Nowcasting private consumption: traditional indicators, uncertainty measures, and the role of internet search query data[1]

María Gil
Banco de España, Madrid, Spain

Javier J. Pérez*
Banco de España, Madrid, Spain, javierperez@bde.es

A. Jesús Sánchez-Fuentes
University Complutense, Madrid, Spain

Alberto Urtasun
Banco de España, Madrid, Spain

Abstract

The aim of this paper is to nowcast quarterly private consumption in Spain. We estimate a suite of mixed-frequency models on a real-time database for the period 2005Q1-2015Q4, and conduct out-of-sample forecasting exercises to assess the relevant merits of different groups of indicators. The selection of indicators is guided by the standard practice ("hard" and "soft" indicators), but also expand this practice by looking at non-standard variables, namely: (i) a suite of proxy indicators of uncertainty, calculated at the monthly frequency; (ii) two additional sets of variables that are sampled at a much lower frequency: Credit card transactions and indicators based on search query time series provided by Google Trends. The latter set of indicators is based on factors extracted from consumption-related search categories of the Google Trends application. We also illustrate how Google data (sampled at a frequency higher than monthly) can be instrumental to perform event studies, by looking at possible anticipation effects related to VAT increases.

Keywords: Consumption; uncertainty; nowcasting; Google Trends.

1. Introduction

Private consumption represents between 60% to 80% of an average OECD country gross domestic product. Thus the importance, for applied forecasters, of having accurate nowcasts for this GDP component in real-time. Timely data useful to approximate private households' spending decisions are available, both covering "hard" and "soft" (survey-based, sentiment indicators) information. The standard leading indicators of private consumption used by practitioners and academics alike, nonetheless, are typically available in real-term with a 1 to 2 months delay, depending on the country, and are available at the monthly frequency.
More recently, other, less standard, sources of advanced information on consumer spending decisions are starting to be explored in the literature. Among other promising avenues, one example is the use of data collected from automated teller machines (ATMs), encompassing cash withdrawalls at ATM terminals and debit card payments. Typically, electronically recorded data are available in a quite timely fashion and are free of measurement errors (see Duarte et al., 2017, and the references quoted therein). Another prominent example is the construction of indicators of consumption behaviour on the basis of internet search patterns as provided by Google Trends. Over the past decade, the number of Internet users has increased dramatically, and also their buying patterns. In this way *intentions to buy*, as reflected in Internet searches of certain categories of goods and services, might be useful to anticipate *actual* buying behaviour. While indicators linked to income indicate the *ability to spend* of consumers, and survey-based indicators capture the *willingness to spend*, Google-searches-based variables based on consumption-related search queries may provide a measure

of consumers' *preparatory steps to spend* (see Vosen and Schmidt, 2011, 2012; Choi and Varian, 2012). Finally, a recent strand of the literature has highlighted the relevance of the level of uncertainty prevailing in the economy for private agents' decision-making, and has offered a wealth of indicators aiming at measuring it (see, among others, Backer et al., 2017; Gil et la., 2017). This is all the more relevant in the field of consumption decisions, as prescribed by the existing theoretical literature.

These new sources of information might be instrumental for the analysis of private consumption developments in real-time, in particular as regards its very short-term dynamics (nowcasting). To exploit the data in an efficient and effective manner, we estimate a suite of mixed-frequency models on a real-time database for the period 1995Q1-2016Q4, at the monthly frequency, and conduct out-of-sample forecasting exercises to assess the relevant merits of different groups of indicators. We focus on the case of Spain.

The selection of indicators starts from standard practice ("hard" and "soft" indicators), and moves forward to incorporate non-standard variables, namely: (i) a suite of proxy indicators of uncertainty, calculated at the monthly frequency; (ii) Credit card transactions; (iii) indicators based on search query time series provided by Google Trends. The latter indicator is based on factors extracted from consumption-related search categories of the Google Trends application. An optimal way to use our data is to build a single model that relates data at all frequencies. In this paper we construct multivariate state space mixed-frequencies models for the quarterly private consumption aggregate, and on the monthly information (by blocks). Our approach is closely related to that of Harvey and Chung (2000), and Pedregal and Prez (2010).[2] These papers use a temporal aggregation method that relies on the information contained in related indicators observed at the desired higher frequency. The statistical treatment of structural time series models is based on the state space form and the Kalman Filter (see Harvey, 1989). The basic model is of the Unobserved Component Model class known as the Basic Structural Model (Harvey, 1989), that decomposes a set of time series in unobserved though meaningful components from an economic point of view (mainly trend, seasonal and irregular). The mixture of frequencies, and the estimation of models at the monthly frequency, implies combining variables that at the monthly frequency can be considered as stocks with those being pure flows. The quarterly private consumption figures cast into the monthly frequency is a set of missing observations for the first months of the quarter (January and February, in the case of Q1) and the observed value assigned to the last month of each quarter (say, March). Theoretically, the quarterly National Accounts series would be obtained from monthly National Accounts series by summation of the 3 months of a quarter (January to March) had them been available.

The rest of the paper is organized as follows. In Section 2 we describe the data sources. In Section 3 we describe the econometric methodology used in the nowcasting exercise, and in Section 4 the results of the empirical exercise. Finally, in Section 5 we provide the main conclusions of the paper.

2. The data

Traditional indicators For our analysis we use a set of hard and soft indicators commonly used for private consumption forecasting. All these indicators are provided on monthly basis, although with different lags between the publication date and the month they refer to, and show a high correlation with private consumption. In this set of indicators we include as "hard" indicators Social Security registrations (Ministry of Social Security), the Retail Trade Index and the Services Activity Index (both provided by the National Statistical Institute), while as "soft" indicators we consider PMI Services (provided by Markit Economics) and the Consumer Confidence Index of the European Commission (Business and Consumer Surveys).

From a real-time perspective, Social Security registrations, the PMI Services and the Consumer confidence Index are available with a one-month lag, while the Retail Trade Index presents a lag of two months, and the Services Activity Index of three months.

Uncertainty indicators Since the end of the financial crisis, a number of geopolitical events have brought

[2] Other approaches for modeling data at different sampling intervals are the methods based on regression techniques (Chow and Lin, 1971, Guerrero, 2003), the MIDAS (MIxed DAta Sampling) approach (see Ghysels et al., 2006, Clements and Galvo, 2008), the state space approaches of Liu and Hall (2001) and Mariano and Murusawa (2003), or the ARMA model with missing observations of Hyung and Granger (2008).

to the forefront of the policy discussion the risks that heightened economic uncertainty may pose for global economic prospects. Most recent events in the second half of 2016 include the (mostly) unexpected results of the UK referendum on the EU (the so-called *Brexit*), the Constitutional referendum in Italy, or the victory of President Donald Trump in the US. By now, it is well established in the theoretical and empirical literature that heightened economic uncertainty has the potential to harm economic activity (see, among others, Bloom, 2014). In the recent empirical literature, a number of works have dealt with the hurdle of finding proxy measures of economic uncertainty, being the later a non-observable variable. The extant studies tend to focus on one specific proxy or method, the most popular ones being: (i) stock market volatility (see, e.g. Leahy and Whited, 1996; Bloom, 2009; Caggiano et al, 2014); (ii) the variance of forecasters' expectations, in many cases approximated by a concept of disagreement (see, e.g., D'Amico and Orphanides, 2008; Bachmann et al., 2013; Balta et al., 2013; Popescu and Smets, 2010) ; (iii) the frequency of news related to policy uncertainty to form a proxy of policy uncertainty (Baker et al., 2016); (iv) the common components of forecast errors from several macroeconomic time series (see e.g. Jurado et al., 2013); on related grounds, some authors compute uncertainty measures on the basis of real-time forecasting models (see, e.g. Scotti, 2016).

In the current paper we focus on measures covering (ii) and (iii), the reason being that we consider them to be more related to our topic of study, namely, private consumption decisions. In particular, we use the textual indicator known as Economic Policy Uncertainty Index (EPU) for Spain elaborated by Baker, Bloom y Davis (2015), and construct measures of disagreement about private consumption (based on forecasts taken from a national panel[3]), and measures of uncertainty based on some European Commission's consumer survey forward looking questions, namely unemployment prospects and uncertainty about major expected purchases over the next 12 months.

Regarding the forward-looking indicators on "Unemployment perspectives over next 12 months" and "Perspectives of major purchases over next 12 months", we follow the approach of Bachmann et al. (2013) to construct measures of uncertainty that exploit the information contained in the dispersion of responses. Specifically, respondents to the above-mentioned questions can be grouped in three answers: "decrease", "unchanged" or "increase". Following Bachmann et al. (2013), let $Frac_t^+$ denote the weighted fraction of consumers in the cross section with "increase" responses at time t, and $Frac_t^-$ the weighted fraction of consumers with "decrease" responses. Then the "uncertainty indicator" is computed as

$$\sqrt{Frac_t^+ + Frac_t^- - \left(Frac_t^+ - Frac_t^-\right)^2}$$

As to the measure of disagreement about private consumption forecast constructed, we take as starting point the month t cross section of current and one-year-ahead forecasts about national accounts' private consumption produced by analysts that do respond to the "FUNCAS panel". FUNCAS is an independent institution that has been compiling forecaster's views since 1999. At each point in time, the measure of "disagreement" is computed as the standard deviation of such cross-section of n forecasters $\frac{1}{n}\sum_{i=1} n\left(\hat{C}_i - \hat{C}_A\right)^2$. Given that each analyst provides growth rates of two fixed-event forecasts (current and year-ahead) m months ahead, it is necessary to correct each time-t value by the fact that it is computed on an evolving information set. For that, we follow the methodology of Dovern, Fritsche y Slacalek (2012):

$$F_{y_0,m,12}^{fh}(x) = \frac{12-m+1}{12} F_{y_0,m,y_0}^{fe}(x) + \frac{m-1}{12} F_{y_0,m,y_{0+1}}^{fe}(x)$$

Credit card data We use the monthly amounts spent by households that have been paid by means of credit/debit cards. This is a widespread means of payment by Spanish consumers.

Internet search query data (Google Trends) The paper introduces a new indicator for private consumption which is constructed using data on internet search behavior provided by Google Trends. Due to the increasing popularity of the internet it is certain that a substantial amount of people also use web search

[3]Bi-monthly FUNCAS panel.

Table 1: Consumption categories according to the National Accounts and Google Trends searches.

Classification by national product and income accounts (NIPAs)	Google categories
Durable goods	
Motor vehicles and parts	Automotive, auto financing, automotive parts, auto insurance, Seat, Mercedes Benz, Mercedes offer, second hand car, car, to buy a car
Furnishing and durable household equipment	Electrical appliance, home insurance, home remodel, home furnishing, interior decoration, interior design
Recreational goods and vehicles	Online movie, to buy a movie, watch online movie, video games
Other durable goods	Telecommunications, router wifi, mobile phone, electronic book, novel
Nondurable goods	
Food and beverages	Food and beverages, food, drink
Clothing & footwear	Clothing, second hand clothing, footwear, second hand footwear, male & female lingerie, undergarments, T-shirts
Gasoline and energy goods	Electricity, energy, gaoline, gas
Other nondurable goods	Personal care, beauty, chemicals, medications, face & body care, beauty products, newspapers, tobacco
Services	
Household consumption expenditures	
Housing and utilities	Home & auto insurance, house remodel, interior decoration, interior design, real estate agency
Health care	Health, health insurance, medical services, mobile phone, wireless
Transportation services	
Recreational services	Leisure, video games, online movie, to buy a movie, watch online movie, ticket sales
Food services and accommodation	Hotels, accommodation, restaurant, restoration, terrace, welfare
Financial services and insurance	
Other services	Telecommunications, life insurance, social services

SOURCE: National Accounts, Google Trends.

engines to collect information on goods they intend to buy. To use Google Trends data for forecasting private consumption, common unobserved factors are extracted from time-series of web search categories provided by the Google Trends application.

Google Trends provides an index of the relative volume of search queries conducted through Google. The application provides aggregated indexes of search queries which are classified into a total of 605 categories and sub-categories using an automated classification engine. Google Trends provides a time series index of the volume of queries users enter into Google in a given geographic area. The query index is based on query share: the total query volume for the search term in question within a particular geographic region divided by the total number of queries in that region during the time period being examined. The maximum query share in the time period specified is normalized to be 100, and the query share at the initial date being examined is normalized to be zero. Google Trends data are provided on a weekly basis. We compute monthly averages since data consumption are only available on monthly basis. The Google time-series are not seasonally adjusted. All data are available only since 2004. We use this data in levels.

We select 72 consumption-relevant categories that in our view are best matches for the product categories of personal consumption expenditures of the BEA's national income and product accounts, described in Table 1. Then we use principal components analysis to extract the common factors to these 72 consumption-relevant categories. We include in our database the first four components that explain almost 50% of the total variance of the whole set.

3. Methodology

Econometric methodology The exposition in this section follows closely Pedregal and Pérez (2010). The starting point of the modeling approach is to consider a multivariate Unobserved Components Model known as the Basic Structural Model (Harvey, 1989). A given time series is decomposed into unobserved components which are meaningful from an economic point of view (trend, T_t, seasonal, S_t, and irregular, e_t). Equation (1) displays a general form, where t is a time sub-index measured in months, z_t denotes the variable in National Accounts terms expressed at a quarterly sampling interval for our objective time series (private consumption), and u_t represents the vector of monthly indicators.

$$\begin{bmatrix} \mathbf{z}_t \\ \mathbf{u}_t \end{bmatrix} = \mathbf{T}_t + \mathbf{S}_t + \mathbf{e}_t \tag{1}$$

The general consensus in this type of multivariate models in order to enable identifiability is to build SUTSE models (Seemingly Unrelated Structural Time Series). This means that components of the same type interact among them for different time series, but are independent of any of the components of different types. In addition, statistical relations are only allowed through the covariance structure of the vector noises, but never through the system matrices directly. This allows that, trends of different time series may relate to each other, but all of them are independent of both the seasonal and irregular components. The full model is a standard BSM that may be written in State-Space form as (see Harvey, 1989)

$$\mathbf{x}_t = \mathbf{\Phi x_{t-1}} + \mathbf{E w_t} \tag{2}$$

$$\begin{bmatrix} \mathbf{z}_t \\ \mathbf{u}_t \end{bmatrix} = \begin{bmatrix} \mathbf{H} \\ \mathbf{H}^u \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} \epsilon_t \\ \mathbf{v}_t \end{bmatrix} \tag{3}$$

where $\epsilon_t \sim N(0, \Sigma_\epsilon)$ and $\mathbf{v}_t \sim N(0, \Sigma_{\mathbf{v}_t})$. The system matrices $\mathbf{\Phi}$, $\mathbf{E}$, $\mathbf{H}$ and $\mathbf{H}^u$ in equations (2)-(3) include the particular definitions of the components and all the vector noises have the usual Gaussian properties with zero mean and constant covariance matrices (ϵ_t and $\mathbf{v}_t$ are correlated among them, but both are independent of $\mathbf{w}_t$). The particular structure of the covariance matrices of the observed and transition noises defines the structures of correlations among the components across output variables. The mixture of frequencies, and the estimation of models at the quarterly frequency, implies combining variables that at the quarterly frequency can be considered as stocks with those being pure flows. Thus, given the fact that our objective variables are observed at different frequencies, an accumulator variable has to be included

$$C_t = \begin{cases} 0, & t = firstquarter \\ 1, & otherwise \end{cases} \tag{4}$$

so that the previous model turns out to be

$$\begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{bmatrix} = \begin{bmatrix} C_t \otimes \mathbf{I} & \mathbf{H\Phi} \\ \mathbf{0} & \mathbf{\Phi} \end{bmatrix} \begin{bmatrix} \mathbf{z}_{t-1} \\ \mathbf{x}_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & \mathbf{HE} \\ \mathbf{0} & \mathbf{E} \end{bmatrix} \begin{bmatrix} \epsilon_t \\ \mathbf{w}_t \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} \mathbf{z}_t \\ \mathbf{u}_t \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^u \end{bmatrix} \begin{bmatrix} \mathbf{z}_t \\ \mathbf{x}_t \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{v}_t \tag{6}$$

Given the structure of the system and the information available, the Kalman Filter and Fixed Interval Smoother algorithms provide an optimal estimation of states. Maximum likelihood in the time domain provides optimal estimates of the unknown system matrices, which in the present context are just covariance matrices of all the vector noises involved in the model. The use of the models selected and the estimation procedures described in the previous paragraph, allows the estimation of models with unbalanced data sets, i.e. input variables with different sample lengths. This is a feature of relevance for the construction of the database at hand, given occasional differences in temporal coverage of indicators.

The nowcasting exercise The real time nowcasting problem is illustrated in Table 2. Consider, for example, the information available at the third month of each quarter. Approximately one and a half months before the official QNA release is published by the statistical agency, monthly employment figures and various surveys corresponding to the second month of the quarter. The retail trade index and the EPU correspond to the first month and the services activity index to the last month of the previous quarter. Finally, indicators based on search query time series are available, at least partially, for the third month (lags are the same for the first and the second month).

Different information sets within the quarter are available, so the information that can be exploited to nowcast consumption is different in each month of the quarter. In this respect, we consider three set of forecasts: (i) first month of the quarter $m1$, (ii) second month of the quarter $m2$ and (iii) third month of the quarter $m3$. Several models are considered, that differ in the set of indicators include in each model. The complete list of models used can be found in Table 3.

4. Empirical results

Table [to be added] provides different metrics regarding the forecast performance of the models. In the last block of the table we provide the relative RMSE to a nave model (random walk, computed by repeating the latest q-q available growth rate). As seen in the table, most of the models outperformed the random walk nowcast.

Regarding the contribution of the different variables, the models that do include the indicator related to credit card transactions deliver the best forecasting performance overall. Indeed, model 3 (social security registrations, retail trade index and credit card transactions) always gives the best results. It is important to note the good forecasting performance of model 10, that only includes credit card transactions. As regards the value added of other "non-standard" indicators, we look at models 3, 4 and 5, than sequentially include credit cards, EPU and Google Trends to a model that contains as baseline variables Social Security Registrations and the retail Trade Index. As we said above model 3 provides the highest forecast accuracy over the sample and within quarters. Model 4 (with EPU) also outperforms the random walk. On the other side, model 5 (including Google indicators) only outperform the random walk in the first month of the quarter, when some information of the QNA is still missing, but apparently these indicators seem to be introducing noise in the forecast exercise, delivering relative bad results in $m2$ and $m3$ forecast origins. These results certainly indicate that the computed principal components based on Google variables have to be revised, in order to find ways to maximize the information content of such indicators, trying alternative ways of aggregation and/or exploiting the granular information available, given that the literature for other countries/macro variables has shown that Google-based indicators tend to incorporate substantial nowcasting power.

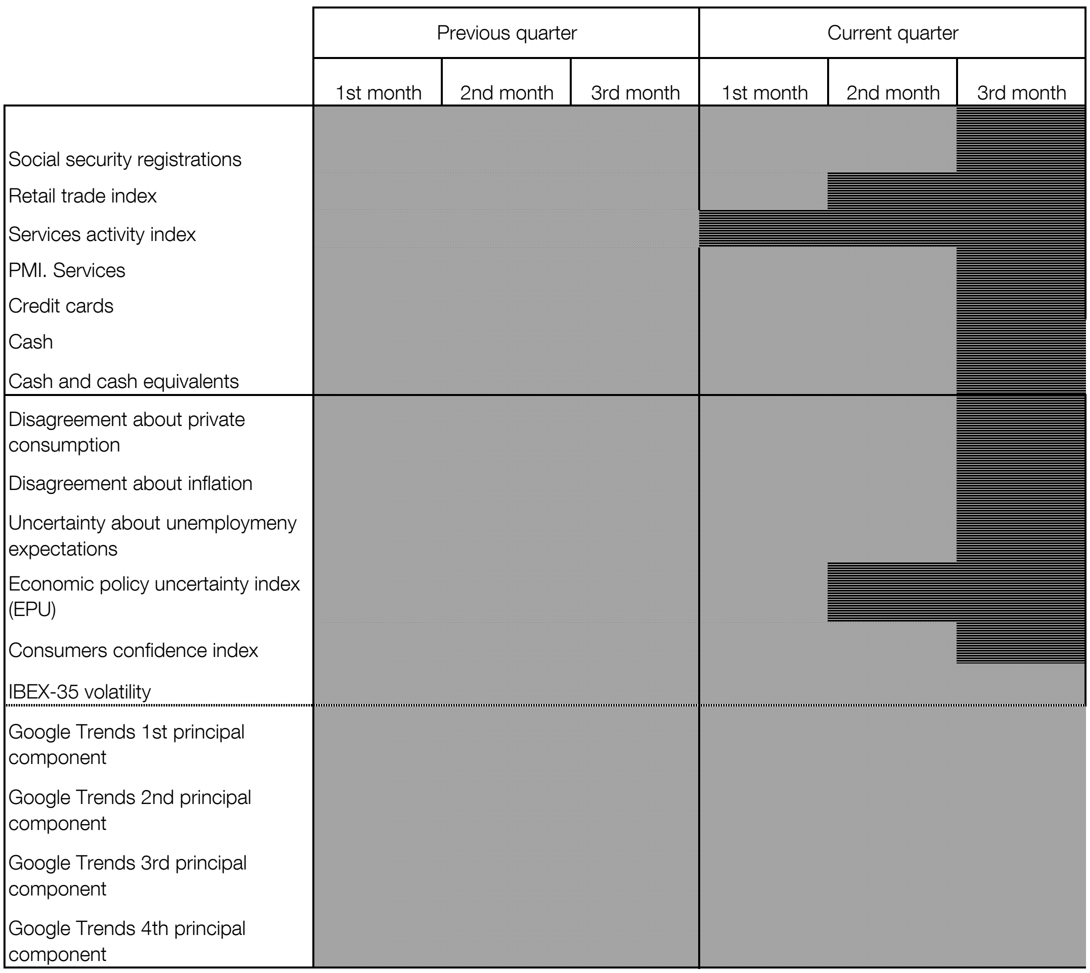
5. Conclusions

We estimate a suite of mixed-frequency models on a real-time database for the period 2005Q1-2015Q4, and conduct out-of-sample forecasting exercises to assess the relevant merits of different groups of indicators. The selection of indicators is guided by the standard practice ("hard" and "soft" indicators), but also expand this practice by looking at non-standard variables, namely: (i) a suite of proxy indicators of uncertainty, calculated at the monthly frequency; (ii) two additional sets of variables that are sampled at a much lower frequency: Credit card transactions and indicators based on search query time series provided by Google Trends. The latter set of indicators is based on factors extracted from consumption-related search categories of the Google Trends application. Our very preliminary results highlight the usefulness of credit card spending variables for nowcasting NA private consumption. At the same time, the principal components extracted from the set of Google information do not seem to add much forecasting power, which suggest that the procedure we used to summarize the granular information coming from internet searches needs to be refined.

References

Bontempi, M. E., R. Golinelli, and M. Squadrani (2016), "A new index of uncertainty based on Internet searches: a friend or foe of other indicators?", Quaderni Working Paper DSE N 1062, Department of Economics, Universit di Bologna.

Table 2: High frequency variables used in the study and information flow (information available at nowcasting origin $m3$.



SOURCE: own elaboration.

a. Horizontal lines denote lack of availability of the indicator in a particular point in time within the quarter.

Table 3: Indicators included in the different models.

M1	Social security registrations	Retail trade index	Activity services index	PMI. Services	Consumer confidence index
M2	Social security registrations	Retail trade index			
M3	Social security registrations	Retail trade index	Credit cards		
M4	Social security registrations	Retail trade index	Economic policy uncertainty index (EPU)		
M5	Social security registrations	Retail trade index	Google Trends (1st principal component)	Google Trends (2nd principal component)	
M6	Activity services index	PMI. Services	Consumer confidence index		
M7	Activity services index	PMI. Services	Consumer confidence index	Credit cards	
M8	Activity services index	PMI. Services	Consumer confidence index	Economic policy uncertainty index (EPU)	
M9	Activity services index	PMI. Services	Consumer confidence index	Google Trends (1st principal component)	Google Trends (2nd principal component)
M10	Credit cards				
M11	Disagreement about consumption forecasts	Uncertainty about major purchases	Uncertainty about unemployment expectations	Economic policy uncertainty index (EPU)	
M12	Disagreement about consumption forecasts	Uncertainty about major purchases	Uncertainty about unemployment expectations		
M13	Economic policy uncertainty index (EPU)				
M14	Google Trends (1st principal component)	Google Trends (2nd principal component)	Google Trends (3rd principal component)	Google Trends (4th principal component)	
M15	Google Trends (1st principal component)	Google Trends (2nd principal component)			
M16	Social security registrations	Retail trade index	Activity services index	PMI. Services	Consumer confidence index
	Credit cards	Disagreement about consumption forecasts	Economic policy uncertainty index (EPU)	Google Trends (1st principal component)	Google Trends (2nd principal component)

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

Choi, H. and H. Varian (2012), “Predicting the present with Google Trends”, *Economic Record*, 88, pp. 2-9.

Degrer, Ch. and K. A. Kholodilin (2013), “Forecasting private consumption by consumer surveys”, *Journal of Forecasting*, 32, pp. 10-18.

Duarte, C., P. M. M. Rodrigues, and A. Rua (2017), “A mixed frequency approach to the forecasting of private consumption with ATM/POS data”, *International Journal of Forecasting*, 33, pp. 61-75.

Edelman, B. (2012), “Using Internet data for economic research”, *Journal of Economic Perspectives*, 26, pp. 189-206.

Einav, L. and J. D. Levin (2013), “The data revolution and economic analysis”, Working Paper 19035, NBER Working Paper Series, National Bureau of Economic Research.

Faberman, J. and M. Kudlyak (2016), “What does online job search tell us about the labor market?”, Economic Perspectives, Federal Reserve Bank of Chicago, 1.

Fondeur, Y. and F. Karamé (2013), “Can Google data help predict French youth unemployment?”, *Economic Modelling*, 30, pp. 117-125.

Gil, M., J. J. Pérez and A. Urtasun (2017), “Macroeconomic uncertainty and its effects: the case of Spain Spain”, forthcoming, Bank of Spain Working Paper.

Guzmán, G. (2011), “Internet search behaviour as an economic forecasting tool: The case of inflation expectations”, *Journal of Economic and Social Measurement*, 36, pp. 119-167.

Harvey, A. (1989), Forecasting structural time series models and the Kalman Filter. UK: Cambridge University Press.

Kholodilin, K., M. Podstawski, and B. Siliverstovs (2010), “Do Google searches help in nowcasting private consumption? A real-time evidence for the US”, DP 997, DIW Berlin, German Institute for Economic Research.

Koop, G. and L. Onorante (2013), “Macroeconomic nowcasting using Google probabilities”, mimeo.
Li, X. (2016), “Nowcasting with big data: is Google useful in presence of other information?”, mimeo, London Business School.

Mazzi, G.L. (2016), “Some guidance for the use of Big Data in macroeconomic nowcasting”, First International Conference on Advanced Research Methods and Analytics, Universitat Politecnica de Valencia, Valencia.

Pedregal, D. J. and J. J. Pérez (2010), “Should quarterly government fi-nance statistics be used for fiscal surveillance in Europe?”, *International Journal of Forecasting*, 26, pp. 794-807.

Penna, N. della and H. Huang (2009), “Constructing consumer sentiment index for U.S. using Google searches”, Working Paper n 2009-26, Department of Economics, University of Alberta.

Seabold, S. and A. Coppola (2015), “Nowcasting prices using Google trends”, Policy Research Working Paper n. 7398, Macroeconomics and Fiscal Management Global Practice Group, World Bank Group.

Schmidt, T. and S. Vosen (2012), “Using Internet to account for special events in economic forecasting”, Ruhr Economic Papers, 382.

Smith, P. (2016), “Google’s MIDAS touch: predicting UK unemployment with Internet search data”, *Journal of Forecasting*, 35, pp- 263-284.

Taylor, L., R. Schroeder and E. Meyer (2014), “Emerging practices and perspectives on Big Data analysis in economics: bigger and better or more of the same?”, Big Data and Society, 1-10.

Tomczyk, E. and T. Doligalski (2015), “Predicting new car registrations: nowcasting with Google search and macroeconomic data”, [in:] Sl. Partycki (ed.), e-Society in Middle and Eastern Europe. Present and Development Perspectives, Wydawnictwo KU, 228-236.

Toth, I. J. and M. Hadju (2012), “Google as a tool for nowcasting household consumption: estimations on Hungarian data”, Institute for Economic and Enterprise Research, HCCI.

Varian, H. R. (2014), “Big Data: new tricks for econometrics”, *Journal of Economic Perspectives*, 28, pp. 3-28.

Vosen, S. and T. Schmidt (2011), “Forecasting Private Consumption: Survey-Based Indicators vs. Google Trends”, *Journal of Forecasting*, 30, pp. 565-578.

Vosen, S. and T. Schmidt (2011), “A monthly consumption indicator for Germany based on Internet search query data”, *Applied Economics Letters*, 19, pp. 683-687.