and size were correlated against the percentage change in their FiveThirtyEight chances. The correlation of changes in probabilities of winning with Clinton's page views ($r = 0.019$), revisions ($r = 0.018$), and size ($r = 0.41$) can be contrasted with Trump's page views ($r = 0.053$), revisions ($r = 0.006$), and size ($r = -0.110$).

Related article dynamics

This section tests H2 that the successful candidate mobilized greater Wikipedia activity than the opponent by examining the dynamics of information production and consumption unfolded for the candidates' *related articles*. After removing the 16 related articles present in both candidates' sets, there are 1,320 Clinton-related articles and 933 Trump-related articles. There is a consistent "enthusiasm gap" among editors of Clinton's and Trump's related articles. Figure 3 plots the cumulative activity for five distinct behaviors over the course of the 2016 campaign. Examining the activity for the set of all related articles (solid lines) as well as the set of new articles created over the course of the campaign (dashed lines), there is a common pattern of all and new Trump-related articles having higher levels of activity than Clinton-related articles by Election Day (point *J*).

New article creation. Most related articles predate the start of the campaign, but many were created after the campaign began as Wikipedia editors fill in additional details about the campaign, people, events, and controversies beyond the scope of the candidates' biographies. In the period starting after 1 January 2015, 84 Clinton-related articles and 692 Trump-related articles were created. This disparity in new article creation is not surprising: Clinton's previous history of high-profile public service meant much of this content already existed while corresponding content needed to be created for Trump over the course of his campaign. Examples of these newly-created related articles for each candidate are given in the next sub-section.

Revisions. Among all (new) related articles between 1 January 2015 and 7 November 2016, Clinton's articles received 43,010 (3,602) revisions and Trump's articles received 55,232 (26,557) revisions. From January 2015 until Trump's June 2015 announcement (point *B*), Clinton's all (new) related articles had accumulated an average of 5.9 (1.7) revisions compared to Trump's 4.4 (0.2) revisions ($H_{all} = 410.9, p_{all} < 0.001, H_{new} = 39.2, p_{new} < 0.001$). Clinton's most-revised *pre-existing* articles in the pre-campaign phase include her 2016 campaign article (518), Secretary of Defense Ash Carter (207), and NATO bombing of Yugoslavia (174). The most-revised *new* articles about Clinton in the pre-campaign phase include the derogatory book *Clinton Cash* (57), Executive Order 13175 (26), and *Clinton: The Musical* (15). Trump's most-revised *pre-existing* articles in the pre-campaign phase include Miss Universe (684), The Apprentice's 14th season (437), and GOP primary opponents like Ben Carson (350), Mike Pence (168), and Rick Perry (142) who would end up joining his administration's Cabinet. Trump's most-revised *new* articles in the pre-campaign phase include the Russian Internet Research Agency (56), internet meme Pepe the Frog (36), and future-former Deputy Attorney General Sally Yates (20).

Over the course of the campaign itself, from Trump's 2015 announcement through 7 November 2016 (to exclude the bursts of activity surrounding the day of the election), Clinton's all (new) related articles had accumulated an average of 28.2 (41.1) revisions compared to Trump's 55.0 (38.2) revisions ($H_{all} = 267.0, p_{all} < 0.001, H_{new} = 82.6, p_{new} < 0.001$). Clinton's most-revised *pre-existing* articles over the campaign include her campaign article (1,520), her political positions (1,005), her Vice President pick Tim Kaine (964), and her senior advisor Huma Abedin (896). The most-revised *new* articles about Clinton over the campaign include Gold Star parents Khizr and Ghazala Khan (475), the derogatory book *Hillary's America* (365), the vice presidential candidate

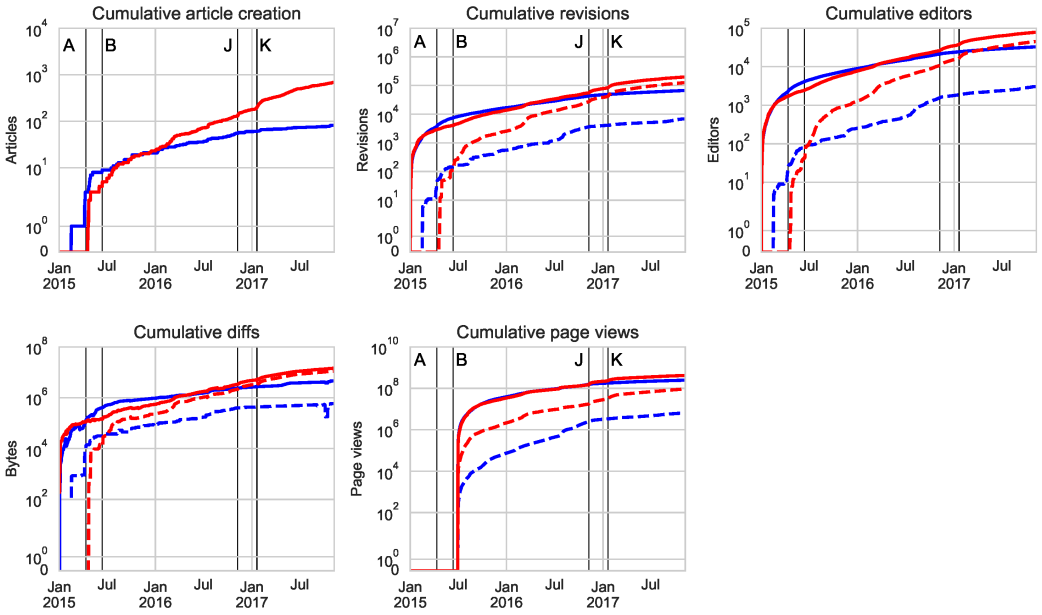

Fig. 3. Cumulative activity on related articles for Clinton (blue) and Trump (red). Solid lines for *all* child articles and dashed lines for *new* articles created after 1 January 2015. Page view data is only available after 1 July 2015.

selection article (240), and Clinton Foundation-State Department controversy (225). Trump's most-revised related articles over the campaign are dominated by the new articles and include a list of campaign endorsements (5,135), campaign article (4,775), Miss Universe (2,216), political positions (1,962), and sexual assault allegations against him (1,517).

In the year since the 8 November 2016 election, revision activity on Trump-related articles accelerated. Clinton's all (new) related articles had accumulated an average of 19.2 (38.8) revisions compared to Trump's 154.6 (144.1) revisions ($H_{all} = 351.5, p_{all} < 0.001, H_{new} = 11.9, p_{new} < 0.001$). For both Clinton and Trump, the most-revised articles in this post-campaign phase are similar with a focus on contemporary events. Clinton's most-revised articles after the campaign include her book *What Happened* (504), a list of her presidential non-political endorsements (444), and her campaign article (442). Trump's most-revised articles after the campaign include his political appointments (4,331), the Women's March (4,238), the Unite the Right rally (3,432), and Executive Order 13769 (3,179).

Editors. Trump's related articles attracted significantly more unique editors over the course of the campaign than Clinton's related articles. Among all (new) related articles from 1 January 2015 through 7 November 2016, Clinton's articles had 20,460 (1,067) unique editors and Trump's articles had 19,505 (6,492) unique editors. From January 2015 until Trump's June 2015 announcement, Clinton's all (new) related articles had accumulated an average of 5.0 (6.8) editors compared to Trump's 11.1 (8.5) editors ($H_{all} = 45.4, p_{all} < 0.001, H_{new} = 0.6, p_{new} = 0.44$). Over the course of the campaign itself, from Trump's 2015 announcement through 7 November 2016, Clinton's all (new) related articles had accumulated an average of 14.1 (49.2) editors compared to Trump's 18.5 (50.1) editors ($H_{all} = 129.6, p_{all} < 0.001, H_{new} = 4.3, p_{new} = 0.037$). In the year since the 8 November 2016 election, Clinton's all (new) related articles had accumulated an average of 10.6 (47.3) editors

compared to Trump's 47.3 (38.0) editors ($H_{all} = 240.5, p_{all} < 0.001$, $H_{new} = 3.8, p_{new} = 0.053$). Between June 2015 and November 2016, the editors of all Trump articles made an average of 2.3 revisions per editor, significantly more than the 1.8 average revisions per editor on all Clinton articles ($H = 59.3, p < 0.001$).

Size. Trump's related articles accumulated significantly more content than Clinton's related articles over the course of the campaign. Among all (new) related articles from 1 January 2015 through 7 November 2016, Clinton's articles accumulated 2,389 (397) kilobytes and Trump's articles accumulated 3,315 (2,137) kilobytes of content. From January 2015 until Trump's June 2015 announcement, Clinton's all (new) related articles had accumulated an average of 320.9 (372.8) bytes of content compared to Trump's 155.1 (19.1) editors ($H_{all} = 97.9, p_{all} < 0.001$, $H_{new} = 24.2, p_{new} < 0.001$). Over the course of the campaign itself, from Trump's 2015 announcement through 7 November 2016, Clinton's all (new) related articles had accumulated an average of 1,573 (4,360) bytes of content compared to Trump's 3,406 (3,070) bytes ($H_{all} = 0.5, p_{all} = 0.48$, $H_{new} = 81.9, p_{new} < 0.001$). In the year since the 8 November 2016 election, Clinton's all (new) related articles had accumulated an average of 1,740.1 (2,286) bytes of content compared to Trump's 12,069 (13,135) bytes ($H_{all} = 147.5, p_{all} < 0.001$, $H_{new} = 21.3, p_{new} = 0.053$). Between June 2015 and November 2016, the editors of all Trump's related articles added an average of 179 bytes per editor, significantly more than the 117.8 average bytes per editor on all Clinton articles ($H = 34.0, p < 0.001$).

Page views. Trump's related articles attracted significantly more page views over the course of the campaign than Clinton's related articles. Among all (new) related articles from 1 July 2015 through 7 November 2016, Clinton's articles received 153.5 million (2.47 million) page views and Trump's articles accumulated 147.8 million (17.3 million) page views. Over the course of the campaign, Clinton's all (new) related articles had accumulated an average of 123,519 (30,498) total page views compared to Trump's 163,686 (26,015) total page views ($H_{all} = 352.3, p_{all} < 0.001$, $H_{new} = 81.0, p_{new} < 0.001$). In the year since the 8 November 2016 election, Clinton's all (new) related articles had accumulated an average of 75,875 (50,004) page views compared to Trump's 300,260 (114,707) page views ($H_{all} = 37.9, p_{all} < 0.001$, $H_{new} = 1.1, p_{new} = .30$).

Editor dynamics

RQ2 asked "Who are the editors revising information about candidates during campaigns?" The attributes of the editors who revised both candidates' biographical and related articles are analyzed for their experience and characteristics, changes in commitment by cohort, changes in overlapping membership, and differences in the behavior of active users after they began contributing during a political campaign.

Registered and bot editors. How did the makeup of the editor pools on each article change over time? The top subplot of Figure 4 plots the (cumulative) fraction of revisions from bot editors contributing to each candidate's biographical (solid) and related (dashed) article(s). Bot editors were responsible for a substantially larger fraction of revisions on Trump's biographical article than Clinton's article for most of its history, although this difference rapidly diminished over the course of the campaign (points *A* and *J*). Conversely, bots were responsible for a larger fraction of revisions to Clinton's related articles than Trump's.

The bottom subplot of Figure 4 visualizes the (cumulative) fraction of revisions from registered editors contributing to each candidate's biographical (solid) and related (dashed) articles. Registered editors are more likely to be aware of editing norms, editorial consensus on a given page, and generally less likely to engage in vandalism. Clinton's article consistently had more revisions from registered editors until the campaign begins (point *A*). Each article's early history shows a rise

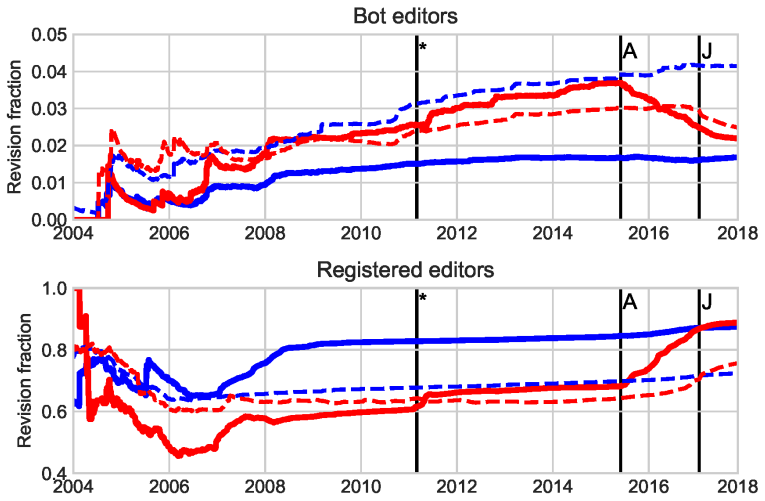

Fig. 4. Fraction of revisions coming from bots and registered users to the Clinton (blue) and Trump (red) biographies (solid) and related articles (dashed).

and decline in the fraction of registered editors, followed by stabilization. Trump's article had fewer registered editors than Clinton's article throughout its history, but this fraction increased substantially over the course of the campaign. Clinton's related articles again had a higher fraction of revisions from registered editors than Trump's related articles preceding the 2016 campaign, but this difference disappeared by November 2016.

User characteristics. What kinds of editorial experience did the active editors of Clinton and Trump's biographical and related articles over the course of the campaign have beforehand? The characteristics of the active editors who exclusively edit Clinton's ($N = 47$) and Trump's ($N = 549$) articles (henceforce, the *partisan editors*) are compared against the active editors who contribute to both candidates' articles ($N = 2,420$ and henceforth, *bipartisan editors*). The Jaccard score for these editors is 0.802.

Account age is the number of days elapsed between an editor's account creation and their first contribution to the corpus of either candidate's articles. There are neither significant differences ($H = 1.57, p = 0.21$) between the account ages of active editors contributing only to Trump (777 ± 931) and only Clinton articles (1025 ± 1170) nor between ($H = 1.12, p = 0.29$) the bipartisan editors (893 ± 1033).

Edit count is the number of revisions made by the editor over their whole history. There are no significant differences ($H = 0.0, p = 0.97$) between the partisan Clinton ($14,644 \pm 27,805$) and Trump ($13,787 \pm 34,025$) editors. Bipartisan editors ($44,104 \pm 166,617$) have significantly greater counts than the partisan Clinton ($H = 14.8, p < 0.001$) and Trump ($H = 162.6, p < 0.001$) editors.

Blocked fraction is the number of active editors whose accounts are blocked. A simple logistic regression model using editor type (Clinton partisan, Trump partisan, and bipartisan) as a predictor identifies a marginally significant ($\beta = 0.37, p = 0.038$) effect of Trump partisans (7.8%) having a greater likelihood of being blocked than bipartisan editors (5.4%) but no significant difference ($\beta = 0.46, p = 0.36$) between Clinton partisans (8.7%) and bipartisan

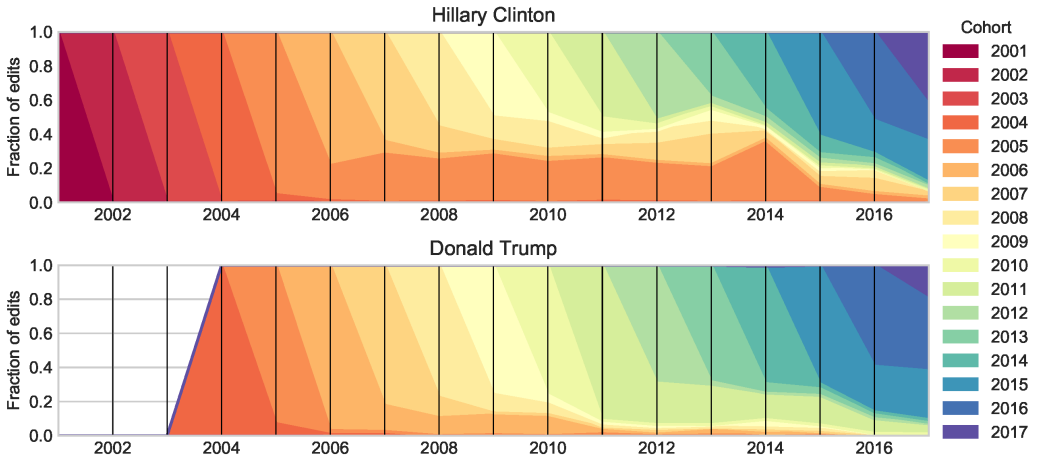

Fig. 5. Fraction of revisions made to the Clinton (*top*) and Trump (*bottom*) biographical articles by annual editor cohorts.

editors. The statistical significance of this effect is not robust under model specifications using additional features.

Topical focus is a fraction with the revisions an editor made to either Clinton or Trump related articles divided by their total edit count: values closer to 1 indicate an editor focused on articles related to the candidates to the exclusion of other content on Wikipedia. There are no significant differences ($H = 0.03, p = 0.863$) in the topical focus between the partisan Clinton (5.7% ± 15.4%) and Trump (8.2 ± 61.8%) editors. However, bipartisan editors had a significantly ($H = 78.2, p < 0.001$) greater topical focus on Trump articles (1.9% ± 5.6%) than on Clinton articles (0.83% ± 2.8%).

Editor cohorts. We performed a "time-aware analysis" [2] of revisions made to each candidate's article by editors who made their first revisions to the article across different years. Figure 5 plots the fraction of revisions made to the article in each from each annual cohort of editors. Both candidates' articles show high levels of editor attrition year-over-year initially as the editors who were active in the previous year do not re-engage in editing the article in the subsequent year. While ownership where editors assert control over content they have authored has been observed in Wikipedia [22, 57], this turnover in revision activity across cohorts suggests limited territoriality and enthusiasm for contributing to political content.

There is a notable exception: the cohort of editors who began contributing to Clinton's article in 2005 shows a comparatively high level of persistence, contributing more than 20% of revisions from 2005 through 2013. This is largely attributable to a single user, "Wasted Time R", whose hundreds of annual contributions for more than a decade focused on policing Clinton's article alongside his contributions to other politicians' articles including John McCain, Mitt Romney, and Joe Biden [12, 56].

For the revisions made over the course of the campaign in 2016, the largest fractions came from editors in the 2016 cohort (51.8% for Clinton, 59.2% for Trump) followed by the 2015 cohort (19.4% for Clinton, 26.7% for Trump). The next largest cohort for Clinton was from 2007 (7.7% of revisions in 2016), corresponding to editors who first became engaged during Clinton's 2008 presidential campaign. The next largest cohort for Trump was from 2011 (7.7% of 2016 revisions), corresponding to editors who first became engaged during Trump's discredited allegations about

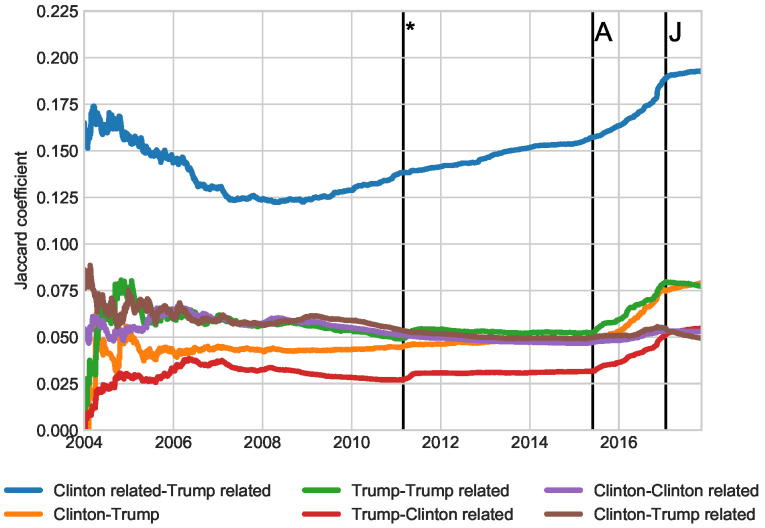

Fig. 6. Cumulative Jaccard coefficient among editor sets.

President Obama's birth certificate. In 2017, the most active cohorts for Clinton was the 2017 cohort (42.4%), followed by the 2015 (24.1%) and 2016 (21.6%) cohorts. Trump's 2016 cohort had the highest level of 2017 activity (42.1%), followed by the 2015 (28.6%) and 2017 (19.8%) cohorts.

Editor overlaps. How did the composition of the editors collaborating on Clinton and Trump's articles change over time? Figure 6 plots the cumulative Jaccard coefficient for the permutations of the editors contributing to the biographical as well as related articles for both candidates from 2004 through 2017. Jaccard coefficients closer to 1 indicate more overlap between the sets of editors. The largest overlap is observed for the editors contributing to the related articles for both candidates (blue), followed by the overlap between the Clinton and Trump biographical articles (orange), the Trump biographical and Trump-related articles (green). These articles saw an accelerated rise over the course of the campaign (points *A* and *J*), but the overlap between the Clinton biographical and related articles (purple) does not change as dramatically over the course of the campaign.

Figure 7 plots the cumulative fraction of revisions to the candidates' made by editors from each of these different editor sets over time. These fractions exceed 100% because of the overlaps among the sets (as described in the previous paragraph), so editors from both sets of related articles are also reported. Editors of Clinton's related articles were responsible for the majority of revisions (67.4%) to her biographical article over time before the campaign began, which rose to 71.2% by the end of the campaign. The largest set of editors to Trump's biographical article before the campaign came from Trump's (43.7%) and Clinton's related articles (41.4%) rising to 67.7% and 62.8% of revisions (respectively) at the end of the campaign.

The composition of collaborators contributing to Trump's biographical article change dramatically at three distinct points in time. The first discontinuity was in 2011 (annotated with $*$ in Figure 7) during Trump's discredited accusations about Obama's birth certificate. The fraction of revisions from Trump and Clinton related articles increased substantially reflecting an influx of editors to his biographical article. The second discontinuity happened following the announcement of his 2016 candidacy (point *A*). The shift in the composition of the editors on his biographical article accelerated as contributors to Trump related, Clinton related, and Clinton's biographical

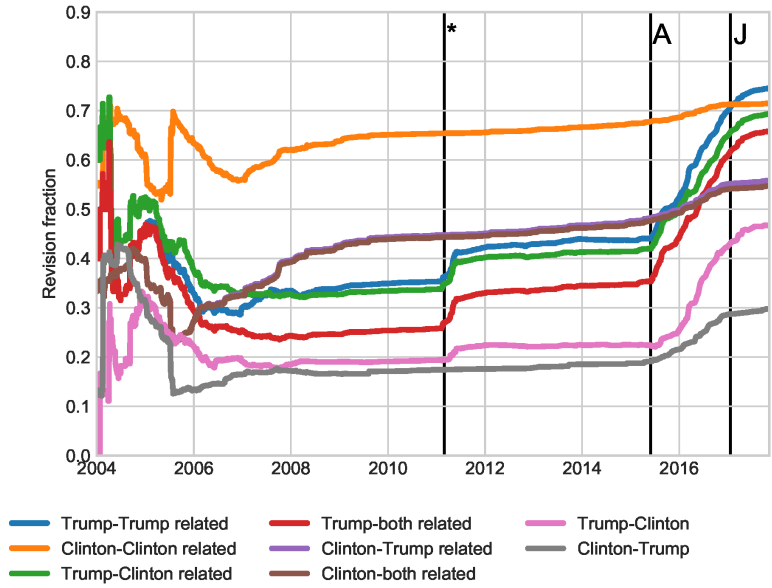

Fig. 7. Revision fraction on biographical articles among editors sets.

	Before	*After*	*H-test*
Comments	1852 ± 5294	5399 ± 11952	215.4 *
Dates active	153 ± 181	264 ± 214	246.8 *
Entropy	4.7 ± 1.8	5.9 ± 1.8	198.5 *
Latency (days)	0.39 ± 5.4	0.11 ± 1.5	24.0 *
Pages created	183 ± 1437	497 ± 1622	77.4 *
Pages edited	944 ± 4098	3096 ± 8033	217.5 *
Namespaces	6.5 ± 3.6	8.4 ± 4.2	103.8 *
Revisions	2187 ± 5772	5980 ± 12479	226.3 *
Rev. size (bytes)	69.7 ± 384.8	60.4 ± 125.1	0.7

Table 2. Average user contribution behavior before and after first revision to campaign article ($*, p < 0.001$)

article. The third discontinuity corresponds with Trump's victory at the end of the campaign (point *J*) as the contributions from these other sets of editors stabilized above 60% of the total revisions.

Predecessor and successor collaborations. How did active editors' behavior change after their first revision to a campaign article? The contribution histories from 1 January 2014 through 9 November 2017 for 1,075 active users who made their first revision to a candidate's biographical or related article after 1 January 2015 were retrieved and analyzed to compare their contribution history *after* their first revision to a Clinton or Trump article to their contributions *before* this first revision. This analysis uses an active editor's first contribution to these candidates' biographical or related articles as a discontinuity to test the changes in behavior before and after this expression of interest in editing political content during a campaign. Table 2 summarizes the average active editors' contribution behavior before and after their first revision to a candidate's biographical or related article. Following their first "political" contribution, active editors make significantly

more comments, are active on more days, increase the entropy of revisions made across articles, reduce the latency between successive edits, create more pages, edit more pages, contribute in more namespaces, and make more revisions in the period afterwards. The average size of their individual revisions does not change significantly.

DISCUSSION

How was information about political candidates produced and consumed on Wikipedia during and following the 2016 U.S. presidential campaign? We explored this research question through three levels of analysis: the dynamics of Clinton and Trump's biographical articles, the dynamics of their related articles, and the dynamics of the editors who contributed to them during the campaign. This analysis was motivated by (1) providing a descriptive understanding of these article and editor dynamics, (2) understanding whether systems justification behavior explained how editors shifted their labor with candidates' likelihood of winning, and (3) measuring the enthusiasm gap on the part of users producing and consuming information about each candidate. While the goal of this study was not to forecast the outcome of the election, there was significant and diverse evidence that Wikipedia editors allocated significantly more attention towards Trump than towards Clinton during the campaign.

Biographical article dynamics. The biographical articles showed high levels of information production and consumption activity responded to major events over the course of the 2016 campaign with a focus on announcements, primary elections, and the party conventions. Clinton's article was in much better shape at the start of the campaign but Trump's attracted significantly more revisions and unique editors, overtook Clinton's article in size by March 2016 ultimately tripling in size. Perhaps most critically, Trump's biographical article received significantly more page views than Clinton's article indicating a greater demand for information about this candidate. Weekly changes in the production and consumption of biographical information had very weak relationships with the changing likelihoods of winning the election according to FiveThirtyEight's forecasting model. The absence of even correlational evidence between changes in each candidate's likelihood of winning and Wikipedia article activity does not support our hypothesis that systems justification behavior explains changes in Wikipedia editors' and readers' engagement with content related to the presidential candidates.

Related article dynamics. The gap in information production and consumption on the candidates' biographical articles likewise extended to the related articles about each candidate. Over the course of the campaign between January 2015 and November 2016, Wikipedians created more new articles, made more revisions, generated larger collaborations, made larger articles, and viewed more articles about Trump than Clinton. A similar and significant gap in information production and consumption favoring Trump over Clinton also unfolded for new articles created during the campaign. While these differences in related article dynamics could be ascribed to "incumbency effects" (there is little to add to pre-existing articles), the gaps in information production and consumption about *new* articles (likely to have more similar demands for production and consumption) instead suggest much greater attention and/or enthusiasm for Trump-related content on the part of Wikipedia editors and readers. This does not appear to be overtly partisan editing: some of the most revised new articles during the campaign reflected controversies implicating both candidates.

Editor composition dynamics. The composition of the editors on the candidates' biographical and related articles likewise showed substantial changes over the course of the campaign. The fraction of revisions on Trump's articles coming from bots decreased over the course of the campaign as more registered editors began to contribute (Figure 4). The number of overlapping editors between

Clinton and Trump's articles increased substantially over the course of the campaign (Figure 6) and the contributions from these overlapping editors made up a majority of the revisions made to these articles by the end of the campaign (Figure 7). A time-aware analysis of the user cohorts contributing to the articles found significant turn-over year-over-year in the editors contributing to each candidate's biographical article, but some select editors do persist in contributing to these articles for multiple years. The most active partisan and bipartisan editors of these articles do not show significant differences in their account ages or tendency to be blocked, but the bipartisan editors contributing to both candidates' articles tended to have significantly larger edit counts as well as a significantly greater topical focus on contributing to Trump-related articles. Finally, there were significant differences (Table 2) among active editors who began editing during the campaign between their contribution behavior before and after their first edit to a candidate's biographical or related article.

Implications and explanations

Wikipedia's coverage of major political campaigns is of interest to the social computing and CSCW communities for several reasons. First, the motivations for contributing to peer production projects are typically assumed to be stable over time, but rational choice theories of voters' information seeking imply that exogenous events in the campaign could drive major shifts in the composition of participants and their motivations for participating. This suggests the need for theorizing, methods, and strategies for capturing these shifts in participant motivations and their potential for conflict as well as socialization and organizational learning. Second, Wikipedia is increasingly imbricated within other platforms' content moderation systems. Facebook and YouTube both ingest and amplify Wikipedia content to combat their own disinformation. These forms of interoperability introduce significant risks of blowback as bad faith agents on other platforms migrate their influence efforts to Wikipedia as a newly central information hub and diminish its capacity to police disinformation. Third, Wikipedia's relative resilience to 2016-era disinformation campaigns compared to other social media platforms suggests there may be valuable strategies for other platforms to adopt. For example, moderator attention on Wikipedia is concentrated around common artifacts while moderator attention on Facebook or Twitter is fragmented across personalized newsfeeds. Wikipedia also delegates the authority to moderate content to a much wider set of users, who draw on clear rules and precedents, unlike the seemingly *ad hoc* and automated moderation strategies employed by traditional platforms.

The "systems justification" and "enthusiasm gap" theories are distinct and irreconcilable mechanisms for explaining voter behavior itself. Under a systems justification framework, voters are *extrinsically* motivated turn out for a candidate because of their beliefs about the candidate's likelihood of winning: they act on their beliefs about other voters' behavior. Under an enthusiasm gap framework, voters are *intrinsically* motivated to turn out for a candidate to influence the election outcome: a campaign generates sufficient motivation to overcome some threshold or inertia among voters to get them to turn out and vote. The lack of observed support for H1 about Wikipedians' systems justification behavior are compatible with prior findings about Wikipedians' generally non-partisan editing behavior [8, 20, 46]. This non-partisanship combined with the limited predictive power of Wikipedia activity on election outcomes [18, 63] suggests the enthusiasm gap is not a persuasive mechanism either.

Two alternative explanations for the differences in political content production and consumption are (1) Wikipedians' responsiveness to media coverage and (2) Wikipedians' biased population. The observed enthusiasm gaps may not be signals of partisan enthusiasm, but rather an echo of the biases in the the quantity and quality of media coverage about each candidate. By one estimate, Trump received $4.96 billion in "free" media coverage during the 2016 campaign compared to

Clinton's $3.24 billion across online news, broadcast, blogs, Twitter, and print media [9, 24]. On the consumption side, greater coverage of Trump than Clinton could drive more information-seeking behavior by Internet users towards Trump, some large fraction of whom are likely to end up on Wikipedia. On the production side, Wikipedia editors operating under norms about neutrality, reliable sources, Biographies of Living People as well as routines for responding to current events could use the proportionally greater amount of "raw material" from the media's coverage of Trump to generate more content, reproducing the same coverage biases in Wikipedia.

Wikipedia editors are also not representative of the population at large and these in turn introduce systematic biases into Wikipedia's coverage of topics [23, 25]. The advantages Trump saw along all dimensions of information production and consumption could be echoes of well-documented biases Wikipedia has to topics related to women [44, 60]. It is difficult to generalize from a single case study, but a closer content analysis or examination of editors' contribution histories could reveal substantive biases in expertise, interests, and motivations to contribute—or not—to candidates with non-dominant identities.

Limitations and future work

These analyses were primarily descriptive and used revision activity data from the English Wikipedia. The analysis did not examine page protection events, which have the effect of preventing unregistered and new editors from making contributions [27]. Although both candidate's articles may have been prone to similar levels of vandalism attempts and the analysis of editor dynamics focused on active users who are less likely to be prevented from editing, any differences in the type or duration of page protection spells between candidates' articles would still bias the central constructs of editor and revision activity. Because the content of every revision to Wikipedia articles is archived, an analysis of the content dynamics would enable a better understanding of framing and agenda-setting processes and the evolution of the structure and prominence of hyperlinked concepts around each candidate. More critically, a content analysis would also highlight Wikipedia's susceptibility to "fake news" by examining the prevalence and durability of links to known sources of misinformation. Interviewing active contributors to these articles using retrospective methods could also provide rich insights into practices, contexts, and motivations that are not apparent in the event log data used here. Wikipedia's openness and prominence makes it an obvious target for computational propaganda efforts, but its apparent resilience may offer important lessons for other socio-technical systems like Facebook or Twitter to empower users to help govern the quality of online information. Finally, tools for tracking the online media like Media Cloud could substantiate the hypothesized relationship between the volume of media coverage and Wikipedia content production and consumption [13].

CONCLUSION

Wikipedia's centrality as an online reference provides important context about the supply and demand for information during a political campaign. This analysis demonstrated a significant and sustained enthusiasm gap for creating, editing, and viewing content favoring Trump over Clinton existed on Wikipedia. In almost every metric examined, content related to Trump attracted more editors, more revisions, more new articles, larger articles, and more page views than content related to Clinton. These differences emerged relatively early in the primaries, were amplified over the course of the campaign, and are not explained by changes in the likelihood of each candidate winning. Instead, this enthusiasm gap in political information production and consumption may be an effect of Wikipedia editors responding to coverage disparities in the broader media ecosystem rather than partisan agendas. These findings have implications for researchers examining the effects of political communication, candidates and their campaigns understanding political information

seeking behavior of large populations, as well as the Wikipedia community reflecting on its responsibilities within a complex media ecosystem influencing political outcomes.

ACKNOWLEDGMENTS

Thank you to Simon Munzert and the participants of the 2017 workshop on "Citizen and Elite Activity on Wikipedia" at the Hertie School of Governance for their feedback. Posters of this work were presented at the 2018 AAAI International Conference on Web and Social Media (ICWSM) and the 2018 International Conference on Computational Social Science (IC^2S^2). Thank you to the Wikimedia Foundation and its Research team for supporting the open technical infrastructures for research like this.

REFERENCES

[1] Hunt Allcott and Matthew Gentzkow. 2017. *Social Media and Fake News in the 2016 Election.* Working Paper 23089. National Bureau of Economic Research. https://doi.org/10.3386/w23089
[2] Samuel Barbosa, Dan Cosley, Amit Sharma, and Roberto M. Cesar, Jr. 2016. Averaging Gone Wrong: Using Time-Aware Analyses to Better Understand Behavior. In *Proc. WWW 2016.*
[3] Yochai Benkler. 2006. *The Wealth of Networks.* New Haven, CT, USA.
[4] Yochai Benkler, Aaron Shaw, and Benjamin Mako Hill. 2015. Peer production: A form of collective intelligence. *Handbook of collective intelligence* 175 (2015).
[5] Samuel J. Best, Kristin Johnson, and Brian S. Krueger. 2012. Americans' Interest and Enthusiasm in Election Campaigns. *Public Opinion Quarterly* 76, 4 (2012), 783–814.
[6] John Borland. 2007. See Who's Editing Wikipedia – Diebold, the CIA, A Campaign. *WIRED* (2007).
[7] Sarah Boxer. 2004. Mudslinging Weasels Into Online History. *The New York Times* (Nov. 2004).
[8] Adam R. Brown. 2011. Wikipedia as a Data Source for Political Scientists: Accuracy and Completeness of Coverage. *PS: Political Science & Politics* 44, 2 (2011), 339–343. https://doi.org/10.1017/S1049096511000199
[9] Philip Bump. 2017. Assessing a Clinton Argument That the Media Helped to Elect Trump. *The Washington Post* (2017). https://www.washingtonpost.com/news/politics/wp/2017/09/12/assessing-a-clinton-argument-that-the-media-helped-to-elect-trump/?utm_term=.1aa449cc7c64
[10] Noam Cohen. 2007. Chinese Government Relaxes Its Total Ban on Wikipedia. *The New York Times* (2007).
[11] Noam Cohen. 2008. Don't Like Palin's Wikipedia Story? Change It. *The New York Times* (Aug. 2008).
[12] Eve Fairbanks. 2008. Wiki Woman. *The New Republic* (April 2008). https://newrepublic.com/article/63288/wiki-woman
[13] Robert Faris, Hal Roberts, Bruce Etling, Nikki Bourassa, Ethan Zuckerman, and Yochai Benkler. 2017. *Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election.* Technical Report. Berkman Klein Center for Internet and Society.
[14] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. The Rise of Social Bots. *Comm. ACM* 59, 7 (June 2016). https://doi.org/10.1145/2818717
[15] Michaela Ferron and Paolo Massa. 2014. Beyond the Encyclopedia: Collective Memories in Wikipedia. *Memory Studies* 7, 1 (Jan. 2014), 22–45. https://doi.org/10.1177/1750698013490590
[16] Daniel Gayo-Avello. 2013. A Meta-Analysis of State-of-the-Art Electoral Prediction From Twitter Data. *Social Science Computer Review* 31, 6 (Dec. 2013), 649–679. https://doi.org/10.1177/0894439313493979
[17] R. Stuart Geiger. 2014. Bots, Bespoke, Code and the Materiality of Software Platforms. *Information, Communication & Society* 17, 3 (March 2014), 342–356. https://doi.org/10.1080/1369118X.2013.873069
[18] Sascha Göbel and Simon Munzert. 2017. Political Advertising on the Wikipedia Marketplace of Information. *Social Science Computer Review* (May 2017), 1–19. https://doi.org/10.1177/0894439317703579
[19] Robert Gorwa. 2017. Twitter Has a Serious Bot Problem and Wikipedia Might Have the Solution. *Quartz* (Oct. 2017).
[20] Shane Greenstein and Feng Zhu. 2012. Is Wikipedia Biased? *American Economic Review* 102, 3 (May 2012), 343–348. https://doi.org/10.1257/aer.102.3.343
[21] Shane Greenstein and Feng Zhu. 2017. Do Experts or Crowd-Based Models Produce More Bias? Evidence from Encyclopaedia Britannica and Wikipedia. *Management Information Systems Quarterly* (2017).
[22] Aaron Halfaker, Aniket Kittur, Robert Kraut, and John Riedl. 2009. A Jury of Your Peers: Quality, Experience and Ownership in Wikipedia. In *Proc. WikiSym '09.* ACM. https://doi.org/10.1145/1641309.1641332
[23] Eszter Hargittai and Aaron Shaw. 2015. Mind the skills gap: the role of Internet know-how and gender in differentiated contributions to Wikipedia. *Information, Communication & Society* 18, 4 (2015), 424–442.

[24] Mary Harris. 2016. A Media Post-Mortem on the 2016 Presidential Election – mediaQuant. https://www.mediaquant.net/2016/11/a-media-post-mortem-on-the-2016-presidential-election/
[25] Benjamin Mako Hill and Aaron Shaw. 2013. The Wikipedia gender gap revisited: characterizing survey response bias with propensity score estimation. *PloS one* 8, 6 (2013), e65782.
[26] Benjamin Mako Hill and Aaron Shaw. 2014. Consider the Redirect: A Missing Dimension of Wikipedia Research. In *Proc. OpenSym '14*. ACM. https://doi.org/10.1145/2641580.2641616
[27] Benjamin Mako Hill and Aaron Shaw. 2015. Page Protection: Another Missing Dimension of Wikipedia Research. In *Proc. OpenSym 2015*. ACM, New York, NY, USA. https://doi.org/10.1145/2788993.2789846
[28] Seth J. Hill. 2014. A Behavioral Measure of the Enthusiasm Gap in American Elections. *Electoral Studies* 36 (Dec. 2014), 28–38. https://doi.org/10.1016/j.electstud.2014.06.012
[29] J. D. Hunter. 2007. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* 9, 3 (2007), 90–95. https://doi.org/10.1109/MCSE.2007.55
[30] Dariusz Jemielniak. 2014. *Common knowledge?: An ethnography of Wikipedia.* Stanford University Press.
[31] Eric Jones, Travis Oliphant, Pearu Peterson, et al. 2001–. SciPy: Open source scientific tools for Python. http://www.scipy.org/
[32] John T. Jost, Mahzarin R. Banaji, and Brian A. Nosek. 2004. A Decade of System Justification Theory: Accumulated Evidence of Conscious and Unconscious Bolstering of the Status Quo. *Political Psychology* 25, 6 (2004), 881–919. https://doi.org/10.1111/j.1467-9221.2004.00402.x
[33] Andreas Jungherr, Harald Schoen, Oliver Posegga, and Pascal Jürgens. 2016. Digital Trace Data in the Study of Public Opinion An Indicator of Attention Toward Politics Rather Than Political Support. *Social Science Computer Review* 35, 3 (Feb. 2016), 336–356. https://doi.org/10.1177/0894439316631043
[34] Joshua L. Kalla and Peter M. Aronow. 2015. Editorial Bias in Crowd-Sourced Political Information. *PLoS One* 10, 9 (Sept. 2015). https://doi.org/10.1371/journal.pone.0136327
[35] Aaron C. Kay, Maria C. Jimenez, and John T. Jost. 2002. Sour Grapes, Sweet Lemons, and the Anticipatory Rationalization of the Status Quo. *Personality and Social Psychology Bulletin* 28, 9 (2002), 1300–1312. https://doi.org/10.1177/01461672022812014
[36] Brian Keegan. 2013. A History of Newswork on Wikipedia. In *Proc. WikiSym '13*. ACM. https://doi.org/10.1145/2491055.2491062
[37] Brian Keegan. 2015. Emergent Social Roles in Wikipedia's Breaking News Collaborations. In *Roles, Trust, and Reputation in Social Media Knowledge Markets*. Springer, 57–79.
[38] Brian Keegan and Jed R. Brubaker. 2015. 'Is' to 'Was': Coordination and Commemoration in Posthumous Activity on Wikipedia Biographies. In *Proc. CSCW '15*. ACM. https://doi.org/10.1145/2675133.2675238
[39] Brian Keegan, Darren Gergle, and Noshir Contractor. 2011. Hot off the Wiki: Dynamics, Practices, and Structures in Wikipedia's Coverage of the Tōhoku Catastrophes. In *Proc. WikiSym '11*. ACM. https://doi.org/10.1145/2038558.2038577
[40] Brian Keegan, Darren Gergle, and Noshir Contractor. 2013. Hot Off the Wiki: Structures and Dynamics of Wikipedia's Coverage of Breaking News Events. *American Behavioral Scientist* 57, 5 (2013), 595–622. https://doi.org/10.1177/0002764212469367
[41] Srijan Kumar, Robert West, and Jure Leskovec. 2016. Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes. In *Proc. WWW '16*. 591–602. https://doi.org/10.1145/2872427.2883085
[42] Drew B. Margolin, Sasha Goodman, Brian Keegan, Yu-Ru Lin, and David Lazer. 2016. Wiki-Worthy: Collective Judgment of Candidate Notability. *Information, Communication & Society* 19, 8 (Aug. 2016), 1029–1045. https://doi.org/10.1080/1369118X.2015.1069871
[43] Wes McKinney. 2011. pandas: a foundational Python library for data analysis and statistics. *Python for High Performance and Scientific Computing* 14 (2011).
[44] Amanda Menking, David W. McDonald, and Mark Zachry. 2017. Who Wants to Read This?: A Method for Measuring Topical Representativeness in User Generated Content Systems. In *Proc. of CSCW '17*. ACM, 2068–2081. https://doi.org/10.1145/2998181.2998254
[45] Panagiotis T. Metaxas and Eni Mustafaraj. 2012. Social Media and the Elections. *Science* 338, 6106 (2012), 472–473. https://doi.org/10.1126/science.1230456
[46] Jessica J. Neff, David Laniado, Karolin E. Kappler, Yana Volkovich, Pablo Aragón, and Andreas Kaltenbrunner. 2013. Jointly They Edit: Examining the Impact of Community Identification on Political Interaction in Wikipedia. *PLOS ONE* 8, 4 (April 2013). https://doi.org/10.1371/journal.pone.0060584
[47] Sabine Niederer and José van Dijck. 2010. Wisdom of the Crowd or Technicity of Content? Wikipedia as a Sociotechnical System. *New Media & Society* 12, 8 (Dec. 2010), 1368–1387. https://doi.org/10.1177/1461444810365297
[48] Oded Nov. 2007. What Motivates Wikipedians? *Commun. ACM* 50, 11 (2007), 60–64. https://doi.org/10.1145/1297797.1297798

[49] Mike Pearl. 2016. A Wikipedian Explains How Wikipedia Stays Reliable in the Fake News Era. *Vice* (Nov. 2016).
[50] Christian Pentzold. 2009. Fixing the Floating Gap: The Online Encyclopaedia Wikipedia as a Global Memory Place. *Memory Studies* 2, 2 (Jan. 2009), 255–272. https://doi.org/10.1177/1750698008102055
[51] Jacob Ratkiewicz, Michael Conover, Mark R. Meiss, Bruno Gonçalves, Alessandro Flammini, and Filippo Menczer. 2011. Detecting and Tracking Political Abuse in Social Media. In *Proc. ICWSM '11*. AAAI.
[52] Joseph Michael Reagle. 2010. *Good faith collaboration: The culture of Wikipedia*. MIT Press.
[53] Aaron Shaw and Benjamin M. Hill. 2014. Laboratories of Oligarchy? How the Iron Law Extends to Peer Production. *Journal of Communication* 64, 2 (2014), 215–238. https://doi.org/10.1111/jcom.12082
[54] Herbert A. Simon. 1954. Bandwagon and Underdog Effects and the Possibility of Election Predictions. *The Public Opinion Quarterly* 18, 3 (1954), 245–253.
[55] Kate Starbird. 2017. Examining the Alternative Media Ecosystem Through the Production of Alternative Narratives of Mass Shooting Events on Twitter.. In *Proc. ICWSM '17*. AAAI.
[56] Maxwell Tani. 2015. Meet the guy who has protected Hillary Clinton's Wikipedia page for almost a decade. *Business Insider* (May 2015). http://www.businessinsider.com/meet-hillary-clintons-wikipedia-editor-2015-5
[57] Jennifer Thom, Dan R. Cosley, and Geri Gay. 2009. What's Mine Is Mine: Territoriality in Collaborative Authoring. In *Proc. CHI '09*. ACM. https://doi.org/10.1145/1518701.1518925
[58] Zeynep Tufekci and Christopher Wilson. 2012. Social media and the decision to participate in political protest: Observations from Tahrir Square. *Journal of communication* 62, 2 (2012), 363–379.
[59] Marlon Twyman, Brian Keegan, and Aaron Shaw. 2017. Black Lives Matter in Wikipedia: Collective Memory and Collaboration Around Online Social Movements. In *Proc. CSCW '17*. ACM. https://doi.org/10.1145/2998181.2998232
[60] Claudia Wagner, David Garcia, Mohsen Jadidi, and Markus Strohmaier. 2015. It's a man's Wikipedia? Assessing gender inequality in an online encyclopedia. In *Ninth international AAAI conference on web and social media*.
[61] Morten Warncke-Wang, Vivek Ranjan, Loren G. Terveen, and Brent J. Hecht. 2015. Misalignment Between Supply and Demand of Quality Content in Peer Production Communities.. In *Proc. ICWSM '15*. AAAI.
[62] Heng-Li Yang and Cheng-Yu Lai. 2010. Motivations of Wikipedia content contributors. *Computers in Human Behavior* 26, 6 (2010), 1377–1383. https://doi.org/10.1016/j.chb.2010.04.011
[63] Taha Yasseri and Jonathan Bright. 2016. Wikipedia Traffic Data and Electoral Prediction: Towards Theoretically Informed Models. *EPJ Data Science* 5, 1 (June 2016), 1–15. https://doi.org/10.1140/epjds/s13688-016-0083-3

Received April 2019; revised June 2019; accepted August 2019