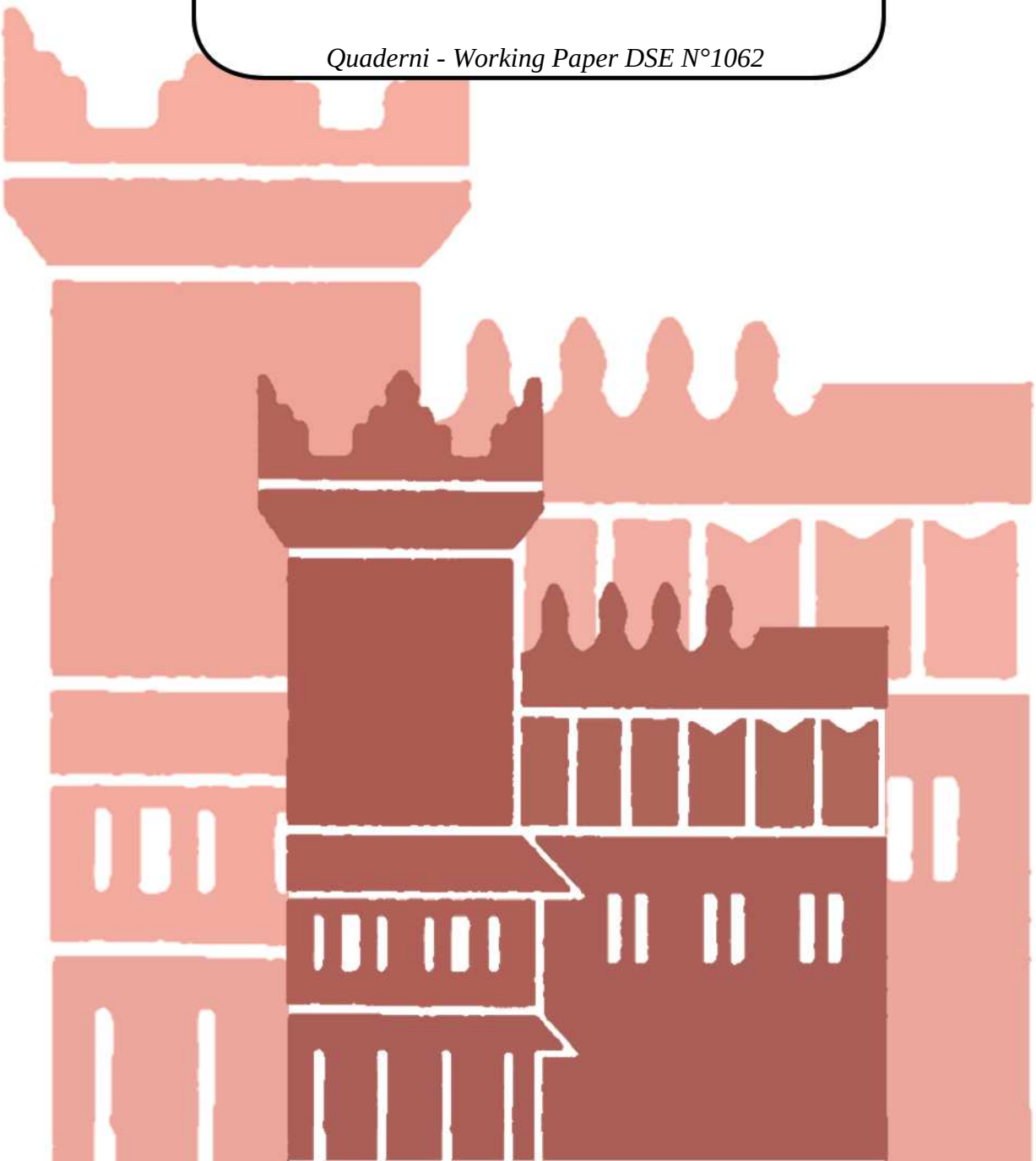
ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

A New Index of Uncertainty Based on Internet Searches: A Friend or Foe of Other Indicators?

M. E. Bontempi
R. Golinelli
M. Squadrani

Quaderni - Working Paper DSE N°1062



A New Index of Uncertainty Based on Internet Searches: A Friend or Foe of Other Indicators?

M. E. Bontempi[*], R. Golinelli[**] and M. Squadrani[*]

Abstract

The preliminary evidence in the literature suggests that changes in uncertainty have a role in shaping the U.S. economic cycle. But what is effectively measured by the different available indicators of uncertainty still remains an "uncertain" issue. This paper has two aims: (*i*) to introduce a new uncertainty indicator (*GT*) based on Internet searches; and (*ii*) to compare the main features and the macroeconomic effects of alternative measures of uncertainty, including our own. Results suggest that *GT* shocks embody timely information about people's perception of uncertainty and, in some cases, earlier than other indexes. Furthermore, the effect of uncertainty shocks on output is more influenced by parameter breaks due to in-sample events than by model specification. The consequence is that an all-comprehensive indicator able to weight different sources of uncertainty is preferable.

Keywords: Uncertainty measures; finance-based, survey-based, news-based; Internet searches; Google; comparison across indicators; macroeconomic effects of uncertainty.

JEL codes: D80, E32, E27, E37, C32.

[*] Department of Economics, University of Bologna, Italy

[**] *Corresponding author*
Department of Economics, University of Bologna
Strada Maggiore, 45; 40125 Bologna (Italy)
roberto.golinelli@unibo.it
https://sites.google.com/site/robigolinelli/

This version: March, 8th 2016

Index

1. Introduction[1]

How do consumption and investment respond to uncertainty shocks? What role does uncertainty play in macroeconomic fluctuations? Is macroeconomic uncertainty a possible cause of weak recovery after adverse shocks, or is it rather a consequence of negative economic conditions? Over the last decade, questions like these have become increasingly important, as an attempt is made to better understand the role of uncertainty in shaping the global financial crisis.

Quantitative results rely on different estimates of the unobservable, time-varying macro-economic uncertainty, the measure of which is neither unique nor objective. Hence the question of how to proxy uncertainty has attracted increasing interest in recent years, as witnessed, for example, by the papers of Bachmann et al. (2013), Bekaert et al. (2013), Bloom (2009), Julio and Yoox (2012), Jurado et al. (2015), Rich and Tracy (2010), and Rossi and Sekhposyan (2015). The different measures of uncertainty formulated to date broadly follow three alternative approaches (see also Bloom, 2014).

The first approach is “finance-based”, as it utilizes somewhat sophisticated methods to process the financial information thrown up by stock market volatility (see e.g. Bekaert et al., 2013, Bloom, 2009, Gilchrist et al., 2014, and Knotek and Khan, 2011), under the assumption that financial volatility can be a guide to the state of macroeconomic uncertainty even though not all individuals invest in the stock market (Romer, 1990), or share the same information that financial market actors have access to.

The second approach is “forecasts-based”, as it estimates uncertainty by relying on the concept of the economy’s predictability and on the measurement of the discrepancy between professional forecasts, under the assumption that the lack of predictability and a large discrepancy between forecasters' views reflect (and are both perceived as signs of) a more uncertain economy. As with the previous approach, a large variety of methods are employed: see Bachmann et al. (2013), Jurado et al. (2015), Rich and Tracy (2010), Rossi and Sekhposyan (2015), Scotti (2013), and Segal et al. (2014).

The third approach is “news-based”. Given the question "how does the average citizen comprehend the implications of stock market volatility and economic predictability underlying his/her uncertainty?", the answer is "media is the messenger" (see Alexopoulos and Cohen, 2009). Therefore, the degree of uncertainty in a given period can be proxied by the frequency with which a lengthy list of words related with uncertainty appears in journalistic articles. The assumption is that,

[1] Paper presented at the 35th International Symposium on Forecasting of the International Institute of Forecasters, Riverside, June 21-24, 2015. We are grateful, without implication, to Nick Bloom, Joshy Easaw, Carlo Favero, Alessandro Girardi, Giuseppe Parigi, and Barbara Rossi for their comments, as well as to conference participants.

when certain causes of uncertainty matters, they are likely to be reported by journalists through the use of certain words. More specifically, it is assumed that the media are able to gauge any uncertainty indicated by market outcomes, professional economists and political debate, and to draw the general public's attention to this uncertainty through the recurrent use of specific words. This approach - which is similar to the narrative analysis designed to identify monetary and fiscal policy shocks (see, for example, Romer and Romer, 2004, and Ramey, 2009) - leads to the news-based uncertainty measures formulated, for example, by Alexopoulos and Cohen (2009), Baker et al. (2015), and Knotek and Khan (2011).

The three previous approaches all have their pros and cons. On the one hand, the pre-selection of directly observable specific events are easy to perceive but somewhat arbitrary. On the other hand, the methodological approaches that extract uncertainty estimates from latent processes are statistically and economically sound, but are also very complex black boxes not strictly related to observable indicators of uncertainty.

This paper has two aims. The first is to introduce a new uncertainty indicator (*GT*) that methodologically belongs to the class of news-based measures, but that replaces the more traditional counting of words reported by the press with the intensity of Internet searches. In short, *GT* shifts from word counts to individual searches, i.e. it focuses on the receiver (individuals) rather than on the channel through which the message is conveyed (press, media). While the traditional news-based measure uses journalists' feeling about uncertainty and how they communicate it by using specific words, our measure is based on how Internet users explicitly manifest their uncertainty by searching for specific words with greater/less frequency. Higher levels of uncertainty are likely to result in a greater appetite for knowledge, and consequently in a more intense use of tools capable of gathering further information. In particular, we assume that the intensity of Internet searches is related to the uncertainty about the outcome of a specific political or economic event weighted by its subjective relevance as perceived by individuals. In this way, the intensity of Internet searches reflects uncertainty regarding the occurrence of an event, and the consequences of that event, which in principle may also be certain to occur. Google Trends is the tool we use to assess the search intensity of terms linked with economic and policy news, and to understand the level of information that individuals need in order to increase their awareness. To the best of our knowledge, this is the first time that Google Trends has been used to measure in a comprehensive way the need for information within the context of the measurement of uncertainty.[2]

[2] Our choice of Google Trends is motivated in Section 2.2. Preliminary realizations were independently implemented by Squadrani (2013) and BBVA (2012); a simplified use of Google Trends, based on the single search term "economy" can be also found in Dzielinski (2012). By contrast, Google Trends is regularly used within the context of forecasting in order to provide predictors: see, for example, Choi and Varian (2012), and D'Amuri and Marcucci (2012).

The second aim is to rectify the lack of comparisons, over a common time span, of the main features of the uncertainty measures for the U.S., to disentangle the signals that each of them uses to depict macroeconomic uncertainty, and to better understand their relationship with the economic cycle. In doing so, we also include our brand new Google-based index in the comparative analysis.

Results show that Internet searches can provide the measures of macroeconomic uncertainty with fresh information. Therefore, the answer to the main question of the paper is that Internet search activity is a friend of the alternative uncertainty measures because it is able to capture individuals' perceived uncertainty not embodied in the other indicators. For example, we show that, in a situation of increasing uncertainty, Internet searches and press coverage follow different time paths. Internet searches sometimes peak before press coverage (as in the case of fiscal policy issues), while at other times their causality is either reversed (as in the case of monetary policy issues), or vanishes when they are for the best part contemporarily related (as in the case of questions of entitlement programs).

The comparison of the macroeconomic effects of alternative uncertainty measures (i.e. finance-, forecasts- and news-based) shows that they are influenced by model parameter breaks due to in-sample events.

The paper is organised as follows. Section 2 presents both the conceptual framework and the technical issues of our new "search-based" Google Trends uncertainty index (*GT*). Section 3 analyses the dynamic interaction between our series of Internet searches and those news-based searches in Baker et al. (2015), in order to understand whether media coverage drives Internet searches, and whether such searches continue even after press coverage wanes. Section 4 surveys the "fantastic five" uncertainty indicators (i.e. those that are most frequently used in the literature, and that are publicly available on a monthly basis) by allocating them to the aforementioned three methodological approaches. Moreover, the section compares their statistical features at both univariate (persistence, seasonality and variability over time) and multivariate (Granger causality and contemporaneous correlation) levels. Section 5 focuses on the interaction between alternative uncertainty measures (including *GT*) and certain key macroeconomic variables (such as output and employment). Section 6 offers our conclusions.

2. The making of *GT*, the new index of uncertainty based on Google Trends

This section motivates our new uncertainty measure based on Google's search volumes (in Section 2.1) and describes the steps taken to create it (in Section 2.2). Given that search volumes reflect individuals' need for information about specific terms, their listing and wording is of paramount importance. We have chosen to extract Google Trends series with 184 queries closely

related to the 210 search terms used by Baker et al. (2015), henceforth *BBD*, to make "Newsbank", that is the Categorical Newsbank version of *BBD* news-based economic policy uncertainty index (EPU, see more details in Section 4.1). Newsbank is exclusively based on news data, and is developed starting from specific policy area uncertainty sub-indexes.[3] The 184 single series can be aggregated into 8 Newsbank policy categories, and the latter are summed up into the macro-economic index, *GT*.[4]

2.1 - The conceptual framework of the GT index

The science of uncertainty quantification (see, among others, Der Kiureghian and Ditlevsen (2009)) establishes that uncertainty may be either aleatory (statistical) or epistemic (systematic). Aleatory uncertainty is irreducible: it arises naturally from perceiving real-life facts (that is, from "observing the system"), since it represents unknowns that randomly differ each time we examine those facts. Epistemic uncertainty represents the lack of knowledge about potentially knowable things that are unknown in practice.[5] In both cases, we are faced with a lack of knowledge. In order to measure this lack of knowledge, we have to assume that there are "hidden" variables which can better inform individuals about the issues in question, and enable them to distinguish the truth from among the possible alternatives: the larger the set of possible alternatives, the greater the non-specificity resulting in uncertainty. Moreover, uncertainty may be also driven by ambiguity and conflict (that is, by a lack of consensus or by disagreement): the more mutually exclusive items of evidence individuals have, or the more even the strength of the disagreeing items of evidence is, the greater the conflict.[6]

Both non-specificity and conflict are mostly due to the poor quality of the information available (Harmanec, 1999), and this fact inevitably fuels individuals' need to gather further information in order to bridge the gap between what is known and what needs to be known so as to be able to make better decisions; see also the discussion in Bloom (2014). This point is widely acknowledged in various fields. For example, in the field of economic psychology (see, for example, Lemieux and Peterson, 2011), individuals respond to greater uncertainty by intensifying

[3] This index is based on the count of articles from the Access World News Newsbank ("Newsbank") - a database of about 2,000 national and regional newspapers in the US. See the complete list of queries in the appendix to Baker et al. (2015) and on their website (www.policyuncertainty.com/categorical_terms.html). The 184 Google Trend queries we defined after rearrangements and validation of BBD search terms are listed in Appendix A1.

[4] Macroeconomic *GT* and data by category are obtained using the common-term aggregation procedure. Technical details are in Appendix A2.

[5] For example, with regard to uncertainty in official economic statistics, Manski (2015) considers uncertainty as the lack of knowledge deriving from an incomplete understanding of the information provided about an economic concept, or from the lack of clarity of the concepts themselves.

[6] The economic rationale of this point is modelled in Ilut and Schneider (2014).

their search for more information. Similarly, sticky information models predict that "more volatile shocks [in other words, greater uncertainty] lead to the more frequent updating of information, since inattentiveness is more costly in a world that is rapidly changing", Reis (2006, p. 803).

If uncertainty can be reduced by increasing knowledge, the intensity of search for more knowledge can be used to measure the level of uncertainty. In other words, higher levels of uncertainty lead to a greater appetite for increased knowledge, and consequently result in a more intense use of tools capable of gathering further information. In recent years the Internet has become an effective means of collecting and divulging information for an increasing number of people, and thus data regarding internet searches could prove useful in measuring uncertainty. Google Trends is one possible instrument with which to gather data on Internet searches. [7]

We think that this framework can motivate our proposal of transitioning from a "news-based" to a "search-based" approach (intensity of Internet searches).

2.2 – Pros and Cons to the utilization of Google Trends data and how we dealt with

Google Trends provides an index of the volume of queries on Google since 2004. This index is called the Search Volume Index - in symbols SVI_{st} – and it measures the volume of queries for a generic search term (or group of terms) *s* in a given region (in our case the U.S.A.) at time *t:*

$$SVI_{st} = \frac{sv_{st}}{sv_{Gt} \times MSV_{[o,T]}} \times 100 = \frac{sv_{st}}{sv_{Gt} \times \max_{t=[0,T]} \left\{ sv_{st} / sv_{Gt} \right\}} \times 100 \qquad (1)$$

where sv_{st} is the number of searches for *s* within period *t*; the periodicity of observations is weekly or monthly, depending on the search term's popularity.[8] The division by sv_{Gt}, the total number of Google searches within the same period *t*, is done in order to prevent SVI_{st} from being affected by the overall increase in Internet users over time. Moreover, SVI_{st} time series are bounded between 0 and 100 because time series are scaled by the maximum value of sv_{st}/sv_{Gt} from 0 to *T* (i.e. over the download period), and then are multiplied by 100. Google Trends only provides observations for those search terms exceeding a minimum threshold of popularity, otherwise they are set to zero; therefore zero SVI_{st} indicates either no or insignificant searches for *s* at time *t*.

[7]With regard to the composition and relevance of Internet usage over the US population, note that in 2015 Internet users in North America represented 88% of the country's population (see www.internetworldstats.com). Furthermore, Google is the leading search engine, with a worldwide market share of 68.5% in May 2015 (see www.netmarketshare.com). However, the risk of incurring in population bias exists, and there are also other issues relating to the use of social media data: see Ruths and Pfeffer (2014).

[8]Search terms with low popularity are only available on a monthly basis. In our database, weekly Google Trends series are converted into monthly format by averaging them on the basis of the month in which the week begins.

SVI indicators have their pros and cons. On the positive side, they are freely available, measured at very high frequency, regularly updated, and early released. In addition, being peak-normalised, the *SVIs'* sensitivity to extreme values is *per se* sharply reduced, while other ways of treating extreme values (such as outlier-trimming) can bias the original data structure (see e.g. Konig and Ohr, 2012). The normalization approach, in fact, produces indexes the changes in which proxy distances from their peaks, while differences across *s* are independent of their relevance over total Google traffic. An increase in the required information about term *s* is not measured by an increase in its share, but rather by an increase in its level towards 100. Therefore, SVI_{st} indexes are short-term indicators, measuring how close the need for information about *s* at time *t* is to its highest point, rather than structural indicators identifying the most searched-for terms.

On the negative side, such indicators are subject to sampling variability, which depends on the day of the download (Choi and Varian, 2012). Therefore, it is impossible to replicate the search volumes exactly due to the fact that they may differ slightly from one download day to the next. Following Da et al. (2011), who report that correlation is usually greater than 0.97 for series downloaded several times, we argue that the *SVI*s for a search term only change very slightly from one day to another, especially when considering highly popular terms.[9] Moreover, the lack of information about the relevance of each *SVI* is a problem involving the aggregation of individual search terms' volumes into groups of overall means, as we do not know the weights of those sums.[10]

Here we just note that both the aggregation of our 184 search terms in 8 categories representing 8 policy areas comparable with *BBD* index, and the sum of these categories into the macroeconomic *GT* is made using the common-term approach.[11] The intermediate level of aggregation in categories is motivated by the need of comparing by component news- and search-based indicators (i.e. *BBD* and *GT*) without available data on single *BBD* search terms.[12] Hence, the information flows between alternative media (newspapers, *BBD*, and Google searches, *GT*) will be studied, in the next Section 3, at the level of aggregation by categories.

[9] In order to address this sampling issue, we downloaded *SVI*s for more than 200 terms from Google Trends on 13 consecutive days (from the 8th to the 20th of June, 2014), and then computed sample correlation among the series in relation to different days. We found that sampling problems (identified as a correlation lower than 0.9) arise mainly in the presence of series containing zeros (i.e. low popular series).Therefore, we do not consider this as a relevant issue, because in the aggregation process the weight of low popularity series is negligible.

[10] Alternative *SVI* aggregation procedures are described in Appendix A2.

[11] Details are in Appendix A2.

[12] In fact, the category is the highest level of disaggregation available for both sources (BBD and GT), apart from two terms for which we report comparisons at single search term level in Appendix A4.

3. How uncertainty flows? Comparing Google Trends and Newsbank by category

Little attention has been paid to analysing the relationships between press coverage and the information found on the Internet.[13] In this section, by taking advantage of the availability of two consistent data sources (BBD's Newsbank series by category and the corresponding Google Trends), we try to shed some light on how newspaper reports and Internet searches interact in response to uncertainty shocks in order to deepen the knowledge by component of the news-based and *GT* indicators of macroeconomic uncertainty.

3.1 – The methodological framework

The basic ingredients are two sets of series measuring searches for the same terms from a particular policy category: BBD_{ct} (BBD's news-based counts), and GT_{ct} (Google Trends search volumes); where c = 1, 2, .., 8 (policy categories), and t are monthly observations over the period 2004m1-2014m12, i.e. T = 132 months. Although referring to the same search terms, these two variables are different: BBD_{ct} measures uncertainty as the frequency with which certain terms are mentioned in the newspapers (reflecting journalists' decision to focus their readers' attention on certain issues rather than on others), while GT_{ct} defines uncertainty as the frequency with which the same terms are used in Google searches. Therefore, according to the BBD_{ct} series, journalists are the messengers of the state of uncertainty, which they communicate by using specific words whereas, according to GT_{ct}, Google users manifest their uncertain mood by searching in a more/less intense manner using those same words.

The dynamic relationships between these variables can be assessed within the context of the VAR model. Suppose that, for the c^{th} category, the k-dimensional stationary VAR(p) process y_{ct} consists of the m-dimensional process z_{ct} and the $(k-m)$-dimensional process x_{ct} with non-singular white noise covariance matrix $\Sigma_{\varepsilon c}$

$$y_{ct} = \begin{bmatrix} z_{ct} \\ x_{ct} \end{bmatrix} = \begin{bmatrix} \mu_{c1} \\ \mu_{c2} \end{bmatrix} + \begin{bmatrix} A_{c11,1} & A_{c12,1} \\ A_{c21,1} & A_{c22,1} \end{bmatrix} \begin{bmatrix} z_{ct-1} \\ x_{ct-1} \end{bmatrix} + \ldots + \begin{bmatrix} A_{c11,p} & A_{c12,p} \\ A_{c21,p} & A_{c22,p} \end{bmatrix} \begin{bmatrix} z_{ct-p} \\ x_{ct-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{c1t} \\ \varepsilon_{c2t} \end{bmatrix} \qquad (2)$$

where, in our bivariate context: k=2 and m=1; $y_{ct} = (BBD_{ct}, GT_{ct})'$ is the vector of the variables of interest for the c^{th} category; scalars μ_c and matrices A_c are heterogeneous parameters (they are allowed to differ across categories); and $\varepsilon_{ct} = (\varepsilon_{c1t}, \varepsilon_{c2t})' = (\varepsilon_{ct}^{BBD}, \varepsilon_{ct}^{GT})'$ is the vector of

[13] Some empirical research has focused on modelling the ways information is divulged via the Internet, and the speed at which this happens in matters far removed from the question of uncertainty: see, for example, Eberth et al. (2014).

the random shocks to *BBD* and *GT* uncertainty measures for *c*. The lag length *p* is first set by using the AIC criterion (starting from *p*=13), and the corresponding residuals are tested for white noise errors (otherwise *p* is slightly altered until the white noise residuals target is met). The joint stationarity of all the variables listed in the VAR model, i.e. that the system (2) has full rank, can be assessed by using the Johansen (1995) trace test: if the sequence of null hypotheses of reduced rank is always rejected, then the data support the stationarity assumption. Finally, seasonal dummies are included in the VAR models when they are significant.

The analysis using the VAR system (2) relies on two basic concepts: Granger- and instantaneous-causality. Granger causality involves the assessment of the null hypotheses: $A_{c12,i} = 0$ for i = 1, 2, ..., p (which implies that *BBD* is not Granger caused by *GT*), and $A_{c21,i} = 0$ for i = 1, 2, ..., p (which implies that *GT* is not Granger caused by *BBD*).[14] In our context, the Granger causality from *BBD* to *GT* for a certain category *c* implies that past news-based shocks are related to present Google searches: past newspaper headlines lead people to increase their knowledge of the events in question even after the news shock has occurred, i.e. people still feel - after the shock - the need for further knowledge about *c*. In other words, the news-based measure of the uncertainty category *c* - informing people about what is happening now - leads to Google searches. Conversely, the Granger causality from *GT* to *BBD* for a certain category *c* implies that past Google searches inform journalists about their readers' persistent interest in past headlines regarding *c*, and thus newspapers tend to satisfy this continued need for knowledge about the evolution of *c*. In this second case, Google searches – signalling readers' interest in *c* – drives the news-based measure of the uncertainty category *c*.

Although we must be very cautious in interpreting the outcomes of Granger causality statistical tests in behavioural terms, for the sake of brevity we shall now label the Granger causality from *BBD* to *GT* as the "news-pooled" uncertainty, and the Granger causality from *GT* to *BBD* as the "query-driven" uncertainty. Having found the direction of Granger causality for each category (where such direction exists), the outcomes may be expanded by replacing, in the bivariate VAR, the Google Trends measure for *c* with its sub-groups and/or single search terms, thus utilizing the information concerning the most relevant terms within each sub-group.[15]

Given that VAR residuals are not orthogonal, i.e. the covariance matrix $\Sigma_{\varepsilon c}$ is usually not diagonal because these models are in reduced form, the presence of a significant instantaneous

[14] When the Johansen's test does not reject the null of reduced rank (i.e. that not all the variables in VAR are stationary), the Granger causality must be tested by following the Toda and Yamamoto (1995) approach. This is the way we followed for the sub-groups 20-21-22 of the Entitlement Programs category, and for the sub-groups 23-24 of the Trade Policy category; see Appendix A3.

[15] Details are in Table A2.1 summarising the results of the Bayesian aggregation approach in Appendix A2.

correlation between *BBD* and *GT* shocks ($E(\varepsilon_{c1t}\varepsilon'_{c2t}) \neq 0$) means that the two measures of uncertainty for *c* are coincident, i.e. they reflect the same timing.

3.2 – Empirical results by policy category

The main results produced by using VAR system (2) are summarised in Table 1.[16] The two columns and three rows of the table delimit six areas (cases) in which the 8 categories are allocated. Along the columns, we classify the categories as cases of high/low instantaneous correlation (degree of coincidence), depending on whether the value of such correlation is greater or less than 0.25 (a figure denoting its "high" 1% significance). Along the rows, we classify the 8 categories into three cases: the case of news-pooled uncertainty (i.e. when *BBD* Granger causes *GT*), the case of query-driven uncertainty (i.e. when *GT* Granger causes *BBD*), and finally the case of no-dynamically-related uncertainty (i.e. when Granger causality is not significant in either direction).

Table 1 here

The results shown in Table 1 can be summarized as follows. "Fiscal policy" and "Sovereign debt and currency crisis" are carefully monitored by people, as Internet searches increase as soon as shocks occurs, even though newspapers do not place much emphasis on them. Moreover, Internet activity and newspaper mentions also overlap significantly, since the instantaneous correlation coefficient is always high (in the 0.40-0.50 range). The same result is broadly found also in the case of "Healthcare", albeit with considerably lower significance and intensity.

If we further the VAR analysis by trying to establish what the main drivers (at the level of individual search term) of the detected query-driven uncertainty are, the results in Appendix A3 suggest the following answers. With regard to fiscal policy, the most reactive Google queries are "Debt ceiling"[17] and "Government deficits", while the "Sovereign debt and currency crisis" *GT* dynamics is mostly driven by the term "Sovereign debt", although "Currency devaluation" and "Euro crisis" also play a significant role. In this context, the "Health care" *GT* category leads the news-based shock mainly because of the one search term "Affordable Care Act".

The direction of the Granger causality is reversed in the case of "Monetary policy" and "Regulation", since it seems that people start looking for more information about the search terms in these categories only after newspapers have begun to mention them, i.e. news-based shocks drive Google searches. In this context, the term "Trade policy" behaves rather similarly, albeit with a considerably lower level of instantaneous correlation, which probably denotes the general public's

[16] The detailed outcomes are reported in Appendix A3.

[17] The analysis of the search term "debt ceiling" is studied further in Appendix 4.

limited attention to the news shocks in this category (the correlation coefficient here is about 0.10, against values of 0.55 and 0.31 recorded in the previous two cases, respectively).

Finally, "Entitlement programs" and "National security and war" do not show any Granger causality from press and media reports to Google searches or *viceversa*. However, they behave differently as far as the degree of simultaneity is concerned. In fact, for "Entitlement programs", the news and Google are strongly correlated (above 0.40), thus denoting a substantial overlap, in the same month, of press reports and internet searches; in the case of "National security and war", on the other hand, the instantaneous correlation is about half the previous figure, suggesting that the need for knowledge that feeds the Google searches is not related to newspaper headlines to any significant degree, probably because readers already know a lot about the matter in question.

BBD recently released data regarding news coverage of two specific search terms: "Government shutdown" and "Debt ceiling". This offered us the possibility for testing whether the comparison between news coverage (*BBD*) and Internet searches (*GT*) indicators at the category level is affected by problems deriving from aggregating possible heterogeneous search terms. Overall, in this context of highly disaggregated comparisons, some findings resulting from analyses conducted at the category level are confirmed. The first finding is the coincidence of news coverage and Internet searches in relation to many uncertainty shocks. The second one concerns journalists' attention to their reader's needs: the persistence of news coverage following large shocks, can indeed be interpreted not only as a consequence of journalists' individual tastes, but also as an attempt to feed their readers' desire for information with increasingly more fresh news.[18]

4. A state-of-the-art assessment of macroeconomic uncertainty measures

This section compares our *GT* index with what we called the "fantastic five" uncertainty indicators, because they are among the most popular, commented-on measures in the press and in the literature. Moreover, they are comparable with our *GT* index because they have also been publicly available on a monthly basis for a sufficiently long time to cover a large number of historical events.

After presenting the "fantastic five" (in Section 4.1), we analyze, comparatively with *GT*, their statistical features at both univariate (persistence, seasonality and variability over time) and multivariate (Granger causality and contemporaneous correlation) levels (in Sections 4.2, and 4.3 respectively).

[18] Details are reported in Appendix A4.

4.1 – Introducing alternative proxies: the "fantastic five"

We consider, as the "fantastic five": [19]

- two finance-based indexes, (*i*) the CBOE Volatility Index, VIX (see Chicago Board Options Exchange, 2009), and (*ii*) the corporate bond spreads, SPREAD, of Bachmann et al (2013);
- two forecast-based measures, (*iii*) the one-months ahead macro-uncertainty indexes, PREDICT1, of Jurado et al. (2015), and (i*v*) the measure of disagreement in business economists, FDISP, of Bachmann et al. (2013);
- one news-based policy index, (*v*) the Economic Policy Index, EPU, of Baker et al. (2015), in the specific form of the HNEWS. This index is coupled with our search-based policy index *GT*.

Although often considered as alternatives, these measures were introduced in the original under specific, different names. For example, SPREAD is defined in general terms as an "uncertainty measure", FDISP as a "business-level uncertainty" measure, PREDICT1 as a measure of "macroeconomic uncertainty", and the EPU as a measure of "policy-related economic uncertainty".

While proxy (*v*) was essentially commented on, together with *GT*, in Sections 2 and 3, some observations can be made with regard to the other four proxies. The (*i*) VIX is used in many empirical studies, for example in Bloom (2009). But one caveat that emerges from the use of the VIX to proxy uncertainty concerns its ability to represent macroeconomic uncertainty, since it is based on stock market information alone. According to Bekaert et al. (2013), the VIX does not only reflect uncertainty but can be broken down into uncertainty and risk-aversion.[20] Since risk aversion accounts for a sizeable part of the VIX,[21] even more caution should be taken when considering it as a proxy of macro-economic uncertainty. Of course, the same criticisms of the use of financial information to proxy whole-economy uncertainty features (which also embody non-financial information) also apply to the (*ii*) SPREAD proxy as well. Moreover, Gilchrist et al. (2014) use credit spreads to measure the degree of financial market friction - rather than uncertainty - and in this context their results suggest that credit spreads are an additional channel through which uncertainty fluctuations may spread to the real economy.[22]

---

[19] Details of their computation are given in Appendix A5.

[20] In particular, they show that the uncertainty component of VIX can be defined as the expected stock market variance estimated using a projection model which includes the lagged squared VIX and past realized volatility, whilst risk-aversion can be defined as the difference between the squared VIX and the estimated VIX uncertainty component.

[21] For this reason, the VIX is usually referred to as the "investor fear gauge" (see, for example, Whaley 2000).

[22] Gilchrist et al. (2014) assume that credit spreads reflect the endogenous effects of informational and contractual friction deriving from an (exogenous) uncertainty shock, which they measure using high-frequency firm-level stock market data. However, their Figure 4 suggests that the latter measure is highly countercyclical and moves in tandem with credit spreads. Nodari (2014) provides empirical evidence of the direct effects of financial regulation policy uncertainty on credit spreads, and subsequently on the real economy.

Although independent of any single observable economic indicator or event, the (*iii*) PREDICT1 proxy is a computationally intensive black box that is not directly linked to the uncertainty perceived by the general public. The main caveat associated with the (*iv*) FDISP proxy is whether or not the disagreement of manufacturing c-levels – i.e. high-ranking executive members of organizations (a clearly defined, albeit narrow, category of economic actors) – is somehow reflected in US macroeconomic uncertainty. Furthermore, it is worth remembering the results in Lahiri and Sheng (2010). By breaking down the individual forecast error into a component related to shocks common to all forecasters over the forecast period, and another component concerning forecaster-specific idiosyncratic errors, Lahiri and Sheng (2010) show that the ex ante forecast uncertainty can be expressed as the sum of the expected volatility of future (common) aggregate shocks over the forecast horizon, and the observed disagreement among forecasters. In this context, large common shocks occurring in periods of economic instability or long forecast horizons, increase the difference between disagreement and uncertainty.

4.2 – Univariate analysis of the "fantastic five" and GT

Figure 1 depicts the "fantastic five" and *GT*: the two plots in the first row show the finance-based measures (VIX and SPREAD), those in the second row show the forecast-based measures (PREDICT1 and FDISP), and the news- and search-based measures (HNEWS and *GT*) are displayed in the third row. The first five alternative measures of uncertainty cover the common sample 1970m1-2014m12 (T = 540 months, of which 83 of NBER downturns shown by the shaded areas), while our Google Trends indicator is available since 2004m1.

Figure 1 here

Besides different scales (summary statistics are in Table 2 below), the eyeball comparison of the uncertainty measures highlights quite different patterns of variability that are not very closely related to the nature of the indicators. In fact, some series (SPREAD and PREDICT1) clearly spike in recession periods only, while others (FDISP and HNEWS) seem more affected by random fluctuations over time. The VIX series is affected by certain spikes even outside downturns, when financial crises occur (e.g. the Black Monday of October 19th 1987).

The upper panel in Table 2 reports the unconditional statistics for the uncertainty measures, revealing the larger variability of VIX, SPREAD and HNEWS. Apart from FDISP and *GT* (which seem more Gaussian), all the other series show positive skewness and excess of kurtosis, indicating tails on the right side of the distribution that are longer than those on the left side (i.e. the bulk of the density and the medians lie to the left of the means). This is the effect of positive spikes during recession phases, which results in counter-cyclical uncertainty, also supported by means and

variances that are often higher during downturns. These shifts are less pronounced for forecats-based and news/search-based indicators, while the opposite holds for financial and macroeconomic uncertainty measures. Individual answers, newspaper information and Internet search fluctuations are all more weakly associated with the macroeconomic cycle, and thus their shocks may disentangle negative social and political events (see Bloom, 2014) from the mix of other occurrences that could over-represent macroeconomic conditions.

Table 2 here

Given that persistence is a relevant feature of theories where uncertainty is a driving force of long-lasting economic downturns (Scaal, 2011), the lower panel of Table 2 provides more in-depth analysis of the persistency properties of the various series. The results of the unit root tests (p-values) over the sample 1970m1-2014m12 suggest that the fantastic five are stationary, as the null hypothesis is always rejected at 5% (with the exception of HNEWS where it is only rejected at 10%). However, their persistence is different: after the occurrence of an uncertainty shock, the period of time it takes to halve the gap between actual data and the steady state, ranges between 5-7 months for VIX, FDISP and HNEWS, and stands at about one year for SPREAD, and more than three years for PREDICT1 (the slowest to adjust).[23] The shorter time span available for *GT* explains the statistical non-rejection of the unit-root null hypothesis, while the point estimate of its persistence suggests the behaviour of a stationary series with short-lived shocks.

These results are confirmed by estimating ARFIMA models for the uncertainty level *y*:

$$\rho(L)(1-L)^{d}(y_t-\mu)=\vartheta(L)\varepsilon_t \qquad (3)$$

Where *d* is the fractional-integration parameter, $\rho(L)$ is the AR polynomial in the lag operator *L*, $\vartheta(L)$ is the MA polynomial, and ε is the *iid* innovation term.[24] Since Granger and Joyeux (1980), it has been widely acknowledged that the ARFIMA model provides a parsimonious parameterization of long memory processes that nests the ARMA model and, by allowing for fractional degrees of integration through *d* parameter estimates, it also generalises the ARIMA model with integer degrees of integration. The ARFIMA model is covariance stationary for

[23] Note that in Jurado et al. (2015), the point estimate of the half-life of PREDICT1 after an uncertainty shock is 53 months, whilst the half-life for the VXO is 4 months. These estimates are only slightly different from ours because they are based on an AR(1) model and refer to the period 1960m7–2011m12.

[24] All the univariate models in this section account for potential seasonal effects using dummy variables rather than by imposing a filter on data, because seasonal filters (e.g. the Census X13) might change the pattern and timing of the original measures. A battery of tests for stable and moving seasonality (see section D8.A in the X-13-ARIMA-SEATS output, see Lothian and Morry, 1978) support our choice in this context. Significant seasonal effects always emerge, with the exception of the FDISP. Note that Bachmann et al. (2013) acknowledge that the FDISP is based on seasonally adjusted data, while in all the other cases the issue is at best only partially discussed. Overall, seasonal dummies' significance suggests that in modeling uncertainty, the seasonality required to be handled carefully.

$|d| < 1/2$. The ARFIMA outcomes in Table 2 support the unit-root tests' findings: all series behave in a stationary manner (i.e. varying towards a time-invariant steady state) at different speeds.

The similar persistence estimated by the ARMA/unit-root and ARFIMA approaches is remarkable. However, the null hypothesis of identically distributed uncertainty shocks is always rejected regardless of the modelling context (ARMA or ARFIMA), due to the presence of residuals' autoregressive conditional heteroskedasticity, mainly deriving from shocks increasing in size during downturns. Therefore, a data congruent representation of uncertainty measures needs to account for this feature.

We now use the result shown in Tsay (1987), according to which the GARCH effects in the innovations can also be explained by uncertainty models where the parameters are allowed to change and by *iid* innovation shocks. Given the close similarity in the dynamics of the ARMA-ARFIMA shown by the results in Table 3, the common starting point of the two approaches is that of representing the stationary process driving the linear dynamics of the conditional-mean of uncertainty in the context of the MA(∞) Wold's decomposition of the short-term ARMA fluctuations of the ARFIMA model (3), by assuming that $d = 0$.

The first approach, exploits the class of ARMA models with GARCH innovations as a parsimonious and flexible way of representing the dynamics of conditional mean and variance (see Diebold and Lopez, 1995), and of preventing the insurgence of residuals' heteroscedasticity:

$$y_t = \mu + B(L)\varepsilon_t \tag{4}$$

where μ is the unconditional mean of y_t and $B(L) = 1 + \sum_{i=1}^{\infty} b_i L^i$ is a lag L polynomial obtained from $\rho^{-1}(L)\vartheta(L)$. The shocks ε_t are serially uncorrelated uncertainty innovations with time-varying heteroskedasticity h_t conditional to the information set in *t-1*, Ω_{t-1}:

$$\varepsilon_t | \Omega_{t-1} \sim N(0,\ h_t) \tag{5}$$

where $h_t = \omega + C(L)\varepsilon_t^2$, and $C(L) = \sum_{i=1}^{\infty} c_i L^i$ is a lag polynomial. If the two infinite-ordered autoregressive lag operator polynomials *B(L)* and *C(L)* are approximated by ratios of low-ordered lag operator polynomials, we obtain the classical ARMA model with GARCH errors (see e.g. Bollerslev et al., 1994).

The innovations ε_t of the ARMA/GARCH models (4)-(5) affect the dynamics of both the conditional mean and the variance of uncertainty, to represent the pattern of highly or slightly

uncertain periods, that is, when significant changes in uncertainty are likely to be followed by other large changes, and when small changes are likely to be followed by other small changes, respectively. We choose an ARMA(2, 3) model for the conditional mean and a GARCH(1, 2) model for the conditional variance; although simple, these models can detect, within a suitable non-linear framework, clusters of time where the uncertainty shocks are particularly large, and they nest all the uncertainty measures:

$$y_t = \gamma_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \vartheta_1 \varepsilon_{t-1} + \vartheta_2 \varepsilon_{t-2} + \vartheta_3 \varepsilon_{t-3} + \varepsilon_t \quad (6)$$

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \beta_1 h_{t-1} \quad (7)$$

The second approach starts once again from the MA(∞) Wold's decomposition of the ARMA fluctuations of uncertainty in equation (4); however, here we now assume that the unconditional mean of uncertainty can vary across m+1 different regimes (denoted with r) of T_r observations:

$$y_t = \mu_r + B(L)\varepsilon_t \quad (8)$$

where μ_r is the unconditional mean of y_t in the r^{th} regime (with r = 0, 1, 2, ..., m), $B(L)$ is a polynomial of infinite order as in equation (4), and ε_t are zero-mean, serially uncorrelated and homoskedastic (*iid*) uncertainty innovations. In order to estimate the unknown number of break dates, we follow a procedure (see Bai and Perron, 1998) which starts from m breaks under the null (obtained by a global minimization of the sum of squared residuals) and tests it against the alternative of m+1 breaks, for m = 0, 1, 2, ..., $m_{\max}$ (here we set the maximum number of breaks over the 2004m1-2015m12 period at five).

Table 3 reports the Maximum Likelihood estimates of the ARMA/GARCH representation (6)-(7) of the "fantastic five" and *GT*.

Table 3 here

As regards the persistence of the conditional mean of the uncertainty shocks, the OLS estimates of π in Table 2 are confirmed by the estimates of $\phi_1 + \phi_2 - 1$ of the ARMA/GARCH model in Table 3, where the conditional mean specification is always characterized by large autoregressive parameter estimates with cyclical effects that are particularly evident for PREDICT1 and HNEWS, while only HNEWS clearly shows a first-order moving average estimate, mirrored by the extremely long lag order (k = 15) of the unit root test equation in Table 2.

With regard to the conditional variance, GARCH components are always strongly significant, with the financial SPREAD revealing the highest persistence in volatility (thus the

Integrated GARCH is the best representation). At the other extreme, the lowest persistence of the volatility of the financial VIX and *GT* shocks is coupled with the low persistence of their first-moments.

The patterns of the estimated conditional variances h_t of the different ARMA-GARCH models are plotted in the upper graph of Figure 2, and suggest that the clusters of time-varying volatility (i.e. the periods in which uncertainty is prone to larger/smaller shocks) are not much different across indicators. The variability of uncertainty shocks is more pronounced at the beginning of the sample (i.e. from the 1970s to the first half of the 1980s), before weakening during the phase of great moderation (that began in around 1984), with the exception of the financial blip in 1987m11. At the time of the Great Recession of 2008, variability rose again and continued to maintain high levels until the end of the sample (most obviously in the case of the HNEWS and *GT* conditional variance estimates).

Figure 2 here

The clustering volatility of uncertainty is mirrored in the pattern of $\hat{\varepsilon}_t$ (i.e. the ARMA-GARCH models' residuals) reported in the lower graph of Figure 2. These $\hat{\varepsilon}_t$ can be seen as our best estimates of the unpredictable macroeconomic uncertainty innovations.[25] Their dispersion broadly corresponds to the phases before/during/after the Great Moderation, thus lending support to the belief that the size of uncertainty innovations and the variability of the macroeconomic cycle are somehow related. The prevalence of positive shocks during recession phases confirms the counter-cyclicality of uncertainty.

Not all peaks of the different uncertainty shocks overlap, since some measures are affected by more shocks than others are; this feature is noted in Jurado et al. (2015) as well, who report fewer extreme values in PREDICT1 than in VIX. We further investigate the issue in the lower part of Table 3, by counting the number of unpredictable adverse innovations $\hat{\varepsilon}_t$ above 3 alternative thresholds: one, one and a half, and three times their time-varying standard deviations $\sqrt{h_t}$. Peaks in uncertainty are considerably fewer for forecast-based indicators (PREDICT1 and FDISP) than for finance-based indicators (VIX and SPREAD). HNEWS and *GT* peak quite frequently, and in this they are similar to the finance-based indicators because news and Internet searches are related to important financial shocks even when such occur outside of large macroeconomic downturns. Our

[25]The results presented in the following sections using $\hat{\varepsilon}_t$ are qualitatively similar to those obtained using the standardized residuals, i.e. $\hat{\varepsilon}_t$ scaled with their corresponding time-varying standard deviations.

results confirm the findings of Jurado et al. (2015), and suggest a possible dichotomy between finance-based and forecast-based indicators.

Table 4 reports, for the six uncertainty measures listed down the columns, the F-statistics of the largest significant number of breaks (i.e. the number of breaks for which the null hypothesis of *m-1* breaks is rejected against the alternative of *m* breaks), together with the corresponding estimates of *m*, and with Bai and Perron (2003)'s 5% critical values.

Table 4 here

The OLS estimates of model (8) conditional on its own break dates deliver residuals $\hat{\varepsilon}_t$ that never reject the null hypotheses of not serially correlated and homoskedastic errors. The evidence of homoskedastic errors supports the ARMA-with-breaks model (8) as a viable alternative to the ARMA model (6) with GARCH errors (7) to prevent the insurgence of residuals' autoregressive conditional heteroscedasticity. In other words, we are able to identify the high/low uncertainty regimes by allowing for shifts in uncertainty states, rather than by assuming GARCH variability of the uncertainty shocks over time.

4.3 – Multivariate results across the "fantastic five" and GT

Previous outcomes can be investigated further by extending the analysis from the univariate to the multivariate VAR context. Below we exploit two different VAR models, depending on the use of ARMA/GARCH residuals or of the uncertainty measures in a context with multiple breaks.

According to the first univariate approach, we assess the simultaneity and the dynamic causation across shocks $\hat{\varepsilon}_t$ (i.e. the ARMA/GARCH residuals) of different uncertainty indicators. First, the time series of $\hat{\varepsilon}_t$ are modelled in a VAR system of the second order.[26] Secondly, the VAR(2) parameter estimates are tested for Granger causality. Results are reported in Table 5.

Table 5 here

Given that data availability changes if we include *GT* in the VAR or otherwise, Table 5 is divided into two parts: in panel A the results from the largest sample period (1970m3-2014m12) are reported for our "fantastic five" benchmark measures, while in panel B the results refer to an extended VAR where *GT* also is included, and thus the sample period is shorter (2004m4-2014m12). The upper section of each of the two panels shows the p-values of the Granger causality

[26] As suggested by both the AIC criterion and by the non-serially correlated VAR residuals. These unpredictable innovations $\hat{\varepsilon}_t$ modelled by a VAR model, can be interpreted as uncertainty measures following pre-whitening to remove their predictable parts such as inertia.

test, where each column refers to a VAR equation, and each row refers to a block of explanatory lags. In the lower part of each panel, the sample correlation matrix of VAR residuals is reported in order to assess the degree of simultaneity between different indicator innovations.

In panel A, one- and two-month VIX innovation lags drive the dynamics of SPREAD innovations, while past changes in SPREAD do not lead VIX uncertainty: financial returns variability reacts more quickly to spread fluctuations within the same month, as shown by the VAR residuals correlation matrix.[27] The dynamic causation of VIX extends not only to the innovations in SPREAD, but also to those in FDISP and HNEWS: an uncertainty shock in the stock market, measured by VIX innovations, also leads forecasters' disagreement and (to a lesser extent) HNEWS. SPREAD innovations, on the other hand, simply lead FDISP shocks and no other uncertainty innovations.

Along the FDISP column, FDISP innovations are led by all the other shocks in the VAR. This striking causation could be accounted for by the low degree of information updating by survey respondents: despite being high level executives, they are not professional forecasters, and as such their disagreement shocks can be caused by past shocks to other uncertainty measures bearing more recent information about the state of the forecasting environment. This interpretation, in line with Lahiri and Sheng (2010), is supported by the limited evidence of Granger causality along the column of PREDICT1 obtained using a large, updated dataset of economic indicators, and which as such is more difficult to be led than FDISP, while it anticipates all other uncertainty shocks. The Granger causality from FDISP to PREDICT1 is by and large apparent: being both indicators based on forecasts, past FDISP data - computed using six-month forecasts - overlap (rather than genuinely driving) the one-month horizon of PREDICT1. Overall, FDISP shocks do not cause any other uncertainty shock because FDISP forecasters target a specific area of the US, the economic indicators of which may not be perfectly in line with US-wide financial and macroeconomic indicators in the long term.

The HNEWS shocks significantly Granger cause PREDICT1 because, despite being based on a large dataset of predictors, PREDICT1 may be lacking in news-based information; moreover, HNEWS innovations fuel shocks to the variability of forecasts (FDISP) as well. On the other hand, innovations to HNEWS cannot lead financial uncertainty shocks since financial markets instantaneously react to policy news, as shown by the significant contemporaneous correlation

[27] In particular, SPREAD shocks are simultaneously related to those in VIX by a significant correlation of about ¼: financial market returns variability is positively and immediately associated to the interest rates differential shocks.

between HNEWS and VIX(about 0.3).[28] The HNEWS column reveals that it is almost impossible to lead HNEWS uncertainty shocks (apart from in the PREDICT1 case). This substantiates the idea that journalists are quite responsive to economic news: the joint Granger causality test p-value of HNEWS is slightly below 5%, while it is well above those of the other measures, which on the contrary are strongly Granger caused (1% significant). High news responsiveness is further corroborated by the significant contemporaneous correlation between HNEWS idiosyncratic uncertainty shocks and almost all other sources of shock arising within the same month, as reported by the residuals' correlation matrix. Idiosyncratic SPREAD innovations represent the only case where the media are not very good at detecting them within the current month, thus suggesting that newspapers are not very interested in reporting on interest rate differentials, and this fact recalls the issue of journalists' "taste" when it comes to choosing what news to highlight or otherwise (see the discussion in Section 3).

Overall, two main results emerge from the sample period 1970m1-2014m12. Firstly, PREDICT1 shocks anticipate almost all the other measures, and they are hardly ever Granger-caused. The systematic utilization of large information sets increases PREDICT1 timeliness in detecting new shocks of different origin. Probably for this reason, and also because it is much smoother than other measures, PREDICT1 is not coincident with the other uncertainty measures (see also the next finding). Secondly, our results suggest that journalists are timely messengers, and consequently the news-based uncertainty measures bear a lot of coincident information coming from other sources of uncertainty.

In panel B of Table 5 the reduction in the time span increases the sample dominance of the Great Recession (virtually the only large-scale event in this temporal window), and consequently weakens the evidence of Granger causality between uncertainty shocks, in favour of an increase in their degree of simultaneity. In fact, over the 2004m4-2014m12 sample the Granger causality is still significant only for finance-based indicators, while it vanishes for forecast- news- and search-based ones. At the same time, there is an increase in the simultaneous correlation between shocks, from the levels estimated with the longer sample; owing to the dominance of financial facts in this short sample, the correlation between VIX and SPREAD shocks rises from 0.25 in panel A to about 0.60 in panel B.

Overall, the results of the comparison of different sample periods in which the same facts have different weights, emphasise the importance of the time span, and consequently of the

[28] This evidence of the ability of journalists to detect financial uncertainty contrasts with Bachmann et al.'s (2013, p. 240) hypothesis that "asset market variables (stock volatility and the corporate bond spread) pick up a kind of uncertainty that is not captured by survey-based and news-based uncertainty indices and that triggers "wait and see" dynamics".

size/number/nature of the shocks during that time span, to the shaping of the dynamic relationship between uncertainty measures. However, the two main findings regarding the larger time span are strong enough to survive over the shorter period.

Summing up, results are listed below.

(1) PREDICT1 is again the indicator that most frequently leads several other uncertainty measures, and it remains simultaneously related to HNEWS.

(2) Journalists continue to act as the messengers of many different uncertainty shocks, as HNEWS is still strongly simultaneously correlated with all other sources of uncertainty innovation (even though we must acknowledge that over the shorter span, a journalist's job was "made easier" by the advent of the Great Recession).

Even though the short time span calls for great care to be taken when interpreting results, the inclusion of *GT* in the VAR in panel B produces two additional findings of interest.

(3) Shocks in Internet search activity may lead shocks to the finance-based VIX and SPREAD indexes, and such evidence of *GT* Granger causality complements (rather than substituting) the one of PREDICT1. This fact, together with the weekly frequency with which virtually unrevised Google Trends are released, may open the way to the latter's use in tracking the volatility of stock returns in real time.

(4) The simultaneous correlation between HNEWS and *GT* is one of the strongest in panel B (about 0.30). Besides the advent of the Great Recession, such a coincident effect is due to the aggregation of different policy categories in overall indexes that can sweep away much of the leading and lagging information which – shown by category in Table 1 of Section 3 – is nevertheless significant despite being estimated over the same (short) sample.

Table 6 presents the results delivered by the second approach with parameters shifting over time. We follow a two-step procedure able to extend the analysis from the univariate ARMA-with-breaks model (8) to the multivariate VAR. Firstly, we estimate a VAR(2) model for the vector of six uncertainty measures (VIX, SPREAD, PREDICT1, FDISP, HNEWS and *GT*) conditional on all the break dates detected at univariate level[29], and test for the significance of the step-dummies' parameters measuring the break dates; at the same time, non-significant dates are eliminated. Secondly, the restricted VAR(2) - which is only conditional on the jointly significant break dates - is estimated and examined for Granger causality and the simultaneity of residuals.

Table 6 here

[29] At the univariate stage, we found 20 break dates measured by 19 step-dummies, because one of the estimated dates was the same in both the PREDICT1 and the *GT* equations.

The 9 surviving break dates in the restricted VAR are listed in the top of the first column of Table 6, while the subsequent columns show parameter estimates of the corresponding step dummies as ratios over the average uncertainty measures, in order to better understand their economic relevance and size.

Certain differences emerge compared with the univariate break outcomes. This is because within the multivariate context, some uncertainty measures can co-break (i.e. some breaks can be cleared across measures), while in other equations the breaks in significant explanatory variables' can induce further shifts. For example the financial-based uncertainty measures show, within the multivariate context, a greater number of significant breaks than those detected at the univariate stage (shown in the first row of Table 4), because significant explanatory lags in PREDICT1 and *GT* can result in further breaks in their equations. On the other hand, the FDISP equation shows fewer multivariate breaks than univariate ones since it probably co-breaks with significant explanatory lags of HNEWS and *GT*. The equations for the news-based and search-based measures show barely the same number of breaks as those present in the univariate models. Finally, the case of the PREDICT1 equation remains puzzling, probably due to the excess smoothness of its series in the 2004m1-2014m12 period, which may make it difficult for the breaking model to detect clear shift-points.

The bottom section of Table 6 reports both the p-values of the Granger causality test and the VAR's residual correlation matrix. These depict outcomes that are perfectly in line with those shown in Table 5, obtained from a VAR modelling the GARCH uncertainty innovations. The estimates of the simultaneous correlations in the breaking VAR shown in Table 6 are only slightly lower than those in Table 5, because part of the simultaneity between variables is captured by common breaking dates such as the general upward shift in 2008m9 that the restricted VAR identifies as the increase in uncertainty realized at the beginning of the Great Recession.

5. Uncertainty measures and macroeconomic dynamics

Findings in literature suggest that every time uncertainty is modelled within the macro-economic VAR context, it always displays a significant negative relationship with economic activity, as uncertainty shocks are broadly found to exert a negative impact on output and employment. Similarly, albeit by means of a different approach, Stock and Watson (2012, p. 81) find that "the shocks that produced the 2007-2009 recession were primarily associated with financial disruptions and heightened uncertainty".

However, this key finding is only robust in regard to the uncertainty impact in the short run, whereas in the long run different works have pointed to somewhat heterogeneous output responses.

For example, the results in Bloom (2009) sustain the over-shooting effect of a VIX uncertainty shock on real activity: following the shock, the economy suffers in the short term, but in the long run the initial level of output is surpassed. The evidence in Figure 6 of Bachmann et al. (2013) suggests that Bloom's over-shooting is more due to his use of finance-based indexes rather than to any genuine uncertainty effect.[30] The latter fact reinforces our caveats about the reliability of measures of macroeconomic uncertainty based solely on financial information, and suggests that researchers need to be careful when proxying uncertainty with these finance-based measures, as they may label certain transitory financial crises as uncertainty shocks.

Jurado et al. (2015) and Bachmann et al. (2013) instead utilize forecast-based measures. Their VAR models reveal that the dynamic response of output to uncertainty shocks sharply reduces the level of production with effects that persist well beyond the horizons considered in their exercises (i.e. more than 4-5 years after the shock).

Baker et al. (2015) model the economy by slightly reducing the number of variables in Bloom's (2009) VAR (from 8 to 5 macroeconomic variables, uncertainty included), and use their news-based economic policy uncertainty index. They report a negative dynamic response of manufacturing production to a shock. However, unlike in Jurado et al. (2015) and Bachmann et al. (2013), these output responses are significantly negative for only the first 15-18 months after the shock, before gradually declining to zero, i.e. without overshooting.

Overall, these results lead to three, strictly related questions: (1) are uncertainty shocks temporary or permanent? (2) Is the degree of persistence of these negative uncertainty effects on output related to the number of variables in the VAR and/or to the uncertainty measure adopted? (3) Is any role played by the time span over which the models were estimated?

The answer to question (1) must take account of the fact that unrestricted VAR models can only represent reduced form shocks, and thus they provide little or no guidance in regard to structural causality. However, as Baker et al. (2015, p. 19) emphasise, "VAR estimations are helpful for getting a sense of the potential magnitudes of policy uncertainty effects, and what other factors are correlated with this". Therefore, in this section we use VAR models to analyse the comparison of the dynamic responses of output to shocks of alternative forecast-based and news-based uncertainty measures.

Given that we will be utilizing data and sample periods similar to Jurado et al. (2015) and Bachmann et al. (2013), any difference between our results and theirs can be ascribed to the reference model, and this fact leads to question number (2) regarding the choice of VAR and

[30]Jurado et al. (2015) argue that Bloom's over-shooting is a data figment mainly due to his HP filtering uncertainty, since with raw data the over-shooting dynamics vanish.

uncertainty proxies to be used. The literature has answered this question by offering a wide range of options: from the small bivariate VAR models (with alternative measures of uncertainty and output) of, for example, Bachmann et al. (2013), to the VAR with 11 variables of Jurado et al. (2015) which mimics the macroeconomic model of Christiano et al. (1995). Bivariate VARs have the advantage of parsimony, but at the cost of potential biases due to the omission of relevant macroeconomic channels; while large VARs have the advantage of a satisfactory theoretical basis, albeit at the cost of inefficient estimates due to the traditional curse of dimensionality, which is unavoidable with large monthly VAR models.[31] Here, we choose the more parsimonious five-variable VAR of Baker et al. (2015).

The translation of the Baker et al. (2015) VAR in our system (2) notation leads to a five-variable VAR created as follows. Orthogonal shocks to identify the impulse-responses functions are recovered using the Cholesky decomposition, and therefore the order of the variables matters. As in Baker et al. (2015), our first variable, z_t, measures the uncertainty which in three different VARs is proxied by PREDICT1 (i.e. the one-month predictability of Jurado et al., 2015), by the log-levels of EPU (i.e. the economic policy uncertainty index of Baker et al., 2015), and by the log-levels of *GT*.[32] Our second VAR's component is the vector $x_t = (sp_t, ff_t, emp_t, ipman_t)'$, where: *sp* is the log-levels of the S&P 500 index; *ff* is the logs of one plus the federal funds rate; *emp* represents the log-levels of employment; and *ipman* is the log-levels of the manufacturing production index.[33]

With regard to the answer to question number (3), we have to consider that, *a priori*, the pattern of dynamic responses to uncertainty shocks could vary over time: breaking VAR parameters could embody different mixtures of heterogeneous periods. If this is true, the permanent effects in Jurado et al. (2015) and Bachmann et al. (2013) could differ from those in Baker et al. (2015) simply because in the first two cases the sample period is long (it begins in the 1960s), while in the third case it is considerably shorter (it begins in the mid-1980s), thus excluding all the noisy observations prior to the Great Moderation. Moreover, the sample period in Baker et al. (2015) is also permeated to a greater degree by Great Recession observations, when "large shocks were not

[31] This problem is further exacerbated by short span, as we want to use our *GT* uncertainty measure.

[32] The decision to take logs-levels of EPU and *GT* was made in order to reduce the weight of outliers, which we showed are mainly present in the news-based and search-based measures. The main reason for reducing the extent of larger shocks is that when they are present, the linear approach could break because of nonlinearities. In fact, it is quite unlikely that doubling the size of a given uncertainty shock consequently doubles the size of the response, as is assumed by linear VAR models. In regard to this point see the appendix in Knotek and Khan (2011), for example.

[33] Given the evidence of seasonality discussed above, for each uncertainty measure we seasonally adjusted the data with the Census X13 filter, when needed. The decision whether to seasonally adjust uncertainty or not was based on a number of tests that are reported in Table A6.1 of Appendix A6; outcomes there suggested the adjustment of the PREDICT1 and GT series, while all the other VAR variables were seasonally adjusted by the source.

simply feeding through the usual dynamics. [...]. The usual dynamics did not explain what was going on for several months around the peak of the crisis." (Sims, 2012, p. 143).

To assess the effect of the sample period, we estimate our 5-variable VARs over different time span: first, the whole time sample available for each indicator; next, the shorter common sample periods. With equal VAR settings, different results are due to the different time spans used to estimate the model, and to different uncertainty indicators.[34]

Figure 3 shows the dynamic responses over 48 months of the manufacturing production index to one standard deviation of the uncertainty innovations. In particular, the upper section of Figure 3 shows the plots from the VAR model measuring uncertainty with the seasonally adjusted PREDICT1[35] and estimated over three different sample periods: (1) the largest available (starting from 1960m7, with responses shown in black together with the two dashed lines showing 90% standard errors bands); (2) the reduced sample period available with EPU data (starting from 1985m1 with responses shown in blue); (3) the shortest sample period available with *GT* data (starting from 2004m1, with responses shown in red). In a similar way, the plots reported in the lower section of Figure 3 are obtained from the VAR where uncertainty is measured by log(EPU) and which is estimated over two sample periods: (1) the largest available with EPU data (starting from 1985m1, with responses shown in blue together with the two dashed lines showing 90% standard errors bands); (2) the shortest sample period available with *GT* data (starting from 2004m1, with responses shown in red). No EPU data are available for the period prior to 1985.

Figure 3 here

Shocks to PREDICT1 progressively reduce the output log-levels, with effects persisting over the full 48-month horizon. In the long run, one standard deviation of PREDICT1 innovations (i.e. an uncertainty shock corresponding to about the 8th decile of the PREDICT1innovations' distribution) entails a loss of about 1% in the manufacturing production level. This outcome is qualitatively the same as those reported in Figures 6 and 7 of Jurado et al. (2015), and in Figure 6 of Bachmann et al. (2013), despite the fact that we use a smaller VAR.[36] Therefore, regardless of the VAR settings, the use of forecast-based indicators qualifies the uncertainty shocks as significant and persistent determinants of output fluctuations. In particular, the negative response of output first monotonically increases up to 24-25 months after the shock, and then stabilizes.

[34] For each sub-sample, the VAR lag length *p* is first set using the AIC criterion, and the corresponding residuals are tested for white noise errors (otherwise *p* is slightly altered until the target of white noise residuals is met). As reported in Appendix A5 (Table A5.1), the optimal length slightly changes across samples.

[35] Results are substantially unchanged even when using the non seasonally adjusted Index.

[36] Although output response patterns are very similar, the standard error bands here are smaller because of the greater parsimony of our VAR.

Shocks to EPU (blue line in the lower plots of Figure 3) initially reduce output sharply during the first 12 months after the shock, then this is followed by a mild recovery towards zero. In the long run, one standard deviation of log(EPU) innovations (an uncertainty shock corresponding to about the 8th decile of its innovations' distribution) entails a loss of about 0.1% in the manufacturing production level. This outcome is qualitatively the same as that reported in Figure 8 of Baker et al. (2015), despite the fact that we use a measure of impulse based on the estimated innovations' variance, unlike them who base their measure on the change in average uncertainty levels before/after the Great Recession. Very similar results are also obtained over the 1985m1-2011m12 sample period by Bachmann et al. (2013, Figure 6), in a bivariate system with log manufacturing production and GOOGLE, a previous version of the economic policy uncertainty measure of Baker et al. (2015) based on Google News rather than on newspapers. Unlike in cases of forecast-based uncertainty, the key finding here seems another: the use of news-based indicators qualifies the uncertainty shocks as significant determinants of output fluctuations only in the short run. In particular, the dynamic response of output sharply decreases for a period of one year after the shock, followed by a tendency towards recovery.

Both cases above reject the long-run overshoot effect in Bloom (2009), and lead to the following generalization: each uncertainty measure has its own dynamic effects on output, as finance-based uncertainty can induce overshooting effects, forecast-based uncertainty can induce very persistent effects, while news-based uncertainty can only induce temporary effects.

However, this view is highly disputable if we tackle the issue regarding the stability of VAR estimates (and hence dynamic responses) over different sample periods. In fact, the upper plots of Figure 3 show that the dynamic response (in blue) of output to a PREDICT1 shock in the estimates over a sample starting in 1985 (the same sample as that for the VAR with EPU uncertainty), is very similar to that shown in blue in the lower plots in Figure 3: first a sharper decline than that with the longer sample (in black), followed by a recovery. This additional impulse response raises the question of whether the output dynamics is driven more by the sample period than by alternative uncertainty measures. The latter possibility is also supported by the findings of Nodari (2014), where impulses coming from *only financial* news-based uncertainty shocks induce an output response perfectly in line with the one we obtain by the use of EPU uncertainty.

Furthermore, we estimated both our VAR (the upper part with PREDICT1 and the lower part with log(EPU)) over the shortest 2004-2014 sample corresponding to *GT* uncertainty data availability. Although data limitation requires great care to be taken when interpreting the results, we note that again the change in sample period affects the dynamics of output responses much more than the use of different uncertainty indicators does. In fact, both the output responses shown in red

further emphasise the sharpness of the short-run effects of uncertainty shocks (regardless of the uncertainty measure), while in the longer run the output always recovers. Although not significant, the dynamics of output responses in both VARs estimated over the shorter span (permeated by large financial shocks) are rather similar to the long-run overshoot found in Bloom (2009) deriving from estimates formulated over a longer sample period and using a finance-based measure of uncertainty.

As a consequence of this additional evidence from the same VAR over alternative sample periods, the previous statement regarding the specificity of output responses in relation to different uncertainty measures is completely reversed. In fact, according to this new view, the historical events in the models' estimation sample, and the corresponding parameter shifts, largely account for the different output responses, more so than the use of alternative uncertainty measures (further evidence of parameter instability in a different context is reported in Rossi, 2006).

On the one hand, longer spans (including the 1960-1970s period, i.e. the era of large, noisy real shocks) reveal sustained long-term effects of uncertainty (either forecast-based or news-based). On the other hand, shorter spans (almost entirely centred around the era of the large, noisy financial shocks that occurred during the Great Recession) tend to reveal stronger short-run effects followed by recovery. In general, the intensity of both the short-run output slowdown and its longer-run recovery, vanishes the more real shocks occur over the sample period, or alternatively the intensity increases the more the sample period is centred on the financial shocks of the Great Recession.

Figure 4 further corroborates this new view: the same sample period (starting in 2004) is utilized by three VARs that only differ in terms of their respective measures of uncertainty (PREDICT1, EPU, *GT*). The results are broadly the same: uncertainty significantly curbs output in the short run, while in the longer run recovery prevails, although the overshooting effects are always clearly not significant.

Importantly, when considering the dynamic relationship with economic activity, the Google Trends index and the Economic Policy Uncertainty index behave very similarly in spite of certain differences in their respective time series.

Figure 4 here

Overall, the evidence in this section supports the idea that the shocks that occurred over the sample period imply different output recovery paths simply because different periods of time embody shocks of different kinds.

6. Concluding remarks

Macroeconomic uncertainty is made of many and unobservable components. As such, it might not be quantitatively assessed by any specific measure which qualitatively relies more on some of these components. The heterogeneous nature of uncertainty emerges from the literature where alternative indicators behave differently regarding time patterns, information contents and macroeconomic effects. We could conclude that measuring uncertainty is a very "uncertain" activity.

Under this upshot, our paper has two main aims. The first one is to introduce Internet searches as a tool to make a new uncertainty indicator able to capture the general public's need to gather more information when they feel uncertain. To do so, we use Google Trends to measure the intensity of such searches and to obtain a new uncertainty index, *GT*. The second aim is to investigate the underlying reasons for the differing results obtained in the literature with regard to the interaction between uncertainty and economic activity, and to explore comparatively whether they can be ascribed to the specific series of uncertainty that are exploited.

The information content of six uncertainty measures (including our own) is compared in order to assess whether some shocks related to some indicator come first than others or they simply move together. The joint analysis of news-based uncertainty measures and our brand new *GT* allows for understanding that journalists not only meet, and in a certain sense "drive", the general public’s interest (and uncertainty), but they also pay considerable attention to the interests of their readers. Indeed, different categories of economic policy uncertainty entail alternative dynamic relationships of the nexus between newspapers headlines and Internet activity. Hence, the different attention paid by general public to different categories of economic policy issues explains the tendency of *GT* to either lead or lag the newspapers' uncertainty. This finding, together with the evidence about the features of other indicators, suggests two additional and related points. First, although the process of data collection about the intensity of Internet searches is not subject to the same care and transparency as that usually devoted in the traditional sample surveys, Google Trends series deliver patterns and tell histories that, at least in the context of uncertainty measurement, are consistent with those of their counterpart indicators. Second, and in general, measures of intensity of Internet activity can represent a genuinely novel source of information, a sort of open-source survey which at any time can deliver useful information about general public moods, provided that the "appropriate" set of queries is posed.

Therefore, Internet activity is definitely a useful complement to the other measures, which can provide original and early updated information about uncertainty. As such, many of its uses are still unexplored, and could support a number of future advances in the literature.

The comparative analysis suggests that the effect of uncertainty on the economic cycle is strongly related to the nature of the shocks embodied by the indicator used. For example, financial uncertainty (e.g. measured by the VIX) is prevalently driven by short-lived financial shocks which, in principle, should not entail very persistent responses. Forecast-based uncertainty measures reflect all those shocks that make the art of forecasting difficult; during the 1970s there were plenty of real shocks: these shocks were unpredictable, and at the same time had a long-lasting impact on output. News-based and search-based uncertainty measures are driven by all kinds of shock, provided that such shocks are deemed to be topical by journalists (news) and/or by newspaper readers (queries); as such, they record a wide range of historical events occurred over the sample period in a simple way.

For the first time our comparative analysis points to evident instabilities in the relationship between uncertainty shocks and macroeconomic dynamics, although our VAR models are in line with those used by other papers. In our parameters' shift context, the shape of the output responses to an uncertainty shock can change (and actually does, as we showed), depending on the events occurred in the sample period. Therefore, a further step forward in this field would be to try and shed some light on the macroeconomic nature of these *miscellanea* of uncertainty shocks, even though finding exogenous variation in uncertainty in order to identify them is a real challenge, as Stock and Watson (2012) results, and the criticisms thereof, suggest.

References

Alexopoulos, M. and J. Cohen (2009), "Uncertain times, uncertain measures", *Working Paper, Department of Economics, University of Toronto*, No. 352.

Bachmann, R, S. Elstner and E. R. Sims (2013), "Uncertainty and the economic activity: evidence from business survey data", *American Economic Journal: Macroeconomics*, Vol. 5, No. 2, pp. 217-249.

Bai, J. and P. Perron (1998), "Estimating and testing linear models with multiple structural changes", *Econometrica*, Vol. 66, pp. 47-78.

Bai, J. and P. Perron (2003), "Critical values for multiple structural change tests", *Econometrics Journal*, Vol. 18, pp. 1-22.

Baker, S. R., N. Bloom and S. J. Davis (2015), "Measuring economic policy uncertainty", *working paper downloadable at*: http://www.policyuncertainty.com/index.html

BBVA (2012), "Uncertainty: models and impacts", *Economic Watch* https://www.bbvaresearch.com/KETD/fbin/mult/120321_EconomicWatchEEUU_155_tcm348-297922.pdf

Bekaert, G., M. Hoerova and M. Lo Duca (2013), "Risk, uncertainty and monetary policy", *Journal of Monetary Economics*, Vol. 60, No. 7, pp. 771-788.

Bloom, N. (2009), "The Impact of uncertainty shocks", *Econometrica*, Vol. 77, No. 3, pp. 623-685.

Bloom, N. (2014), "Fluctuations in uncertainty", *Journal of Economic Perspectives*, Vol. 28, No. 2, pp. 153-176.

Bollerslev, T., R. F. Engle and D. B. Nelson (1994), "ARCH Models" in R. F. Engle and D. L. McFadden (eds.), *Handbook of Econometrics*, Vol. 4, Ch. 49, Elsevier.

Chicago Board Options Exchange, (2009), "The CBOE Volatility Index- VIX", *CBOE White Paper*.

Choi, H. and H. Varian (2012), "Predicting the present with Google Trends", *Economic Record*, Vol. 88, No. 1, pp. 2– 9.

Christiano, L. J., M. Eichenbaum and C. L. Evans (2005), "Nominal rigidities and the dynamic effects of a shock to monetary policy", *Journal of Political Economy*, Vol. 113, No. 1, pp. 1-45.

Da, Z., J. Engelberg and P. Gao (2011), "In search of attention", *The Journal of Finance*, Vol. 66, pp 1461-1499.

D'Amuri, F. and J. Marcucci (2012), "The predictive power of Google searches in forecasting unemployment", *Bank of Italy Working paper* (Temi di Discussione), No. 891.

De Luca, G. and J. R. Magnus (2011), "Bayesian model averaging and weighted-average least squares: equivalence, stability, and numerical issues", *The Stata Journal*, Vol. 11, No. 4, pp. 518-544.

Der Kiureghian, A and O. Ditlevsen (2009), "Aleatory or epistemic? Does it matter?", *Structural Safety*, Vol. 31, No. 2, pp. 105-112.

Dickey, D. A. and W. A. Fuller (1979), "Distribution of the estimators for autoregressive time series with a unit root", *Journal of the American Statistical Association*, Vol. 74, No. 366, pp. 427-431.

Diebold, F. X. and J. A. Lopez (1995), "Modeling volatility dynamics", in K. D. Hoover (ed.), *Macroeconometrics: Developments, Tensions and Prospects*, Ch. 11, Springer.

Dzielinski, M. (2012), "Measuring economic uncertainty and its impact on the stock market", *Finance Research Letters*, Vol. 9, pp. 167-175.

Eberth, J. M., K. N. Kline, D. A. Moskowitz, J. R. Montealegre and M. E. Scheurer (2014), "The role of media and the internet on vaccine adverse event reporting: a case study of human papillomavirus vaccination", *Journal of Adolescent Health*, Vol. 54, pp. 289-295.

Efron, B., T. Hastie, I. Johnstone and R. Tibshirani (2004), "Least angle regression", *The Annals of Statistics*, Vol. 32, No. 2, pp. 407-499.

Gilchrist, S., J. W. Sim and E. Zakrajsek (2014), "Uncertainty, financial frictions and investment dynamics", *NBER Working Papers Series*, No. 20038.

Granger, C. W. J. and R. Joyeux (1980), "An introduction to long memory time series models and fractional differencing", *Journal of Time Series Analysis*, Vol. 1, pp. 15-29.

Hacamo, I. and T. Reyes (2012), "Negativity bias in attention allocation: retail investors' reaction to stock returns",http://finance.uc.cl/docs/conferences/6th/Negativity%20Bias%20in%20Attention%20Allocation%20-%20November%202012.pdf

Harmanec, D. (1999), "Measures of uncertainty and information", *Documentation Section, Society for Imprecise Probability Theory and Applications*, http://www.sipta.org/

Ilut, C. L. abd M. Schneider (2014), "Ambiguous business cycles", *American Economic Review*, Vol. 104, No. 8, pp. 2368-2399.

Johansen, S. (1995), *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press.

Julio, B. and Y. Yook (2012), "Political Uncertainty and Corporate Investment Cycles", *The Journal of Finance*, Vol. 67, pp. 45–83.

Jurado, K., S.C. Ludvingson and S. Ng (2015), "Measuring uncertainty", *American Economic Review*, Vol. 105, No. 3, pp. 1177-1216.

Knotek, E. S. and S. Khan (2011), "How do households respond to uncertainty shocks?", *Economic Review of the Federal Reserve Bank of Kansas City*, Second Quarter, pp. 5-34.

Konig, J and R. Ohr (2012), "Different efforts in European economic integration: implications of the EU index", *Journal of Common Market Studies*, Vol. 51, No. 6, pp. 1074-1090.

Lahiri, K. and X. Sheng (2010), "Measuring forecast uncertainty by disagreement: The missing link", *Journal of Applied Econometrics*, Vol.25, pp. 514-538.

Leamer, E. E. (1978), *Specification Search: Ad Hoc Inference with Nonexperimental Data*, Wiley.

Lemieux, J. and R. A. Peterson (2011), "Purchase deadline as a moderator of the effects of price uncertainty on search duration", *Journal of Economic Psychology*, Vol. 32, pp. 33-44.

Lothian, J. and M. Morry (1978), "A test for the presence of identifiable seasonality when using the X-11-ARIMA program", *Research papers*, Statistics Canada.

Magnus, J. R., O. Powell and P. Prufer (2010), "A comparison of two model averaging techniques with an application to growth empirics", *Journal of Econometrics*, Vol. 154, pp. 139-153.

Mander (2006), "LARS: Stata module to perform least angle regression", *Statistical Software Components S456860*, Boston College Department of Economics, revised 22 July 2014.

Manski, C. F. (2015), "Communicating uncertainty in official economic statistics: an appraisal fifty years after Morgenstern", *Journal of Economic Literature*, Vo. 53, No. 3, pp. 631-653.

Ng, S. and P. Perron (2001), "Lag length selection anf the construction of unit root tests with good size and power", *Econometrica*, Vol. 69, No. 6, pp. 1519-1554.

Nodari, G. (2014), "Financial regulation policy uncertainty and credit spreads in the US", *Journal of Macroeconomics*, Vol. 41, pp. 122-132.

Ramey, V. A. (2009), "Identifying government shocks: it's all in the timing", *NBER Working Paper Series*, No. 15464.

Reis, R. (2006), "Inattentive producers", *Review of Economic Studies*, Vol. 73, pp. 793-821.

Rich, R and J. Tracy (2010), "The relationships among expected inflation, disagreement, and uncertainty: evidence from matched point and density forecasts", *The Review of Economics and Statistics*, Vol. 92, No. 1, pp. 200-207

Romer, C. D. (1990), "The great crash and the onset of the great depression", *Quarterly Journal of Economics*, Vol. 105, No. 3, pp. 597-624.

Romer, C. D. and D. H. Romer (2004), "A new measure of monetary shocks: derivation and implications", *American Economic Review*, Vol. 94, No. 4, pp. 1055-1084.

Rossi, B and T. Sekhposyan (2015), "Macroeconomic Uncertainty Indices Based on Nowcast and Forecast Error Distributions", *American Economic Review*, forthcoming.

Rossi, B. (2006), "Are exchange rates really random walks? Some evidence robust to parameter instability", *Macroeconomic Dynamics*, Vol. 10, No. 1, pp. 20-38.

Ruths, D. and J. Pfeffer (2014), "Social media for large studies of behavior", *Science*, Vol. 346, No. 6213, pp. 1063-1064.

Schaal, E. (2012), "Uncertainty, Productivity and Unemployment in the Great Recession", *mimeo*.

Scotti, C. (2013), "Surprise and Uncertainty Indexes: Real-time Aggregation of Real-Activity Macro Surprises", *International Finance Discussion Paper*, Federal Reserve Board, No. 1093.

Segal, G., I. Shaliastovich and A. Yaron (2014), "Good and Bad Uncertainty: Macroeconomic and Financial Market Implications",*mimeo*, https://www.economicdynamics.org/meetpapers/2014/paper_488.pdf

Sims, C. A. (2012), "Comment by", *Brookings Papers on Economic Activity*, Spring, pp. 141-148.

Squadrani, M. (2013), "A Google Trends-based approach for measuring economic policy uncertainty", *MSc dissertation*, Department of Economics, University of Bologna.

Stock, J. H. and M. W. Watson (2012), "Disentangling the channels of the 2007-09 recession", *Brookings Papers on Economic Activity*, Spring, pp. 81-135.

Toda, H. Y. and T. Yamamoto (1995), "Statistical inference in vector autoregressions with possibly integrated processes", *Journal of Econometrics*, Vol. 66, pp. 225-250.

Tsay, R. S. (1987), "Conditional heteroscedastic time series models", *Journal of the American Statistical Association*, Vol. 82, No. 398, pp. 590-604.

Whaley, R. (2000), "The Investor Fear Gauge", *Journal of Portfolio Management*, Vol. 26, pp. 12-17.

Tables and Figures

Tab. 1 - The dynamics of news-based (*BBD*) and Google Trends (*GT*) relationships ([a])

	Instantaneous correlation:	
Granger causality *from*/*to*:	*High* (>0.25)	*Low* (<0.25)
Query-pushed, *GT*/*BBD* ([b])	Fiscal policy (**FP**), Sovereign debt and currency crisis (**SDCC**)	Health care (**HC**)
News-pooled, *BBD*/*GT* ([c])	Monetary policy (**MP**), Regulation (**RE**)	Trade policy (**TP**)
No Granger causality	Entitlement programs (**EP**)	National security and war (**NS**)

([a]) This table summarizes the VAR results detailed in Appendix A.3, where also disaggregate multivariate inspections are reported.

([b]) Query-pushed = Google Trends indicators Granger-cause news-based ones.

([c]) News-pooled = news-based indicators Granger-cause Google Trends ones.

Tab. 2 - Univariate analysis of uncertainty proxies ([a])

	VIX	SPREAD	PREDICT1	FDISP	HNEWS	GT
Summary statistics						
Mean, μ	20.0	1.890	0.685	0.670	138.0	112.0
Median	18.4	1.760	0.656	0.691	133.8	111.9
Maximum	62.6	5.560	1.129	0.880	309.5	150.8
Minimum	10.8	0.860	0.553	0.350	54.9	88.0
Std. Dev., σ	6.8	0.653	0.103	0.101	39.2	12.8
Coeff. Variation, σ/μ	0.342	0.346	0.151	0.151	0.284	0.115
Skewness	1.985	1.641	1.711	-0.839	1.078	0.543
Kurtosis	9.785	8.034	6.034	3.382	5.276	2.764
T[a]	540	540	540	540	540	132
Cyclicality						
Downturn/upturn μ ratios	1.406	1.458	1.262	1.068	1.003	1.058
Downturn/upturn σ ratios	1.652	2.077	1.714	0.997	1.121	0.999
Persistence testing and measures						
Through unit-root tests ([b])						
- augmentation, *k*	8	2	2	6	15	11
- tests' p-values	0.001	0.002	0.017	0.002	0.076	0.584
- speed of adjustment, π	-0.120	-0.047	-0.017	-0.108	-0.098	-0.190
- half-life months, *m*([c])	5	14	41	6	7	3
- ARCH(1) test, p-values ([d])	0.001	0.022	0.000	0.000	0.032	0.091
Through fractional integration ([e])						
- *d* estimate	0.339	0.411	0.410	-0.293	0.427	0.463
- highest AR root	0.55	0.87	0.90	0.95	0.12	-
- highest MA root	-	0.46	-0.35	-	-	0.86
- half-life impulse-response, *m*	5	28	86	4	3	2
- ARCH(1) test (p-values) ([d])	0.000	0.021	0.000	0.001	0.002	0.052

([a]) Sample period: 1970m1-2014m12 (2004m1-2014m12 for GT). T=540, of which 83 downturn periods and 457 upturns (see NBER dating); only 18 downturn periods and 114 upturns for GT.

([b]) Dickey and Fuller (1979) test equation: $\Delta y_t = c + \pi\, y_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta y_{t-1} + u_t$; where *k* is selected by using the MAIC criterion of Ng and Perron (2001) starting from a given maximum number of lags (k_{MAX}).

([c]) Months for closing 50% of the disequilibria. In general, $m = \frac{\ln(1-p)}{\ln(1+\pi)}$, where *p* is the part of the initial gap to be closed between actual y_t and its long run forecast (if y_t is stationarity, it is its unconditional mean), and π is the speed of adjustment (see above). Here, *p*=0.50, i.e. 50%.

([d]) Significant ARCH tests highlight residuals' conditional heteroskedasticity of the first order for the estimated ARMA/ARFIMA model.

([e]) In this block, the first row reports the ML estimate of the fractional integration parameter *d* in the ARFIMA (p, *d*, q) model: $\rho(L)(1-L)^d(y_t - \mu) = \vartheta(L)\varepsilon_t$ which is covariance stationary for |*d*|<1/2; the second/third rows respectively report the highest roots of the $\rho(L)$ and $\vartheta(L)$ polynomials (to check for stationary and invertible ARMA processes); the fourth row reports the number of months *m* required for 50% of the initial impulse to accumulate.

Tab. 3- ARMA/GARCH model estimates ([a])

	VIX	SPREAD	PREDICT1	FDISP	HNEWS	*GT*
Parameter estimates([b])						
ϕ_1	0.836 ***	0.944 ***	1.571 ***	0.604 ***	1.378 ***	0.724 ***
ϕ_2			-0.599 ***	0.271 ***	-0.395 ***	
ϑ_1		0.329 ***			-0.879 ***	
ϑ_2			-0.135 **			
ϑ_3		0.033 ***				
α_1	0.285 ***	0.204 ***	0.498 ***	0.180 ***	0.054 ***	0.202 **
α_2				-0.150 **		-0.044 ***
β_1		0.796 ***	0.338 ***	0.965 ***	0.909 ***	
Persistence in						
conditional mean ([c])	-0.164	-0.056	-0.028	-0.125	-0.017	-0.276
conditional variance	0.285	1.000	0.836	0.994	0.963	0.158
No. of adverse shocks $\hat{\varepsilon}_t$ *above* ([d])						
$\hat{h}_t$	60	101	97	77	73	*62*
$1.5\times\hat{h}_t$	36	50	47	29	38	*33*
$3\times\hat{h}_t$	11	9	2	1	9	*12*

([a]) General ARMA(2,3)-GARC(1,2) equation: $y_t = \gamma_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \vartheta_1 \varepsilon_{t-1} + \vartheta_2 \varepsilon_{t-2} + \vartheta_3 \varepsilon_{t-3} + \varepsilon_t$, and conditional variance: $h_t = \omega + \alpha_1 \varepsilon^2_{t-1} + \alpha_2 \varepsilon^2_{t-2} + \beta_1 h_{t-1}$.

([b]) Unreported parameter estimates are restricted to zero. ***, **, and * denote 1%, 5% and 10% significance. Sample period 1970m1-2014m12 (except for *GT* whose sample is 2004m1-2014m12).

([c]) This measure is the speed of adjustment (see π estimates in Table 2), and is obtained as: $\hat{\phi}_1 + \hat{\phi}_2 - 1$.

([d]) The number of counts for *GT* is multiplied times 540/131 in order to account for the shorter *GT* sample (only 131 observations against 540 for the other series).

Tab. 4 – Modelling uncertainty shifts ([a])

Testing for the number of breaks

Equation for	VIX	SPREAD	PREDICT1	FDISP	HNEWS	*GT*
- number of breaks, *m*	3	4	4	3	3	3
- F-statistic	21.97	26.79	36.31	16.13	12.64	22.01
- Bai-Perron (2003) 5% *cv.*	11.14	11.83	11.83	11.14	11.14	11.14

([a]) Sample period: 2004m1-2014m12.

Tab. 5 - Dynamic relationships between uncertainty innovations over different periods ([a])

Panel A, sample period 1970m3 - 2014m12

Block Granger causality tests (p-values) in a VAR(2) of uncertainty innovations

		Equation for the innovation to:				
Two lags of explanatory innovations to:	*df*	**VIX**	**SPREAD**	**PREDICT1**	**FDISP**	**HNEWS**
VIX			**0.0000**	0.1394	**0.0255**	**0.0753**
SPREAD	*2*	0.5860		0.6423	**0.0126**	0.6966
PREDICT1	*2*	**0.0000**	**0.0000**		**0.0090**	**0.0155**
FDISP	*2*	0.8076	0.1042	**0.0137**		0.5002
HNEWS	*2*	0.4572	0.6211	**0.0069**	**0.0583**	
Joint	*8*	**0.0000**	**0.0000**	**0.0084**	**0.0002**	**0.0401**

Correlation between unpredictable uncertainty innovations, i.e. between VAR(2) residuals

SPREAD	**0.2567**			
PREDICT1	0.0978	0.0707		
FDISP	0.0523	-0.0039	0.0177	
HNEWS	**0.2956**	0.0386	**0.1199**	**0.1313**

Panel B, sample period 2004m4 - 2014m12

Block Granger causality tests (p-values) in a VAR(2) of uncertainty innovations

		Equation for the innovation to:					
Two lags of explanatory innovations to:	*df*	**VIX**	**SPREAD**	**PREDICT1**	**FDISP**	**HNEWS**	***GT***
VIX			**0.0045**	**0.0650**	0.4463	0.7382	0.7054
SPREAD	*2*	0.5897		0.8009	0.9860	0.1646	0.2983
PREDICT1	*2*	**0.0000**	**0.0000**		0.2964	0.2255	**0.0312**
FDISP	*2*	0.8866	0.8796	0.9645		0.3716	0.6046
HNEWS	*2*	0.8448	0.5947	0.6356	0.1261		0.3636
GT	*2*	**0.1044**	**0.0970**	0.7566	0.3525	0.7633	
Joint	*10*	**0.0000**	**0.0000**	0.2140	0.2354	0.6429	0.3331

Correlation between unpredictable uncertainty innovations, i.e. between VAR(2) residuals

SPREAD	**0.5945**				
PREDICT1	-0.0382	0.0810			
FDISP	**0.1702**	0.1203	0.0527		
HNEWS	**0.3169**	**0.1963**	**0.1451**	0.0208	
GT	-0.0467	-0.1128	0.0640	-0.0178	**0.2993**

([a]) In bold, 10% significant estimates.

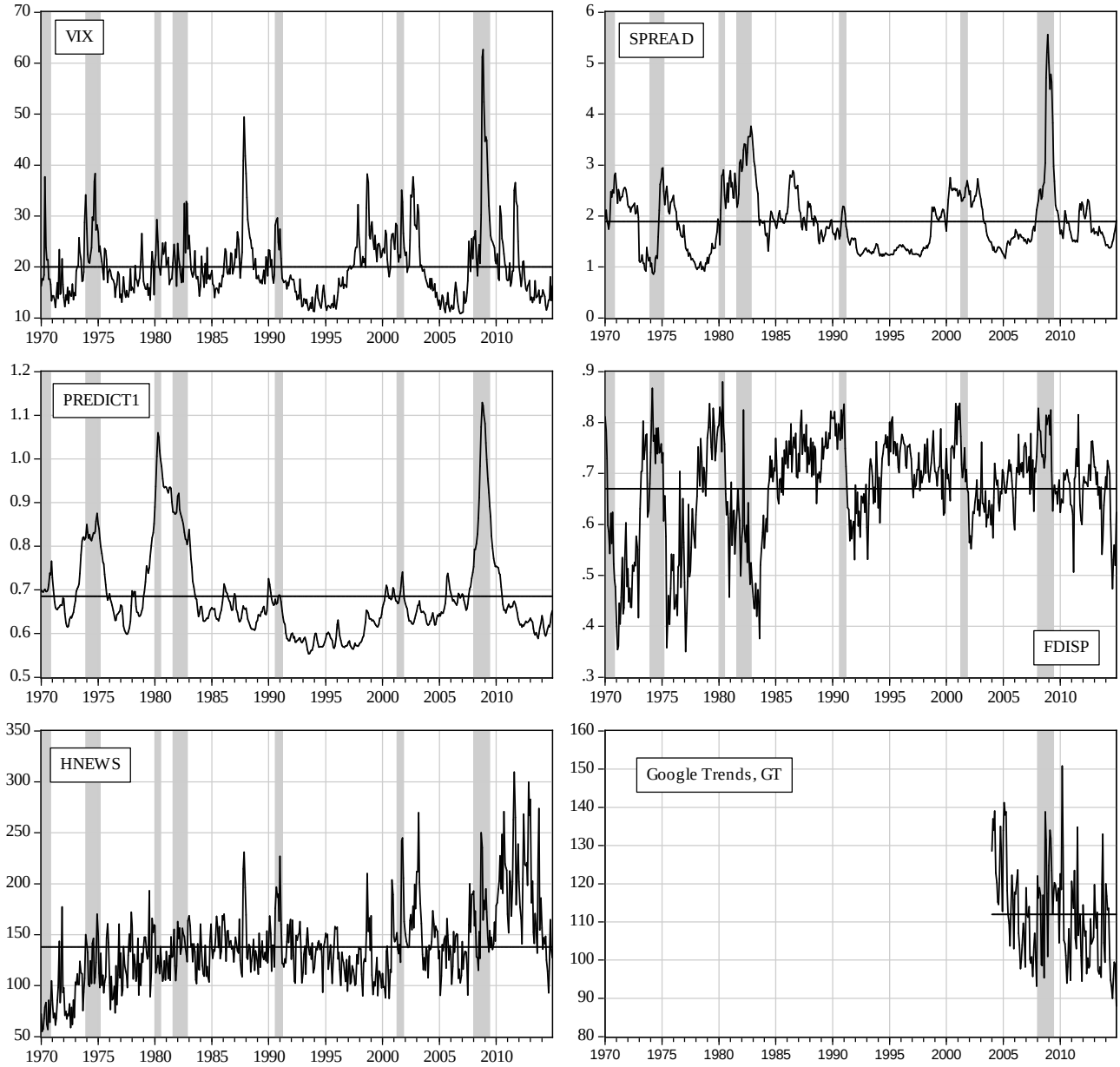
Tab. 6 - Dynamic relationships between uncertainty measures in the multiple-break context[a]

Estimates of the VAR(2) conditional on joint significant break dates, sample period 2004m1 - 2014m12

VAR equation for:	**VIX**	**SPREAD**	**PREDICT1**	**FDISP**	**HNEWS**	***GT***
Break dates and step-dummy estimates (reported as ratios over uncertainty sample averages)						
2006m3	**0.1160**	0.0296	0.0039	**0.0694**	0.0072	**-0.0886**
2008m9	**0.6206**	**0.2341**	-0.0049	0.0121	**0.4869**	**0.2566**
2009m4	-0.1786	**-0.1791**	-0.0007	**-0.1734**	-0.1509	-0.0942
2009m7	**-0.4120**	**-0.1628**	-0.0006	0.0836	-0.2065	-0.0496
2010m3	**-0.3319**	-0.0630	-0.0010	-0.0495	**0.2234**	**0.1528**
2010m5	**0.4287**	**0.1497**	0.0040	0.0608	-0.0493	**-0.2438**
2011m11	0.1172	**0.1838**	-0.0027	0.0030	0.0016	-0.0563
2012m10	**-0.2230**	**-0.1206**	0.0035	0.0028	0.1379	0.0350
2013m4	0.0695	0.0295	-0.0049	-0.0297	**-0.3870**	-0.0418
Block Granger causality tests (p-values)						
				VAR equation for:		
Two lags of explanatory innovations to:	*df*	**VIX**	**SPREAD**	**PREDICT1**	**FDISP**	**HNEWS**	***GT***
VIX			**0.0004**	0.1803	0.1334	0.1087	0.9419
SPREAD	*2*	**0.0041**		0.9017	0.7415	0.1200	0.7514
PREDICT1	*2*	**0.0000**	**0.0000**		0.9340	0.2771	0.5584
FDISP	*2*	0.8994	0.9888	0.9326		0.4792	0.7222
HNEWS	*2*	0.4633	0.1910	0.5978	**0.0160**		0.4730
GT	*2*	**0.0503**	0.5095	0.1383	**0.0167**	0.1466	
Joint	*10*	**0.0000**	**0.0000**	0.1183	**0.0365**	0.1349	0.7317
Correlation between residuals							
SPREAD		**0.4904**					
PREDICT1		-0.0164	0.1303				
FDISP		0.1217	0.0637	0.0684			
HNEWS		**0.3075**	**0.1393**	**0.1673**	0.0606		
GT		-0.0988	-0.1192	0.0217	0.0119	**0.1734**	

([a]) In bold, 10% significant estimates.

Fig. 1 – Alternative uncertainty measures over the common 1970m1-2014m12 period [a]



([a]) Shaded areas denote NBER downturn phases. Horizontal lines measure the sample averages.

Fig. 2 – GARCH conditional variance estimates (standardized, top), and residuals (down)

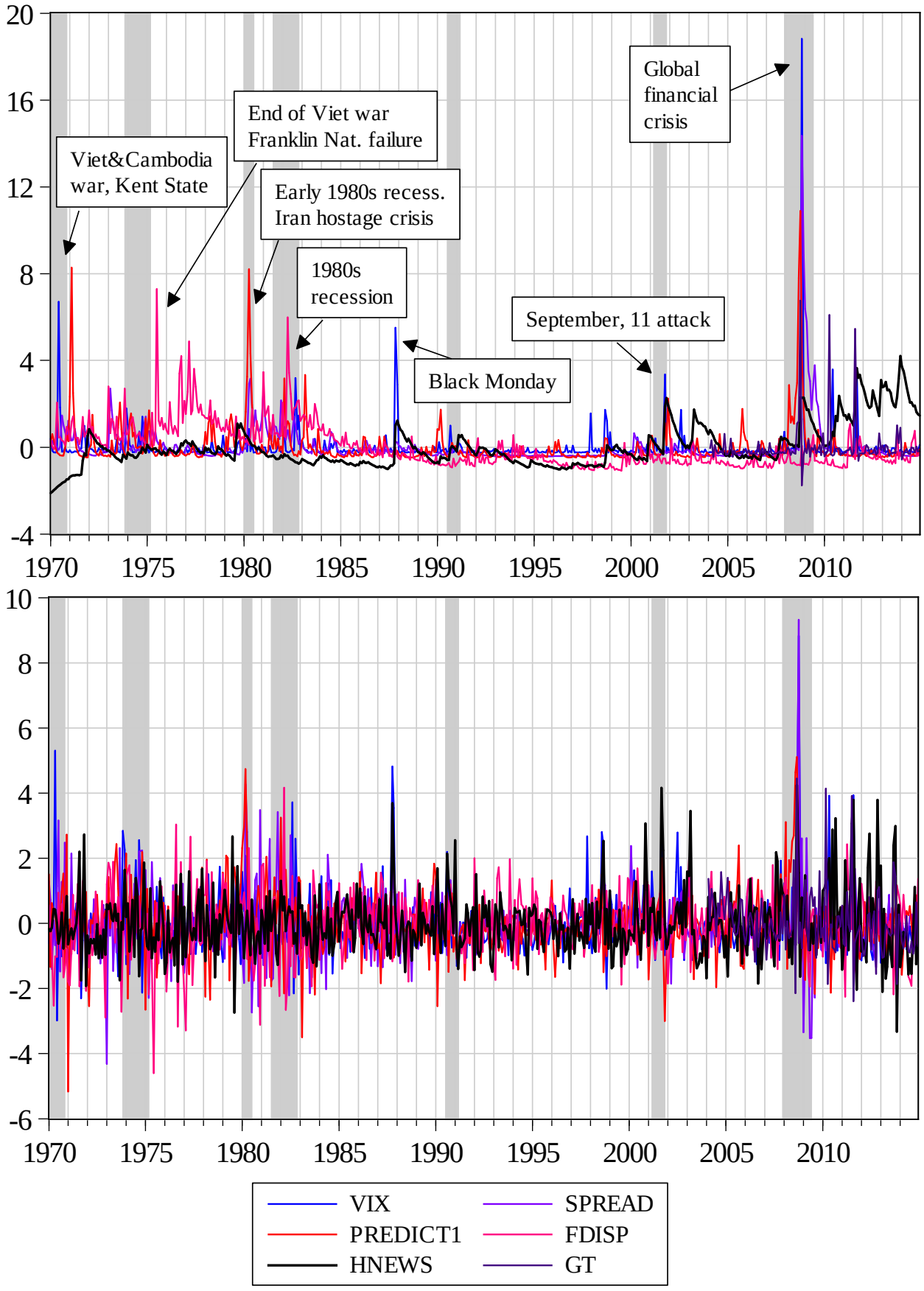
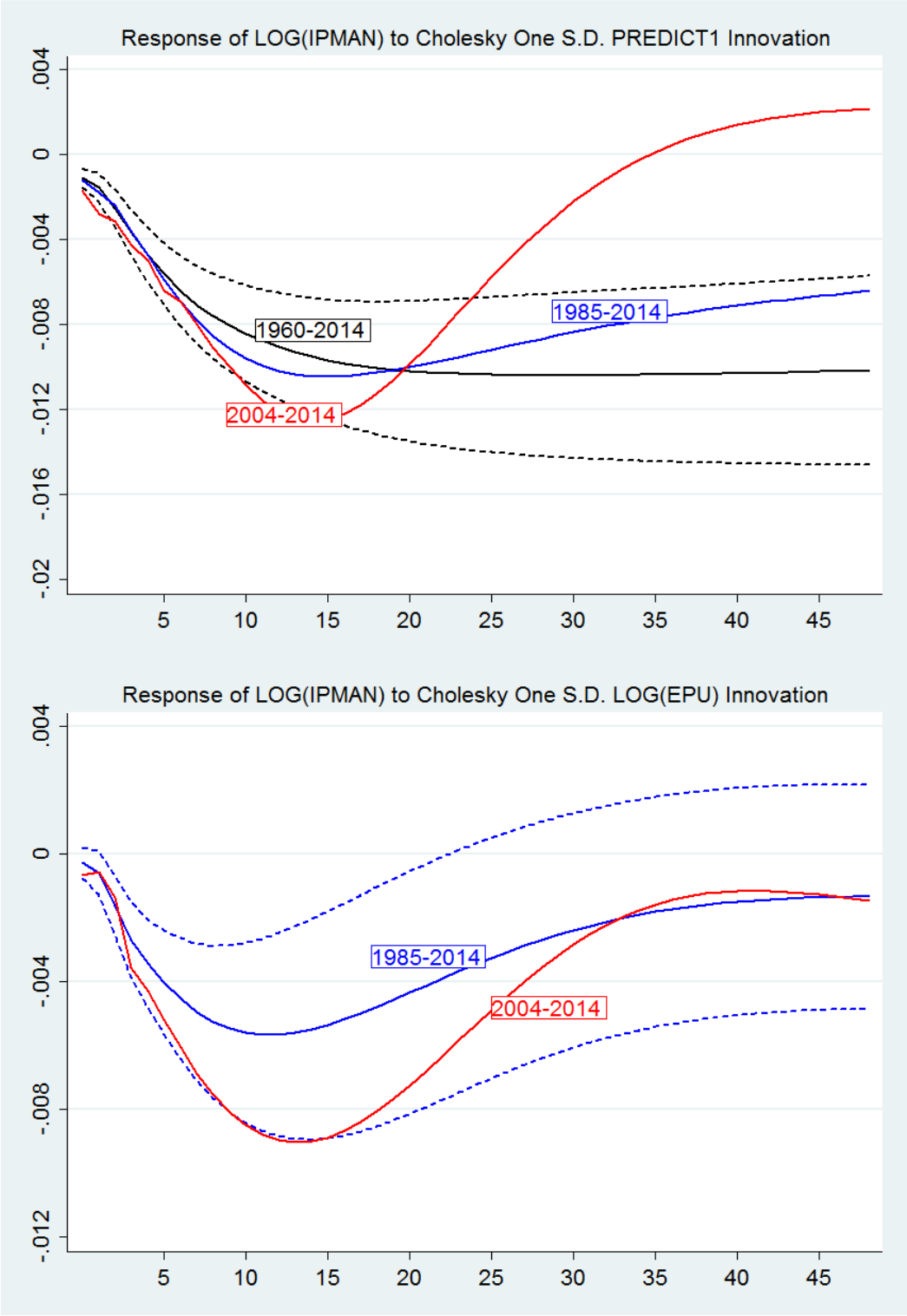
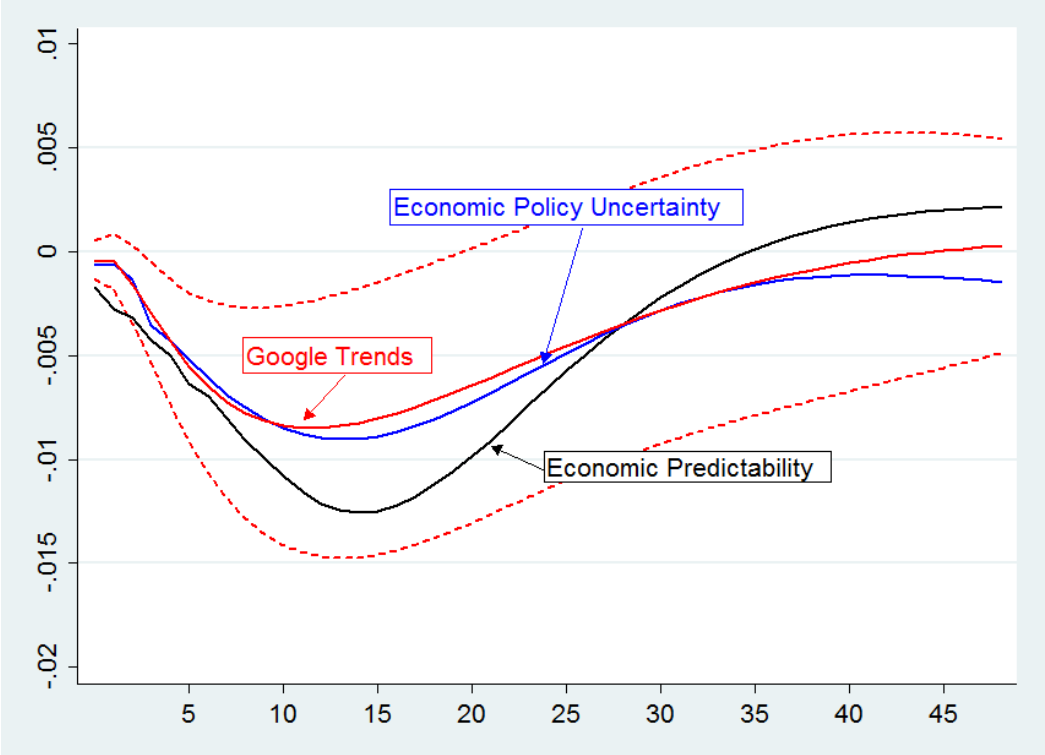


Fig. 3 – Impulse response of production (IPMAN) from the estimation of 5-variable VARs with alternative uncertainty measures (upper: PREDICT1, lower: EPU) and sample periods (black: starting from 1960; blue: from 1985: red: from 2004)



Response of log manufacturing production to Cholesky One S.D. impulse in PREDICT1 (seasonally adjusted) and log(EPU) for different estimation sample periods: (1) 1960m7 - 2014m12, (2) 1985m1 - 2014m12, (3) 2004m1 - 2014m12. Identification is based on a 5-variables VAR(p), ordered as: *uncertainty*, log(SP500), log(1+Fed funds effective rate/100), log manufacturing employment, log manufacturing industrial production. According to AIC criteria, for each sample, *p* is set equal to the optimal value reported in Table A6.2. Estimations on the sample 2004-2014 are performed with a small-sample degrees-of-freedom adjustment. Short-dashed lines represent 90% standard error bands of the IRF for the VAR estimated on sample (1) for PREDICT1 and (2) for log(EPU).

Fig. 4 – Impulse response of production (IPMAN) from the estimation of 5-variableVARs with alternative uncertainty measures (PREDICT1, EPU and *GT*) over the 2004-2014 sample period



Response of log manufacturing production to Cholesky One S.D. impulse in PREDICT1, log(EPU) and log(*GT11*). PREDICT1 and *GT11* are seasonally adjusted. Estimation sample period is 2004m1 - 2014m12. Identification is based on a 5-variables VAR(p), ordered as: *uncertainty*, log(SP500), log(1+Fed funds effective rate/100), log manufacturing employment, log manufacturing industrial production. According to AIC criteria, for each sample, p is set equal to the optimal value reported in Table A6.2. Estimations are performed with a small-sample degrees-of-freedom adjustment. Short-dashed lines represent 90% standard error bands for the IRF for an impulse in Google Trends.

Appendix A1 – The list of the 184 search terms by policy category

Taking the original list of 210 *BBD* search terms as our starting point, below we classify our selected 184 queries (*Sj*, *j*=1, .., 184) in 8 policy areas (*c* = 1, …, 8). We used only 184 terms of the 210 ones of *BBD*: of the 26 dropped terms, 8 of them were repeated several times in the list, while the other 18 never reached the minimum popularity threshold.

184 included queries of BBD

(1) Fiscal policy, FP (16 queries)

1. "taxes rates"
2. "tax rate"- "calculator"
3. "taxation"
4. "taxed"
5. "government spending"
6. "us federal budget"
7. "budget battle"
8. "balanced budget"
9. "fiscal stimulus"
10. "us budget deficit"
11. "federal debt"
12. "national debt"
13. "Gramm Rudman"
14. "debt ceiling"
15. "government deficits"
16. "balance the budget"

(2) Monetary Policy, MP (25 queries)

17. "the federal reserve"
18. "the fed"
19. "money supply"
20. "open market operations"
21. "quantitative easing"
22. "monetary policy"
23. "fed funds rate"
24. "Bernanke"
25. "Paul Volcker"
26. "Alan Greenspan" - "Mitchell" -"wife"
27. "the central bank"
28. "interest rates" - "calculator" - "best"
29. "fed chairman"
30. "fed chair"
31. "lender of last resort"
32. "fed discount window"
33. "European Central Bank"
34. "ECB bank"
35. "Bank of England"
36. "Bank of Japan"
37. "BOJ" - "xem" - "anglers" - "jamaica"
38. "Bank of China"
39. "Bundesbank"
40. "Bank of France"
41. "Bank of Italy"

(3) Health care, HC (15 queries)

42. "health care reform"
43. "Medicaid program"
44. "Medicare program"
45. "health insurance reform"
46. "malpractice tort reform"
47. "malpractice reform"
48. "prescription drug program"
49. "drug policy" - "nfl"
50. "food and drug administration"
51. "FDA regulation"
52. "medical malpractice law"
53. "liability reform"
54. "Medicare Part D" - "humana" - "aarp"
55. "affordable care act"
56. "Obamacare law"

(4) National security and war, NS (15 queries)

57. "national security strategy"
58. "us war"
59. "military conflict"
60. "terrorism"
61. "war on terror"
62. "after 9/11"
63. "defence spending"
64. "military spending"
65. "police action"
66. "us armed forces" -"ranks"
67. "military base closure"
68. "saber rattling"
69. "naval blockade"
70. "no-fly zone"
71. "military invasion"

(5) Regulation, RE (65 queries)

72. "federal regulation"
73. "banking supervision"
74. "Glass Steagall"
75. "tarp program"
76. "thrift supervision"
77. "Dodd frank" - "form" - "certification"
78. "financial reform"
79. "commodity futures trading commission"
80. "cftc"
81. "house financial services committee"
82. "Basel Accord"

83. "capital requirement" - "working capital"
84. "Volcker rule"
85. "bank stress test"
86. "securities and exchange commission"
87. "us sec"
88. "deposit insurance"
89. "fdic" - "jobs"
90. "fslic"
91. "office of thrift supervision"
92. "Office of the Comptroller of the Currency"
93. "firrea"
94. "truth in lending"
95. "union rights"
96. "union card check"
97. "collective bargaining law"
98. "national labor relations board"
99. "minimum wage"
100. "living wage" - "calculator"
101. "right to work"
102. "closed shop"
103. "wages and hours"
104. "workers compensation law"
105. "affirmative action"
106. "at-will employment"
107. "trade adjustment assistance"
108. "davis bacon"
109. "equal employment opportunity"
110. "eeo laws"
111. "osha safety"
112. "antitrust"
113. "competition policy"
114. "monopoly power"
115. "patent law" - "firm" - "firms" - "school" - "schools" - "lawyer" - "attorney" - "group" - "bar" -"jobs"
116. "copyright law"
117. "federal trade commission"
118. "the ftc" - "complaint"
119. "unfair business practice"
120. "competition law"
121. "price fixing" - "adm" - "apple"
122. "class action law"
123. "healthcare lawsuit"
124. "tort reform"
125. "punitive damages" - "punitive definition" - "definition of punitive"
126. "energy policy"
127. "energy tax"
128. "carbon tax"
129. "cap and trade"
130. "cap and tax"
131. "offshore oil drilling"
132. "clean air act"
133. "clean water act"
134. "environmental protection agency"
135. "the epa" - "jobs"
136. "immigration policy"

(6) Foreign sovereign debt and currency crisis, SDCC (14 queries)

137. "sovereign debt"
138. "currency crisis"
139. "currency devaluation"
140. "currency revaluation"
141. "currency manipulation"
142. "euro crisis"
143. "Eurozone crisis"
144. "European financial crisis"
145. "European debt"
146. "Asian financial crisis"
147. "Asian crisis"
148. "Russian financial crisis"
149. "Russian crisis"
150. "exchange rate policy"

(7) Entitlement programs, EP (20 queries)

151. "entitlement program"
152. "entitlement spending"
153. "government entitlements"
154. "social security" - "office" - "number" - "my" - "calculator" - "online" - "jobs"
155. "government welfare"
156. "welfare reform"
157. "unemployment insurance"
158. "unemployment benefits" - "online"
159. "food stamps" - "application" - "online"
160. "afdc"
161. "tanf program"
162. "wic program"
163. "state disability insurance"
164. "oasdi"
165. "Supplemental Nutrition Assistance Program"
166. "Earned Income Tax Credit"
167. "eitc tax"
168. "head start program" - "jobs"
169. "public assistance" - "application" - "apply"
170. "government subsidized housing"

(8) Trade policy, TP (14 queries)

171. "import tariffs"
172. "import duty" - "calculator"
173 "government subsidies"
174 "government subsidy"
175 "wto" - "howto"
176 "world trade organization"
177 "trade treaty"
178 "trade agreement"
179 "trade policy"
180 "trade act"
181 "doha round"

182 "uruguay round"
183 "gatt" -"joseph" - "josh" - "stefan"
184 "anti dumping"

26 excluded queries of BBD

Fiscal policy:

1. "defence spending" → Already included in "National Security and War"
2. "military spending" → Already included in "National Security and War"
3. "entitlement spending" → Already included in "Entitlement programs"
4. "fiscal footing" → Not enough search volume to show graphs.

Monetary Policy:

5. "overnight lending rate" → Not enough search volume to show graphs.
6. "the fed" → Already included in "Monetary Policy"

Health care:

7. "prescription drug act" → Not enough search volume to show graphs.
8. "medical insurance reform" → Not enough search volume to show graphs.

National security and war:

9. "military procurement" → Not enough search volume to show graphs.
10. "military embargo" → Not enough search volume to show graphs.

Regulation:

11. "bank supervision" → Not enough search volume to show graphs.
12. "nlrd" → Not enough search volume to show graphs.
13. "advance notice requirement" → Not enough search volume to show graphs.
14. "overtime requirements"→ Not enough search volume to show graphs.
15. "merger policy" → Not enough search volume to show graphs.
16. "cartel" → Not possible to find an alternative query with enough search volume to show graphs.
17. "tort policy" → Not enough search volume to show graphs.
18. "medical malpractice" → Already included in "Health Care".
19. "drilling restrictions" → Not enough search volume to show graphs.
20. "pollution controls" → Not enough search volume to show graphs.
21. "environmental restrictions" → Not enough search volume to show graphs.

Foreign sovereign debt and currency crisis

22. "currency crash" → Not enough search volume to show graphs.

Entitlement programs:

23. "Medicaid" → Already included in "Health Care".
24. "medicare" → Already included in "Health Care".
25. "part d" → Already included in "Health Care".

Trade policy:

26. "import barrier" → Not enough search volume to show graphs.

How the single search terms are obtained and validated

The SVI_{st} can measure the volume of searches for a given term or phrase:

- if more search terms are listed together (up to a maximum of 30 words), Google Trends counts all searches that contain those words in any order;
- if the same list of search terms is placed between inverted commas, Google Trends counts all searches that include the words in the exact order they are entered (searches including other words before or after the term or phrase in inverted commas are also considered in the count);
- the Boolean operators "+" and "–" can combine search terms (by considering searches that contain any of several terms separated by the "+" sign), or exclude from the search volumes those queries that have nothing to do with the target in question (by listing them after "–").

Given that the main problem with obtaining reliable search terms is the potentially ambiguous identification of meanings that are close to the *BBD* search terms,[37] each term reported in *BBD* needs to be validated before retrieving the corresponding search volume series. This validation procedure is made easier since for each query entered, Google Trends also reports the "top" and "rising" related queries. This further information can help us establish whether we are referring exclusively to the actual search in question, or also to extraneous topics.

Our judgmental procedure is structured as follows. Each *BBD* search term (either single terms or specific phrases) is placed in inverted commas, and the corresponding list of "top" and "rising" queries is inspected. If nothing suspicious emerges, the search term is left unchanged (this happened in 64% of our cases). Otherwise, if any queries result as being suspicious, the solution to the problem depends on the number of such queries. When there are only a few of them, the Boolean "–" is used to exclude them.[38] However, if there are a great number of such suspicious queries, we add a number of words to the original *BBD* search terms in order to narrow results.[39]

The list above of our 184 queries reveals no suspicious results among the corresponding "top" or "rising" queries.

---

[37] Ambiguity problems can arise when there are terms with multiple meanings not all associated with the chosen topic (for example, some of the searches for "gatt", an acronym for General Agreement on Tariffs and Trade, relate to the soccer player Joshua Gatt, or to the actor Joseph Gatt); or when the search purpose is ambiguous (for example, some searches for "interest rates" relate to searches for an interest rate calculator).

[38] For example, if the search terms are "interest rates"–"calculator"–"best", the results will include searches containing the words "interest rates" etc. in this precise order, but will exclude searches with the words "calculator" or "best" reflecting queries related to anything other than monetary policy issues.

[39] For example, we often added words like "act", "law" or "program" to the original search terms.

Appendix A2 – The sum of the 184 series into sub-groups, policy categories and the aggregate

Since the 184 single series are scaled by their maximum value, each series is scaled by a different value, and thus different series cannot be directly summed into policy categories.

Ideally, we would simply merge the individual search terms listed in Appendix A1 into groups of queries so as to create 8 new search terms representing the combination of the individual ones, by using the Boolean "+". However, due to the Google Trends 30-word limit for each search term, the aim of replicating the 8 *BBD* policy categories using Google Trends cannot be accomplished in this way. The merging of single-series search terms is only viable if we split each of the 8 *BBD* categories into smaller sub-groups in order to keep the number of words in the resulting combined search terms below the upper limit of 30. Following this route we come to the 24 sub-groups (Gi, i = 1,.., 24), homogeneous in terms of the number of series within each sub-group, allocated to the 8 policy areas and listed below.

(1) Fiscal policy, FP (2 *sub-groups*)

1. query_1 – query_8
2. query_9 – query_16

(2)Monetary Policy, MP (3 *sub-groups*)

3. query_17 – query_24
4. query_25 – query_32
5. query_33 – query_41

(3) Health care, HC (2 *sub-groups*)

6. query_42 – query_49
7. query_50 – query_56

(4)National security and war, NS (2 *sub-groups*)

8. query_57 – query_64
9. query_65 – query_71

(5)Regulation, RE (8 *sub-groups*)

10. query_72 – query_80
11. query_81 – query_89
12. query_90 – query_98
13. query_99 – query_107
14. query_108 – query_114
15. query_115 – query_118
16. query_119 – query_126
17. query_127 – query_136

(6)Foreign sovereign debt and currency crisis, SDCC (2 *sub-groups*)

18. query_137 – query_143
19. query_144 – query_150

(7)Entitlement programs, EP (3 *sub-groups*)

20. query_151 – query_157
21. query_158 – query_164
22. query_165 – query_171

(8) Trade policy, TP (2 *sub-groups*)

23. query_172 – query_178
24. query_179 – query_184

Obviously, the merging of individual search terms to form sub-groups prevents the insurgence of those missing data that, in specific single queries, would result from the paucity of search volumes in certain weeks/months. Furthermore, given that the peak-normalization of the search volumes' series is made at sub-group level, the relevance of each search term within its sub-group is accounted for when we extract the 24 aggregate series by sub-group. Of course, this (partial) outcome (i.e. 24 sub-groups series from 184 individual search terms) still does not meet our requirement of having 8 Google Trends aggregate search volumes accounting for the weight of each search term within the category. This result can be achieved in one of two ways: Bayesian and common-term aggregation.

Bayesian aggregation

The first way (henceforth "Bayesian aggregation") is to shrink the number of individual search terms included in each sub-group by using the Bayesian model averaging (BMA) method.This gives us the list of the "most relevant" single search terms of those sub-groups belonging to the same category that can be successfully merged without exceeding the 30-word limit. The shrinking issue is tackled as a problem of the choice of the "best" of several explanatory variables (i.e. the single search volumes) in linear regressions where each sub-group is the dependent variable. In doing so, BMA provides a coherent method of inference of the model's parameters by taking explicit account of the uncertainty surrounding both the estimation and the steps of model selection: see, for example, Leamer (1978) and Magnus et al. (2010).[40]

The outcomes of the Bayesian aggregation approach are reported in Table A2.1, where the first column shows the descriptions of the sub-groups (i.e. our target variables), the second column lists the individual search terms that are parsimoniously selected by BMA, the third column reports the statistical significance of the selected search terms, and the fourth column shows the posterior inclusion probability (which must be greater than 50%).

[40]To perform BMA analysis we used the bma Stata code of De Luca and Magnus (2011). We also tried alternative model selection algorithms, such as Lars, Lassoand Least Angle algorithms, by using the lars procedure of Adrian Mander (2006). These alternative regularization methods utilize additional information in order to prevent overfitting; this information usually involves the introduction of a penalty for complexity, with different penalties involved by different methods; see Efron et al. (2004).The selection results of Lars, Lasso and Least Angle are not reported because they deliver outcomes that are very similar to those of BMA, albeit less parsimonious.

Tab. A2.1 - List of single search terms selected by BMA

SUB GROUP (*Gi*)	SEARCH TERM (*Sj*)	t	PIP
	(1)Fiscal policy, FP (16 queries)		
G1	2. "tax rate"- "calculator"	26.88	1.00
	3. "taxation"	22.07	1.00
	5. "government spending"	4.70	1.00
	8. "balanced budget"	1.15	0.66
G2	11. "federal debt"	1.57	0.80
	12. "national debt"	12.46	1.00
	14. "debt ceiling"	40.41	1.00
	(2)Monetary policy, MP **(25 queries)**		
G3	17. "the federal reserve"	4.69	1.00
	18. "the fed"	3.77	0.99
	21. "quantitative easing"	6.01	1.00
	22. "monetary policy"	4.03	1.00
	23. "fed funds rate"	6.26	1.00
	24. "Bernanke"	19.23	1.00
G4	25. "Paul Volcker"	4.19	1.00
	26. "Alan Greenspan" - "Mitchell" -"wife"	15.10	1.00
	28. "interest rates" - "calculator" - "best"	76.44	1.00
G5	33. "European Central Bank"	10.46	1.00
	35. "Bank of England"	1.84	0.86
	36. "Bank of Japan"	9.51	1.00
	38. "Bank of China"	8.40	1.00
	(3)Health care, HC (15 queries)		
G6	42. "health care reform"	76.15	1.00
	49. "drug policy" - "nfl"	3.43	0.99
G7	50. "food and drug administration"	8.52	1.00
	54. "Medicare Part D" - "humana" - "aarp"	18.22	1.00
	55. "affordable care act"	54.13	1.00
	(4) National security and war, NS (15 queries)		
G8	60. "terrorism"	42.30	1.00
	61. "war on terror"	4.77	1.00
	63. "defense spending"	3.49	0.99
G9	66. "us armed forces" -"ranks"	2.81	0.96
	67. "military base closure"	8.84	1.00
	70. "no-fly zone"	7.91	1.00
	(5) Regulation, RE (65 queries)		
G10	74. "Glass Steagall"	13.97	1.00
	76. "thrift supervision"	12.42	1.00
	77. "Dodd frank" - "form" - "certification"	25.12	1.00
	78. "financial reform"	29.26	1.00
	80. "cftc"	5.57	1.00
G11	84. "Volcker rule"	3.12	0.98
	85. "bank stress test"	3.15	0.98
	86. "securities and exchange commission"	7.65	1.00
	89. "fdic" - "jobs"	87.01	1.00
G12	91. "office of thrift supervision"	18.55	1.00
G13	99. "minimum wage"	35.60	1.00
	100. "living wage" - "calculator"	4.89	1.00
	101. "right to work"	11.26	1.00
	105. "affirmative action"	10.02	1.00
	107. "trade adjustment assistance"	1.59	0.81
G14	108. "davis bacon"	9.85	1.00
	109. "equal employment opportunity"	6.62	1.00
	112. "antitrust"	18.46	1.00

G15	116. "copyright law"	11.49	1.00
	117. "federal trade commission"	15.50	1.00
G16	122. "class action law"	6.05	1.00
	124. "tort reform"	28.89	1.00
	126. "energy policy"	20.72	1.00
G17	127. "energy tax"	16.23	1.00
	128. "carbon tax"	1.99	0.89
	129. "cap and trade"	25.80	1.00
	134. "environmental protection agency"	24.13	1.00
	135. "the epa" - "jobs"	5.91	1.00
	136. "immigration policy"	5.09	1.00
	(6) Foreign sovereign debt and currency crisis, SDCC (14 queries)		
G18	137. "sovereign debt"	11.04	1.00
	139. "currency devaluation"	6.64	1.00
	141. "currency manipulation"	5.66	1.00
	142. "euro crisis"	8.03	1.00
G19	145. "European debt"	9.96	1.00
	(7) Entitlement programs, EP (21 queries)		
G20	154. "social security" - "office" - "number" - "my" - "calculator" - "online" - "jobs"	91.23	1.00
	157. "unemployment insurance"	13.12	1.00
G21	158. "unemployment benefits" - "online"	36.49	1.00
	159. "food stamps" - "application" - "online"	64.23	1.00
	162. "wic program"	1.27	0.70
G22	165. "Supplemental Nutrition Assistance Program"	1.79	0.85
	166. "Earned Income Tax Credit"	14.33	1.00
	168. "head start program" - "jobs"	4.04	1.00
	(8) Trade policy, TP (13 queries)		
G23	173. "government subsidies"	-3.96	1.00
	175. "wto" - "howto"	9.40	1.00
	176. "world trade organization"	4.88	1.00
	177. "trade treaty"	2.42	0.94
	178. "trade agreement"	8.83	1.00
G24	183. "gatt" -"joseph" - "josh" - "stefan"	12.18	1.00

Reported search terms *Sj* are those whose t ratio is greater, in absolute value, than one and whose posterior inclusion probability (PIP) is greater than 0.5

The outcome in Table A2.1 has the considerable advantage of delivering a limited list of the "most relevant" individual terms within each sub group: the larger the weight of a single search term within its sub-group, the greater the likelihood it will be picked.[41] However, in the *BBD* categories (2) "monetary policy", (5) "regulation" and (7) "entitlement programs" the reduced set of individual search terms is still too large to merge the individual Bayesian-selected search terms into a compound search term category of less than 30-words. Therefore, in order to implement a Bayesian aggregation in the case of these three categories, we would need extra restrictions excluding further individual search terms.

[41] This information is utilized in Section 3 when we compare certain Google Trends search volumes within a category with the Newsbank categories of BBD.

Common-term aggregation

The second approach helps us and utilizes a feature of Google Trends that enables up to five different terms or groups of terms (henceforth, "common-term aggregation") to be compared. Starting from equation (1), the use of the function “compare” on Google Trends ensures that for each search term (or group of search terms) included in the comparison, we have the same scaling value ,corresponding to the maximum value of over the period *0-T*, among all the single series that we are comparing. Therefore, by scaling all the series for a common term, we can easily aggregate them by summing or averaging.[42]

For example, considering two generic series $\{j,\ n\}$ included for the purpose of comparison in Google Trends, and summing their *SVI*s, we obtain a series defined as follows:

$$SVI_{SUMt} = \left(\frac{sv_{jt}}{sv_{Gt} \times MSV_{[o,T]}} + \frac{sv_{nt}}{sv_{Gt} \times MSV_{[o,T]}}\right) \times 100 = \\ = \frac{100}{MSV_{[o,T]}} \times \left(\frac{sv_{jt} + sv_{nt}}{sv_{Gt}}\right) \tag{A2.1}$$

The volumes by sub-group obtained using the common-term aggregation (A2.1) are equivalent to what we would obtain by merging the single terms j and n in Google Trends equation (1) of the main text using the Boolean operator "+", except for small differences due to rounding and the fixed difference in the constant scaling term. However, it is worth stressing that the outcome of common-term aggregation has the advantage of bypassing the 30-word limit by merging search terms.

The search volume series, obtained as a result of Bayesian aggregation, deliver outcomes that are very similar to those obtained by means of common-term aggregation, since the number of search terms selected using BMA accounts for a very large portion of the variability of the full set of search terms by category (i.e. their weights are significant within the category).[43]

Therefore, hereinafter we shall only measure the Google Trends volumes by category as obtained using the common-term aggregation approach, and the same approach is followed to sum the 8 categories into the aggregate *GT* indicator of macroeconomic uncertainty.

[42] As far as we know, this approach has been only used by Hacamo and Reyes (2012).

[43] Note that the correlation coefficients by category between Bayesian aggregation (when feasible) and common-term aggregation (always feasible) are greater than 0.96 in four cases out of five, and equal to 0.70 in the remaining cases.

Appendix A3 – *BBD* and *GT* comparisons: detailed VAR outcomes

([a])	Fiscal policy, FP				Monetary policy, MP				Health care, HC				National security and war, NS			
	i	*j*	***BBD***	***Gi, Sj***	*i*	*j*	***BBD***	***Gi, Sj***	*i*	*j*	***BBD***	***Gi, Sj***	*i*	*j*	***BBD***	***Gi, Sj***
BBD				**0.5719**				**0.0163**				**0.4714**				**0.9711**
GT			**0.0245**				**0.1447**				**0.3422**				**0.7675**	
corr shock([b])			*0.4873*				*0.5321*				*0.1299*				*0.2013*	
Stationarity([c])			0.0162				0.0048				0.0075				0.0048	
BBD				**0.2455**				**0.0379**				**0.5628**				**0.9061**
Gi	**1**		**0.5980**		**3**		**0.6626**		**6**		**0.3248**		**8**		**0.9518**	
corr shocks			*0.3260*				*0.4965*				*0.0183*				*0.2315*	
Gi	**2**		**0.0174**		**4**		**0.3270**		**7**		**0.0176**		**9**		**0.4658**	
corr shocks			*0.4892*				*0.4617*				*0.2136*				*0.0455*	
Gi					**5**		**0.7396**									
corr shocks							*0.2009*									
Stationarity			0.0218				0.0300				0.0122				0.0218	
BBD				**0.6784**								**0.8654**				**0.7414**
Gi	**1**		**0.1513**						**6**		**0.4566**		**9**		**0.4936**	
corr shocks			*0.2808*								*0.0291*				*0.0517*	
Sj		**14**	**0.0035**							**50**	**0.1394**			**60**	**0.9663**	
corr shocks			*0.3879*								*0.0565*				*0.1432*	
Sj		**15**	**0.0165**							**54**	**0.9537**			**61**	**0.6608**	
corr shocks			*0.5062*								*-0.1048*				*0.1892*	
										55	**0.0089**			**63**	**0.0330**	
											0.2400				*0.1910*	
Stationarity			0.0274								0.0366				0.0057	

	Regulation, RE			Sovreign debt and crisis, SDCC			Entitlement programs, EP			Trade policy, TP			
	i	*j*	***BBD*** / ***Gi, Sj***		*i*	*j*	***BBD*** / ***Gi, Sj***		*i*	*j*	***BBD*** / ***Gi, Sj***		*I*	*J*	***BBD*** / ***Gi, Sj***	
BBD			**0.1204**				**0.0778**				**0.7728**				**0.0196**	
GT			**0.5375**				**0.0000**				**0.6957**				**0.9546**	
corr shocks			*0.3128*				*0.3775*				*0.4386*				*0.1149*	
Stationarity			0.0211				0.0180				0.0034				0.0098	
BBD			**0.0155**				**0.0633**				**0.6512**				**0.3342**	
Gi	**10**		**0.0001**		**18**		**0.0000**		**20**		**0.7332**		**23**		**0.7431**	
corr shocks			*0.1922*				*0.3116*				*0.2703*				*0.2265*	
Gi	**11**		**0.0004**		**19**		**0.2672**		**21**		**0.6438**		**24**		**0.0680**	
corr shocks			*0.3478*				*0.2803*				*0.4393*				*0.1299*	
Gi	**12**		**0.4497**						**22**		**0.8384**					
corr shocks			*0.1780*								*0.1930*					
Gi	**13**		**0.1379**													
corr shocks			*-0.0043*													
Gi	**14**		**0.4644**													
corr shocks			*0.0971*													
Gi	**15**		**0.6887**													
corr shocks			*0.0695*													
Gi	**16**		**0.2602**													
corr shocks			*-0.0666*													
Gi	**17**		**0.9977**													
corr shocks			*-0.0282*													
Stationarity			0.0466				0.0252				0.1332				0.1318	
BBD							**0.0030**								**0.0189**	
Gi					**19**		**0.3814**						**23**		**0.3321**	
corr shocks							*0.2472*								*0.1321*	
Sj						**137**	**0.0002**							**183**	**0.2868**	
corr shocks							*0.4034*								*0.0883*	
Sj						**139**	**0.0718**									
corr shocks							*0.0642*									
Sj						**141**	**0.3956**									
corr shocks							*-0.1101*									
Sj						**142**	**0.0639**									
corr shocks							*0.2094*									
Stationarity							0.0091								0.0020	

([a]) In each block of columns (one for each of the 8 categories *c*), rows are grouped in 3 sets (delimited by grey shadows) corresponding to the results obtained from 3 different VAR specifications: the upper set reports the results of the bivariate VAR for aggregate-by-*c BBD* and *GT* indicators; the middle one is for the VAR with the aggregate-by-*c BBD* indicator and the *GT* index disaggregated by sub-groups (*Gi*) belonging to the considered category *c*; the lower set deepens middle results by substituting significant *Gi* indexes with the single *GT* search terms (*Sj*) which emerged as the most relevant in the *Gi* sub-group (according to the BMA analysis in Table A2.1). In each of the 3 sets, figures in bold report the p-values of the Granger causality tests. In particular, those in the central column of each block (labelled as "***BBD***") assess the causality from the single Google Trends series (either *GT*, or *Gi*, or *Sj*) to the news-based index (*BBD*); figures in the last column of each block (labelled as "***Gi, Sj***") assess the joint reverse causality from *BBD* to all the Google Trends series.

([b]) "corr shocks" = instantaneous correlations between the residuals of the *BBD* equation and each of the Google Trends indicator equations (either *GT*, or *Gi*, or *Sj*).

([c]) "Stationarity" = p-values of the Johansen (1995) trace test for assessing the full rank: if the null hypothesis is rejected, all the variables in VAR are jointly stationary. When such null is not rejected (i.e. in the middle VAR of EP and TP), the non-standard Granger causality is assessed as suggested in Toda and Yamamoto (1995).

Appendix A4 – A focus on "government shutdown" and "debt ceiling"

The individual search terms “government shutdown” and “debt ceiling crises” can be interpreted as policy uncertainty shocks, since during the standoff people are concerned about, and discuss, "what economic policy actions will be undertaken and when, and the economic effects of past, present and future policy actions".[44] For example, with reference to the Government shutdown on October 8 2013, a White House statements reports: "After a discussion about potential paths forward, no specific determination was made […] The President looks forward to making continued progress with members on both sides of the aisle" (source: Reuters). The thousands of comments on online newspapers articles reporting this news testify to the uncertainty surrounding this situation.[45]

As far as the methodological approach is concerned, we will not rely on the time series properties with VAR models, but rather on their behaviour after the trigger event, because we are only faced with a debt-ceiling crisis or (even more so) a Government shutdown - i.e. two important but rare events - "once in a blue moon". In Figures A4.1 and A4.2, the relative frequency of newspaper mentions and the Google Trends series are in fact near to zero over a large portion of the sample, with the exception of a limited number of significant spikes.

Figures A4.1 and A4.2 here

As expected, *BBD* and *GT* indicators substantially overlap, as the same trigger events drive both of them. However, interestingly Google searches decline more rapidly after the peak, as if in the case of important events, most people seek information about what is going on when the standoff is at its peak; after this peak, many of them stop caring about the uncertainty question well before newspapers do. In particular, with regard to the “Government shutdown” series in Figure A4.1, the highest spike of Internet activity (with a value of 100, considering the weekly raw series downloaded from Google Trends) was during the week from 29/09/2013 to 05/10/2013, which was to be expected given that from October 1 to October 16 certain federal agencies were temporarily closed or worked part-time. Google figures for the following two weeks were 28 and 26 respectively, and after these 3 weeks Internet searches declined rapidly to a mean value of 1.5 between 20 October and 30 November 2013. Media coverage, on the other hand, remained high for a longer period, declining more gradually: considering the standardized time-series reported in Figure A4.1, after a peak of 10.45 in October (associated with a value of 5.76 for the Google Trends

[44]To quote BBD’s definition of what an article about policy uncertainty should be concerned with.

[45]As an example, in response to this news we discovered: 12,265 comments on Yahoo News *(“They're talking? Really? Republicans, Obama have 'productive' meeting on debt, shutdown”* by C. Moody), 5,697 in The Washington Post (*“House, Senate Republicans offer competing plans on debt-limit, government shutdown”*, P. Kane, Z. A. Goldfarb and L. Montgomery*)*, and 1,716 in The Wall Street Journal (*“Obama, GOP Open Talks Over Temporary Debt Fix”;* J.Hook and P. O’Connor).

index in the same month), the following two months' values were respectively 1.89 (0.23 the corresponding *SVI*) and 1.31 (0.15 the corresponding *SVI*), well above the one standard deviation threshold, unlike the Internet searches series, the observed values of which were close to the mean of zero. A similar trend, with Internet searches declining more rapidly than media coverage in the aftermath of a shock, was observed after interest peaked with regard to the debt ceiling debate of July 2011.

Given these findings, if Internet activity is perceived as a proxy of the number of people influenced by a policy uncertainty shock, then we can tentatively conclude that many people just glance at media reports, while few of them give much attention to them for any length of time. Following less severe shocks, however, news reports and Internet searches spike together, and also decline together, since in the wake of a moderate shock even newspapers quickly stop concerning themselves with the issue in question.

Fig. A1 - News coverage and Internet searches for the term "Government Shutdown"

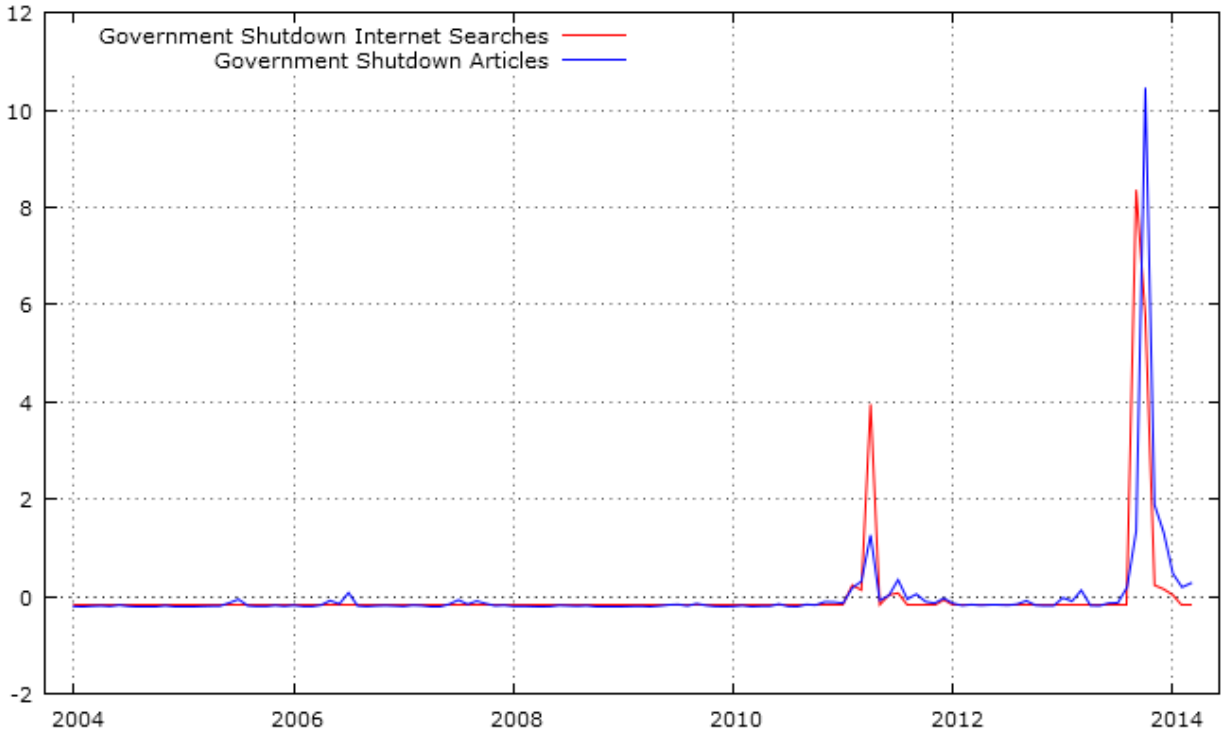
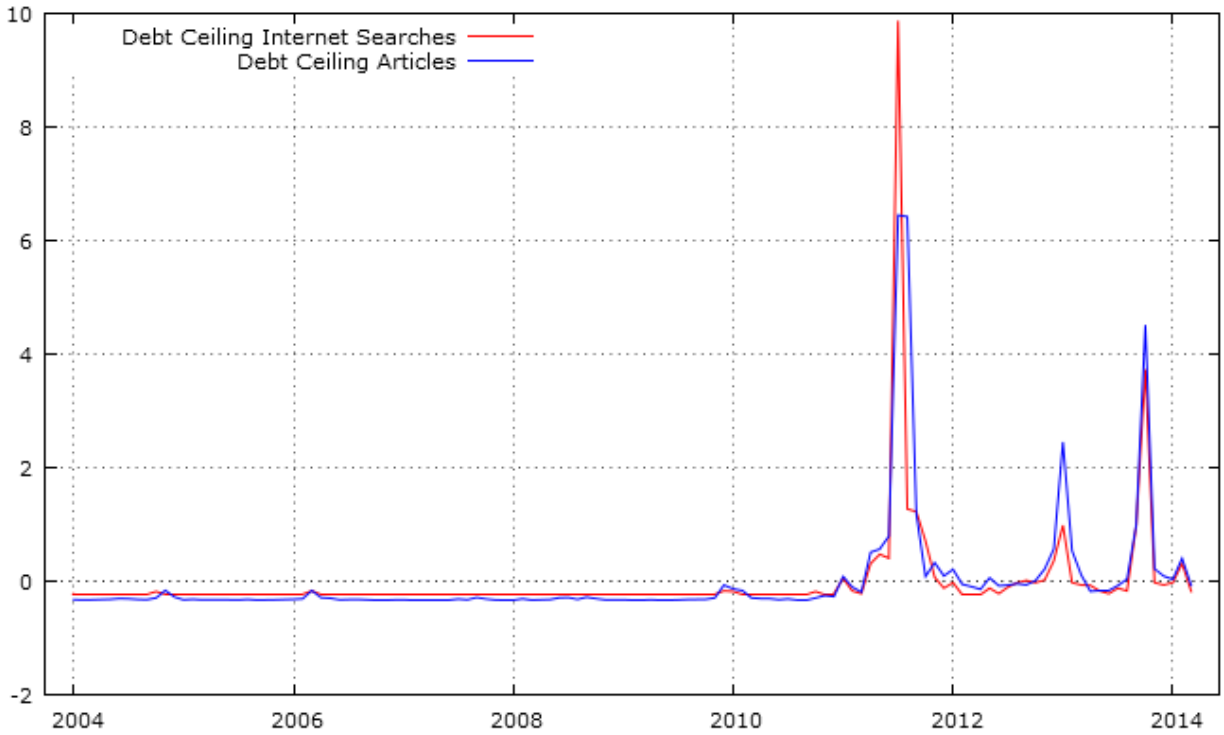


Fig. A2 - News coverage and Internet searches for the term "Debt Ceiling"



Monthly relative frequencies of mentions of "government shutdown" and "debt ceiling" in the US newspapers included in the Access World New's NewsBank service (data retrieved from www.policyuncertainty.com/categorical_epu.html) and Google Trends search volume indexes for the same search terms. The series are standardized to have a mean of 0 and a standard deviation of 1.

Appendix A5 – The measurement of the "fantastic five"

The (*i*) VIX proxy is obtained from the forecast 30-day volatility of the S&P 500 Index, and is computed by averaging the weighted prices of put and call options on the S&P 500 stock market index (for details, see the CBOE White Paper, 2009). Given that VIX monthly data have been available since January 1990, we used realized (rather than implied) stock market return volatility to proxy the pre-1990 missing VIX data going back to the beginning of the 1960s (see Bloom, 2009, and Bachmann et al, 2013).

The (*ii*) SPREAD proxy is defined as the difference between Moody's BAA-rated corporate bond yield and the 30-year Treasury constant maturity rate (monthly). Following Bachmann et al (2013), we used the 20-yearTreasury bond data when the 30-year Treasury bond data were not available, as in the 2002-2005 period and from the 1950s to 1977.

The (*iii*) PREDICT*h* proxy is forecast-based: Jurado et al. (2015) estimate uncertainty by assuming that a more (or less) uncertain economy is less (or more) predictable, and thus uncertainty can be measured as a lack of predictability. The estimation of PREDICT*h* dates back to 1960m7, and is obtained by aggregating the individual uncertainties of 132 economic and financial time series over alternative forecast horizons *h*. Three macro-uncertainty indexes (with $h = 1$, 3 and 12) are obtained from the following steps: (1) *h*-months ahead forecasts are estimated using diffusion index models based on both lagged and contemporaneous common factors, autoregressive terms and other predictors such as simple non-linear transformations of factors;[46] (2) the uncertainty regarding each single variable is defined as the volatility of the *h*-months ahead prediction error (conditional on the information set in period *t*), and is computed using stochastic volatility models; (3) all these uncertainties are combined into one single measure of overall macro uncertainty by means of simple averaging (alternative aggregation approaches deliver similar outcomes). We chose *a priori* to use the PREDICT1 ($h = 1$) in order to prevent the forecast horizon from overlapping the sample frequency of data. Results using the PREDICT1 are robust to the use of $h = 3$ and 12.

The (*iv*) FDISP proxy of uncertainty is the standard deviation of forecasts from two or more analysts. Bachmann et al. (2013) have estimated FDISP since 1968m5 by using the cross-sectional disagreement between the forecasts of large firms' management to be found in the Philadelphia Fed's Manufacturing Business Outlook Survey (BOS). The FDISP is based on the cross-section

[46] Factors are estimated through the principal components analysis of a set of 279 indicators that includes the 132 series on which individuals uncertainties are computed, plus 147 additional financial time series.

forecast dispersion of the fraction of each category of response.[47] The BOS sample composition reflects the industrial mix of the Third Federal Reserve district (i.e. eastern Pennsylvania, southern New Jersey, and Delaware), thus it might not perfectly reflect the general US outlook. Bachmann et al. (2013) acknowledge this caveat, by reporting similar results obtained using forecast data from surveys not restricted to any particular area or industrial sector.

The (*v*) proxy EPU is the news-based policy-related economic uncertainty index. Baker et al. (2015) obtain the EPU index by mixing three basic ingredients: (1) the newspaper coverage of policy-related economic issues (the news-based component NEWS); (2) the number of federal tax code provisions set to expire in future years; and (3) the disagreement among economic forecasters. Forecasters' disagreement is measured by the interquartile range of individual one-year forecasts reported in the Federal Reserve Bank of Philadelphia's quarterly Survey of Professional Forecasters. Among the SPF's wide range of predicted variables, BBD focus on the consumer price index, purchases of goods and services by state and local governments, and purchases of goods and services by the federal government. In order to obtain their global EPU index of policy-related economy uncertainty, BBD compute the average value of the three normalized components, using weights of 1/2 on the broad NEWS index (1), and 1/6 on each of the other three measures (the tax provision expiration index (2), and the disagreement of forecasters (3) about both the CPI and the federal/state/local purchases measures.

Of the above three EPU components, the one with more weight and of main interest here is NEWS, which is a restricted version of the Newsbank index we used in previous Section 2: actually NEWS and Newsbank have a correlation of 0.85, as reported by BBD. But NEWS is based on the search results from only 10 large newspapers: BBD perform month-by-month searches in each paper for terms relating to economic and policy uncertainty[48], and compute the ratio between the raw count of policy uncertainty articles and the total number of articles published in the same paper in the same month. They then normalize the resulting series, by newspaper, in order to obtain a unit standard deviation over the period from 1985m1 to 2010m12. Finally, they sum the values over papers, and re-normalize the multi-paper index to an average value of 100 from 1985m1 to 2009m12.[49]

---

[47] The responses to the BOS survey about the state of "general business conditions" over the following six months, are classified into three separate categories (up, no change, down). The three categories of interviewed forecasters are publicly available. In addition to Bachmann et al. (2013), see also http://www.phil.frb.org/research-and-data/regional-economy/business-outlook-survey/index.cfm

[48] In order to be included in the count, the news articles must include the words "uncertain or uncertainty", "economy or economics" and specific terms relating to economic and policy topics.

[49] Further details are available at: http://www.policyuncertainty.com/us_monthly.html

Since the sample period of EPU index is shorter than the periods covered by previous indexes (available well before the 1980s), our long-term comparison shall refer to a recent outcome (still in progress) of the BBD project: the Historical News-Based Policy Index (HNEWS), that extends NEWS back to 1900 by performing month-by-month searches for specific terms in 6 newspapers.[50] Since HNEWS is presently available up to 2014m2, we updated it to 2014m12 by using its "twin" NEWS index (over the overlapping period their correlation is about 0.9).

[50] See details at: http://www.policyuncertainty.com/us_historical.html.

Appendix A6 – Tests of seasonal adjustment of uncertainty and VAR lag selection

Tab. A6.1 -Seasonality tests (X13 filter context)

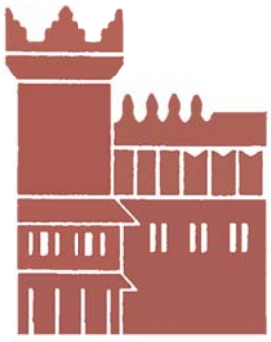
	Stable Seasonality		Moving Seasonality	Identifiable Seasonality	Monitoring and Quality Assessment Statistics
Variable	*F test*	*K*	*F test*[c]	*M7*[d]	*Q (without M2)*
PREDICT1	28.553[a]	245.3526[b]	4.325[**]	0.591[**]	0.65 [e]
GT	47.454[a]	108.5798[b]	2.171[*]	0.377[**]	0.53 [e]
EPU	6.087[a]	61.5886[b]	1.301	0.946[*]	1.27

(a) Seasonality present at the 0.1% level
(b) Seasonality present at the one percent level
(c) [**] and [*] denote moving seasonality present at the one and five percent levels
(d) [**] and [*]denote *Identifiable seasonality present* and *probably not present*
(e) Accepted

Tab. A6.2 -Test for the optimal lag length over each sample period [a]

		PREDICT1[b]			**log(EPU)**			**log(*GT*[b])**		
Sample	*Obs*	*FPE*	*AIC*	*SBIC*	*FPE*	*AIC*	*SBIC*	*FPE*	*AIC*	*SBIC*
1960m7 - 2014m12	648	6	6	2						
1985m1 - 2014m12	354	4	4	2	4	4	3			
2004m1 - 2014m12	126	5	5	2	3	4	2	3	3	2

(a) Maximum lags allowed: 6. Test based on a 5-variables VAR with *uncertainty*, log(SP500), log(1+Fed funds effective rate/100), log manufacturing employment and log manufacturing industrial production.
(b) Seasonally adjusted using the X13 procedure



Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
http://www.dse.unibo.it