SUPPLEMENTAL MATERIALS FOR:

Did an Ebola Outbreak Influence the 2014 U.S. Federal Elections? (Hint: Only if you ignore autocorrelation).

Leonid Tiokhin[1,2] ,Daniel Hruschka[1,2]

[1] Arizona State University School of Human Evolution and Social Change

[2] Arizona State University Center for Evolution & Medicine

MATERIALS AND METHODS

Sample. We use the same data and variables analyzed by BHS in their analyses. Study 1 considered national-level polling data and Ebola search volume data between September 1 and November 1 aggregated across the U.S. BHS analyzed the relationship between U.S. country-level voter intentions and U.S. country-level Google searches for "Ebola". In the time frame of September 1, 2014 to November 1, 2014, aggregate nationwide polling results were available for 24 days – 9 days preceding the initial Ebola outbreak and 15 days following the initial outbreak. BHS conducted analyses using this entire sample, as well as an 8-value subset of this sample (the last week of September and first week of October).

Study 2 considered the same variables and time frame as study 1, but aggregated at the state Senate-election level for 34 out of 36 elections in which data was available (Kansas and Alabama had insufficient data and were excluded). The variables were either measured at the level of country (Google searches for "Ebola") or individual Senate elections (Within-state voter intentions for each senate election). Of these 34 states, two (Hawaii and Rhode Island) had outlier state-Senate Voter Intention Index scores, and BHS conducted analyses both with and without them. BHS excluded Virginia from moderation analyses of Partisan Voter Index, at it had score of 0 (See "Variables").

While we attempted to use the same dataset as BHS, there are some differences between our dataset and theirs. While the original BHS analyses for study 1 contain data points for each of 24 days, our study 1 analyses contain change scores for 23 pairs of days. Controlling for first-order autocorrelation requires calculating changes from one data value to the next. This results in the loss of the first data-point in the time series, for which changes cannot be calculated. In study 2, calculating changes to control for first-order autocorrelation results in the exclusion of six state-Senate elections (Idaho, Mississippi, Nebraska, Tennessee, West Virginia, and Wyoming). These have insufficient data points to calculate correlations (i.e. less than three) and are excluded from our analyses. Because our resulting sample of state-senate elections differs from that of BHS, we also re-conduct the original BHS analyses, excluding these states. In the main commentary, Table 1 summarizes exclusions, inclusions, and results across all analyses

Variables. *Voter Intention Index (VII).* Voter intentions were estimated from nationwide polling data from Pollster, a poll-aggregation website. For each day on which data was available, The

Pollster website specified the percentage of potential voters within the United States who indicated an intention to vote for each candidate from either the Republican or Democratic Party. To generate VII values, BHS subtracted the percentage of voters who intended to vote Democrat from the percentage of voters who intended to vote Republican. As such, positive VII values indicated nationwide preference for Republican candidates and negative VII values indicated nationwide preference for Democrats. For study 2, BHS used the same procedure as in study 1 to generate the VII for each state-Senate election on the days such polling data was available.

Ebola Search Volume Index (ESVI). For the time period of August 26, 2014 to November 1, BHS obtained internet search volume data for the term "Ebola" from Google Trends. For each day from September 1, 2014 to November 1, 2014, BHS took the arithmetic mean "Ebola" daily search volume from the previous 7 days, ending on and including the specified day. This resulted in an Ebola search-volume index (ESVI). The ESVI was used in both study 1 and study 2 as the measure of Ebola's psychological salience.

Candidate Leading Polls at Time of Ebola Outbreak: In study 2, BHS used VII values to categorize each state-Senate election as being led by a Democratic (n=11) or Republican (n=22) candidate at the time of the Ebola outbreak. This categorization was based on the most recent poll preceding the outbreak.

Partisan Voter Index (PVI): BHS use data from the 2014 Cook Political Report's Partisan Voter Index (httw://www.cookpolitical.com/story/5604) as a measure of a state's "enduring political norms". While this link only lists district-specific PVI scores for 2014, BHS use state-level PVI scores in their analyses. How state-level PVI scores were generated was not reported in BHS's paper, but it appears that a state's PVI score is the arithmetic mean of the PVI scores of all districts in that state. States with positive PVI scores were categorized as generally Republican (n=19) and states with negative PVI scores were categorized as generally Democratic (n=12). Virginia had a PVI score of 0 and was excluded from PVI analyses.

ANALYTICAL STRATEGY

In study 1, BHS examined correlations between national-level VII time series: 1) ESVI time series for the 24 days in September and October for which this data was available and 2) raw Ebola search volume for the 8 days during the two-week period that included the last week of September and the first week of October. In study 2, BHS assessed correlations between state-specific VII values and ESVI values, and examined whether these correlations were greater in states that had higher initial support for Republicans.

These initial analyses did not investigate first-order autocorrelation or fully detrend the data to address it. To assess levels of first-order autocorrelation, we re-analyzed the same variables by lagging an observation from each variable (x) to the temporally adjacent observation (x + 1), and then correlated x with x+1. The resulting correlation is a measure of how predictive the value of a variable at timepoint x is of the value of that same variable at timepoint x + 1.

To remove first-order autocorrelation from the variables, we subtracted the value of that variable at time x from its value at time x+1. We then repeated the BHS analyses with these detrended variables. For study 1, we created a lagged VII variable (lagged from the temporally prior observation), and subtracted the original VII data from this lagged VII variable. This resulted in a new variable, “VII Changes”, which contained changes for 23 pairs of adjacent days compared to 24 days for which the original VII had data (1 day lost due to calculating the changes). “VII Changes” measures the marginal increase or decrease in the VII during any given time period. We used this same strategy to calculate the changes in ESVI, for each day that data on “VII Changes” was also available. This resulted in an “ESVI Changes” variable, which contained data for the same 23 days for which “VII Changes” data was available. Once we created these variables, we ran bivariate correlations, for both the entire time period for which data was available and for just the two-week period that included the last week of September and the first week of October, replicating BHS’s analysis.

For study 2, we used the same strategy as study 1 to calculate the changes in VII and ESVI for each state-Senate election. Because daily changes are calculated by subtracting a variable’s value at time x from its value at time x+1, this resulted in “State-Specific VII Changes” and “State-Specific ESVI Changes” variables of size n-1, where n is the total number of data points for that election in the original data. We then re-conduct BHS’s correlational time-series analyses, controlling for first-order autocorrelation, by using “State-Specific VII Changes” instead of “VII” and “State-Specific ESVI Changes” instead of “ESVI”.

We also directly replicate BHS’s results and report 95% Confidence Intervals for their correlations.

RESULTS

Is there temporal autocorrelation in the time series?

In study 1, first order autocorrelation is high for both ESVI ($r = 0.99$, $N = 64$, $p < 0.001$, 95% CI [0.98, 0.99]) and VII ($r = 0.95$, $N = 23$, $p < 0.001$, 95% CI [0.89, 0.98]). Further, BHS’s use of the 7-day aggregate ESVI inflates the autocorrelation above that which is present in the raw Ebola search volume data ($r = 0.90$, $N = 64$, $p < 0.001$, 95% CI [0.84, 0.94]). This is because each ESVI value is the arithmetic mean of the raw Ebola searches over the last 7 days, meaning that each ESVI value is made up of 6 of the 7 same values as the previous ESVI value. Substantial second-order autocorrelation (similarity in the direction of changes in a variable) still remains for both VII Changes ($r = 0.43$, $N = 22$, $p = 0.046$, 95% CI [0.1 0.72]) and ESVI Changes ($r = 0.60$, $N = 22$, $p = 0.003$, 95% CI [0.23, 0.81]), but is substantially attenuated. Study 2 used the same ESVI data as study 1, and first order autocorrelation for the state-Senate election VII was ($r = 0.76$, $N = 26$, $p < 0.001$, 95% CI [0.61, 0.92]). There is no evidence for second-order autocorrelation after calculating state-Senate election VII Changes ($r = 0.03$, $N = 23$, $p > .250$, 95% CI [-0.16, 0.22]). 3 state-Senate elections (Montana, South Carolina 1, and South Carolina 2) had insufficient data points to calculate autocorrelation in VII Changes and were excluded from this analysis. There is no evidence that these exclusions affect the level of autocorrelation in the raw state-Senate election VII data ($r = 0.75$, $N = 23$, $p < 0.001$, 95% CI [0.58, 0.92]).

Do internet searches for "Ebola" predict people's intentions to vote for Republican candidates at the national level?

In study 1, BHS find that the ESVI is positively and significantly correlated with the VII ($r = .51$, $p = 0.012$, $N = 24$ days, 95% CI [0.13, 0.76]). They also find a large, positive but non-significant correlation between raw Google Trends "Ebola" search volumes and the VII for the two-week period that included the last week of September and the first week of October ($r = 0.61$, $p=0.111$, $N = 8$ days, 95% CI [-0.17, 0.92]).

When removing first-order autocorrelation, the correlation between VII Changes and ESVI Changes is substantially lower, and no longer significant ($r = 0.30$, $p = 0.159$, $N = 23$, 95% CI [-0.12, 0.64]) for the months of September and October. If we replicate the additional BHS correlation of raw Google Trends "Ebola" search volume with VII over the two-week period that included the last week of September and the first week of October, the correlation of detrended variables is still non-significant, but also no longer has the suggestively high, positive magnitude ($r = -0.19$, $p > .250$, $N = 8$ days, 95% CI [-0.79, 0.60].

Do internet searches for "Ebola" predict people's intentions to vote for Republican candidates at the state-level?

In study 2, BHS find that, across 32 Senate elections, the arithmetic mean correlation between ESVI and the state-specific VII is statistically significant (mean $r = 0.31$, $p = 0.016$, $N = 32$, CI [0.06, 0.55]), though it drops to (mean $r = 0.24$, $p = 0.057$, $N = 34$, 95% CI [-0.01, 0.49]) if outliers Hawaii and Rhode Island are included. When considering only those 28 states with sufficient data for later comparison with a detrended analysis (at least 3 time points), the arithmetic mean correlation between ESVI and the state-specific VII is lower with the two outliers excluded (mean $r = 0.19$, $p = 0.18$, $N = 26$, 95% CI [-0.09, 0.47]) and included (mean $r = 0.12$, $p = 0.18$, $N = 28$, 95% CI [-0.16, 0.39]).

Once detrending the two key variables, the arithmetic mean correlation between "ESVI Changes" and "VII Changes" across 28 Senate elections is reduced to nearly zero and is no longer significant (mean $r = 0.04$, $p > .250$, $N = 26$, CI [-0.159, 0.232]) and is (mean $r = 0.03$, $p > .250$, $N = 28$, CI [-0.149, 0.212]) when outliers Hawaii and Rhode Island are included. Neither of these are significantly different from 0.

Does a state-level bandwagon effect moderate the relationship between internet searches for "Ebola" and people's intentions to vote for Republican candidates?

In the second part of study 2, BHS find that, across 32 Senate elections, the arithmetic mean correlation between ESVI and the state-specific VII is positive for elections in which a Republican led the polls at the time of the initial Ebola Outbreak (mean $r = 0.51$, $N = 21$, $p < 0.001$, CI [0.26, 0.75]) but not for elections in which a Democrat led the polls at that time (mean $r = - 0.08$, $N = 11$, $p > .250$, CI [-0.60, 0.44). This difference is statistically significant ($d = 0.92$, $t(30) =2.52$, $p = 0.017$). With outliers Hawaii and Rhode Island included, they find a similar pattern for Republican leading states (mean $r = 0.51$, $N = 21$, $p < 0.001$, 95% CI [0.26, 0.75]) and for Democratic leading states (mean $r = - 0.19$, $N = 13$, $p > .250$, 95% CI [-0.65, 0.27]). When considering only those 28 states that had at least 3 time points (for comparison with a later detrended analysis), the arithmetic mean correlation between ESVI and the state-specific VII is still positive and statistically significant (mean r= 0.39, N = 15, p = .021, 95% CI [0.07, 0.70]) for Republican leading states but not for Democratic leading states (mean r = - 0.08, N = 11, p > .250, CI [-0.60, 0.44]). The difference between these is not statistically significant (Welch Two Sample t-test (t = 1.69, df = 17.64, p = 0.109)). With outliers Hawaii and Rhode Island included, the pattern is similar for Republican (mean r = 0.39, N = 15, p = .021, 95% CI [0.07, 0.70]) and Democratic (mean r = - 0.19, N = 13, p > .250, 95% CI [-0.65, 0.27]) leading states.

In contrast, when removing 1st order autocorrelation, we find no evidence of a moderating effect of party leading the polls at the time of the outbreak on the correlation between Ebola searches and people's voting intentions. The arithmetic mean correlation between "ESVI Changes" and "VII Changes" across 26 Senate elections is (mean r = 0.09, N = 15, p >.250, 95% CI [-0.18, 0.35]) for elections in which a Republican led the polls at the time of the outbreak and (mean r = - 0.03, N = 11, p > .250, 95% CI [-0.38, 0.31]) for elections in which a Democrat led the polls at the time of the outbreak. This pattern holds when outliers Hawaii and Rhode Island are included for Republican leading elections (mean r = 0.09, N = 15, p >.250, 95% CI [-0.18, 0.35]) and Democrat leading elections (mean r = - 0.03, N = 13, p >.250, 95% CI [-0.32, 0.25]). Neither of these are significantly different from 0. Further, they are not significantly different from each other (Welch Two Sample t-test, outliers excluded: t = 0.59, df = 20.60, p >.250; Welch Two Sample t-test, outliers included: (t = 0.66, df = 25.49, p >.250)).

In a related analysis, BHS find that, across 31 Senate elections, the arithmetic mean correlation between ESVI and the state-specific VII is positive for elections with Republican-leaning PVI scores (mean r = 0.55, N = 19, p < 0.001, 95% CI [0.30, 0.81]) but not for those with Democrat-leaning PVI scores (mean r = -0.12, N = 12, p >.250, 95% CI [-0.58, 0.34]). This difference is statistically significant (d = 1.11, t(29) = 3.00, p = 0.005). With outliers Hawaii and Rhode Island included, they find a similar pattern (mean r = 0.55, N = 19, p < 0.001, 95% CI [0.30, 0.81]) for positive PVI states and (mean r = -0.22, N = 14, p >.250, 95% CI [-0.63, 0.19]) for negative PVI states. When considering only those 27 states that had at least 3 time points for comparison with a later detrended analysis, the arithmetic mean correlation between ESVI and the state-specific VII is positive, and statistically significant for Republican-leaning states (mean r = 0.43, N = 13, p = .020, 95% CI [0.08, 0.78]) but not for Democrat-leaning states (mean r = - 0.12, N = 12, p > .250, 95% CI [-0.58, 0.34]). This difference is statistically significant (Welch Two Sample t-test: (t = 2.11, df = 21.17, p = 0.047). With outliers Hawaii and Rhode Island included, the pattern is similar for Republican-leaning (mean r= 0.43, N = 13, p = 0.020, 95% CI [0.08, 0.78])) and Democrat-leaning (mean r = - 0.22, N = 14, p >.250, 95% CI [-0.63, 0.19) states.

In contrast, when controlling for first-order autocorrelation, we find that the moderating effect of state PVI on the correlation between Ebola searches and people's voting intentions disappears. The arithmetic mean correlation between "ESVI Changes" and "VII Changes" across 25 Senate elections for is small and non-significant for both Republican-leaning states (mean r = 0.13, N = 13, p >.250, 95% CI [-0.19, 0.45]) and Democrat-leaning states (mean r = - 0.05, N = 12, p >.250, 95% CI [-0.34, 0.24]). This pattern holds when outliers Hawaii and Rhode Island are included (Republican-leaning states: mean r = 0.13, N = 13, p >.250, 95% CI [-0.19, 0.45]; Democrat-leaning states: mean r = - 0.05, N = 14, p >.250, 95% CI [-0.29, 0.19]). Neither of these are significantly different from zero. Further, they are not significantly different from each other (Welch Two Sample t-test, outliers excluded: t = 0.92, df = 22.94, p > .250; Welch Two Sample t-test, outliers included: (t = 0.97, df = 23.07, p >.250).

DETAILED DISCUSSION

In their paper, BHS use time-series data to test the hypothesis that the 2014 Ebola outbreak influenced the 2014 U.S. Federal elections. They find that the volume of Google searches for "Ebola" strongly covaries with support for Republican candidates, at both the national and state level, and that the correlation between support for Republican candidates and Ebola search volume is strongest in states with greater support of Republican candidates and with longstanding Republican voting norms. However, BHS assumed that time series observations were independent from each other, whereas we show that all analyzed variables exhibited extremely high degrees of temporal autocorrelation (See Table S1). Here we reanalyze the BHS data, controlling for first-order autocorrelation, and find that essentially all relationships between the Ebola outbreak and people's voting intentions become attenuated and non-significant (See Table 1). Because the BHS findings are not robust to such a basic controls, this strongly suggests that either: (1) many of the initial findings were either spurious or (2) the study design used by BHS is insufficiently powered to detect any associations that may actually exist.

In study 1, we find that the positive correlation between ESVI and VII in September and October drops substantially and becomes non-significant after removing first-order autocorrelation. If we focus on only the two-week period that includes the last week of September and first week of October, and instead analyze the relationship between raw Google Trends "Ebola" search volumes and VII, we find that the positive correlation between these two variables disappears entirely after removing first-order autocorrelation.

In study 2, we find that the positive correlations between national ESVI and state-Senate election VII across all available states becomes indistinguishable from zero after removing those states that don't have sufficient data for detrending, and drops to values very close to zero when removing first-order autocorrelation. Further, we also find that the moderating effects of candidate leading the polls at time of the outbreak and state-level PVI disappear after removing first-order autocorrelation. These results are robust to the composition of the sample (including/excluding outliers; excluding/including six states with insufficient data on daily changes). The state-Senate election data is the most fine-grained of all the BHS data and contains the largest number of data points. As such, it affords the best test of the hypothesis that the psychological salience of Ebola affected people's voting intentions in during the 2014 U.S. elections. Our analyses reveal no evidence for this purported relationship: there is no evidence from this data that increases in national-level Google searches for "Ebola" are associated with people's tendency to favor Republican or Democratic candidates in state-Senate elections. Further, there is no evidence for an increased inclination to conform to popular opinion with increases in national-level Google searches for "Ebola".

Caveats. Here, we focused our analyses on the more fine-grained time-series data analyzed by BHS, as it provides the strongest test of their hypotheses. The fact that support for Republican candidates was higher after announcement of the Ebola epidemic than before the announcement is obviously suggestive, but is not particularly informative for the proposed hypotheses given that it represents 2 observations (before and after) from a single data point—National-level means. BHS also constructed a *voter-intention-change index*, which assessed the difference between that day's VII and the VII 7 days before, resulting in 16 data points. This was a useful start toward a fully detrended analysis, but it only detrended one of the variables, and also still permitted overlap in the observations being aggregated (creating more temporal autocorrelation). Finally, the BHS analysis of how state-level pre-post changes differ across states that are Republican- or Democrat-leaning is suggestive. However, it also raises interesting questions about spatial and cultural autocorrelation which are not the direct topic of this paper, but deserve additional exploration (Hruschka & Henrich, 2013).

Our re-analysis of the BHS data has a number of limitations. Because of the small sample size in the original data, the analyses were insufficiently powered to detect potential associations. Of course, this is a concern with any analyses of the existing BHS data, suggesting the current data may not be sufficiently powered to test these hypotheses. Indeed, some guidelines suggest a minimum of 50 data points for time series analyses (Jebb, Tay, Wang, & Huang, 2015). After controlling for first-order autocorrelation, none of the purported associations between Google searches for "Ebola" and people's voter intentions remained significant, though it is possible that such associations can only be detected in a longer time-series. Thus, while the current analyses have a number of limitations, these are the same caveats that apply to the original BHS analyses.

References

Hruschka, D. J., & Henrich, J. (2013). Institutions, parasites and the persistence of in-group preferences. *PloS One*, *8*(5), e63642.

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: examining and forecasting change. *Frontiers in Psychology*, *6*. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4460302/

Tables and Figures

Table S1: 1st order autocorrelation of key variables.

	BHS Data	**1st-order Autocorrelation Removed**
Raw Ebola Search Volumes (national)	0.90	-0.37

ESVI (national)	0.99	0.60
VII (national)	0.95	0.43
VII (state-level average)	0.75	0.03

Table S1 demonstrates that all variables in the original BHS data have extremely high levels of first-order autocorrelation and that calculating the changes between data values strongly attenuates levels of autocorrelation.

Figure S1 – Ebola's Psychological Salience across Time, Operationalized via Google Search Volumes for the term "Ebola"

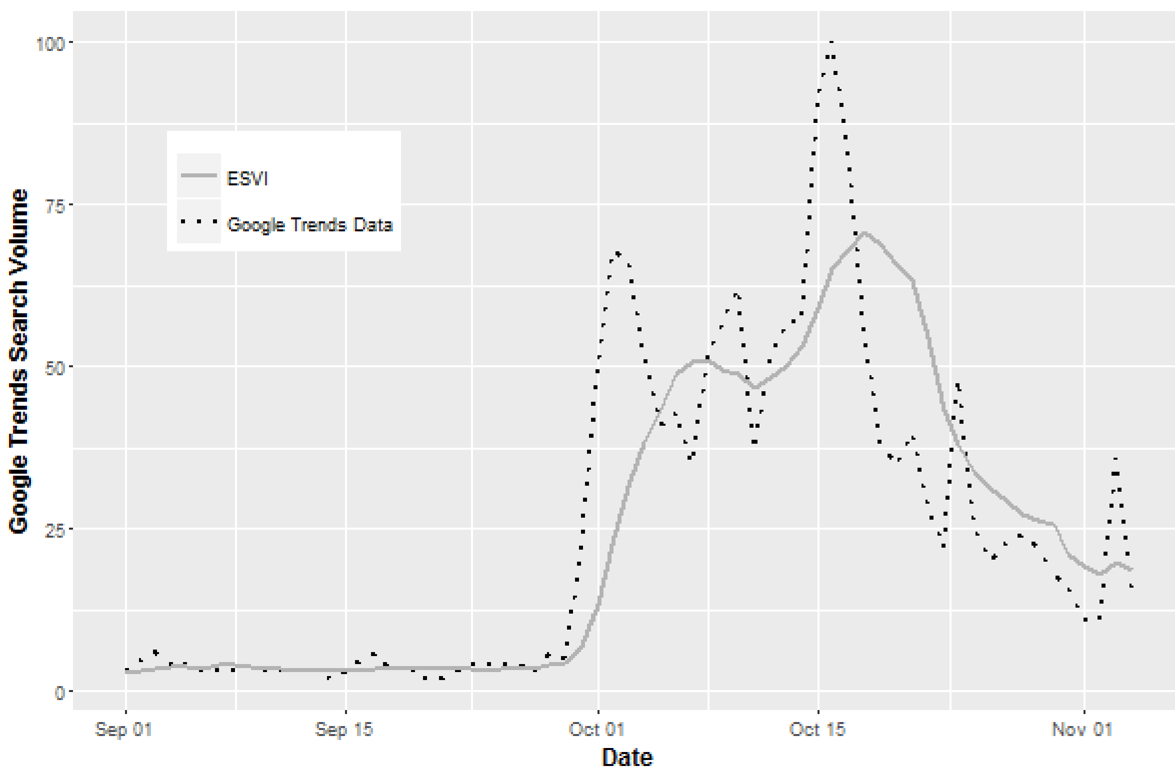

Figure 1 shows that the Ebola Search Volume Index (ESVI) is insensitive to daily fluctuations in raw Google search volumes for "Ebola"

Figure S2 – Nationwide Voting Intentions across Time, Operationalized via the Voter Intention Index (VII)

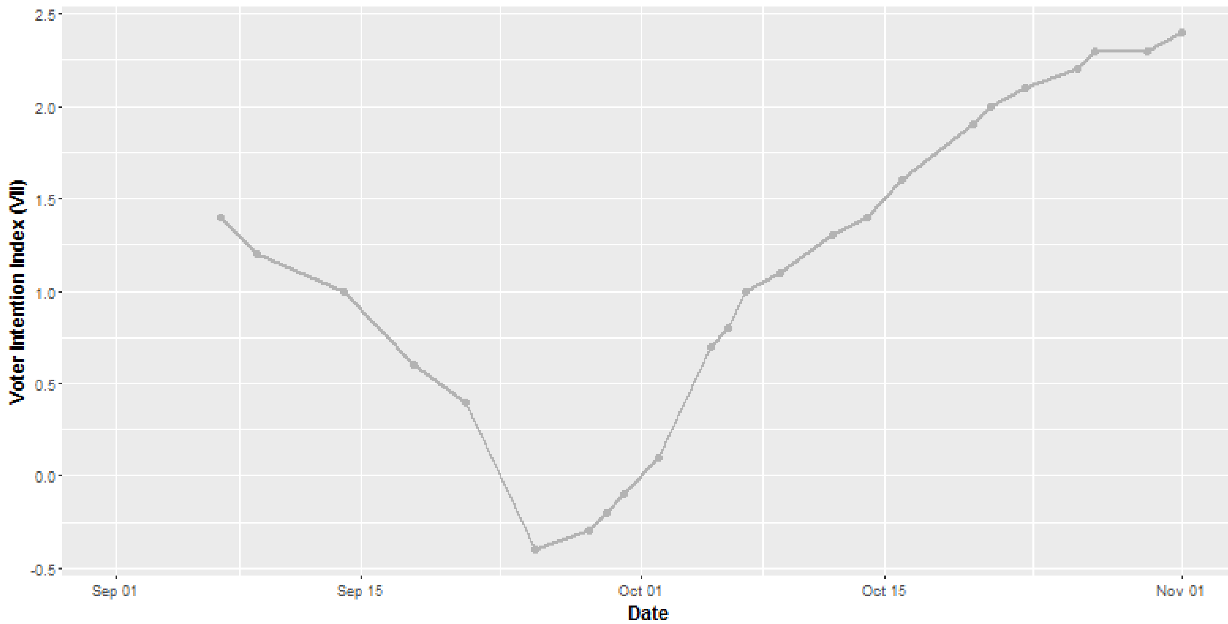

Figure 2 shows that the Voter Intention Index (VII) steadily decreases throughout most of September, and steadily increases from late September to November 1.

Figure S3 – Distribution of Correlations between Ebola Search Volume Index (ESVI) and Voter Intention Index (VII) for 1) A 26-election subset of the BHS data and 2) The same 26-election subset of the BHS data after removing 1st-order autocorrelation

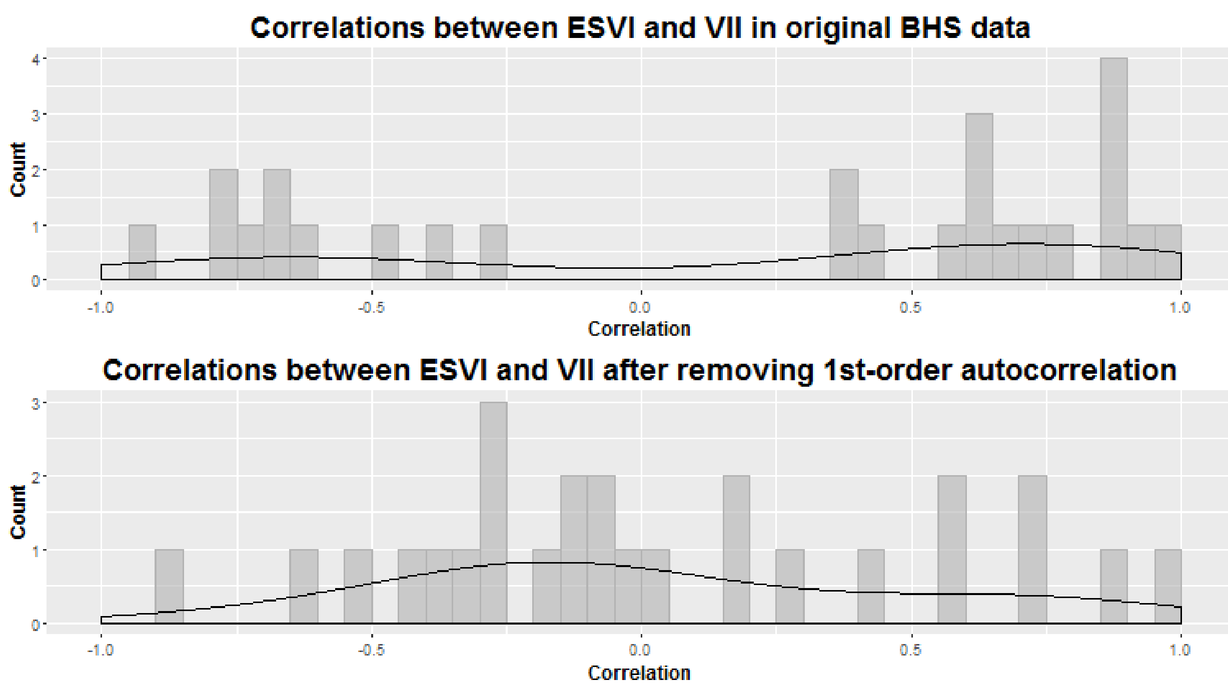

Figure S3 demonstrates that correlations in a 26-election subset of the BHS data take on extreme values and are bimodally distributed. Removing 1st-order autocorrelation results in a distribution of correlations more closely resembling a normal distribution.

Figure S4 – Distribution of Correlations between Ebola Search Volume Index (ESVI) and Voter Intention Index (VII) for 1) The original 32-election BHS data and 2) A 26-election subset of the BHS data after removing 1st-order autocorrelation

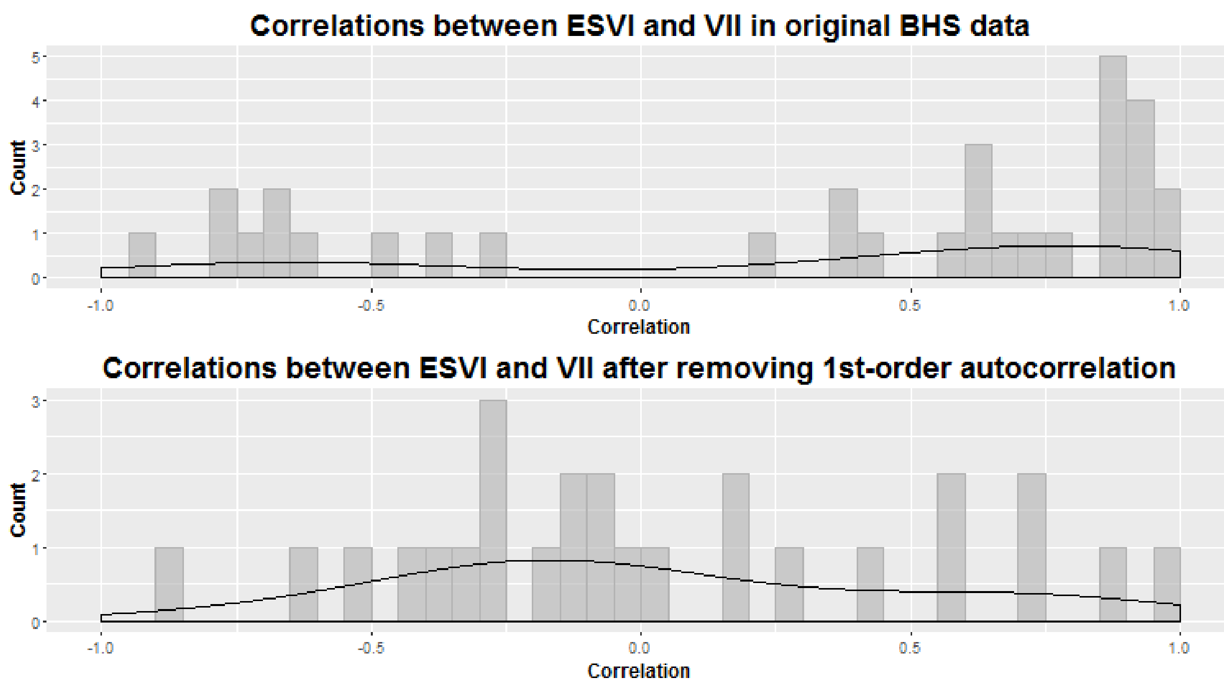

Figure S4 demonstrates that correlations in the original 32-election BHS data take on extreme values and are bimodally distributed. Removing 1st-order autocorrelation results in a distribution of correlations more closely resembling a normal distribution.