The 2016 US presidential election and media on Instagram: Who was in the lead?

Harald Schmidbauer [a, b, *], Angi Rösch [c], Fabian Stieler [d]

[a] Shanxi University of Finance and Economics, Taiyuan, Shanxi Province, China
[b] Telostat Pte Ltd, Singapore
[c] FOM University of Applied Sciences, Munich, Germany
[d] University of Augsburg, Augsburg, Germany

A R T I C L E I N F O

Article history:
Received 6 April 2017
Received in revised form
8 November 2017
Accepted 14 November 2017
Available online 5 December 2017

Keywords:
2016 US presidential election
Instagram
Media hashtag
Wavelet analysis
Media upload timing

A B S T R A C T

The present study is an effort to analyze the timing of media postings related to candidates Clinton and Trump on Instagram before and after the 2016 US presidential election. Hashtags are used to determine whether a posting was intended to support or oppose either candidate. We thus obtain four hourly time series: Clinton vs. Trump, supporters vs. opponents. Based on cross-wavelet analysis, we find that, at the 12-h period, Trump supporters were leading Trump opponents as well as Clinton supporters the days before the election, while Clinton opponents were often leading Clinton supporters: Trump supporters and Clinton opponents were eager to post media, while Trump opponents and Clinton supporters were sluggish. Considering election forecasts, our results come as a surprise.

1. Introduction

"Instagram users post black squares after Trump wins US presidency", an online magazine headlined late at night on the day of the 2016 US presidential election.[1] According to John Quelch and Thales Teixeira from Harvard Business School, online social media receive a fresh boost every four years from the US presidential campaign, and, alluding to Donald Trump's fondness for this service, they dubbed the most recent campaign the "Twitter Election".[2] "But does it have the power to determine which candidate will win?" the authors asked two months before Election Day and, pondering the two candidates' Twitter performances and its potential effect on actual voter behavior, concluded: "Probably not."

While the presidential race of 2004 embraced mainly websites and blogs, the 2008 campaigns were the first to fully integrate Web 2.0 technologies, i.e. those providing an open forum for social interactivity and networking, like Facebook, Twitter, and YouTube. The way online social media influenced the course of the 2008 electoral process — "The Facebook Election" — was addressed by several authors, e.g. Johnson and Perlmutter (2010), Kushin and Yamamoto (2010), Fernandes, Giurcanu, Bowers and Neely (2010), and Woolley, Limperos and Oliver (2010) in a special issue of *Mass Communication and Society*. Aronson (2012) concluded: The influence was through enhancing the flow of information, suggesting campaign focuses, increasing opportunities for fundraising, profiling candidates in public opinion, stimulating political participation and the turnout of young voters, and "in some cases, impacting election results themselves".

The goal of the aforementioned large-scale study conducted by Woolley et al. (2010) was to understand the overall image of the two then presidential candidates Barack Obama and John McCain generated by politically motivated Facebook groups. Pages of sampled Facebook groups were captured the day before the election and subjected to quantitative content analysis. For Obama-focused groups, the researchers found significantly higher levels of group membership as well as group activity, which they measured as sum of the number of items posted. This finding was

* Corresponding author.
E-mail addresses: harald@hs-stat.com (H. Schmidbauer), angi@angi-stat.com (A. Rösch), stielerfabian@gmail.com (F. Stieler).

[1] Article released on November 9, 2016, by the US magazine "Wired"; available online at https://www.wired.com/2016/11/instagram-trump-win/. Accessed on December 21, 2016.

[2] "The Twitter Election", online article by J. Quelch and T. Teixeira, released on September 9, 2016. Harward Business School Working Knowledge; available online at http://hbswk.hbs.edu/item/the-twitter-election. Accessed on December 2, 2016.

more or less expected, given the demographics of Facebook users and the fact that younger people predominantly supported Obama. However, the content analysis of references to the two candidates offered some surprises too. Classifying references as conveying a positive, negative, or neutral connotation, the researchers not only found significantly more positive references to Obama than to McCain across groups, but also that McCain-focused groups were "overwhelmingly negative". According to Woolley et al. (2010), tentative explanations may include Obama's reported commitment to a "positive" campaign, as well as potential differences in perceived social capital between the two groups of supporters which may have driven the motivation to express political opinions some way or other. Their findings may indicate a higher extent of third-person and hostile media perception among Republican supporters (as found by Banning (2006) and Lee (2005)), and the researchers doubt the benefits of online social media as a tool for promoting dialogue.

Adding to a growing body of studies questioning the democratic potential of online social media, the research by Yamamoto and Kushin (2014) suggests that consuming online social media for political campaign information decreases skepticism. Moreover, their study of online survey data of college students in the 2008 presidential election showed empirical evidence for increased cynicism and apathy among mere consumers. These latter, negative, aspects of political disaffection, were, however, not found associated with those who use online social media to express political views and interact with others. Instead, skepticism was increased. In another article, c.f. Kushin and Yamamoto (2010), the authors relate online expression among young adults to situational political involvement, though not to political self-efficacy.

Mass media (i.e. US and international newspapers, broadcast media, blogs, online media and magazines) coverage in the 2012 US presidential election was studied by Sudhahar, Veltri and Cristianini (2015), generating a network of political actors and issues, and quantifying relations of support and opposition between them by means of automated text mining. Among other things, their network analysis revealed that mass media reported positive statements more frequently for the Democrat than for the Republican campaign, and the latter was portrayed more negatively. An observation made by Jahanbakhsh and Moon (2014) in a large-scale sentiment analysis among Twitter users reflects a corresponding hostile media effect: the prominence of negative tweets for Obama in the run-up to the election day (although Obama also received more positive tweets than his competitor Mitt Romney). In an experimental study carried out a few months before the election, Iyengar and Westwood (2015) observed that the "polarization [in terms of social identity and affect towards candidates] of the American electorate has dramatically increased".

The predictive power of social media is, however, an issue of debate. Jahanbakhsh and Moon (2014) in the aforementioned study used geo-tagged tweets to estimate the 2012 candidates' popularities across states, scoring a 76% success rate in predicting election results. Prediction accuracy was lower for states won by Romney, which may relate to an underrepresentation of Twitter users in those states; nevertheless, the researchers conclude, social media could be an "important source of information for opinion mining". Gayo-Avello, Metaxas and Mustafaraj (2011) and Metaxas, Mustafaraj and Gayo-Avello (2011) discuss limitations to the predictability of election results using online social media. They propose a set of standards a prediction method should follow in order to be consistently competitive with state-of-the-art techniques, in particular addressing the sampling bias, and request a "need for a deeper understanding of the dynamics of political conversation in social media". Bovet, Morone and Makse (2016) develop analytics to predict election polls concerning the 2016 US presidential election based on "opinion mining" among Twitter users, claiming that Twitter can be an "early warning signal of global opinion trends" since their results anticipate the New York Times National Polling Average.

In scientific literature, little attention has been payed to Instagram so far. This online photo- and video-sharing platform,[3] launched in October 2010 and pioneering the "visual social media" trend, is today among the social networks with the highest growth rates. It had 10 million users one year after it had been founded, exceeding 500 million in June 2016, with about 100 million living in the US.[4] Instagram enables its users to upload media in order to share them privately or publicly, allowing access from other social networking sites (Facebook, Twitter, Tumblr, Flickr; it was taken over by Facebook in 2012), add captions or comments and categorize uploaded media using hashtags. Unlike Facebook, Twitter, or Pinterest, Instagram doesn't allow the reposting of media; therefore, the number of media postings with a certain hashtag cannot be inflated that way.

A survey of college students was used by Sheldon and Bryant (2016) to investigate motives why people use Instagram. Their study suggests four main motives: "surveillance/knowledge about others" which emerged as the most influential, "documentation", "coolness", and "creativity". Correlation studies including indicators of interpersonal interaction, life satisfaction, social activity and narcissism identified social activity as the most important pedictor of Instagram usage, apart from gender; females are more likely to be on Instagram.

Hu, Manikonda and Kambhampati (2014) analyzed, among other things, the content of posted photos on the basis of a random sample of Instagram users. They found that the two photo categories "Selfies" and "Friends" are the most prominent ones (each accounting for more than 20% of all photos), emphasizing the role of Instagram for self-promotion and social networking. Furthermore, the user proportions with respect to "engagement" (in terms of the number of photos posted) were the most balanced ones for these two categories. Giannoulakis and Tsapatsoulis (2016) examined the descriptive power of Instagram hashtags in a survey, asking whether other users would tag the image in question with the same annotation. They found an average match of 66%, suggesting initial evidence that Instagram image-hashtag pairs might be used for training automated image annotation models.

The present study is an effort to contribute to the understanding of the dynamics of media postings on Instagram in connection with the 2016 presidential election. Our research questions are:

RQ1: Do media postings supporting (or opposing) Clinton differ in number from media postings supporting (or opposing) Trump?

RQ2: Do media postings supporting (or opposing) Clinton differ with respect to their *timing* (that is, *when* they are posted) from media postings supporting (or opposing) Trump?

The second question is more interesting and also more difficult to answer, because it explicitly refers to the dynamic aspect of the media posting process. A selection of hashtags is used as a criterion to classify media postings as either being in support of (or neutral towards) a candidate, or opposing a candidate. Our data collection method is therefore non-participant observation; it amounts to monitoring the number of Instagram hashtag prevalence in the course of time, starting in September 2016. An application

[3] See https://www.instagram.com/instagram/. Accessed on January 17, 2017.

[4] See http://blog.instagram.com/post/146255204757/160621-news. Accessed on January 17, 2017.

Table 1
Selection of Instagram hashtags, Trump and Clinton, pos/neutral and negative.

hashtag	number of media with this hashtag		
	2016-09-01	2016-11-08	2016-11-30
#donaldtrump	1,074,451	1,486,197	1,840,775
#makeamericagreatagain	453,909	684,828	883,888
#trump	1,334,783	2,031,761	2,798,253
#trumpforpresident	77,878	118,523	134,175
#trumptrain	212,354	340,673	421,529
Trump pos/neutral	3,153,375	4,661,982	6,078,620
#clinton	220,489	354,168	447,463
#hillary	277,318	445,411	543,804
#hillary2016	240,413	310,105	330,859
#hillaryclinton	559,744	985,405	1,232,371
Clinton pos/neutral	1,297,964	2,095,089	2,554,497
#dumptrump	141,300	238,131	307,840
#fucktrump	177,169	248,934	393,438
#nevertrump	92,299	184,250	224,173
#stoptrump	13,720	21,464	24,524
Trump neg	424,488	692,779	949,975
#killary	64,251	106,784	119,566
#neverhillary	110,046	199,241	215,643
#stophillary	16,151	21,699	21,714
Clinton neg	190,448	327,724	356,923

programming interface (API) provided by Instagram enables us to extract hourly readings of the number of tagged media postings with a given hashtag. We use wavelet and cross-wavelet transformations to study the periodic properties of the hourly time series of media postings obtained in this way. Considering the findings obtained in studies investigating social media in connection with previous US presidential elections, our working hypotheses are:

H1. Media posted in favor of (or neutral towards) Clinton outnumber those in favor of (or neutral towards) Trump.

H2. The media posting process relating to either candidate is periodic; it follows a 24-h pattern.

H3. Media in favor of (or neutral towards) Clinton are posted more eagerly than Trump-related media; as a consequence, the series of media supporting (or being neutral towards) Clinton is leading either (positive/neutral or negative) Trump-related series; the latter series are lagging.

H4. Media opposing Trump are posted more eagerly than media in support of (or neutral towards) Trump; as a consequence, the series of media opposing Trump is leading the series of media supporting (or being neutral towards) Trump; the latter series is lagging.

H5. The periodic properties of the time series mentioned above are time-dependent: Their behavior before and after the three presidential debates may differ.

To the best of our knowledge, this is the first study on Instagram media postings in connection with a US presidential election, and also the first to apply wavelet methodology in the context of social media.

This paper is organized as follows. Section 2 specifies the data on which this study is based, documents some aspects of data retrieval, and provides a first glance at the data. Section 3 introduces some concepts of wavelet and cross-wavelet analysis. Empirical results are presented in Section 4, followed by a discussion in Section 5. Section 6 summarizes and concludes the paper. — All computations were carried out with scripts written in R (R Core Team, 2016); wavelet computations and plots were accomplished with R package *WaveletComp* (Rösch & Schmidbauer, 2014).

2. The data

2.1. Data retrieval

Instagram's built-in API allows snapshots of the current number of tagged media postings with a given hashtag, whether publicly shared or not, but it does not provide historical data.[5] We used a cron demon (with scripts written in Python) to monitor the number of tagged media with certain hashtags in the course of time. The API rate limit of 500 calls per hour allows the query of a total of 50 hashtags every 5 min. Hourly readings, however, were found more appropriate to capture the uploading process; 1-h time intervals are short enough to provide a sufficient level of granularity in order to detect significant patterns in real time and long enough to provide a sufficient data flow for the analysis, avoiding spurious patterns caused by lack of new media uploads in too short time intervals.[6] In compliance with the data protection regulations, no private data of any kind (user names, profiles, media, comments, likes, number of followers, location) were collected.

The search for hashtags relating to either Clinton or Trump proceeded in two steps. The first step, taken in June 2016, involved scrutinizing (online) press releases and results of Instagram's search routine for potentially relevant keywords (e.g. #election, #america, #usa, #trump, #clinton, #obama, #nobama, #vote, etc.). The latter provided a hit list of associated media postings together with a list of "similar hashtags". In this vein, a preliminary set of hashtags was identified, suggesting positive, neutral, and negative annotations to either presidential candidate. Some hashtags were excluded because they had only a weak connection with the slogans of the Clinton campaign at that time, for instance, the hashtags

[5] See the documentation at https://www.instagram.com/developer/endpoints/tags/. Accessed on February 23, 2017.

[6] An hourly granularity for trending purposes is also what Instagram suggests; see https://engineering.instagram.com/trending-on-instagram-b749450e6d93#.cjjakufgp. Accessed on February 23, 2017.

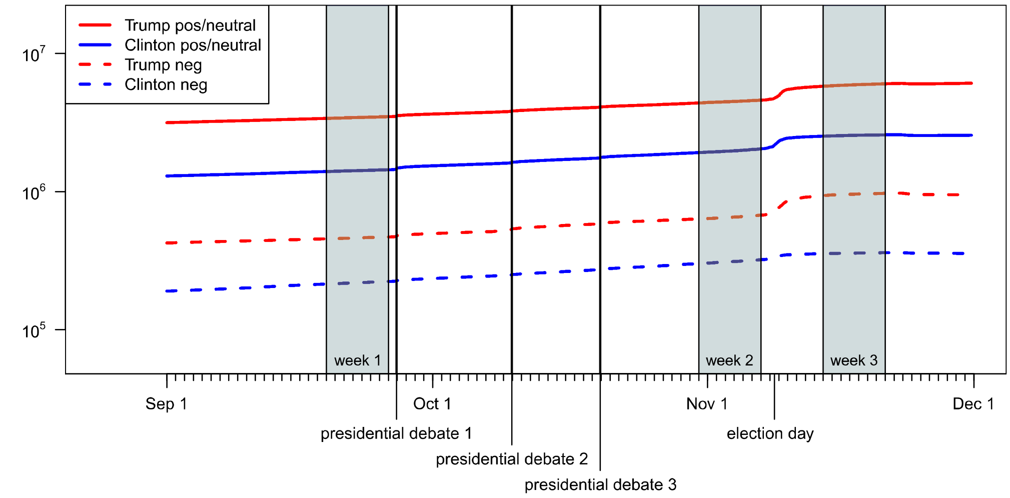

Fig. 1. Number of media, pos/neutral and negative hashtags (logarithmic scale).

#strongertogether and #imwithher. In contrast, the hashtag #makeamericagreatagain was included because it matched the Trump campaign very well. In the second step, the number of media postings with each hashtag in this set was monitored for testing purposes, and a hashtag was included in the final set only if at least 10,000 Instagram media were tagged with it at that time. Since the present study rests on the availability of consistent time series data, this set of hashtags was not revised later on. Thus, we identified a total of 16 relevant Instagram hashtags, which were monitored on an hourly basis from September through November 2016, seven Clinton-related and nine Trump-related. The larger number of the latter reflects the abundance of Trump-related hashtag phrases and media uploads from June 2016 onward. It will be seen below that the difference in the numbers of candidate-related hashtags does not impair the analysis of periodic patterns. The time stamp of the series is Eastern Standard Time (EST). Since Eastern Daylight Time (EDT) was in effect until November 6, 2:00 a.m., both EDT- and EST-stamped data were recorded for the time slot of fall back.

2.2. *A first glance at the data*

Table 1 lists the hashtags on which our investigations are based. Assuming that a hashtag can tell us what intentions media uploaders had, the hashtags were classified into four categories: Trump pos[itive]/neutral, Clinton pos[itive]/neutral, Trump neg [ative], and Clinton neg[ative]. Numbers of media posted on Instagram with these hashtags (as recorded at 0:00 a.m.) are reported for three days in Table 1: the first day of September, about three weeks before the first presidential debate; the day of the election, November 8; and the last day of November. Fig. 1 shows a plot of the four relevant time series on a logarithmic scale.

Although total media numbers per category should be compared with caution, Table 1 demonstrates that Trump-tagged outnumbered Clinton-tagged media by far at all three points in time. There are three hashtag pairs for which a direct comparison is meaningful: #donaldtrump vs. #hillaryclinton, #nevertrump vs. #neverhillary, and #stoptrump vs. #stophillary. On September 1, media with hashtag #hillaryclinton counted a good half of those with hashtag #donaldtrump. On election day, #hillaryclinton-tagged media had increased by 76%, but even this strong increase was not enough to mitigate the dominance of hashtag #donaldtrump. We therefore find hypothesis H1 confirmed in terms of net media increase during the pre-election period, at least when considering this hashtag pair. However, in terms of levels, H1 is rejected. In contrast, media with hashtag #neverhillary (#stophillary) always outnumbered, albeit slightly, their Trump-related counterparts; these passed by only after the election.

Fig. 1 gives an idea of the dominance of Trump-related media in Instagram and shows the increasing trend of media counts. It is also visible that the three presidential debates[7] as well as the election itself boosted media uploading. In order to capture the fundamental dynamics of the media uploading process relating to the two candidates, we will focus on three non-overlapping weeks (each one extending from Monday through Sunday), excluding these exceptional boosts. The weeks under investigation, also highlighted in Fig. 1, are:

- **Week 1:** September 19–25, the week preceding the first presidential debate;
- **Week 2:** October 31 – November 6, after the presidential debates, before the election[8];
- **Week 3:** November 14–20, the week following the election.

In order to assess the periodicity of the media posting process (see hypothesis H2), analyzing the time series of hourly differences is more suitable, because hourly differences reflect media increase (or decrease) more precisely than the level series of Fig. 1. For each of the three weeks under investigation, Fig. 2 displays the time series of hourly differences referring to the four hashtag categories; Table 2 reports some distributional characteristics.

A glance at Fig. 2 reveals a basic daily pattern of media upload activity among Instagram users, confirming hypothesis H2: low in morning hours, high in evening hours. (We shall see in Section 4 below that the periodic pattern is actually richer than what is visible to the naked eye.) According to Table 2, all average differences are positive, in line with the increasing trends. Between weeks 1 and 2, average differences per hour more or less tripled in the case of Clinton pos/neutral (not making up leeway with respect to Trump, as we have seen) and Trump neg, while they only doubled for the other two cases. Maybe not surprisingly, the only

[7] See http://www.uspresidentialelectionnews.com/2016-debate-schedule/2016-presidential-debate-schedule/. Accessed on January 30, 2017.

[8] Concerning the fall back to standard time on November 6, we ignore the time change; in terms of standard time, week 2 actually ends on November 7, 1:00 a.m.

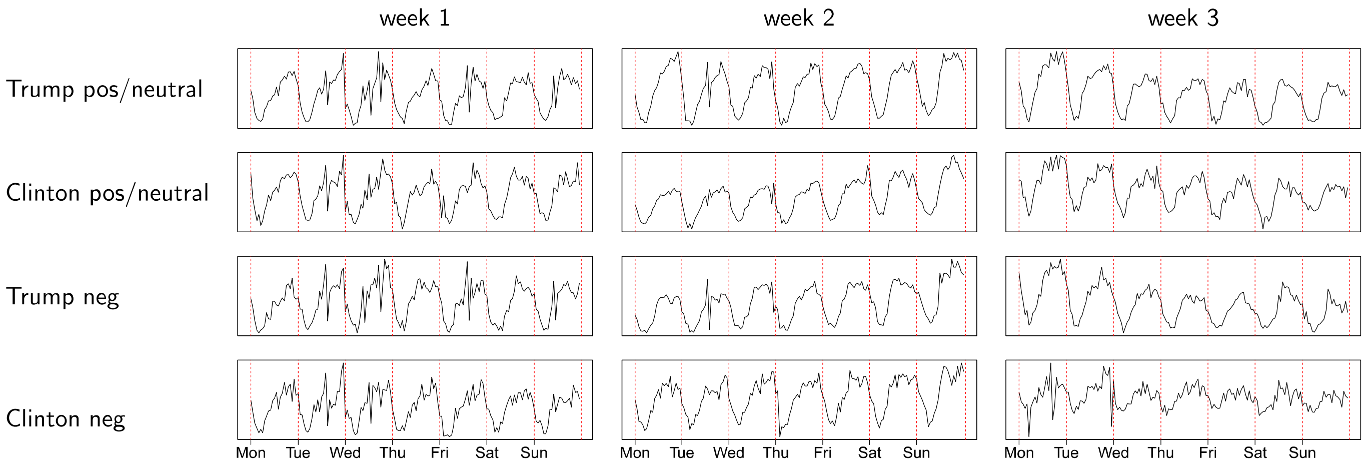

Fig. 2. Time series of differences, three weeks.

Table 2
Minimum, maximum, arithmetic mean of hourly differences; three weeks.

	week 1			week 2			week 3		
	2016-09-19–2016-09-25			2016-10-31–2016-11-06			2016-11-14–2016-11-20		
	min	max	mean	min	max	mean	min	max	mean
Trump pos/neutral	19	1402	596.7	−19	2420	1235.4	273	2928	1424.4
Clinton pos/neutral	−21	542	249.5	−111	1799	727.0	−184	802	334.8
Trump neg	−2	200	83.4	−14	616	245.6	−5	516	199.8
Clinton neg	−7	143	54.5	−27	238	117.7	−74	127	33.5

further increase in hourly average was with Trump pos/neutral in week 3 after the election. — The following sections will analyze the periodic pattern of upload activity among supporters and opponents of either candidate and thus provide insight into the timing of uploads.

3. Wavelet methodology

The core argument of the present paper rests on wavelet transformations of the time series displayed in Fig. 2; it is therefore in order to give a very brief account of the aims of wavelet methodology, as far as needed in our context. Wavelet methodology is a reasonable choice to study periodic phenomena in a time series. Its advantage, particularly in the presence of periods potentially changing over time, has been widely documented in applications in natural sciences, in signal and image processing and more recently also economics; see e.g. the textbooks by Carmona, Hwang and Torrésani (1998), Gencay, Selcuk and Whitcher (2001), or the research articles also providing an introduction to the methodology by Torrence and Compo (1998) in the field of geophysics; Aguiar-Conraria, Azevedo and Soares (2008) with a focus on economic time series, to name but a few.

As an example, the left-hand part of Fig. 3 shows how the series z, of length 400 (400 observations), is "synthesized" using three constituents, namely: a series x with period 100, a series y with period 40, and Gaussian white noise. The resulting series $z = x + y +$ noise still has periods 100 and 40, but this is not easily visible to the naked eye. Wavelet analysis of z reveals the periodic properties of z and thus helps to recover the hidden structure of z. A typical result of the wavelet transformation of a **univariate** time series is the wavelet power spectrum. The right-hand side of Fig. 3 shows a "heat map" of the wavelet power spectrum of z. The heat

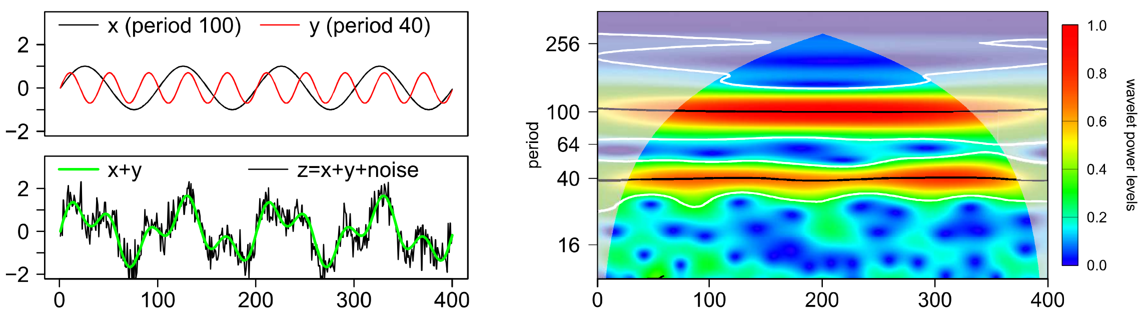

Fig. 3. Time series z and its wavelet power spectrum.

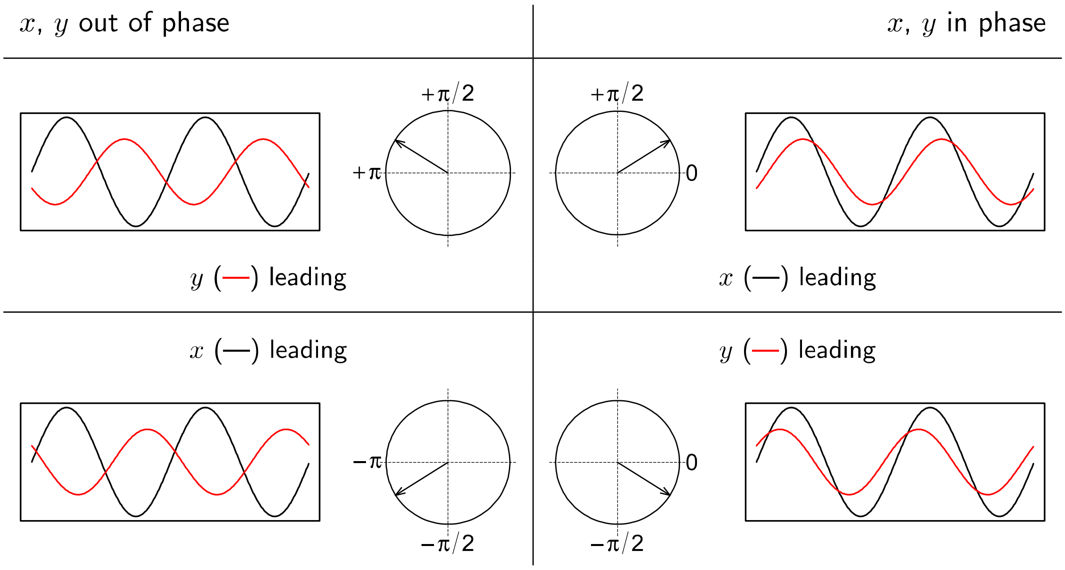

Fig. 4. Phase shifts.

map shows which periods (on the vertical axis, in logarithmic scale) are important at a given time (the horizontal axis). The degree of importance (the wavelet power or the "heat") is visualized by means of colors: Throughout the observed time span, periods 40 and 100 are important in z, indicated by the high power levels and the ridges marking local maxima (the black lines). Periods outside of the area delineated by the white lines for a given point in time are not statistically significant. The shaded area in the upper part of the heat map indicates those time-period combinations which are outside of the so-called cone of influence. For example, it is not possible to obtain certain knowledge about the importance of period 300 in a time series when only 400 observations are available (this is an example of one variant of the uncertainty principle).

In our context, the essential tool is the wavelet transformation of a **bivariate** — that is, two values are observed at the same time, for example: Trump neg and Clinton neg, see Fig. 2 — time series. Similar to the univariate case, the typical outcome is a heat map of what is now called the cross-wavelet power spectrum, as shown in Figs. 5 and 6 below. The heat map in this case shows which periods are important *in both series* at a given point in time. For a period shared by both series, further relevant information to be retrieved by wavelet analysis concerns the relative positions of the series with respect to each other: Which series is leading, which one is lagging? This is information about the phase shift, and the four possible cases are illustrated in Fig. 4. For example, series x and y in the top right of Fig. 4 are in phase because they both reach their maxima within the *same* half-cycle (or half-period); x is leading (and y is lagging) because it reaches its maximum before y. With the convention that a full cycle corresponds to the circumference of a unit circle (which equals 2π), the lead time (the phase shift) of x over y can be expressed as a fraction of a full cycle; this is depicted in the top right circle in Fig. 4. Thus, the arrow plotted in the circle indicates the relative positions of the two series with respect to each other. For the cases where x and y are in phase, an upward (downward) arrow in the cross-wavelet transformation of x over y indicates that x (y, respectively) is leading and y (x, respectively) is lagging; the magnitude of the phase shift can be inferred from the angle between the arrow and the horizontal line.

Two further properties of wavelet analysis are relevant in our context: (i) Wavelet analysis will detect a change in the periodic properties of a time series. For example, if another important period emerges in the course of time, it will show up as "hot area" in the heat map. Another example: If the phase shift between two series evolves, the angles of corresponding arrows will change too. (ii) The magnitude of a series (or two series in cross-wavelet analysis) does not matter. For example, transforming $2z$ instead of z in Fig. 3 will lead to an identical heat map. Another example: Cross-wavelet transforming the pair $2x + 5$ and y instead of the pair

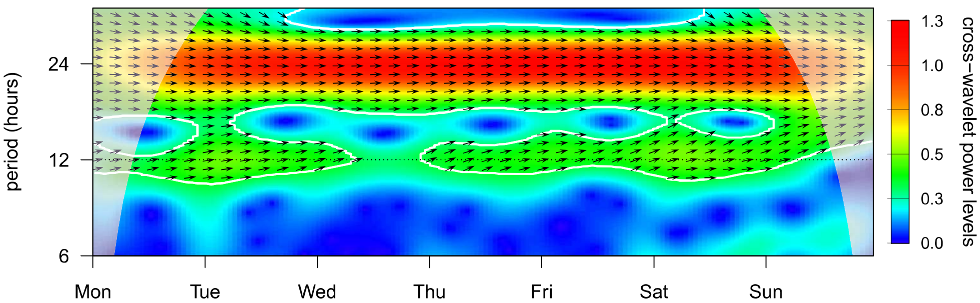

Fig. 5. Cross-wavelet power spectrum, Trump pos/neutral over Clinton pos/neutral, week 2.

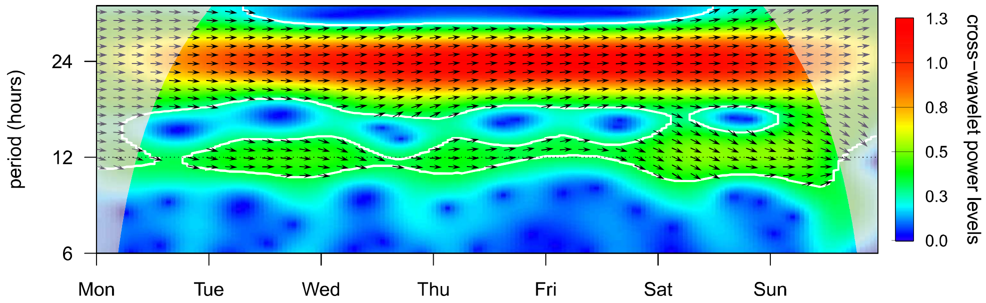

Fig. 6. Cross-wavelet power spectrum, Trump neg over Clinton neg, week 2.

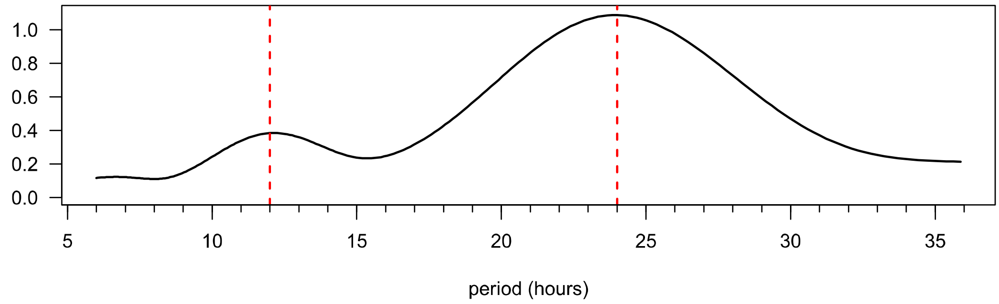

Fig. 7. Average wavelet power, Trump pos/neutral, week 2.

x and *y* in Fig. 4 will lead to the same arrow angle.

4. Empirical results

4.1. Cross-wavelet transformation

Cross-wavelet methodology, as outlined in Section 3, was applied to the six pairs of hourly media postings time series (the series of differences, displayed in Fig. 2), for each week separately. Two examples of resulting cross-wavelet power spectra are shown in Figs. 5 and 6; both of them refer to week 2 (the week before the election). These heat maps indicate periods which are jointly significant in both series, as well as the phase shift between the two series at a given period between 6 h and 36 h. It is obvious that period 24 h is very powerful throughout the week, underpinning the validity of hypothesis H2 from the wavelet perspective. This 24-h period could be said to mirror the daily routine behavior of Instagram users. Figs. 5 and 6 also bring to light that period 12 h is another important constituent of the series of differences, and, for the example of the Trump pos/neutral series in week 2, this observation is confirmed by the plot of the average (taken over the entire week, within the cone of influence) power of the wavelet transform in Fig. 7. Wavelet analysis reveals that similar remarks apply to all twelve time series of differences displayed in Fig. 2: All of them have strong and statistically significant periods of 12 as well as 24 h (Fig. 7 shows one of the twelve plots of average wavelet power; the other plots, which are not shown here, are all very similar in shape). This leads to the conclusion that each series of differences displayed in Fig. 2 is essentially a synthesis of two series, one with a period of 12 h and another with a period of 24 h. Compared with the 24-h period, the 12-h period reflects a more intense media traffic on Instagram.

The arrows plotted in Figs. 5 and 6 provide some first insight into the relative behavior of the pair of series in question: At period 24 h, the arrows are aligned almost exactly horizontally, so that the series are in phase and perfectly synchronized, with a phase shift of zero — Trump pos/neutral and Clinton pos/neutral Instagram media uploadings reach their respective maxima at the same time each day (i.e. at the 24-h period), and so do Trump neg and Clinton neg Instagram media uploadings. This is not true, however, for the 12-h period: The corresponding arrows point upward in Fig. 5, indicating that Trump pos/neutral is leading Clinton pos/neutral, while they mostly point downward in Fig. 6, indicating that Clinton neg is leading Trump neg. In other words: The days before the election, the intense regular (periodic) media traffic on Instagram was such that media opposing Clinton were uploaded faster than media supporting (or neutral towards) Clinton, and media supporting (or neutral towards) Trump were uploaded faster than media opposing Trump. Clinton supporters and Trump opponents were relatively sluggish the days before the election. Therefore, as far as week 2 is concerned, there is evidence neither in favor of hypothesis H3 nor hypothesis H4; while they are pending at the 24-h period, the 12-h period suggests the rejection of both.

For each week, six pairs can be formed from the four series under investigation,[9] resulting in a total of 18 heat maps of cross-wavelet powers, such as Figs. 5 and 6. These heat maps do reveal information about the phase shift between two periodic time series; it is, however, more practical to extract the time series of arrow angles directly from the cross-wavelet transformation in order to systematically analyze the phase shift between two series. This is why we do not show the entire set of 18 heat maps in this paper — the arrow angle time series (see Fig. 8) provides more precise information in our context.

4.2. Phase shifts: who was when ahead?

The goal in this section is to analyze phase shifts of pairs of time series (shown in Fig. 2) at period 12 h; this will reveal which series was leading.

[9] The binomial coefficient "4, choose 2" equals 6.

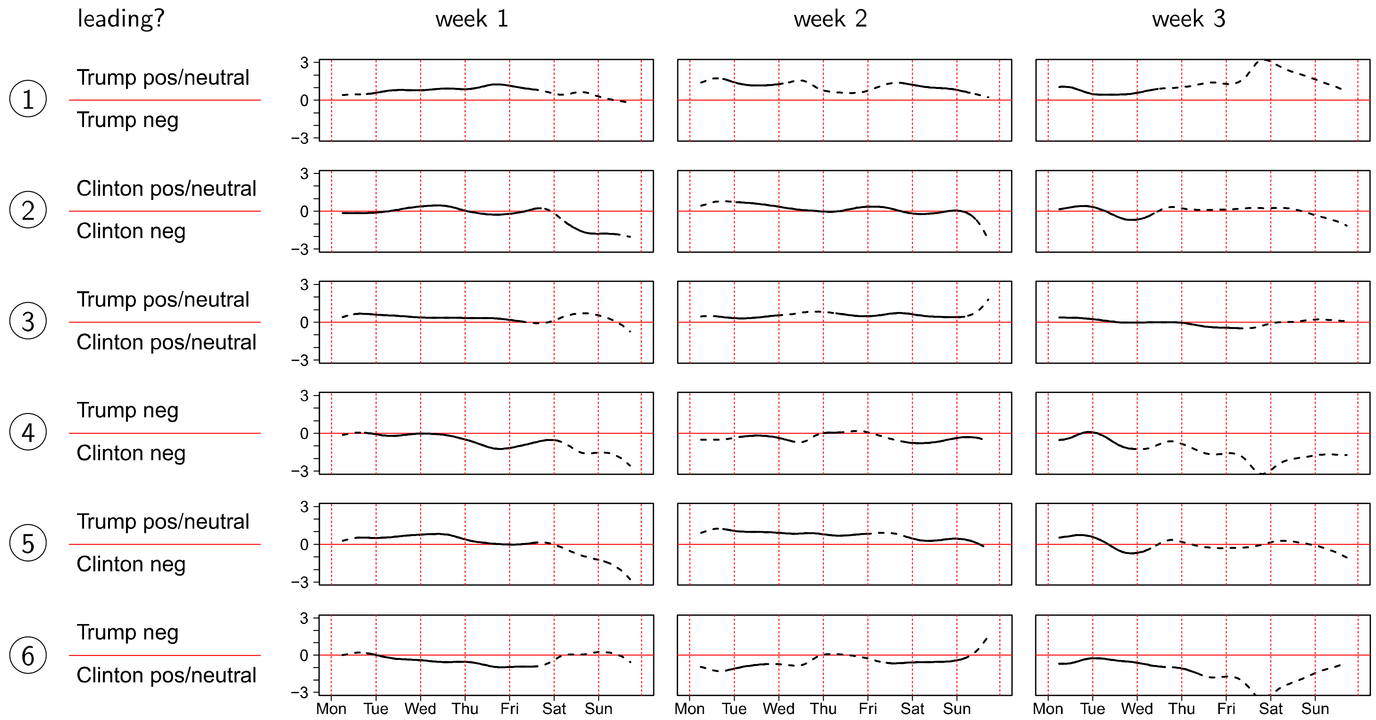

Fig. 8. Time series of phase differences (in hours) at period 12 h.

At period 12 h, a quarter-cycle ($= \pi/2 = 90^\circ$, as displayed in Fig. 4) corresponds to 3 h. For each point in time, indicated on the horizontal axis of Figs. 5 and 6, the angle of an arrow at period 12 h thus translates into a phase shift measured in hours. The result is displayed in Fig. 8: There are six pairs, indicated by the label, times three weeks, each line representing the time series of angles now expressed in hours (the vertical axis). Whenever a line is above (below) zero, the series designated by the upper (lower) label is leading by the number of hours indicated on the vertical axis. The lines are plotted only within the cone of influence. A dashed line means that period 12 h is not statistically significant at the 10% level; nevertheless these parts can be interpreted in the spirit of descriptive statistics. The series in Fig. 8 are arranged as follows:

- ①, ②: same candidate, cross-attitude
- ③, ④: cross-candidate, same attitude
- ⑤, ⑥: cross-candidate, cross-attitude

For example, the angles of the arrows in Fig. 5 (Fig. 6) at period 12 h can be found in row ③ (row ④, respectively), week 2. The series in Fig. 8 permit the following description and interpretation:

Week 1: Monday, September 19, through Sunday, September 25; the week before the presidential debates.

- Monday through Friday: Trump pos/neutral is mostly leading Trump neg (average phase shift: 50 min), Clinton pos/neutral (21 min), and Clinton neg (24 min); see ①, ③, ⑤.
- Towards the end of week 1 (Saturday and Sunday): Clinton neg is leading Clinton pos/neutral (average phase shift: 94 min), Trump neg (102 min), and Trump pos/neutral (88 min); see ②, ④, ⑤.
- The main features of week 1 are: Trump pos/neutral is mostly leading; Clinton neg is gaining momentum towards the end of week 1; Trump neg is lagging behind.

Week 2: Monday, October 31, through Sunday, November 6; the week after the presidential debates and before the election.

- Throughout the week, Trump pos/neutral is leading Trump neg (average phase shift: 64 min), Clinton pos/neutral (35 min), and Clinton neg (44 min); see ①, ③, ⑤.
- Most of the time, Clinton pos/neutral as well as Clinton neg are leading Trump neg (average phase shift: 23 and 30 min); see ①, ④, ⑥.
- The main features of week 2 are similar to those of week 1: Trump pos/neutral is mostly leading; Trump neg is again lagging behind.

Week 3: Monday, November 14, through Sunday, November 20; the week after the election.

- Except for ①, period 12 h is significant only for the first two or three days of the week, suggesting less intense upload traffic in week 3 — in spite of the still growing number of media postings, see Fig. 1 and Table 1.
- Trump neg is now lagging behind any other series; see ①, ④, ⑥.
- The main feature of week 3 is the reduced media upload traffic at period 12 h; Trump pos/neutral is still leading Trump neg; the latter is again lagging behind.

Summing up, Trump pos/neutral is almost always leading Trump neg; an analogous phenomenon cannot be observed for Clinton. Trump neg is mostly sluggish; Clinton neg is much faster. In addition, Trump pos/neutral is also leading Clinton pos/neutral, albeit its advance is less in comparison and limited to the pre-election period. This leads to the conclusion that both hypotheses H3 and H4 are rejected at the 12-h period. Even though details of the media upload process differ with respect to their periodic properties across the three weeks under investigation, hypothesis

H5 cannot be fully rejected: the crucial features at the 12-h period are very similar throughout the three weeks.

It seems like the approaching election day was a wake-up call for Clinton supporters (and Trump opponents); they began posting media more eagerly. But they were still rather sluggish in doing that: Trump supporters (and Clinton opponents) won the "Instagram battle".

4.3. Robustness of the results

Several issues of robustness should be addressed to make sure that the results presented above are valid and not artifacts produced by random phenomena.

4.3.1. Different magnitude of upload series

As shown in Section 2 (see in particular Table 2), wavelet transformation relates time series of differences which have different magnitudes. This does not affect the results of cross-wavelet analysis, however (see Section 3). This also implies that, from the perspective of magnitude, it is immaterial how many hashtag series are added up to yield a series to be analyzed (five for Trump pos/neutral, four each for Clinton pos/neutral and Trump neg, three for Clinton neg; see Table 1). On the other hand, a pertinent question is: Is there any *one* constituent series which is responsible for our results, while other series produce very different outcomes? This aspect of robustness will be discussed next.

4.3.2. Is there one series of differences with a dominating impact?

For Monday through Friday of week 1, it was found that Trump pos/neutral is mostly leading Trump neg, Clinton pos/neutral, and Clinton neg. Repeating this analysis when replacing Trump pos/neutral with one of its constituents (see Table 1), it is found that similar statements hold: Any one constituent of Trump pos/neutral is almost always leading the other three (aggregate) series, with positive average phase shifts; the latter were uniformly largest in the case of #trumpforpresident (the "most aggressive one" in this respect) and uniformly smallest in the case of #trumptrain (the "most sluggish one"). In the same vein, it turns out that Saturday and Sunday, each of the three constituents of Clinton neg is leading Clinton pos/neutral, Trump neg, and Trump pos/neutral.

The results for week 2 are confirmed as well: Each constituent of Trump pos/neutral is leading the other three series most of the time; all average phase shifts are positive. A similar picture appears for the constituents of Trump neg, all of which lag behind the three other (aggregate) series most of the time, with the sole exception of #dumptrump, which is leading Clinton neg about half the time in week 2. It could be said that Clinton neg is about as sluggish as #dumptrump in week 2; all other constituents of Trump neg are more sluggish than Clinton neg.

For week 3, each constituent of Trump pos/neutral is leading Trump neg all the time, but they are more or less on a par with Clinton pos/neutral and Clinton neg. An exception is #trumpforpresident, which is still leading both aggregate Clinton series.

These considerations show that the results reported in Sections 4.1 and 4.2 reflect a systematic pattern and are not produced by random effects. Substituting a constituent series instead of the aggregate series would have led to very similar results.

4.3.3. Methodological alternatives

The analysis undertaken in the present study rests essentially on wavelet analysis of the time series of hourly differences. The 12 series under consideration (see Fig. 2) display strong autocorrelation at lags 12 h and 24 h, which confirms the results obtained via wavelet analysis. It would then also be possible to analyze these series using, for example, a seasonal ARMA model. There are two reasons why this approach was not taken here: (i) ARMA and similar models are confined to situations in which it seems appropriate that time proceeds in discrete steps. Even though this is the case for our input data, wavelet analysis — with its capability to smooth a given time series in continuous time — was found more appropriate. (ii) Again due to their discrete-time nature, ARMA and related models provide no means to measure the phase shift between two periodic time series.

4.3.4. Analyzing other popular — non-political — hashtags

For purposes of comparison, we have analyzed time series of Instagram media uploads with hashtags referring to thirteen popular fashion labels[10] with the same methodology and for the same three weeks. The periodic behavior of these series is heterogeneous. Only three of them (in contrast, all four candidate-related series) have a significant 12-h period during all three weeks. For example, consider the series of hourly differences of hashtags #louisvuitton and #gucci[11] during week 2, which have about half the magnitude of the series Trump pos/neutral (the latter is reported in Table 2).[12] Both have a significant period of 24, but not 12, hours. This is visible in Figs. 10 and 11 (for #louisvuitton) in Appendix B, which are analogous to Figs. 5–7. The two series are more or less synchronous at period 24 h, which is not the case for a majority of pairs among fashion label hashtags. The periodic pattern of fashion-related hashtag series is thus not homogeneous and can differ markedly from that of the candidate-related series, which confirms that the patterns we detected in the latter are not commonly found in arbitrary time series.

5. Discussion

The present study surveys the dynamics of candidate-related Instagram media postings around the US presidential elections of 2016. An implicit assumption is made when trying to interpret our results beyond a merely technical aspect, namely that the media postings in question are actually associated with Instagram users expressing their political opinion and having the right to vote. However, retrieving the geo-tagging of media was not possible for this study, and bot-induced interference cannot be ruled out either. In addition, voting intentions of Instagram users may differ from their actual voting behavior.

A glance at the demographic characteristics confirms that Instagram users are probably not representative of the US electorate. A rough idea of the age distribution of Instagram users can be inferred indirectly from Facebook's "Ads Manager"[13]: In February 2017, ages 18–30 account for more than 47%, ages 18–35 (the generation labeled "Millennials" by demographers) for more than 60% of users; female users outnumber male users by 5–10 percentage points, depending on the age group. This contrasts with about 31% of voters in the age group 18–35 in the 2016 US

[10] These are (in alphabetical order): #adidas, #burberry, #cartier, #dior, #gucci, #hermes, #hm, #louisvuitton, #nike, #prada, #ralphlauren, #tiffany, #zara. The corresponding thirteen brands were the most valuable brands worldwide in 2016 in the sectors apparel, luxury, and sporting goods, according to Interbrand; see http://interbrand.com/best-brands/best-global-brands/2016. Accessed on August 4, 2017.

[11] Louis Vuitton and Gucci both belong to the luxury segment of Interbrand's classification.

[12] This remark should only emphasize that the series have enough "volume" for a meaningful wavelet analysis and do not contain sub-series of zeros. As we have seen above, the absolute magnitude of a series does not affect its wavelet transform.

[13] See https://www.facebook.com/business/help/200000840044554; age data were retrieved on February 5, 2017.

presidential election.[14] In this regard, it is telling to connect election predictions with our results.

As of early October, the CIRCLE phrased[15]: "Although the overall level of support for Trump is very low, his supporters may be more energized. The Millennials who say that they would vote for Donald Trump were more likely to say that they will cast a ballot (76%) than Clinton supporters (68%)." They add that Clinton led Trump by 21 percentage points (49% vs. 28%) among likely young voters. Results from the monthly GenForward surveys,[16] putting racial and ethnic heterogeneity of Millennials into particular perspective, had shown a persistent (September through November) majority of young African and Asian Americans, and a plurality of Latino/as, saying that they would vote for Clinton. Non-Hispanic white youth, in September still more or less evenly divided, had favored Clinton over Trump by 14 percentage points in early October, her advantage dwindling to 3 percentage points by early November.

Our study reveals that Trump supporters (and Clinton opponents), from September 2016 onward, were posting Instagram media more energetically than Clinton supporters (Trump opponents) — not in terms of numbers alone, but particularly in terms of speed, and it involved a substantial segment of the population that had actually been expected to vote for Clinton rather than for Trump.

If a minority — Trump supporters early October — becomes very active on social media, can it change the outcome of an election? This question may be discussed on the basis of a random walk problem posed in a 1940 article by McCrea and Whipple (1940). Only recently, the framework has been related to the consensus problem in multi-agent systems by Chatzigiannakis, Dolev, Fekete, Michail and Spirakis (2009). This model assumes that there is a population of size N and there are two candidates, X and Y. Each voter is in one of the states "support X", "support Y", "undecided". The process of consensus finding is then modeled as the movement of a particle in a two-dimensional grid. The horizontal (vertical) axis gives the current number of votes for X (Y, respectively), and absorbing state $(0, N)$ (and $(N, 0)$) means that Y (or X, respectively) has won unanimously. (Of course, no unanimous vote needs to be achieved in a presidential election.) If X gains (or loses) a vote, the particle makes a horizontal step to the right (left, respectively), and similar remarks apply for Y. No candidate is preferred if each movement has probability $1/4$. Then, given initial votes of $N/2$ for X, $N/4$ for Y (the minority), Y will win the unanimous vote in the end with probability $3/8$, which is larger than Y's initial share. If state transition rules prefer Y, this probability will obviously grow. The power of an active minority in social media could be discussed along these lines, and our findings may provide a motivation to do so. This may also contribute to the discussion initiated by John Quelch and Thales Teixeira (see Introduction).

6. Summary and conclusions

The present study is an effort to analyze the timing of Instagram media postings in connection with the 2016 US presidential election. A selected set of 16 hashtags is used to determine whether a posting was intended to support or oppose a candidate, constituting four categories: Hillary Clinton vs. Donald Trump, supporters vs. opponents. From September through November 2016, hourly readings of the number of Instagram postings with each hashtag were recorded. The main goal of our study was to compare the resulting time series, especially the net number of media uploads per hour, and to investigate their periodic structure using cross-wavelet methodology. We focused on three full weeks: week 1, before the presidential debates; week 2, between the last presidential debate and the election day; and week 3, after the election.

Throughout the time period under consideration, the number of Instagram media postings in favor of, or neutral towards, Trump was massively higher than any other category. Clinton-related postings rose significantly in number, but were still behind on the election day. By comparison, there were far fewer media uploads opposing either candidate, although the number of uploads opposing Trump also grew significantly.

It turns out that each time series has significant 12- and 24-h periods. While the latter mirrors daily routine behavior, the period of 12 h reflects high "upload traffic intensity". We find that, at the 12-h period, the time series of Trump supporters was almost always leading Trump opponents: Trump supporters were faster, or more eager, to upload media. In addition we found that Trump supporters as well as Clinton opponents were also leading Clinton supporters the days before the election. The approaching election day may have given Clinton supporters a wake-up call; they started to post more media. Nevertheless, they were still quite sluggish when compared to the other groups.

Keeping in mind that 60% of Instagram users belong to the age group 18—35 years, our results are not in line with expectations, enunciated before the election on November 8, that a majority of young people is likely to vote for Clinton, and not in line with models forecasting "a big victory for Hillary Clinton"[17] either. It seems that Trump supporters (and Clinton opponents) won the "Instagram battle".

One lesson to learn from our findings might be that the behavior of Instagram users contains important clues about the effectiveness of political campaigns prior to elections. Indeed, a recent study by Pittman and Reich (2016) found that visual social media, as realized by platforms like Instagram, conform to the need of individuals for communicating thoughts and feelings quicker and more effectively than text-based media on platforms like Twitter. Monitoring media uploads on Instagram and the analysis of the upload behavior could provide a real-time barometer of public opinion and sentiments for policymakers. Detecting media uploads by "fake accounts", run by computer-generated "social bots" that have the potential to fool trending algorithms, poses an intricate challenge in this context that has to be tackled. Instagram has already taken action: In 2014, it has "purged millions of fake accounts, in an effort to provide more accurate numbers to marketers." Nevertheless, it was said that about 8% of Instagram accounts still appeared to be run by computer-generated bots.[18] For decades, computer scientists have tried to write software that passes the Turing test.[19] "Social bots"

[14] See http://www.pewresearch.org/fact-tank/2016/05/16/millennials-match-baby-boomers-as-largest-generation-in-u-s-electorate-but-will-they-vote/. Accessed on February 7, 2017.

[15] CIRCLE (The Center for Information & Research on Civic Learning and Engagement) at Tufts University, Massachusetts, USA; see http://civicyouth.org/exclusive-circle-poll-on-millennial-attitudes-about-presidential-election-contact-by-campaignsparties/. Accessed on February 23, 2017.

[16] GenForward: A survey of the Black Youth Project with the AP-NORC Center for Public Affairs Research; see http://genforwardsurvey.com/. Accessed on February 7, 2017.

[17] See http://money.cnn.com/2016/11/01/news/economy/hillary-clinton-win-forecast-moodys-analytics/. Accessed on February 10, 2017.

[18] "Fake accounts still plague Instagram despite purge, study finds", The Wall Street Journal, June 15, 2015; https://blogs.wsj.com/digits/2015/06/30/fake-accounts-still-plague-instagram-despite-purge-study-finds/. Accessed on September 20, 2017.

[19] Turing's imitation game, see Turing (1950); commonly known as the Turing test.

may have reached this goal — maybe temporarily. In spite of these achievements, the Turing test is not out of date: Today's challenge may well be to make sure a social actor is actually human.

Our methodology in this paper — essentially, to analyze the timing of media uploads to a social network on the basis of wavelets — has a vast potential beyond applications in the political arena. Assuming that public interest in a timely topic manifests itself also in media upload intensity, the same approach to investigate human-computer interaction can be useful to understand human behavior in many other fields as well. More specifically, the method lends itself very well to gauge which among several competing hashtags (representing, for example, competing ideas or competing brands) is attracting interest more intensely. Our brief example of the analysis of fashion-related hashtags (here only used for purposes of comparison) may have given a glimpse into further possibilities of the methodology, for example in the realm of business analytics.

Acknowledgments

The authors would like to thank three anonymous referees whose constructive criticism helped to improve especially the discussions in this paper.

Appendix A. The mathematics of wavelet and cross-wavelet transformation

When studying periodic phenomena in time series, cycles of different frequencies and of limited duration may overlap, necessitating a decomposition of the series in the time and frequency domains simultaneously. Wavelets are utilized to tackle the arising time and frequency resolution dilemma (resulting from the Heisenberg uncertainty principle). We will briefly outline the mathematical concepts of wavelet analysis as far as relevant for this study, using the notation of the R package *WaveletComp* (Rösch & Schmidbauer, 2014).

Morlet wavelet transformation

We adopt the continuous Morlet wavelet transform, which yields a finer resolution than discrete wavelet transforms, and is complex-valued and therefore highly redundant and information-preserving with any careful selection of parameters; cf. Morlet, Arens, Fourgeau and Giard (1982a, 1982b). It provides information on both amplitude and phase. The latter is a prerequisite for the analysis of two time series with respect to synchronicity at a given period. It also provides a method to reconstruct the original series.

The "mother" Morlet wavelet, defined by

$$\psi(\eta) = \pi^{-1/4}\, \mathrm{e}^{\mathrm{i}6\eta}\, \mathrm{e}^{-\eta^2/2} \tag{1}$$

(with a particular choice of six oscillations per one revolution of 2π (radians), which is the preferred value for computational purposes in literature since it makes the Morlet wavelet approximately analytic), is depicted in Fig. 9.

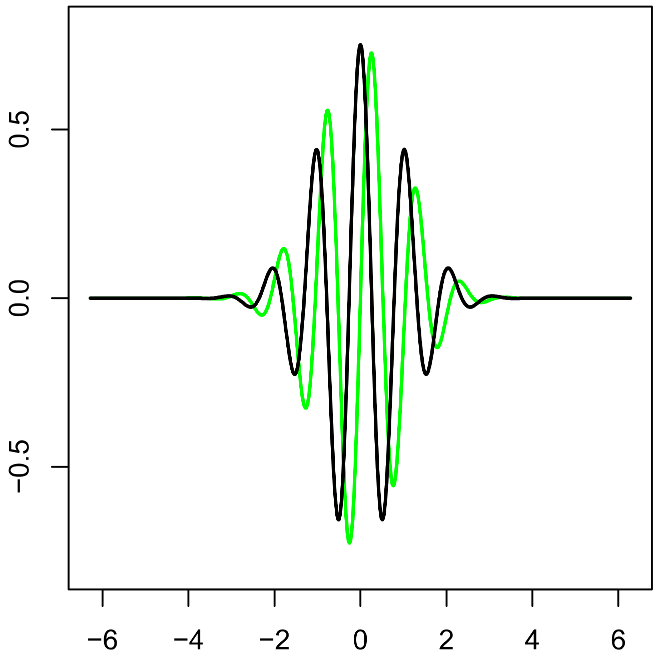

Fig. 9. The Morlet mother wavelet — real part (black) and imaginary part (green).

Characteristics of any wavelet are its compact support across time and zero area underneath the curve (so as to make its energy equally distributed, and the multiplication by a constant signal, without periodic fluctuations, results in an area of zero). Intuitively, wavelet transformation decomposes the time series at hand into a set of base functions, called the "wavelet daughters", which are generated from the mother wavelet by translation in time and by scaling (compressing and stretching). Accordingly, an increase in scale, which corresponds to a decrease in frequency, is at the expense of time resolution, while a decrease in scale and increase in frequency sacrifices on frequency resolution. The transformation results in a matrix of (complex-valued) coefficients, namely the (complex) wavelet transform of the time series (x_t) as a function of translation τ and scale s:

$$\mathrm{Wave}(\tau, s) = \sum_t x_t \frac{1}{\sqrt{s}} \psi^\star\left(\frac{t-\tau}{s}\right) \tag{2}$$

(taking the rectification factor $1/\sqrt{s}$ into account, see Liu, Liang and Weisberg (2007)) with $\star$ denoting the complex conjugate. The shift of the wavelet daughters' translation is determined by an increment $\mathrm{d}t$, the sampling resolution in the time domain. The scale is usually set to a fractional power of 2, a "voice" in an "octave" (according to octaves in music denoting intervals between pitches with half or double of each other's frequency). An increment of $\mathrm{d}j$ determines the sampling resolution in the frequency domain, and $1/\mathrm{d}j$ is the number of voices per octave. In our study, adopting the Morlet wavelet with 6 oscillations, the Fourier factor $2\pi/6$ is used to convert scales to periods, and, for ease of interpretation, it is the period which is set to a fractional power of 2, entailing the corresponding values of scale (this is the default setting in *WaveletComp* (Rösch & Schmidbauer, 2014)); sampling resolutions in our study are set to $\mathrm{d}t = 1$ and $\mathrm{d}j = 1/100$ (i.e. 100 voices per octave in scale direction).

The local amplitude of any periodic component of (x_t) and how it evolves with time can then be retrieved from the modulus of its wavelet transform. Its square has an interpretation as wavelet energy density, which is called the wavelet power spectrum and is usually displayed as a "heat map" in the time-scale (or, correspondingly: time-frequency or time-period) domain:

$$\text{Power}(\tau, s) = \text{Ampl}(\tau, s)^2 = \frac{1}{s} \cdot \left|\text{Wave}(\tau, s)\right|^2 \tag{3}$$

In wavelet applications, it is common to raise the squared local amplitude to a further power in order to accentuate contrast in the corresponding heat map; see, e.g. Percival and Walden (2000).

In case of a white noise process, the expected wavelet power at each time and scale, disregarding the proportionality factor $1/s$, corresponds to the process variance. Therefore, in applications of wavelet methodology, it is conventional to standardize the time series at hand, after detrending it, to obtain a measure of the wavelet power which is relative to unit-variance white noise and directly comparable to results of other time series. (The detrending of an input time series by local polynomial regression is another optional feature of *WaveletComp* (Rösch & Schmidbauer, 2014). In our application to hourly media postings, detrending was not necessary.)

The complex nature of the Morlet wavelet bears information about the local wavelet phase, that is: $\text{phase}(\tau, s) = \text{Arg}(\text{Wave}(\tau, s))$, which is an angle in the interval $[-\pi, \pi]$ measuring displacements of any periodic component of (x_t) relative to a localized origin in the time domain; this is utilized in our study of time series synchronicity.

Cross-wavelet transformation and phase differences

The concepts of cross-wavelet analysis are appropriate for a comparison of the frequency contents of two time series, and conclusions about their synchronicity. The cross-wavelet transform of two time series, say (x_t) and (y_t), with respective wavelet transforms Wave.x and Wave.y, decomposes the Fourier co- and quadrature-spectra in the time and frequency domains simultaneously:

$$\text{Wave.xy}(\tau, s) = \frac{1}{s} \cdot \text{Wave.x}(\tau, s) \cdot \text{Wave.y}^{\star}(\tau, s), \tag{4}$$

(taking the rectification factor $1/s$ into account, see Veleda, Montagne and Araujo (2012)) with translation parameter τ and scale parameter s Its modulus has the interpretation as cross-wavelet power (sometimes called cross-wavelet energy) and lends itself to an assessment of the similarity of the two series' wavelet power with respect to any periodic component and how it evolves with time:

$$\text{Power.xy}(\tau, s) = |\text{Wave.xy}(\tau, s)| \tag{5}$$

Again, a heat map is the usual way to visualize the cross-wavelet power spectrum. Power averages illustrate the prominence of certain periodic components across time.

In a geometric sense, the cross-wavelet power is the analog of the covariance, and like the latter, it depends on the unit of measurement of the series involved and may not be ready for interpretation with regard to the degree of association of the two series. The concept of wavelet coherency, which is analogous to correlation, may remedy this. For the purpose of this study, we abstained from developing this step any further since conclusions about periods drawn from the cross-wavelet power spectrum are supported by comparisons with the power spectra in the univariate case.

It is crucial for our purposes in the present study that the cross-wavelet transform carries information about the synchronicity of the two series in terms of the local phase advance of any periodic component of the one series with respect to the corresponding component of the other:

$$\begin{aligned} \text{Angle.xy}(\tau, s) &= \text{Arg}(\text{Wave.xy}(\tau, s)) \\ &= \text{phase.x}(\tau, s) - \text{phase.y}(\tau, s) \end{aligned} \tag{6}$$

This so-called phase difference of x over y at each time and scale equals the difference of individual local phase displacements (relative to a localized origin) when converted into an angle in the interval $[-\pi, \pi]$. An absolute value less (or larger) than $\pi/2$ indicates that the two series move in phase (anti-phase, respectively) at the scale (or, equivalently: frequency, period) in question, while the sign of the phase difference shows which one is the leading series in this relationship. Information on phase differences at certain periods can be retrieved and analyzed separately in *WaveletComp* (Rösch & Schmidbauer, 2014).

Assessing the statistical significance of periodic components

The statistical significance of the patterns emerging from (cross-) wavelet transformation is assessed by comparison with simulated white noise (500 surrogates for each time series). In our context, the null hypothesis of white noise to be tested reflects an agnostic statement: there is no periodicity in the series at hand. Localized p-values in the time and scale domains are derived from simulated shares of exceedances of power levels attained by the time series to be tested, following the approach by Aguiar-Conraria and Soares (2011). In addition, selective time series reconstruction tools support the identification of "powerful", i.e. substantial, periodic constituents of the series.

Appendix B. Periodic properties of hashtag series #louisvuitton and #gucci, week 2

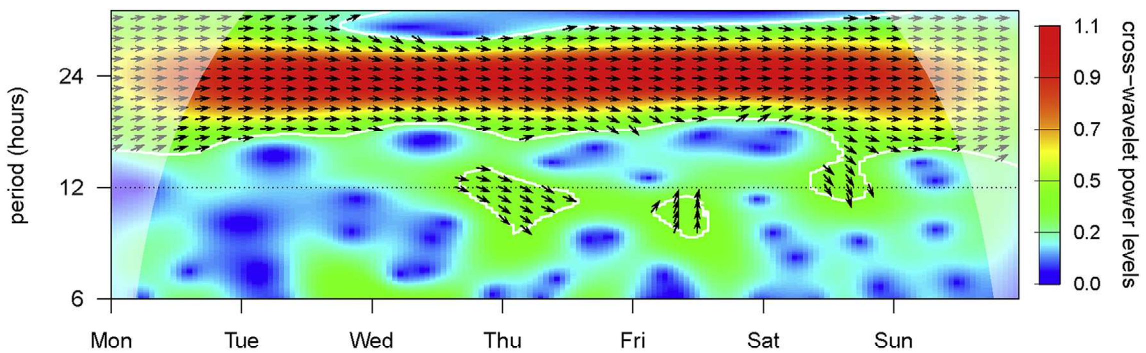

Fig. 10. Cross-wavelet power spectrum, #louisvuitton over #gucci, week 2.

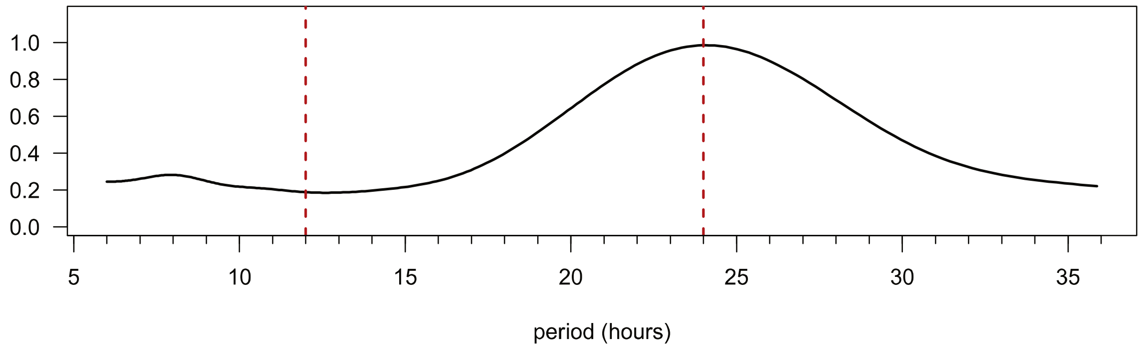

Fig. 11. Average wavelet power, #louisvuitton, week 2.

References

Aguiar-Conraria, L., Azevedo, N., & Soares, M. J. (2008). Using wavelets to decompose the time–frequency effects of monetary policy. *Physica A, 387*, 2863–2878.

Aguiar-Conraria, L., & Soares, M. J. (2011). *The continuous wavelet transform: A primer*. NIPE Working Paper Series 16/2011. Braga, Portugal: Universidade do Minho http://www.eeg.uminho.pt/economia/nipe.

Aronson, E. D. (2012). Cyber-politics: How new media has revolutionized electoral politics in the United States. *Colgate Academic Review, 9*. Article 7. http://commons.colgate.edu/car/vol9/iss1/7

Banning, S. A. (2006). Third person effects on political participation. *Journalism and Mass Communication Quarterly, 83*, 785–800.

Bovet, A., Morone, F., & Makse, H. A. (2016). *Predicting election trends with Twitter: Hillary Clinton versus Donald Trump*. Working Paper. New York, USA: City College of New York https://arxiv.org/abs/1610.01587.

Carmona, R., Hwang, W.-L., & Torrésani, B. (1998). *Practical time frequency analysis. Gabor and wavelet transforms with an implementation in S*. San Diego, USA: Academic Press.

Chatzigiannakis, I., Dolev, S., Fekete, S. P., Michail, O., & Spirakis, P. G. (2009). Not all fair probabilistic schedulers are equivalent. In T. F. Abdelzaher, M. Raynal, & N. Santoro (Eds.), *Lecture notes in computer science: Vol. 5923. Principles of distributed systems. OPODIS 2009*. Berlin,Germany: Springer.

Fernandes, J., Giurcanu, M., Bowers, K. W., & Neely, J. C. (2010). The writing on the wall: A content analysis of college students' Facebook groups for the 2008 presidential election. *Mass Communication and Society, 13*(5), 653–675.

Gayo-Avello, D., Metaxas, P. T., & Mustafaraj, E. (2011). Limits of electoral predictions using Twitter. In *Proceedings of the fifth international AAAI conference on weblogs and social media*. Barcelona, Catalonia, Spain, July 17-21, 2011.

Gencay, R., Selcuk, F., & Whitcher, B. (2001). *An introduction to wavelets and other filtering methods in finance and economics*. San Diego, USA: Academic Press.

Giannoulakis, S., & Tsapatsoulis, N. (2016). Evaluating the descriptive power of Instagram hashtags. *Journal of Innovation in Digital Ecosystems, 3*(2), 114–129. https://doi.org/10.1016/j.jides.2016.10.001.

Hu, Y., Manikonda, L., & Kambhampati, S. (2014). What we Instagram: A first analysis of Instagram photo content and user types. In *Proceedings of the eigth international AAAI conference on weblogs and social media*. Ann Arbor, Michigan, USA, June 1-4, 2014.

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science, 59*(3), 690–707.

Jahanbakhsh, K., & Moon, Y. (2014). *The predictive power of social media: On the predictability of U.S. presidential elections using Twitter*. Research Report, arXiv: 1407.0622. Ithaka, New York, USA: Cornell University Library https://arxiv.org/abs/1407.0622.

Johnson, T. J., & Perlmutter, D. D. (2010). Introduction: The Facebook election. *Mass Communication and Society, 13*(5), 554–559.

Kushin, M. J., & Yamamoto, M. (2010). Did social media really matter? College students' use of online media and political decision making in the 2008 election. *Mass Communication and Society, 13*(5), 608–630.

Lee, T.-T. (2005). The liberal media myth revisited: An examination of factors influencing perceptions of media bias. *Journal of Broadcasting and Electronic Media, 49*(1), 43–64.

Liu, Y., Liang, X. S., & Weisberg, R. H. (2007). Rectification of the bias in the wavelet power spectrum. *Journal of Atmospheric and Oceanic Technology, 24*, 2093–2102.

McCrea, W. H., & Whipple, F. J. W. (1940). XXII. — Random paths in two and three dimensions. *Proceedings of the Royal Society of Edinburgh, 60*(3), 281–298.

Metaxas, P. T., Mustafaraj, E., & Gayo-Avello, D. (2011). How (not) to predict elections. In *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing 2011, Boston, Massachusetts, USA, October 9-11*. https://doi.org/10.1109/PASSAT/SocialCom.2011.98.

Morlet, J., Arens, G., Fourgeau, E., & Giard, D. (1982a). Wave propagation and sampling theory — Part I: Complex signal and scattering in multilayered media. *Geophysics, 47*, 203–221.

Morlet, J., Arens, G., Fourgeau, E., & Giard, D. (1982b). Wave propagation and sampling theory — Part II: Sampling theory and complex waves. *Geophysics, 47*, 222–236.

Percival, D. B., & Walden, A. T. (2000). *Wavelet methods for time series analysis*. New York, USA: Cambridge University Press.

Pittman, M., & Reich, B. (2016). Social media and loneliness: Why an Instagram picture may be worth more than a thousand Twitter words. *Computers in Human Behavior, 62*, 155–167.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria.: R Foundation for Statistical Computing. http://www.R-project.org/.

Rösch, A., & Schmidbauer, H. (2014). *WaveletComp: Computational wavelet analysis*. R package version 1.01 https://cran.r-project.org/web/packages/WaveletComp/index.html.

Sheldon, P., & Bryant, K. (2016). Instagram: Motives for its use and relationship to narcissism and contextual age. *Computers in Human Behavior, 58*, 89–97.

Sudhahar, S., Veltri, G. A., & Cristianini, N. (2015). Automated analysis of the US presidential elections using Big Data and network analysis. *Big Data & Society*. January–June 2015.

Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society, 79*(1), 61–78.

Turing, A. M. (1950). *Computing, machinery and intelligence*. *Mind* LIX No. 236 (pp. 433–460).

Veleda, D., Montagne, R., & Araujo, M. (2012). Cross-wavelet bias corrected by normalizing scales. *Journal of Atmospheric and Oceanic Technology, 29*, 1401–1408.

Woolley, J. K., Limperos, A. M., & Oliver, M. B. (2010). The 2008 presidential election, 2.0: A content analysis of user-generated political Facebook groups. *Mass Communication and Society, 13*(5), 631–652.

Yamamoto, M., & Kushin, M. J. (2014). More harm than good? Online media use and political disaffection among college students in the 2008 election. *Journal of Computer-Mediated Communication, 19*, 430–445.