Bank Deposits and Google Searches in a Crisis Economy: Bayesian Non-linear Evidence for Greece (2009-2015)

Konstantinos N. Konstantakis
National Technical University of Athens, Greece

Despoina Paraskeuopoulou
National Technical University of Athens, Greece

Panayotis G. Michaelides
National Technical University of Athens, Greece

Mike G. Tsionas
Lancaster Business School, UK

ABSTRACT

Due to a variety of reasons, the Greek economy faced a severe crisis being a member of EMU. Nonetheless, the country's banking system experienced a dramatic outflow of deposits, in the period 2009-2015. The present paper attempts to shed light on the possibility of forecasting bank deposits, based on the keyword "Grexit" of Google searches. In this context, apart from standard forecasting models like AR (p) and ARDL (p, q) we estimate a novel Neural Network ARDL (p, q, G) model and its respective Bayesian modification. We show that extending standard autoregressive models with the information provided by Internet Searches leads to significant improvement in the forecasting accuracy of the Bank Deposits, compared with other standard models. Furthermore, the forecasting performance of the models, which are extended with Google searches, is shown to be better than models containing only the well-known indicators. Our findings are robust and econometrically sound.

Keywords: Forecasting, Crisis, Deposits, Greece, Grexit, Google Trends

JEL classifications: C22, C53

1. Introduction

Nowadays, the internet has changed the dissemination of information by making it very easily accessible to almost everyone at a very low cost or even freely available. However, obtaining the appropriate information can be a difficult task because of the enormous size of the Internet and this is the reason why nowadays people rely on search engines to locate information on the web (Vlastakis and Markellos 2012).

The global financial crisis of 2007 and the structural debt European crisis of 2009 have dictated the evolution of the banking system as well as of the financial linkages among institutions and economies over the last years. In this turbulent environment, the banking sector, and especially commercial banks, continue to operate under a lot of pressure, and their main income-generating activity, i.e. bank deposits, are subject to great instability. Bank deposits constitute probably the most important indicator of a bank's viability; especially in the European Monetary Union (EMU), where the periodic banks stress tests demand an increasing quantity of bank deposits. This increasing demand for large quantities of bank deposits or large ratio of bank deposits over loans, is attributed to the fact that the quantity or percentage of bank deposits is often associated with bank failures and financial crises in both developing and developed countries, the so-called "bank- runs".

A critical issue regarding the "bank-runs" is the fact that "herding" behavior of depositors could potentially lead sound banks to fail. Of course, this type of behavior is difficult to be quantified using *ex-ante* measures. Nonetheless, technological advancement has led to the increasing use of internet, in general, and social platforms, in specific, in a daily context. Here, we argue that the evolution of internet usage could provide useful data for financial institutions that could be used to uncover future trends and potentially "herding" behavior.

In this context, in case the information extracted from the internet usage is "negative", in the sense that it would negatively affect the future behavior of banks through their systemic risk exposure, then it would be possible to witness "herding" behavior. Therefore, in this work we suggest that internet searches of key words of distress that are associated with the banking system could act as valuable proxies for capturing the potential "herding" behavior among depositors *ex-post*.

The exploration of the determining factors of bank deposits is an issue of substantial importance for regulatory authorities concerned about financial stability. In this context, the utilization of internet media-based information for forecasting key developments in the financial sector has attracted limited attention in the relevant literature, so far. In this work, we show that the last decades' advances in internet technology, which permit us to have direct access to a vast amount of information such as internet searches, offers the possibility of forecasting key measures in the economy's banking system, such as the amount of bank deposits, which is of crucial importance.

According to Kolapo et al. (2018): "The leading role played by banking institutions in developing economies cannot be undermined. In the last decade, institutional credit to the private and public sectors for investment purposes has increased significantly; thereby helping to build up huge infrastructural facilities, capital project backing as well as meeting other recurrent expenditures of the government respectively". Bank deposits act as lenders of short-and medium term loans to both private and public sector borrowers, hence they create the money they lend out to customers collectively as a system. Given the importance of bank deposits for both the public and private economy, the focus of the present work is on the impact of the volume of web search on the volume of bank deposits, in the period 2009-2015.

The Greek crisis could act as a "clean" prototype case for the empirical investigation of bank deposits under capital controls for econometrically examining the role of web searches in determining the amount of bank deposits, over the potential impact of the standard macroeconomic and bank-specific variables. As the Greek crisis was deepening, the term "Grexit" was added to the world's everyday vocabulary. Since then, the Greek economy was "bailed out" several times. In fact, in this work, we argue that the World Wide Web, with its property of being a storehouse of a huge amount of valuable and important information, could contribute to the forecasting of upcoming events in banking and finance (see e.g. Dergiades et *al.*, 2015).

We believe that such an approach is nowadays indispensable, since the amount of media coverage of the various events perfectly captures the general interest of the public about this specific event and could easily act as a tracking device of the unfolding mechanism of this specific situation. In this context, this paper focuses on the research question: *"Can information extracted on the internet be useful in forecasting the future trends in the banking sector?"* In order to tackle this research question we will employ autoregressive models, which will be augmented with non-linear terms to account for non-linearities. More precisely, we will combine Bayesian ARDL modeling with non-parametric terms whose critical feature is that they let the dataset itself serve as evidence to support the model's approximation of the functional form of the bi-variate relationship, instead of imposing some linear functional form, as is the case with standard ARDL models.

The dependent variable will be bank deposits in Greece over the period 2009-2015. The crucial explanatory variable expresses the internet searches for the word "Grexit". The mechanism that is in place behind the connection between the quantity of bank deposits and “Grexit” is the following: Greek consumers are informed through the media regarding the potential exit of Greece from the EMU. This information is communicated to Greek depositors via the acronym “Grexit”. In the light of this new information, Greek depositors start to investigate the consequences of “Grexit” regarding the general economy and thus their savings that are deposited in the Greek banking system. Due to the fact that the Greek deposits were guaranteed up to the amount of one hundred thousand euros (100,000€) by the Greek Government, they realized that if “Grexit” occurs, then their deposits might be in jeopardy. As a result, increased searches of the word “Grexit”, signifies a change in the information set of the depositors regarding the safety of their deposits leading to a potential change in the actual quantity of deposits. The present paper examines exactly the validity of this mechanism based on a robust econometric framework.

“The concept of “Grexit" became popular in December 2014 when the Greek Parliament failed to elect a new president, leading to Parliament’s dissolution and the announcement of new elections in January 2015. In the aftermath of the elections, the possibility of Greece’s exit from the euro area began to solidify in popular imagination and become viewed as a likely outcome and, even, as an extreme remedy to the ongoing crisis” (Wildmer and Sacchi 2017). A potential “Grexit” would have led to the issuing of a “drachma” which would then be depreciated as it could not be supported by the fundamentals of the Greek economy.

In brief, the paper contributes to the literature in the following ways: (i) It is the first in the relevant literature to relate the pattern of bank deposits in the economy with Google searches. (ii) It is the first that develops a novel Bayesian ARDL scheme based on non-parametric terms, which allows the researcher to condition on a set of relevant economic and / or financial variables, and (iii) it allows the identification of relationships even if the true character between two variables is non-linear in nature.

The contribution of the paper is in the description of the role of Google searches for approximating the herding behavior. The empirical application of the paper focuses on the Greek economy and does not attempt to provide an analytical model for bank-runs.

The paper has the following structure: Section 2 offers a review of the literature; Section 3 sets out the methodological framework; Section 4 presents the empirical results; Section 5 discusses the findings and, Section 6, concludes the paper.

2. Background Literature

As we have mentioned, a critical issue regarding the "bank-runs" is the fact that the "herding" behavior of depositors could potentially lead sound banks to fail. Of course, this type of behavior is difficult to be quantified using *ex-ante* measures. Nonetheless, technological advancement has led to the increasing use of internet, in general, and social platforms, in specific, in a daily context. Here, we argue that the evolution of internet usage could provide useful data for financial institutions that could be used to uncover future trends and potentially "herding" behavior.

At the next step, it can now be safely argued that *Google Searches* are triggered on the arrival of information for a country's economic fundamentals or even the social or political situation. Put differently, this news are assumed to express information about the country's economic, financial and social future. In this context, in case the information extracted by these searches is "negative" in the sense that it affects negatively the future expectations of public opinion regarding the economy's fundamentals, then it is possible that these searches could be used in order to proxy "herding" behavior of individuals.

The Greek banking sector reflects the macroeconomic environment in Greece, which has traditionally been regulated, especially during the 80s (OECD 1986) and credit granting decisions were often based on personal contacts (Gibson and Tsakalotos 1992, p. 61). However, the recent global recession led to distress for the Greek banking sector.

The banking system of Greece since the 1990s represents a so-called "clean" prototype case for the empirical investigation of bank deposits under capital controls. Specifically, the banks in Greece operate within a liberalized institutional environment, in the context of an advanced and relatively closed economy, which was growing rapidly, until the outbreak of the crisis. Furthermore, banks follow a traditional business model involving mainly deposit-taking and loan-granting, while their trading activities are relatively limited and the shadow banking sector is not developed. Finally, the value of the currency is stable due to the participation of Greece in the Eurozone. The aforementioned features of the macroeconomic and banking environment ensure that there is no significant impact by additional complicating factors (e.g. banks being highly involved in trading activities, or swings in international trade, or exchange rates affecting the macroeconomic environment).

Now, turning to the phenomenon of "bank-runs", as put by Diamond and Dybvig (1983, p. 401): "Bank runs are a common feature of the extreme crises that have played a prominent role in monetary history. During a bank run, depositors rush to withdraw their deposits because they expect the bank to fail. In fact, the sudden withdrawals can force the bank to liquidate many of its assets at a loss and to fail. In a panic with many bank failures, there is a disruption of the monetary system and a reduction in production".

In order to understand the complexity of the various banking products, we elaborate on the mechanism that dictates the evolution of bank operations. Banks make loans at a high price that cannot be sold expeditiously. Banks issue deposits of demand allowing depositors to withdraw at any time. This liquidity mismatch, in which the liabilities of a bank are more liquid than their assets, has caused problems for banks when an extravagant quantity of depositors try to withdraw at once (a situation termed a bank run).

Diamond and Dybvig (1983) developed a model to explain why banks choose to issue more liquid deposits than their assets and why banks are subject to runs. The model was commonly used for understanding bank runs and other types of financial crises, as well as ways of avoiding these crises. In the words of Prescott (2010, p. 1), the Diamond-Dybvig (1983) model has three basic elements:

- Long-term investments that are more productive than short-term investments;
- A random need for liquidity on the part of an individual; and
- Private information about an individual's need for liquidity".

Economists and researchers have used many variations of the so-called Diamond-Dybvig model to explore banking issues (e.g., Qi [1994]; Jacklin [1993]; Russell [1993]; Haubrich and King [1990]; Engineer [1989]; Chari and Jagannathan [1988]; Freeman [1988]; Jacklin and Bhattacharya [1988]; Jacklin [1987]; Postlewaite and Vives [1987]) and continues to be widely cited as providing a definitive theoretical case for government deposit insurance (Mcculloch and Min-Teh, 1998).

Of course, based on economic theory, banks are responsible for determining money supply in the economy. In the words of Goodhart (2017, p. 33): "During the last two centuries there have been four main approaches to analyzing the determination of the money supply, to wit: (1) Deposits cause Loans, (2) The Monetary Base Multiplier, (3) The Credit Counterparts Approach and (4) Loans cause Deposits. All four approaches are criticized, especially (2) which used to be the standard academic model, and (4) which is now taking over as the consensus approach". In fact, Goodhart (2017) argues that banking is a service industry, which sets the terms and conditions whereby the private sector can create additional money for itself.

In this context, Central banks and commercial banks create new money when they grant loans or purchase assets and pay in their notes or credit the amount as a sight deposit. However, this money-creation mechanism implies that because banks can create money they can also determine the money supply. Nonetheless, it is argued that this one-dimensional interpretation of money creation exaggerates the role of banks in initiating private-sector credit expansions and fails to account for the influences that bank debtors and creditors exert over the money supply determination, including both the non-bank private and public sector (Decker and Goodhart, 2018).

As explained further by Goodhart (2017, p. 52): "The part of banking business where the ethos and approach of narrow banking would be most valuable relates to property finance. If banks were forbidden from issuing long-term mortgages until they had arranged appropriate backing from long-term funding, banking would become much safer. Moreover, the process of house purchases is generally so long drawn-out that delays caused by the need for banks to arrange appropriate funding should not be too onerous, (in contrast, most credit card and business overdraft borrowers need to make immediate payments). Perhaps because of the political sensitivities involved, proponents of the narrow banking idea rarely apply that approach specifically to housing finance, where it would do most good". Financial analysts have utilized numerous varieties of Goodhart's influential paper (e.g., Werner, 2014; Westbrook, 2017; Decker, 2018; Hartmann, Huang and Schoenmaker, 2018; Hiermeyer, 2018; Hiermeyer, 2019) and proceeds to be broadly cited as giving a definitive hypothetical case on money supply.

Also, as we have seen, only limited research has been done, thus far, on the analysis of commercial bank deposits despite being a very crucial part of the overall private financial savings, which are of outmost importance, especially in time of financial crisis. In the meantime, despite the fact that financial market activity is related to the information in the market (see, inter alia, Fama et al., 1969; Clark, 1973; French and Roll, 1986, Epps and Epps, 1973; Ederington and Lee, 1993; Tauchen and Pitts, 1983; Berry and Howe 1994; Mitchel and Mulherin 1994), limited work has been done on the role of information for banking issues.

The importance of online search activity has received some attention in the literature. Smith (2012) found evidence that information can contribute to the explanation of upcoming movements in banking and finance, and more precisely for currency markets. In this context, Da et *al.* (2011), Joseph et *al.* (2011), Beracha and Wintoki (2013), showed that online search activity or access to social media can predict price movements in the US equity market (Tett 2013). In fact, in a seminal paper by Hasan et *al.* (2013), it has been shown that financial markets are more influenced by press rumors than by the fundamentals. For the impact of Tweets and Google trends on the GIIPS sovereign spreads see Dergiades et *al.* (2015)

Thompson et *al.* (1987), Bessembinder et *al.* (1996) and Ryan and Taffler (2004) showed that firm-specific information are a significant determinant of stock prices and trading volume and some researchers even derived a relevant measure (Mitchel and Mulherin, 1994; Berry and Howe, 1994). On the relationship between risk aversion and information demand, see Willinger (1989); Eeckhoudt and Godfroid (2000); Freixas and Kihlstrom (1984) and Verrecchia (1980, 1982).

Another strand in the literature examines the relationship between bank risk and either deposit interest rates or interest costs, such as Hannan and Hanweck (1988), Cargill (1989), Ellis and Flannery (1992), Kutner (1992), Brewer and Mondschean (1994), Hess and Feng (2007), and Uchida and Satake (2009). In another strand, the disciplinary effect of reduced deposit availability is examined. See Billett et *al.* (1998), Park and Peristiani (1998), Jordan (2000), Jagtiani and Lemieux (2000), Goldberg and Hudgins

(2002), Maechler and McDill (2006), and Shimizu (2009). For relevant studies in developing or emerging economies, see Hosono (2005), Hadad et *al.* (2011), Barajas and Steiner (2000), Peria and Schmukler (2001), Calomiris and Powell (2001), Ioannidou and Penas (2010), Mondschean and Opiela (1999), Peresetsky (2008) and Karas et *al.* (2013), Kouassi et *al.* (2011), Jackowicz (2004), Kraft and Galac (2007), Onder and Ozyildirim (2008). In total, deposit insurance systems in developing countries are not always fully credible (Prean and Stix 2011).

Thus far, the existing literature on private savings includes, among others, the seminal works by Modigliani and Brumberg (1954) (*Life Cycle Hypothesis)* and by Friedman (1957) (*Permanent Income Hypothesis).* More recently, there are also the seminal articles by Deaton (1991) and Carroll (1992). However, except for these seminal works, which focus on private savings and income, no adequate research exists on commercial bank deposits.

Levy-Yeyati et *al.* (2004) showed that during crisis periods, depositors' sensitivities to macroeconomic variables increases. Also, Opiela (2004) demonstrated that in Thailand, in the period preceding the 1997 crisis, depositors monitored banks quite closely. Kraft and Galac (2007) showed that during the Croatian crisis (1998–1999), the interest-rate elasticity of deposits practically vanished. Another study on the topic is by Finger and Hesse (2009) who explain deposit evolution in Lebanon by means of a vector error correction model (VECM).

In a similar vein, see also Haron and Wan Azmi (2006) who focused on commercial banks' deposits in Malaysia. As for Brazil, Oliveira et *al.* (2011) found that during the crisis, most banks recorded a substantial increase in uninsured deposits. Using a large sample of banks from developed and emerging economies, they provided no evidence for augmented market discipline during crisis periods. The majority of the aforementioned studies, including Forssback (2011), found that depositors exhibit low sensitivity to bank fundamentals in crisis periods. Hosono (2005) demonstrated that in

South Korea, Malaysia, and Thailand, the sensitivity of deposit volumes and interest costs to bank fundamentals actually declined after 1998.

In brief, the existing literature on the effects of crisis on depositor discipline showed that market participants do monitor the risk-taking activities of banks but they are relatively poorly understood and the literature could be characterized as being inconclusive. See further Hasan et *al.* (2012) and the references therein.

Finally, as far as the data source employed in this article is concerned, it has already been used in epidemiology (Ginsberg et *al.*, 2009) and in different fields of economics (Choi and Varian 2012; Edelman, 2012). For instance, Preis et *al.* (2013) found that query volumes are directly related to stock market moves. Da et *al.* (2011) showed the importance of Google data for a sample of 3,000 stocks, whereas Vosen and Schmidt (2011) showed that Google-based indicators help in forecasting consumption. Einav and Levin (2013) showed that internet-related data have a great influence on economic research, while central banks also suggest using Google-based data along with traditional economic indicators (Artola and Galan, 2012; McLaren and Shanbhorge, 2011; Troy et *al.* 2012). Recently, D'Amuri, and Marcucci (2017) studied the crucial role of Google searches in forecasting US unemployment. In a similar vein, see also Ettredge et *al.* (2005), Antenucci et *al.* (2014), Tuhkuri (2016).

3. Econometric methodology

3.1 Linear Time Series Model

In this work, we will make use of the autoregressive (AR) approach, which can also be influenced by other variables, which are formed outside of the time-series system, called exogenous variables. Such an AR(p) model of order *p* with exogenous variables with *s*=0 order of exogenous variables, has the following general representation:

$$\Delta y_t = \sum_{i=1}^{p} k_i \Delta y_{t-i} + \sum_{k=1}^{N_x} \beta_k \Delta x_{t-k} + \varepsilon_t$$

where Δ is the first difference operator used in case of non-stationary variables in levels, y_t, y_{t-1},…,y_{t-p} are the observations in periods $t, t-1, \dots\ t-p$ is the number of lags of the AR term, k_i are the AR parameters, ε_t is the disturbance for period t, x is the exogenous variables matrix where each column is a time series, β_k are the exogenous variables parameters and N_x is the number of exogenous variables.

A natural extension of the AR(*p*) model is the standard ARDL model, which has the following general form:

$$\Delta y_t = \sum_{l=1}^{L_1} \gamma_l \Delta y_{t-l} + \sum_{l=0}^{L_2} \boldsymbol{\delta}_l' \Delta \boldsymbol{x}_{t-l} + u_t,$$

: where we allow for up to $L_1 \equiv p > 1$ lags of the dependent variable y and up to $L_2 \equiv q > 0$ lags of the explanatory variable $\boldsymbol{x}$.

3.2 Non-linear Time Series Model

Of course, due to the hidden non-linearities that are present in many time series variables, exploration of non-linear models is imperative for a sound econometric analysis. In this context, in order to allow for the possibility of nonlinearities we use an artificial neural network (ANN) of the following form. Suppose $\boldsymbol{z}_t = \left[y_t, \boldsymbol{x}_t'\right]'$. Then, our specification is:

$$\Delta y_t = \sum_{l=1}^{L_1} \gamma_l \Delta y_{t-l} + \sum_{l=1}^{L_2} \delta_l' \Delta x_{t-l} + \sum_{g=1}^{G} \lambda_g \varphi_g\left(\beta_g' z_{t-1}\right) + u_t \qquad (1)$$

Where G is the number of nodes, β_g and l_g are parameters, $g = 1,...,G$. When the functions $\varphi_g(\cdot)$ are known, we have an artificial neural network (ANN).

The main advantages of the proposed non-linear ANN-ARDL scheme are the following: (i) it allows the researcher to condition on a set of relevant economic and / or financial variables, and (ii) it allows the identification of relationships even if the true character between two variables is non-linear in nature.

In the ANN literature, it is common to use parametric activation functions of the form:

$$\varphi_g\left(w\right) = \varphi\left(w\right) = \frac{1}{1+\exp(-w)}, w \in \mathbb{R}, g = 1,...,G.$$

The assumption of common activation parametric functions cannot be adequate empirically and it would be best to allow for non-parametric alternatives in which both G and the functions φ_g are treated as unknown.

3.3 Bayesian non-linear Time Series Model

A well-known problem of traditional Neural Networks estimation is that of over-fitting. In order to robustly overcome this we make use of Bayesian Neural Network estimation. The main advantage of Bayesian estimation is the possibility of mixing different pieces of information (sample information, prior information, etc) in order to construct a model that accounts for the stochastic character of the variables.

In general, the semi—parametric analysis is inspired by Projection Pursuit Regression (Cook et al, 1993, Friedman, 1987 and Friedman and Stuetzle, 1981) as well as Gradient Boosting (GB), see Buhlmann and Yu (2003), Friedman (2001), Natekin and Knoll (2013).

Turning back to Neural Networks estimation, if we assume that the functions are twice continuously differentiable, as they are univariate, suppose $w_{g,(t)} \equiv \beta_g' z_{t-1}$ assuming β_g has been estimated and let $G = 1$. For simplicity we can re-order them so that $w_{g,(1)} \leq \ldots \leq w_{g,(T)}$. Then, it is well known that the assumption of twice continuously differentiable functions, implies a spline-like smoothness prior, Berger (2006), of the form:

$$\varphi_{g,(t)} - 2\varphi_{g,(t-1)} + \varphi_{g,(t-2)} \sim N\left(0, \sigma_\varphi^2 (w_{g,(t)} - w_{g,(t-1)})^2\right)$$

where σ_φ^2 controls the degree of smoothness, and $\varphi_{g,(t)} = \varphi_g\left(w_{g,(t)}\right)$. Given a Bayesian (posterior) estimate, say $\tilde{\varphi}_{g,(t)}$, we can compute an estimate$\hat{\varphi}_g\left(\beta_g' z_{t-1}\right)$ using interpolation.

Our baseline prior for the semi-parametric ARDL model is the following. Given the model:

$$\Delta y_t = \sum_{l=1}^{L_1} \gamma_l \Delta y_{t-l} + \sum_{l=1}^{L_2} \delta_l' \Delta x_{t-l} + \sum_{g=1}^{G} \lambda_g \varphi_g\left(\beta_g' z_{t-1}\right) + u_t,$$

denote the vector of parameters by:

$$\theta = \left[\gamma_1, \ldots, \gamma_{L_1}, \delta_1', \ldots, \delta_{L_2}', \beta_1', \ldots, \beta_G'\right]', \theta \in \mathbb{R}^D.$$

We use proper but relatively uninformative priors of the form:

$$\theta \sim N_D(\bar{\theta}, \bar{\Sigma}_\theta)$$

where:$\bar{\theta} = \mathbf{0}$ and $\bar{\Sigma}_\theta = \bar{h}\boldsymbol{I}_D$ for $\bar{h} = 10^3$. We need the notation in the prior sensitivity analysis section below.In the proposed model, for $G = 1$, the posterior of model (1) can be analyzed easily using:

i) A block Gibbs step for $(\gamma, \delta', \lambda')' | \beta_1, y, X$,

ii) A Metropolis-Hastings step for $\beta_1 | (\gamma, \delta', \lambda')', y, X$.

Here, $\gamma = [\gamma_1, \ldots, \gamma_{L_1}]'$and $\delta = \left[\delta_1', \ldots, \delta_{L_2}'\right]'$. To decide whether it is worthwhile to increase the number of nodes to $G = 2$we consider the model:

$$\Delta y_t = \sum_{l=1}^{L_1} \gamma_l \, \Delta y_{t-l} + \sum_{l=1}^{L_2} \delta_l' \, \Delta x_{t-l} + \lambda_1 \varphi_1\left(\beta_1' z_{t-1}\right) + \lambda_2 \varphi_2\left(\beta_2' z_{t-1}\right) + u_t.$$

We apply again a MCMC procedure using steps (i) and (ii) above but we update only $\lambda_2$and β_2, keeping all other parameters fixed at their posterior means from $G = 1$. We repeat the same procedure up to $G = G^*$ and we choose the model with the highest marginal likelihood. The algorithm resembles both Projection Pursuit Regression (Cook et al, 1993, Friedman, 1987 and Friedman and Stuetzle, 1981) as well as Gradient Boosting (GB), see Buhlmann and Yu (2003), Friedman (2001), Natekin and Knoll (2013).

3.4 Model comparison

We will compare forecasts of *h*>0 steps ahead generated by means of a multivariate autoregressive model with additional information (i.e. Google searches), with those generated by the aforementioned models, *without* the explanatory variable of Google Searches. The forecasting accuracy of the various models will be assessed based on two out-of-sample forecasts measures widely adopted in the relevant literature, namely the Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE).

A model's MAE for any given forecast horizon (*h*) is given by the following formula:

$$MAE = 1/h \sum_{t=0}^{h} |F_t - A_t|$$

where: h is the forecasting horizon of the model, F_t are the out-of-sample forecasted values of the model, and A_t are the actual values. The smaller the MAE values of a model the better its forecasting ability.

A model's MAPE is given by the expression:

$$MAPE = (100/h)\sum_{t=0}^{h}\left|\frac{F_t - A_t}{A_t}\right|$$

In order to econometrically test for the difference in forecasts accuracy between the various models employed we make use of the Diebold and Mariano (1995) test. The test is based on the following expression:

$$d_t = g(\varepsilon_{1,t}) - g(\varepsilon_{2,t})$$

where: g is a loss function, $\varepsilon_{i,t}$, $i = 1,2$ are the forecasts errors for the competing models 1 and 2, respectively. The null hypothesis for the test is equal predictive accuracy between the two models i.e. $H_0: d_t = 0, \forall t \in T$.

Lastly, both one-step-ahead and dynamic out-of-sample forecasting were evaluated with the results to be significantly consistent. In addition, in all forecast horizons the rolling window, the fixed window and the expanding window methods were utilized[1], with the results being robust in every case, since no statistically significant changes were present among the various methods.

4. Empirical Analysis

4.1 Variables

To examine the dependent variable of bank deposits, we put forward a specification that focuses on real economic activity, prices and interest rates. We choose the Industrial Production (IP) index as a proxy for economic activity, given its very close correlation with GDP and its high frequency. We also use the Spread (SPR) of the Greek deposit

[1] The sample size in all the forecasting exercises exceeded sixty (60) observations.

interest rate relative to the German interest rates downloaded from Bloomberg. To this core set, we added another variable that could potentially affect deposits in Greece. In particular, the dataset also contains Bloomerg's economic sentiment indicator (ESI). Also, we choose to use the Consumer Price Index (CPI), given the availability of a sufficiently long time series, but its use in the model did not yield significant results. As a result, is was excluded from the models.

Furthermore, we use the monthly number of Google searches for the term "Grexit", through its service called *Google Trends*. The term "Grexit" is an artificial term created specifically to compactly describe the exit of Greece from the European Monetary Union (EMU) and of course its potential consequences. As a result, the word "Grexit" characterizes and summarizes at the same time the adverse consequences of the Greek crisis, a fact that makes the term an ideal candidate to act as proxy for bad news for the whole economy and, thus, for its banking sector. It is worth noticing that alternative or roughly similar terms such as the phrase "Greek Crisis", and "Greek bank-runs", "Greek capital controls" showed very high correlation with our baseline term and were thus excluded from the analysis.

We focus on Google, because it is the most popular Internet search engine in the globe where Internet users around the world input several billion search queries per month onto the website. Also, on Google Trends, users can get time series data on the number of times a particular keyword search term is entered into the Google search engine. By inputting a keyword search term on Google insights, users can observe the actual flow of worldwide Internet searches for that particular keyword over time (Smith 2012).

Google trends are publicly available at https://www.google.com/trends/. Once the data have been accessed and saved, they can be downloaded. The data on Google Searches may vary according to the day and/or the IPs of download. In this work, we took the raw data coming from a single download since the various time series (over different days and IPs) were almost identical, with correlations that are equal to 0.99.

Finally, the monthly deposits for all Greek banks, that refer to deposits of domestic corporations and households, come from the Bank of Greece are used as the dependent variable, over the period 2009-2015. For a concise graphical representation of Greek bank deposits, see Figures 1 and 2.

Figure 1: Plot of Greek Bank Deposits in Levels

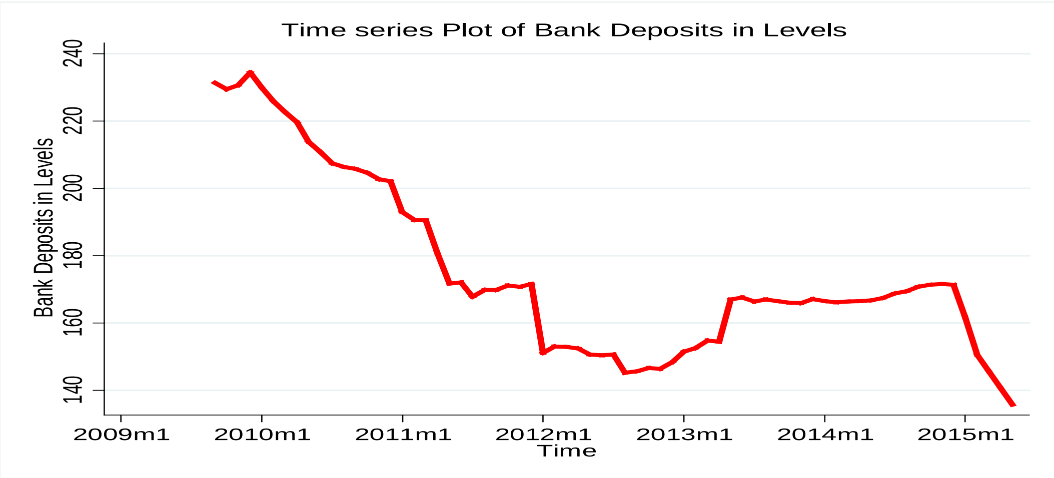
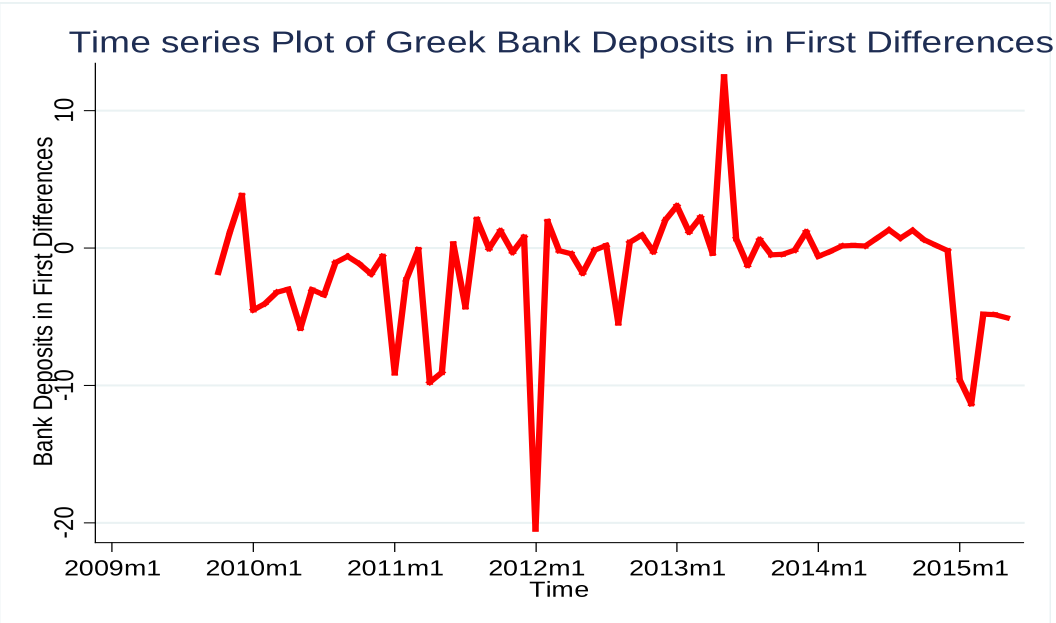


Figure 2: Plot of Greek Bank Deposits in First Differences



4.2 Model development

In this work, we will evaluate the out-of-sample performances of the various competing models relative to a benchmark model by comparing formally the MAPE with that of the benchmark, and we will test for equal forecast accuracy using the Diebold and Mariano (1995) test. Moreover, we will also plot the Root Mean Squared Forecast Error (RMSFE) of the best linear and non-linear models for visual inspection.

- **Step A: Degree of Integration**

At first, we check the stationarity properties using relevant unit root tests. We employ the Phillips-Perron unit root test with the null hypothesis (H0) that the time-series contain a unit root, i.e. non-stationary, and the Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) with the null hypothesis of time-series stationarity; if non-stationarity is present, we use first order differencing and the results are available upon request.

- **Step B: Order Selection**

The second step is order selection. Several models of the AR (*p*) family were applied and their performance was compared on the basis of their prediction accuracy with the respective AR with exogenous variables (ARX(*p*)) with no lags in the exogenous variables and for various horizons *h*. The exercise is repeated using both a set of variables which contain all the well- known indicators *excluding* Google Searches (named ARX(*p*)) and one set which contains all the known indicators, *including* Google Searches (named ARXGR(*p*)). Suitability of the models was determined using the MAE and the MAPE. Table 1 summarizes the models developed. Based on the results presented in Table 1, the ARXGR(1) model which corresponds to an AR(1) model with the exogenous set of variables being augmented with the *Google trends* of the word "Grexit", is the best model in terms of predictive ability using both the MAE and MAPE measures, respectively.

Table 1: Out of Sample Forecasts using AR(p) models for Horizon (h), h=1,…,6

Model Specification	h=1		h=2		h=3		h=4		h=5		h=6	
	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
AR(1)	135.798	1.000	72.968	0.519	51.862	0.357	41.290	0.276	37.555	0.236	34.479	0.206
ARX(1)	129.083	0.951	73.078	0.512	51.241	0.353	36.363	0.243	35.542	0.223	31.900	0.191
ARXGR(1)	**126.505**	**0.932***	**71.918**	**0.511***	**50.676**	**0.349***	**34.468**	**0.231***	**34.415**	**0.216***	**30.797**	**0.184***
AR(2)	135.798	1.000	72.968	0.519	51.514	0.355	40.969	0.274	37.008	0.231	33.815	0.202
ARX(2)	135.722	0.999	75.990	0.540	52.961	0.364	36.784	0.246	35.731	0.224	32.023	0.191
ARXGR(2)	133.217	0.981	74.840	0.532	52.251	0.359	34.936	0.234	34.615	0.217	30.929	0.185

* denotes rejection of H0 of equal forecasting accuracy ability, at a 5% level or higher, between the designated model and the competing ARX.

This process helps us determine the optimal number of lags *p**. And, in order to confirm our findings regarding the optimal order of the model *p**, we resort to the seminal paper by Diebold and Mariano (1995) to access the differences in the forecasting performance among the models with various lags. In this context, based on Table 1, both out-of-sample forecast measures produce statistically significant different forecasts from the baseline model.

- **Step C: Lag Selection**

Once the order *p** has been determined, according to Table 1, lags *(q)* in the exogenous variables will be allowed, transforming the ARX model to an ARDL model. Now, a similar investigation will take place where the ARDL models excluding and *including* the variable of *Google Searches* will take place, named ARDL (*p**,*q*) and ARDLGR(*p**,*q*), respectively, and for various horizons *h*. Again, suitability of the models is determined using the MAE and the MAPE criteria. Please note that a benchmark model was specified, where the lag length *p* was set equal to *p*= 1, as it produced lower MAE and MAPE. In all cases, we maintained *p*= 1 for the benchmark forecasts. Based on the results presented in Table 2, the ARDLGR(1,2) model which corresponds to an ARDL model with one (1) autoregressive term and two (2) lags for the exogenous variables is the best model in terms of predictive ability.

Table 2: Out of Sample Forecasts using ARDL(p,q) models for Horizon (h), h=1,…,6

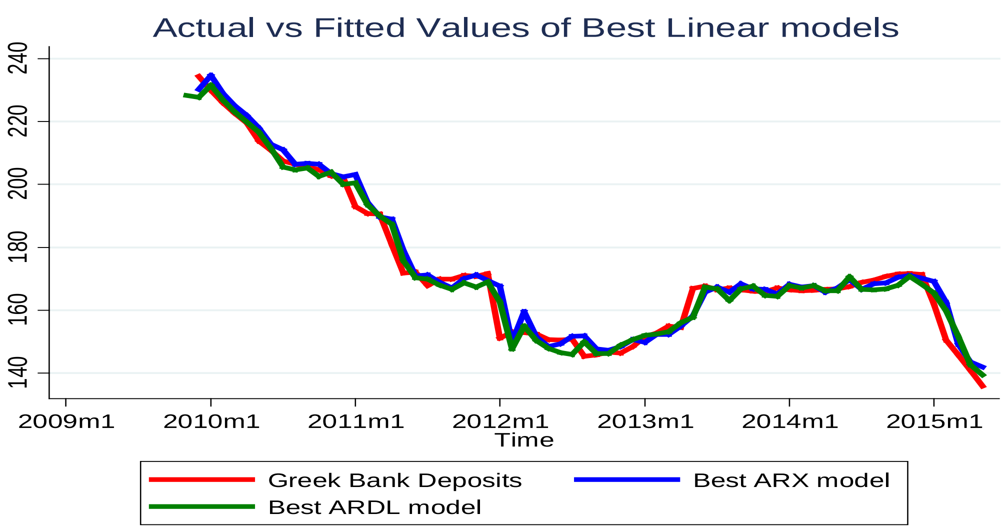
Model Specification	h=1		h=2		h=3		h=4		h=5		h=6	
	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE
ARDL(1,1)	7.441	0.053	7.152	0.051	7.784	0.055	8.921	0.062	8.771	0.059	7.454	0.051
ARDLGR(1,1)	7.168	0.052	6.932	0.050	7.271	0.052	8.280	0.058	8.357	0.058	7.067	0.048
ARDL(1,2)	5.490	0.041	5.967	0.043	6.834	0.049	8.297	0.058	8.498	0.058	7.195	0.049
ARDLGR(1,2)	**4.897**	**0.036***	**5.024**	**0.036***	**6.738**	**0.047***	**8.182**	**0.056***	**8.205**	**0.056***	**6.997**	**0.047***
ARDL(1,3)	5.510	0.041	5.981	0.043	7.330	0.052	8.967	0.062	9.112	0.062	7.669	0.052
ARDLGR(1,3)	4.590	0.038	6.067	0.042	7.152	0.051	8.751	0.061	8.976	0.062	7.542	0.052

* denotes rejection of H0 of equal forecasting accuracy ability, at a 5% level or higher, between the designated model and the competing ARDL.

Table 2 summarizes the models developed and the process helps us determine the optimal number of lags *q**. In order to confirm our findings regarding the optimal number of lags of the exogenous variables *p**, we will resort again to the seminal paper by Diebold and Mariano (1995).

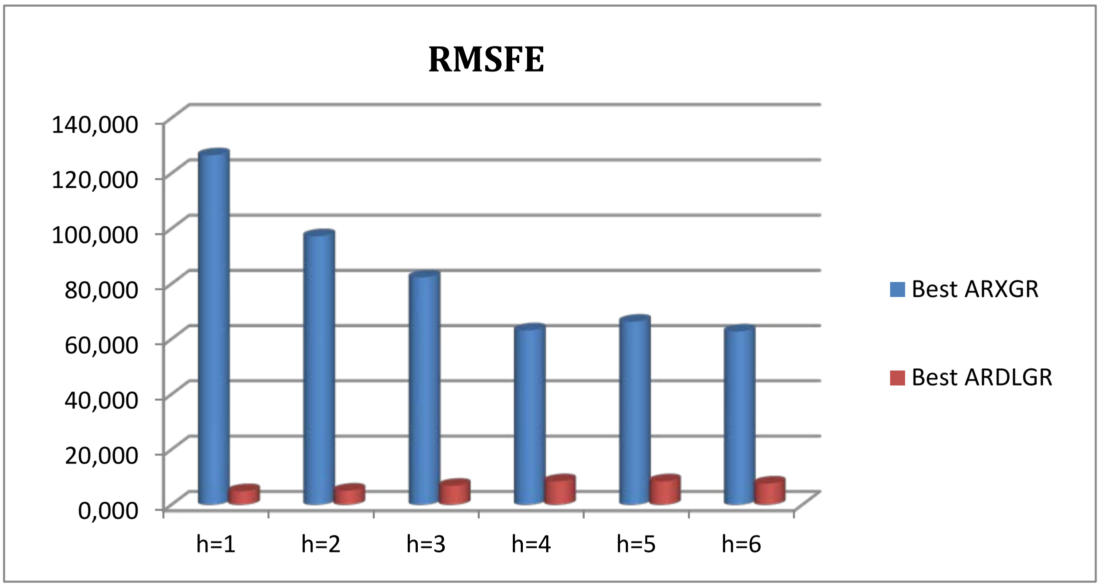
Next, for the best linear models employed so far, Figure 3, presents their fitting ability, while Figure 4, presents another popular measure, the so called Root Mean Squared Forecasting Error (RMSFE).

Figure - 1: Actual Vs Fitted Values for Best Linear Models



As we can be inferred from Figure 4, the best ARDLGR model employed, significantly outperforms the forecasting ability of the best ARXGR model, a fact which is also verified by the respective Diebold and Mariano (1995) tests.

Figure - 2: RMFSE for Best Linear models



- **Step D: Number of Nodes Selection**

Once the optimal order of the model p^* and the optimal lags of the exogenous variables q^* have been determined according to Tables 1 and 2, respectively, the different non-linear Neural Network models will be compared, *excluding* and *including* the variable of *Google Searches*, named ANN-ARDL (p^*, q^*, G) and ANN-ARDLGR (p^*, q, G), respectively, and for various horizons h. Here, the nodes G of the model are allowed to vary and the optimal number of nodes G* in the model will be determined using the MAE and the MAPE criteria. Table 3 summarizes the ANN-ARDL models developed. Again, the differences in the forecasting performance, is assessed by means of the Diebold and Mariano (1995) test.

Table 3: Out of Sample Forecasts using Neural-ARDL(p,q,G) models for Horizon (h),G=1,…, 3 and h=1,…,6												
Model Specification	**h=1**		**h=2**		**h=3**		**h=4**		**h=5**		**h=6**	
	MAE	**MAPE**	**MAE**	**MAPE**	**MAE**	**MAPE**	**MAE**	**MAPE**	**MAE**	**MAPE**	**MAE**	**MAPE**
NNARDL(1,2,1)	2.657	0.02	2.772	0.02	3.648	0.025	4.276	0.029	3.854	0.026	3.854	0.026
NNARDLGR(1,2,1)	2.463	0.018	2.284	0.017	2.667	0.02	4.176	0.028	3.855	0.025	3.855	0.025
NNARDL(1,2,2)	2.336	0.017	2.527	0.018	2.949	0.021	4.212	0.029	3.886	0.026	3.838	0.025
NNARDLGR(1,2,2)	**1.461**	**0.011***	**1.234**	**0.001***	**1.293**	**0.009***	**2.08**	**0.014***	**1.785**	**0.012***	**2.199**	**0.014***
NNARDL(1,2,3)	2.412	0.018	2.186	0.016	2.839	0.02	4.072	0.028	3.849	0.026	3.795	0.023
NNARDLGR(1,2,3)	2.189	0.016	1.84	0.013	2.779	0.019	3.197	0.025	3.318	0.023	3.495	0.025

* denotes rejection of H0 of equal forecasting accuracy ability, at a 5% level or higher, between the designated model and the competing NNARDL.

Based on the results presented in Table 3, ANN-ARDLGR (1,2,2) outperforms in terms of forecasting ability the rest of the Neural Network specifications. In this context, the optimal number of Neural terms employed were two (2). An interesting fact in our analysis is the increased accuracy in absolute terms of the Neural Network ARDL employed, since all MAPEs are less than 3%, irrespective of the forecasting horizon. This increased accuracy of the Neural ARDL models is also confirmed by the results of the Diebold and Mariano (1995) test, which are available upon request by the authors.

- **Step E: Bayesian Non-linear Model Selection**

Once: (a) the order, (b) the number of lags in the exogenous variables, and (c) the number of nodes has been established, a Bayesian version of the models will be estimated and their forecasting performances will be again compared formally on the basis of the Diebold and Mariano (1995) test.

Analytically, the main reason for using a Bayesian approach here is that it takes full account of the uncertainties related to model and parameter values. In contrast, most decision analyses based on least squares (L.S.) estimation involve fixing the values of parameters that may, in actuality, have an important bearing on the final outcome of the analysis and for which there is considerable uncertainty. One of the major benefits of the Bayesian approach is the ability to incorporate prior information. Also, MCMC, along with other numerical methods, makes computations tractable for virtually all parametric models. See Carlin and Lewis (2000), Robert (2001) and Wasserman (2004).

4.2 Bayesian Non-Linear Model Estimation

We use the rolling window methodology in a Bayesian non-linear ARDL set-up. The rolling window is a methodology that repeats estimations using sub-samples of the total data by shifting the start (and/) or end-points with a fixed window (Zivot and Wang 2006). The aforementioned process (i.e. rolling estimation window) was repeatedly used.

For the Bayesian non-linear ARDL (BANN-ARDL) models, we compare the forecasting performance of a BANN-ARDL model with lags of IP relative to a BBANN-ARDL without IP. Then, we compare a BANN-ARDL model with IP and SPR to the previous ARDL model which includes only IP, and so on. Finally, we augment the set of alternative explanatory variables by including the most crucial variable, i.e. the number of *Google Search* queries containing the word "Grexit", called BANN-ARDLGR.

Next, forecasts were generated based on the explanatory variables-augmented. In Table 4, columns contain the *p*-values of the tests of equal predictive ability between explanatory variable-augmented models and a benchmark nonlinear-ARDL model.

Please note that for the non-linear cases, a benchmark model was specified, where the lag length $p > 0$ was set equal to $p = 1$, and contained two nodes G=2 because it led to a model with the highest marginal likelihood.

In order to be able to compare the forecasting performance of the various models against the benchmark one, we used the Diebold and Mariano (1995) test, as mentioned earlier, to test H0, that each of the explanatory-variables augmented model has equal predictive ability as the benchmark models specified above.

Table 4: Out-of-sample forecasting: Rolling window

Semi-parametric ARDL	*p*-value	G
IP	0.071	2
IP, SPR	0.054	3
IP, SPR, ESI	0.031	3
IP, SPR, ESI, GS	0.008	4

Note 1: IP is Industrial Production Index, SPR is spreads relative to German interest rates, ESI is the economic sentiment indicator GS is Google searches.

Probably, the most important observation is that the out-of-sample forecasting performance of models accounting for Google Searches is superior to the models, which contain all the well-known explanatory variables *except for* Google Searches. In other words, we notice an increase in the forecasting performance of the BNNARDLGR model augmented with GS.

i) Expanding Window

Furthermore, we need to ensure the robustness of our results, in the sense that they do not depend critically on the assumptions or on the computations on which they were based. Hence, we will use an Expanding Window to estimate the forecasting performance of the models. As expected, the latter technique results in a slight increase in their performance, as reported in Table 5. However, the main conclusions reached previously remain, qualitatively, the same.

Table 5: Out-of-sample forecasting: Expanding Window		
Semi-parametric ARDL	***p*-value**	**G**
IP	0.024	2
IP, SPR	0.023	3
IP, SPR, ESI	0.019	3
IP, SPR, ESI, GS	0.002	4

Note 1: IP is Industrial Production Index, SPR is spreads relative to German interest rates, ESI is the economic sentiment indicator GS is Google searches.

ii) *Priors*

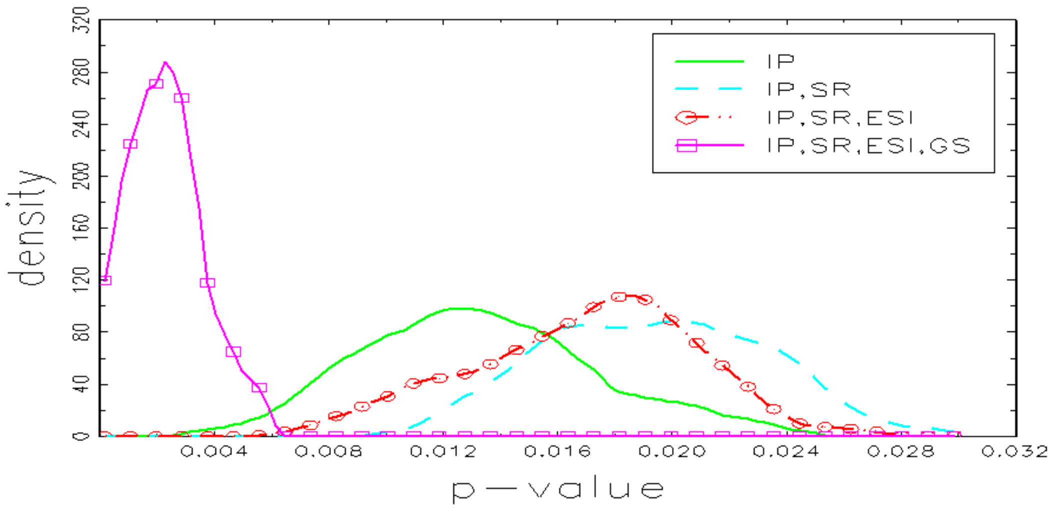
Our analysis is complemented with numerous empirically plausible priors selected from relevant classes of priors (Berger, 1985). We produced 10,000 alternative priors as follows:

$$\bar{q}^{(i)} \sim N_D\left(\mathbf{0}, 10^3\mathbf{I}_D\right), \ \ \log h^{(i)} \sim N(3,10), i = 1,\ldots,10000.$$

This allows for widely differing views about the coefficients. Depending on the priors the number of non-parametric functions, G, changes and therefore, D, the length of the parameter vector changes as well. For each new prior we employ our MCMC algorithm using 25,000 passes the first 5,000 of which are discarded to mitigate start up effects.

For each new prior we compute expanding window out-of-sample forecasts and we record their *p*-value. To examine sensitivity to prior information we report the results in Figure 3, for the Bayesian Non-linear ARDL models.

Figure - 3: Mariano Diebold test of equal forecasting accuracy between competing models (10,000 iterations)



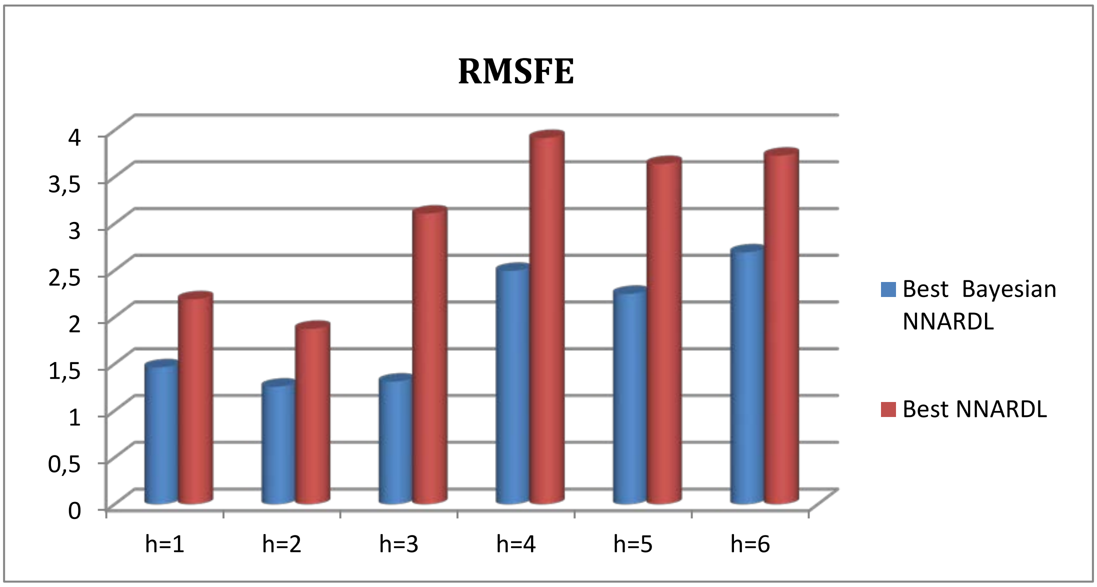
Our results were not found to be critically sensitive to the alternative priors used. This clearly implies that we can safely proceed based on these findings. For a detailed discussion on the theoretical foundations of prior selection see, for instance, Kass and Wasserman (1996).

To sum up, the improvement in out-of-sample forecasting performance models, which are extended to account for Google Searches, is shown to be more significant than in models that contain only the well-known indicators.

At this point, it should be noted that Bayesian non-linear ARDL models do not predict well when GS is omitted, as *p*-values are marginal. As a result, GS needs to be used to capture the hidden non-linearities.

The overall forecasting performance of the best BNNARDL employed as opposed to the best standard ANN-ARDL is assessed through the RMSFEs, which are presented in Figure 4, for the various horizons.

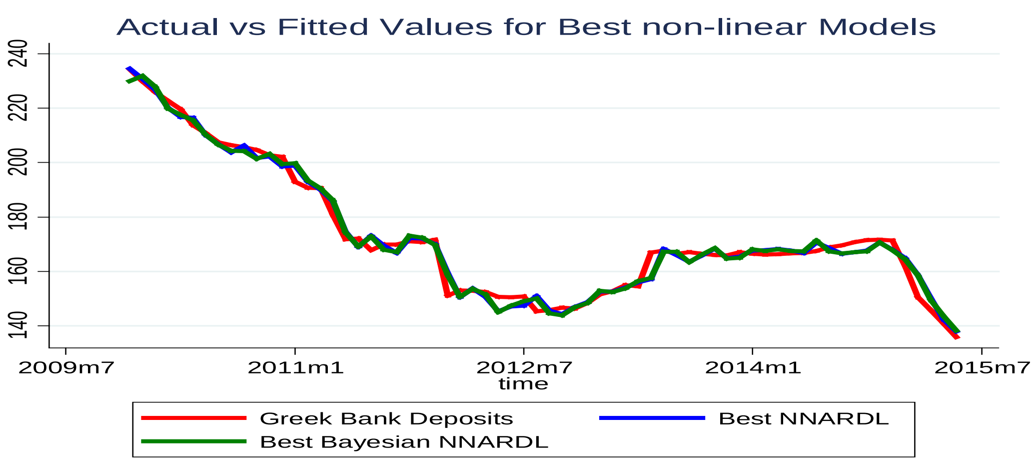
Figure - 4: RMSFE for Best Non-Linear Models



Based on Figure 4, the Bayesian ANN-ARDL outperforms the standard one in all forecasting horizons. This finding coincides with the results of the Diebold and Mariano (1995) test, which is available upon request by the authors.

Finally, the excellent fitting ability of the best non-linear models is presented in Figure 5.

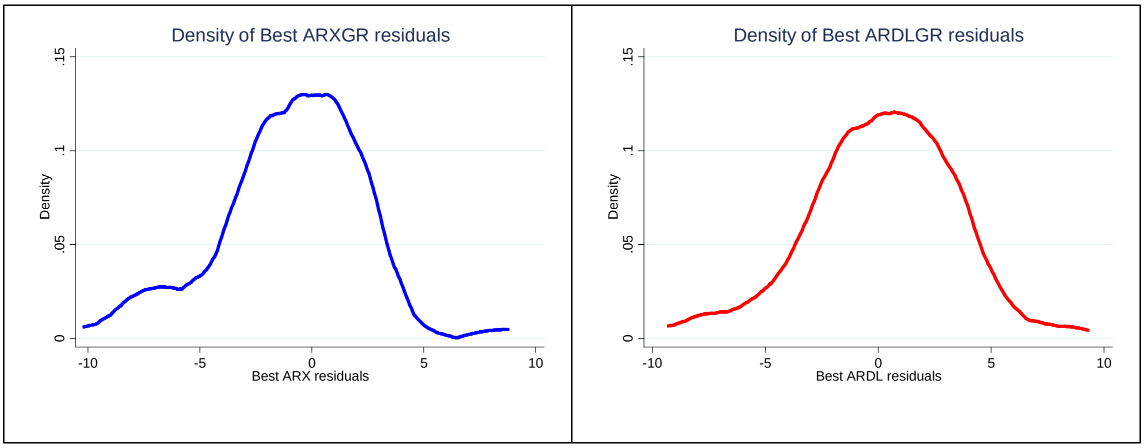
Figure – 5: Actual Vs Fitted Values Best Non-Linear Models



- **Testing for Gaussianity**

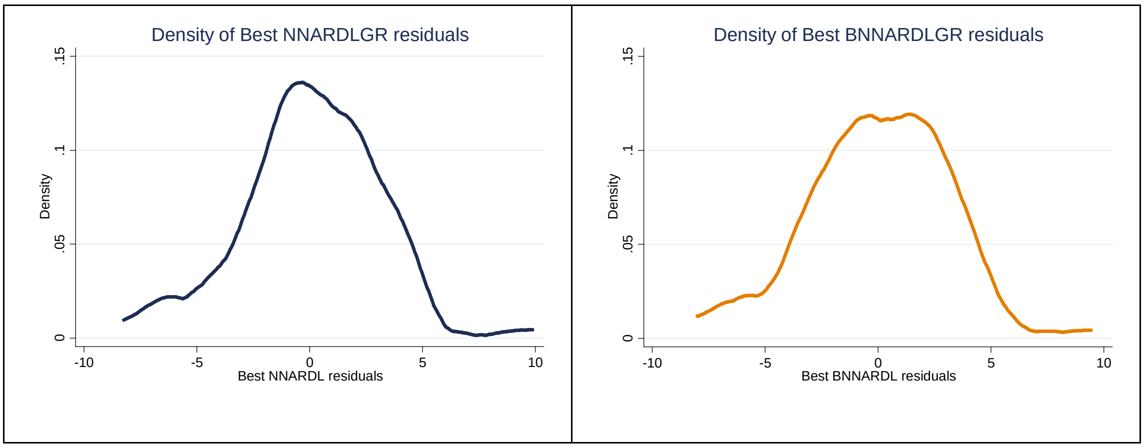
Now, all the model disturbances were studied, where errors had to be random. We also used the Ljung-Box statistic to check the closeness of the residuals to white noise and no evidence of *non*-Gaussianity was found in all models examined. For the sake of brevity, Figures 6 and 7 illustrate the density of the residuals for the best linear and non-linear models.

Figure - 6: Density for the residuals of the Best Linear Models



As we can infer easily from visual inspection, in all models the residuals follow the normal distribution, a fact which coincides with the results of the Ljung-Box test which are available upon request by the authors.

Figure - 7: Density for the residuals of the Best Non-Linear Models



5. Discussion

Based on the empirical analysis, it can be argued that, there is significant improvement in the out-of-sample forecasting performance of the various models, which are extended to account for *Google Searches*, when compared to the ones that contain only the well-known indicators. In fact, competing models tend to outperform the benchmark model at each forecast horizon, irrespective of the forecasting method used (one step ahead or dynamic forecasts) and the sample utilized, i.e. rolling window, fixed window, expanding window. In other words, the added predictive ability of the *Google searches* is a robust finding from an econometric perspective.

Also, we can see that the Diebold and Mariano (1995) test of equal forecasting accuracy always rejects the null hypothesis at horizons from 1 to 6 months ahead between the GS -augmented models and the respective benchmark models.

Based on our analysis, now a brief description of the economic mechanism can take place: Among the potential causes of the drawdown in Greek bank deposits are: (i) the deep recession, which shrunk the disposable income and forced domestic firms as well as individuals to resort to their existing pool of savings to finance daily consumption and operation; (ii) high risk and uncertainty about the stability of the Greek economy, which has caused an outflow of deposits abroad.

At the next step, it can now be safely argued that *Google Searches* are triggered on the arrival of information for a country's economic fundamentals or even the social or political situation. Put differently, this news is assumed to express information about the country's economic, financial and social future. In this context, in case the information extracted by these searches is "negative" in the sense that it affects negatively the future expectations of public opinion regarding the economy's fundamentals, then it is possible that these searches could be used in order to proxy "herding" behavior of individuals.

If the information extracted by the use of the term "Grexit" implies that the country is close to severe capital controls (or even default), this would lead immediately most of the potential bank account holders to the (herding) behavior of withdrawing money in attempt to take money out of the banking system, where a part of them was invested in assets outside the so-called M3 (aggregate) or simply held in the form of "under-the-mattress" cash.

Also, another reason for this situation was that there was a rumor that the government was planning to cross-check deposits with domestic commercial banks and income declared for tax purposes. Of course, this plan led several firms and individuals to send money abroad. What happened next? There came the downgrades of Greece's credit, influencing seriously its banks' ratings, which -in turn- limited their access to external markets. As a result, the liquidity position of Greek banks came under extreme pressure.

Now, more bank deposit holders will withdraw their money and, indeed, as the crisis deepens day–by–day, and the levels of risk increase dramatically, then an increasing number of bank account holders mimic this behavior and react in the same way (herding behavior). Given that deposits have always been a significant channel of banks' ability to "transfer" loans to the economy we can see that this mechanism harms the real economy, as well. Clearly, if we study a case where a negative economic environment exists, such as the ones just described, then we could safely argue that the increased amount of Google search is linked to decreasing bank deposits, i.e. a "bank-run".

Our approach of taking into consideration the number of Google searches containing the keyword "Grexit" is based on the implicit assumption that it summarizes, in a nutshell, the public sentiment worldwide about the situation of the Greek economy. It also has the advantage of being extremely fast, accurate and costless.

6. Concluding Remark

The evolution of the banking sector in Greece has always reflected wider macroeconomic developments in the Greek economy. In this framework, a prominent victim of the Greek economy's very deep recession was the country's banking system which faced a dramatic outflow of deposits, over the period 2009-2015.

In this context, we shed light on the possibility of forecasting bank deposits, based on keyword Google searches on a daily basis. Our approach of taking into consideration the number of Google searches containing the keyword "Grexit" is based on the implicit assumption that it summarizes, in a nutshell, the public sentiment worldwide about the situation of the Greek economy. It also has the advantage of being extremely fast, accurate and costless and could thus serve as an *ex-ante* "proxy" for bank holder's potential herding behavior.

In this framework, we showed that extending standard autoregressive models with the keyword "Grexit" lead to improvement in the forecasting ability of the Bank Deposits in the country, compared to models that contain just some standard explanatory variables. Our findings are checked for robustness in the data samples, in the activation function used and the prior distributions employed, and are found to be econometrically robust and economically sound, confirming the crucial role of *Google Searches* for predicting bank deposits.

The use of *Google Searches* for improving the predictive ability of other fundamental measures would be of great interest and constitutes a good example for future investigation. Also, the keyword, the country, the model employed as well as the period examined could be further investigated and expanded.

References

Antenucci, D., Cafarella, M., Levenstein, M.C., Re, C., Shapiro, M.D. (2014), Using social media to measure labor market flows, *NBER Working Papers*, 20010.

Artola, C., Galan, E. (2012), Tracking the future on the web: construction of leading indicators using Internet searches, *Documentos Ocasionales (1203*), Bank of Spain.

Barajas, A., Steiner, R. (2000), Depositor behaviour and market discipline in Colombia, Working Paper WP/00/214, *International Monetary Fund.*

Berger, J. (2006), The case for objective Bayesian analysis. Bayesian Analysis, 1, 385–402.

Berry, T. D. and Howe, K.M. (1994), Public Information Arrival, *The Journal of Finance*, vol. 49, pp. 1331-1346.

Bessembinder, H., Chan, K. and Seguin, P. J. (1996), An empirical examination of information, di_erences of opinion, and trading activity, *Journal of Financial Economics*, vol. 40, pp. 105-134

Billett, M.T., Garfinkel, J.A., O'Neal, E.S., 1998. The cost of market versus regulatory dis-cipline in banking, *Journal of Financial Economics,* 48, 333-358.

Brewer III, E., Mondschean, T.H. (1994), An empirical test of the incentive effects of de-posit insurance: The case of junk bonds at savings and loan associations, *Journal of Money, Credit and Banking* 26, 146-164.

Calomiris, C.W., Powell, A. (2001), Can emerging market bank regulators establish credi-ble discipline? The case of Argentina, 1992-1999, in: Mishkin, F.S., (Ed.), *Prudential Supervision: What Works and What Doesn't.* The University of Chicago Press, Chicago, pp. 147-196.

Cargill, T.F. (1989), CAMEL ratings and the CD market, *Journal of Financial Services Research* 3, 347-358.

Carlin, B.P. and Louis, T.A. (2000), *Bayes and empirical Bayes methods for data analysis*, second ed. Chapman & Hall, London.

Carroll, Christopher D. (1992), The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence, *Brookings Papers on Economic Activity*, 1992(2), 61–156.

Chari V.V., and Jagannathan R., (July 1988), "Banking Panics, Information, and Rational Expectations Equilibrium," Journal of Finance, 43, pp. 749–761.

Choi, H., Varian, H. (2012), Predicting the present with Google Trends, *Economic Record*, 88, 2–9.

Clark, P. K. (1973), A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices, *Econometrica*, vol. 41, pp. 135-155.

Cook. D., Buja, A. and Cabrera, J. (1993), Projection pursuit indexes based on orthonormal function expansions, *Journal of Computational and Graphical Statistics***2**(3): 225–250.

Da, Z., Engelberg, J., Pengjie, G. (2011), In search of attention, *Journal of Financ*e, 5, 1461–1499.

Deaton, Angus S. (1991), Saving and Liquidity Constraints, *Econometrica*, 59, 1221–1248.

Decker F. and Goodhart C.A.E., (2018), "Credit mechanics: A precursor to the current money supply debate", Vox Cepr Policy Portal.

Dergiades, T., Milas, C. and Panagiotidis, T (2015), Tweets, Google Trends, and Sovereign Spreads in the GIIPS, *Oxford Economic Papers,* Volume 67, Issue 2, pp. 406–432.

Diamond Douglas W., (Spring 2007), "Banks and Liquidity Creation: A Simple Exposition of the Diamond-Dybvig Model", Economic Quarterly Volume 93, Number 2, Spring 2007, pp.189–200.

Diamond Douglas W., and Philip H. Dybvig, (1983), “Bank Runs, Deposit Insurance, and Liquidity”, Journal of Political Economy 91 (5), pp. 401–19.

Diebold, F. X., Mariano, R. S. (1995), Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 253–263.

Edelman, B. (2012), Using Internet data for economic research, *The Journal of Economic Perspectives*, 26(2), 189–206.

Ederington, L. H. and Lee, B.H. (1993), How Markets Process Information - News Releases and Volatility, *The Journal of Finance*, vol. 48, pp. 1161-1191.

Eeckhoudt, L. and Godfroid, P. (2000), Risk aversion and the value of information, *The Journal of Economic Education*, vol. 31, pp. 382-388.

Einav, L., Levin, J. (2013), Economics in the age of big data, *Science*, 346 (6210), 715–721.

Ellis, D.M., Flannery, M.J. (1992), Does the debt market assess large banks’ risk? Time se-ries evidence from money centre CDs, *Journal of Monetary Economics* 30, 481-502.

Engineer M., (1989), “Bank Runs and the Suspension of Deposit Convertibility,” Journal of Monetary Economics, 24, pp. 443–454.

Epps, T. W. and Epps, M. L. (1973), The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture of Distributions Hypothesis, *Econometrica*, vol. 44, pp. 305-321.

Ettredge, M., Gerdes, J., Karuga, G. (2005), Using web-based search data to predict macroeconomic statistics, *Communications of the ACM*, 48(11), 87–92.

Fama, E. F., Fisher, L., Jensen, M.C., Roll, R. (1969), The Adjustment of Stock Prices to New Information, *International Economic Review*, vol. 10, pp. 1-21.

Forssbæck, J. (2011), Ownership structure, market discipline, and banks’ risk-taking incen-tives under deposit insurance, *Journal of Banking and Finance*, 35, 2666−2678.

Francesco D'Amuri, Juri Marcucci (2017), The predictive power of Google searches in forecasting US Unemployment, *International Journal of Forecasting* 33, 801–816.

Freeman S., (1988): "Banking as the Provision of Liquidity," Journal of Business, 61, pp. 45–64.

Freixas, X. and Kihlstrom, R. "Risk aversion and information demand", in *Bayesian Models of Economic Theory* (M. Boyer and R. Kihlstrom, Eds.), pp. 93-104, Elsevier, Amsterdam, 1984.

French, K. R. and Roll, R. (1986), Stock Return Variances - The Arrival of Information and the Reaction of Traders, *Journal of Financial Economics*, vol. 17, pp. 5-26.

Friedman J. (2001), Greedy boosting approximation: a gradient boosting machine, *Annals of Statistics*,29, 1189–1232.

Friedman, J. H. (1987), Exploratory projection pursuit, *J. Am. Statist. Assoc.* 82, 249-266.

Friedman, J. H., and Stuetzle, W. (1981), Projection pursuit regression.*J. Am. Statist. Assoc.* 76, 817-823.

Friedman, M. (1957), *A Theory of the Consumption Function*, Princeton: Princeton University Press.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009), Detecting influenza epidemics using search engine query data, *Nature*, 457, 1012–1014.

Goldberg, L.G., Hudgins, S.C. (2002), Depositor discipline and changing strategies for regulating thrift institutions, *Journal of Financial Economics* 63, 263-274.

Goodhart, C. A. E., (2017), "The Determination of the Money Supply: Flexibility Versus Control", The Manchester School 85, pp. 33-56.

Goodhart, C. A. E., (September 2017), "A Central Bank's Optimal Balance Sheet Size?", London School of Economics & Political Science (LSE) - Financial Markets Group.

H. Gibson, E. Tsakalotos (1992), *Economic Integration and Financial Liberalization: Prospects for Southern Europe*, MacMillan in Association with St Antony's College, Oxford.

Hadad, M.D., Agusman, A., Monroe, G.S., Gasbarro, D. (2011), Market discipline, finan-cial crisis and regulatory changes: Evidence from Indonesian banks, *Journal of Banking and Finance* 35, 1552-1562.

Hannan, T.H., Hanweck, G.A., (1988).Bank insolvency risk and the market for large certifi-cates of deposit, *Journal of Money, Credit and Banking* 20, 203−211.

Hartmann P., Huang H., Schoenmaker D., (2018), "The Changing Fortunes of Central Banking", Cambridge University Press.

Haubrich J.G. and King R.G., (December 1990), "Banking and Insurance," Journal of Monetary Economics, 26, pp. 361–386.

Herald Finger and HeikoHesse (2009), Lebanon-Determinants of commercial banks Deposits in a Regional Financial Center, *IMF Working paper* , WP/09/195.

Hess, K. and Feng, G. (2007), Is there market discipline for New Zealand non-bank financial institutions?, *Journal of International Financial Markets, Institutions and Money* 17, 326-340.

Hiermeyer M., (2018): "A Tautologies-Founded IS-LM Model", Munich Personal RePEc Archive (MPRA).

Hiermeyer M., (February 2019), "An Improved IS-LM Model To Explain Quantitative Easing", Munich Personal RePEc Archive (MPRA).

Hosono, K. (2005), Market discipline to banks in Indonesia, the Republic of Korea, Malaysia and Thailand, *Asian Development Bank Institute Conference*, Corporate Governance of Banks in Asia, January 20-21.

Iftekhar Hasan, Krzysztof Jackowicz, Oskar Kowalewski and Lukasz Kozlowski (2013), Market discipline during crisis: Evidence from bank depositors in transition countries, *Journal of Banking & Finance,* Vol. 37, pp. 5436–5451

Ioannidou, V.P., Penas, M. F. (2010), Deposit insurance and bank risk-taking: Evidence from internal loan rating, *Journal of Financial Intermediation*, 19, 95-115.

Jacklin C.J., (1987) "Demand Deposits, Trading Restrictions, and Risk Sharing," in Contractual Arrangements for Intertemporal Trade, Edward C. Prescott and Neil Wallace (Eds.), Minnesota Studies in Macroeconomics, Vol. 1. University of Minnesota Press.

Jacklin C.J., (1993), "Market Rate Versus Fixed Rate Demand Deposits," Journal of Monetary Economics, 32, pp. 237–258.

Jacklin C.J. and Bhattacharya S., (1988), "Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications," Journal of Political Economy, 96, pp. 568–592.

Jackowicz, K. (2004). Dyscyplinarynkowa w bankowości.Rodzaje i możliwościzastoso-wania [Market discipline in banking.Types and applications], WydawnictwoWyższejSzkołyPrzedsiębiorczości i Zarządzaniaim. L. Koźmińskiego, Warszawa.

Jagtiani, J., Lemieux, C. (2000), *Stumbling Blocks to Increasing Market Discipline in the Banking Sector: A Note on Bond Pricing and Funding Strategy Prior to Failure*, Federal Reserve Bank of Chicago, Chicago.

Jordan, J.S. (2000), Depositor discipline at failing banks, *New England Economic Review*, 15-28.

Kass, R.E., Wasserman, L.A. (1996), The selection of prior distributions by formal rules, *J. Am. Stat. Assoc.* 91, 1343–1370.

Kolapo, F. T., Oke, M. O., & Olaniyan, T. O. (2018). Deposit Money Banks' Credit to Private-Public Sectors and Economic Development Nexus: Nigeria in Focus. International Journal of Academic Research in Business and Social Sciences, 8(5), 215–233.

Kouassi, T., Distinguin, I., Tarazi, A. (2011), *Bank deposit insurance, moral hazard and market discipline: Evidence from Central and Eastern Europe.* http://dx.doi.org/10.2139/ssrn.1830265.

Kraft, E., Galac, T. (2007), Deposit interest rates, asset risk and bank failure in Croatia, *Journal of Financial Stability* 2, 312-336.

Kutner, G.W. (1992), CD Market response to continental Illinois' liquidity crisis and its subsequent bailout, *Mid-Atlantic Journal of Business*, 28, 133-146.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992), Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root, *Journal of Econometrics*, vol. 54, pp. 159-178.

Levy-Yeyati, E., Peria, M.S.M., Schmukler, S.L. (2004), *Market discipline under systemic risk: Evidence from bank runs in emerging economies*, Working Paper 3440, World Bank Policy Research.

Panopoulou, M. (2005), *Technological Change and Corporate Strategy in the Greek Banking Industry*, KEPE, Athens.

Maechler, A.M. and McDill, K.M. (2006), Dynamic depositor discipline in US banks, *Journal of Banking and Finance* 30, 1871-1898.

Mcculloch H. and Min-Teh Yu, (1998), "Government Deposit Insurance and the Diamond-Dybvig Model", The Geneva Papers on Risk and Insurance Theory, pp. 139–149.

McLaren, N., and Shanbhorge, R. (2011), Using Internet data as economic indicators, *Bank of England Quarterly Bulletin*, Second Quarter.

Mitchel, M. L. and Mulherin, J. H. (1994), The Impact of Public Information on the Stock Market, *The Journal of Finance*, vol. 49, pp. 923-950.

Modigliani, Franco, and Richard H. Brumberg (1954), Utility analysis and the consumption function: an interpretation of cross-section data, in Kenneth K. Kurihara, ed., *Post- Keynesian Economics*, New Brunswick, NJ. Rutgers University Press, pp. 388–436.

Mondschean, T.S. and Opiela, T.P. (1999), Bank time deposit rates and market discipline in Poland: The impact of state ownership and deposit insurance reform, *Journal of Financial Services Research* 15, 179-196.

N. Kamberoglou, E. Liapis, G. Simigiannis, P. Tzamourani (2004), *Cost Efficiency in Greek Banking*, Working Paper Bank of Greece.

Natekin, A. and Knoll, A. (2011), Gradient boosting machines, a tutorial, *Front. Neurorobot.*,7, 21.

Oliveira, R.F., Schiozer, R.F., Barros, L.A.B. (2011), *Financial crisis and cross-border too big to fail perception*, Midwest Finance Association 2012 Annual Meetings Paper, http://dx.doi.org/10.2139/ssrn.1787661.

Önder, Z., Özyildirim, S. (2008), Market Reaction to risky banks: Did generous deposit guarantee change it? *World Developments* 36, 1415-1435.

Opiela, T. (2004), Was there an implicit full guarantee at financial institutions in Thailand? Evidence of risk pricing by depositors, *Journal of Comparative Economics* 32, 519-541.

Buhlmann P. and B. Yu. (2003), Boosting with the *l*2loss: regression and classification, *J. Amer. Statist. Assoc.*, 98:324–339.

Park, S. and Peristiani S. (1998),.Market discipline by thrift depositors, *Journal of Money, Credit, and Banking* 30, 347-364.

Peria, M.S.M. and Schmukler, S.L. (2001), Do depositors punish banks for bad behavior? Mar-ket discipline, deposit insurance, and banking crisis, *The Journal of Finance* 56, 1029-1051.

Postlewaite A., and Vives X., (1987), "Bank Runs as an Equilibrium Phenomenon," Journal of Political Economy, 95, pp. 485–491.

Prean, N. and Stix, H. (2011), The effect of raising deposit insurance coverage in times of fi-nancial crisis – Evidence from Croatia microdata, *Economic Systems* 35, 496−511.

Preis, T., Moat, H. S. and Stanley, E. (2013), Quantifying trading behavior in financial markets using Google Trends, *Nature Scientific Reports*, Article 1684.

Prescott E., (First Quarter 2010), "Introduction to the Special Issue on the Diamond-Dybvig Model", Economic Quarterly, Federal Reserve Bank of Richmond, pp. 1-9.

Qi, J., (1994), "Bank Liquidity and Stability in an Overlapping Generations Model," Review of Financial Studies, 7(2), pp. 389–417.

Robert, C.P. (2001), *The Bayesian Choice*, second ed. Springer-Verlag, New York, NY.

Russell S., (January-February 1993), "The Government's Role in Deposit Insurance," Federal Reserve Bank of St. Louis Review, 75, pp. 3–9.

Ryan, P. and Taer, R. J. (2004), Are Economically Signicant Stock Returns and Trading Volumes Driven by Firm-specific News Releases?,*Journal of Business Finance & Accounting*, vol. 31, pp. 49-82

Shimizu, K. (2009), Is the information produced in the stock market useful for depositors? *Finance Research Letters* 6, 34-39.

Smith (2012), Economics of Money Banking and Financial Markets, Seventh Edition, Frederic S. Mishkin, Columbia University.

SudinHaron and Wan Nursofiza Wan Azmi (2006), Deposit Determinants of Commercial Banks in Malaysia, *Finance India*, 2006, Vol. XX, No. 2.

Tauchen, G. E. and Pitts, M. (1983), The Price Variability-Volume Relationship on Speculative Markets, *Econometrica*, vol. 51, pp. 485-505

Tett G. (2013), 'Markets insight: wake up to the Twitter effect on markets', Financial Times, 18 April.

Thompson, R. B., Olsen, C., Dietrich, R. J. (1987), Attributes of News about Firms: An Analysis of Firm-Specific News Reported in the Wall Street Journal Index, *Journal of Accounting Research*, vol. 25, pp. 245-274

Troy, G., Perera, D. and Sunner, D. (2012), Electronic indicators of economic activity, *Australian Reserve Bank,Economic Bulletin*, June 1–12.

Tuhkuri, J. (2016), Forecasting unemployment with Google searches, *ETLA Working Paper* No. 35.

Uchida, H, Satake, M. (2009), Market discipline and bank efficiency, *Journal of International Financial Markets, Institutions and Money* 19, 792-802.

Verrecchia, R. E. (1980), Consensus Beliefs, Information Acquisition, and Market Information Efficiency, *The American Economic Review*, vol. 70, pp. 874-884

Verrecchia, R. E. (1982), Information Acquisition in a Noisy Rational Expectations Economy, *Econometrica*, vol. 50, pp. 1415-1430

Vlastakis, Nikolaos and Markellos, Raphael (2012), Information demand and stock market volatility, *Journal of Banking & Finance*, 36, issue 6, p. 1808-1821.

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

Wasserman, L. (2004), *All of Statistics: A Concise Course in Statistical Inference*, Springer-Verlag, New York, NY.

Werner R., (2014), “How do banks create money, and why can other firms not do the same? An explanation for the coexistence of lending and deposit-taking”, International Review of Financial Analysis, 36(C) pp. 71-77.

Westbrook D. A., (Winter 2017), “The Paradigm Sways: Macroeconomics Turns to History (reviewing three titles) ” University at Buffalo School of Law.

Willinger, M. (1989), Risk Aversion and the Value of Information, *The Journal of Risk and Insurance*, vol. 56, pp. 104-112

Zivot, E. and J. Wang (2006), *Modeling Financial Time Series with S-Plus*, 2nd edition, Springer, New York.