Nowcasting and the role of Big Data in short-term macroeconomic forecasting: a critical review

Pete Richardson
UN Headquarters, New York, October 2015
Presentation outline

Background and Motivation
The relevance of Big Data to Economic Forecasting?
Background to Internet Search based studies
Google Trends search data tool
Quick demo of how to use Google Trends
A review of Internet Search-based studies
The limitations of Google Trends data
Other Big Data and related studies
Some broad conclusions
A key limitation of indicator and now-cast models is the lag in availability of hard statistical information. Typically goodness-of-fit and out-of-sample predictive performance improve significantly the more information is available for monthly hard indicators during the quarter in question. Might the availability of more timely data from unconventional sources assist short-term assessment?
Big Data - a broad term for data sets so large or complex that traditional data processing applications are (or were) inadequate.

Typically the availability and use of Big Data has been made feasible by the near exponential growth of data storage and processing capacities.

Challenges include analysis, capture, data curation, search, sharing, storage, transfer, visualization, and information privacy.

A number of recent, mostly post-crisis, studies have focussed on the possible usefulness to forecasting of two such new sets of information:

- Internet search statistics and
- Detailed micro-level transactions data available from economic and financial systems

both of which are, in principle, available on a near real-time basis.
Internet Search based studies

Rationale:

- Internet search has become a widespread means for economic agents to obtain information relevant to their immediate economic activities and decisions which get reflected in their behaviour and the wider set of official economic statistics.

- Such data embodies relevant additional information which is available quickly, at high frequency and possibly with a significant lead time on transactions being recorded.

A growing body of studies has evolved on the use Internet Search statistics in models, following Ettredge et al (2005), Choi and Varian (2009a and 2009b) and Wu and Brynjolsson (2009).

Typically involve the construction of time series indicators of “frequency” of Internet searches for specific keywords/phrases relevant to a specific category of economic activity by country.

For example, searches for terms such as “welfare and unemployment benefits” or “mortgage foreclosure” or “car loans”, car scrapping schemes etc. for specific countries.

The relevant indicator is then typically added to and tested for significance within a baseline forecasting model on within- and out-of-sample bases.
Google Trends Search Tool

Earliest studies used fairly raw internet search statistics from diverse search engines.

In 2009, Choi and Varian of Google Labs launched fairly refined facilities within the Google Trends/Google Insights for Search website.

Google Trends enables researchers to recover tailor-made sample statistics on the frequency of searches for specific keywords by location and on a near real-time basis, starting from around 2004. Essentially mining Google’s **Big data** search archives.

Sample sizes pose a specific limitation on their general usefulness for modelling, as is the sampling method which is inevitably variable over time.

A wide range of studies: originally focussing on unemployment indicators, but then widened to include housing, tourism, retail sales and consumption, housing markets, inflation expectations and stock and financial markets, and for a variety of countries.
- Go to Google trends: https://www.google.com/trends/
- Log into a Google account (essential if you intend to download data)
• Pick country/region e.g. UK
- Pick keyword/topic e.g. UK unemployment
Pick country option (was Worldwide choose UK instead)
Google Trends

- UK Search stats for UK unemployment
Google Trends

- Also gives related search topics and regional stats (over time)
- Select download option to get csv data file
- Accessing and saving report file
### Data file gives weekly data points for the search index

<table>
<thead>
<tr>
<th>Date</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>03-02 - 2007-09-08</td>
<td>36</td>
</tr>
<tr>
<td>03-09 - 2007-09-15</td>
<td>29</td>
</tr>
<tr>
<td>03-16 - 2007-09-22</td>
<td>29</td>
</tr>
<tr>
<td>03-23 - 2007-09-29</td>
<td>29</td>
</tr>
<tr>
<td>03-30 - 2007-10-06</td>
<td>27</td>
</tr>
<tr>
<td>10-07 - 2007-10-13</td>
<td>27</td>
</tr>
<tr>
<td>10-14 - 2007-10-20</td>
<td>27</td>
</tr>
<tr>
<td>10-21 - 2007-10-27</td>
<td>27</td>
</tr>
<tr>
<td>10-28 - 2007-11-03</td>
<td>27</td>
</tr>
<tr>
<td>11-04 - 2007-11-10</td>
<td>27</td>
</tr>
<tr>
<td>11-11 - 2007-11-17</td>
<td>26</td>
</tr>
<tr>
<td>11-18 - 2007-11-24</td>
<td>26</td>
</tr>
<tr>
<td>11-25 - 2007-12-01</td>
<td>26</td>
</tr>
<tr>
<td>12-02 - 2007-12-08</td>
<td>33</td>
</tr>
<tr>
<td>12-09 - 2007-12-15</td>
<td>28</td>
</tr>
<tr>
<td>12-16 - 2007-12-22</td>
<td>28</td>
</tr>
<tr>
<td>12-23 - 2007-12-29</td>
<td>28</td>
</tr>
<tr>
<td>12-30 - 2008-01-05</td>
<td>28</td>
</tr>
<tr>
<td>01-06 - 2008-01-12</td>
<td>28</td>
</tr>
<tr>
<td>01-10 - 2008-01-19</td>
<td>27</td>
</tr>
<tr>
<td>01-20 - 2008-01-26</td>
<td>32</td>
</tr>
<tr>
<td>01-27 - 2008-02-02</td>
<td>30</td>
</tr>
<tr>
<td>02-03 - 2008-02-09</td>
<td>26</td>
</tr>
<tr>
<td>02-10 - 2008-02-16</td>
<td>26</td>
</tr>
<tr>
<td>02-17 - 2008-02-23</td>
<td>29</td>
</tr>
<tr>
<td>02-24 - 2008-03-01</td>
<td>25</td>
</tr>
<tr>
<td>03-01 - 2008-03-08</td>
<td>37</td>
</tr>
<tr>
<td>03-08 - 2008-03-15</td>
<td>35</td>
</tr>
<tr>
<td>03-15 - 2008-03-22</td>
<td>27</td>
</tr>
</tbody>
</table>
- A wide range of period options available with pull down box (NB they are scaled to Max/Min values for the period chosen).
The early study by Ettredge et al (2005), predates Google Trends, looks at US monthly unemployment 2001-2004, uses Internet search indicator of job-search within a simple autoregressive forecasting model. Finds a significant relationship with published unemployment data for adult males, and superior to the official weekly claims data.

Broadly similar results are reported for monthly total unemployment for:

- Choi and Varian (2009b) for the United States,
- D’Amuri and Marcuccio (2009) for the United States at aggregate and state level
- Suhow (2009) for Israel,
- D’Amuri (2009) for Italy,
- Anvik and Gelstad (2010) for Norway
- McLaren and Shanbhogur (2011) for the United Kingdom

Most of these studies use a similar method of adding the search indicator to fairly naïve time series models in level or first-differenced terms

Of note, D’Amuri and Marcuccio (2009), use more sophisticated models which include other economic variables and leading indicators relevant to unemployment.

Most studies find the relevant search indicator to be 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.
Studies of consumption, retail sales and car sales include:

- Choi and Varian (2009a and 2011) and Kholodilin et al (2010) and Schmidt and Vosen (2011) for the United States
- Chamberlin (2010) for the United Kingdom
- Toth and Hadju (2012) for Hungary.

Both methods used and results obtained vary somewhat across these studies. Some follow a similar strategy to those used for unemployment by adding relevant internet search indicators to relatively naïve time series forecasting models. Others include search indicators in combination with other measures of consumer sentiment or, in the case of Schmidt and Vosen (2011), more fully specified reduced form economic models 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.
The results of Schmidt and Vosen (2011) are particularly telling:

- individual significance of such variables to be greatest in simple AR(1) models.
- In more semi-structural consumption function specifications they are found to perform as well as or well in combination with the Conference Board Indicator, though the best one-month-ahead nowcasts are given by models including the Google Indicator.
- An interesting bi-product of this study is the finding that the Michigan Consumer Sentiment indicator is found to have no additional value.

Schmidt and Vosen (2012) find Google-based indicators generally useful in modelling and predicting the effects of motor vehicle scrapping schemes (so called “cash for clunkers”) for new car sales the United States, France, Germany and Italy in the period 2002-2009.

This suggests a 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 major challenges lie in identifying significant irregular events and constructing an appropriate indicator from Google Insights.

Most recently an INSEE paper by Borfoli and Combes (2015) reviews the usefulness of Google indicators for modelling French consumption (agg and disagg). Results are mixed and they that Google trend searches improve household expenditure forecast in only a limited way.
For the US housing market, Webb (2009) finds high correlations between searches for “foreclosure” and recorded foreclosures, whilst Wu and Brynjolfsson find a Google-based housing search indicator significant/strongly predictive for housing market sales and prices and the sales of home appliances. Hellerstein and Middledorp (2012) find similar improvements for predicting mortgage refinancing, though the gains are found to be insignificant beyond a lead time of one week.

McLaren and Shanbhogur (2011) report relatively strong results for UK house prices, with an internet search indicator outperforming other indicators over the period 2004-2011.

Tourism, Choi and Varian (2011) report significant results for Hong Kong tourism. Artola and Galen (2012) find similar results when adding Google based indicators to ARIMA models of the UK demand for holidays in Spain, though the latter also report considerable sensitivity to the choice of both baseline model and search keywords, particularly when used in other languages.

Inflation expectations indicators for the US: Guzman (2011) finds that higher frequency Google-based indicators to generally outperform lower frequency traditional measures.
Internet Search-based studies: Financial sector

Financial sector studies have not typically been in a forecasting context.

- Andrade et al (2009) use such measures in identifying market volatility in the run up to the 2007 Chinese stock market bubble,

- Vlastakis and 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, Engelberg and Gao (2010 and 2011) find similar correlations between product search variables and revenue surprises and investor attention for 3000 US companies

- Preis, Reith and Stanley find strong correlations between name searches and transactions volumes for S&P 500 companies.

- Dimpf and Jank (2012) report strong co-movements between Google company name searches and US stock market movements and volatility, providing better out of sample forecasts than ARIMA models.

- Hellerstein and Middledorp (2012) find a Google search indicator to be significant in modelling movements in dollar-Renminbi forward market variables, but with low predictive power.

The lack of firm evidence or applications on the forecasting side is perhaps of less importance given the availability of high frequency indicators for financial markets.
Limitations of the Google Trends approach: the data

Results are mixed across topics and subject to specific limitations and biases.

- Search indicators do not represent the **absolute** number of searches but the **proportion** of searches for a specified keyword at any one time, suitably scaled. So there is sometimes a need for data to be “cleaned” for specific outliers or aberrant search terms.

- High frequency Google-based indicators draw on a variable and non-stratified sample, one which changes on a fairly continuous basis over time. Both are likely to add noise and make the indicators more qualitative than at first sight.

- The shortness of sample limits the scope for the testing within a range of existing models. Most studies rely on relatively short sample high frequency data which are subject to strong seasonality, with the risk of swamping the underlying relationships.

- At least visually, this seems to be the case for a number of studies claiming to illustrate close historical relationships between the search indicator and variable in question.

Many authors note the sensitivity of results to the choice of keywords and language

- Implies the need for care in the construction of an indicator targeted for a specific use.

- A lack of standardised published measures specific to general macroeconomic monitoring.
Limitations of the Google Trends approach: the model framework

- Studies reporting high significance or superior out-of-sample forecasts usually do so by comparison with relatively naïve univariate time series models -- AR(1) or low order ARIMA.
- So the results are probably not surprising to the extent that such models are rarely 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 major turning points.
- A relatively small subset 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 the literature claims 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.
Big Data and other indicators: SWIFT (1)

- SWIFT = Society of Worldwide International Financial Telecommunications messaging system


- Both report sharp year-on-year decline in SWIFT trade-related messages from end-2008 to end-2009 and in early 2011 and their relationship with related global and regional trends in trade.

- Looking at different electronic indicators of wholesale and retail payments Gill, Perera and Sunner (ABS 2012) find that a SWIFT payments indicator combined with conventional short-term macro indicators, improves short-term predictive performance for GDP relative to naïve autoregressive baseline models. Other retail payments indicators including credit card transactions do less well.
• SWIFT (2012) with CORE Louvain, use an OECD aggregate index of filtered transactions in a suite of GDP bridge models, showing significant results for quarterly movements in OECD real GDP for the period 2000 to 2011. The underlying baseline model is a relatively simple statistical ARMA model, taking account of no other relevant information.

• SWIFT and CORE now produce quarterly notes on the nowcasting results based on various SWIFT indicators covering the US, the UK, Germany and the EU27.

• More importantly this data can now be downloaded free from the SWIFT site.

• Overall results to date are generally supportive of the broad approach, and with the advantage of being available for a longer sample period, merit further investigation.

• Important caveat: SWIFT indicators relate to the volume of messages not the levels of transactions and need to be filtered for content and coverage, as between trade, financial and other activity-related transactions.
Other Big Data indicators: ADP, ADS and Ceridian Pulse

Automatic Data Processing Inc’s National Employment Report (2012) for the United States takes monthly and bi-weekly payroll data processed by the ADP’s system -- covering approximately 20% of U.S. private sector workers -- 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 used, in conjunction with the Philadelphia Federa ADS Business Conditions Index (see Aruoba, Diebold and Scotti, 2009), to estimate a system of VAR equations to predict monthly changes in BLS private employment data by sector, since April 2001. Significance of individual variables is not reported but 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.

Ceridian-UCLA Anderson Pulse of Commerce Index (PCI) - based on Ceridian electronic card payment services for US transportation industry, diesel sales for freight haulage. 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.

To date no published analytical studies appear to be available using the PCI, 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.
Some broad conclusions

Internet-based search indicators appear to provide a novel and useful means of measuring various aspects of consumer and business behaviour in an almost real-time way, which may not be embodied in other indicators or available on such a timely basis.

They are subject to important limitations in terms of quality, form, sample sizes and their “qualitative” nature.

Available empirical studies provide interesting insights and evidence of predictive performance across a range of topics.

But the results to date are generally mixed and the models used need further refinement to be more generally useful for macroeconomic forecasting.

To date, other Big Data sources seem to be relatively unused by existing studies. They also show some promising features but are equally limited in terms of information content and transparency.

The overall message is that Big Data indicators are useful additions to the forecasters toolkit, but warrant further development and monitoring in parallel with other macroeconomic indicators and forecasting techniques.