EURQ: A New Web Search-based Uncertainty Index

EURQ: A New Web Search-based Uncertainty Index

Maria Elena Bontempi[*] Michele Frigeri[**,a] Roberto Golinelli[**,b] and Matteo Squadrani[**,c]

Abstract

Measuring economic uncertainty is extremely important for evaluating its role in economic activity. Nevertheless, measuring uncertainty is a difficult task since we do not know when economic agents perceive uncertainty and which type of uncertainty affects them. This paper introduces the economic uncertainty-related queries (*EURQ*) index, computed for both the USA and Italy, which measures economic, political, and normative uncertainty through large-scale searches on the Internet. We show that the *EURQ* captures economic agents' need for information in response to uncertainty shocks. Moreover, we show that this need for information is not just curiosity triggered by press coverage but rather captures individuals' genuine interest, particularly in specific topics subject to uncertainty. Hence, the *EURQ* can be fruitfully exploited to measure the level of uncertainty perceived by economic agents and to assess the role of specific types of uncertainty in economic activity.

Keywords: economic uncertainty measurement; perception of uncertainty; Internet searches; Google Trends; finance-, survey- and news-based indices.

JEL classification numbers: D8, E, C3, C8.

[*] Corresponding author, Department of Economics, University of Bologna, Piazza Scaravilli 2, Bologna IT-40126, Italy, mariaelena.bontempi@unibo.it

[**] Department of Economics, University of Bologna, Piazza Scaravilli, 2, Bologna IT-40126, Italy
[a] michele.frigeri@studio.unibo.it, CRIF SPA, Via M. Fantin 3, Bologna IT-40131, Italy
[b] roberto.golinelli@unibo.it, University of Bologna, Piazza Scaravilli, 2, Bologna IT-40126, Italy
[c] matteo.squadrani@studio.unibo.it, CRIF SPA, Via M. Fantin 3, Bologna IT-40131, Italy

1. Introduction

This paper introduces the economic uncertainty-related queries (*EURQ*) index, a new monthly index that uses Internet searches to measure economic agents' interest in topics affected by uncertainty. We present the theoretical and empirical fundamentals underlying the implementation of the *EURQ* index for the USA and Italy. The monthly updated series of *EURQ* for both the USA and Italy are hosted by the EPU (economic policy uncertainty) index website of Steven Davis (University of Chicago), Nick Bloom (Stanford University) and Scott Baker (Northwestern University).[1]

The use of Internet searches is motivated by two main pieces of evidence.

On the one hand, although uncertainty is a fundamental determinant of economic activity (e.g., Bloom, 2014), a unique or objective metric has not yet been defined. There are measurement difficulties stemming from the fact that uncertainty is determined by events whose heterogeneous nature varies depending on the period considered and is characterized by an ununiform and unknown distribution among economic agents. Relatedly, the uncertainty indicators developed in the literature (Bekaert et al., 2013, Bloom, 2009, Jurado et al., 2015, Bachmann et al., 2013, Ludvigson et al., 2020, Rich and Tracy, 2010, Rossi and Sekhposyan, 2015, Scotti, 2016, and Baker et al., 2016) focus on well-defined segments of the economy rather than on the average individual. Specifically, these indices could be classed as finance-based, forecast-based and news-based indices, which account for the risk aversion and feelings of investors, the feelings and disagreements of professional forecasters responding to surveys, and the perceptions of journalists, respectively.

On the other hand, even though online search technology was introduced quite recently (it first emerged in 1993), Internet users currently make trillions of online searches worldwide each year. According to Sirotkin (2012), these queries are composed of navigational queries (when the user looks for a specific web page that is known or supposed to exist), accounting for 12-15%; transactional queries (when the user looks to perform a transaction such as buying or downloading), accounting for 22-27%; and informational queries, accounting for 58-66% of total queries and characterized by an average query consisting of two or three terms with no phrase operators. The percentage share of the latter type of query suggests that many people view the Internet as an effective way of collecting information. This is also confirmed in a Pew Research Center (2016) survey revealing that the online channel is now the second most important information source after television and the most popular among people who prefer to read news stories rather than watch or listen to the news.

Remarkably, the literature has recently started to use Internet search data with different aims and interpretations: as predictors in forecasting (Vosen and Schmidt, 2011, Carrière-Swallow and

[1] The link to the EURQ vintage data is http://policyuncertainty.com/EURQ_monthly.html.

Labbé, 2013, D'Amuri and Marcucci, 2017, Bulut, 2018, Gotz and Knetsch, 2019, and Ferrara and Simoni, 2019), as an index of well-being (Algan et al.,2016), as an index of job search activity (Baker and Fradkin, 2017), or as a measure of individual moods (investors' sentiment in Da et al., 2011, 2015, the interest that the municipal balance sheet generates among voters in Repetto, 2018, and investors' need for information about earnings announcements in Drake et al., 2012).

However, online searches are still not widely exploited in the construction of uncertainty indices. Our main contribution is to show that this information is important for measuring various aspects of uncertainty.

The volume of *informational queries* specifically related to economic and political issues is the main ingredient of our *EURQ* index: it measures the quantity of searches on uncertainty-related topics to quantify the uncertainty perceived by economic agents. We demonstrate how Internet search volumes can be used to obtain a measure of interest/confidence/feelings/worries/fears expressed by people and *driven by* uncertainty. In this regard, we show that the few papers that have tried to measure uncertainty by using Internet searches (BBVA, 2012, Dzielinski, 2012 and Donadelli, 2015) are affected by methodological limitations due to the limited dictionary of terms that they use.

Our new *EURQ* index offers five main advantages over the uncertainty measures currently available. First, it is based on Google Trends, which is publicly available, free, and very easy to access and download. Therefore, the first appealing aspect of the *EURQ* index is its reliance on a freely available survey of web searchers. Second, connected to the first advantage, the *EURQ* is downloadable in real time, timely updated and easy to compute. This means that the *EURQ* can detect changes in people's moods and feelings at an early stage. Third, the *EURQ* reveals attitudes rather than inquiring about them, and consequently, it may disclose more personal information in cases when non-response rates in surveys are particularly high or the incentive for truth-telling is low (Da et al., 2015). Fourth, the *EURQ* can refer to different geographic locations, both at the regional level within the same country and at the country level inside areas characterized by heterogeneous degrees of development. The possibility of computing uncertainty indices for countries that are usually not covered by other uncertainty measures represents a great advantage of our approach. A fifth interesting aspect of the *EURQ* is the possibility that it offers to compute specific indices for various sub-categories of uncertainty, such as financial, economic, political, and normative, once the appropriate lists of search terms are drawn up. Moreover, the wording of the specific queries can be easily updated to fit changes occurring in the world.

To gauge the feeling of uncertainty among people, the definition of appropriate search terms that individuals usually ask Google when they need further information is of paramount importance. In the case of the USA, we selected 183 queries closely related to 210 search terms that Baker et al.

(2016) used to create the Newsbank version – based exclusively on news data – of their EPU index. To construct the *EURQ* for Italy, we adjusted these 210 search terms to fit the Italian case and ended up with a list of 136 queries. Hence, the *EURQ* index is related to the news-based approach, but the replacement of the frequency of newspaper articles containing specific terms with the frequency of individual queries involving similar search terms represents *a shift in focus* from the channel through which the message is conveyed (the press, the media) towards the receivers of the message (individuals). This shift in perspective implies that the index may also be made for countries/regions where press coverage is incomplete, lacking and/or substantially biased (on media bias, see Groseclose and Milyo, 2005, Puglisi and Snyder, 2016, and Ban et al. 2019). In addition, the *EURQ*'s effectiveness in capturing uncertainty – unlike that of news-based measures – does not depend on the intensity of newspaper use, as web-search activities can also refer to *local sources of information* that embody a wide set of tools, such as social networks, for spreading information gleaned "chatting with neighbours over the garden fence"; see Lahiri and Zhao (2017) and Banerjee et al. (2019).

To clarify the nature and value of our new index, we carry out two types of comparisons.

The first comparison is between our *EURQ* and other uncertainty measures for the USA: finance-based, forecast-based, and news-based indices, ex ante and ex post measures of uncertainty and the principal component of all the uncertainty measures over common sample periods of different lengths. In the case study of Italy, we also include in the comparison the handful of available search-based indices (BBVA, 2012, Dzielinski, 2012, and Donadelli, 2015).

The second comparison examines whether the interest manifested by economic agents in the USA is only driven by the press's emphasis on specific events or whether certain specific topics can spontaneously attract people's genuine interest. Since within a month, even on a given day, it is highly possible that individuals' web searches are prompted by what they heard on the news, we provide evidence at both monthly and daily frequencies.

The paper is organized as follows. Section 2 presents the conceptual framework and the technical issues associated with the construction of the *EURQ* index. Section 3 compares the *EURQ* with alternative uncertainty indicators for the USA and presents Italy as a case study. Section 4 assesses which components of uncertainty generate spontaneous interest among Americans. Section 5 discusses the results and offers our conclusions.

2. Using Internet search volumes to construct a new uncertainty index

2.1. The conceptual framework of the EURQ 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). While

aleatory uncertainty is irreducible, as it arises naturally from our perceptions of real-life facts or from "observing the system", epistemic uncertainty represents a lack of knowledge about potentially knowable facts.[2] In the latter case, uncertainty depends on a narrow information set (Harmanec, 1999) and fuels individuals' need to gather more information when they want to make decisions. According to Kim et al. (2020), "Diffusion is a subset or specific type of communication in that what is communicated is always perceived to be new. Newness, when combined with perceived relevance or importance, creates uncertainty in potential adopters. Uncertainty leads to a desire to resolve it through a search for more information, especially in instances of cognitive inconsistency."[3]

We suggest that Internet search volumes can be exploited to build an indicator of epistemic uncertainty, meaning that economic agents' interest in a larger information set, specifically their need for more information, arises when they are worried about something that is uncertain and could have consequences affecting them. In recent years, the Internet has become an effective means of collecting and divulging information for an increasing number of people in the USA. Approximately 85% of Americans in 2016 (95% in 2020) obtained at least some of this information through websites, apps and social networks, and the online channel is third among the six major news platforms, behind local television news and national or cable television news.[4] Our hypothesis is also reinforced by the query-type distribution reported by Sirotkin (2012), where 58-66% of Internet search activity consists of informational queries revealing people's collective interest in and desire for greater knowledge.

Among the alternative search engines, we chose Google Trends because Google is currently the leading search engine, boasting a worldwide market share in February 2017 of 80.5% for desktops and 98.9% for laptops, tablets, and other mobile devices. Sirotkin (2012) claims that since users are unlikely to be experts in traditional information retrieval systems and query languages, web search engines target the average Internet user or, to be more precise, any Internet user, whether new to the web or a seasoned Usenet veteran. Bontempi et al. (2019) examine the role of the changing nature of Internet use and the introduction of social networks since the beginning of Google Trends data availability in 2004, corroborating our choice of this large data provider.

---

[2] For example, regarding uncertainty in relation to official economic statistics, Manski (2015) considers uncertainty a lack of knowledge deriving from an incomplete understanding of the information provided about an economic concept or from a lack of clarity on the concepts themselves.

[3] Examples come from various disciplines. In the field of economic psychology, individuals respond to greater uncertainty by intensifying their search for more information (Lemieux and Peterson, 2011). In economics, imperfect (noisy) and sticky information models predict that "more volatile shocks [greater uncertainty] lead to the more frequent updating of information, since inattentiveness is costlier in a world that is rapidly changing" (Reis, 2006, p. 803) and that "more tranquil times should be ceteris paribus associated with greater information rigidities" (Coibion and Gorodnichenko, 2015, p 2674).

[4] For further information regarding Internet users in North America, see the results from the survey on search engines' market shares at http://www.internetworldstats.com/stats2.htm and http://www.netmarketshare.com/.

2.2. The conditions supporting the EURQ as an uncertainty index

Three conditions must be satisfied to support the case for using the *EURQ* as an economic uncertainty index, i.e., the idea that the more the economic system is uncertain, the more economic agents need information and make searches on the Web.

Condition #1 (C1) regards *selection of the terms* to be included in the queries and *evaluation* of whether the searches peak in conjunction with periods of high economic and political uncertainty.

Condition #2 (C2) involves validation of the index through the study of its *statistical properties* and *relationship* with other uncertainty indices.

Condition #3 (C3) concerns identification of the *specific components of uncertainty* representing individuals' *genuine interest* that is not triggered by press coverage.

C1, discussed in more detail in Section 2.3, aims to guarantee that the *EURQ* is interpretable as capturing the need for more information of all economic agents driven by uncertainty. In a nutshell, the *EURQ* must magnify the signal against the noise by excluding motives – such as curiosity or the desire to know more about something – that have nothing to do with economic and policy uncertainty. Hence, the balance between exhaustivity and arbitrariness in the selection process of the search terms is of paramount importance. In terms of exhaustivity, a long list of terms takes advantage of the statistical averaging effect across many different queries and encompasses a variety of diverse sources and symptoms of uncertainty, so it minimizes arbitrariness in both the selection of the list and the specific wording of the queries. For example, in Section 3.3, we exploit the case study of Italy to show the problems arising from an inappropriate and excessively short list of queries. In terms of arbitrariness, we assume, in line with the epidemiological model of Carroll (2003), that the wording of the queries used by web searchers is affected by the jargon of journalists because the news published represents the main mode of propagation of the need for further information among the entire population. Thus, the *EURQ* is related to the news-based approach of Baker et al. (2016), henceforth BBD.

The list of queries that we used is reported in Appendix A1. As examples, it excludes the query "baseball", while it includes the query "European Central Bank". Even if fans may wonder about whether their team's games are cancelled due to the COVID-19 pandemic, the relationship of the term "baseball" with economic and policy uncertainty would be spurious and not stable over time. In reality, pandemic-related economic and political uncertainty can be better captured through queries associated with the political and economic situation driven by the pandemic, such as searches on income and social assistance policies. In addition, the volume of searches for "baseball" is generally stable over time (apart from obvious seasonal fluctuations), invariably peaks during the World Series, and fell by 67% from June 2019 to June 2020 because of the halting of public Major League Baseball

games, a pattern unrelated to the one expected during periods of increasing uncertainty. In regards to the second example term, some people might search for "European Central Bank" because they need to analyse Christine Lagarde's speeches and to better understand the bank's views regarding a possible sovereign debt crisis: this is *interest due to uncertainty* in monetary policy issues. Simultaneously, other people might search for "European Central Bank" because of their individual, extemporaneous interest in the bank's research agenda and their wish to examine the most recently published working papers: *this motive is unrelated to uncertainty*. Sporadic interest due to either curiosity or trivial reasons, and hence unrelated to uncertainty, is mere noise, uncorrelated across individuals and randomly fluctuating without any specific pattern.[5] Instead, collective uncertainty-driven interest is the signal that the *EURQ* must capture. This signal is characterized by timing and dynamics derived from common factors induced by millions of simultaneous Internet searches due to the diffusion of uncertainty through conversations between agents, imitating behaviour of people (Sims, 2003), and news divulged by the media.

These examples make clear that the volume of *the selected* queries (containing terms such as "European Central Bank", "spread", "unemployment", "inflation rate", and "public debt") expresses the uncertainty deliberatively manifested by economic agents if, and only if, such queries peak at the time of corresponding episodes characterized by a high degree of uncertainty. Section 2.3 below shows this connection for the *EURQ*.

After selecting the queries and downloading the index, we can verify **C2** and **C3** and reinforce our evidence that the *EURQ* distinguishes noise from signal (curiosity from uncertainty) with clear examples. **C2** is satisfied in Section 3 through the comparison of the *EURQ* with the risk aversion and sentiment of investors (finance-based index of uncertainty); the feelings of respondents expressed in a survey (as captured by forecast-based measures of uncertainty); the counts of specific words reported by journalists (the news-based measure of uncertainty); ex ante and ex post measures of uncertainty; and the principal components of all the uncertainty indices proposed by the literature.[6] This comparison, at both the univariate and vector autoregressions (VAR) levels and with Italy as a case study, furthers our knowledge of unobservable uncertainty and of what the *EURQ* adds to our comprehension of uncertainty.

[5] A spike in the search volumes may also occur after the ECB changes interest rates, i.e., in a moment that should reflect the resolution of uncertainty rather than uncertainty itself. However, this occurrence points to the fact that despite the implemented policy change, the state of agents' knowledge is not yet perceived as fully satisfactory – hence the need to gather additional information to temper uncertainty.

[6] Bontempi et al. (2019) show that the opposite procedure of selecting those terms most closely correlated with the uncertainty indices available in the literature (the so-called correlate approach) delivers *few, generic, spurious search terms* that have nothing to do with uncertainty. These results also underlie our scepticism regarding the reliability of uncertainty indices based either on few terms or on terms found using the correlate approach, such as those proposed by BBVA (2012), Dzielinski (2012), and Donadelli (2015). More details on this appear in Section 3.3.

C3 is satisfied in Section 4 through the comparison of web searches against press coverage on certain topics to quantify, at both monthly and daily frequencies, the timing and the importance of Internet searches for specific components of uncertainty over media coverage. Discovering whether journalists' opinions/feelings affect the degree to which agents become concerned about something and submit specific queries on the web considerably improves our understanding of the diffusion of uncertainty.

2.3. The practical implementation of the EURQ index

To construct the *EURQ* index, we extracted Google Trends series of queries closely related to the search terms employed by BBD when creating the *Newsbank* version – based exclusively on news data – of their *EPU* (economic and policy uncertainty) index.[7] To be included in BBD's *Newsbank* uncertainty index, newspaper articles must include the words "uncertain" or "uncertainty" (U), "economy" or "economics" (E), and one of the following policy terms (P): "congress", "deficit", "Federal Reserve", "legislation", "regulation" or "White House", along with 210 specific terms related to economic and policy topics (L). All this may be symbolized as "U&E&P&L".

Our *EURQ* index starts from the BBD *Newsbank* "L" list of 210 search terms and comes to a final list of 183 queries that people are likely to search when seeking information to overcome their feelings of uncertainty. To reduce ambiguity and obtain a list reflecting the vocabulary and expressions of Internet users, we made some wording adjustments. For example, regarding the search term "healthcare", BBD only include those articles containing both the terms "uncertain" or "uncertainty" *and* the terms "economy" or "economics" *and* "healthcare", whereas we use the term "healthcare reform". According to the BBD approach, it is important to focus exclusively on newspaper coverage of specific health-care issues related to economic uncertainty while excluding generic newspaper articles about medicine. In the *EURQ* case, adding the word "reform" disambiguates the overly generic "healthcare" and makes the search term appropriate for identifying, *per se*, the need to gather information about healthcare legislation instead of just general healthcare-related curiosity. As another example, consider the crisis arising from the coronavirus pandemic. It is important to highlight that terms such as "COVID-19", "pandemic" and "virus" are not included in

[7] BBD's complete list of words can be found in the appendix of Baker et al. (2016) and on their website at http://www.policyuncertainty.com/categorical_terms.html. For more information on the audit process regarding BBD's selected words, see the Audit Guide at https://www.policyuncertainty.com/media/Coding_Guide.pdf. Interestingly, this audit examination (designed to ascertain whether the uncertain mood is pervasive in articles with the listed words) supports the terms that we started from in our query definition. Without an audit, the selection of terms could be criticized as arbitrary. For example, Castelnuovo and Tran (2017) subjectively selected search terms "referring to words that are connected to uncertainty" reported in sentences of "various editions of the Beige Book and the Monetary Policy Statements". Similarly, the indices of Donadelli and Gerotto (2019) and Kupfer and Zorn (2020) derive from two specific features of the search volume extraction on Google Trends (search topics and search categories) that do not depend on specific lists of search terms and, for this reason, cannot be validated.

the list of search terms to construct the *EURQ*. What matters, in fact, is how the pandemic affects searches for terms related to a number of economic and policy issues, such as fiscal and monetary policies, healthcare and social protection, income support measures and unemployment benefits, the environment, foreign trade and sovereign debt. Of course, our methodology allows time-varying dictionaries with new terms to be incorporated to capture future drivers of economic and policy uncertainty.[8] The list of queries that we used to construct the *EURQ* is in Appendix A1. The plot of the *EURQ* with the timing of its peaks is shown in Figure 1.

Figure 1 here

The *EURQ* shows marked increases during crises, elections, and legislative debates, which are indeed periods of high uncertainty. For example, the *EURQ* clearly spikes in correspondence with important episodes of uncertainty, such as the stock market crash and the passing of the Emergency Economic Stabilization Act in late 2008, the passing of the Affordable Care Act in 2010, the debt ceiling dispute in mid-2011, the US government shutdown in late 2013, and the election cycle at the end of 2016. Note that the index also reaches a peak of 235.8 in February 2005, an event not commonly cited in the empirical literature on uncertainty but that is driven by searches for the term "social security" and is consistent with the debate over social security during the Bush administration (more on this in Section 3.1). The *EURQ* index hit a new record high during the COVID-19 pandemic, when it reached a level of 16% above the peak from September 2008. Again, it is important to note that the large and broad dictionary of words and expressions used for the *EURQ* allows us to capture future uncertainty events without including terms such as "COVID-19", "pandemic" and "virus".

3. The performance of the *EURQ* index in measuring uncertainty

3.1. The main features of the EURQ in comparison with other uncertainty proxies

Measures of uncertainty cannot be unique or objective, as the literature has shown (Julio and Yoox, 2012, Rich and Tracy, 2010, Rossi and Sekhposyan, 2015, Rossi et al., 2020, Altig et al., 2020). Thus, many insights about unobservable uncertainty can be drawn from the comparison of our *EURQ* index with the other indices proposed and most frequently used in the literature. Specifically, we select five uncertainty measures that represent the finance-, forecast- and news-based approaches. Like the *EURQ*, these measures are periodically updated and freely downloadable.

The finance-based approach uses information from the stock market (see, for example, Bekaert et al., 2013, Bloom, 2009, Gilchrist et al., 2014, and Knotek and Khan, 2011). Here, the assumption is that financial volatility can be a guide to the state of economic uncertainty, despite the fact that not everyone

[8] For example, a term that in future could enter our list is "helicopter money", which people in the USA started to use intensively in searches from mid-March 2020, when the COVID-19 pandemic begun to manifest its economic effects.

invests in the stock market (Romer, 1990) or shares the same information available to stock market actors. Within this approach, the uncertainty measure that we select is the CBOE Volatility Index (*VIX*) (Chicago Board Options Exchange, 2009), which reflects forward-looking volatility implied by 30-day options on the S&P 500 index. The *VIX* is used in many empirical studies (such as Bloom, 2009), but its ability to capture economic uncertainty is questionable, as it is based on stock market information only.[9]

Forecast-based indices estimate uncertainty by relying on the concept of economic predictability and on the measurement of disagreement across professional forecasters (Bachmann et al., 2013, Henzel and Rengel, 2014, Jurado et al., 2015, Ludvigson et al., 2020, Rich and Tracy, 2010, Rossi and Sekhposyan, 2015, Rossi et al., 2020, Scotti, 2016, and Segal et al., 2015). Here, the assumption is that a lack of predictability and/or disagreement across forecasters reflects a more uncertain economy. Within this approach, we select three measures derived from statistical models fitted on standard macroeconomic data. The first is the *SCOTTI* real uncertainty index related to the state of the economy (Scotti, 2016). The other two series are the macro/real (*MPRED*) and the financial (*FPRED*) components of the monthly macroeconomic uncertainty index of Jurado et al. (2015). As shown by Ludvigson et al. (2020), this decomposition is relevant because the two components are characterized by different degrees of exogeneity.[10] The uncertainty index *MPRED* represents macro uncertainty captured by the common component in the time-varying volatilities of 1-month-ahead forecast errors across many macroeconomic series from real activity. The uncertainty index *FPRED* represents financial uncertainty obtained with the same methodology as *MPRED* but based solely on numerous financial market series. The *SCOTTI*, *MPRED* and *FPRED* uncertainty measures come from computationally intensive statistical procedures, which are inevitably affected by lags in the availability of many real and financial data inputs. This complex process is a serious limitation on producing timely updates to series, and in this paper, the data in these indices are less up to date than those in the other indices.[11]

---

[9] According to Bekaert et al. (2013), the *VIX* is a mixture of uncertainty and risk-aversion, with the latter accounting for a sizeable part of the index. Whaley (2000) refers to the *VIX* as the "investor fear gauge"; Da et al. (2015) describe it as a market-based measure with the disadvantage of being the outcome of many economic forces. We used the monthly averages of daily *VIX* data downloaded from https://fred.stlouisfed.org/series/VIXCLS (Federal Reserve Bank of St. Louis Economic Data [FRED]). Given that the forward-looking CBOE measure starts in 1990, it is often backward estimated using the realized volatility of daily returns, which measures the variability of historical (or known) data; see, e.g., Bloom (2009).

[10] While the macro component is an endogenous response to other shocks that cause business cycle fluctuations during recessions, the financial component is an exogenous source of the fluctuations.

[11] Although, in principle, daily *SCOTTI* data can be updated every time new information becomes available, its latest vintage currently available ends on November 29th, 2019. We used the monthly averages of this vintage downloaded from https://sites.google.com/site/chiarascottifrb/research. The *MPRED* and *FPRED* have recently been updated to April 2020 (from the previous vintage that ended in December 2019) because of the increasing interest in uncertainty during the pandemic. Monthly data are downloadable from https://www.sydneyludvigson.com/macro-and-financial-uncertainty-

The news-based approach answers the question "How does the average citizen comprehend the implications of stock market volatility and economic predictability underlying her uncertainty?" as follows: "The media is the messenger" (Alexopoulos and Cohen, 2015). Here, the assumption is that journalists are likely to report on uncertainty by using specific words when certain causes of uncertainty matter. In other words, the media are assumed to be able to gauge the uncertainty indicated by market outcomes, professional economists, and political debate and to communicate it to the public through the recurrent use of specific words. The degree of uncertainty in each period is thus proxied by the frequency with which a lengthy list of words related to uncertainty appears in newspaper articles.[12] This approach leads to news-based uncertainty measures formulated, for example, by Alexopoulos and Cohen (2015), BBD, and Knotek and Khan (2011). Within this approach, we use the economic policy uncertainty (*EPU*) index of BBD.[13]

Figure 2 allows for visual comparison of the standardized temporal patterns over the common period 2004m1-2020m5 of the finance-based measure *VIX*, the forecast-based measures *SCOTTI*, *MPRED* and *FPRED*, and the news- and Internet-based measures *EPU* and *EURQ*. NBER downturns are shown by the shaded areas. As expected, quite a heterogeneous picture emerges.

Figure 2

While the *MPRED* and *FPRED* are smooth and clearly spike only in recession periods, the other four indices are affected by noisy fluctuations over time. Specifically, the *VIX* shows additional spikes outside downturns that follow financial market events, while the *SCOTTI, EPU* and *EURQ* indices are affected by short-run noisiness related to randomly occurring surprises and news. For example, in 2004-2005, the debate over social security produced an above-average increase in both the *EURQ* (as also noticeable in Figure 1) and the *SCOTTI* that is not evident in the levels of *VIX*, *FPRED* and *EPU* (which were below average). This debate did not represent either bad news or fear for financial markets; hence, it is not captured by the financial indices *VIX* and *FPRED*. Additionally, this debate was not perceived as relevant by journalists, as suggested by the low levels of *EPU*, while it both affected the macroeconomy (*SCOTTI*) and attracted the interest of economic agents (*EURQ*). This evidence supports the idea that the *EURQ*, by focusing on how individuals perceive situations rather than relying on how the media convey the issue, contains additional and valuable information

indexes (we label the time series h=1 inside the macro, real and financial uncertainty files as *PREDICT*, *MPRED* and *FPRED*, respectively).

[12] In a similar way, narrative analysis can identify monetary and fiscal policy shocks (see, for example, Romer and Romer, 2004, and Ramey, 2011).

[13] The monthly three-component *EPU* index is a weighted average of news coverage, tax code expiration, and disagreement between forecasts of inflation and public purchases. BBD use weights of 1/2 on the news-based component and 1/6 on each of the other measures of tax code expiration and disagreement. The latest available monthly *EPU* data are downloadable from https://www.policyuncertainty.com/us_monthly.html.

for measuring economic uncertainty. In Section 4, we deepen this discussion by comparing web searches and press coverage for specific search terms at both monthly and daily frequencies.

From March up to May 2020, all the series in Figure 2 show enormous uncertainty jumps in reaction to the COVID-19 pandemic and the economic fallout. However, the amplitude of their peaks differs substantially. The *VIX* and *FPRED* peaks occurred in March 2020 and are not significantly different from those related to the financial turmoil of 2008. The peak of the *MPRED* in March 2020 is 3-4 standard deviations higher than that of 2008, reflecting the nature of the COVID-19 shock: an extraordinarily massive common shock hitting all economic agents and hence different from the smaller and idiosyncratic shocks commonly occurring during recessions (see Altig et al., 2020). Additionally, the peaks of the *EPU* in May 2020 and of the *EURQ* in March 2020 are one standard deviation higher than those in 2008.

These facts suggest two points. First, the financial proxies *VIX* and *FPRED* peak relatively early but are likely not representative of the overall uncertainty spreading across individuals and companies, while the *MPRED*, *EPU* and *EURQ* reach their highest values since 2004 during the months of the pandemic, hence fully capturing the unprecedented and widespread uncertainty.[14] Second, the *EPU* continues to rise to peak in May (two months later than the *MPRED* and the *EURQ*), while the *EURQ* shows a substantial decline in May 2020 after the peaks of March-April. The impression is that over time journalists repeatedly cover events that cause uncertainty, like the COVID-19 outbreak, while individuals react very quickly to the same events and, after having converted their uncertainty to knowledge, reduce their queries. This point is deepened by analysing the persistence of the impulse responses in Section 3.2 and comparing the *EURQ*- and *EPU*-specific components in Section 4.

The univariate evaluation of the six uncertainty indices is based on the outcomes in Table 1.

Table 1 here

The first part of Table 1 confirms a well-known stylized fact (e.g., Bloom, 2014): all the uncertainty indices are counter-cyclical, and the ratios between their averages in downturn and upturn periods (as dated by the NBER) are always larger than one. Compared with the other indices, the *EPU* and the *EURQ* are less associated with the cycle and more affected by spikes capturing various episodes of uncertainty (the same also broadly occurs in terms of the standard deviation ratios over the cycle). Hence, the overall variability of the *EPU* and the *EURQ* is considerably less clustered over the cycle than that of the other indices, as already noted in the comments on Figure 2.

[14] On 26 August 2020, the IMF wrote that the disconnect between the performance of stock markets and the real economy—starker in the United States but also present in Europe—had become a topic of much interest and debate. For example, the large response by all the major central banks to the COVID-19 shock had a role in the lower sensitivity of stock markets (Caballero and Simsek, 2020).

To shed further light on the nature of these measures, we also check whether they are representative of ex ante or ex post uncertainty. A recent work by Rossi et al. (2020) shows that their uncertainty measure based on the forecast densities of the Survey of Professional Forecasters can be deconstructed into ex ante and ex post components. Ex ante uncertainty quantifies the amplitude of the future shocks expected by markets; hence, it is not affected by actual realizations of data but is a function of the standard deviation of the density forecast.[15] Ex post uncertainty, instead, quantifies the amplitude of shocks that actually occur; it includes the ex post realizations of data and, thus, depends on the misspecification of statistical predictions and the discrepancy between what agents expected and what happened. These ex ante and ex post measures of uncertainty broadly correspond to the concepts of forward- and backward-looking uncertainty of Altig et al. (2020).

The second part of Table 1 presents the correlations of the six uncertainty indices with the estimates of the ex ante and ex post uncertainty components provided by Rossi et al. (2020). The *FPRED*, *MPRED* and *SCOTTI* – which Altig et al. (2020) classify as backward looking – are mostly correlated with ex post uncertainty, while the *EPU* – which Altig et al. (2020) classify as forward looking – is mostly correlated with ex ante uncertainty. We also provide new evidence for the *VIX* and *EURQ* series. Although *VIX* dynamics should measure investors' forward-looking perceptions of S&P 500 volatility, they are instead more correlated with ex post uncertainty. We suggest that the *VIX* may just be a (smart) reflection of past realizations instead of the real forward-looking measure that it is generally claimed to be.[16] Finally, similar to the *EPU*, the *EURQ* is also a measure that is relatively more correlated with ex ante uncertainty. Hence, both the *EPU* and the *EURQ* highlight uncertain situations where the predictive density of agents becomes more spread out and when, accordingly, both journalists perceive more uncertainty and web searchers need more information.

Since the ex ante and ex post components of uncertainty are not orthogonal (the correlation coefficient is 0.61), the six uncertainty indices could also be correlated with each other, even if they embody relatively more ex post or ex ante uncertainty. In the third part of Table 1, we report the correlations computed over two periods: starting from 2004m1, the first period ends in 2019m11 to exclude the COVID-19 months, while the second period ends in 2020m5 to include the COVID-19 months. The outbreak of the pandemic could, in fact, have changed the correlations. The positive correlation coefficients support the idea that there is a sizeable degree of co-movement across the indices, although some differences emerge. As expected, the uncertainty measures based on financial data (the *VIX* and the *FPRED*) are the most correlated with each other and with the other indices, with

[15] Similar measures of ex ante uncertainty have been used by, e.g., Guiso and Parigi (1999) and Bontempi et al. (2010).
[16] It should also be noted that unavailable past *VIX* data are often backwards interpolated using the realized volatility (Bloom, 2009) which, of course, is closer to the ex post uncertainty. For example, in the VAR context of Rossi et al. (2020), the macroeconomic effects of the ex post component are like those of the *VIX*.

coefficients in the 0.5-0.8 range that are almost unaffected by the COVID-19 crisis. Only the correlations with the *EURQ* are lower than 0.5, as the *EURQ*, compared to the *EPU*, seems to be more correlated with macro uncertainty (the *MPRED* and *SCOTTI* when available) than with financial uncertainty. The simultaneous correlation between the *EPU* and the *EURQ* is, in fact, quite weak (0.09 up to 2019 and 0.18 up to 2020), suggesting the possibility of fruitful integrations between the views of newspapers and web searches in measuring different aspects of ex ante (forward-looking) uncertainty. When the period of the pandemic is included in the sample, the increases in the *EPU*/*EURQ*, *EPU*/*MPRED* and *MPRED*/*EURQ* correlations are noticeable: both journalists and web searchers seem to perceive promptly and clearly the widespread uncertainty shock due to the COVID-19 pandemic.

The last step of the evaluations is presented in the fourth part of Table 1 and aims to exploit the correlation between the indices by summarizing them in a single composite indicator obtained through principal component analysis (see, among the others, Haddow et al., 2013). We use the procedure of Solberger and Spanberg (2020), which implements a dynamic factor model in the state space representation to estimate – through the Kalman filter – the unobservable state of the first common factor across the six uncertainty measures. With this approach, we accommodate missing observations in some of the series for the COVID-19 period of 2004m1-2020m5. In the last three rows of Table 1, we report the correlations of each index with the first common factor. The positive and high correlations suggest the ability of the common factor to capture the common variation of all the indices as an aggregate uncertainty measure. When we extend the sample to the COVID-19 months, we do not find the expected increase in all correlations: they significantly increase only for the *EPU* and the *EURQ*, while the correlations with the other indices are approximately the same. This means that the COVID-19 outbreak increased the correlations between the common factor and the ex ante measures of uncertainty, confirming that the pandemic not only worsened ex post uncertainty but also increased ex ante uncertainty. Finally, to facilitate interpretation of the common factor as suggested by Stock and Watson (2002), the last row of Table 1 (labelled "share of each index explained") reports the R^2 of bivariate regressions of each uncertainty index against the common factor over the 2004m1-2020m5 sample. The results show that the common factor mostly explains the ex post financial uncertainty of the *VIX* and the *FPRED* and, to a slightly lesser extent, the ex post macro uncertainty of the *MPRED*, while it is less able to explain the ex ante uncertainty captured by the *EPU* and the *EURQ*.[17] Despite the recent increase in the correlation with the ex ante uncertainty

[17] The *SCOTTI*'s low R^2 may be related to the lack of data.

measures, the uncertainty proxied by the common factor can be broadly labelled ex post financial uncertainty and not as "general" uncertainty "common" to all the indices, as we might expect a priori.

3.2. Macroeconomic dynamics and the EURQ

Although the empirical literature suggests that uncertainty shocks exert a negative impact on economic activity[18], the evidence about the magnitude, relevance and persistence of this impact is discrepant and often depends on the measure adopted to proxy for uncertainty.

Regarding the magnitude and relevance of the effect of uncertainty shocks, Stock and Watson (2012, p. 81) claim that the shocks producing the 2007-2009 recession were primarily associated with a heightened degree of uncertainty (together with financial disruptions), and Ferrara and Guerin (2018) argue that labour market and credit variables are the indicators that react most negatively to uncertainty shocks. However, Born et al. (2018) show that increased macro and financial uncertainties can explain only up to 10% of the drop in GDP at the height of the Great Recession.

Regarding the persistence of the impact of uncertainty shocks, the heterogeneous picture that emerges from the literature can be summarized as follows. In a seminal contribution, Bloom (2009) sustains the overshooting effect on the real economy of a financial uncertainty shock (the "wait and see" dynamics): shocks generate a short-run drop in output (lasting for several periods) and a long-run overshoot. However, Bachmann et al. (2013, Figure 6) find different outcomes and suggest that Bloom's overshoot is due to the use of a finance-based measure rather than to genuine uncertainty effects. In addition, Choi (2013) and Beetsma and Giuliodori (2012) show that the impact on real activity of stock market volatility shocks is not robust over time.[19] This first set of conflicting findings goes against the use of only financial information to proxy for uncertainty, since certain transitory financial crises and other random events could be mistaken for uncertainty shocks.[20]

However, the evidence obtained with other uncertainty measures is still mixed. With forecast-based measures, Jurado et al. (2015) and Bachmann et al. (2013) show that the dynamic response to uncertainty shocks is a sharp reduction in output in the short run, with effects that persist in the long run (i.e., for more than 4-5 years after the shock). With their news-based economic policy uncertainty measure, BBD find that uncertainty shocks induce negative dynamic responses of output only in the short run, as output responses are significantly negative only in the first 15-18 months, as in Bloom (2009), and then vanish (without overshooting). A possible explanation of the permanent effects

[18] Within the context of DSGE models, see, e.g., Mumtaz and Theodoridis (2018) and Leduc and Liu (2016).
[19] Further, Jurado et al. (2015) argue that Bloom's overshooting is a data relic attributable mainly to his HP filtering of uncertainty, as these dynamics vanish in the raw data.
[20] This point is related to Carriero et al. (2015), where financial uncertainty shocks are modelled in the context of VARs with measurement errors in uncertainty. More details on this appear below.

found by Jurado et al. (2015) and Bachmann et al. (2013), in contrast to the temporary effects of BBD, could be related to their sample periods, which are longer than that of BBD, as they start in the 1960s, while the BBD sample starts in the mid-1980s. The shorter BBD sample excludes all noisy shocks due to events occurring before the Great Moderation and is more heavily influenced by the Great Recession sub-period, when large financial shocks were not simply feeding through the usual dynamics (Sims, 2012).

To further our understanding of the role of in-sample events and how they are depicted by alternative uncertainty indicators, we run a global comparison of output responses to shocks of *alternative* uncertainty measures (including our *EURQ*), but estimated within a *common* empirical framework, over *common and updated* monthly spans, and measuring the uncertainty impulses in two ways: with the customary one standard deviation of the uncertainty shocks (to ease comparison with much of the literature) and with an increase in uncertainty caused by a *common* event: the Lehman bankruptcy. This mode of comparison has clear advantages: it can encompass the bulk of evidence in various papers about the role of uncertainty in economic activity and prevents outcomes from being affected by the specific events that occurred during different periods of index availability.

As is typical in the literature, we use reduced-form VAR models as the common empirical framework to capture many macroeconomic channels without imposing many parameter restrictions. Regarding the specification, we select, from a wide range of options[21], a fixed five-variable VAR, summarized as in model (1):

$$\begin{bmatrix} z_t \\ x_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} B_{11,1} & B_{12,1} \\ B_{21,1} & B_{22,1} \end{bmatrix} \begin{bmatrix} z_{t-1} \\ x_{t-1} \end{bmatrix} + \cdots + \begin{bmatrix} B_{11,p} & B_{12,p} \\ B_{21,p} & B_{22,p} \end{bmatrix} \begin{bmatrix} z_{t-p} \\ x_{t-p} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \qquad (1)$$

In model (1), the uncertainty measure z_t is alternatively proxied by seven variables: the forecast-based macro (*MPRED*) and financial (*FPRED*) uncertainty indices proposed by Ludvigson et al. (2020), the news-based policy uncertainty index (*EPU*) of BBD, Scotti's (2013) uncertainty measure (*SCOTTI*), the *VIX*, our index based on Internet searches (*EURQ*), and the macroeconomic uncertainty index (*PREDICT*) of Jurado et al. (2015). The four-variable vector x_t is always the same across different levels of z_t and is defined as $(sp_t, ff_t, emp_t, ipman_t)'$, where *sp* is the S&P 500 index in logs, *ff* is the log of one plus the federal funds rate, *emp* is manufacturing employment in

[21] On the one hand, small bivariate VARs with only uncertainty and output, like those in Bachmann et al. (2013) and Scotti (2016), offer the advantage of parsimony but are subject to biases due to the omission of relevant macroeconomic channels. On the other hand, large VARs, like the 11-variable specification in Jurado et al. (2015), offer the advantage of a satisfactory theoretical basis but suffer from inefficient estimates due to the curse of dimensionality, exacerbated in our case by the short span of data availability for the *EURQ*.

logs, and *ipman* is the manufacturing production index in logs. These variables, and their order, are the same as those in BBD.[22]

Given that the orthogonal shocks originating from the impulse-response functions are recovered by means of a Cholesky decomposition, the order of the variables is relevant for identification. Our selected order in VAR model (1) implies that the uncertainty shocks impact all the other variables in the first period, while uncertainty is assumed to be contemporaneously exogenous to the shocks in the other variables, an assumption that is coherent with the exogeneity of uncertainty found in Carriero et al. (2018). [23]

The first three plots of Figure 3 show the dynamic responses of output (*ipman*) to one standard deviation of the uncertainty shocks obtained from VAR model (1), with parameters estimated over three alternative time spans, the longest span being the 1963-2019 period, the medium-length span the 1985-2019 period, and the shortest span the 2004-2019 period, and using all the uncertainty indices for which the data are available. Hence, over the 1963-2019 sample, only the *VIX*, *FPRED*, *MPRED* and *PREDICT* are used; over the 1985-2019 sample, we can add the *EPU*; and over the 2004-2019 sample, we can use all the measures, the *EURQ* and *SCOTTI* included. In the fourth plot of Figure 3, we add the impulse response function of output when uncertainty is measured by means of the single composite indicator (labelled *FACTOR*) obtained through the principal component analysis described in Section 3.1.

Figure 3 here

Over the 1963-2019 temporal span (the first plot of Figure 3), shocks to the forecast-based *MPRED, FPRED* and *PREDICT* gradually reduce output, with initial responses monotonically decreasing up to 18-20 months after the shock before stabilizing thereafter at statistically significant levels. In the long run, a one-standard-deviation innovation in the *FPRED* or the *PREDICT* entails an output loss of approximately 1%, one in the *MPRED* of approximately 0.6% and one in the *VIX* of approximately 0.5%. This outcome is in line with Jurado et al. (2015, Figures 6 and 7)[24], demonstrating that independent of the VAR size, the use of either forecast- or finance-based indicators finds that uncertainty shocks are significant and persistent determinants of output fluctuations. However, this outcome changes if we measure output dynamic responses to uncertainty shocks over shorter sample periods.

[22] The variables are also the same as those in Rossi et al.'s (2020) quarterly VAR, with the sole exception of GDP instead of manufacturing industrial production. In Rossi et al. (2020), the order of the variables is the same as in Jurado et al. (2015), except for uncertainty, which Rossi et al. (2020) put first.
[23] In Appendix A2, we report some robustness checks based on three alternative VAR identification schemes.
[24] Our standard error bands are narrower because our VAR is more parsimonious.

The switch from the long (1963-2019) period to the medium-length (1985-2019) period (in the second plot of Figure 3) does not relevantly affect the short-term output responses (in the range of -0.4/-0.9%), but it weakens them after the initial downturn, with the sole exception of the output response to the *VIX* shock, which is persistent in the long run. The exclusion from the sample of the noisy real shocks of the 1960s and 1970s makes the dynamic impact of uncertainty shocks on output not qualitatively much different from the one obtained using the *EPU*, which shows a decline in production over the first year and then a return to pre-shock conditions (in line with the findings of BBD).

The estimates over the shorter (2004-2019) period confirm that changing the sample period affects the dynamics of output responses much more than the use of different uncertainty indicators. In fact, all the output responses in the third graph of Figure 3 display negative short-run effects that peak at -0.4/-1% after 7-15 months, depending on the measure of uncertainty. However, after the peaks, a clear and rapid output recovery is always evident, sometimes with a tendency to overshoot. In general, the output responses to uncertainty shocks estimated over the shortest sample are no longer significant after their negative short-run peaks.

In the fourth graph of Figure 3, we deepen the 2004-2019 analysis. Differently from the *EURQ* and *EPU*, *FACTOR* shocks lead to output dynamics that are close to those obtained with the ex post measures of uncertainty (the *VIX*, *MPRED*, and *FPRED*). This evidence corroborates the finding in the final part of Table 1 that *FACTOR* – despite being an indicator summarizing *all* our uncertainty measures – is more driven by the ex post uncertainty measures. Additionally, these outcomes confirm Rossi et al. (2020) in finding that the ex post measures have a higher impact in magnitude than ex ante uncertainty.

Although the size of the one-standard-deviation shocks is quite substantial,[25] it is not possible to link these "statistical" magnitudes to past events and assess the relevance of specific uncertainty-carrying episodes in explaining output fluctuations. For this purpose, in Figure 4, we focus on the Lehman bankruptcy case, and we show the impact on the output of uncertainty shocks equal to the difference between the average of each uncertainty measure in September–November 2008 and that in June–August 2008. Given that we estimate the size of the Lehman bankruptcy shocks for each uncertainty measure, the scale of the output responses in Figure 4 provides alternative estimates of the effect of the Lehman bankruptcy on economic activity over alternative uncertainty indices, while their dynamics over time are the same as discussed above.

[25] Considering a close to normal distribution of the shocks of each measure over time, shocks larger than one standard deviation can occur with a 15% probability only.

Figure 4 here

The output responses to shocks in the forecast-based measures (the *FPRED*, *MPRED* and *PREDICT*) show the strongest impact on economic activity (in a wide 5-15% range), while the effects of the other uncertainty measures (the *EPU*, *SCOTTI* and *EURQ*) are significantly lower, always below 5% (the *VIX* effect stays at an intermediate level). The fourth graph in Figure 4 stresses that the assessment of the magnitude of the Lehman bankruptcy shock is larger with the ex post uncertainty measures (summarized by *FACTOR*) than with the ex ante measures (the *EPU* and *EURQ*). Hence, Figure 4 highlights the large heterogeneity of outcomes attested to in the literature and suggests that the assessment of the macro relevance of a single event is dominated by the way in which each uncertainty proxy quantifies that event.[26] In short, the spread of the estimated magnitudes across uncertainty measures is so wide that any other factor can play only a minor role in describing the macro fluctuations.

Taken together, the outcomes in Figures 3 and 4 suggest that the assessment of magnitude and relevance of the impact on economic activity of episodes of high uncertainty is strictly specific to the measure used to proxy for uncertainty, while the persistence over time of this impact is strongly related to the sample period over which the model is estimated, with a minor role left to the single uncertainty measures. Given that different historical events can be incorporated or excluded in each sample period, their number and nature affect the output responses.[27]

On the one hand, the 1963-2019 temporal span, comprising an era of substantial, noisy real shocks (in the 1960s and 1970s), reveals the significant long-term effects of uncertainty on output (regardless of whether it is measured with a forecast- or finance-based indicator). On the other hand, the 2004-2019 temporal span is almost entirely centred on large and noisy financial shocks (those that occurred during the Great Recession) and reveals strong short-term effects that quickly (in approximately a year) affect economic activity and are then followed by a period of recovery.

On the basis of two recent strands of research, we provide two possible explanations for this outcome. The statistical view suggests that different sample periods cover different events whose shocks are measured by uncertainty proxies that capture a mixture of genuine uncertainty signals and noise (measurement errors). The economic view suggests that different policy and structural events in the past have induced large changes in reduced-form VAR parameters.

According to the statistical view, measurement errors create an attenuation bias in the responses of macro variables to uncertainty shocks, while accounting for measurement errors in

[26] The estimated magnitudes of the Lehman bankruptcy shock with our proxies are the following: 14 times the average monthly increase in *PREDICT*, 7 times that of *MPRED*, 11 times that of *FPRED*, 6 times that of log(*VIX*), 4.5 times that of log(*EPU*), 4 times that of *SCOTTI*, 5 times that of log(*EURQ*), and 13 times that of *FACTOR*.

[27] A similar effect is documented in a different context by Rossi (2006).

uncertainty proxies—as in Carriero et al. (2015)—produces a larger and more persistent estimated impact of financial uncertainty than that estimated by Bloom (2009).[28] Therefore, the extent of the bias in the estimate of the impact of uncertainty depends on the ratio of the signal conveyed by the uncertainty proxies (which depends on the size and composition of the shocks) to noise (i.e., measurement errors). Noise is probably less important over our large sample period (1963-2019) because the amount of signal brought to the cycle by the shocks of the 1960s and 1970s should prevail over any measurement errors. Therefore, in the long sample, the impulse-response patterns are only slightly downward biased. Instead, measurement errors are probably more relevant over our short sample period (2004-2019) because noisy financial shocks should jeopardize the signal brought by any uncertainty measure. Therefore, in the shorter (2004-2019) sample, the impulse-response patterns are significantly downward biased and less persistent.

According to the economic view, our evidence of decreasing persistence of the output response to uncertainty shocks from the long to the short sample, coupled with the FAVAR evidence in Mumtaz and Theodoridis (2018) of the decline over time of the effect of uncertainty shocks, can be explained in the context of a DSGE model only by allowing for changes in parameters: (1) an increase in the Federal Reserve's anti-inflationary stance (in line with the shift before and after Volcker's appointment as Fed chairman in 1979; see Clarida et al., 2000) and (2) a change in the parameters of the Phillips curve, implying a rise in price stickiness and a fall in indexation to past inflation (see Stock and Watson, 2007, and Cogley et al., 2010).[29]

Finally, our results also extend those in Scotti (2016) and Caggiano et al. (2014, 2017). Based on a bivariate VAR exercise with employment (instead of output) and alternative uncertainty measures over the 2003m5-2016m3 period (very close to the span of our short sample), Scotti (2016) finds that the uncertainty measures strictly related to real activity (like her real-activity uncertainty index, the *SCOTTI*) produce macro responses to uncertainty shocks that are weaker than those produced by indices related to the stock market. Although our results over the shortest sample in Figure 3 support Scotti's view (the effects are stronger with the *VIX* and *FPRED* and weaker with the *EPU*, *EURQ* and *SCOTTI*), the picture emerging over the longest sample is completely different: the *VIX* effect is the weakest, while the financial *FPRED* and macroeconomic *PREDICT* effects are very close and stronger.

[28] See the last two plots of Figure 3 in Carriero et al. (2015). Measurement errors are accounted for in their proxy SVAR by using instrumental variables. Although interesting, the implementation of their approach faces the difficulty of finding valid instruments to identify genuine macro responses, as evidenced by Stock and Watson (2012) and the ensuing strong criticism of their results.

[29] This view squares well with the results of Choi (2017) and Berg et al. (2018).

By using a smooth transition VAR with different parameters in recession and expansion phases, Caggiano et al. (2014, 2017) find that the response of output to uncertainty shocks is greater during recessions. Since the sample for our short period is permeated by Great Recession data, our results in Figure 4 are in line with the "wait and see" output dynamics reported in Figure 6 of Caggiano et al. (2014).

3.3. A case study: the EURQ index for Italy

Key factors for a successful country-specific EURQ index: the number and adequacy of the search terms. Given the multifaceted nature of uncertainty, a high-quality *EURQ* index must be based on an accurate list of several clearly explicable search terms specific to the country under analysis. The construction of the *EURQ* index for Italy is based on the same approach used in the American case, taking the Italian equivalents of the 210 BBD terms used to construct the *EURQ* index for the USA. More specifically, starting from the BBD list, the definition of an adequate Italian list only partly relies on the translation of terms from English to Italian, as what truly matters is careful consideration of the features and facts that directly concern the economy and politics in Italy. The following examples may clarify this point. While, for example, the terms "terrorism" and "taxation" have been translated into the Italian "terrorismo" and "tassazione", the English term "WTO", for "World Trade Organization", has not been translated, as it is a term of common use in Italy and is more popular than the Italian acronym OMC, "Organizzazione Mondiale del Commercio". Moreover, the term "collective bargaining law", which in the USA refers to negotiations between an employer and a group of employees to determine employment conditions, has been substituted by the term "CCNL", the Italian acronym for "contratto collettivo nazionale di lavoro" ("national [collectively bargained] labour agreement"), which is a national agreement between trade unions and employers specific to Italy.

This assessment process led to a final list of 136 search terms (see Appendix A1) used to create the *EURQ* index for Italy, May 2020 vintage, depicted in the annotated graph of Figure 5.

Figure 5 here

The *EURQ* for the Italian data shows an overall increasing trend with a break in correspondence with the European debt crisis (2011); spikes due to political disputes about the labour market and severance pay reforms (2012m6 and 2015m3); and high uncertainty around the calling of elections and, specifically, the constitutional referendum dispute (in 2016m11). The massive spike due to COVID-19 is noticeable also for Italy, even though specific terms related to the pandemic are not included in the list of queries. In April 2020, the index shows an increase of over 70% against its

levels in both April 2019 and February 2020. This increase is the largest in the series, well above the past peaks due to the Great Recession, the Jobs Act, and the constitutional referendum.

In Figure 6, we compare, over the common period 2004m1-2020m5, the Italian *EURQ* with the three Italian equivalents of the finance-, forecast- and news-based indicators: the volatility index of the Italian stock market, the *SVOL;* the macroeconomic uncertainty index, the *MUI*, of Jurado et al. (2015); and the *NEWS* index for Italy.[30] To extend the number of indices made for the Italian case, we added four other search-based measures. We compute the first, labelled the *EPUGT*, by using those queries corresponding to the 9 terms of BBD to obtain their *NEWS* index for Italy. The other three indices are the Italian versions of Google-based measures proposed in the literature: the *GSI* of Donadelli (2015), the *ECON* of Dzielinski (2012) and the *UI* of BBVA (2012) obtained using 3 terms, only 1 term and 15 terms, respectively.

Figure 6

The comparison of the *EURQ* with these latter four (Google) search-based indices confirms the importance of query selection and the danger of using too few queries. Disregarding **C1** produces measurement errors, unreliability, and odd temporal paths. For example, *ECON* and *UI* show a continuous decline from their initial creation in 2004; a trend is also observed in the *GSI* except for some very high, and unaccounted for, spikes in the middle of the period; the *EPUGT* does not exhibit any increase in 2020 due to the COVID-19 crisis, while the 2004-2006 period emerges as the most affected by uncertainty.

Macroeconomic dynamics and the EURQ for Italy. We estimate the impulse response functions of output to uncertainty shocks measured by alternative indices. To enhance comparability with the USA case in Section 3.2, we use a VAR similar to that in equation (1): z_t is uncertainty, proxied in turn by the *EURQ*, *SVOL*, *MUI* and *NEWS* indices; the four-variable vector x_t includes the benchmark Italian stock exchange index in logs, the log of one plus the Euribor rate, and manufacturing employment and production, both in logs. The impulse response functions are plotted in Figure 7.

Figure 7 here

Following a shock in the *EURQ*, industrial production shows an immediate decline, followed by a rapid recovery without overshooting. On the other hand, a shock in either the *MUI* or the *SVOL* yields negative mid-term effects followed by a slow recovery. Finally, a shock in the *NEWS* involves a puzzling short-run increase in industrial production, followed by a sequence of negative responses

[30] Both the *SVOL* and the *MUI* for Italy come from Meinen and Roehe (2017). The *NEWS* for Italy is downloadable from BBD's economic and policy uncertainty index web page, http://policyuncertainty.com/europe_monthly.html.

that decline in the medium term. This odd pattern of output responses can be ascribed to measurement biases due to the few newspapers and, especially, the few search terms used, as confirmed in Figure 6 by the *EPUGT,* the search version of the *NEWS*. Overall, as for the USA case in Figure 3, the ex post uncertainty measures (the *MUI*, macroeconomic, and the *SVOL*, financial) induce a less sharp production decline in the short run than the *EURQ*, while their lowest troughs occur after approximately 2 years, when the *EURQ* effect is approximately zero.

4. The *EURQ* as a means of assessing how specific uncertainty components are perceived

Pew reports that "While the Internet is growing as a news platform, it has not displaced completely offline news sources for most American adults: 59% of Americans get news from a combination of online and offline sources on a typical day. Just over a third (38%) rely solely on offline sources, while just 2% rely exclusively on the Internet for their daily news" (Pew Research Center, 2016). Given this description, the relationship between press coverage and the information that people want to obtain from the web can offer important insights into what type of uncertainty is perceived by economic agents *before* the press begins to divulge specific news and, conversely, what kind of news *attracts* the interest of people and encourages them to look for information online. Whether a specific type of uncertainty is a harbinger of fear, about which related searches are conducted on the web even before press mentions start, will certainly have repercussions in terms of economic policy.

The present section thus offers a comparison between our *EURQ* index and the BBD's *Newsbank* series at both monthly and daily frequencies.[31]

4.1 – A monthly check of web searches against press coverage

The basic ingredients of our experiment are two sets of series measuring searches for the same terms belonging to policy category c: $Newsbank_{ct}$ (BBD news-based counts) and $EURQ_{ct}$ (Google Trends search volumes), where c = 1, 2,…, 8 indicates the policy categories "Fiscal policy" (**FP**), "Monetary policy" (**MP**), "Health care" (**HC**), "National security and war" (**NS**), "Regulation" (**RE**), "Sovereign debt and currency crises" (**SDCC**), "Entitlement programmes" (**EP**), and "Trade policy" (**TP**), and t is monthly observations. Although referring to the same search terms, $Newsbank_{ct}$ and $EURQ_{ct}$ capture two different aspects: according to $Newsbank_{ct}$, journalists are the messengers of uncertainty, which they convey by using specific words; according to $EURQ_{ct}$, web users are the ones

[31] Some empirical research (e.g., Eberth et al., 2014) has focused on modelling the ways that information is divulged via the Internet and the speed at which this happens, albeit on topics not concerned with assessing the role that uncertainty plays in generating people's interest and concern.

who manifest their interest/uncertain mood by searching more/less intensively for the same words used by newspapers.

The dynamic relationship between $Newsbank_{ct}$ and $EURQ_{ct}$ can be assessed within the context of the VAR model. Let us suppose that for the cth 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 $\Sigma_{\varepsilon c}$:

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

where, in our bivariate context, k=2 and m=1, $y_{ct} = (Newsbank_{ct}, EURQ_{ct})'$ is the vector of the uncertainty indices for the cth category (therefore, $z_{ct} = Newsbank_{ct}$ and $x_{ct} = EURQ_{ct}$), scalars μ_c and matrices A_c are heterogeneous parameters (they are allowed to differ across categories), and $\varepsilon_{ct} = (\varepsilon_{c1t}, \varepsilon_{c2t})' = \left(\varepsilon_{ct}^{Newsbank}, \varepsilon_{ct}^{EURQ}\right)'$ is the vector of the random shocks to the *Newsbank* and *EURQ* uncertainty measures for c.

The analysis conducted using VAR system (2) relies on two basic concepts: Granger causality and contemporaneous causality. Granger causality involves the assessment of the null hypotheses $A_{c12,i} = 0$ for i = 1, 2, ..., p (which implies that *Newsbank* is not Granger-caused by *EURQ*) and $A_{c21,i} = 0$ for i = 1, 2, ..., p (which implies that *EURQ* is not Granger-caused by *Newsbank*).

Although we must be very careful when interpreting the outcomes of statistical tests in behavioural terms, Granger causality from *Newsbank* to *EURQ* for a certain category c implies that past news-based shocks are related to present web searches: past newspaper headlines lead people to seek further knowledge about c after the news shock has occurred. In other words, the news-based measure of uncertainty category c – informing people about what is happening now – drives web searches that, over time, propagate following their own dynamics $A_{c22,i}$. We label this case "news-pooled" uncertainty.

Conversely, Granger causality from *EURQ* to *Newsbank* for a certain category c can be explained as if journalists feed readers' constant need for information regarding subject c and continue to satisfy that need in their newspapers. In this second case, web searches – signalling readers' interest in c – "drive" the news-based measure of this category. We label this second case "query-driven" interest.

Given that the VAR residuals are not orthogonal (the covariance matrix $\Sigma_{\varepsilon c}$ is not diagonal), the presence of a significant contemporaneous correlation between *Newsbank* and *EURQ* shocks, $E(\varepsilon_{c1t}\varepsilon_{c2t}^{'}) \neq 0$, means that in addition to possible Granger causality in one direction or the other,

the two measures of uncertainty for category *c* are coincident: the news and web search shocks can also occur in the same month.

The results from the VAR system (2) are summarized in the upper part of Table 2, where two columns and three rows delimit six areas (cases) containing the 8 categories. The columns are used to classify the categories as cases of high/low contemporaneous correlation (degree of coincidence), depending on whether the value of this correlation is higher or lower than 0.25 (the level denoting 1% statistical significance). The 8 categories are classified into three cases along the rows: the case of news-pooled uncertainty (when *Newsbank* Granger-causes *EURQ*), the case of query-driven interest (when *EURQ* Granger-causes *Newsbank*), and finally the case of no-dynamics-uncertainty interest (when Granger causality is not statistically significant in either direction).

Table 2 here

"Fiscal policy" (**FP**) and "Sovereign debt and currency crisis" (**SDCC**) are carefully monitored by people: the number of web searches for such terms increases as soon as shocks occur, even if newspapers do not give the same importance to them. Internet activity and newspaper mentions overlap significantly, with a high contemporaneous correlation. "Health care" (**HC**) also leads news-based uncertainty but with a lower contemporaneous correlation. When investigating the individual search terms within query-driven uncertainty, we find that "Debt ceiling" and "Government deficits" are the most relevant search terms inside the **FP** category, while **SDCC** searches are mostly accounted for by the term "Sovereign debt" (although "Currency devaluation" and "Euro crisis" also play a significant role). Finally, the **HC** result is driven mainly by the search term "Affordable Care Act".

The "Monetary policy" (**MP**) and "Trade policy" (**TP**) categories are news pooled: people start searching the web for more information about these categories after the newspapers have begun to mention them. The "Regulation" (**RE**) category behaves rather similarly, albeit at a considerably lower level of contemporaneous correlation, denoting quite unrelated dynamics of web searches and news for this category.

Finally, the "Entitlement programmes" (**EP**) and "National security and war" (**NS**) categories do not display Granger causality in either direction. However, they behave differently in terms of the degree of contemporaneous correlation. In fact, for **EP**, the *Newsbank* and *EURQ* are highly correlated, denoting in the same month a substantial overlap of press reports and Internet searches. On the other hand, for **NS**, the correlation is low, suggesting that the need for knowledge feeding web searches is not related to newspaper headlines, as readers are already aware of the matter in question (on the terrorism issue, see the recent Jetter, 2019).

4.2 – An intra-daily check of web searches against press coverage

Within a month, it is highly possible that the daily web searches of individuals are caused by what they hear on the news. According to Enke (2020), people form beliefs by consuming what is going on around them, and the news media clearly play a large role in information dissemination. Hence, we also investigate whether the *EURQ* measures individuals' genuine interest or instead just curiosity triggered by press coverage of certain topics related to specific daily events. Following an approach like that of Piffer and Podstawski (2017), who argue that the gold price can be used as a proxy for uncertainty by looking at news coverage of certain events and how stock markets react around the same time, in Figures 8a and 8b, we look at daily search behaviour and press coverage for some relevant terms around selected events. We use data from the *New York Times* search API[32], which can be accessed free for non-commercial uses, as a proxy for general daily news coverage. Of course, caution must be exercised in interpreting the results, as the representativeness of news data is related to the specific newspaper that we select to count words.[33] The correlation between news and the intensity of searches through which Internet users manifest their interest is also related to the circulation of the newspaper and the regularity with which it is read.

Figure 8a here

The first two plots of Figure 8a investigate the query-driven "Health care" (**HC**) category through illustrative examples regarding the query "Affordable Care Act" from June 1 to December 27, 2013, and from September 1 to March 30, 2017. Although some articles in the *New York Times* discussed the issue, people's attention was not caught during the June-August 2013 and September 2016 periods. Instead, web searches rose rapidly and peaked in September 2013 during the protracted standoff over the Affordable Care Act that culminated in the federal government experiencing a lack of funding that forced it to shut down on October 1, 2013. Web searches also anticipated the news in October 2016 just before the presidential elections and in January 2017 before the Senate voted to pass a budget resolution to repeal the Affordable Care Act. This is an example of a topic carefully monitored by people, for which the newspapers do not impact web searches; rather, economic agents start to search when they perceive a change in the situation, even when journalists do not give much weight to the shock.

The second two plots of Figure 8a illustrate the examples "Dodd-Frank" and "FDIC" (Federal Deposit Insurance Corporation) within the "Regulation" (**RE**) category. During the promotion of

[32] https://developer.nytimes.com/. The API offers searchable access to articles from the *New York Times*, retrieving headlines, abstracts, number of words, authors, and other attributes.

[33] The *New York Times* is reported to be a "newspaper of record" with a large circulation (among the top three newspapers in the USA) and authoritative editorial functions. Groseclose and Milyo (2005) estimate the media bias for several major media outlets and classify the *New York Times* as a newspaper with a balanced perspective.

changes to the Dodd-Frank law (Chon, 2016) in the pre-electoral period of 2016, the *New York Times* published articles mentioning the term, which produced the reaction of people searching for additional information online. The same was even more evident in February 2017, after a White House meeting with executives from Wall Street during which the president signed a directive on the Dodd-Frank Act (Protess and Hirschfeld, 2017). The *New York Times* published news referring to the term 'Dodd-Frank' and triggered a considerable increase in corresponding queries on the Internet. Additionally, for the term "FDIC", searches peaked after federal regulators seized IndyMac Bank on 12 July 2008 and the *New York Times* reported on this news. A similar situation is observed for the 30th of September 2008, the day after the stock market crashed and the House of Representatives rejected a bailout package. This news increased uncertainty and worries about the strength of the banks among depositors.

As another example of news-pooled uncertainty, in the last two plots of Figure 8a, we present the query "Interest Rates" within the "Monetary policy" (**MP**) category in relation to the interest rate increases by the Federal Reserve from September 2016 to March 2017 and the interest rate cuts by the Federal Reserve from October 2019 to April 2020. The *New York Times* published many articles on the topic, particularly in 2020 regarding the Fed's possible future measures to fight the recession sparked by the COVID-19 crisis (see, for example, Irwin, 2020). In 2016, web users' attention was caught only after the 14th of December. In 2020, searches increased particularly after the 15th of March, when, having cut its benchmark interest rate by 50 basis points in a surprise move on the 3rd of March, the Federal Reserve further cut rates to zero and launched quantitative easing programmes. A common feature of the news-pooled case is that unlike journalists who repeatedly cover events producing uncertainty, individuals' one-time attempt to convert their uncertainty to knowledge is enough. Of course, the life span of agents' interest depends on how much the issue or situation reported on by the press could change their lives and, hence, how much it triggers their uncertainty.

Figure 8b here

Finally, the plots in Figure 8b support the evidence of how searches tend to move independently from the news in the case of the "Entitlement programmes" (**EP**) category through the analysis of the specific queries "Unemployment benefits" during both the 2008 and COVID-19 crises and "Food stamps" during the pandemic. Irrespective of the Unemployment Compensation Extension Act in 2008, the American Recovery and Reinvestment Act in 2009 and the Coronavirus Aid, Relief and Economic Security Act in 2020, and independent of the messages conveyed by the press, many workers who probably had become unemployed began to seek information on subsidies and other government payments. The duration of their searches, even during periods when the *New York Times* did not report any news, confirms that Internet searches capture the increased uncertainty at the time

of the Great Recession in 2008 and the COVID-19 crisis in 2020 and reflects the need for more information from households and businesses regarding public measures to combat the effects of the shock. The low correlation between news and searches for the "Food stamps" query during the COVID-19 crisis is noteworthy, and although the correlation for "Unemployment benefits" is high, this does not mean that the web searches were triggered by the press.

Overall, Figures 8a and 8b provide evidence that web searches react to uncertainty at daily frequency. Moreover, even though some queries are likely influenced by news coverage, some other searches on the Internet exhibit behaviour independent of the press. The different perspectives provided by web searches (information consumption) and news articles (information production) support the idea that the *EURQ* can add useful information to better assess the level of uncertainty perceived by economic agents and to measure individuals' genuine interest in specific economic policy topics.

5. Results and discussion

Our results can be discussed in relation to our paper's two main aims. The first aim is to use Internet searches to make a new *EURQ* index measuring the volumes of "economic uncertainty-related queries". The *EURQ* is based on the effective behaviour of all economic agents and represents their need for information when they are concerned and uncertain about political and economic events. Being based on people's moods, the *EURQ* can quantify additional important qualitative aspects of uncertainty that are not easily accounted for in the uncertainty indices proposed in the literature. The literature shows that each uncertainty measure has its own dynamic effects on output: finance-based uncertainty induces overshooting effects, forecast-based uncertainty induces highly persistent effects, and news-based uncertainty induces transitory effects. These differences in the output responses to uncertainty are less marked if they are obtained from common time spans and VAR models, and our evidence supports the idea that the different uncertainty measures produce different magnitudes of output responses to shocks simply because each of them accounts for historical events differently. All the uncertainty proxies are measured with errors (Carriero et al., 2015), but the period-specific signal-to-noise ratios affect the finance-based and forecast-based indices more than the news-based index and our *EURQ*, while the search-based indices consisting of few terms are heavily biased. Additionally, the events occurring in certain periods imply structural changes in the parameters establishing the role of uncertainty (Mumtaz and Theodoridis, 2018). We show that news- and search-based measures have great practical relevance and appeal in comparison with other measures. They are ex ante measures of uncertainty that are able to capture widespread uncertainty shocks in a timely manner. They are model-free and hence able to better track uncertainty

if their underlying data-generating process is highly non-linear (while finance- and forecast-based indices are model-based measures necessarily implying smooth and rigid approximations). They can track concepts such as political and economic uncertainty and geopolitical risk that are broader than those captured by the other indices (for example, finance- and forecast-based measures are available for only a limited number of variables). Specifically, the *EURQ* index improves our understanding of the heterogeneous nature of uncertainty since the queries it uses could be differentiated into specific components of uncertainty and computed for different geographical zones within the same country.

The second aim of our paper is to establish whether the interest manifested by economic agents is driven by certain events and triggered by specific news or whether there are types of uncertainty that are able to generate the spontaneous interest of economic agents. The identification and measurement of specific types of uncertainty is extremely important since the output responses to uncertainty are influenced by the nature (and intensity) of the shocks that occurred. The joint analysis of news-based uncertainty and the *EURQ* suggests that distinct categories of economic and policy uncertainty entail alternative dynamic relationships between newspaper headlines and web searches. Topics relating to taxes, health and economic crises induce spontaneous and conscious interest regardless of any stimulus from the press. This suggests that macro-real uncertainty may manifest its impact on economic variables at a very early stage. On the other hand, topics only affecting people's lives after changes have been made to rules/regulations and monetary/foreign policies tend to stimulate the interest of economic agents only after the press has reported on such changes and journalists have driven the public's general attention towards such issues. This suggests that financial shocks are amplified and produce a more pronounced reduction in economic activity because newspapers can feed the worries of agents and journalists' intervention produces the multiplicative effect of the web. Extremely important issues concerning employment and terrorism simultaneously generate the interest of both web users and the press, without the one influencing the other.

5.1 Conclusions

Economic uncertainty embodies several unobservable components. Thus, it is difficult to fully quantify it by using any specific measure based on only a few of the aforementioned components. The heterogeneous nature of uncertainty is substantiated by the literature, which proposes a variety of different indicators: finance-, forecast- and news-based. In short, measuring uncertainty is a very uncertain activity. Moreover, we do not know how and when people perceive uncertainty and which components of uncertainty may have a strong impact on economic agents. "The measures of uncertainty tend to combine economic uncertainty with other notions. For example, stock return volatility combines information about stock market volatility with economic uncertainty and forecast

disagreement could measure a divergence of opinions among forecasters rather than just the underlying uncertainty about the economy" (Scotti, 2016, p. 2). Moreover, "Agents base decisions on their perceived uncertainty rather than on an objective uncertainty that they do not observe" (Scotti, 2016, p. 16).

Our *EURQ* index delivers patterns that in the context of uncertainty measurement are both interesting and useful. Provided that the appropriate set of queries is used, the *EURQ* index, which is based on large-scale data from a freely available survey, delivers updated high-frequency information on people's moods. We believe that the *EURQ* index furthers our knowledge of the dynamics of the *perception* of uncertainty by all economic agents and of the *specific components* of uncertainty that worry economic agents the most. In the future, we aim to create a series of disaggregated *EURQ* indices: one based on macroeconomic queries, another based on financial queries, another based on normative queries, and a fourth based on political queries. This disaggregation may help address measurement errors and the endogeneity of the mainstream uncertainty indices proposed in the literature (for a recent example of estimates of the impact of different types of uncertainty on the US economy, see Mumtaz and Surico, 2018). We also shed some light on the *miscellanea* of empirical results in other papers (Angelini et al., 2019, Ludvigson et al., 2020, and Shin and Zhong, 2020) and offer new interpretations incorporating the recent statistical and economic views of Born et al. (2018), Carriero et al. (2016), Choi (2017), Mumtaz and Theodoridis (2018), and Scotti (2016).

Acknowledgements
Our research started in 2013 and since then has greatly improved thanks to the comments of participants at the 35th International Symposium on Forecasting organized by the International Institute of Forecasters, Riverside, June 21-24, 2015; the 2nd workshop on macroeconomic uncertainty held at University College London, May 12-13, 2016; the Stanford Institute for Theoretical Economics (SITE) session on the macroeconomics of uncertainty and volatility, Stanford University, September 7-9, 2016; the International Association for Applied Econometrics (IAAE) Conference, Montreal, Canada, June 26-29, 2018; and the 12th International Conference on Computational and Financial Econometrics held at the University of Pisa, Italy, December 14-16, 2018. We are grateful to Scott Baker, Nick Bloom, Wojciech Charenza, Steven Davis, Joshy Easaw, Luca Fanelli, Kajal Lahiri, Svetlana Makarova, Anna Orlik, Barbara Rossi, the editor and two anonymous referees for their comments and suggestions. The usual disclaimers apply.

References

Alexopoulos, M. and J. Cohen (2015), "The power of print: uncertainty, markets, and the economy", *International Review of Economics and Finance,* Vol. 40, pp. 8-28.

Algan, Y., E. Beasley, F. Guyot, K. Higa, F. Murtin and C. Senik (2016), "Big data measures of well-being: Evidence from a Google well-being index in the US", *OECD Statistics Working Paper*, No. 3.

Altig, D., S. R. Baker, J. M. Barrero, N. Bloom, P. Bunn, S. Chen, S. J. Davis, J. Leather, B. H. Meyer, E. Mihaylov, P. Mizen, N. B. Parker, T. Renault, P. Smietanka and G. Thwaites (2020), "Economic uncertainty before and during the COVID-19 pandemic", NBER Working Paper, No. 27418.

Angelini, G., E. Bacchiocchi, G. Caggiano and L. Fanelli (2019), "Uncertainty across volatility regimes", *Journal of Applied Econometrics*, Vol. 34, No. 3, pp. 437-455.

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.

Baker, S. R., N. Bloom and S. J. Davis (2016), "Measuring economic policy uncertainty", *The Quarterly Journal of Economics*, Vol.131, No. 4, pp. 1593-1636.

Baker, S. R. and A. Fradkin (2017), "The impact of unemployment insurance on job search: evidence from Google search data", *The Review of Economics and Statistics*, Vol. 99, No. 5, pp. 756-768.

Ban, P., A. Fouirnaies, A. B. Hall and J. M. Snyder (2019), "How newspapers reveal political power", *Political Science Research and Methods*, Vol. 7, No. 4, pp. 661–678.

Banerjee, A., A. Chandrasekhar, E. Duflo and M. O. Jackson (2019), "Using gossips to spread information: Theory and evidence from two randomized controlled trials", *The Review of Economic Studies*, Vol. 86, No. 6, pp. 2453–2490.

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

Beetsma, R. and M. Giuliodori (2012), "The changing macroeconomic response to stock market volatility shocks", *Journal of Macroeconomics*, Vol. 34, pp. 281-293.

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.

Bontempi, M. E., M. Frigeri, R. Golinelli and M. Squadrani (2019), "Uncertainty, Perception and the Internet", *University of Bologna DSE Working Paper*, No. 1134.

Bontempi, M. E., R. Golinelli and G. Parigi (2010), "Why demand uncertainty curbs investment: Evidence from a panel of Italian manufacturing firms", *Journal of Macroeconomics*, Vol. 32, pp. 218-238.

Born, B., S. Breuer and S. Elstner (2018), "Uncertainty and the Great Recession", *Oxford Bulletin of Economics and Statistics*, Vol. 80, No. 5, pp. 951-971.

Bulut, L. (2018), "Google Trends and the forecasting performance of exchange rate models", *Journal of Forecasting*, Vol.37, pp. 303-315.

Caballero R. J. and A. Simsek (2020), "Asset prices and aggregate demand in a "COVID-19" shock: a model of endogenous risk intolerance and Lsaps", *NBER Working Papers Series*, No. 27044.

Caggiano, G., E. Castelnuovo and N. Groshenny (2014), "Uncertainty shocks and unemployment dynamics in US recessions", *Journal of Monetary Economics*, Vol. 67, pp. 78-92.

Caggiano, G., E. Castelnuovo and N. Groshenny (2017), "Economic policy uncertainty and unemployment in the United States: a nonlinear approach", *Economics Letters*, Vol. 151, pp. 31-34.

Carrière-Swallow Y. and F. Labbé (2013), “Nowcasting with Google Trends in an emerging market”, *Journal of Forecasting*, Vol. 32, pp. 289–298.

Carriero, A., H. Mumtaz, K. Theodoridis and A. Theophilopoulou (2015), "The impact of uncertainty shocks under measurement error: a proxy SVAR approach", *Journal of Money, Credit and Banking*, Vol. 47, No. 6, pp. 1223-1238.

Carriero, A., T. E. Clark and M. Marcellino (2018), "Endogenous Uncertainty", *Federal Reserve Bank of Cleveland*, Working Paper No. 18-05.

Carroll, C. D. (2003), Macroeconomic expectations of households and professional forecasters”, *The Quarterly Journal of Economics*, Vol. 118, No. 1, pp. 269-298.

Castelnuovo, E. and T. D. Tran (2017), "Google It up! A Google Trends-based uncertainty index for the United States and Australia", *Economics Letters*, Vol. 161, pp. 149-153.

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

Choi, S. (2013), "Are the effects of Bloom's uncertainty shock robust?", *Economics Letters*, Vol. 199, pp. 216-220.

Choi, S. (2017), “Variability of the effects of uncertainty shocks: new stylized facts from the OECD countries”, *Journal of Macroeconomics*, Vol. 53, pp. 127-144.

Chon, G. (2016), “Critics are lining up to oppose changes to Dodd-Frank law”, *The New York Times*, September 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.

Clarida, R., J. Galì and M. Gertler (2000), "Monetary policy rules and macroeconomic stability: evidence and some theory", *The Quarterly Journal of Economics*, Vol. 115, No. 1, pp. 147-180.

Cogley, T., G. E. Primiceri and T. J. Sargent (2010), "Inflation-gap persistence in the US", *American Economic Journal: Macroeconomics*, Vol. 2, pp. 43-69.

Coibion, O. and Y. Gorodnichenko (2015), “Information rigidity and the expectations formation process: a simple framework and new facts”, *American Economic Review*, Vol. 105, No. 8, pp. 2644-2678.

Da Z., J. Engelberg and P. Gao (2011), “In Search of Attention”, *The Journal of Finance*, vol. 66, No. 5, pp. 1461-1499.

Da Z., J. Engelberg and P. Gao (2015), “The Sum of All FEARS Investor Sentiment and Asset Prices”, *Review of Financial Studies*, vol. 28, No. 1, pp. 1-32.

D'Amuri, F. and J. Marcucci (2017), "The predictive power of Google searches in forecasting unemployment", *International Journal of Forecasting*, Vol. 33, pp. 801-816.

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

Donadelli, M. (2015), "Google search-based metrics, policy-related uncertainty and macroeconomic conditions", *Applied Economic Letters*, Vol. 22, No. 10, pp. 801-807.

Donadelli, M. and L. Gerotto (2019), “Non-macro-based Google searches, uncertainty and real economic activity”, *Research in International Business and Finance*, Vol. 48, pp. 111-142.

Drake M. S., D. T. Roulstone and J. R. Thornock (2012) “Investor Information Demand: Evidence from Google Searches Around Earnings Announcements”, *Journal of Accounting Research*, Vol. 50, No. 4, pp. 1001-1040.

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.

Enke, B. (2020), "What you see is all there is", *The Quarterly Journal of Economics*, Vol. 135, No. 3, pp. 1363-1398.

Ferrara, L. and P. Guérin (2018), "What are the macroeconomic effects of high-frequency uncertainty shocks?", *Journal of Applied Econometrics*, Vol. 33, pp. 662–679.

Ferrara, L. and A. Simoni (2019), "When are Google data useful to nowcast GDP? An approach via pre-selection and shrinkage", *CREST Serie de Documents de Travail*, No. 4.

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

Gotz, T. B. and T. A. Knetsch (2019), "Google data in bridge equation models for German GDP", *International Journal of Forecasting*, Vol. 35, pp. 45-66.

Groseclose, T. and J. Milyo (2005), "A measure of media bias", *The Quarterly Journal of Economics*, Vol. 120, No. 4, pp. 1191-1237.

Guiso, L. and G. Parigi (1999), "Investment and demand uncertainty", *Quarterly Journal of Economics*, Vol. 114, No. 1, pp. 185-227.

Haddow, A., C. Hare, J. Hooley and T. Shakir (2013), "Macroeconomic uncertainty: what it is, how can we measure it and why does it matter?", Bank of England Quarterly Bulletin, No. 2, pp. 100-109.

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

Henzel, S. R. and M. Rengel (2014), "Dimensions of macroeconomic uncertainty: a common factor analysis", *CESifo Working Papers*, No. 4991.

Irwin, N (2020). "Why the Stock Market Isn't Too Worried About Coronavirus" *The New York Times*, February 20.

Jetter, M. (2019), "The inadvertent consequences of al-Qaeda news coverage", *European Economic Review*, Vol. 119, pp. 391-410.

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.

Kim, D. K. D., K. K. Kee and J. W. Dearing (2020), "Applying the communication theory of diffusion of innovations to economic sciences: a response to the 'Using gossips to spread information' experiments conducted by the 2019 Nobel Laureates", *Journal of Applied Communication Research*, Vol. 48, No. 2, pp. 157-165.

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.

Krugman, P. (2011), "To the limit", *The New York Times*, June 30.

Kupfer, A. and J. Zorn (2020), “A language-independent measurement of economic policy uncertainty in eastern European countries”, *Emerging Markets Finance and Trade*, Vol. 56, No. 5, pp. 1166-1180.

Lahiri, K. and Y. Zhao (2017), “Determinants of consumer sentiment over business cycles: evidence from the US Survey of Consumers”, *Journal of Business Cycle Research*, Vol. 12, No. 2, pp. 187-2015.

Leduc, S. and Z. Liu (2016), "Uncertainty shocks are aggregate demand shocks", *Journal of Monetary Economics*, Vol. 82, pp. 20-35.

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.

Ludvigson, S. C., S. Ma and S. Ng (2020), "Uncertainty and business cycles: exogenous impulse or endogenous response?", *American Economic Journal: Macroeconomics*, forthcoming.

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.

Meinen, P. and O. Roehe (2017), “On measuring uncertainty and its impact on investment: Cross-country evidence from the euro area”, *European Economic Review*, Vol. 92, pp. 161-179.

Mumtaz, H. and K. Theodoridis (2018), "The changing transmission of uncertainty shocks in the US: An empirical analysis", *Journal of Business & Economic Statistics*, Vol. 36, No. 2, pp. 239-252.

Mumtaz, H and P. Surico (2018), "Policy uncertainty and aggregate fluctuations". *Journal of Applied Econometrics*, Vol. 33, No. 3, pp. 319-331.

Pew Research Center (2016), “The modern news consumer: News attitudes and practices in the digital era”, downloadable at http://www.journalism.org/2016/07/07/the-modern-news-consumer/

Piffer, M and M. Podstawski (2017), “Identifying uncertainty shocks using the price of gold”, *The Economic Journal*, Vol. 128, No. 616, pp. 3266-3284.

Protess, B and Hirschfeld, J.D. (2017), “Trump moves to roll back Obama-era financial regulations”, *The New York Times*, February 3

Puglisi R. and Snyder J. M. (2016), “Empirical studies of media bias”, in S. P. Anderson, J. Waldfogel and D. Stromberg (eds.), *Handbook of Media Economics*, Vol. IB, Elsevier.

Ramey, V. A. (2011), "Identifying government shocks: it's all in the timing", *The Quarterly Journal of Economics*, Vol. 126, No. 1, pp. 1-50.

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

Repetto, L. (2018), "Political budget cycles with informed voters: evidence from Italy”, *The Economic Journal*, Vol. 128, No. 616, pp. 3320-3353.

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

Rossi, B and T. Sekhposyan (2015), “Macroeconomic Uncertainty Indices Based on Nowcast and Forecast Error Distributions”, *American Economic Review*, Vol. 105, No. 5, pp. 650-655.

Rossi, B., T. Sekhposyan and M. Soupre (2020), “Understanding the sources of macroeconomic uncertainty”, *mimeo*

Scotti, C. (2016), “Surprise and uncertainty indexes: real-time aggregation of real-activity macro surprises”, *Journal of Monetary Economics*, Vol. 82, pp. 1-19.

Segal, G., I. Shaliastovich and A. Yaron (2015), “Good and bad uncertainty: macroeconomic and financial market implications”, *Journal of Financial Economics*, Vol. 117, No. 2, pp. 369-397.

Shin, M. and M. Zhong (2020) “A new approach to identifying the real effects of uncertainty shocks”, *Journal of Business & Economic Statistics*, Vol. 38, No. 2, pp. 367-379.

Sims, C. A. (2003), "Implications of rational inattention", *Journal of Monetary Economics*, Vol. 50, pp. 665–690.

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

Sirotkin P. (2012), *Metrics On Search Engine Evaluation*, PhD Thesis, Heinrich-Heine-Universität Düsseldorf, April.

Solberger, M. and E. Spanberg (2020), “Estimating a dynamic factor model in EVIEWS using the Kalman filter and smoother”, *Computational Economics*, Vol. 55, pp. 875-900.

Stock, J. H., and M. W. Watson (2002), “Macroeconomic forecasting using diffusion indexes”, *Journal of Business & Economic Statistics*, Vol. 20, No. 2, pp. 147-162.

Stock, J. H. and M. W. Watson (2007), "Why has US inflation become harder to forecast?", *Journal of Money, Credit and Banking*, Vol. 39, No. 1, pp. 3-33.

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.

Vosen S. and T. Schmidt (2011), “Forecasting private consumption: survey-based indicators vs. Google trends”, *Journal of Forecasting*, Vol. 30, pp. 565-578.

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

Tab. 1 – Correlation analysis of alternative uncertainty indices

	logVIX	SCOTTI	FPRED	MPRED	logEPU	logEURQ
(1) Cyclicality						
Downturn/upturn *mean* ratios	1.189	1.349	1.310	1.119	1.012	1.011
Downturn/upturn *std dev* ratios	1.755	1.877	2.172	3.217	1.469	1.247
(2) Correlation with measures of						
- ex ante uncertainty	0.2858	0.2073	0.2402	0.2885	**0.3668**	**0.3632**
- ex post uncertainty	**0.4876**	**0.5073**	**0.5319**	**0.6291**	0.2654	0.2814
(3) Correlation across indices						
Sample without COVID-19 months						
SCOTTI	0.4325					
FPRED	0.8184	0.4383				
MPRED	0.5531	0.5235	0.7465			
logEPU	0.5768	0.1232	0.5527	0.1791		
logEURQ	0.3410	0.3800	0.2950	0.3801	0.0902	
Sample with COVID-19 months						
SCOTTI	-					
FPRED	0.8090	-				
MPRED	0.5530	-	0.7679			
logEPU	0.6114	-	0.5986	0.3276		
logEURQ	0.4121	-	0.3452	0.4309	0.1860	
(4) Links with the 1st common factor						
- correlation coefficients (No C.)	0.8790	0.5505	0.9707	0.8036	0.5470	0.4098
- correlation coefficients (Yes C.)	0.8581	-	0.9747	0.8225	0.6181	0.4637
- share of each index explained	0.7363	0.3029	0.9422	0.6458	0.3821	0.2150

The *VIX*, *EPU* and *EURQ* are taken in logs to mitigate the effect of many outliers; the *FPRED*, *MPRED* and *EURQ* are seasonally adjusted using the Census X13 filter when seasonality tests are significant (for details, see Bontempi et al., 2019).

(1) Cycle measured using NBER dating since 2004m1.

(2) Ex ante and ex post measures are from Rossi et al. (2020). The higher correlation for each index are in bold.

(3) Sample periods: 2004m1-2019m11 (without COVID-19 months) and 2004m1-2020m5 (with COVID-19 months). The latter correlations exclude *SCOTTI,* as its data are not available after 2019m11.

(4) The correlation coefficients are computed using both the common sample 2004m1-2019m11 (without COVID-19 months, "No C.") and all the available observations over the whole sample 2004m1-2020m5 (with COVID-19 months, "Yes C."). The "share of each index explained" row reports the R^2 of bivariate regressions of each uncertainty index against the common factor over the "unbalanced" sample.

Tab. 2 - The dynamics of news-based (*Newsbank*) and search-based (*EURQ*) index relationships ([a])

	Contemporaneous correlation ([d]):	
Granger causality *from*/*to*:	*High* (>0.25)	*Low* (<0.25)
Query-driven, *EURQ/Newsbank* ([b])	Fiscal policy (**FP**), Sovereign debt and currency crisis (**SDCC**)	Health care (**HC**)
News-pooled, *Newsbank*/*EURQ* ([c])	Monetary policy (**MP**), Trade policy (**TP**)	Regulation (**RE**)
No-dynamics-uncertainty interest	Entitlement programmes (**EP**)	National security and war (**NS**)

([a]) This table summarizes the VAR model (2) results. The joint stationarity of all the variables is assessed by the Johansen (1995) trace test. When the Johansen test does not reject the null of reduced rank (not all the variables are stationary), Granger causality is tested by the Toda and Yamamoto (1995) approach. Seasonal dummies, if significant, are included in the VAR model (2).
([b]) Query-driven = *EURQ* index Granger-causes *Newsbank* index.
([c]) News-pooled = *Newsbank* index Granger-causes *EURQ* index.
([d]) The 0.25 threshold of correlation coefficients (in absolute value) corresponds to the statistical significance of the null hypothesis that the correlation is 1% significant; coefficients below 0.25 ("Low") are not significantly different from zero.

Most searched words:

EP Social security job
RE FDIC jobs
HC Health care reform
MP Interest rate
EP Food stamps
HC Affordable Care Act
NS Terrorism
FP Debt ceiling
RE Minimum wage
EP Unemployment benefits
RE Financial reform and tort reform
FP Tax rate
RE Cap and trade
RE Environmental Protection Agency
FP Taxation
RE Energy policy
MP Bernanke
RE Office of thrift supervision
HC Medicare
….
SDCC Sovereign debt

Query-driven:
FP (high corr.)
SDCC (high corr.)
HC (low corr.)

News-pooled:
MP (high. corr.)
RE (low corr.)
TP (high corr.)

No Granger causality:
EP (high corr.)
NS (low corr.)

Column on the left: words ordered from the most to the least searched according to Bayesian analysis. Column on the right: categories that are query driven, news pooled and without Granger causality.

Fig. 1 – The *EURQ* for the USA: annotated chart

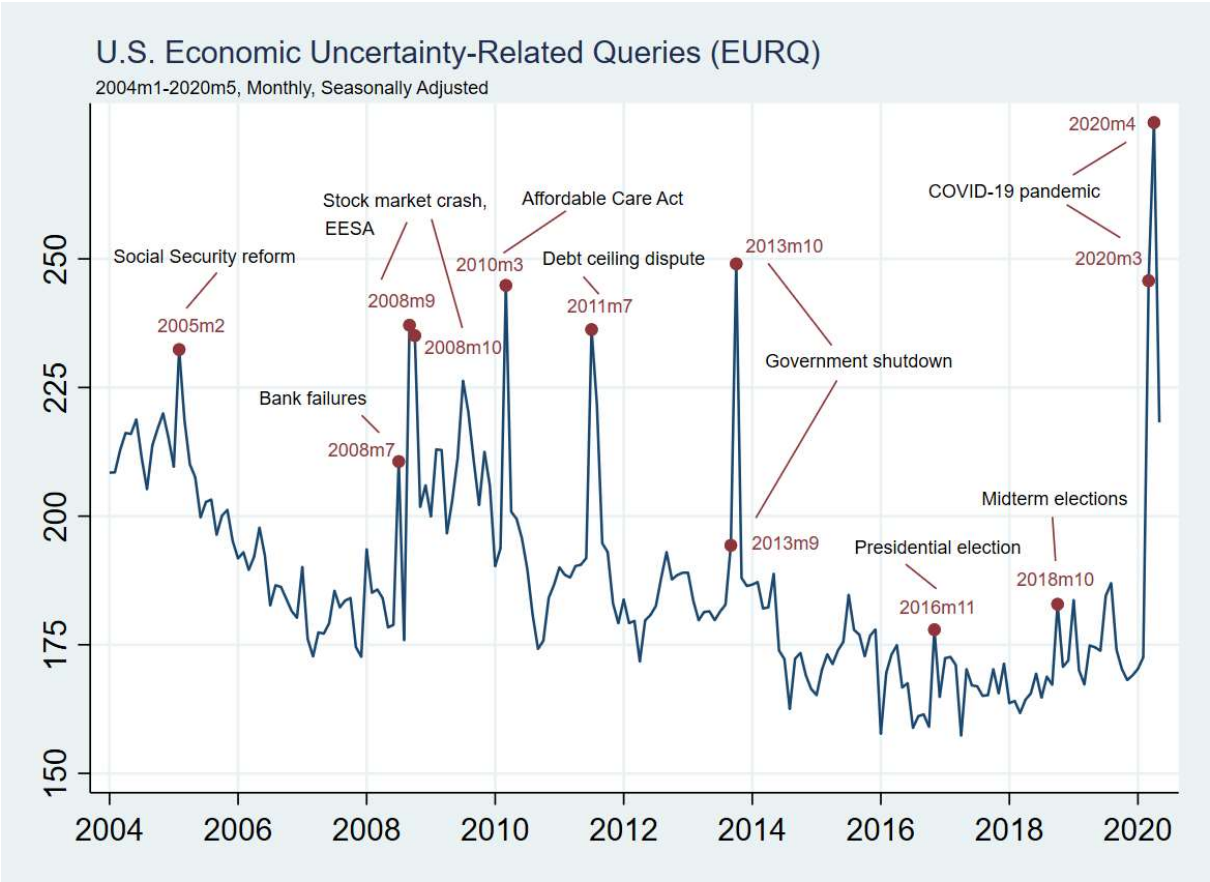
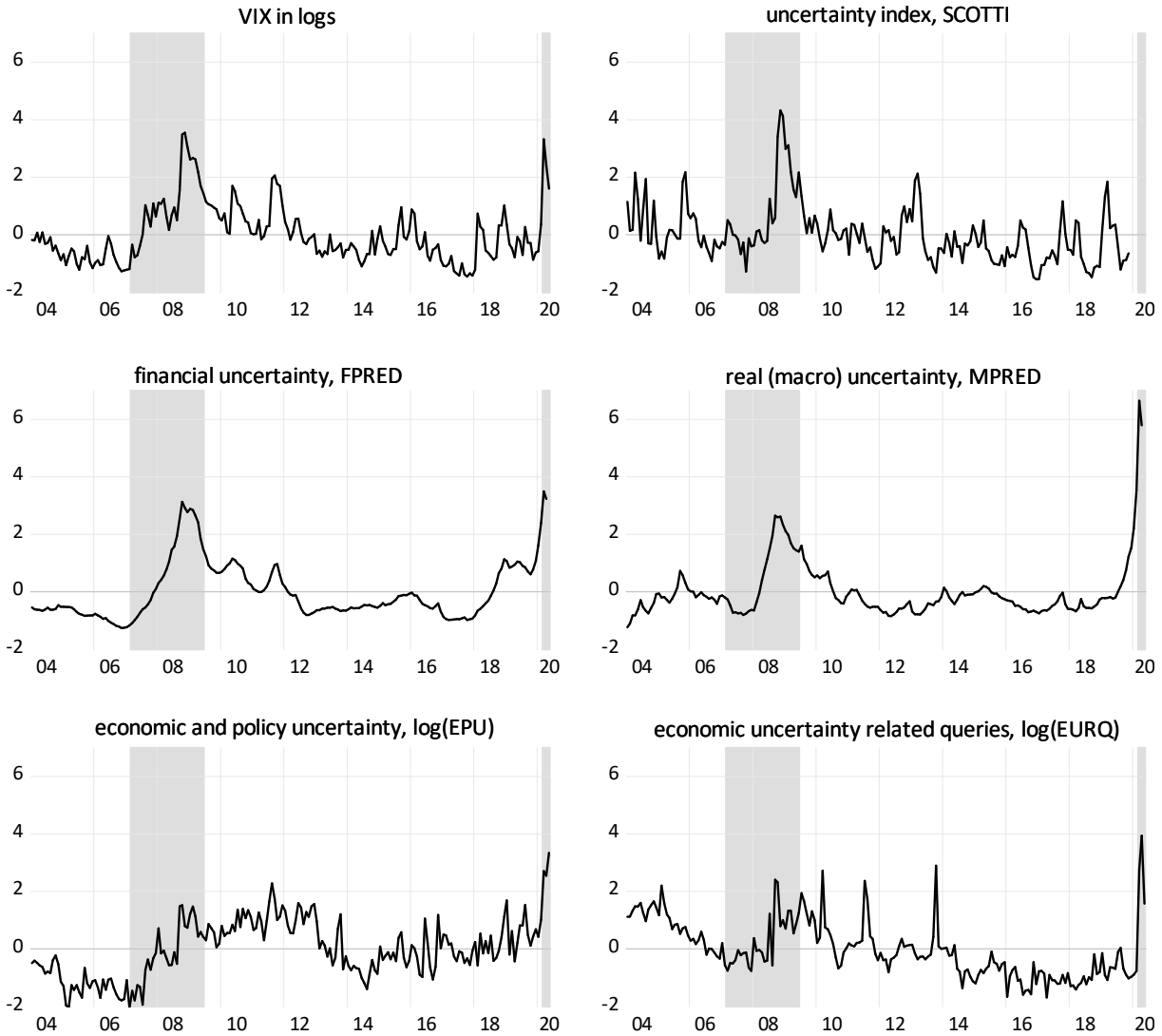
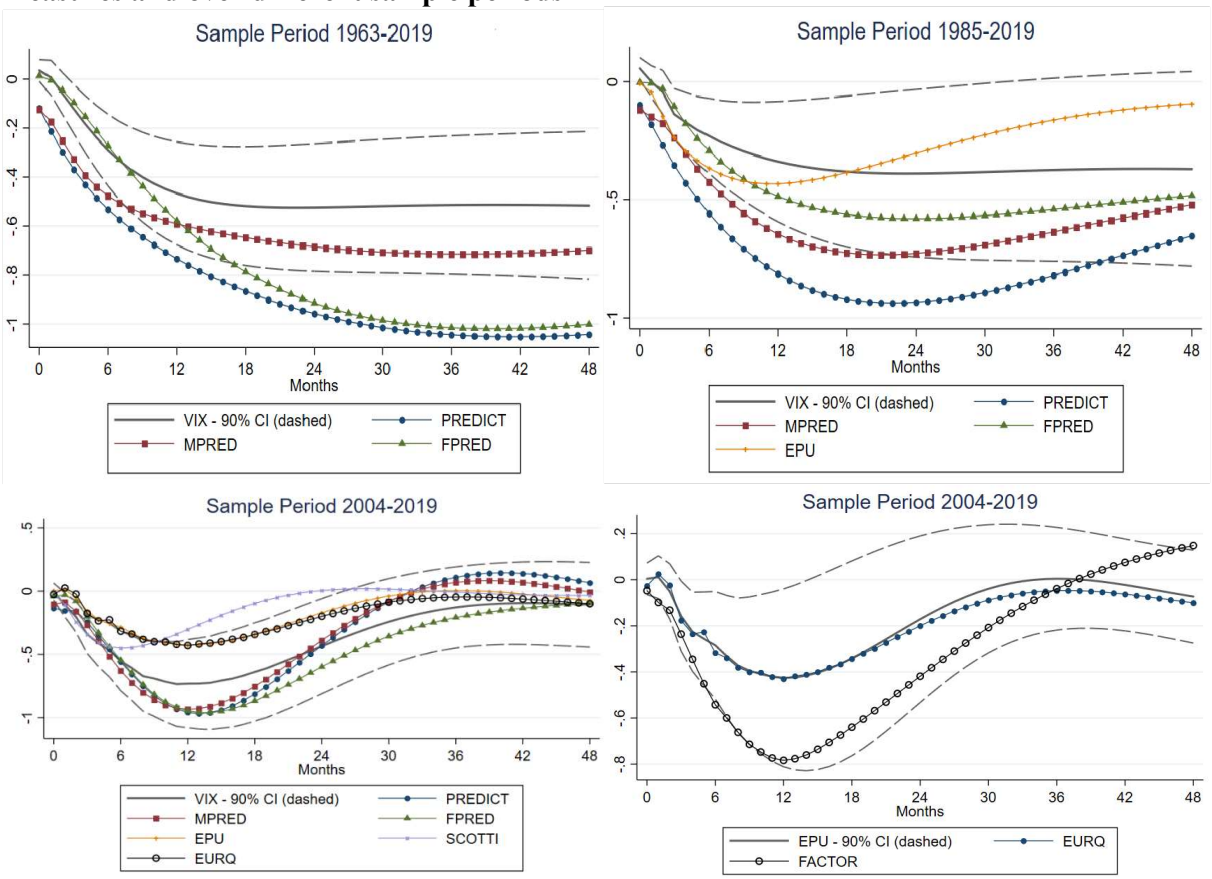


Fig. 2 – Alternative uncertainty measures since 2004m1



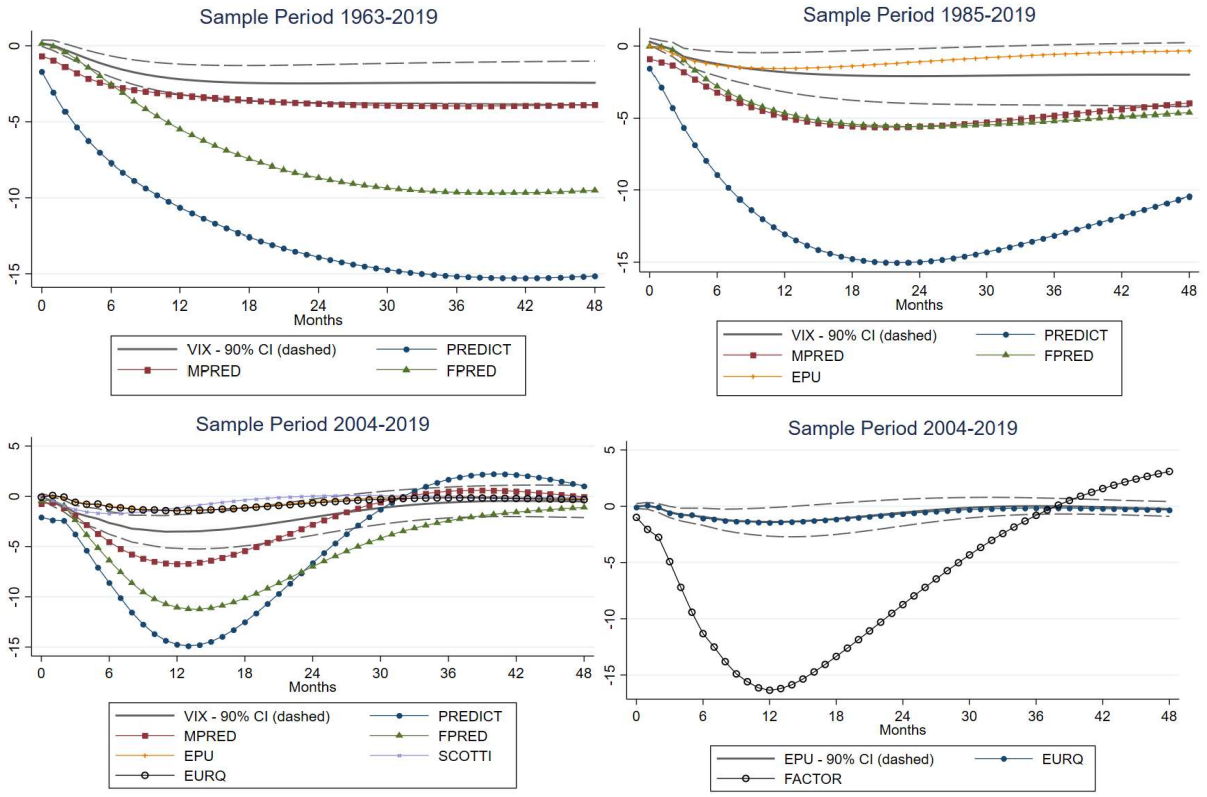
While the *VIX*, *EPU* and *EURQ* are promptly updated and downloadable at the beginning of each month, the *SCOTTI* ends in November 2019, and the *MPRED* and *FPRED* have been recently updated to April 2020. The *VIX*, *EPU* and *EURQ* are taken in logs to mitigate the effect of many outliers (e.g., Baker et al., 2016); the *MPRED*, *FPRED* and *EURQ* are seasonally adjusted (e.g., Jurado et al., 2015) using the Census X13 filter when seasonality tests are significant (for details, see Bontempi et al., 2019). Measures are standardized to ease comparisons. Shaded areas denote NBER downturn phases.

Fig. 3 – Output responses to uncertainty shocks in VARs with alternative uncertainty measures and over different sample periods



Response (%) of log manufacturing production to a Cholesky one s.d. impulse in *MPRED*, *FPRED* and *PREDICT* (all seasonally adjusted), log*VIX*, log*EPU*, *SCOTTI* and log*EURQ* (seasonally adjusted) for three different estimation samples: 1963–2019 (long), 1985–2019 (medium), and 2004–2019 (short). In each sample, we use different indices according to their data availability. Identification is based on a 5-variable VAR(p), ordered as follows: *uncertainty*, log*SP500*, log(1+fed funds effective rate/100), log of manufacturing employment, and log of manufacturing industrial production. The VARs' lag order is set to:
Long sample: *MPRED* = 3, *FPRED* = 2, log*VIX* = 2, *PREDICT* = 2
Medium sample: *MPRED* = 3, *FPRED* = 3, log*VIX* = 3, *PREDICT* = 2, log*EPU* = 3
Short sample: *MPRED* = 2, *FPRED* = 4, log*VIX* = 6, *PREDICT* = 4, log*EPU* = 4, *SCOTTI* = 2, *EURQ* = 6, *FACTOR*=5
Dashed lines represent the 90% standard error bands obtained in the VAR using log*VIX* as the uncertainty measure.

Fig. 4 – Output responses to uncertainty shocks with alternative measures and sizes corresponding to the Lehman bankruptcy



Response (%) of log manufacturing production to Cholesky impulses of different sizes. For each uncertainty proxy, transformed in log and/or seasonally adjusted as indicated in Figure 3, the size of the shock is set equal to the increase from the average value in June–August 2008 to the average in September-November 2008, i.e., before and after the Lehman bankruptcy. Other details are the same as those in the notes of Figure 3.

Fig. 5 – *EURQ* for Italy: annotated chart

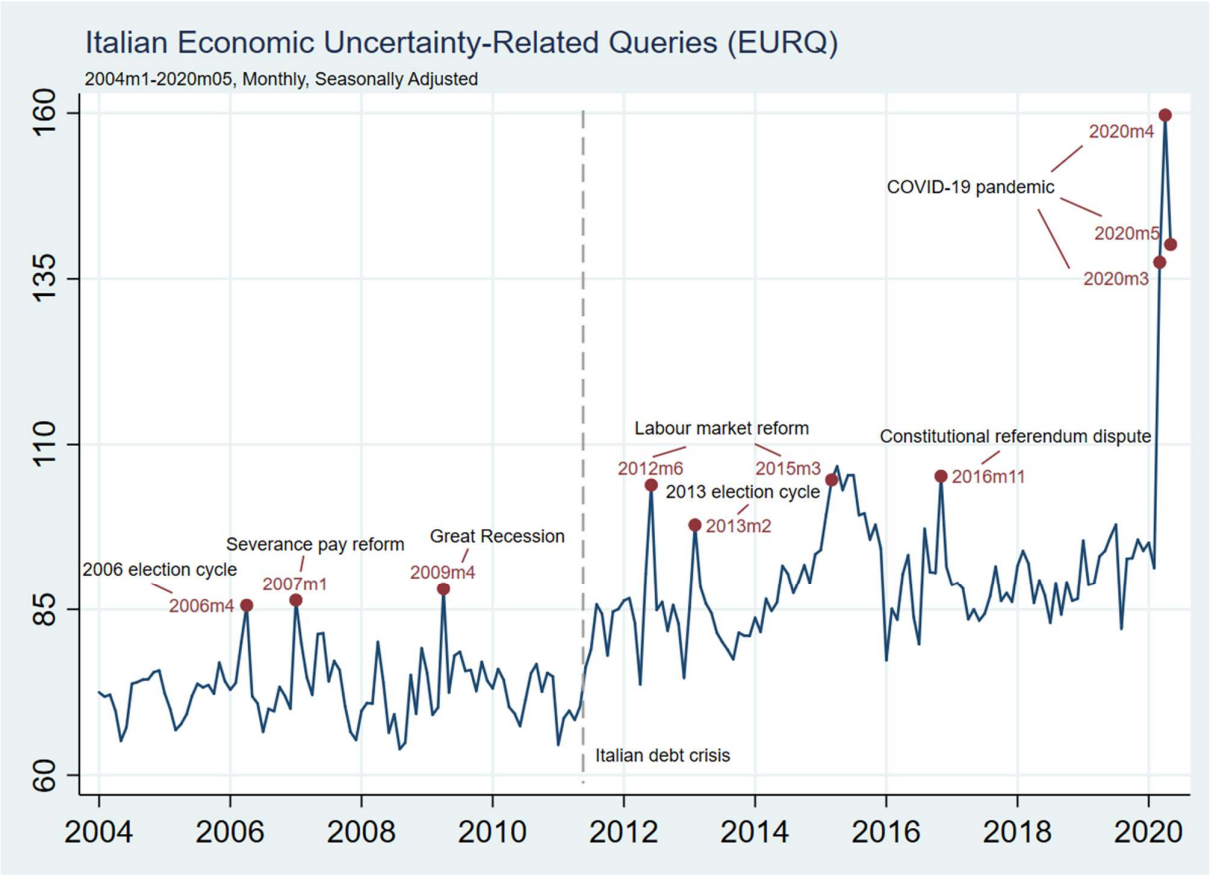
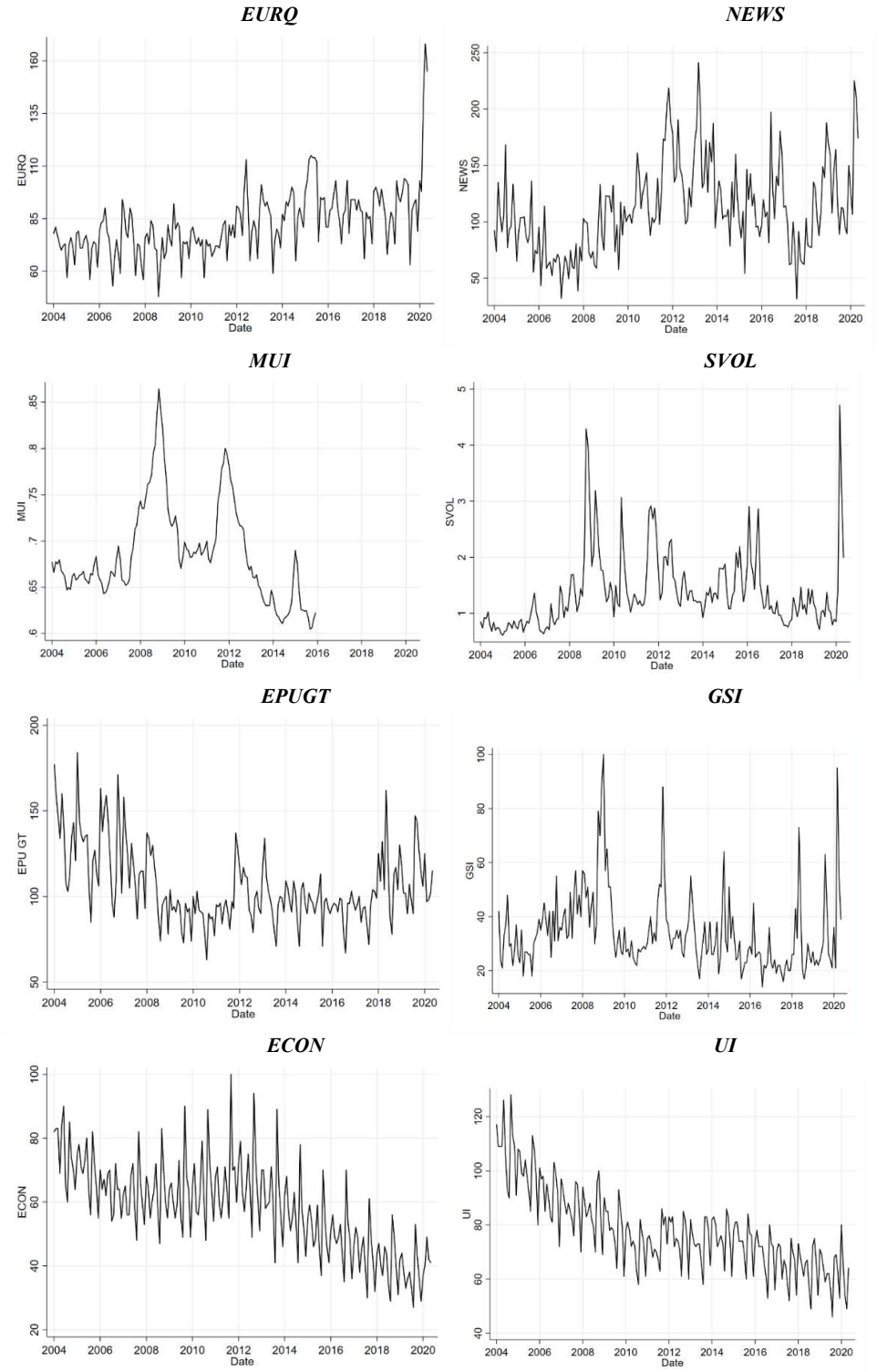
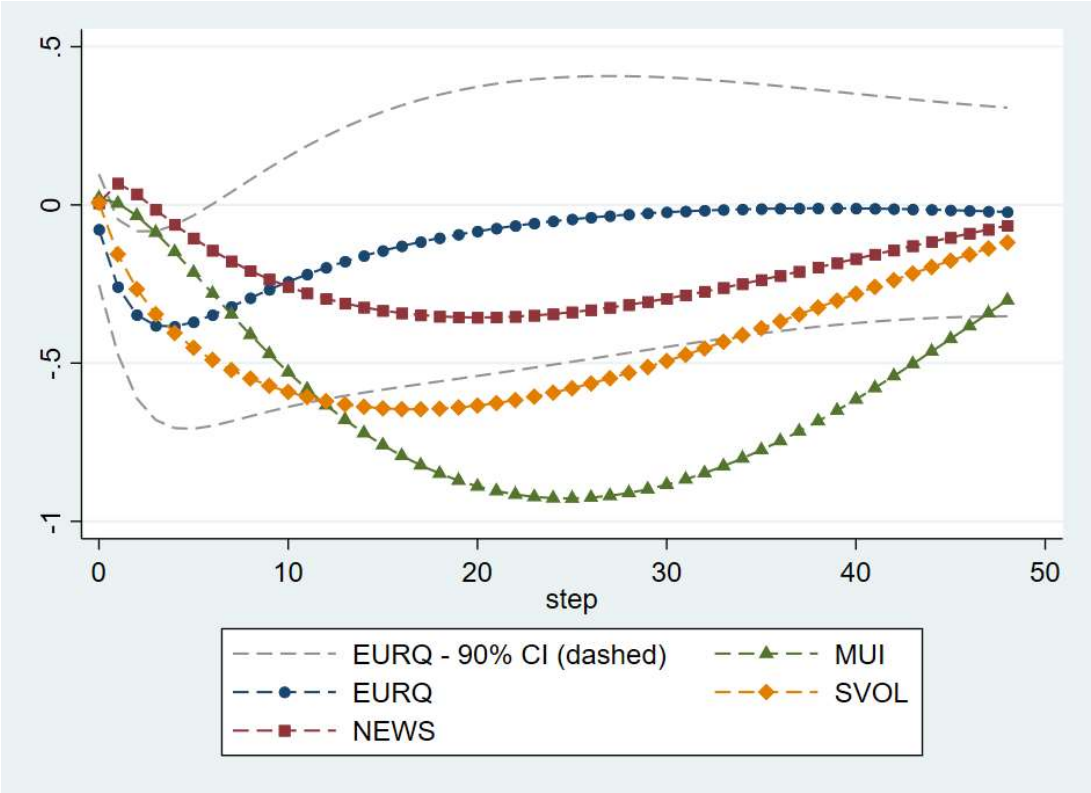


Fig. 6 – Alternative uncertainty measures since 2004m1 for Italy



EURQ is the same as in Figure 5; *NEWS* is the news-based index (BBD approach); *MUI* is the forecast-based index (Jurado et al., 2015 approach); *SVOL* is the finance-based index (from Meinen and Roehe, 2017); *EPUGT* is the search-based index that we create using the nine terms from BBD to compute their *NEWS* index for Italy; *GSI* is the search-based index of Donadelli (2015, three terms); *ECON* is the search-based index of Dzielinski (2012, one term); *UI* is the search-based index of BBVA (2012, fifteen terms). Period 2004m1-2020m5 (2004m1-2015m12 for *MUI*).

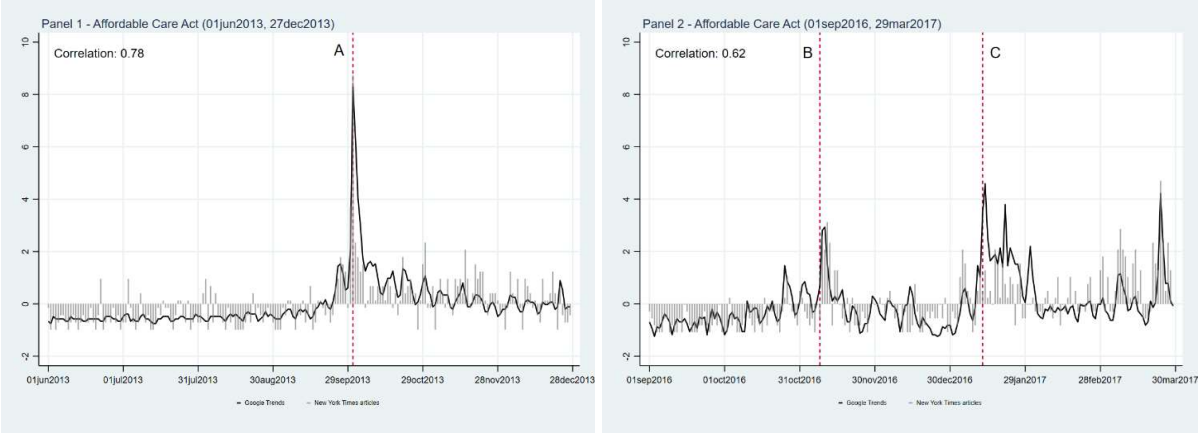
Fig. 7 – Output responses to uncertainty shocks in VARs with alternative uncertainty measures for Italy



Response (%) of log industrial production (excluding the construction industry) to a Cholesky one s.d. impulse in log(*EURQ* - seasonally adjusted), log(*NEWS*), log(*MUI* - seasonally adjusted), log(*SVOL*). Estimated periods: 2004m1–2019m12 for *EURQ*, *NEWS* and *SVOL*; 2004m1–2015m12 for *MUI*. Identification is based on a 5-variable VAR(*p*), ordered as follows: *uncertainty*, log(*FTSE-MIB*), which is the benchmark Italian stock exchange index, log(1+one-month Euribor), log(employment), log(industrial production). The number of lags in the VAR is equal to 1. Estimates are performed with a small-sample degree-of-freedom adjustment. Dashed lines represent the 90% standard error bands of the VAR with log(*EURQ*) as the uncertainty measure.

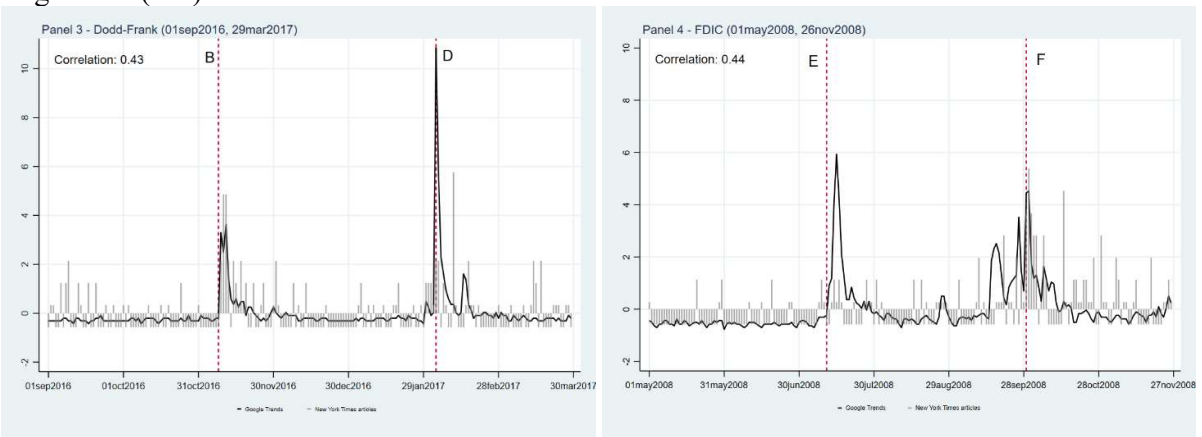
Fig. 8a – Web searches against NYT coverage for some relevant terms ([a])

Health care (**HC**) "Affordable Care Act"



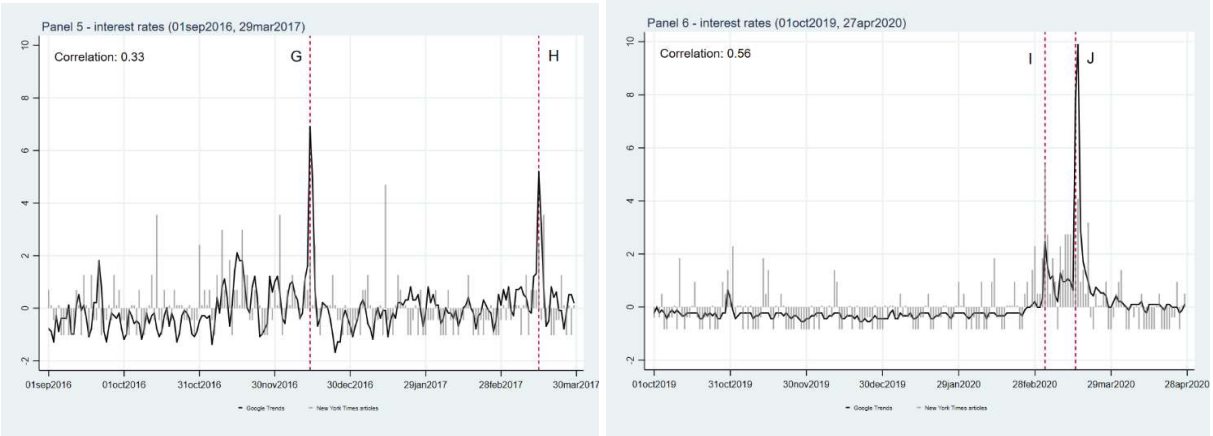
A: October 1, 2013, federal government shutdown; B: November 8, 2016, presidential elections; C: January 12, 2017, Senate vote to pass a budget resolution to repeal to the Affordable Care Act.

Regulation (**RE**) "Dodd-Frank" and "FDIC"



B: November 8, 2016, presidential elections; D: February 3, 2017, Trump's order to review Dodd-Frank; E: July 11, 2008, IndyMac's failure; F: September 29, 2008, stock market crash.

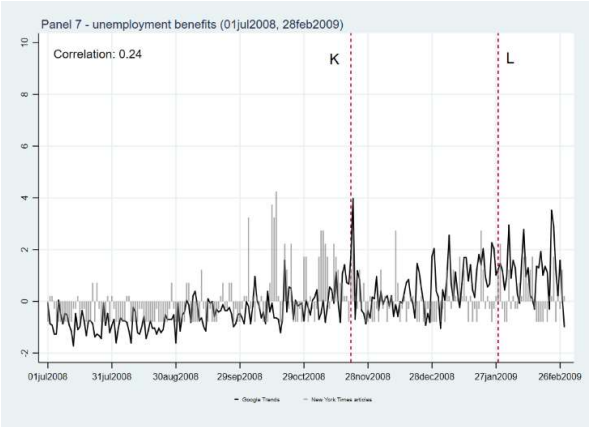
Monetary policy (**MP**) "Interest Rates"

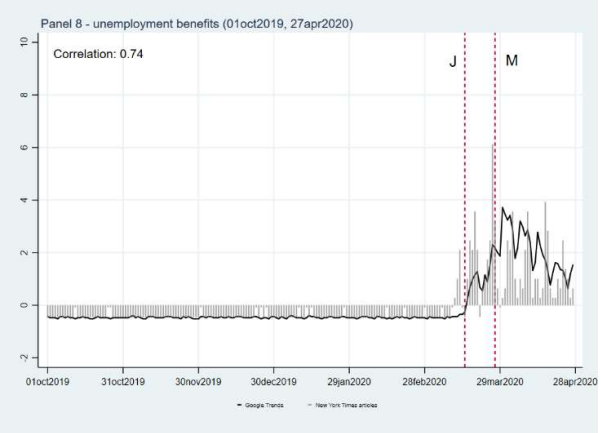


G: December 14, 2016, Fed increase in target for short-term interest rates by 0.25 percentage points (only the second time in a decade that the Fed raised rates). H: March 15, 2017, Fed increase in the key interest rate by 0.25 percentage points (only the third time that the Fed increased rates since the financial crisis). I: March 3, 2020, surprise cut of the Fed's benchmark interest rate by 50 basis points. J: March 15 2020, Fed rate cut to zero and launch of a massive quantitative easing programme.

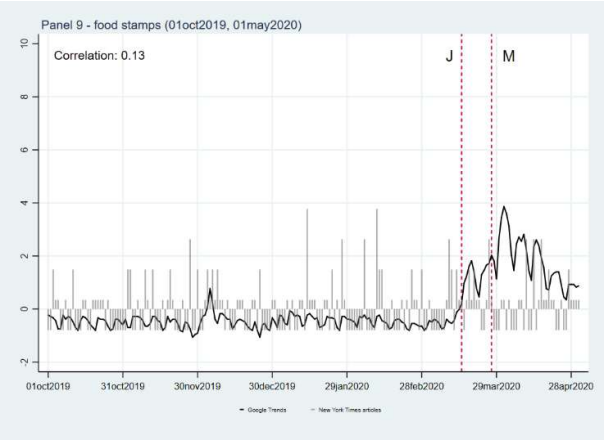
Fig. 8b – Web searches against NYT coverage for some relevant terms ([a])

Entitlement programmes (**EP**) "Unemployment Benefits"





Entitlement programmes (**EP**) "Food Stamps"



K: November 20, 2008, Announcement of White House support for legislation to extend unemployment benefits/Unemployment Compensation Extension Act of 2008. L: January 28, 2009, passage of the American Recovery and Reinvestment Act of 2009 in the House of Representatives. J: March 15, 2020, Fed rate cut to zero and launch of massive quantitative easing programme. M: March 27, 2020, signing into law of the Coronavirus Aid, Relief, and Economic Security Act, a $2.2 trillion economic stimulus bill.

Standardized measures to ease comparisons. "New York Times articles" reflects the daily number of articles from the *New York Times* containing each selected term; data are from the *New York Times* search API (https://developer.nytimes.com/).

Appendix A1 – The technical implementation of the *EURQ*

Google Trends provides an index of the volume of Google searches that is freely available, measured at high frequency and released quickly (almost in real time). This index is called the search volume index – in symbols SVI_{st} – and measures the volume of searches for a query s in each country or region at time t:

$$SVI_{st} = \frac{sv_{st}}{sv_{Gt} \times MSV_{[o,T]}} \times 100 = \frac{sv_{st}}{sv_{Gt} \times \max_{t=[0,T]} \left\{ {sv_{st}}/{sv_{Gt}} \right\}} \times 100 \qquad \text{(A2.1)}$$

where sv_{st} is the number of searches for s within period t.[34] The division by sv_{Gt} – the total number of Google searches within the same period t – should prevent SVI_{st} from being significantly affected by the extensive margin in Internet searches. Moreover, the SVI_{st} series are bounded between 0 and 100 since they are scaled by the maximum value of sv_{st} / sv_{Gt} from 0 to T (i.e., over the entire time span) and then multiplied by 100. The aggregate *EURQ* index is obtained by summing the relative *SVIs*. As they are peak-normalized, the *SVI*s' sensitivity to extreme values is *per se* sharply reduced, as this avoids the use of various methods of treating extreme values (such as outlier trimming) that could bias the genuine data structure. However, differences in SVI_{st} are consequently independent of the relative relevance of s over total Google traffic: an increase in the required information about term s is not measured as an increase in its share but rather as an increase in its level towards 100. Therefore, the SVI_{st} indices are short-term indicators measuring how close the need for information about s at time t is to its highest point rather than indicators of the most searched-for terms. Of course, the *SVI* indicators are subject to sampling variability since it is impossible to exactly replicate the search volumes, which differ slightly from one download to the next.[35] In Bontempi et al. (2019), we conducted many robustness analyses. We used the sequence of real-time vintages of *EURQ* to analyse the informational content of the downloads in different months; the results suggest that data revisions do not mix up the real *EURQ* signal.[36] We also assessed the sensitivity of the *EURQ* series to the presence or absence of subsets of search terms by constructing alternative *EURQ* indices that do not take into account blocks of search terms; the robustness of all the resulting series to the omission of given terms is confirmed by the high correlations (in the 0.89-0.95 range) with the *EURQ* index based on the full list of terms. We also investigated how the intensity of Internet use over time could affect the search volumes used to construct the *EURQ*. Since the advent of Google Trends in 2004, Internet

[34] In our extraction period, the frequency of the Google Trends series was weekly, and we converted this to monthly frequency by averaging based on the month in which the week begins. Google Trends only provides *SVIst* observations for those search terms exceeding a minimum threshold of popularity in period t; otherwise, they are set to zero. Therefore, zero SVI_{st} indicates either no searches or a nonsignificant number of searches for s at time t.

[35] Regarding this point, we noticed even intra-day changes.

[36] Confirming Da et al. (2011), who report that the correlation is usually greater than 0.97 for series downloaded several times, we found that the *SVI*s for a search term change very slightly from one download to another, especially when considering highly popular terms.

penetration rates in the USA have been increasing (the extensive margin), while search activity has evolved depending on the size and coverage of issues on the web (the intensive margin). As the volumes of searches for each individual term are scaled to total traffic, the *EURQ* index cannot be significantly affected by the extensive margin. Even though quantification of the impact of the intensive margin is more complex due to changes in the composition of searches, our checks show that spurious low-frequency unit root-like fluctuations can be excluded: the *EURQ* tends to spike during periods of considerable uncertainty rather than showing smoothed local trends.

Taking the original list of 210 search terms of BBD as our starting point, below, we classify our selected 183 queries (s = 1, ..., 183) in 8 policy categories (c = 1, …, 8). We used only 183 terms of the 210 in BBD:[37]

183 included queries for the USA

(1) Fiscal policy, FP (16 queries)

1. taxes rates - calculator
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 (26 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. "Bank of England"
35. "Bank of Japan"
36. BOJ - xem - anglers - jamaica
93. "union rights"
94. "union card check"
95. "collective bargaining law"
96. "national labor relations board"
97. "minimum wage"
98. living wage - calculator
99. "right to work"
100. "closed shop"
101. wages and hours
102. "workers compensation law"
103. "affirmative action"
104. "at-will employment"
105. "trade adjustment assistance"
106. "davis bacon"
107. "equal employment opportunity"
108. "eeo laws"
109. "osha safety"
110. "antitrust"
111. competition policy
112. "monopoly power"
113. patent law - firm - firms - school - schools - lawyer - attorney - group - bar - jobs
114. "federal trade commission"
115. the ftc - complaint
116. "competition law"
117. price fixing - adm - apple
118. "class action law"
119. "healthcare lawsuit"
120. "tort reform"
121. punitive damages - definition - define - what
122. "energy policy"
123. "energy tax"
124. "carbon tax"
125. "cap and trade"
126. "cap and tax"
127. "offshore oil drilling"
128. "clean air act"
129. "clean water act"
130. "environmental protection agency"

[37] Of the 26 dropped terms, 8 of them were repeated several times in the list of included queries, while the other 18 never reached the minimum popularity threshold.

37. "Bank of China"
38. "Bundesbank"
39. "Bank of France"
40. "Bank of Italy"
41. "ECB"
42. overnight lending rate

(3) Health care, HC (13 queries)

43. "health care reform"
44. "Medicaid program"
45. "Medicare program"
46. "health insurance 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. Medicare Part D - humana - aarp
54. "affordable care act"
55. "Obamacare law"

(4) National security and war, NS (16 queries)

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

(5) Regulation, RE (64 queries)

72. "federal regulation"
73. "Glass Steagall"
74. "tarp program"
75. "thrift supervision"
76. Dodd Frank - form - certification
77. "financial reform"
78. "commodity futures trading commission"
79. "cftc"
80. "house financial services committee"
81. "Basel Accord"
82. "Volcker rule"
83. "bank stress test"
84. "securities and exchange commission"
85. "us sec"
86. "deposit insurance"
87. fdic - jobs
88. "fslic"
89. "office of thrift supervision"
90. "Office of the Comptroller of the Currency"
91. "firrea"
92. "truth in lending"
131. the epa - jobs
132. "immigration policy"
133. nlrd
134. pollution controls
135. "copyright law"

(6) Foreign sovereign debt and currency crisis, SDCC (15 queries)

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

(7) Entitlement programmes, EP (20 queries)

151. "entitlement program"
152. "entitlement spending"
153. "government entitlements"
154. social security - office - number - my - calculator - online - jobs - application
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 (13 queries)

171. "import tariffs"
172. import duty - calculator
173. "government subsidy"
174. "government subsidies"
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. "anti dumping"

The Italian *EURQ* index has been obtained selecting and adapting from the 210 BBD terms, the following 136 queries, presented as sorted from the highest to the lower peaks:[38]

136 included queries for Italy	
1. inps - orari - numero - pin - in	69. assegno sociale"
2. ingv	70. "assegno familiare"
3. "agenzia delle entrate"	71. "politica monetaria"
4. "elezioni politiche"	72. "concorrenza sleale"
5. riforma - protestante	73. "dazi doganali"
6. inail	74. "corte di giustizia europea"
7. tfr	75. banca d italia - concorso - concorsi
8. "protezione civile"	76. "carbon tax"
9. isee	77. "assicurazione vita"
10. sanità - istituto - rione - quotidiano	78. "Bank of England"
11. ccnl	79. "contrattazione collettiva"
12. arpa - orari	80. "bonus sociale"
13. terrorismo	81. "autorità vigilanza contratti pubblici"
14. "garanzia giovani"	82. "diritto alla disoccupazione"
15. "pubblica amministrazione"	83. "Bank of Japan"
16. tassazione	84. "copyright law"
17. disoccupazione	85. "sicurezza nazionale"
18. caf	86. Bundesbank
19. "assegni familiari"	87. "emission trading"
20. corte dei conti - concorso	88. ECB
21. anac	89. agcm
22. "spread btp bund"	90. "sicurezza sociale"
23. fallimenti	91. "legge droghe"
24. invalidità - punteggio	92. "sussidio di disoccupazione"
25. consob	93. "pareggio di bilancio"
26. "debito pubblico"	94. "uruguay round"
27. aams	95. "spesa pubblica"
28. servizi sociali - Berlusconi	96. "assegno di disoccupazione"
29. protocollo di Kyoto - riassunto	97. "Bank of China"
30. "titoli di stato"	98. tassato
31. "bail in"	99. Trichet
32. pari opportunità - carfagna	100."legge immigrazione"
33. concorrenza - esercizi	101."vigilanza bancaria"
34. "cuneo fiscale"	102."organizzazione mondiale del commercio"
35. "detassazione straordinari"	103."spese militari"
36. "reddito minimo garantito"	104."no-fly zone"
37. bankitalia - concorso	105.trattato internazionale
38. wto - wikipedia - significato	106."salario minimo"
39. antitrust - significato - wikipedia	107."mutuo surroga"
40. Mario Draghi - moglie - stipendio	108."politica energetica"
41. "detrazione fiscale"	109."centro per l'impiego"
42. "exchange rate"	110."fondo interbancario di tutela dei depositi"
43. "detrazioni fiscali"	111."derivati finanziari"
44. "class action"	112."aliquota fiscale"
45. sussidio	113.military spending
46. "banca centrale europea"	114."forze armate italiane"
47. welfare state - keynes - significato - definizione - beveridge	115."crisi russa"
	116."Crisi asiatica"
48. rivalutazione monetaria - Andreani - avvocati	117."crisi euro"
49. "patto di stabilità"	118."accordi di Basilea"
50. banca centrale - sede	119."anti dumping"
51. "diritto di abitazione"	120."prestito senza busta paga"

[38] The other queries were excluded because of their low relevance in the Web searches over the 2004m1-2020m5 period.

52. "aliquote fiscali"	121."pensione di invalidità civile"
53. "tasso di interesse"	122."svalutazione monetaria"
54. brevetti e marchi	123."offerta di moneta"
55. "agenzia per il lavoro"	124."guerra al terrorismo"
56. "Ignazio Visco"	125."doha round"
57. embargo	126."pratiche commerciali scorrette"
58. "quantitative easing"	127."trivellazioni petrolifere"
59. "portale fallimenti"	128."blocco navale"
60. "fondo di garanzia"	129."prestatore di ultima istanza"
61. "federal reserve"	130."tasso overnight"
62. "indennità di disoccupazione"	131."deficit di bilancio"
63. svalutazione	132."rapporto debito pil italia"
64. "stress test banche"	133."operazioni di mercato aperto"
65. "tassi bce"	134."assicurazione infortuni e malattia"
66. "tasso di cambio"	135."debiti sovrani"
67. "Bank of America"	136."Presidente Federal Reserve"
68. "trattato di Schengen"	

The *EPUGT* index derives from the translated *policy* terms defined by BBD to implement their Italian uncertainty index based on two newspapers, Corriere Della Sera and La Repubblica:

1. tassa	7. spese
2. tasse	8. deficit
3. politica	9. “Banca Centrale”
4. regolamento	10. “Banca d’Italia”
5. regolamenti	11. Legge di bilancio
6. spesa pubblica	12. Bilancio

The *GSI*, *ECON* and *UI* uncertainty indices have been obtained through the Italian translation of the American terms used, respectively, by Donadelli (2015), Dzielinski (2012) and BBVA (2012) for their own indices:

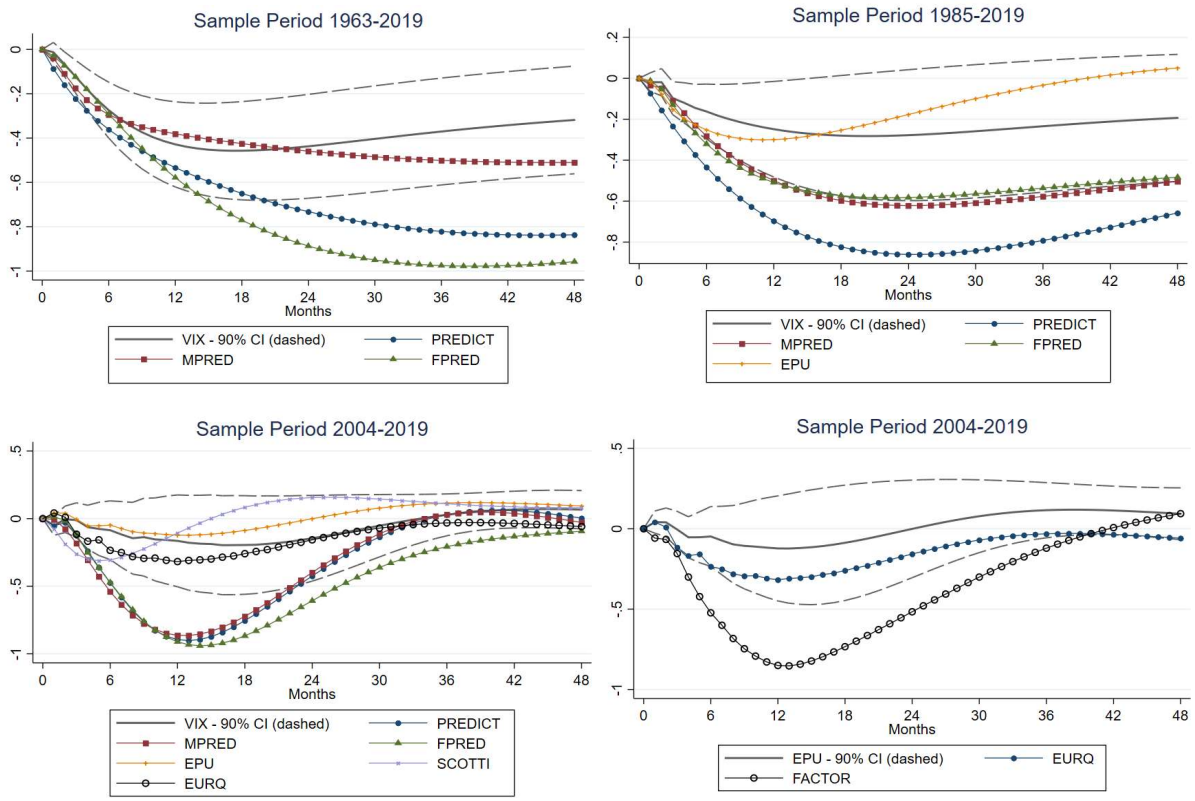
Index	**English term**	**Italian term**
GSI (Donadelli)	US stock market	mercato azionario italiano
	US fed	Bce
	US politics	politica italiana
ECON (Dzielinski)	Economy	Economia
UI (BBVA)	Tax	Tasse
	Debt	Debito
	Fiscal	Fiscale
	Medicare	Riforma
	social security	sicurezza sociale
	Iran	Iran
	Israel	Israele
	Terrorism	Terrorismo
	Revolution	Rivoluzione
	Iraq	Iraq
	Inflation	Inflazione
	Economy	Economia
	Jobs	Lavoro
	Fed	Bce
	stock market	mercato azionario

Appendix A2 – Robustness of results based on alternative VAR identification schemes

In this appendix, we report some robustness checks of VAR model (1), presented in Section 3, based on three alternative identification schemes. In particular, the outcomes in Figure A2.1 are obtained from a VAR identified as in Jurado et al. (2015), where in line with Christiano et al. (2005), variables are ordered from slow-moving industrial production (first place) and employment (second place) to fast-moving SP500 (fourth place) and uncertainty (fifth place). The outcomes in Figure A2.2 are obtained from a VAR identified as in Rossi et al. (2020), where in line with Jurado et al. (2015), variables are ordered from slow moving to fast-moving, except for uncertainty, which Rossi et al. (2020) order first. Finally, the outcomes in Figure A3.3 are obtained from a VAR identified as in Bloom (2009), where the variables are listed as in BBD (i.e., as in our baseline VAR in the main text), except for uncertainty, which Bloom (2009) puts in second place after SP500.

The comparison of the three figures with our baseline Figure 3 stresses the remarkable robustness of the findings of Section 3, specifically that the shape of the output responses to uncertainty shocks is basically due to the different sample periods over which the VAR models are estimated, rather than to the use of different uncertainty proxies. It is also worth remembering that many outcomes that we obtained with our baseline VAR in Section 3 are in line with those reported by the literature using VAR models with a larger number of variables (for example, Jurado et al. (2015) use an 11-variable VAR, but their results are perfectly in line with those from our 5-variable VAR) and estimated over slightly different sample periods (our data are more up to date than those used by the published literature).

Fig. A2.1 – Output responses to uncertainty shocks in VARs ordered as in Jurado et al. (2015)



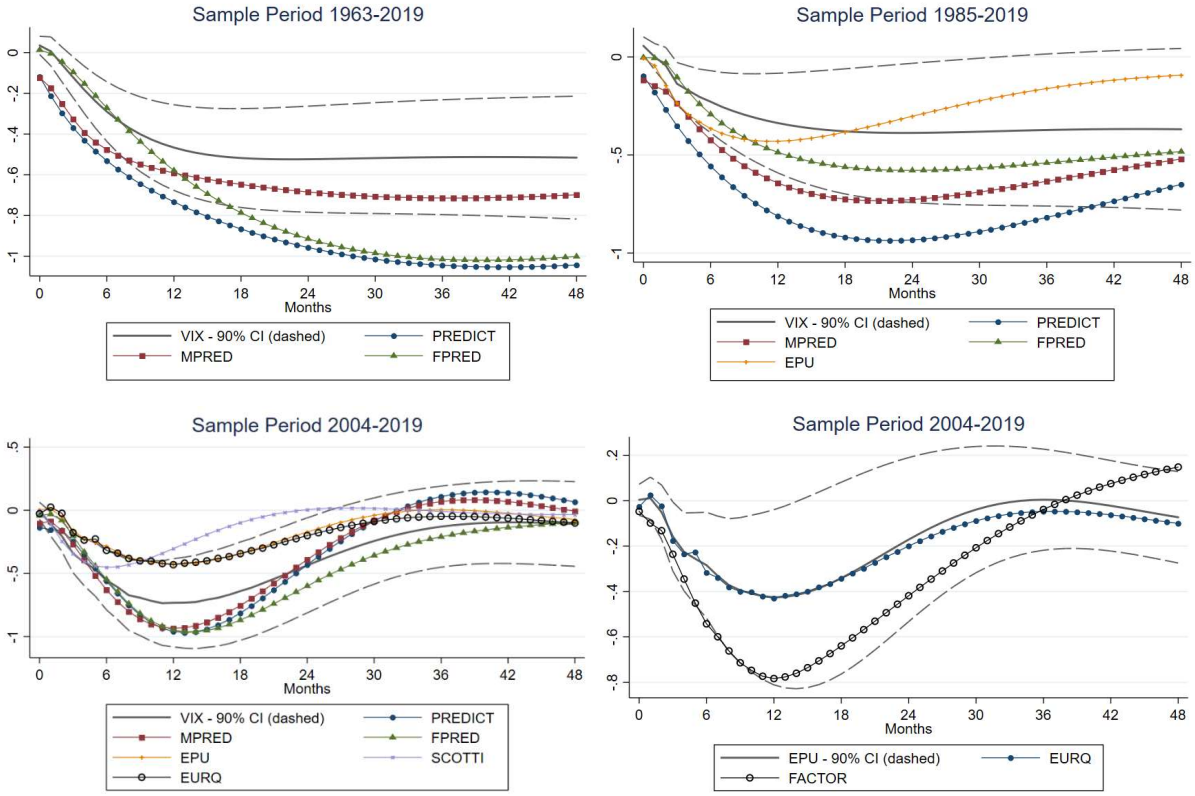
Identification is based on a 5-variable VAR(p), ordered as follows:

-log of manufacturing industrial production
-log of manufacturing employment
-log (1+Fed funds effective rate/100)
-log *SP500*
-alternative uncertainty measures

With respect to Figure 3 in Section 3, the order is completely reversed, with the industrial production at the top of the list, and the alternative uncertainty measures at the end.

All other details are the same as those reported in the notes of Figure 3 in Section 3.

Fig. A2.2 – Output responses to uncertainty shocks in VARs ordered as in Rossi et al (2020)



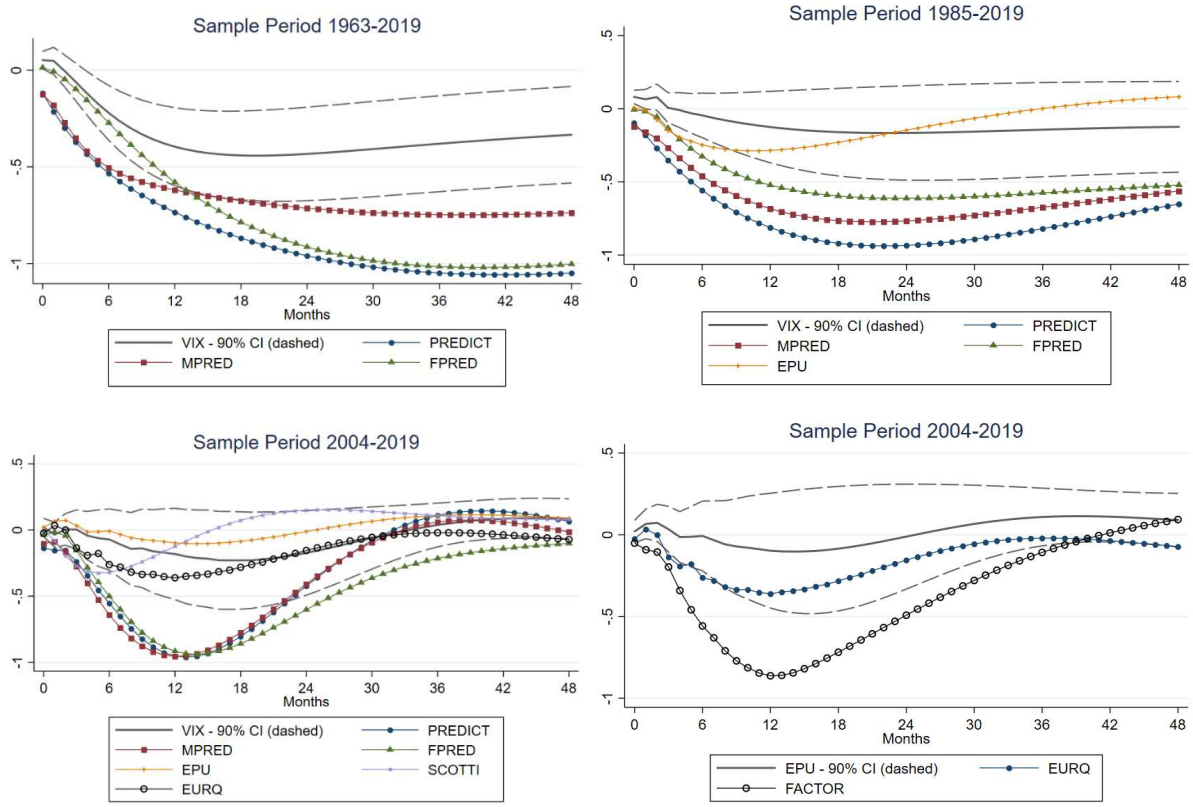
Identification is based on a 5-variable VAR(p), ordered as follows:

-alternative uncertainty measures
-log of manufacturing industrial production
-log of manufacturing employment
-log (1+Fed funds effective rate/100)
-log*SP500*

With respect to Figure A2.1, the alternative uncertainty measures are at the top of the list.

All other details are the same as those reported in the notes of Figure 3 in Section 3.

Fig. A2.3 – Output responses to uncertainty shocks in VARs ordered as in Bloom (2009)



Identification is based on a 5-variable VAR(p), ordered as follows:

-log*SP500*
-alternative uncertainty measures
-log (1+Fed funds effective rate/100)
-log of manufacturing employment
-log of manufacturing industrial production

With respect to Figure 3 in Section 3, SP500 is at the top of the list, while the alternative uncertainty measures are in the second place.

All other details are the same as those reported in the notes of Figure 3 in Section 3.