



# Forecasting unemployment with internet search data: Does it help to improve predictions when job destruction is skyrocketing?



María Rosalía Vicente <sup>\*</sup>, Ana J. López-Menéndez <sup>\*</sup>, Rigoberto Pérez

*Applied Economics, University of Oviedo, Oviedo, Spain*

## ARTICLE INFO

### Article history:

Received 23 April 2014

Received in revised form 8 December 2014

Accepted 17 December 2014

Available online 6 January 2015

### Keywords:

Google indicators

Forecasting

Unemployment

Spain

ARIMA

ARIMAX

## ABSTRACT

As more and more daily activities take place online, data on internet behaviour is becoming a key information source. In this sense, several papers have explored the usefulness of internet search data in order to improve the nowcasting and forecasting of economic indicators. Special attention has been paid to predicting unemployment. Nonetheless, most of the empirical evidence on this field has focused in countries with low/moderate rates of unemployment.

This paper follows this line of research and explores the use of Google Trends data in order to improve the forecasting of the figures of unemployment in Spain. This country reveals as a very interesting case due to the sharp increases in unemployment caused by the economic crisis.

With the aim of providing an extensive approach on the Spanish unemployment, we use ARIMA models, also introducing as explanatory variables the Google searches for “job offers” and a business confidence indicator referred to employment perspectives. In this way we combine the time series perspective with qualitative indicators from the supply and the demand sides, leading to a more comprehensive description and also improving forecasting performance.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

In the last few years the term “big data” has gained popularity in order to refer to the vast amount of digital data which information and communication technologies are making possible to gather, store, diffuse and share in unprecedented ways. Such phenomenon is leading both companies and governments to think carefully over the ways to exploit this data as efficiently as possible.

Hence, economists have been paying special attention to the data coming from online search behaviour since the internet has become a major source of information. Recent figures from Eurostat (2014b) report that 59% of Europeans use the internet to search for information on goods and services; furthermore, among the unemployed Europeans almost half (48%) use the internet to search for a job or to

submit an application, a figure which raises over 80% in countries such as Finland and the United Kingdom.

In this sense, Ettredge et al. (2005, p. 87) early pointed out that “people reveal useful information about their needs, wants, interests, and concerns via their internet behaviour, and that terms submitted to search engines reflect this information”. In fact, search engines are the main entry for internet users when they look for any kind of information as these engines “provide rapid and easy access to websites and information sources” (Eurostat, 2013).

Hence, several papers have addressed the use of the data coming from online search engines in order to improve the nowcasting and forecasting of economic indicators (Choi and Varian, 2009a, 2009b; D’Amuri, 2009; Della Penna and Huang, 2009; Guzman, 2011; McLaren and Shanbhogue, 2011; Vosen and Schmidt, 2011, 2012a, 2012b; Wu and Brynjolfsson, 2010), given that “official economic statistics are generally published with a lag” (McLaren and Shanbhogue, 2011).

Forecasting unemployment with online search-related data has particularly drawn researchers’ attention (Anvik and Gjelstad, 2010; Askitas and Zimmermann, 2009; D’Amuri, 2009;

<sup>\*</sup> Corresponding authors at: Applied Economics, University of Oviedo, Campus del Cristo s/n, Oviedo, 33006 Asturias, Spain.

E-mail addresses: mrosalia@uniovi.es (M.R. Vicente), anaj@uniovi.es (A.J. López-Menéndez), rigo@uniovi.es (R. Pérez).

D'Amuri and Marcucci, 2009; Fondeur and Karamé, 2013; McLaren and Shanbhogue, 2011; Suhoy, 2009). The available empirical evidence shows the usefulness of online search data for nowcasting and forecasting unemployment figures. Nevertheless, it has focused in countries with low/moderate rates of unemployment – France 10.8%, Germany 5.1%, the United Kingdom (UK) 7.2% (Eurostat, 2014a) and the United States (US) 6.7% (Bureau of Labour Statistics, 2014), among others<sup>1</sup>.

Within this context, the aim of this paper is to explore whether online search data still helps to improve the nowcasting and forecasting of unemployment in the context of high rates of employment destruction over time. In particular, the present analysis focuses on Spain. This country is an interesting case to study since the economic crisis has been badly hitting the labour market, with rates of unemployment that double the European Union's average. In fact, recent figures released by Eurostat (2014a) show an unemployment rate for Spain of 25.8% compared to the rate of 10.7% estimated for the whole European Union. Only Greece exhibits a higher rate: thus, 27.8% of the Greek economically active population is unemployed.

The analysis presented in this paper is based on a time series approach (ARIMA models), also introducing a business sentiments index (the employment confidence indicator) and the intensity of internet search (through Google Trends). The obtained results confirm the usefulness of internet search data as both an economic indicator and a forecasting tool.

## 2. Literature review

Over the last few years, several attempts have been made to explore the potential of internet-based search data in order to forecast economic variables. Ettredge et al. (2005) were the first to address such issue and investigated the link between online job searches and the official rates of unemployment in the United States.

Choi and Varian (2009a, 2009b) put forward this line of research by describing and illustrating how internet search data and, in particular, Google's, could be used in order to improve the predictions of several economic indicators such as unemployment claims, retail sales, real estate demand, and holiday destinations' demand.

These two papers have stimulated a lot of recent research in this field. Special attention has been paid to both nowcasting and forecasting unemployment rates with internet job search-related data. Evidence has been provided for different countries: France (Fondeur and Karamé, 2013), Germany (Askatas and Zimmermann, 2009), Israel (Suhoy, 2009), Italy (D'Amuri, 2009), Norway (Anvik and Gjelstad, 2010), the United Kingdom (McLaren and Shanbhogue, 2011) and the United States (D'Amuri and Marcucci, 2009).

In addition to unemployment figures, the usefulness of internet search-related data has also been explored in other areas such as the housing market. Hence, Chamberlin (2010), McLaren and Shanbhogue (2011) and Wu and Brynjolfsson (2010) have focused on housing prices and sales in the United Kingdom and the United States, while Webb (2009) has examined US home foreclosures.

Private consumption and retail sales have also been of interest to several authors (Carrière-Swallow and Labbé, 2013; Chamberlin, 2010; Della Penna and Huang, 2009; Kholodilin et al., 2010; Jun et al., 2014; Lindberg, 2011; Vosen and Schmidt, 2011, 2012a, 2012b). In this sense, Vosen and Schmidt (2012b) have illustrated how internet search data can be especially useful to predict private consumption in the presence of some irregular events, such as “stimulus” public programmes (i.e. “cash for clunkers” programmes).

Moreover, Guzman (2011) has explored the predictive power of internet queries' data in forecasting inflation expectations, while Artola and Galán (2012) and Bing Pan and Song (2012) have used this data to improve the forecasting performance regarding tourist flows.

Several studies have also investigated whether internet search data might improve the prediction of financial indicators such as investors' attention (Da et al., 2011), stock trade volume and volatility (Joseph et al., 2011; Vlastakis and Markellos, 2012), firms' share prices (Siganos, 2013), foreign currencies (Smith, 2012), and financial market fluctuations (Preis et al., 2010), among others.

Overall, this literature highlights the multiple applications that online search-related data has for forecasting some key economic variables.

## 3. Data and methodology

### 3.1. Data

This paper combines variables coming from several data sources. The variable of interest is given by registered unemployment in Spain. This information comes from the Registered Unemployment Statistics, monthly published by the Spanish Ministry of Employment and Social Security (2014), and which provides data on the number of job demands registered by the public employment services. The Spanish labour market can also be monitored through the Economically Active Population Survey (EAPS), which collects quarterly information from a sample of 65,000 households on population' participation in the labour market (National Institute of Statistics, 2004). This paper uses the data from the former database since it allows a larger sample size than the latter (monthly versus quarterly data). The period of analysis is from January 2004 to December 2012, taking the year 2013 as forecasting horizon.

Fig. 1 plots the evolution of registered unemployment in Spain in the period of analysis, January 2004–December 2012. The effect of the economic crisis over the labour market is clearly seen: in mid-2008, unemployment showed a sharp increase that accelerated over the end of that year. While Spain had started the year 2008 with 2,261,925 registered job demands, that figure increased in 1,065,876 along that year, so January 2009 registered a total of 3,128,963 job demands. The increase in job demands over 2007 had only been at 179,417 (January 2008/January 2007).

With the aim of properly analysing (and then forecasting) the evolution of unemployment in Spain, explanatory variables from both the demand and the supply sides of the labour market are taken into account.

On the supply side, data on online searches for job vacancies is incorporated into the model. Following some previous

<sup>1</sup> All these figures refer to December 2013's rates of unemployment.

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات