

Internet Search Data and Issue Saliience: the properties of Google Insights for Search as a measure of issue saliience

Jonathan Mellon¹ University of Oxford

Abstract: This article examines the conditions required for using Internet search data as measures of aggregate issue saliience. Internet data have clear advantages over survey data in terms of cost, availability and frequency. These advantages have led the media and some researchers to use Internet search data as proxies for public opinion. However, these analyses do not present systematic evidence that search data tell us about the general public's views rather than those of an unrepresentative subset. This paper outlines a general method for assessing the validity of search data against existing measures, including content validity and criterion validity. To this end, weekly Google search data are tested against Gallup's 'most important problem' question. The paper finds the saliience of four issues, fuel prices, the economy, immigration and terrorism, can be measured using search data. Weekly measures of issue saliience are generated for these issues, from 2004-2010, for empirical analysis. The saliience of three other issues – the Iraq conflict, healthcare and dissatisfaction with government – could not be measured using Internet searches.

¹ Jonathan Mellon is a DPhil Student at Nuffield College and the Department of Sociology at the University of Oxford. All correspondence to jonathan.mellon@nuffield.ox.ac.uk. This work was supported by the Economic and Social Research Council [grant number ES/H011765/1].

Data based on the actions of individuals on the Internet have become increasingly available to researchers. Aggregate trends for search terms have been used by researchers (Granka 2010; Graefe and Armstrong 2012; Brunn, Devriendt, and Boulton 2010; Reilly, Richey, and Taylor 2012; Pelc 2011; Rickard; Ellis, Swearingen, and Ripberger 2011; Oltra 2011; Reis and Brownstein 2010; Weeks and Southwell 2010; Dube and Kaplan 2012; Kahn and Kotchen 2010) and the media (for instance Weiner 2012; Silver 2012) to make a variety of claims about public opinion. However, these claims are based on the assumption that the trends in search data track those in the wider population. This article aims to provide a method by which researchers can test whether the Internet search data they use are valid proxies for the phenomena they are interested in by comparing the search data with existing survey data.

Issue salience has been understood in a number of ways but is most commonly seen as the importance placed on an issue by an individual in their decision making on political matters (Wlezien 2005). Measures of issue salience are used widely in political science, particularly in agenda-setting research (Soroka 2002; Wlezien 2005)²; however, these studies have often been restricted by the limitations of available data. Data for the commonly used “most important problem” question (MIP) are only available monthly, at irregular intervals and usually only at the national level in the US. In most other countries the situation is worse, with only sporadic survey data obtainable. Using Internet search indices to measure issue salience could potentially solve these problems, as weekly search data are available from Google between 2004 and 2010 for almost every country, state and city in the world – and for any possible search term. For instance, one can measure the change in the weekly frequency of searches for “illegal immigration” over time in the US and across any period.

To assess when Internet data can be used, a process is outlined that tests the reliability and validity of a candidate search index as measures of a particular issue’s salience. Content validity is tested by looking at whether the Internet searches in each index actually refer to what they are assumed to measure, i.e. the jobs search index is measuring the salience of employment issues and not Steve Jobs (the former CEO of

² For an overview of the agenda-setting literature see McCombs, 2005

Apple Inc). Criterion validity is assessed by quantitatively comparing search indices with the existing MIP measures of issue salience. Although this paper focuses on issue salience, the approach could be equally applicable to aggregate measurements of other attitudes such as approval ratings and other Internet data where representativeness is in question, such as Twitter mentions.

THEORY

This section discusses how to conceptualise issue salience. Firstly, it has been seen as the importance an individual attaches to an issue, i.e. the weight placed on an issue when making political decisions such as voting (Wlezien 2005). Secondly, it has been viewed as how prominent an issue is in the mind of an individual, that is, how accessible the topic is when the individual is prompted to make a political decision (Wlezien 2005). Both conceptualisations fit well with search data, in that perceived issue importance is likely to have a strong effect on decisions about information seeking, and an issue must reach a certain level of prominence in a person's mind before they will search for information about it. So far, MIP has been used widely as a measure of issue salience and its properties are well understood (Soroka 2002). Therefore, to assess the validity of an Internet search index with regards to issue salience, it makes sense to corroborate it against the better understood MIP data.

PREVIOUS WORK

Despite being a relatively new data source, many political scientists have already begun to make use of Google Insights for Search as a measure of issue salience or attention (Granka 2010; Graefe and Armstrong 2012; Brunn, Devriendt, and Boulton 2010; Reilly, Richey, and Taylor 2012; Pelc 2011; Rickard; Ellis, Swearingen, and Ripberger 2011; Oltra 2011; Reis and Brownstein 2010; Weeks and Southwell 2010; Dube and Kaplan 2012; Kahn and Kotchen 2010). A selection of these studies are referenced in Table 5 in Appendix 1 along with the search terms they use in the study.³ These have been

³ There are many more studies that make brief use of search indices to illustrate increased public interest in the author's topic and a large number of working papers in progress.

used to address problems from constituent reactions to WTO trade disputes (Pelc 2011) to the context of the Occupy movement (Dube and Kaplan 2012).

However, these studies generally do not address the question of whether the search indices are representative of the public agenda. For instance, in agenda setting studies it is plausible that a small subset of the population who search using a certain term may be more or less responsive to media coverage than the public as a whole. So, findings of agenda setting are entirely consistent with the alternative explanation that a small politically interested minority follow each development of a political story obsessively and react strongly to media coverage, while the vast majority of the population pay no attention, or hear about the story through different mechanisms. A good proxy variable must be shown to closely track the theoretical variable of interest (in this case aggregate issue salience in the public).

A few papers try to address this problem. Ripberger claims to validate Internet search indices on the basis that some closely match the coverage of issues in the *New York Times* (Ripberger 2011). However, correlations with media coverage are insufficient to validate a search index as a measure of an issue's salience to the public: the *New York Times* is not a representative measure of issue salience in the general population (although correlations with media coverage can provide an additional line of evidence through convergent validity).

Two other articles looked at search data's relationship to survey measures of issue salience. A recent research note (Scheitle 2011) briefly presented correlations between MIP and Internet search data to illustrate a call for more research using search data but without examining any methodological concerns regarding time series or content validity. Scharnow and Vogelgesang attempted to validate search engine queries as a measure of issue salience (Scharnow and Vogelgesang 2011). Survey responses relating to a controversial 2005 German politician, Paul Kirchof, were compared to the search volumes of "Paul Kirchof". The authors found a weak relationship between the two series ($R^2 = 0.24$), concluding that "search engine queries are a somewhat different but still related measure of the public agenda". However,

the evidence for the relationship between the series is weak: due to the small sample size (500 respondents per time point, with a maximum of 3% mentioning Paul Kirchof), variations observed in the time series cannot be distinguished from sampling error.⁴

The above discussion shows that there are problems with drawing conclusions from previous work using Internet search data. This paper seeks to provide a method for assessing when a search index can be used as a proxy for public opinion and when this claim is unjustified. It is not necessary that the people searching for a term are perfectly representative of the population, but the dynamics of their interest needs to closely track that of the whole population.

DATA

This section outlines the properties of Gallup MIP data and Google Internet search data. It then discusses data issues that arise when comparing the two. The data are for the US between January 2004 and March 2010. The Google data are aggregated weekly across the time period whilst the Gallup data are available once a month at irregular intervals. The US was chosen due to having the highest number of Google searches in the world and its centrality in the agenda-setting literature. This paper attempts to measure the following issues: macroeconomics, Iraq, healthcare, dissatisfaction with government, immigration, terrorism and fuel prices, which were the MIP issues with sufficient numbers of respondents choosing them over the time period to allow analysis.

Google Internet search data

The search indices were accessed through the Google Insights for Search website (Google 2004-2010). The Google data are not given in absolute volumes but are indexed to the highest observed search volume, which is set to 100. Consequently, it is not possible to ascertain the frequency of searches that took place at any given time, but only how the searches have changed over time. As this scale has no

⁴ An additional problem is that the authors use only 34 data points across the time period, and these do not form a continuous time series because no surveys were conducted on weekends. This means that they are unable to perform any tests for stationarity and autocorrelation. These tests are particularly necessary in light of the low fit between the series. In the presence of autocorrelation, both R^2 values and levels of significance are likely to be greatly inflated (Granger and Newbold 1974) and spurious relationships are likely to be found.

intrinsic meaning, one important aim of this paper is to convert these indices into units with intuitive meaning, namely the predicted percentage of respondents citing an issue as important.

Search indices are measures of the frequency of searches including a given word or set of words, referred to as “search terms”. If the term “jobs” is entered into Google Trends, the resulting index is the change in frequency of searches including the word “jobs” over time (including searches for the word “jobs” and ones for “government jobs”, “secretary jobs”, etc.) It is also possible to combine indices to include all searches containing either of two keywords. For instance, one index in this paper comprises all searches including either the word “terrorism” or the phrase “al Qaeda”.

Representativeness of Internet users

A potential problem with Internet data is that users are not fully representative of the population as a whole demographically and, as such, some groups are underrepresented in the population of potential searchers. Analysis of a 2010 Pew Internet use poll (Pew 2010) shows that younger age, higher education and higher income predict increased probability of using the Internet. To test whether these differences could bias the Google search data, the individual level Gallup MIP data was reweighted to reflect the demographics of Internet users.⁵ The MIP time series resulting from the reweighted data was correlated against the original MIP series for each issue. The correlations between the series are extremely high ($R^2 > 0.9$ in all but one case), indicating that the demographic differences between Internet users and non-Internet users are unlikely to introduce significant biases.⁶ It is possible that Internet users differ in other ways that are not weighted for, but it is unlikely that these effects would be larger than the effect of demographic characteristics, so they should not seriously bias the results. Even given this analysis, the people searching with some search terms are likely to be much less typical of the public than Internet users in general. Hence, it is necessary to demonstrate the correspondence between each search index and data that is representative of issue salience in the wider population.

⁵ Weights were created from predicted probabilities of Internet usage derived from the Pew survey using age group, sex, education and income level as independent variables. These weights were then combined with the Gallup survey weights to derive the time series for each issue. Only 30 of the 75 Gallup surveys across the time period in this study had individual level data available through iPoll.

⁶The one exception to this were MIP responses for healthcare issues: the two series only correlated with an R^2 of 0.79.

Gallup MIP data

MIP has been used to measure issue salience in more than 100 papers on the agenda-setting effect of the media (Soroka 2002). It has also been used to examine the interplay between public policy and issue salience (Jennings 2009; Jones, Larsen-Price, and Wilkerson 2009).⁷ For this paper, summaries of Gallup's MIP responses were collated between 2004 and 2010 for all seventy-five months (Gallup 2010). The open-ended survey responses were coded in line with the Public Policy Agendas (PPA) schema (Baumgartner and Jones 2006), with the exception of terrorism, which is included as a single category. The PPA category of macroeconomics includes the economy in general, unemployment, taxes, budget deficits and inflation.

This resulted in seven substantive MIP series, as only issues mentioned by at least 10% of respondents in at least one survey were included. As such, this paper attempts to find Internet search measures for the following issues: macroeconomics, Iraq, healthcare, dissatisfaction with government, immigration, terrorism and fuel prices.

Differences between MIP and search data

For the validation of Google searches against Gallup's MIP, a week of Google searches is chosen that corresponds most closely to each Gallup poll's timing. If the days on which Gallup interviewed respondents do not fit into a single week of Google searches, the week with the most interview days is selected as corresponding to that Gallup survey.

Another difference between Gallup's MIP and Google's search indices is that Gallup respondents are at least partially restricted in the number of responses they can give to the MIP question. Consequently, the responses in each category are correlated negatively with responses in other categories (Wlezien

⁷ MIP (and "most important issue") measures are also used widely in voting behaviour studies (Bélanger and Meguid 2008; Green and Hobolt 2008) but these normally use individual level analysis which is less relevant for comparing to Internet search data.

2005). By contrast, Google searches are not structurally constrained by searches for other issues,⁸ which may lead to lower correlations between search indices and MIP series. It is preferable to have unconstrained salience measures for testing theories that refer to the level of importance the public gives to an issue, so this paper will not adjust the models used to create constrained series.

METHODOLOGY

This article has two main aims: 1) to assess whether any Internet search indices can function as measures of issue salience and 2) to generate a set of issue salience measures based upon these indices. To fulfil these aims it is important that the measures satisfy two of the key requirements for measurement:

reliability and validity. The validity of each search index is assessed in three steps:

1. **Face validity** – Intuitive plausibility of the search index.
2. **Content validity** – Intuitive plausibility of the search terms within the search index. That is, whether the search terms used are relevant to the issue being measured.
3. **Criterion validity** – Analysis of the association between search indices and MIP series after taking into account time series properties such as trending, which can lead to spurious regressions. Search indices showing seasonality have the seasonal component removed prior to comparison.

At each of these steps, the search index is excluded if it lacks validity at that stage, reducing the number of search indices throughout the process. For the remaining search indices an OLS regression model is fitted to the MIP and search index for an issue. This model is used to predict MIP values for every week across the period, for valid search indices. In addition to these validity checks, the reliability of the generated issue salience measures is estimated through the use of out-of-sample prediction.

4. **Out-of-sample testing** – Models usually fit better to the data used to construct them than to unobserved data. In order to estimate the reliability of the weekly issue salience measures, an

⁸ Empirically, concern over one issue may genuinely decrease concern about others due to cognitive capacity and prioritisation in the public. However, the relationship is likely to be more complex than simply dividing a permanently fixed quantity of importance between issues.

OLS model is fitted between the search indices and MIP series for half the data. The model is then used to predict the MIP values for the other half of the data and the success of the model is assessed.

Face Validity

In this context, face validity means that the word or phrase used in the search index sounds like something people are likely to search with, if they are concerned about the issue. For instance, the words “immigration” and “immigrant” have face validity for measuring concern about immigration. By contrast, the word “tree” would not have face validity for measuring concern over immigration because it is not plausible that people concerned about immigration would use it to search.

This paper begins by selecting a number of indices with face validity for each potential issue. This list is not intended to be exhaustive but just a reasonable selection of terms with face validity, which contrasts with the approach of some previous papers that looked at all possible search indices (around 50 million), selecting only those that fitted statistically with the dependent variable of ‘flu outbreaks (Ginsberg et al. 2009). That approach risks finding significant relationships simply by chance due to the high number of tests being performed (Ioannidis 2005).

Content Validity

Carmines and Zaller define content validity as the extent to which “an empirical measurement reflects a specific domain of content” (Carmines and Zeller 1979). In this study content validity is the extent to which a search index contains search terms that reflect the issue, in other words searches that someone who is concerned about a given issue would make.

Content validity is assessed by examining the top fifty search terms in each index for each year, which allows for the identification of terms not plausibly related to an issue’s salience. For instance, the search index of Iraq might include searches for “Iraq soccer”, which is not in the domain of searches made by someone concerned about the issue of the Iraq conflict. These confounding searches can then be removed from the search index by adding “-[confounding term]” to the query on the Google Insights website. In

the example, the query would read “Iraq –soccer”. However, no more than five terms can be removed, so an index would be rejected if there were more confounding terms than this example.

The need for these content validity checks is demonstrated by a previous use of Google data. A depression charity reported “a 488% growth in searches for [suicidal thoughts] between January 2004 and January 2010” (Schimelpfening and Nancy 2010), although a less alarming interpretation becomes apparent by looking at the top terms associated with the search index, as a song titled “Suicidal Thoughts” was re-released around the time of the huge increase in search volume (IMDB 2009).

Carmines and Zeller argue that “it is not possible to determine the specific extent to which an empirical measure should be considered content valid”, so it should not be used as confirmation of the validity of a measure (Carmines and Zeller 1979). Whilst a measure cannot be proved to be content valid, it is often easier to observe when it is invalid. Furthermore, it is impossible to confirm that the full universe of relevant terms is included in a search index, but it is relatively straightforward to observe the presence of an irrelevant term in the index – as in the “Iraq soccer” example.

Criterion Validity

Criterion validity describes how closely a measure corresponds to an existing measure of the concept under study (Carmines and Zeller 1979). Since MIP is well established as a measure of issue salience, it makes sense to 1) use it as a criterion to compare new search index measures to and 2) distinguish between those indices that can measure issue salience and those that cannot. Criterion validity is assessed by an OLS regression using MIP as a dependent and the relevant search index as the independent variable, with Newey-West standard errors that account for autocorrelation and heteroskedasticity (Wooldridge 2009; Newey and West 1987). Criterion validity is best regarded as a continuous property rather than a binary one (Carmines and Zeller 1979), so the weight that can be placed on the results of criterion validity tests will depend on the strength of association found between the MIP series and Google searches (i.e. R^2 values showing the percentage of variance explained).

Time series complications

Before the regressions can be performed, it is necessary to take account of several time series complications. Determining criterion validity is considerably more complex when using time series data, as seasonal patterns and non-stationarity can generate misleading results. This subsection discusses how to deal with these concerns.

Many of the search indices include a substantial seasonal component not present in the MIP series. It is not possible to definitively say why MIP series show less seasonality than the Internet searches. However, one possibility is that respondents to the Gallup surveys discount seasonal trends in the obtrusiveness of a problem when responding to pollsters; for instance, Internet searches for “jobs” dip in December before rebounding in January. This wouldn’t reflect a stronger economy every December but simply reduced activity in the labour market over the holiday period. As people are aware of this fact, they will discount this reduction when reporting their concerns about the economy to the pollster. Consequently, it is necessary to adjust the search index to remove the seasonal component.

To decide which MIP and search index time series to adjust for seasonality, the X-12 seasonality diagnostic tests were run on each series, as suggested in the US census guidelines for seasonal adjustment (McDonald-Johnson et al. 2010). Seasonal adjustment was performed using Seasonal Trend Decomposition by LOESS (STL) (Cleveland et al. 1990)⁹, which separates a time series into trend, seasonal and remainder components (Wessa 2010; Cleveland et al. 1990), with 52-week periodicity. To explain how this works, an example is shown for the “jobs” search index in Figure 1. The initial series (labelled data) clearly displays a yearly pattern with dips in searches for jobs around December, which is confirmed by the large seasonal component (seasonal) identified. Once the seasonal component is removed, a smooth trend component and the residual variance (remainder) are returned and recombined to create a seasonally adjusted series.

[Figure 1 about here]

Non-stationary time series are prone to spurious correlation (Granger and Newbold 1974), which will compromise criterion validity tests. The MIP series and Google Trends are first tested for stationarity

⁹ For a detailed explanation of the STL procedure see Cleveland et al. 1990

using Augmented Dickey-Fuller tests attempting to reject the null hypothesis of non-stationarity. Where two series are different orders of integration¹⁰ it is not possible to estimate accurately the relationship between them.

If the search index is I(1) and the MIP series I(0), the search index is first differenced to obtain a stationary series which is used in the regression. However, when the MIP series is I(1) and the search index is I(0), the search index is rejected because taking the first difference would mean predicting the change rather than the level of the issue salience, which is not substantively interesting (this paper is concerned with generating time series of *levels* of issue salience). If both the search index and MIP series are I(1) and cointegrated, then OLS estimates of coefficients are consistent and unbiased (Wooldridge 2009). The series are cointegrated if their residuals form a stationary sequence. To test for cointegration, OLS is run and its residuals tested for cointegration using adjusted critical values (Davidson and MacKinnon 1993) to assess whether the null hypothesis of non-stationarity can be rejected for the residuals.

The full process for selecting and modifying the search indices to be used for generating weekly issue salience series is outlined in the flowchart in Figure 2. The process shows the steps in testing for validity and any adjustments to the data made to correct for problems.

[Figure 2 about here]

Prediction of weekly issue salience series

The second aim of this article is to generate weekly issue salience measures for use in empirical investigation, for those search indices that are accepted after following the process outlined in the flowchart in Figure 2. The monthly MIP series for each issue is regressed on the relevant search index (that has been adjusted for seasonality, if appropriate).¹¹ This model is then used to predict MIP responses for each week between January 2004 and March 2010 for that issue, which produces a weekly time series

¹⁰Order of integration is expressed as I(x) where x is the number of first differences required to obtain a stationary sequence. For a wider discussion of stationarity and orders of integration see Wooldridge 2009.

¹¹ Time trends are included as control variables if either the MIP series or the search index is trend stationary.

measured on the same scale as MIP responses, allowing comparisons to assess which issues are most salient at different times.

Out-of-sample testing

The utility of the time series generated in the previous section is dependent upon how well they extrapolate to time points without MIP data. Two different kinds of extrapolation would be valuable when using the instruments generated from search data – firstly, extrapolation to more detailed time points within the time period under study (weekly data) and secondly, it would be useful to predict issue salience values outside of this period in future, without having to fit new models between MIP and search indices. The existing MIP data are split into training and test sets, with a model fitted to the training set, which is then used to predict the test set. Root mean square errors are calculated for the training and test samples to assess the degree to which the error increases when the model is fitted to new data. Apart from the reduced data to which they are fitted, the models are specified in the same way as when predicting weekly values in the previous section.

Two separate versions of these tests are performed to reflect the two kinds of extrapolation performed in this paper. As no weekly MIP data are available, the issue of extrapolating to finer time intervals is addressed by arbitrarily defining the training sample as all even years and the test sample as odd years. This split was chosen in order to ensure the independence of the training and test samples, which could be undermined by autocorrelation if alternate months or a random sample were used. This means the test measures the accuracy of predicted values within the time period. Since the model is fitted to less data than the actual model used to generate the weekly series, this test should produce an estimate of the lower bound of weekly issue salience measure reliability.

To address the issue of extrapolation outside of the time period, the training sample is defined as the years 2004-2006 and the test sample as 2007-2010. This test looks at whether the relationship changes systematically over time. This second out-of-sample test is not performed on time series which are trend

stationary, as they need to include a linear time trend in the regression to be calculated accurately and this would potentially lead to unrealistic predictions, such as negative salience.

RESULTS

This section details the outcomes of each section of the methodology outlined in this article when applied to Google search data and Gallup's MIP series in the US between 2004 and 2010. A search index is only included in each subsequent stage if it passed the requirements of the previous one.

Face validity

Search indices with face validity were chosen for each of the seven issues that could be tracked across the complete period of interest, using MIP: macroeconomics, Iraq, healthcare, dissatisfaction with government, immigration, terrorism and fuel prices. Some potential search indices such as "health reform", used for concern about healthcare, could not be selected because they lacked sufficient search volume to track weekly frequencies across the full period. As 'macroeconomics' is a composite issue of several MIP responses, search indices such as "unemployment", "fiscal" and "taxes" were chosen because they corresponded to the components. In total, twenty search indices with face validity were selected to measure the seven issues that are important across this time period, as shown in Table 1.

Content validity

Each of the twenty search indices was examined for threats to content validity. A good example of the content validity methodology is the jobs search index employed to measure the salience of 'macroeconomics'. The bulk of the searches comprised various types of jobs ("catering jobs", "government jobs", etc.), although two invalid terms were found in the fifty top searches – "blow jobs" and "Steve Jobs" (the former CEO of Apple Inc.) – which were then excluded from the search index. Table 1 summarizes the results of the content validity checks. For each search index (shown in the second column) the table indicates the relevant issue it is intended to measure (first column), which terms within the index threatened content validity (third column) and finally whether the index was deemed to be

content valid (fourth column). Of twenty search indices, six were rejected for lacking content validity and three initially included problematic search terms that were able to be removed. After the content validity checks, fourteen search indices covering all seven issues remained for further testing.

[Table 1 about here]

Time series complications

Before assessing the criterion validity of the remaining search indices, it is necessary to account for the time series properties of the data, namely seasonality and non-stationarity. Eight of the fourteen remaining search indices showed evidence of seasonality and were adjusted,¹² but none of the MIP series exhibited seasonality.

Non-stationary time series can be spuriously correlated, which has to be accounted for before assessing criterion validity. Five indices could not be used in OLS regressions because of differing orders of integration,¹³ while another four were tested for cointegration with the relevant MIP series, although only one of these, “jobs”, was cointegrated, so the other four search indices were rejected.

Criterion validity

The remaining six search indices were tested for criterion validity by regressing each MIP series on the relevant search index using OLS with Newey-West standard errors. The results are shown in Table 2, with coefficients and standard errors shown for each search index predicting the corresponding MIP series. The “terrorism” regressions include time terms because the “terrorism” MIP series is trend stationary. Five search indices were found to predict the relevant MIP series significantly, indicating high levels of association, but searches for “fuel” were selected over those for “gas” as a measure of fuel price

¹²The full results of the seasonality and cointegration tests are available upon request.

¹³Some readers may consider the stationarity tests unduly demanding. Some valid search indices were likely excluded, as augmented Dickey-Fuller tests have low statistical power (Hansen 1995), so often fail to reject non-stationarity even when the series is actually stationary. This paper is intended to provide a set of valid issue salience time series that can be used in empirical research with a high level of confidence. However, if a lower confidence level is acceptable, these criteria can be relaxed to generate more series for empirical use by choosing a lower significance cut-off for the tests

concern because the association was substantially higher at an R^2 of 0.72 compared to 0.49.¹⁴ The “terrorism” search index was not significantly related to the MIP series, but the search index combining searches for “terrorism + al Qaeda” showed a strong fit, which left four issues with search indices that were significantly and substantively related to the relevant MIP series. The R^2 values ranged from 0.67 to 0.75, demonstrating a very strong association with the MIP series. These covered four issues: macroeconomics, immigration, terrorism and fuel prices.

[Table 2 about here]

The four MIP series and corresponding search indices that survived the validity tests are shown in Figure 3 for the weeks where a Gallup survey was available. The left axis shows the proportion of respondents mentioning the particular issue as the MIP that month and the right axis corresponds to the arbitrary scale for that search index. These graphs confirm that the Internet search indices are visually similar to the equivalent MIP time series, with similar trends and peaks visible in both time series.

[Figure 3 about here]

Predicting weekly issue salience series

To create weekly series, OLS models were fitted between the monthly MIP series and the equivalent search index week, and then used to predict MIP scores for every week across the time period. The final four predicted issue salience series for immigration, macroeconomics, fuel prices and terrorism are shown in Figure 4.

These time series allow a comparison of each issue’s salience over time and against other issues. The chart shows terrorism’s salience gradually declined over time, whilst macroeconomic problems increased in salience as the economy worsened. By contrast, the salience of immigration and fuel prices is characterised by short periods of greatly increased salience (each issue briefly overtakes macroeconomics in importance), but generally they elicit only low levels of importance.

¹⁴ A separate model was fitted with both “gas” and “fuel” search indices predicting the fuel MIP series. However the “fuel” search index was no longer significant in this model and the overall fit only slightly increased (adjusted R^2 from 0.711 to 0.718).

[Figure 4 about here]

A comparison of the weekly predictions and monthly survey responses is shown for fuel prices in Figure 5. The predicted fuel series looks immediately suspect, with a large peak in the salience in September 2005, whilst the MIP series gives no indication of such a peak. However, an explanation becomes clear when examining gasoline prices around that time, as for one week the average gasoline price was greater than \$3 per gallon for the first time ever, but it fell sharply the week after (US Energy Information Administration 2010). Since Gallup did not survey in that week, it missed the spike in salience accompanying the fuel price rise.

[Figure 5 about here]

Out-of-sample reliability tests

The generated weekly series were checked for face validity, content validity and criterion validity, but it still needs to be determined how reliable they will be outside the data to which they were fitted. To achieve this end, an out-of-sample test is performed by regressing the search indices on monthly MIP data for even years and testing the resulting model in predicting monthly MIP data for odd years. The results for this out-of-sample test are shown in Table 3.

[Table 3 about here]

Overall, the “jobs”, “illegal immigration”, “terrorism + al Qaeda” and “fuel” search indices show good reliability in predicting MIP series within the 2004-2010 time period, suggesting the weekly measures of issue salience will be reliable within this period. The root mean square error (RMSE) of the models changed between the training sample and test samples. In the case of the “jobs” search index predicting the macroeconomics MIP series, the RMSE increased but is still reasonable given that 37% of respondents choose a macroeconomic problem on average across the time period. The “terrorism”, “illegal immigration” and “fuel” search indices’ RMSEs did not substantially change in the test sample, suggesting a high degree of reliability outside the data the model was fitted to.

The second out-of-sample test fits a model to the years 2004-2006 and tests the predictions against data from 2007-2010 to assess how well the models are likely to perform outside the time period to which they were fitted. The terrorism weekly series is not tested in this section, as it was trend stationary. The results for these reliability tests are shown in Table 4.

[Table 4 about here]

This second set of out-of-sample tests broadly agrees with the first. The “jobs” search index RMSE greatly increases out-of-sample, which suggests that it will not be highly reliable when extrapolated beyond the period to which the model was fitted. The “illegal immigration” and “fuel” search index RMSEs increase moderately but are still broadly reliable. These results suggest that the “illegal immigration” and “fuel” search indices would be reliable measures in the future, even without fitting new models to more MIP data, while the “jobs” search index would be unreliable if extrapolated beyond the period to which it was fitted.

The number of search indices narrowed throughout this paper as follows:

1. Face validity – the method was started with twenty search indices with face validity.
2. Content validity – six indices were removed for lacking content validity, leaving fourteen indices.
3. Seasonality – eight out of fourteen indices showed seasonality and were adjusted accordingly.
4. Criterion validity – eight search indices were rejected because they could not be tested appropriately using time series methods. One further series was rejected because it did not demonstrate criterion validity when tested against MIP data. This left five indices covering four issues.

At both the content validity stage and criterion validity stage a high proportion of the search indices were removed. The reduction from twenty indices to five shows that simply assuming validity would have been misleading and have led to erroneous conclusions: if researchers use invalid measurements in research, they are not answering the questions they believe they are answering. Importantly, however, only one search index out of six was rejected at the criterion validity testing stage, suggesting that the

content validity analysis was successful in removing many time series that were not suitable measures of issue salience. The method also led to the successful measurement of four of the seven issues.

COMPARISON OF INTERNET SEARCH MEASURES WITH MIP MEASURES

The four newly generated issue salience measures are available for use in future empirical work,¹⁵ so it is useful to compare their properties to those of traditional MIP measures. The new procedures use MIP as the standard against which to measure, so it is not possible to test directly whether these Internet search measures can surpass MIP as gauges of issue salience. Consequently, the primary advantage of Internet search measures is their greater availability both in terms of frequency and location.

MIP measures only allow one problem to be reported by the respondent, so when one issue becomes important in the population it will reduce MIP responses for other issues, even if they are also important to respondents. This introduces an artifactual, negative correlation between different issues and thus can lead to problems with interpreting analyses that use MIP, as responses for one issue will be correlated negatively with any variable that is correlated with any other issue. Search indices do not constrain each other directly, so these artificial correlations are not introduced. The difference is apparent when examining the different issue salience time series, in that the MIP series for macroeconomics and immigration correlate strongly ($r = -0.62$), whereas the equivalent series generated from search indices correlate far less strongly ($r = -0.21$).

Another disadvantage of the existing MIP measures is that they violate the time series assumption of uniform intervals between measurements. The time between surveys ranges between three and five weeks, which means that time series models will not be specified correctly because they will not calculate the effects of lags consistently. However, MIP series are still subject to the problems experienced with time series (autocorrelation and non-stationarity), so models that ignore these aspects of the data will draw erroneous conclusions. By contrast, the search indices consist of evenly spaced weekly data, so they can be modelled accurately using time series techniques.

¹⁵ The four issue salience series are available at [URL anonymized].

The major disadvantage of Internet search measures of issue salience is that they will not necessarily capture the importance of all major issues in a time period. This paper was able to measure the salience of fuel prices, macroeconomics, terrorism and immigration but was unable to capture the importance of dissatisfaction with government, healthcare and the Iraq conflict. The salience of the healthcare issue could not be captured using search data for two reasons. The MIP responses for healthcare broadly follow political debates on healthcare reform, whereas searches for “healthcare” are mainly focused on searching for health insurance information and are stable over time. The search term “healthcare reform” was initially considered, as it would probably be driven by political concern over healthcare, but the search volume was too low to track it throughout the whole time period.

The issue of the Iraq conflict had several plausible search indices including “Iraq” and “Iraq war”, but the “Iraq” search index was problematic because a large number of searches were for graphic videos of an American contractor being beheaded in Iraq and pictures of the abuse from Abu Ghraib. These types of searches were therefore not plausible as a measure of concern over the Iraq war, especially as an extremely high proportion of the Iraq searches in the year where the beheading incident took place were for these videos. It is unlikely that most of the concern for the Iraq war was related to that one incident that year.

The third issue, dissatisfaction with government,¹⁶ was difficult to measure because it was not obvious which search terms reflect a generalized concern about government. There were high levels of MIP responses for dissatisfaction with government under both George W. Bush and Barack Obama, but it seems unlikely that the respondents would use similar terms for concern about each administration. The keyword “big government” was examined and appeared conceivable (although it is generally more related to conservative concerns over government), but the time series included a structural break in September 2009 following the opening of a political website called ‘Big Government’ (Big-Hollywood 2009).¹⁷

¹⁶ Note that this issue describes those MIP responses which talk about dissatisfaction with government rather than explicit government satisfaction questions.

¹⁷ This meant that the “big government” search index time series was non-stationary and couldn’t be compared to the stationary time series for MIP responses mentioning government.

Internet search measures still possess the useful properties of MIP. Since search data are used to predict MIP values, the resultant time series can be used to measure the relative importance of issues over time as well as the dynamics of a single issue's salience. MIP is plausible as a measure of issue salience and is derived from a broadly representative sample of the population. As such, the close fit of Internet search measures to MIP indicates that they are also tracking issue salience in the population.

CONCLUSION

This paper set out to assess whether Internet search data could provide measures of issue salience, which would be desirable because they are available more frequently and in more countries than existing MIP measures. To achieve this aim it was necessary to 1) create a process for assessing search data validity and 2) establish whether this process was able to identify search indices that were valid measures of issue salience within the US between 2004 and 2010. Out of seven issues that were salient over the time period, four issues (immigration, macroeconomics, fuel prices and terrorism) were able to be measured using search data. This conclusion was reached after showing that the search indices were generated from qualitatively plausible search terms (content validity) and that they closely tracked the most widely used issue salience measure (criterion validity). The correlations between the Internet search measure and the MIP measure were greater than 0.8 for all four issues. In addition, the search measures were shown, by using out-of-sample testing, to be reliable outside of the time period to which they were fitted. These findings mean that the strength of the relationships between search indices and MIP measures is of a similar magnitude to that found between the two most popular measures of issue salience, namely MIP and MII (Jennings and Wlezien 2011). Indeed, the correlation of 0.86 between "jobs" searches and macroeconomic responses is considerably higher than the 0.44 correlation found between MIP and MII responses on the economy (Jennings and Wlezien 2011).¹⁸

¹⁸It should be noted that the comparison of MIP and MII was over a considerably longer period (1977–2001) possibly making the composition of the economy category more heterogeneous over the twenty-four year period compared to the six year period studied here. Jennings and Wlezien also point to possible differences in coding criteria for these categories between Gallup and Ipsos-MORI. Nevertheless, it is striking that the relationship between MII and MIP would be weaker than that between MIP and Internet searches.

These findings have a number of implications for future research. As a result of the greater frequency of data, it is possible to test more complex hypotheses in agenda-setting. For instance, it should be possible to test whether it takes extended periods of media coverage to change public perceptions or if there is a reciprocal relationship between media coverage and salience. Establishing any of these relationships requires frequent observations of issue salience and time series techniques, both of which assume regular intervals between observations. Furthermore, search data has frequent and regular observations, whereas most MIP series do not.

The results also highlight the importance of testing the validity of Internet search measures. The salience of four out of seven issues was able to be measured, but many potential search measures were rejected because they either had problematic time series properties or because an examination of the search terms within them showed that they were unlikely to reflect the importance of an issue to the person searching. Picking a plausible sounding search term and examining its trends is insufficient to make claims about the general public. For example, Graefe and Armstrong attempted to measure the relative salience of the Iraq war and economic problems using searches for “Iraq” and “economy”, respectively. However both of these were shown to lack content validity. The Iraq searches are dominated by searches for beheading videos and the economy searches largely refer to fuel economy.¹⁹

Where validity can be established for search indices they offer many advantages for a researcher. However, there are several limitations that need to be considered when using these measures. The first limitation is that it may not always be possible to measure all the important issues in the time period. In this study three issues could not be measured: dissatisfaction with government, healthcare and the Iraq conflict. The second limitation is that content validity could become threatened in future by new, confounding search terms, which might invalidate some measures but would be detected easily through repeating the content validity tests to look for new confounding terms.

¹⁹ These search terms were also tested for criterion validity in order to check whether Graefe and Armstrong’s findings are likely to be affected. Unsurprisingly, given the content validity results, both of these series fail criterion validity tests as well.

Finally, although search data is more widely available than survey data, it is still restricted by the need for MIP survey data against which to validate the new measures. Nonetheless, this validation could be achieved with fewer MIP data points than the seventy-five available in this study. This limitation will become less important over time, as even countries where only a few MIP surveys are conducted each year will have enough data for validity testing, as the time period jointly covered by internet search data and surveys increases. Indeed, the benefits are potentially greater for these areas, as researchers may be able to move from a biannual time series to a weekly one.

Using Internet search data might seem complex due to the many ways in which its validity must be assessed. Nevertheless, not using Internet search data limits the scope of problems that can feasibly be studied, as the most important problem can change quickly and processes of agenda-setting may vary across countries. Infrequent data originating from a small group of countries limits causal inferences and the generalizability of any results, but measures of issue salience derived from Internet searches will allow for a more detailed and wider examination of agenda setting.

REFERENCES

- Baumgartner, Frank, and Bryan Jones. *Topics Codebook*. Policy Agendas Project 2006.
- Big-Hollywood. *New Political Blog 'Big Government' Launches Thursday* 2009. Available from <http://bighollywood.breitbart.com/bighollywood/2009/09/09/new-political-blog-big-government-launches-tomorrow/>.
- Brunn, Stan, Lomme Devriendt, and Andrew Boulton. 2010. City networks in cyberspace and time : using Google hyperlinks to measure global economic and environmental crises. In *Mobile and Ubiquitous Urban Infrastructures : surveillance, locative media and global networks*, edited by R. Firmino, F. Duarte and C. Utramari: IGI Global.
- Bélanger, Éric, and Bonnie M. Meguid. 2008. Issue salience, issue ownership, and issue-based vote choice. *Electoral Studies* 27 (3):477-491.
- Carmines, Edward G., and Richard A. Zeller. 1979. *Reliability and validity assessment, A Sage university paper : Quantitative applications in the social sciences*. Beverly Hills, Calif.: Sage Publications.
- Cleveland, Robert, William Cleveland, Jean McRae, and Irma Terpenning. 1990. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. *Journal of Official Statistics* 6 (1):3-73.
- Davidson, Russell, and James G. MacKinnon. 1993. *Estimation and inference in econometrics*. New York: Oxford University Press.
- Dube, A., and E. Kaplan. 2012. Occupy Wall Street and the Political Economy of Inequality. *The Economists' Voice* 9 (3).
- Ellis, Curtis, Colin Swearingen, and Joseph Ripberger. 2011. Examining the Impact of Public Attention on Fundraising in US Senate Elections.
- Gallup. 2010. Gallup Monthly Most Important Problem Questions, 2004-2010. 1st Roper Center for Public Opinion Research.
- Ginsberg, Jeremy, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant. 2009. Detecting influenza epidemics using search engine query data. *Nature* 457 (7232):1012-1014.
- Google. 2004-2010. Google Insights for Search.
- Graefe, Andreas, and J. Scott Armstrong. 2012. Predicting elections from the most important issue: A test of the take-the-best heuristic. *Journal of Behavioral Decision Making* 25 (1):41-48.
- Granger, Clive W. J., and Paul Newbold. 1974. Spurious regressions in econometrics. *Journal of Econometrics* 2 (2):111-120.
- Granka, Laura A. 2010. Measuring Agenda Setting with Online Search Traffic: Influences of Online and Traditional Media. *2010 Annual Meeting of the American Political Science Association, September 2-5, 2010*.
- Green, Jane, and Sara B. Hobolt. 2008. Owning the issue agenda: Party strategies and vote choices in British elections. *Electoral Studies* 27 (3):460-476.
- Hansen, Bruce E. 1995. Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power. *Econometric Theory* 11 (05):1148-1171.
- IMDB. *Notorious (2009)*. Internet Movie Database 2009.
- Ioannidis, John P. A. 2005. Why Most Published Research Findings Are False. *PLoS Med* 2 (8):e124.

- Jennings, Will. 2009. The Public Thermostat, Political Responsiveness and Error-Correction: Border Control and Asylum in Britain, 1994-2007. *British Journal of Political Science* 39 (04):847-870.
- Jennings, Will, and Christopher Wlezien. 2011. Distinguishing Between Most Important Problems and Issues? *Public Opinion Quarterly* 75 (3):545-555.
- Jones, Bryan D., Heather Larsen-Price, and John Wilkerson. 2009. Representation and American Governing Institutions. *The Journal of Politics* 71 (01):277-290.
- Kahn, Matthew E., and Matthew J. Kotchen. 2010. Environmental Concern and the Business Cycle: The Chilling Effect of Recession. *National Bureau of Economic Research Working Paper Series* No. 16241.
- McDonald-Johnson, Kathleen, Brian Monsell, Ryan Fescina, Roxanne Feldpausch, Catherine Hood, and Monica Wroblewski. 2010. Seasonal Adjustment Diagnostics. edited by U. C. Bureau. Washington DC.
- Newey, Whitney K., and Kenneth D. West. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55 (3):703-708.
- Oltra, Christian. 2011. Stakeholder perceptions of biofuels from microalgae. *Energy Policy* 39 (3):1774-1781.
- Pelc, K. 2011. Testing Assumptions of International Political Economy Using Search Engine Query Data.
- Pew. 2010. Social Side of the Internet.
- Reilly, Shauna, Sean Richey, and J. Benjamin Taylor. 2012. Using Google Search Data for State Politics Research. *State Politics & Policy Quarterly* 12 (2):146-159.
- Reis, Ben Y., and John S. Brownstein. 2010. Measuring the impact of health policies using Internet search patterns: the case of abortion. *BMC public health* 10:514.
- Rickard, Matthew. Search Data vs. Election Outcomes. *population* 76 (11.28):12.26.
- Ripberger, Joseph T. 2011. Capturing Curiosity: Using Internet Search Trends to Measure Public Attentiveness. *Policy Studies Journal* 39 (2):239-259.
- Scharkow, Michael, and Jens Vogelgesang. 2011. Measuring the Public Agenda using Search Engine Queries. *International Journal of Public Opinion Research* 23 (1):104.
- Scheitle, Christopher P. 2011. Google's Insights for Search: A Note Evaluating the Use of Search Engine Data in Social Research*. *Social Science Quarterly* 92 (1):285-295.
- Schimelpfening, and Nancy. *Google Trends Indicate Rising Suicidal Thoughts in Americans*. About.com 2010. Available from <http://depression.about.com/b/2010/01/20/google-trends-indicate-rising-suicidal-thoughts-in-americans.htm>.
- Silver, Nate. 2012. Maine Caucus Offers an Opportunity for Paul. In *FiveThirtyEight*. New York Times.
- Soroka, Stuart. 2002. Number of Responses and the Most Important Problem. In *Nuffield College Politics Working Paper*. Oxford: Nuffield College.
- Stephens-Davidowitz, S. 2011. The Effects of Racial Animus on Voting: Evidence Using Google Search Data.
- Weeks, Brian, and Brian Southwell. 2010. The Symbiosis of News Coverage and Aggregate Online Search Behavior: Obama, Rumors, and Presidential Politics. *Mass Communication and Society* 13 (4):341-360.
- Weiner, Rachel. 2012. Google shows how debate hurt Rick Santorum. In *The Fix*. washingtonpost.com.

- Wessa, Patrick. *Free Statistics Software, Office for Research Development and Education, version 1.1.23-r6* 2010. Available from <http://www.wessa.net/>.
- Wlezien, Christopher. 2005. On the salience of political issues: The problem with '[]most important problem'. *Electoral Studies* 24 (4):555-579.
- Wooldridge, Jeffrey M. 2009. *Introductory econometrics : a modern approach*. 4th ed. Mason, OH: South Western, Cengage Learning.

TABLES

Table 1| Content validity problems for each search index and monthly volume of searches

Issue	Search Index	Problematic searches	Passes content validity check
Macroeconomics	Economy	Fuel economy, foreign countries	No
Macroeconomics	US economy	Greece, China	Yes
Macroeconomics	Jobs	Steve, blow	Yes
Macroeconomics	Unemployment	None	Yes
Macroeconomics	Tax burden	None	Yes
Macroeconomics	Taxes	Terms related to financial planning	No
Macroeconomics	Fiscal	Spanish and Portuguese terms related to finance	No
Macroeconomics	Budget Deficit	None	Yes
Iraq	Iraq	Beheading, Abu Ghraib	No
Iraq	Iraq War	None	Yes
Iraq	Iraq Casualties	Afghanistan, Vietnam	Yes
Healthcare/Hospital	Healthcare	None	Yes
Government	Government	Jobs	No
Government	Big government	None	Yes
Immigration	Illegal immigration	None	Yes
Immigration	Immigration	Visas, green cards, etc.	No
Terrorism	Terrorism	None	Yes
Terrorism	Terrorism or Al Qaeda	None	Yes
Fuel Prices	Fuel	None	Yes
Fuel Prices	Gas	None	Yes

Table 2|Criterion validity testing by OLS regression analysis

Independent Variable (Search Index)	Dependent variable (MIP series)	Coefficient	Newey-West Std. Error	R ²
Jobs	Macroeconomics	1.726*	0.132	0.754
Illegal immigration	Immigration	0.284*	0.038	0.668
Terrorism and Al Qaeda (time trend)	Terrorism	0.305*	0.051	0.760
Fuel	Fuel Prices	0.522*	0.063	0.715
Gas	Fuel Prices	0.824*	0.218	0.491
Terrorism (time trend)	Terrorism	0.029	0.037	0.668

* p<0.01 n=75 for all regressions

Table 3|Even/odd year out-of-sample reliability test results

Independent variable (search index)	Dependent variable (MIP series)	RMSE in-sample	RMSE out-of-sample	Mean MIP series
Jobs	Macroeconomics	8.12	11.54	36.7
Illegal immigration	Immigration	2.13	2.67	5.0
Terrorism and Al Qaeda	Terrorism	2.82	2.31	6.3
Fuel	Fuel Prices	3.42	2.33	4.7

Table 4|Contiguous out-of-sample reliability test results

Independent variable (search index)	Dependent variable (MIP series)	RMSE in-sample	RMSE out-of-sample	Mean MIP series
Jobs	Macroeconomics	6.54	14.7	36.7
Illegal immigration	Immigration	1.99	2.80	5.0
Fuel	Fuel Prices	3.01	3.15	4.7

FIGURES

Figure 1|Seasonal Trend Decomposition by LOESS of searches for “jobs” on Google showing the original data and its seasonal, trend and remainder components. The axes are all expressed on the same arbitrary scale.

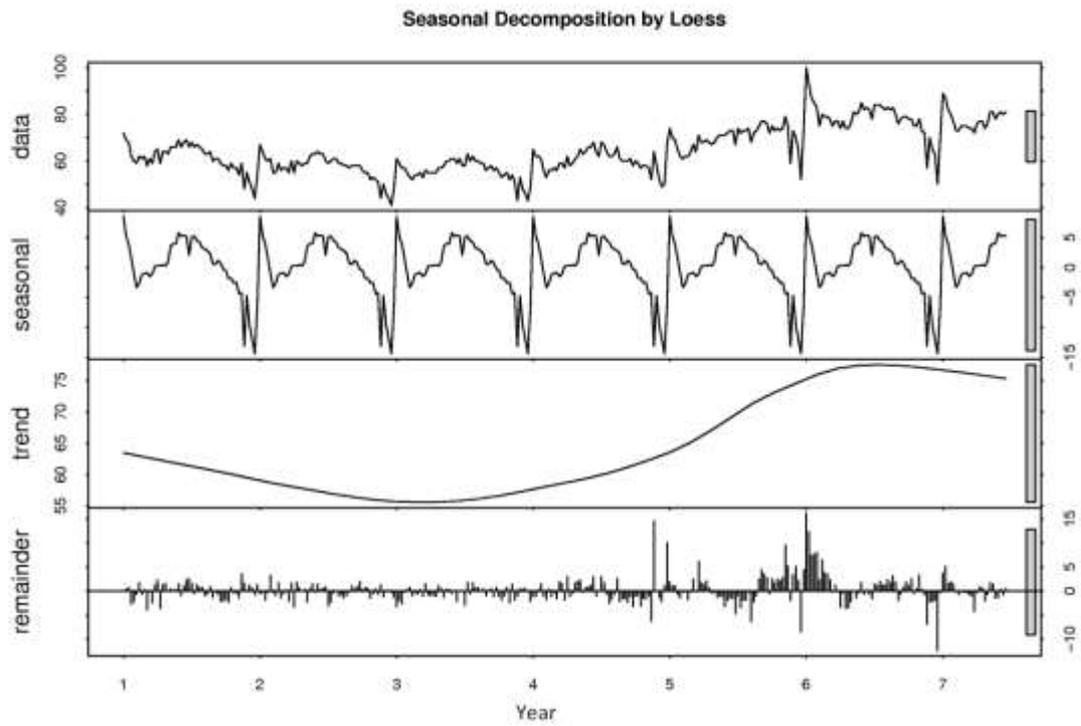


Figure 2| Flowchart showing the process involved in selecting search indices that are valid for measuring issue salience.

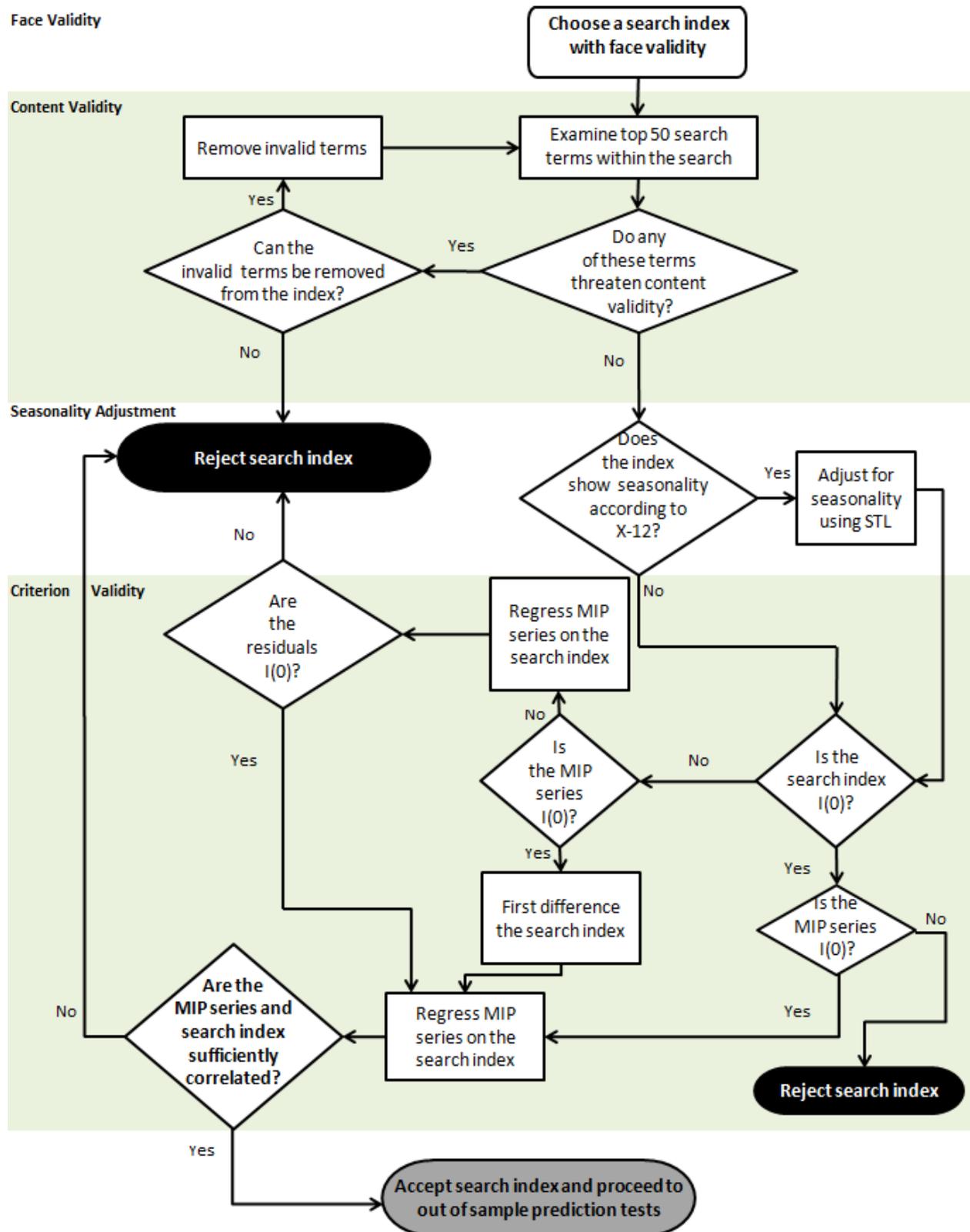


Figure 3|Frequency of Internet searches and MIP responses for terrorism, fuel, immigration and macroeconomics for each week where there is an MIP survey.

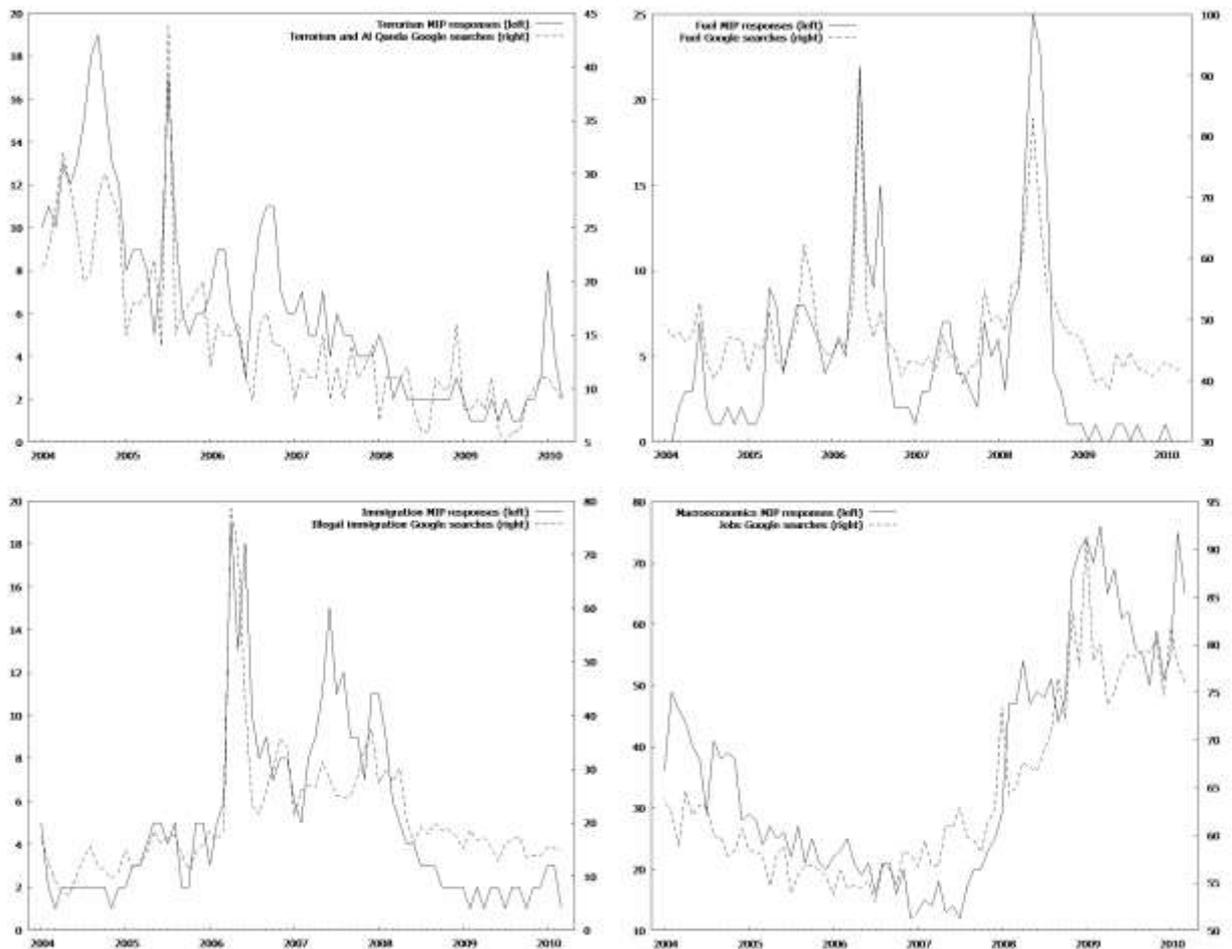


Figure 4|Weekly issue salience series predicted from Google Trends showing the predicted percentage of respondents who would have cited each issue as an MIP in each week between 2004 and 2010.

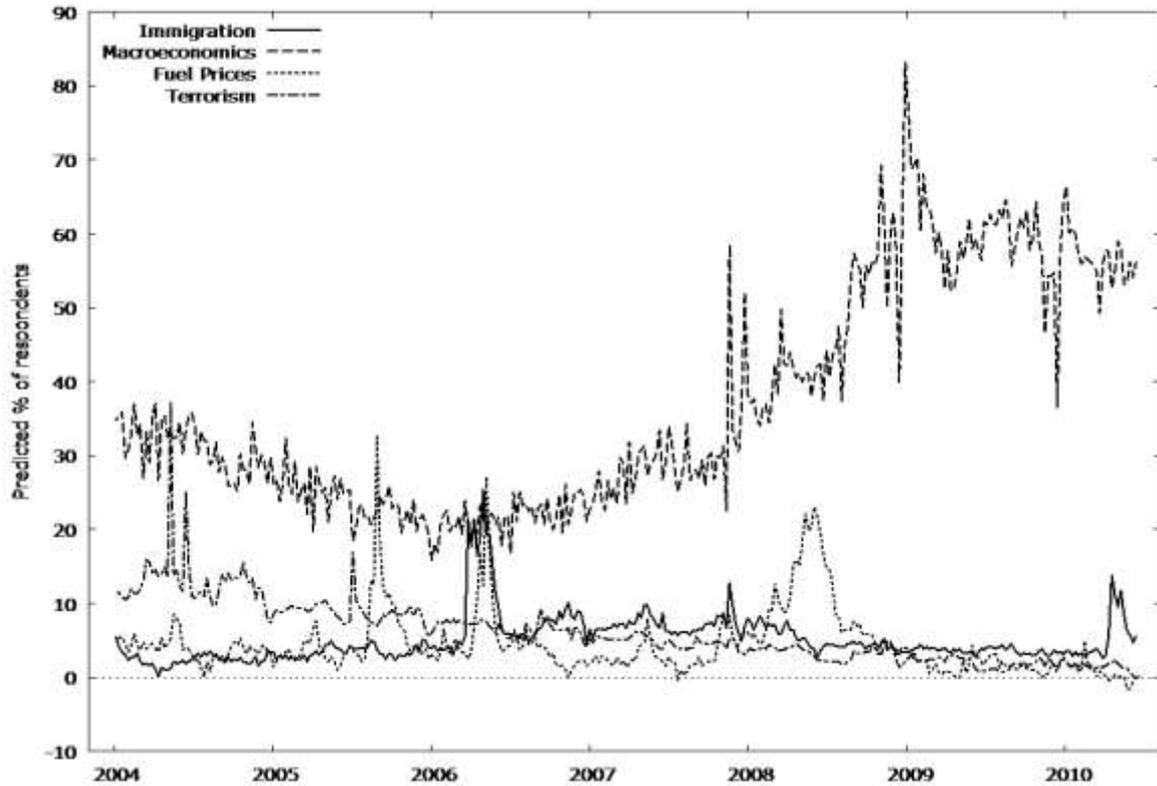
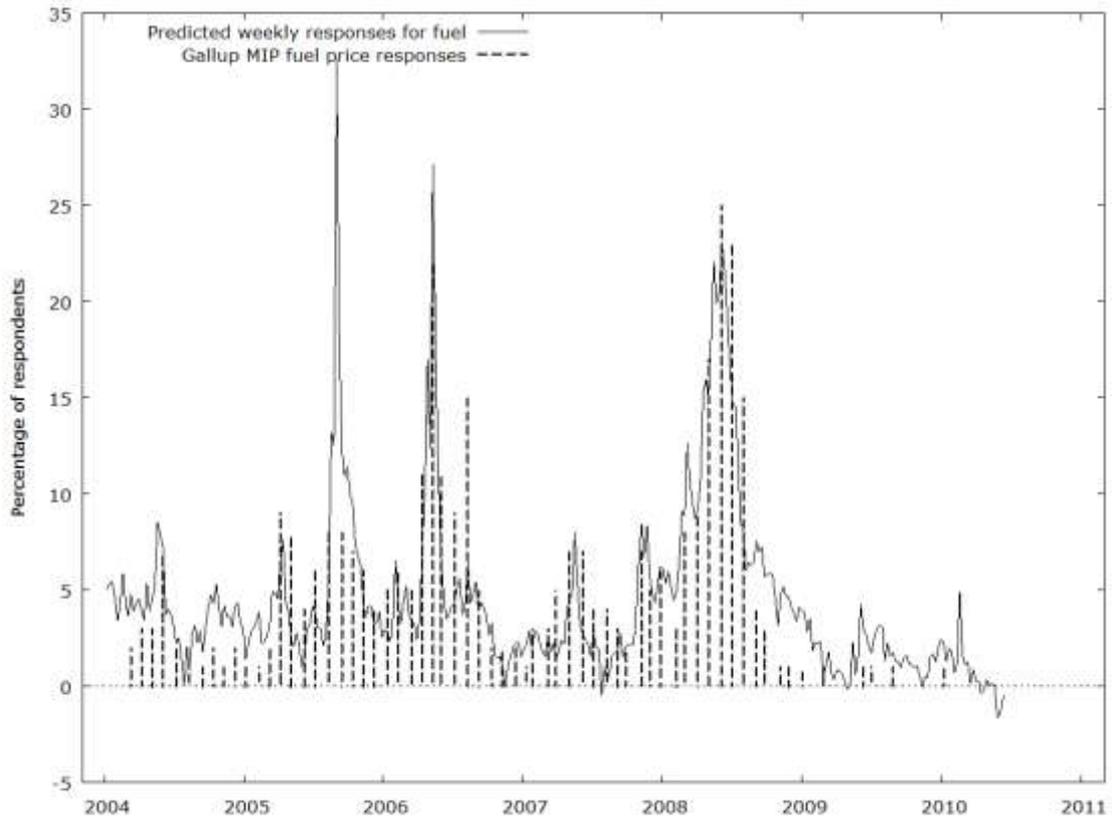


Figure 5| Predicted percentage of weekly MIP responses for fuel prices compared to monthly MIP percentage of responses.



APPENDIX 1

Table 5 shows a number of papers that have already made use of Google Insights for Search data.

This list is not complete and excludes working papers that requested not to be cited.

Table 5 | A selection of papers addressing political topics making use of Google Insights for Search data

Paper Title	Author(s)/Year	Search Terms
Measuring Agenda setting with online search traffic: Influences of online and traditional media	Granka 2010	story specific search terms e.g. "Iran election", "Joe the Plumber", "Wikileaks" etc.
Predicting elections from the most important issue: A test of the take-the-best heuristic	Graefe and Armstrong 2012	"iraq", "economy"
City networks in cyberspace and time: using Google hyperlinks to measure global economic and environmental crises	Brunn, Devriendt, and Boulton 2010	"global financial crisis", "climate change"
Using Google Search Data for State Politics Research An Empirical Validity Test Using Roll-Off Data	Reilly, Richey, and Taylor 2012	[names of ballot measures]
Testing Assumptions of International Political Economy Using Search Engine Query Data	Pelc 2011	"WTO"
Examining the Impact of Public Attention on Fundraising in US Senate Elections	Ellis, Swearingen, and Ripberger 2011	[Candidate names]
Stakeholder perceptions of biofuels from microalgae	Oltra 2011	"microalgae"
Measuring the impact of health policies using Internet search patterns: the case of abortion	Reis and Brownstein 2010	"abortion"
The symbiosis of news coverage and aggregate online search behavior: Obama, rumors, and presidential politics	Weeks and Southwell 2010	"obama muslim"
Occupy Wall Street and the Political Economy of Inequality	Dube and Kaplan 2012	"inequality"
Environmental Concern and the Business Cycle: The Chilling Effect of Recession	Kahn and Kotchen 2010	"unemployment", "global warming"