Nowcasting and the Use of Big Data in Short-Term Macroeconomic Forecasting: A Critical Review

Article · April 2019

DOI: 10.24187/ecostat.2018.505d.1966

1 author:

Pete Richardson
Llewellyn Consulting and formerly the Organisation for Economic Co-operation and Development (OECD) and

Nowcasting and the Use of Big Data in Short-Term Macroeconomic Forecasting: A Critical Review

Pete Richardson*

Abstract – This paper provides a discussion of the use of Big Data for economic forecasting and a critical review of recent empirical studies drawing on Big Data sources, including those using internet search, social media and financial transactions related data. A broad conclusion is that whilst Big Data sources may provide new and unique insights into high frequency macroeconomic activities, their uses for macroeconomic forecasting are relatively limited and have met with varying degrees of success. Specific issues arise from the limitations of these data sets, the qualitative nature of the information they incorporate and the empirical testing frameworks used. The most successful applications appear to be those which seek to embed this class of information within a coherent economic framework, as opposed to a naïve black box statistical approach. This suggests that future work using Big Data should focus on improving the quality and accessibility of the relevant data sets and in developing more appropriate economic modelling frameworks for their future use.

JEL Classification: C53, E27, E37
Keywords: Big Data, internet search, short-term, macroeconomic forecasting, models, nowcasting

** Senior Associate with Llewellyn-Consulting, London (pete.w.richardson@gmail.com)*
The author, formerly Head of the Macroeconomic Analysis Division in the Department of Economics, OECD Paris, is grateful to former colleagues including Nigel Pain, David Turner and Christophe André at the OECD and to Robert Kaufmann of Boston University for comments and suggestions on earlier versions of this paper and its presentation to the OECD's Working Group on New Approaches to Economic Challenges (NAEC) in Paris, January 2016 and to the UN Project LINK Conference in New York, October 2015. Thanks also go to the anonymous reviewers for helpful comments and suggestions on an earlier draft. Much of the study was originally carried out as part of the background research for the OECD's study "OECD Forecasts During and After the Financial Crisis: A Post Mortem", as reported in Pain et al. *(2014) and Lewis & Pain (2015).*

Received on 29 September 2017, accepted after revisions on 11 May 2018

To cite this article: Richardson, P. (2018). Nowcasting and the Use of Big Data in Short-Term Macroeconomic Forecasting: A Critical Review. *Economie et Statistique / Economics and Statistics*, 505-506, 65–87. https://doi.org/10.24187/ecostat.2018.505d.1966

Although much has been made of the possible role and uses of so-called Big Data in macroeconomic forecasting, there appear to be relatively few systematic reviews of related empirical work to date.[1] This paper seeks to redress the balance by providing a discussion of the relevance of Big Data for economic forecasting and a critical review of a number of empirical studies published to date, drawing on a number of different sources, including internet search and social media-related information and financial and other transactions-related statistics. It does so primarily from a practical economic forecasting perspective.

As noted by Bok *et al*. (2017), whilst "Big Data" is currently associated with those very large economic data sets derived from internet and electronic transactions sources, many of the related challenges to economists and statisticians existed well before their collection became feasible and pervasive for economics and other disciplines. Such challenges are exemplified by the pioneering work of Burns and Mitchell at the NBER[2] to identify business cycles using a very large range of data sets, the dedicated work of Kuznets and many others in developing consistent frameworks for the measurement of the National Accounts and related statistical concepts culminating in the large range of data collection and analyses currently undertaken. At the same time, developments in econometrics and time-series methods over past decades now permit the construction of consistent methods and suitable platforms for monitoring macroeconomic conditions in almost real time.[3]

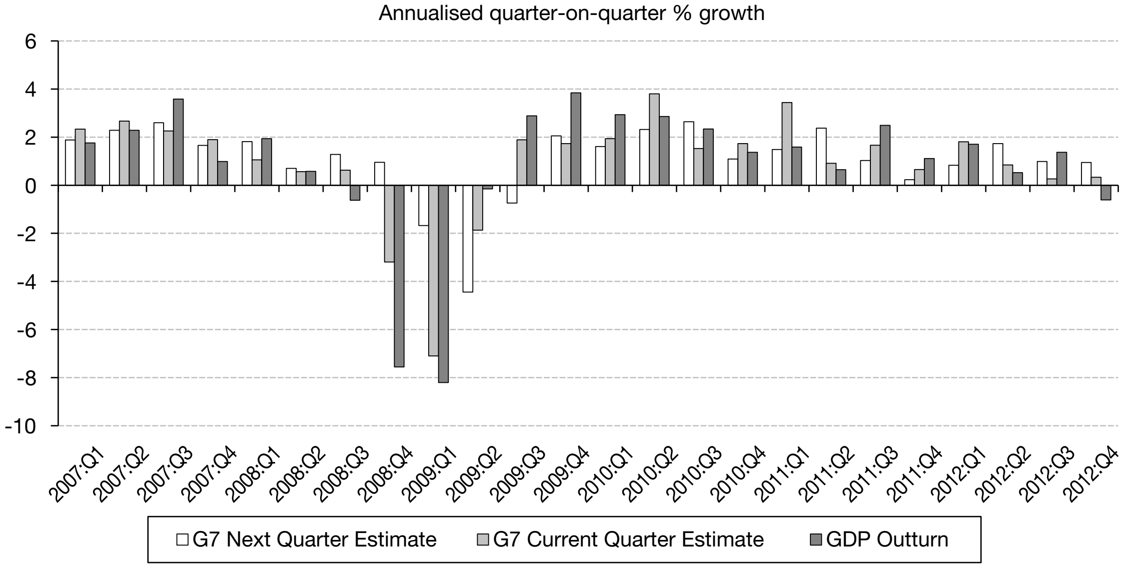
The main starting point and motivation for the present review came from an analysis of the OECD's international forecasting record during and after the financial crisis, as described by Pain *et al*. (2014) and Lewis & Pain (2015). In common with many national international institutions, and in line with more recent developments in so-called "nowcasting" techniques, the OECD's near-term macroeconomic assessments routinely take account of forecasts from a suite of statistical models using high frequency economic indicators to provide estimates of near-term GDP growth for the euro area and individual G7 economies for the current and next quarter.[4] These models typically use a Vector Autoregressive "bridge model" approach to combine information from a variety of "soft" indicators, such as business sentiment and consumer surveys, with "hard" indicators, such as industrial production, retail sales, house prices, etc., using different frequencies of data and a variety of estimation techniques. The associated estimation procedures are relatively automated and can be run as new monthly data are released, allowing also for timely updating and model choice according to the available information set.

Empirically, the main gains from using such an approach are typically found to be largest for current-quarter GDP forecasts made at or immediately after the start of the quarter in question, where estimated indicator models appear to outperform simple autoregressive time series models, in terms of both the size of predictive error and directional accuracy. Thus, the largest gains arise once one month of data is available for the quarter being forecast, typically two to three months before the publication of the first official outturn estimate for GDP. For one-quarter-ahead projections, the performance of the estimated indicator models is only noticeably better than simpler time series models once one or two months of information become available for the quarter preceding that being forecast. Modest gains are nonetheless made in terms of directional accuracy from using indicator models.

The general nature of these gains are illustrated in the Figure below which provides a summary of the successive revisions to the OECD's near-term quarterly GDP forecasts for the aggregate G7 economies during the 2008/9 downturn and subsequent recovery period. On this basis, the pre-recession period comparisons show relatively little systematic difference in predictive accuracy as between current- and next-quarter models (shown by the clear and lightly shaded bars respectively). However, from the second half of 2008, through the downturn and subsequent recovery, the current-quarter model predictions are clearly superior to the initial projections, reflecting the relative importance of hard information. The overall conclusion is that GDP indicator models provided a useful

1. *Useful background to the literature on Big Data sets and their uses in recent empirical studies are also given by Buono* et al. *(2017), Bok* et al. *(2017), Hellerstein & Middeldorp (2012), Hassani & Silva (2015) and Ye & Li (2017).*
2. *See Burns & Mitchell (1946).*
3. *In particular see the recent work of Giannone* et al. *(2008) and others, in developing consistent frameworks for near-term statistical analysis and so-called "nowcasting" by combining models for Big Data with modern filtering and estimation techniques.*
4. *At the OECD these models build on the pioneering work of Sédillot & Pain (2003) and Mourougane (2006) in using short-term economic indicators to predict quarterly movements in GDP by efficiently exploiting available monthly and quarterly information.*

Figure
The OECD's G7 GDP Current and Next Quarter Projections and Outturns Through the Financial crisis



Notes: Current quarter estimates for period 2007-Q1 - 2012-Q4: Mean error = - 0.1; MAE = 1.0; RMSE (actual) = 1.3; RMSE (estimation) = 1.6. For next quarter: Mean error = - 0.2; MAE = 1.6; RMSE (actual) = 2.6; RMSE (estimation) = 2.0.
Figure reports successive OECD forecasts and outturns for quarterly growth in real GDP for the G7 countries, over the period 2007 to 2012, based on the real time OECD short-term indicator models.
Sources: OECD, Pain *et al.* (2014).

basis for assessing current economic conditions through the recession at the point where hard indicator information became available, even though the scale of the global shock was entirely outside of the within-sample experience of the estimated models. Predictive performance was noticeably worse where hard indicator information was absent.

An important limitation in the practical use of indicator and similar nowcast models therefore concerns the lags in availability of hard statistical information, from National Statistical Offices and other statistical and survey agencies. Typically the quarterly goodness-of-fit and out-of-sample predictive performance of such models are found to improve significantly the more information is available for monthly hard indicators during the quarter in question, raising the question of how the availability of more timely from alternative sources might assist the tasks of short-term economic assessment and surveillance.

Big Data, Nowcasting and the Use of Electronic Indicators in Economic Forecasting

Reflecting these concerns, a number of recent academic and institutional studies, mostly post-crisis, have focussed on the possible usefulness of a wider set of data sources than those traditionally provided by the National Statistical authorities and in particular those described as being "Big Data" sets. The term Big Data, has been used in the computing industry field since the early 1990s to describe data sets with sizes which are, or were, typically beyond the ability of commonly used software tools and computer capacities to capture, manage, and process within a tolerable elapsed time, encompassing a wide range of unstructured, semi-structured and structured data sets. However, with the exponential growth of data storage and processing capacities over the recent past, the availability and use of Big Data sets have become increasingly feasible for economists and other analysts.[5]

In this context, a number of recent empirical studies have focussed on the possible usefulness to economic forecasting of three broad sources of such information:

- Internet search statistics, based on the frequency of searches for specific key words or topics;

5. *In economics, Diebold (2000) was first in describing Big Data as "the explosion in the quantity (and sometimes quality) of available and potentially relevant data, largely the result of recent and unprecedented advancements in data recording and storage techniques".*

- Internet-based social media and blog sources, such as Twitter;

- Detailed micro-level transactions data, recorded electronically by rapidly growing and popular financial payments and transactions systems.

The key advantages of using such sources of information lie in the coverage and level of detail they provide (down to individual micro transaction levels) and their timeliness. Being, in principle, available on a near real-time basis they offer a snapshot of current transactions, trends and tendencies well before they become recorded in official statistics. Nonetheless, key challenges remain in their use and development, including interpretation and analysis, as well as traditional concerns about their capture, curation, storage, sharing, visualization, and about privacy.[6]

Against this background, the following sections provide a discussion and critical review of recent studies using data from each of these three main areas for macroeconomic-related analyses and economic forecasting.[7] [8] To complement the review, an annotated summary guide to each of these studies, including their general coverage, the techniques employed and their principal findings and reservations is provided in Appendix.

The Use of Internet Search Information in Macroeconomic Models and Forecasting

Following the pioneering work of Ettredge *et al*. (2005), Choi & Varian (2009a and 2009b) and Wu & Brynjolfsson (2009), a growing body of literature has evolved on the use of internet search statistics in models used for economic forecasting and assessment. Typically such studies involve the construction of weekly, monthly and quarterly time-series indicators related to the "frequency" of internet searches for one or more specific keyword or phrase relevant to a specific topic or category of economic activity for a particular geographic location or country. For example, these might relate to searches for terms such as "welfare and unemployment benefit" or "mortgage foreclosure" or "car loans" etc., for "country A" or "state B". The relevant time-series indicator is then typically added to and tested for significance within a baseline forecasting model on within and out-of-sample bases. The underlying rationale is that internet search has now become a widespread and growing means for economic agents to obtain information relevant to their immediate economic situations, activities and decisions, those which ultimately get reflected in their behaviour and the wider set of economic statistics and accounts for a particular sector, concept or activity. Hence, the value of such an indicator for forecasting lies in its embodying relevant additional information available quickly, at high frequency and with a significant lead time over the transaction being recorded in official statistics.

Whilst earlier studies used raw internet search statistics from diverse search engines, Google Labs have since developed fairly refined tools within the Google Trends/Google Insights for Search website which enable individual researchers to recover tailor-made sample statistics on the frequency of searches for specific keywords and phrases by location and on a near real-time basis, starting from 2004. The relatively restricted historical samples available pose some limitations on their general usefulness for macroeconomic modelling, as does the sampling method which inevitably varies over time, as discussed below. Nonetheless a wide range of studies have emerged, originally focussing mainly on labour market indicators, but then widening to include housing, tourism, retail sales and consumption, housing markets, inflation expectations and financial markets, and for a range of countries.

Labour Market Studies

The earliest and most numerous set of studies using internet search indicators for economic forecasting are those related to labour markets and unemployment. The pioneering study by Ettredge *et al*. (2005), predating the use of Google Trends and other Big Data sources by several years, looks at US monthly unemployment over the period 2001-2004, using an internet search indicator of job-searches from a variety of internet sources. Using a relatively simple autoregressive forecasting model, it finds a significant relationship between search variables and published US unemployment data for adult males, one superior to the alternative use of official weekly claims data. Broadly similar results are reported for monthly total unemployment for Germany by Askitas & Zimmermann

6. *For a thorough and up-to-date description of the various Big Data sets available and their uses see also Buono* et al. *(2017).*
7. *In this respect, this review represents a snapshot of published studies available at or around Spring 2018.*
8. *This review does not include more recent works presented in this issue, which were not available at the time of writing.*

(2009) using Google Search statistics for the period 2004-2008, followed by Choi & Varian (2009b) for the United States, D'Amuri (2009) for Italy, D'Amuri & Marcuccio (2009) and Tuhkuri (2015) for the United States at aggregate and state levels, Suhoy (2009) for Israel, Anvik & Gjelstad (2010) for Norway and McLaren & Shanbhogue (2011) for the United Kingdom.

Most of these studies use a similar method of adding an internet-search indicator to relatively naïve time-series autoregressive models, in level or first-differenced terms. In some cases, most notably D'Amuri & Marcuccio (2009), more sophisticated models including other economic variables and leading indicators relevant to unemployment are used. Though sensitivity to the choice of baseline model and search keywords is often noted, most of these studies find the relevant internet-search indicator to be statistically significant and to provide superior out-of-sample performance compared with naïve baseline models and in some cases other relevant indicators, for example the US Survey of Professional Forecasters.

The more recent US study by Tuhkuri (2015) is generally more thorough in the choice and sophistication of data, statistical models and estimation techniques. The overall finding is that improvements in predictive accuracy from using Google search data appear robust across different model specifications and search terms, but are generally modest compared with previous studies and limited to short-term predictions, and that the informational value of internet search data tends to be somewhat time specific.

Consumption Studies

Studies of consumption, retail sales and car sales using internet search indicators include those by Choi & Varian (2009a, 2011), Kholodilin *et al*. (2010) and Schmidt & Vosen (2011) for the United States, Chamberlin (2010) for the United Kingdom, Bortoli & Combes (2015) for France, Toth & Hajdu (2012) for Hungary, and Carriére-Swallow & Labbé (2010) for Chile. Both the methods used and the results obtained vary considerably across these studies.

Some studies follow modelling strategies similar to those used for predicting unemployment by adding relevant internet search indicators to relatively naïve baseline time series forecasting models, whilst others include search indicators in combination with other measures of consumer sentiment or broad macroeconomic activity. For the United States, Schmidt & Vosen (2011) use more fully specified reduced form economic models of consumption which include lagged income, interest rates and stock market price variables. In most cases, internet-search variables are found to be significant either in their own right or in combination with other variables, though sometimes the gains are found to be relatively small. For car sales in Chile, Carrière-Swallow & Labbé (2010) find the introduction of car brand search indicators to significantly improve goodness-of-fit and predictive performance of baseline autoregressive models and also outperform broader measures of economic activity.

The results of Schmidt & Vosen (2011) in particular tend to show the individual significance of such variables to be greatest with simple AR(1) models, as might be expected (as discussed in a later section). Using more semi-structural consumption function specifications, they are found to perform as well as or in combination with the Conference Board Indicator, and the best one-month-ahead nowcasts are given by models including the Google Indicator. An interesting by-product of this study is the finding that the Michigan Consumer Sentiment indicator appears to have no additional predictive value.

Also of interest is the later study of consumption and new car sales by Schmidt & Vosen (2012), where Google-based indicators are found to be generally useful in modelling and predicting the effects of changes in motor vehicle scrapping schemes (so called "cash for clunkers" schemes) for the United States, France, Germany and Italy over the period 2002-2009. Such a finding suggests the possibly useful role in detecting and predicting the effects of special events or structural change at times when other timely information are not available. However, the authors note that the major challenges in such circumstances often lie in the identification of significant irregular events and constructing an appropriate measure from available search data.

The more recent Insee paper, by Bortoli & Combes (2015), reviews the usefulness of Google indicators for modelling French consumption at different levels of aggregation. The overall results are mixed, and suggest that search statistics improve monthly expenditure forecasts in only a limited way and for a narrow set of goods and services (clothing, food, household durables and transport).

Other Personal Sector Studies

Other largely personal sector-specific studies have involved housing market variables, tourism and inflation expectations. For housing markets, Webb (2009) finds high correlations between searches for "foreclosure" and recorded US home foreclosures, whilst Wu & Brynjolfsson find an internet search-based housing indicator significant and strongly predictive for US house sales and prices, as well as the sales of home appliances. Hellerstein & Middeldorp (2012) find similar improvements for predicting US mortgage refinancing, though the gains are found to be insignificant beyond a lead time of one week. McLaren & Shanbhogue (2011) report relatively strong results for UK house prices, with an internet search indicator out-performing other indicators over the period 2004-2011.

With regard to tourism and travel, Choi & Varian (2011) report significant results for Hong Kong tourism. Artola & Galen (2012) find similar results when adding Google based indicators to ARIMA models of the UK demand for holidays in Spain, although they also report considerable sensitivity to the choice of both baseline model and search keywords, particularly when used in different languages. Examining a range of inflation expectations indicators Guzmán (2011) finds that higher frequency Google-based indicators to generally outperform lower frequency traditional measures in use.

Financial Markets

A considerable number of studies have examined the relevance for search-based indicators for financial markets, though not specifically in a forecasting context. For example, Andrade *et al*. (2009) use such measures in identifying market volatility bubbles in the run up to the 2007 Chinese stock market bubble, Vlastakis & Markellos (2010) show strong correlations between search volume data by company name and trading volumes and excess stock returns for the 30 largest companies traded on the New York Stock Exchange.

Da *et al*. (2010, 2011) find similar correlations between product search variables and revenue surprises and investor attention for 3000 US companies, whilst Preis *et al*. (2012) find strong correlations between name searches and transactions volumes for the S&P 500 companies. Dimpf & Jank (2012) also report strong co-movements between Google company name searches (as a measure of investor attention) and US stock market movements and volatility, with Google search indicators providing better out-of-sample forecasts than ARIMA models. Hellerstein & Middeldorp (2012) find a Google search indicator to be significant in modelling movements in certain dollar-Renminbi forward market variables, but with low predictive power.

Overall, the lack of firm evidence or forecasting applications in the financial market area is perhaps of less importance given the wider availability of high frequency statistics for financial market variables.[9]

Wider Macroeconomic Studies

In contrast to the previous studies, where internet search-based indicators are included directly as explanatory variables in regression models for individual economic variables, Koop & Onorante (2013) use a different approach by introducing Google search-based probability measures into a dynamic model switching (DMS) nowcasting system, one in which current outcomes are regressed on lagged values of the set of dependent variables and Google indicators. That is, instead of using internet search volumes as simple regressors, they also allow them to determine the weights given to alternative nowcasting equation estimates over time. The intuition here is that internet search information may provide the researcher with useful information about which macroeconomic variables are most important to economic agents concerns and expectations at given points in time. This would make sense, for example, in a context where the underlying economic structure is not constant, and are therefore particularly suited to deal with sudden unexpected events like financial crises or recession.

Applying this method to models for monthly US data for a selection of monthly macroeconomic variables (including inflation, industrial production, unemployment, oil prices, money supply and other financial indicators), they find dynamic switching models to be generally superior to others, regardless of whether these models involve search-based probabilities

9. This contrasts with studies of financial markets based on social media indicators, as described in a later section, where high frequency forecasting is a very specific focus of interest.

or not. Firstly, the inclusion of search data is found in many cases to give improvements in nowcast performance, complementing the existing literature by showing that internet search variables are not only useful when dealing with specific disaggregate variables, but can be used to improve nowcasting of broad macroeconomic aggregates. Secondly, they also find that information from search variables is often best included in the form of model probabilities as opposed to simple regressors. The general results are however somewhat mixed across variables, being most positive for inflation, wage, price and financial variables, inconclusive for industrial production and strongly inferior for unemployment.

Limitations of the Use of Internet Search-Based Indicators

Although broadly supportive of the general usefulness of internet search-based measures for short term assessment and nowcasting for a variety of economic variables, many of the above studies note that results tend to be mixed across topics and subject to a number of specific limitations and possible biases, reflecting both the qualitative nature of the data sets and the modelling frameworks in use.

The Data Sets

Firstly, it should be noted that the various measures do not specifically correspond to the absolute number of searches but rather the proportion of searches carried out on a particular subsample using specified keywords or topics over a particular time period, suitably scaled. For this reason the data sets used often need to be "cleaned" subjectively for specific outliers, one-off events or aberrant search terms which might otherwise swamp the data.[10] At the same time, by their very nature, high frequency internet search-based indicators draw on a variable and non-stratified sample, one which changes continuously over time. Both of these factors are likely to add noise to the underlying measures and make them more qualitative than they seem at first sight. Indeed in many cases the qualitative nature of internet search statistics begs the question of the general nature of the underlying relationship e.g. with regard to the scale, linearity or even the sign.[11] Secondly, the shortness of the available samples for internet search information dating from the mid-2000s limits the scope for the stability and testing within a range of existing models, both statistical and structural.[12] Most studies therefore rely on relatively short samples of high-frequency data which are also sometimes subject to strong seasonality, which risks swamping the underlying relationships. At least visually, this seems to be the case for a number of early studies claiming to illustrate close historical relationships between the search indicator and variable in question.

A number of studies also note the sensitivity of results to the choice of keywords and baseline models.[13] The former is necessarily a handicap which implies the need for care in the construction of an indicator targeted for a specific use. Much is left to the individual researcher, to design/construct their own indicators – which has considerable advantages for use in specialist areas – but to date there appear to be no standardised published measures available for specific purposes such as general macroeconomic surveillance or analysis at national or international levels.

The Modelling Frameworks

Concerning sensitivity to the choice of baseline models, it is worth noting that, with few exceptions[14], studies reporting high significance or superior out-of-sample forecasts often do so by comparison with relatively naïve low-order univariate autoregressive time-series baseline models. These results are probably not surprising then, to the extent that without additional information such models are seldom able to provide more than smooth short-term projections adjusting recent out-turns to longer term trends and hence fail to pick up erratic short-term movements or turning points.

Relatively few studies appear to have been done to systematically test or embed internet search-based variables within existing indicator model frameworks used to forecast

10. *For example, the death of Michael Jackson in June 2009 resulted in a huge surge in internet search activity, with a major negative effect on the relative shares of searches for all other topics in that period.*
11. *For example, the intensity of search activity for a range of economic variables might be associated with both positive and negative movements in the variable in question and may be time or episode specific.*
12. *For example, see the comments of Chamberlin (2010), Schmidt & Vosen (2012) and Bortoli & Combes (2015).*
13. *For example, see the comments of Artola & Galen (2012), Askitas & Zimmermann(2015), Chamberlin (2010) and Tkacz (2013).*
14. *Notable exceptions here include D'Amuri (2009), D'Amuri & Marcuccio (2009), Schmidt & Vosen (2011).*

near-term movements or turning points in key GDP or trade aggregates, or to augment or predict other high-frequency indicators significantly ahead of publication. Important exceptions here are found in the work of Koop & Onorante (2013) in combining search information with dynamic probability switching models, and Galbraith & Tkacz (2015) in the testing and use of internet search variables within more extensive indicator systems.

A relatively small subset of studies does however successfully use search-based indicators to augment and improve more conventional economic and/or indicator-based models or to allow for special factors in specific relationships at macro and sectoral levels. Although much of the literature also aims to improve the detection of turning points, very little seems to have been done to systematically test or embed internet search-based variables within existing indicator and bridge-model frameworks used to forecast near-term movements in key GDP or trade aggregates, or to augment/predict other high frequency indicators significantly ahead of publication. Further work in all the above areas would seem necessary to exploit the key advantages of internet search-based indicators over other indicators, as the relevant data sets are extended and improved over time.

The Use of "Social Media" and Twitter Based Information in Macroeconomic Modelling

In many respects social media data sets, such as those embodied in Twitter and other user-based blogs, are potentially richer and therefore have important advantages over indicators based on internet search frequencies:

- Sample sizes are often considerably larger and available on a virtually continuous basis;

- The data are more varied in scope, with greater general and specific detail of posts;

- They permit a more stratified approach, by analysing information coming from selected representative samples or well-defined user groups;

- The absence of pre-preparation/filtering by data proprietors, as with Google Trends, may be an advantage or disadvantage.

Social media blog entries and Tweets can be about any topic, being totally up to the user what they choose to broadcast. For the most part they are publicly available either directly in raw form or indirectly through social media Application Programming Interfaces (API's). This makes them an increasingly accessible and popular source of information for researchers to construct general and specific mood or intentions indicators at a given place and time, and for particular topics of interest.

Financial Markets

To date, the large majority of published empirical studies using social media data as input to economic models and forecasting,[15] are relatively near-term and in the area of stock market prices and finance. Gilbert & Karahalios (2010) for example use a dataset of over 20 million LiveJournal posts, to construct a measure of public anxiety (the Anxiety Index). This is based on a panel of 13 thousand LiveJournal contributors, chosen by linguistic classifiers on the basis of entries for 2004, as a sub-sample known to frequently express varying degrees of general anxiety (not specifically economic events). This sub-sample is then used to construct the Anxiety Index based on their daily blog posts through 2008 and tests carried out for its possible "influence" on the S&P500 stock market index, using a baseline statistical relationship involving lagged index values and the lagged levels and changes in the volume of transactions.

Using a combination of regression and Granger causality tests, the broad conclusion is that the Anxiety Index contains statistically significant information not apparent from the market data. The authors note however that this result is somewhat weakened by further testing for the inclusion of the existing Chicago Board Options Exchange VIX index[16], which in some models tends to dominate the Anxiety Index. Even so, general collinearity with the VIX is seen as a possible validation of the usefulness of the more broadly based Anxiety Index as a measure of stock market uncertainty. The authors nonetheless note that more work

15. *Previous applications based on social media and so-called mood indicators cover a fairly wide range of topics including ; book sales (Gruhl* et al.*, 2005); cinema box office receipts (Mishne & Glance, 2005 and Liu* et al.*, 2007); influenza pandemics (Ritterman* et al.*, 2009); TV ratings (Wakamiya* et al.*, 2011); and election results (O'Connor* et al.*, 2010; Tumasjan* et al.*, 2010).*

16. *The VIX index is a popular measure of the stock market's expectation of volatility implied by S&P 500 index options, calculated and published by the Chicago Board Options Exchange (CBOE), colloquially referred to as the* fear index *or the* fear gauge*, see Brenner & Galai (1989).*

needs to be done in overcoming the inherent difficulties in interpreting blog-based information and potential ambiguities, as well as potential index volatility associated with non-economic external events and, importantly, that the sample year 2008 was exceptional in many respects.

A number of parallel studies have looked only at correlations between the social media-based mood indicators and relevant economic variables, rather than in forecasting models. For example, Zhang *et al*. (2010) examine a very large sample of daily Twitter entries between March and September 2009 to estimate a variety of measures of differing degrees of positive and negative moods, ranging from fear to hope. These are then correlated against corresponding values of the Dow Jones, NASDAQ and S&P500 indices, as well as the VIX index. Statistically significant correlations are found, consistent with negative impacts of lagged mood indicators on current stock market prices and the VIX. However, the authors note that this holds for both positive and negative mood indicators, indicative of the relative importance of emotional outbursts as opposed to specific mood directions over the sample period.

Along similar but more formal lines, Bollen *et al*. (2011), examine the relationship between mood indicators derived from large-scale Twitter feeds and changes in the Dow Jones index over time.[17] Specifically the text content of daily Twitter feeds are analysed using two mood tracking tools from March to December 19, 2008. The first tool, OpinionFinder, analyses the text content of tweets to give a daily time series of the positive vs. negative balance of the public mood. The second tool, the Google-Profile of Mood States (GPOMS), analyses text content to provide a more detailed view of changes in public sentiment using six different mood states (Calm, Alert, Sure, Vital, Kind, and Happy). The resulting indicators are then correlated against the Dow Jones index on a daily basis, using a general autoregressive model and Granger causality testing framework. The authors conclude that results support the view that the accuracy of stock market prediction models is significantly improved (by around 6%) when some but not all mood dimensions are included.[18] In particular, variations along the public mood dimensions of “Calm” and “Happy” as measured by GPOMS appear to have some predictive value, but not the overall balance of optimism and pessimism as measured by OpinionFinder.

Following up on this work, Mao *et al*. (2012), focus more closely on the relevance of finance-specific Twitter information as opposed to general positive and negative mood expressions. Specifically they examine the relationship between the daily number of tweets that mention S&P500 stocks and associated stock prices and traded volumes at the aggregate level, for each of 10 industry sectors and at the individual company level, for Apple Inc. This is done through correlations between daily stock market measures over an approximate 3 month period (February to May 2012) and the Twitter volume indicators. The analysis is then extended using simple linear autoregressive regression models, to predict the stock market indicators with the Twitter data indicator as an exogenous input. The overall results are fairly mixed and vary between different levels of aggregation.

Significant correlations are found at the aggregate level between the Twitter indicator and both levels and changes in prices, though not trading volumes. For 8 out of 10 industry and corporate sectors (notable exclusions being consumer discretionary and stable categories), statistically significant correlations are found for the levels of traded volumes but not prices. For the financial sector and Apple Inc. (the most highly tweeted categories) correlations are statistically significant for both volumes and prices. These results are broadly mirrored in the tests for predictive accuracy with the Twitter indicator improving forecasts for volumes and prices at the aggregate and financial sectors but for volumes only for Apple Inc. Even so, the predictions of directional changes in the sample period are at best 68% accurate for the aggregate and financial sectors and only 52% for Apple Inc. i.e. close to a random walk. The authors conclude that the relevant correlations are statistically significant and help predict some stock market movements on a daily basis, although more work is required to refine the choice of search words, to screen for spurious tweets, to collecting longer-term data, and to combine indicators for the number, relevance and sentiments of individual tweets.

17. For a similar but more micro and higher frequency approach see Wolfram (2010).
18. As discussed further below, this “landmark” result is hotly disputed by other authors, see Lachanski & Pav (2017).

Noting the scope for measurement and classification errors associated with computational machine learning-based processing of blog-related data sources, the further work of Mao *et al*. (2014) focus instead on a simpler set of indicators, based on the frequency of use of terms related to market "bullishness" or "bearishness" in both Twitter posts and Google search queries.[19] These are calculated on daily (for Twitter) over the period 2010 to 2012, and weekly (for Google Trends) bases over the period 2007 to 2012, and then compared with other investor sentiment indicators. Relative predictive powers are then analysed in the context of small dynamic models of the US, UK Canadian and Chinese stock market prices and returns.[20] Comparing between measures and adjusting for frequencies, Twitter-based measures of market bullishness are found to lead and "predict" changes in corresponding Google-based measures, whilst both measures are found to be positively correlated with, and lead established investor sentiment surveys for the United States.[21]

Using a fairly detailed dynamic VAR modelling framework for the United States (one also including trading volumes and other sentiments indicators as explanatory variables), the Twitter-based indicator is found to be both statistically significant and provide better predictions of stock returns on a daily basis. An additional feature is that high levels of Twitter bullishness are found to be associated with changes in daily stock returns over the following days, with there being a reversion to normal levels within the next two to five days. The corresponding Google-based indicator is also found to be statistically significant but with lower predictive power, attributed to its low frequency and lack of relevant dynamics. Similar correlation results are also found for the UK, Canada and China, (within simpler bi-variate models), but with lower predictive power for China. The Google indicator is also found to be significantly correlated with all four stock market prices but with lower predictive power. The authors note that the overall results are promising in terms of predictive correlation but are less clear with regard to causality, which remains a challenging research problem for Big Data analysis and the development of appropriate experimental design methods and machine learning algorithms.

Other notable contributions to the finance literature using Twitter-based indices include Arias *et al*. (2012), which applies complex decision tree computer algorithms to Twitter based information to analyse movie box office sales and stock market prices, and Ranco *et al*. (2015), which examines the impact of Twitter-based measures of so-called "event study" effects on the stock returns of 30 leading companies within the Dow Jones index between 2013 and 2014. On a more detailed basis, Bartov *et al*. (2015), covering 300 companies over the period 2009 to 2012, examine whether aggregate opinion in individual tweets about a firm can help predict the firm's earnings and stock returns around earnings announcements, and whether the ability to predict abnormal returns is greater for firms in weaker information environments.

The Twitter-based literature, emanating initially from computational information science and artificial intelligence studies, is not without critics within the economics and finance world. Indeed, a recent review by Lachanski & Pav (2017) strongly criticises both the general approach and the results of Bollen *et al*. (2011), which they consider incompatible with both information theory and the investor-sentiment based text mining. Attempting to replicate similar mood indicators and models, they find some in-sample but almost no out-of-sample correlation with the Dow Jones index. Whilst this might be attributable to minor differences in data coverage and the selection of the time period studied, they conclude that Bollen's results are very much an outlier and that there is little or no credible evidence that Twitter-based measures of general collective moods can be used to forecast index activity on a daily basis. Overall, they argue that the Bollen *et al*. (2011) study is fundamentally flawed and has contributed to a "growing deadweight loss to the finance literature".

Labour Markets

To date, there appear to be relatively few published Twitter-related economic studies outside the area of financial markets. An important exception has been the work on labour market's by Antenucci *et al*. (2014) at the University of Michigan, in developing measures of labour market flows from social media data. Specifically

19. A specific advantage of such an approach is that these terms are used fairly unambiguously and in a focussed way to refer to financial market conditions.

20. Specifically for the US, they examine a number of market indicators including the Dow-Jones and S&P indices; for the UK, the FTSE100; for Canada the S&P/TSX; and for China the SSE index.

21. These include the Daily Sentiment Index and the US Advisors' Sentiment Report of Investors Intelligence.

large sample Twitter-based data were used to produce indexes of job loss, job search and job posting as a means of analyzing high frequency weekly estimates of job flows from July 2011 to early November 2013. Measures are first derived from the frequency of use of job loss and search-related phrases in the sample of Tweets, which are then combined into composite measures using their principal components to track initial claims for unemployment insurance at medium and high frequencies. The resulting index is found to have a greater signal to noise ratio than initial claims data which might be of value to policy makers needing high-frequency, real-time indicators. Over the sample period, the indicator is found to account for 15 to 20 percent of the variance of the prediction error of the consensus forecast for initial claims. The index was also considered useful in providing realtime indicators of events such as Hurricane Sandy and the 2013 government shutdown, although this body of work is currently said to be under revision since the original model began to deviate in its estimates around mid-2014.

The Limitations in the Use of Social Media in Forecasting Studies to Date

Overall, the challenges in the extraction and use of social media based data sets are considerable and perhaps greater than those involved with internet search material. Typically the researcher has to devise methods of searching across large sets of blog entries to identify within a given sample and timeframe the frequency of the use of specific phrases or keywords by those posting on blogs. For example, this might include looking for phrases indicating concepts like job security and job loss, company and consumer product names or those used more generally to indicate degrees of anxiety or confidence with respect to life in general or more specifically economic and financial conditions. Thus whilst being richer in content, they are also, arguably, more exposed to differences in linguistics, interpretation and nuances in the use of language, than for internet search related data.

For these reasons, and given the huge volume of data being processed, much of the work in this area builds on developments in the informatics, machine learning and artificial intelligence domain, for the design and application of sophisticated automated filters to mine the information content of simple text blog entries. Indeed it is notable that much of the original literature originates in the study of computational, linguistic and machine learning methods as opposed to economics and finance. For this reason, these studies are not always embedded in the sound and familiar theoretical and empirical frameworks more commonly used in other areas of economic research and econometrics. Although these studies may often embody "state of the art" computational machine learning techniques, there is relatively little evidence of testing of one measure or method against another to see whether all the "bells and whistles" are superior to simpler frequency balance measures.

Similarly, there is certain sense of searching for a "Holy Grail" financial market indicator which is both broadly based and able to explain, predict or, at best, correlate with chosen financial variables. Possibly because of the enormous sample sizes involved with raw high frequency data, the chosen time samples often seem to be idiosyncratic and restrictively short, as noted by Lachanski & Pav (2017). In this context, the more recent work of Mao *et al*. (2015) focussing on simpler variables that are more narrowly defined to be of relevance to financial markets over a longer sample period and comparing between measures may be more rewarding. Even so, there is often an excessive focus on very (daily) near-term predictive power and in having more detailed and workable model for US stock prices, as opposed to those for other economies. Both factors clearly limit their overall relevance for macroeconomic analysis as opposed to profit-driven trading applications.

Similar to the body of research based on internet search variables, the models used in many social media-based studies are almost exclusively statistical and, in the absence of other explanatory variables, may be too simple to tell much about the underlying dynamics or relative predictive values of the different indicators being analysed. In addition, a surprising and perhaps important omission in these studies is the fact that financial markets are inherently international and therefore linked to each other and influenced by other global phenomena.

The Structure and Uses of Other Big Data: Electronic Transactions and Confidence Indicators

To the extent that a large and growing share of global financial and commercial transactions are supported by electronic payments and transactions systems, there has also been growing interest in the use of high frequency statistics from a number of such sources as indicators within informal and formal forecasting and assessment frameworks. Typically, these systems cover a range of different detail and frequencies, down to the individual transaction level. As a result, confidentiality and proprietorial ownership rights pose important limitations on their uses beyond the privileged few.

SWIFT Trade Transactions Indicators

In this context, recent ICC Global Surveys of Trade and Finance and recent EBRD blog reports draw particular attention to the use of SWIFT indicators in tracking trade credit and the volume trade transactions.[22] [23] Whilst making a number of important caveats about the form and coverage of this type of data, both reports provide useful illustrations of the sharp year-on-year decline in SWIFT trade-related messages (accounting for a significant share of trade letters of credit) from end-2008 to end-2009 and later in early 2011 and their relationship with global and regional trends in trade over the same periods.

On a more country-specific basis, a recent Australian Reserve Bank study[24] examines the possible use of various electronic indicators of wholesale and retail payments from commercial banks in forecasting a range of macroeconomic aggregates including consumption, domestic demand and GDP. The overall results are mixed in finding that a SWIFT payments indicator used in combination with the principal components of other more conventional short-term macroeconomic indicators, significantly improves short-term predictive performance relative to naïve autoregressive baseline models, whilst other retail payments indicators including credit card transactions do less well.

Following the same general idea, SWIFT, in collaboration with CORE Louvain, has constructed a number of global and regional indicators for use in specific nowcasting applications.[25] In particular, SWIFT (2012) reports the use of an OECD aggregate index of filtered transactions in a suite of GDP bridge models, finding the most significant results using a dynamic mixed frequency forecasting model for quarterly movements in OECD real GDP for the period 2000 to 2011. As with most of the internet search indicator-based studies, the underlying baseline model is a relatively simple statistical ARMA model, taking account of no other relevant information.

An important caveat to these studies is that SWIFT indicators generally relate to the volumes of messages as opposed to the levels or values of transactions and therefore need to be carefully filtered for message versus transactions content and for their coverage, as between trade, financial and other activity-related transactions. Nonetheless, the broad results to date are generally supportive of the broad approach, and having the advantage of being available for a longer sample period, merit further investigation within a wider range of indicators and economic aggregates.

Payments Transactions Statistics

The recent Bank of Canada study by Galbraith & Tkacz (2015) reports an interesting approach combining a range of financial and transactions indicators within a set of mixed frequency GDP indicator models. These models combine measures of the growth in values and volumes of monthly and quarterly Canadian debit, credit and cheque transactions cleared through the Canadian Payments Association (CPA) on a daily basis, with composite leading indicators for the US and Canada, monthly unemployment rates and lagged GDP growth. A key finding is an improvement in the accuracy of the earliest GDP predictions through the inclusion of debit card payments observed for the first two months of the prediction period, although such improvements are not detectable once the previous quarter's GDP value is observed. Overall, this supports the possible value of combining electronic transactions with other data measurable on a daily basis. An obvious limitation to the use of this class of information is its confidentiality and inaccessibility to the general public for research uses, even in processed form.

22. In particular, see the Global and Regional trends sections of the ICC Reports "Global Survey of Trade and Finance: Rethinking Trade Finance", for 2010, 2011 and 2012, and the EBRD blogs "Trade Finance on the Way to Recovery in the EBRD Region", January 2011 and "Rising uncertainty for trade finance as IFI additionality increases", February 2012.

23. The Society of Worldwide Interbank Financial Telecommunication (SWIFT) network covers the financial transactions of over 10,000 financial institutions and businesses worldwide (210 countries).

24. See Gill et al. *(2012).*

25. In particularly see "The SWIFT Index: Technical Description", SWIFT, February 2012.

ADP Employment Indicators

A further example of the use of real-time transactions systems data is given by the work reported in the ADP National Employment Report (2012) for the United States.[26] In this study monthly and bi-weekly payroll data processed by the ADP's system (responsible for the payrolls of establishments covering approximately 20% of U.S. private sector workers) are filtered and classified by size and industry to provide pair-wise matches with the sample used in producing BLS monthly employment data. A set of adjusted sectoral ADP indicators are then used, in conjunction with the Philadelphia Federal Reserve ADS Business Conditions Index[27], to estimate a system of VAR equations to predict monthly changes in BLS private employment data by sector, since April 2001. Although the significance of individual variables is not reported and it is unclear what restrictions are placed across individual parameters/sectoral contributions, the overall in-sample correlations appear to be relatively high (0.83 to 0.95) and the models appear to track overall monthly movements in BLS employment for the total private sector and 5 broad sectors fairly closely.

The Ceridian-UCLA Pulse of Commerce Index

Another "Big Data" indicator of interest for high frequency analysis for United States activity is the Ceridian-UCLA Anderson Pulse of Commerce Index (PCI). Essentially this index is based on Ceridian electronic card payment services for US diesel sales to freight haulage companies. In principle, the transactions data can be tracked and analysed on a yearly, monthly, weekly and daily basis by location and volumes of fuel purchases to provide a detailed high-frequency picture of US road trucking activities including interstate highways and cities, shipping ports, manufacturing centres and border crossings with Canada and Mexico. The PCI's main advantage over other economic indicators is its basis on real-time, actual fuel consumption data in advance of published monthly statistics. Its main disadvantage for economic research is that it is not freely available within the public domain. To date no published analytical studies appear to be available, although UCLA Anderson produce a monthly newsletter 4 to 5 days in advance of the publication of monthly industrial production data and reports that back-testing to 1999 shows the index to closely match growth in real GDP and changes in Industrial Production.

* *
*

The forecasting experience of many national and international forecasters during the 2008-2009 recession and beyond have been similar in that existing models, methods and analyses were not particularly well suited to predict or analyse the scale of the crisis. This very much reflected the nature of the underlying situation, the lack of systematic evidence of the scale of the financial shock, the nature of the international linkages involved and the mechanisms by which financial shocks translated into shocks to the real economy. By contrast, short-term indicator and nowcasting models for world trade and GDP for the G7 economies have sometimes proved to be extremely useful and more accurate, within a current quarter or nowcasting situation. Even so, such models appear to have been limited or poor in going much beyond the current quarter and detecting possible turning points, reflecting the limitations of "soft" survey-based activity indicators and a lack of "hard" information. Overall, such a situation suggests a research priority in improving the timeliness and availability of relevant information.

On the basis of a review of the recent academic literature, a broad conclusion is that internet search and social media based indicators and other Big Data sources provide a novel and possibly useful means of measuring various aspects of consumer and business behaviour in an almost a real time basis. At the same time they may embody information which cannot be captured by other economic indicators or be available on such a timely basis. For this reason, they warrant further development and monitoring in parallel with other macroeconomic indicators.

The range of available empirical studies reviewed provides interesting insights and evidence of significant correlations and predictive performance across a range of topics. Overall, however, the results are generally quite mixed, reflecting both the relative simplicity of the

26. *See "The ADP National Employment Report", Automatic Data Processing Inc. and Moody's Analytics, October 2012.*
27. *The ADS business conditions index is based on the framework developed in Aruoba* et al. *(2009). The index takes on board a combination of high and low frequency indicators, including weekly initial jobless claims; monthly payroll employment, industrial production, personal income less transfer payments, manufacturing and trade sales, and quarterly real GDP.*

models used and important limitations in terms of quality, form, sample sizes and their "qualitative" nature. In these respects more needs to be done to:

- Refine and improve the quality standards for available Big Data sets and their accessibility;

- Develop better methods for extracting relevant economic information relevant to specific fields of economic research;

- Improve the means of comparing and testing between alternative measures;

- Further adapt and improve relevant testing and modelling frameworks, to be more useful to the task of incorporating near-term information in short-term macroeconomic forecasts.

Nonetheless, there are some clear examples where such indicators could usefully augment existing nowcasting and other indicator-based approaches as part of the general selection of variables to be analysed. Now that the first wave of such studies seems to have become less prominent in the literature it would be useful to take advantage of some of these lessons in the design of ongoing future work, rather than see the general topic dismissed as a fad or dead end.

Big Data sources including transactions based and other financial indicators have been more thinly used, until very recently. The results obtained to date also seem rather mixed, although trade finance indicators do appear to have given the right signals prior to and during the financial crisis. They also show some promising features but are equally limited in terms of information content and transparency. As for internet search and social media-based studies, they warrant further investigation in the context of statistical and semi-structural economic and indicator frameworks. In contrast to internet based indicators, this class of data is, in the main, subject to wider concerns about their confidentiality and is therefore available, so far, to a relatively small audience, mostly central bankers and statisticians. Important priorities in this area are therefore to develop suitable quality standards and to improve their accessibility for statisticians and economic researchers in a suitably relevant and condensed form.

The overall message is that Big Data sets provide new and useful sources of information for economic analysis, but also warrant further refinement, development and monitoring in parallel with other macroeconomic indicators and forecasting techniques. As such they are a welcome addition to the economist's and statistician's toolkit for short-term analysis. ☐

BIBLIOGRAPHY

ADP & Moody's Analytics Enhance (2012). *ADP National Employment Report*.
http://mediacenter.adp.com/news-releases/news-release-details/adp-and-moodys-analytics-enhance-adp-national-employment-report/

Andrade, S. C., Bian, J. & Burch, T. R. (2009). Does information dissemination mitigate bubbles? The role of analyst coverage in China. *University of Miami Working Paper.*

Andrade, S. C., Bian, J. & Burch, T. R. (forthcoming). Analyst Coverage, Information, and Bubbles. *The Journal of Finance and Quantitative Analysis*, 48(5), 1573–1605.
https://doi.org/10.1017/S0022109013000562

Antenucci, D., Cafarella, M., Levenstein, C., Ré, C. & Shapiro, M. (2014). Using Social Media to Measure Labour Market Flows. University of Michigan, *NBER Working paper* N° 20010.
https://doi.org/10.3386/w20010

Anvik, C. & Gjelstad, K. (2010). "Just Google It!"; Forecasting Norwegian unemployment figures with web queries. CREAM Publication N° 11.
http://hdl.handle.net/11250/95460

Arias, M., Arratia, A. & Xuriguera, R. (2014). Forecasting with Twitter Data. In: *ACM Transactions on Intelligent Systems and Technology*, 5(1), 1–24.
https://doi.org/10.1145/2542182.2542190

Armah, N. (2013). Big Data Analysis: The Next Frontier. *Bank of Canada Review*, Summer 2013, 32–39.
https://www.bankofcanada.ca/wp-content/uploads/2013/08/boc-review-summer13-armah.pdf

Artola, C. & Galen, E. (2012). Tracking the Future on the Web: Construction of leading indicators using Internet searches. *Bank of Spain Occasional Paper* N° 1203.
https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosOcasionales/12/Fich/do1203e.pdf

Aruoba, S. B., Diebold, F. X. & Scotti, C. (2009). Real-Time Measurement of Business Conditions. *Journal of Business and Economic Statistics,* 27(4), 417–427.
https://doi.org/10.1198/jbes.2009.07205

Askitas, N. & Zimmermann, K. F. (2009). Google Econometrics and Unemployment Forecasting. *Applied Economics Quarterly,* 55(2), 107–120.
https://doi.org/10.3790/aeq.55.2.107

Bartov, E., Faurel, L. & Mohanram, P. (2015). Can Twitter Help Predict Firm-Level Earnings and Stock Returns? Rotman School of Mnagaemnet, *Working Paper* N° 2631421, July 2015.
https://dx.doi.org/10.2139/ssrn.2782236

Bok, B., Caratelli, D., Giannone, D., Sbordone, A. & Tambalotti, A. (2017). Macroeconomic Nowcasting and Forecasting with Big Data. New York Federal Reserve *Staff Report* N° 830, November 2017.
https://www.newyorkfed.org/research/staff_reports/sr830

Bollen, J., Mao, H. & Zeng, X.-J. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1–8.
https://doi.org/10.1016/j.jocs.2010.12.007

Bortoli, C. & Combes, S. (2015). Contribution from Google Trends for forecasting the short-term economic outlook in France: limited avenues. Insee, *Conjoncture de la France*.
https://www.insee.fr/en/statistiques/1408911?sommaire=1408916

Brenner, M. & Galai, D. (1989). New Financial Instruments for Hedging Changes in Volatility. *Financial Analysts Journal*, 45(4), 65–71.
https://www.jstor.org/stable/4479241

Buono D., Mazzi, G. L., Kapetanios, G., Marcellino, M. & Papailas, F. (2017). Big data types for macroeconomic nowcasting. *Eurostat Review on National Accounts and Macroeconomic Indicators*, 1/2017, 93–145.
https://ec.europa.eu/eurostat/cros/system/files/euronaissue1-2017-art4.pdf

Burns, A. F. & Mitchell, W. C. (1946). Measuring Business Cycles. NBER Book Series, *Studies in Business Cycles* N° 2.
https://www.nber.org/books/burn46-1

Carrière-Swallow, Y. & Labbé, J. (2010). Nowcasting with Google Trends in an Emerging Market. *Bank of Chile Working Paper* N° 588. Reprinted (2013) in: *Journal of Forecasting*, 32(4), 289–298.
https://doi.org/10.1002/for.1252

Chamberlin, G. (2010). Googling the present. *Economic and Labour Market Review*, 4(12), 59–95.
https://doi.org/10.1057/elmr.2010.166

Choi, H. & Varian, H. (2009a). Predicting the present with Google Trends. Google, *Technical report*, April 2009.
http://dx.doi.org/10.2139/ssrn.1659302

Choi, H. (2009b). Predicting Initial Claims for Unemployment Benefits. Google, *Technical report*, July 2009.
https://ssrn.com/abstract=1659307

Choi, H. & Varian, H. (2012). Predicting the Present with Google Trends. *Economic Record*, 88, 2–9.
http://dx.doi.org/10.1111/j.1475-4932.2012.00809.x

Cousin, G. & Hillaireau, F. (2018). En attente du titre. *Economie et Statistique / Economics and Statistics* (this issue)

Da, Z., Engelberg, J. & Gao, P. (2010). In Search of Earnings Predictability. University of Notre Dame and University of North Carolina at Chapel Hill, *Working Paper*.
https://pdfs.semanticscholar.org/b68e/aeeac8e5fcd42cff698c7c96dee5e357623a.pdf

Da, Z., Engelberg, J. & Ga, P. (2011). In Search of Attention. *Journal of Economic Finance*, 66(5), 1461–1499.
https://econpapers.repec.org/RePEc:bla:jfinan:v:66:y:2011:i:5:p:1461-1499

D'Amuri, F. (2009). Predicting unemployment in short samples with internet job search query data. *MPRA Paper* N° 18403.
https://econpapers.repec.org/RePEc:pra:mprapa:18403

D'Amuri, F. & Marcucci, J. (2009). "Google It!" Forecasting the US Unemployment Rate with a Google Job Search Index. *ISER Working Paper Series* N° 2009-32.
https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2009-32

Della Penna, N. & Huang, H. (2009). Constructing Consumer Sentiment Index for U.S. Using Internet Search Patterns. University of Alberta, *Working Paper* N° 2009-26.
https://ideas.repec.org/p/ris/albaec/2009_026.html

Diebold, F. X. (2000). "Big Data" Dynamic Factor Models for Macroeconomic Measurement and

Forecasting: A Discussion of the Papers by Lucrezia Reichlin and by Mark W. Watson. In: Dewatripont, M., Hansen, L. P. & Turnovsky, S. (Eds.), *Advances in Economics and Econometrics*, Eighth World Congress of the Econometric Society, pp. 115–122. Cambridge: Cambridge University Press.
https://www.sas.upenn.edu/~fdiebold/papers/paper40/temp-wc.PDF

Dimpfl, T. & Jank, S. (2012). *Can internet search queries help to predict stock market volatility?* New York: Social Science Research Network.

EBRD (2011). Trade Finance on the Way to Recovery in the EBRD Region. EBRD blog, January 2011.

EBRD (2012). Rising uncertainty for trade finance as IFI additionality increases. EBRD blog February 2012.

Ettredge, M., Gerdes J. & Karuga, G. (2005). Using web-based search data to predict macroeconomic statistics. *Communications of the Association of Computing Machinery*, 48(11), 87–92.
https://doi.org/10.1145/1096000.1096010

Galbraith, J. W. & Tkacz, G. (2015). Nowcasting GDP with electronic payments data. ECB *Statistics Paper Series* N° 10.
https://econpapers.repec.org/RePEc:ecb:ecbsps:201510

Giannone, D., Reichlin, L. & Small, D. (2008). Nowcasting: The realtime informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4), 665–676.
https://econpapers.repec.org/RePEc:eee:moneco:v:55:y:2008:i:4:p:665-676

Gilbert, E. & Karahalios, K. (2010). Widespread Worry and the Stock Market. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*.
https://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/view/1513

Gill, T., Perera, D. & Sunner, D. (2011). Electronic Indicators of Economic Activity. *Reserve Bank of Australia Bulletin,* June 2012.
https://www.rba.gov.au/publications/bulletin/2012/jun/1.html

Gruhl, D., Guha, R., Kumar, R., Novak, J. & Tomkins, A. (2005). The predictive power of online chatter. *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, 2005.
https://doi.org/10.1145/1081870.1081883

Guzmán, G. C. (2011). Internet Search Behaviour as an Economic Forecasting Tool: The Case of Inflation Expectations. *The Journal of Economic and Social Measurement,* 36(3), 119–167.
https://ssrn.com/abstract=2004598

Hassani, H. & Silva, E. (2015). Forecasting with Big Data: A Review. *Annals of Data Science*, 2(1), 5–19.
https://doi.org/10.1007/s40745-015-0029-9

Hellerstein, R. & Middeldorp, M. (2012). Forecasting with Internet Search Data. Federal Reserve Bank of New York, *Liberty Street Economics,* January 4, 2012.
https://libertystreeteconomics.newyorkfed.org/2012/01/forecasting-with-internet-search-data.html

ICC (2010). *Rethinking Trade Finance 2010: An ICC Global Survey*. Paris: International Chamber of Commerce.
https://iccwbo.org/publication/icc-global-report-on-trade-finance-2012/

International Institute of Forecasters' Workshop (2014). *Using Big Data for Forecasting and Statistics*. Summary of proceedings of the 11th IIF workshop, April 2014, hosted by the ECB.
https://forecasters.org/wp-content/uploads/11th-IIF-Workshop_BigData.pdf

Jansen, B. J., Ciamacca, C. C. & Spink, A. (2008). An analysis of travel information searching on the web. *Information Technology & Tourism*, 10(2), 101–108.
https://doi.org/10.3727/109830508784913121

Kholodilin, K. A., Podstawski, M. & Siliverstovs, B. (2010). Do Google Searches Help in Nowcasting Private Consumption? Real-Time Evidence for the US. DIW Berlin *Discussion Paper* N° 997.
https://dx.doi.org/10.2139/ssrn.1615453

Koop, G. & Onorante, L. (2013). *Macroeconomic Nowcasting Using Google Probabilities*.
https://www.ecb.europa.eu/events/pdf/conferences/140407/OnoranteKoop_MacroeconomicNowcastingUsingGoogleProbabilities.pdf

Lachanski, M. & Pav, S. (2017). Shy of the Character Limit: "Twitter Mood Predicts the Stock Market" Revisited". *Econ Journal Watch*, 14(3), 302–345.
https://ideas.repec.org/a/ejw/journl/v14y2017i3p302-345.html

Lewis, C. & Pain, N. (2015). Lessons from OECD forecasts during and after the financial crisis. *OECD Journal: Economic Studies*, 5(1), 9–39.
https://doi.org/10.1787/19952856

Liu, Y., Huang, X., An, A., & Yu, X. (2007). *ARSA: a sentiment-aware model for predicting sales performance using blogs*. New York: ACM.
http://doi.org/10.1145/1277741.1277845

Mao Y., Wei, W., Wang, B. & Liu, B. (2012). Correlating S&P 500 stocks with Twitter data. *Proceedings of the 1st ACM Intl. Workshop on Hot Topics on Interdisciplinary Social Networks Research*, 69–72.
http://doi.org/10.1145/2392622.2392634

Mao, H., Counts, S & Bollen, J. (2014). Quantifying the effects of online bullishness on international financial markets. European Central Bank, *Statistics Papers Series* N° 9.
https://www.ecb.europa.eu/pub/pdf/scpsps/ecbsp9.en.pdf?177000b829d4450b007f3d3a612cab18

McLaren, N. & Shanbhogue, R. (2011). Using internet search data as economic Indicators. *Bank of England Quarterly Bulletin,* 51(2), 134–140.
https://econpapers.repec.org/RePEc:boe:qbullt:0052

Mishne, G. & Glance, N. (2005). Predicting Movie Sales from Blogger Sentiment. *Proceedings of the AAAI-CAAW-06, the Spring Symposia on Computational Approaches to Analyzing Weblogs*.
https://www.microsoft.com/en-us/research/publication/predicting-movie-sales-from-blogger-sentiment/

Mourougane, A. (2006). Forecasting Monthly GDP for Canada. OECD Economics Department, *Working Papers* N° 515.
https://doi.org/10.1787/421416670553

O'Connor, B., Balasubramanyan, R., Routledge, B. & Smith, N. (2010). From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. *Proceeding of the International AAAI Conference on Weblogs and Social Media*.
https://www.researchgate.net/publication/221297841_From_Tweets_to_Polls_Linking_Text_Sentiment_to_Public_Opinion_Time_Series

Pain, N., Lewis, C., Dang, T., Jin, Y. & Richardson, P. (2014). OECD Forecasts During and After the Financial Crisis A Post Mortem. OECD Economics Department, *Working Papers* N° 1107.
https://doi.org/10.1787/5jz73l1qw1s1-en

Preis, T., Reith, D. & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: insights from search engine query data. *Philosophical Transactions of the Royal Society* 368(1933), 5707–5719.
https://doi.org/10.1098/rsta.2010.0284

Ranco G., Aleksovski, D., Caldarell,i G., Grcar, M. & Mozeti, I. (2015). The Effects of Twitter Sentiment on Stock Price Returns, *PLoS ONE*, 10(9), 1–21.
https://doi.org/10.1371/journal.pone.0138441

Ritterman, J., Osborne, M. & Klein, E. (2009). Using prediction markets and Twitter to predict a swine flu pandemic. *Proceedings of the 1st International Workshop on Mining Social Media*, pp. 9–17.
https://www.research.ed.ac.uk/portal/en/publications/using-prediction-markets-and-twitter-to-predict-a-swine-flu-pandemic(dcc11feb-77be-44c1-b07a-47da57aba7b8).html

Schmidt, T. & Vosen, S. (2010). Forecasting Private Consumption: Survey-based Indicators vs. Google Trends. *Ruhr Economic Papers* N°155. Also in: *Journal of Forecasting* (2011), 30(6), 565–578.
https://dx.doi.org/10.2139/ssrn.1514369

Schmidt, T. & Vosen, S. (2012). Using Internet Data to Account for Special Events in Economic Forecasting. *Ruhr Economic Papers* N° 382.

Sédillot, F. & Pain, N. (2003). Indicator Models of Real GDP Growth in Selected OECD Countries. OECD Economics Department *Working Papers* N° 364.
http://dx.doi.org/10.1787/275257320252

Suhoy, T. (2009). Query Indices and a 2008 Downturn: Israeli Data. Bank of Israel *Discussion Paper* N° 2009/06.
https://www.boi.org.il/deptdata/mehkar/papers/dp0906e.pdf

SWIFT (2012). The SWIFT index: Technical Description. Society for Worldwide Interbank Financial Telecommunication Inc.

Tkacz, G. (2013). Predicting Recessions in Real-Time: Mining Google Trends and Electronic Payments Data for Clues. *C.D. HOWE Institute commentary* N° 387.
https://ssrn.com/abstract=2321794

Toth, J, & Hajdu, M. (2012). Google as a tool for nowcasting household consumption: estimations on Hungarian data. Institute for Economic and Enterprise Research. Central European University *Research Working Paper*.
https://gvi.hu/files/researches/47/google_2012_paper_120522.pdf

Tuhkuri, J. (2015). *Big Data: Do Google Searches Predict Unemployment?* Masters thesis, University of Helsinki, 2015.
http://urn.fi/URN:NBN:fi:hulib-201703273213

Tumasjan, A., Sprenger, T., Sandner, P. & Welpe, I. (2010). Predicting Elections with Twitter: What 140 Characters Reveal About Political Sentiment. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*?, pp. 178–185.
https://www.researchgate.net/publication/215776042_Predicting_Elections_with_Twitter_What_140_Characters_Reveal_about_Political_Sentiment

Vlastakis, N., & Markellos, R. N. (2012). Information Demand and Stock Market Volatility, *Journal of Banking and Finance,* 36(6), 1808–1821.
https://econpapers.repec.org/RePEc:eee:jbfina:v:36:y:2012:i:6:p:1808-1821

Varian, H. (2014). Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives*, 28(2), 3–28.
https://econpapers.repec.org/RePEc:aea:jecper:v:28:y:2014:i:2:p:3-28

Wakamiya, S., Lee, R. & Sumiya, K. (2011). Crowd-Powered TV Viewing Rates: Measuring Relevancy between Tweets and TV Programs. In: Xu, J., Yu, G., Zhou, S. & Unland, R. (Eds.), *Database Systems for Adanced Applications. DASFAA 2011. Lecture Notes in Computer Science,* vol. 6637, pp. 390–401. Berlin, Heidelberg: Springer.
https://doi.org/10.1007/978-3-642-20244-5_37

Webb, G. K. (2009). Internet Search Statistics as a Source of Business Intelligence: Searches on Foreclosure as an Estimate of Actual Home Foreclosures. *Issues in Information Systems*, X(2), 82–87.
https://scholarworks.sjsu.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1014&context=mis_pub

Wolfram, M. S. A. (2010). *Modelling the stock market using Twitter*. M.S. thesis, School of Informatics, University of Edinburgh, 2010.
http://homepages.inf.ed.ac.uk/miles/msc-projects/wolfram.pdf

Wu, L. & Brynjolfsson, E. (2009 and 2013). The future of Prediction: How Google Searches Foreshadow Housing Prices and Sales. *SSRN papers*.
https://dx.doi.org/10.2139/ssrn.2022293

Ye, M. & Li, G. (2017). Internet big data and capital markets: a literature review. *Financial Innovation,* 3(6).
https://doi.org/10.1186/s40854-017-0056-y

Zhang, X., Fuehres, H. & Gloor, P. (2011). Predicting Stock Market Indicators Through Twitter "I hope it is not as bad as I fear. *Procedia – Social and Behavioral Sciences*, 26, 55–62.
https://doi.org/10.1016/j.sbspro.2011.10.562

APPENDIX

SUMMARY OF RECENT STUDIES USING INTERNET SEARCH AND SOCIAL MEDIA-RELATED INDICATORS FOR MACROECONOMIC FORECASTING AND "NOWCASTING"

Authors	Sector/Topic/Country	Methods and data	Key results	Notes/comments
Andrade *et al.* (2009 and forthcoming)	Analysis of the role of analysts and information dissemination in the run up to the 2007 Chinese stock market bubble.	Correlates different measures of bubble intensity against analyst coverage as measure of information dissemination. Uses a Google search index as check on their timing and intensity.	Significant negative relation found between bubble intensity and analyst coverage. Notes strong positive correlation between the Google search index and volume of new accounts.	This study is mostly tangential to forecasting issues.
Antenucci *et al.* (2014)	University of Michigan study of Twitter based labour market indicators for the period July 2011 to early November 2013.	Estimates indices of job loss, job search and job postings using large sample Twitter as a means of analyzing high frequency weekly estimates. Combines individual measures into composite measures using their principal components to track initial claims for unemployment insurance at medium and high frequencies.	Indicator is found to account for 15 to 20 percent of the variance of the prediction error of the consensus forecast for initial claims. The index also considered useful in providing realtime indicators of events such as Hurricane Sandy and the 2013 government shutdown.	This work is currently under revision since the original model began to deviate in its estimates around mid-2014.
Anvik & Gjelstad (2010)	Forecasting monthly changes in Norwegian unemployment.	Uses Google search indicators related to job search and welfare criteria in simple ARIMA forecasting models of monthly unemployment.	Significant improvements in RMSE found by adding Google search indicators in basic models, and superior to other leading indicators.	Limited to non-economic ARIMA models. Good discussion of the practical limitations of internet search data.
Artola & Galen (2012)	Bank of Spain study of British tourism inflows to Spain.	Uses Google search indicators related to UK search for Spanish holiday destinations in simple ARIMA model of British tourist inflows.	Google search indicator is found to be significant, with improvements in predictive value sensitive to choice of baseline model.	Notes limitations to Google search indicators and sensitivity to choice of language and search criteria.
Askitas & Zimmermann (2009)	Forecasting monthly changes in German unemployment.	Uses Google search indicators in univariate error correction models.	Strong correlations found with models predicting trends and turning points.	Notes limitations in existing data sets and scope for wider use.
Bortoli & Combes (2015)	French Insee study of the use of internet search indicators in predicting consumer's expenditures at the aggregate and detailed disaggregate levels.	Introduces Google search indicators for a wide range of aggregate and disaggregate consumption items into a multivariate indicator model framework.	Finds that search indicators do not improve the forecasting of monthly aggregate household consumption. Results for certain goods (clothing, household durables, and food) and some services (transport) are more positive but generally mixed.	Includes excellent review of the strengths and limitations of internet search variables and their uses. Notes in particular concerns about the continuity and structural stability of internet search based measures.
Bollen *et al.* (2011)	Uses OpinionFinder and Google POMS over the period March to December 19, 2008 to identify 6 Twitter-based measures of mood states. Examines the relationship between mood indicators and changes in the Dow Jones index from March to December 2008.	Correlates mood indicators against the Dow Jones index within a general autoregressive model and Granger causality testing framework on a daily basis.	The broad results suggest that the prediction accuracy of the standard stock market prediction models is significantly improved when some but not all mood dimensions are included.	Notes that variations in measures of Calm and Happiness as measured by GPOMS appear to have some predictive value, but not for general happiness as measured by the OpinionFinder tool.
Carrière-Swallow & Labbé (2010)	Bank of Chile study of the sales of automotive products.	Adds a Google search indicator for the most popular car brands in Chile to simple and high order autoregressive models for year on year car sales in combination with a general indicator of economic activity.	Models including Google search indicator found to significantly out-perform both simple and more complex baseline models within and outside the sample period.	
Chamberlin (2010)	UK NSO study modelling a range of monthly UK statistics, including retail sales, home purchases, car registrations and foreign travel.	Adds Google search indicators to simple monthly first difference autoregressive models.	Results are mixed: significant for detailed expenditures and mortgage approvals but poor for total retail sales, car purchase and travel.	No out-of-sample tests done. Notes sensitivity to search query choice and seasonality.

Authors	Sector/Topic/Country	Methods and data	Key results	Notes/comments
Choi & Varian (2009a)	Google Research team study, modelling a range of different monthly US demand variables.	Adds Google search indicators to simple AR models.	Models including Google search indicators found to generally outperform baseline models, results are mixed with little or no gains for motor vehicles and housing sector.	Innovative and original study. Notes that sampling method may add noise but anticipates improvements over time.
Choi & Varian (2009b)	Forecasting US monthly unemployment benefits claims.	Adds Google search indicators to simple autoregressive models of unemployment claims.	Significant improvements in forecasting accuracy found relative to baseline model.	Results in line with other countries. Models are strictly non-economic.
Choi & Varian (2012)	Consolidates earlier studies and extends methods to include Hong Kong tourism and Australian consumer sentiment.	Adds Google search indicators to simple autoregressive models and tests for out of sample forecasting accuracy.	Significant improvements in forecasting accuracy found.	
Da *et al.* (2010)	Study of US cross-company performance and revenue surprises.	Uses Google search indicators for individual firm products to predict revenue surprises within a time series panel..	Finds significant relationship between search volumes and earnings surprises and company performance.	
Da *et al.* (2011)	Study of a large sample of US company stock performance.	Uses search frequency indicators as measure of investor attention for 3000 US companies.	Finds strong correlations though different from existing proxies for company attention.	
D'Amuri (2009)	Analysis of quarterly Italian unemployment.	Adds Google search indicators of job search enquiries to quarterly multivariate ARIMA models including industrial production and employment expectations variables.	Google search indicators found to be significant and superior to established leading indicators. Small sample results found to be better than for longer samples.	
D'Amuri & Marcuccio (2009)	Analysis of monthly US unemployment in aggregate and at state levels.	Tests the addition of Google search indicator in ARIMA models over a wide range of model forms and specifications, including other relevant US leading indicators.	Combined with Initial Claims indicators, models including Google search indicators found to outperform other models across a wide range of specifications at aggregate and most state levels.	Models including Google search indicators found to be superior to those using the Survey of Professional Forecasters.
Dimpfl & Jank (2012)	Study of daily US stock market volatility.	Uses Google search indicator by company name as measure of investor attention. Tests relationship with stock market prices and volatility within an ARIMA model framework.	Finds strong co-movements between Google searches and market movements and volatility, with search queries providing more precise in and sample prediction.	
Ettredge *et al.* (2005)	Earliest study of US monthly unemployment (2001-2004) using a range of internet search data predating Google Trends.	Constructs and correlates an Internet search based measure of job search within a simple forecasting model.	Finds significant correlation between job-search and unemployment data with significant trade-off between explanatory power and lead time. Index found to be superior to weekly initial claims data. Relationship is only significant for males.	Author strongly promotes future use of internet search statistics as a means of predicting a wider range of macroeconomic data, and proposes related study of consumer confidence.
Galbraith & Tkacz (2015)	Bank of Canada study combining a range of financial and transactions indicators within a set of mixed frequency GDP indicator models.	Models combine measures of the growth in values and volumes of monthly and quarterly Canadian debit, daily credit with composite leading indicators for the US and Canada, monthly unemployment rates and lagged GDP growth.	A key finding is the improvement in accuracy for the earliest GDP nowcasts through the inclusion of debit card payments observed for the first two months of the nowcast period, although such improvements are not detectable once the previous quarter's GDP value is observed (in month 3).	Provides overall support for the need for combining electronic transactions with other data, measured with some accuracy at a daily frequency.
Gilbert & Karahalios (2010)	Twitter based study constructing a broad Anxiety Index based on Live-Journal blog entries. Tests through 2008 data for possible influence of the index on daily changes in the Standard and Poor index (the S&P500).	Estimates a baseline statistical relationship between S&P, its lagged values, and levels and changes in the volume of transactions taking place and the VIX Fear index	Uses a combination of regression and Granger causality tests and finds a statistically significant relationship between the Anxiety Index and future stock market prices.	Notes that result is weakened by inclusion of the VIX index which tends to dominate. Notes difficulties in interpreting blog-based linguistic expressions, index volatility due to external factors and the exceptional nature of 2008.

Authors	Sector/Topic/Country	Methods and data	Key results	Notes/comments
Gill *et al.* (2011)	Reserve Bank of Australia review of the use of various electronic indicators as a means of improving information and forecasts of main Australian macro aggregates.	Use a range indicators of retail and wholesale bank transfer (SWIFT) and card transactions in AR(1) and principal components models for retail sales, consumption, domestic demand and GDP.	Results are mixed. SWIFT payments indicators are found significant in some AR models, but best in principal components models in combination with other measures. Results using retail payments indicators are less significant.	The authors suggest wider use of electronic indicators to improve the real-time measurement of economic aggregates. Suggest that such data are likely to become more useful as payments behaviour and internet use become more stable over time.
Guzmán (2011)	Study of Google search indicators as measure of real-time US CPI inflation expectations.	Tests forecasting performance for search indicators relative to 36 other indicators of inflation expectations and TIPS spreads.	Results suggest higher frequency measures outperform lower frequency measures in use, in terms of accuracy, predictive power. Out-of-sample forecasts using the Google search indicator have lowest forecast errors across the range of indicators used.	
Hellerstein & Middeldorp (2012)	New York Fed blog review of current literature on use of internet search counts in a range of modelling areas includes new work on US financial markets.	Adds Google search indicator for home and mortgage refinancing to a small dynamic model of the refinancing index, also including the influence of market yields. Adds Google search indicator to models of the Renminbi-dollar forward market variables.	Results are mixed. Google search indicator significantly improves forecast performance for mortgage refinancing, but gains are limited by insignificance of lead times. Search indicator found significant in Renminbi forward market analysis although predictive power is low.	Concludes that improvements in predictive power are not universal and do not provide explanatory power beyond more traditional methods, but nonetheless a useful addition to the economist's toolkit.
Kholodilin *et al.* (2010)	Examines usefulness of Google search indicators in nowcasting year-on-year growth in monthly US private consumption (2007-2010).	The Google search indicator-based forecasts are compared to benchmark AR(1) model and others including the consumer surveys and financial indicators.	Google search based forecasts found more accurate than for benchmark model. Similar results found for models including consumer survey and financial variables.	
Koop & Onorante (2013)	Examines the use of Google search probability variables in monthly US dynamic switching models for nine US macroeconomic variables (inflation, industrial production, unemployment, oil prices, money supply and other financial indicators).	Introduces Google search-based probability measures into a dynamic model switching (DMS) nowcasting system in which current outcomes are regressed on lagged values of the set of dependent variables and Google indicators.	Inclusion of internet search data gives improvements in many cases, but best included as model switching probabilities rather than simple regressors. General results are mixed; positive for inflation, wage, price and financial variables, less so for industrial production and inferior for unemployment.	Innovative approach combining search information with a sophisticated DMS nowcasting system.
Lachanski & Pav (2017)	Attempt to replicate Bollen *et al.* (2011) using similar Twitter based data sets methods.	Correlates mood indicators against the Dow Jones index within a general autoregressive model and Granger causality testing framework on a daily basis.	Finds some in-sample but almost no out-of-sample evidence that such a measure contains information relevant to the Dow Jones index.	Concludes that Bollen *et al.* results are an outlier and that there is little/no credible evidence that the collective mood content of raw Twitter text data from the universe of tweets can be used to forecast index activity at the daily time scale.
Mao *et al.* (2012)	Examines the relationship between Tweets mentioning the S&P 500 index and stock prices and traded volume between February and May 2012. Analysis done at the aggregate level, for each of 10 industry sectors and at the company level, for Apple Inc.	Uses simple linear autoregressive regression models, to predict the stock market indicators with the Twitter data an exogenous input.	Generally mixed results. Significant correlations at aggregate level with levels and changes in prices but not trading volumes. Significant correlations for 8 out of 10 sectors with traded volumes but not prices. Significant correlations for both volumes and prices for the financial sector and Apple Inc. Results are broadly mirrored in the tests for predictive accuracy.	Predictions of directional changes in the sample period are at best 68% accurate for the aggregate and financial sectors and only 52% for Apple Inc., close to a random walk.

Authors	Sector/Topic/Country	Methods and data	Key results	Notes/comments
Mao *et al.* (2014)	Analysis of Twitter and Google search-based indicators of "bullishness" or "bearishness" calculated on daily basis (Twitter) over the period 2010 to 2012, and weekly (for Google Trends) over the period 2007 to 2012.	Makes cross comparisons and with other investor sentiment indicators and analyses relative predictive powers in small dynamic models of for US, UK Canadian and Chinese stock market prices and returns. US model is notably more complete. Twitter-based measures found to lead changes in Google-based measures, both are positively correlated with other measures of US investor sentiment.	Twitter-based indicator statistically significant and provides better predictions of stock returns for the US. Google-based indicator also significant but with lower predictive power. Similar correlations for the UK, Canada and China, within simpler bi-variate model, but with lower predictive power for China. Google indicators are significantly correlated for stock market prices but with lower predictive power.	Notes lack of evidence with regard to causality. Notes need to develop appropriate experimental design methods and machine learning algorithms for processing Tweets and for testing causality.
McLaren & Shanbhogue (2011)	Bank of England paper examining use of Internet search data for UK labour and housing markets.	Adds Job Seekers search variable to first-differenced AR models of unemployment including other indicators and house prices, over the period 2004-2011.	Mixed results. Job Seekers indicator significant but outperformed out of sample by Claimant Counts. Stronger results for house prices, Internet search variable in AR(1) model outperforming ones based on other indicators over the period 2004-2011.	Notes limitations in the approach but concludes that search data provide additional insights not covered by business surveys. Bank to monitor search data within range of indicators in reviewing UK economic prospects.
Preis *et al.* (2010)	Examines weekly Google search data looking for possible links between search volume data and weekly US financial market fluctuations.	Complex correlation analysis of company name search and transactions volumes for S&P 500.	Finds evidence of strong correlations. Recurring patterns found using new method for quantifying complex correlations.	
Schmidt & Vosen (2010)	Examines predictive performance of Google search indicator for US private consumption.	Performance assessed relative to Michigan Consumer Sentiment and Conference Board Confidence Index in simple AR models and more conventional consumption functions including lagged income, interest rate and stock market price variables.	Google search indicator outperforms survey based indicators in simple AR models. With an extended consumption function, both Google and Conference Board indicators offer improvements, with the former useful for one month ahead predictions.	Michigan index found to have no additional value.
Schmidt & Vosen (2012)	Examines use of internet search data to predict special events when timely information is not available. Specifically it looks at car scrapping programs in four countries (France, Germany Italy and the United States).	Uses small quarterly dynamic models of changes in consumption over the period 2002 to 2009, including income and a Google search indicator, effectively entering as a shift variable during the relevant programs.	Finds the inclusion of search query data into statistical forecasting models improves the forecasting performance in almost all cases.	Notes that major challenge is to identify irregular events and finding the appropriate time series from Google search statistics.
Suhoy (2009)	Examines use of Google search indicators across a range of sectors and variables for Israel, using query categories including human resources, home appliances, travel, real estate, food and drink and beauty and personal care.	Applies Granger causality tests, first differenced linear and two-state Bayesian models to test for co-movement in indicators and growth cycles.	Labour market indicator found most predictive, improving monthly projection of changes in unemployment rates. Finds weekly frequency useful in real-time monthly monitoring, with query indices preceding official data by up to two months. Co-movements in search queries found useful in assessing economic slowdowns.	
Tkacz (2013)	Canadian study examining the use of Google search indicator for predicting recent turning points and recessions in key macroeconomic indicators.	Examines internet recession-related search indicators alongside other financial and payments variables within probit models to predict turning points in GDP and unemployment.	Finds that Google searches for "recession" and "jobs" would have predicted the 2008 recession up to three months in advance of its onset. Shortness of sample prevents analysis of other turning points.	Provides good review of the nature and limitations of search related variables, noting both advantages in their timeliness but also their qualitative nature and sensitivity to specific choices.

Authors	Sector/Topic/Country	Methods and data	Key results	Notes/comments
Toth & Hajdu (2012)	Examines use of Google search indicators to predict household consumption, retail sales and car sales in Hungary.	Constructs and tests search indicators for retail sales and car sales in simple autoregressive baseline model using monthly data for the period 2004-2011.	A combination of Google variables are found to be significant when used in combination with autoregressive terms. Similar results are obtained for quarterly consumption, though with smaller sample size.	Highly seasonal data set.
Tuhkuri (2015)	PhD thesis study of the use of internet search data in predicting US unemployment at economy wide and state levels.	Introduces internet search frequency information on unemployment benefits into a range of AR benchmark and panel data models at the state level.	Improvements in predictive accuracy using Google data appear robust to different model specifications and search terms, but are generally modest and limited to short-term predictions. The informational value of internet search data also tends to be time specific.	Provides an excellent review of the literature and thorough insights into a variety of tests including causality and stability.
Vlastakis & Markellos (2012)	Study of information demand at market and firm level using data for the largest 30 stocks traded on the NYSE.	Proxies demand by weekly internet Google Trends search volume data by company name.	Results suggest significant relationship with individual stock trading volumes and the conditional variance of excess stock returns. Significance of search indicators diminishes using implied rather than historical measures of volatility at the firm and market levels.	Study confirms theoretical proposition that information demand is positively related to risk aversion.
Webb (2009)	Examines the relationship between Google searches on the keyword “foreclosure” and actual U.S. home foreclosures over the period 2004-2008.	Uses bi-variate correlation and regression analysis.	Finds a high correlation between the two variables providing a reasonably accurate estimate of trends in actual U.S. home foreclosures.	
Wolfram (2010)	Applies Natural Programming Language sampling methods to very high frequency Twitter feeds over a 10 day period in 2010, to predict hourly and daily movements in the individual stock prices for Apple, Google, Intel and other selected stock prices.	Uses automated Support Vector Regression (SVR) methods to model and simulate stock price movements over the very near-term.	Model based predictions were found to be close to baseline for Apple and Google stocks over a very short (15 minute) period, but become unstable as the forecast distance increases (to 30 minutes). Concludes that that relevant information can be extracted to give small but significant advantages in predicting market prices.	Notes the need to improve sampling by more clearly identifying influential users and creating rules specific to the Twitter dataset for focussing more specifically on the topic of financial markets.
Wu & Brynjolfsson (2009 and 2013)	Seminal paper examining the use of Internet search data to predict US housing market trends and sales of house appliances in 2008-2009.	Relevant search indicators are constructed and then introduced into quarterly dynamic joint autoregressive models for house purchases and prices at the state level, including fixed effects variables.	Housing search index found to be significant and strongly predictive of both future housing market sales and prices compared with an underlying baseline model. Out-of-sample predictions and mean absolute errors significantly smaller than baseline model. Similar results found for home appliance sales.	
Zhang *et al.* (2011)	Examine large sample of daily Twitter entries between March and September 2009. Estimates a variety of measures of differing degrees of positive and negative moods, ranging from fear to hope.	Correlates these against corresponding values of the Dow Jones, NASDAQ and S&P500 indices, as well as the VIX index.	Finds statistically significant correlations, consistent with negative impacts of lagged mood indicators on current stock market prices and the VIX.	Notes that the result holds for positive and negative mood indicators, suggesting the relative importance of emotional outbursts as opposed to the specific mood indicator during the sample period.