Increasing knowledge base for nowcasting GDP by quantifying the sentiment about the state of economy

Nikolaos Kourentzes[a,*], Fotios Petropoulos[a]

[a]*Lancaster University Management School*
Department of Management Science, Lancaster, LA1 4YX, UK

Abstract

Predicting the present state of the economy can be challenging. Recently, econometric models that are capable of using multiple sources of hard and soft information, published at various frequencies and dates, to produce nowcasts of policy relevant variables have been proposed. Following the progress in the modelling aspects of nowcasting, the literature has started exploring new useful inputs. The 'Big data' and various Internet sources can be used in order to further enhance nowcasting models, by enriching them with new types of information, as well as being more up to date. This research proposes a methodology to create useful inputs by mining news articles, capturing the current sentiment in the media about the state of the economy. News articles are nowadays easy to access and process, as most news outlets provide online versions of them. Furthermore, the high frequency of publication and the representation of the current economic discourse are invaluable for nowcasts. The case study of the Greek economy is used and rolling nowcasts are produced for the period of 2008 to 2013. Empirical evidence suggests that the inclusion of such information improves the accuracy of nowcasts.

Keywords: Nowcasting, GDP, News, Text mining, Lasso, Greek economy

1. Introduction

In economic policy making we are interested in knowing the current state of the economy. Unlike most forecasting applications, knowledge of

*Correspondance: N Kourentzes, Department of Management Science, Lancaster University Management School, Lancaster, Lancashire, LA1 4YX, UK. Tel.: +44-1524-592911
Email address: n.kourentzes@lancaster.ac.uk (Nikolaos Kourentzes)

the present is a real challenge. Several useful economic indicators, such as GDP, are typically collected at low frequency and with significant time lags. In these cases, nowcasting becomes relevant. The objective of nowcasting is to provide early estimates of important economic variables, based on relevant information that becomes available at a higher frequency and with shorter lags. Furthermore, this information is released in an asynchronous manner causing the 'ragged' edge problem.

Until recently, the process of updating economic variables in the face of new information was based on judgemental adjustments by experts or simple statistical models. The work by Evans (2005) and Giannone et al. (2008) introduced rigorous statistical models for nowcasting. These models focus on capturing the dynamics of the GDP time series together with higher frequency data, typically monthly. At the same time they address the problem of asynchronous publication of the input monthly time series. Since then the area has attracted a lot of interest in improving the modelling aspect further (for example, see Marcellino and Schumacher, 2010; Banbura et al., 2010; Kuzin et al., 2011).

A large number of variables have been used in the literature for econometric nowcast models of GDP. These inputs are typically divided in hard and soft data, the latter being surveys or expert opinions. Examples of lists of useful variables can be found in Evans (2005); Giannone et al. (2008); Banbura et al. (2010). These are published at various frequencies ranging from quarterly to daily, with the vast majority being monthly. Furthermore, these variables may have significant publication delay, often exceeding 30 days.

Recently, with the advent of 'big data' new potentially useful variables for better nowcasting have been considered. For example Choi and Varian (2012) and Scott and Varian (2013) explore the use of aggregated statistics from internet searches as inputs for nowcasting. However, an area that is under-explored is the inclusion of unstructured information that exists online, from non-official sources. An example of this type of information is online articles from news agents that discuss or report on the current state of the economy.

The aim of this work is to contribute to the inputs of nowcasting models. We propose a framework to assimilate information from news outlets into GDP nowcasts, complimenting the soft data inputs currently used in nowcasting models. We use the case study of Greece and provide rolling nowcasts from January 2008 until January 2013, demonstrating that the new information can benefit the accuracy of the nowcasts, during periods of increased

uncertainty for the state of the economy. We collected news items from the leading economic newspaper in Greece, in terms of circulation,[1] which were consequently mined to extract variables capturing the sentiments of the agents in the economy. These variables are updated as news articles are published, asynchronously, and used to nowcast the GDP growth. We find that this new type of data is useful in improving nowcasts, as it captures the hard to quantify sentiments in the economy in a systematic way. This paper does not offer a complete nowcast model, rather provides a proof of concept that the new type of inputs proposed is useful in improving nowcasts and it is envisioned that this should be used to compliment existing nowcasting models.

The rest of the paper is organised as follows: Section 2 presents the proposed methodology to quantify and model the discourse about the state of the economy from news outlets; section 3 outlines the emprical evaluation experimental design and discusses the results, followed by conclusions.

2. Methodology

The proposed methodology contains two different elements, the variable creation, and their use for producing nowcasts. These are discussed separately in sections 2.1 and 2.2 respectively.

2.1. Variable creation from media news items

The underlying idea behind the creation of the news variables is to capture the current discourse about the state of the economy from news outlets. The approach followed here is loosely based on the work by Altheide (1997) and Altheide and Schneider (2012), who explore how media discourse reflect and affect the sentiments in societies. The underlying hypothesis is that the closer particular words appear in the discourse, the more people associate them together. In practice, this hypothesis suggests that if people often use sets of words in the phrasing used to express themselves, these sets can reveal information about their sentiments in the discourse. Focusing on nowcasting economic variables, we want to identify positive or negative sentiments about the state of the economy.

To perform this analysis we first need to produce a text corpus that will contain $\Theta \geq 2$ words of interest. Then the location, $l_{\theta,i}$, of each word

[1] Circulation statistics are available at `http://www.argoscom.gr/`.

within a news article is tracked, where $\theta = 1, \ldots, \Theta$ and i is the number of occurrences of that word in the text. A threshold ϕ is defined, which stores the maximum distance between words that will be evaluated. A small ϕ will demand very strong association between the words in the discourse. The value of ϕ is also related to the syntax rules of the language of the texts analysed. Obviously, languages that use a lot of articles, pronouns, etc. and adjuct part of sentences for their syntax, will require higher ϕ. For the English language $\phi \leq 10$ is a reasonable threshold (Altheide, 1997).

From these words $p = \binom{\Theta}{2}$ pairs are formed. Note that the order of the words is not important in this analysis and therefore we are not interested in the 2-permutations of θ. For each pair, we calculate the distance between two words:

$$d_{\theta,\eta,i,j} = |l_{\theta,i} - l_{\eta,j}|, \tag{1}$$

where θ and η define the word used and $\theta \neq \eta$. Since there may be several occurrences of the same words in a text, the counters i and j iterate across all of them. For each pair of words a vector $\mathbf{O}_{\theta,\eta} = o_{i,j}$ is constructed that compares their respective calculated distances with threshold ϕ:

$$o_{i,j} = \begin{cases} 1 & \text{if } d_{\theta,\eta,i,j} \leq \phi \\ 0 & \text{if } d_{\theta,\eta,i,j} > \phi. \end{cases} \tag{2}$$

The length of these vectors may be different for each pair of words. For all cases where $o_{i,j}^{[\theta,\eta]} = 1$ the word pair of interest was identified in equal or less words from the threshold, thus according to the hypothesis, there is an association between them in the discourse. Finally the evidence from each article is aggregated into a new vector that contains the count of each of the p word pairs, $\mathbf{C} = c_{\theta,\eta}$:

$$c_{\theta,\eta} = \sum_i \sum_j o_{i,j}^{[\theta,\eta]}. \tag{3}$$

Given Ξ news articles, and that the publication date of each article is known, a vector $\mathbf{C}_\xi$ is recorded every time a new article is published, where $\xi = 1, \ldots, \Xi$. These tick data can be aggregated into any desirable time intervals forming time series of counts that the word pairs appeared within the threshold distance. From $\mathbf{C}_\xi$, given the desirable time interval, p time series are constructed with the respective sampling frequency. Assuming that

the sampling frequency of the economic variable of interest is f_1, the resulting time series will have sampling frequency $f_2 \geq f_1$, i.e. it will be of equal or higher frequency.

Over different periods, news outlets may produce a variable number of articles; hence, it is possible to introduce spurious patterns in the time series due to the per-period number of published articles. To avoid this, each X_i is normalised, by divided it by the number of articles published in the respective period. This way, series measures the *'density'* of the word pair in the discourse.

For the purpose of this study, these constructed time series are potential leading indicators for nowcasting the economic variable of interest.

2.2. Nowcasting

To produce nowcasts for economic variables there are two challenges. First, how to use variables of higher sampling frequency that are asynchronously updated to nowcast the target variable. Secondly, how to identify the relevant indicator variables, especially since now they can be arbitrarily many.

Let us assume that variables that were constructed from the news articles are monthly time series, while the target series, $Y = y_{t_{f_1}}$, is quarterly. Let $X_i = x_{i,t_{f_2}}$ be the i^{th} news series, with $i = 1, \ldots, p$ and $t_{f_2} = 1, \ldots, m_i$ the time period for that time series. Note that m_i is not the same for all time series, as they are updated asynchronously. Obviously, due to construction there is relationship between the two frequencies f_1 and f_2, as illustrated in figure 1.

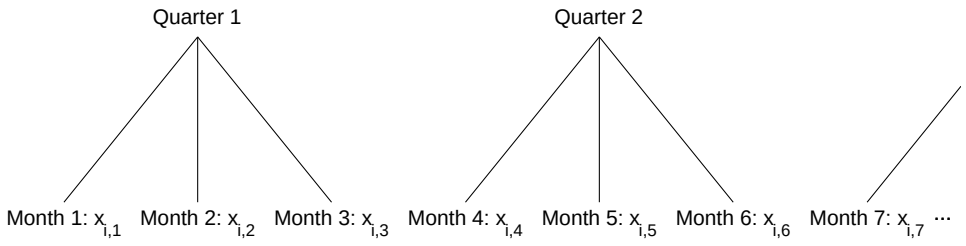

Figure 1: The high frequency time series X_i can be aggregated into the same frequency as Y, for example, quarters.

The conventional modelling in this case is to express Y as a function of the sum of the respective periods of X_i (Mariano and Murasawa, 2003). If the last quarter has not fully elapsed, then we assume any incomplete or unobserved months as missing values. The same is true if the publication date of the predictor variables has not passed yet. This approach has the limitations that one must handle the missing values. Kourentzes et al. (2014) considered the problem of combining estimations from different levels of temporal aggregation and instead rescaled all values in the same level of aggregation; for their case the one corresponding to the highest sampling frequency. For example, since $f_2 \geq f_1$, Y would be rescaled as:

$$Y' = Y\frac{f_1}{f_2}. \tag{4}$$

Since Y is the decision making relevant variable, it is convenient to keep it to the original frequency and scale the high frequency X_i variables. This can be easily done by inverting the ratio of the frequencies. For the case where one would have to consider observations with missing information, either due to date or publication lags, rescaling can be useful. Assume a quarter for which only two months have been observed. To rescale these two months into a full quarter one has to multiply the sum of these months with $f_2/2f_1$.In general, if only z periods have been observed for the higher frequency time series within a period of the lower frequency series, X_i can be scaled to the domain of Y by:

$$x'_{i,t_{f_1}} = \frac{f_2}{zf_1}\sum_{t_{f_2}=v}^{v+z-1} x_{i,t_{f_2}}, \tag{5}$$

where v is the first period of the high frequency time series during the current observation of the lower frequency time series. Note that the variable $X'_i = x'_{i,t_{f_1}}$ is now expressed in units of t_{f_1} and is sampled at the same frequency of Y, it has the same scale, and the same number of observations. Variables X'_i can be interpreted as the rate of publication density of each word pair per period. This transformation assumes that X_i does not have strong seasonal or trend patterns within a period of Y, i.e. that the rate of publication does not change substantially within the period of interest. This limiting assumption affects strongly the approximation at the beginning of the period of t_{f_1} and becomes weaker as time progresses. We will call these incomplete periods as *nowcast periods*, since the nowcasted value of Y will be based on those.

The result of the transformation above is p explanatory variables X_i' that have $n+1$ observations, where n is the sample size of Y and the last observation of X_i' refers to the current unobserved period of Y, the nowcast period. These can be incorporated in the nowcasting statistical model as:

$$Y_{t_{f_1}} = baseline + \sum_{i=1}^{p} w_i x'_{i,t_{f_1}}, \tag{6}$$

where *baseline* is used to denote all other terms of the nowcasting model, including time series dynamics of Y and various other variables that may be included (for examples of variables see Giannone et al., 2008; Banbura et al., 2010). The model is estimated using data up until the last publication date of Y. Once the coefficients w_i are known, it is trivial to include the sentiment in the news for the state of the economy to the nowcast. This is done by using the remaining last observation of X_i' to calculate their effect.

Given a large Θ the resulting number of time series can be more than the available data points of the target economic variable. This introduces a variable selection problem. In the literature there are several approaches to deal with the so called '*fat regression*' problem, where the number of predictors p is larger than the sample size of the target variable n (for example, see Tibshirani, 1996; Stock and Watson, 2002; Jolliffe, 2002; Zou and Hastie, 2005; Stock and Watson, 2006). Here, the use of lasso regression is considered to estimate the coefficients of X_i'. Lasso regression has the ability to drive parameter estimates to zero, due to its $L1$ coefficient penalty function, therefore effectively performing variable selection (Tibshirani, 1996).

Apart from the data driven approach to variable selection, since the word pairs are capturing the sentiment of the press for the current state of the economy, several variables may be possible to be removed simply by considering whether the pairs are meaningful or not. Therefore, instead of providing p variables to the model, several combinations may be removed beforehand. Furthermore, some combinations may have zero occurrences, further reducing the number of variables.

Note that the prescribed approach allows updating the nowcasts every time there is some news item published. Irrespective of the frequency f_2, a new vector $\mathbf{C}$ will be created for each article, which in turn will change X_i and X_i', thus allowing to re-calculate the nowcast considering all information up until that point. It is also possible to use different aggregation frequencies for each word pair. The only difference in that case would be the rescaling

factor in Eq. (5), where f_2 would change accordingly for each variable.

3. Empirical evaluation

3.1. Data

The proposed methodology is applied in nowcasting the GDP growth of Greece. Quarterly GDP data have been downloaded from OECD Library (OECD, 2014), from 2000 until the 1^{st} quarter 2013. Articles from the leading financial newspaper in Greece, *Naftemporiki* have been collected for the same period. In total 394,571 articles have been collected. A text corpus with 42 words have been created using words: i) associated with the state of the economy (see table 1); ii) names of Greek prime ministers and ministers of economy for that period; iii) names of influential EU and IMF officials during the Greek financial crisis. The Greek language uses stresses and different forms of nouns depending on the syntax of sentences. This can cause problems in matching the words appropriately. To avoid this, all articles were preprocessed to remove stresses and the keywords were used without their suffixes to allow for a more generic search.

Table 1: Translated keywords, excluding names and words whose meaning cannot be translated in English with a single word.

austerity	drachma	income	recovery	treasury
budget	economy	memorandum	stock exchange	unemployment
CDS	euro	parliament	surplus	
crisis	GDP	prime minister	tax	
deficit	IMF	recession	troika	

All articles were mined and 861 word pair variables were created using a threshold $\phi = 10$. From these variables, 162 contained only zero observations and were dropped from the analysis. The rest were aggregated to monthly time series. The last 21 GDP growth quarters were used as an out-of-sample set, to evaluate the accuracy of the nowcasts. Figure 2 illustrates the GDP growth series, as well as the different word pair variables that were constructed. Observe that some variables seem to be negatively correlated with GDP growth. In other words, bad news about the economy increase as there is a drop in GDP growth. Essentially, the question that is posed, given the holdout sample, is whether a nowcast model that uses the sentiment in

the economic discourse could accurately describe the significant decline observed, and in particular, the triple dip that can be seen in the GDP growth data.

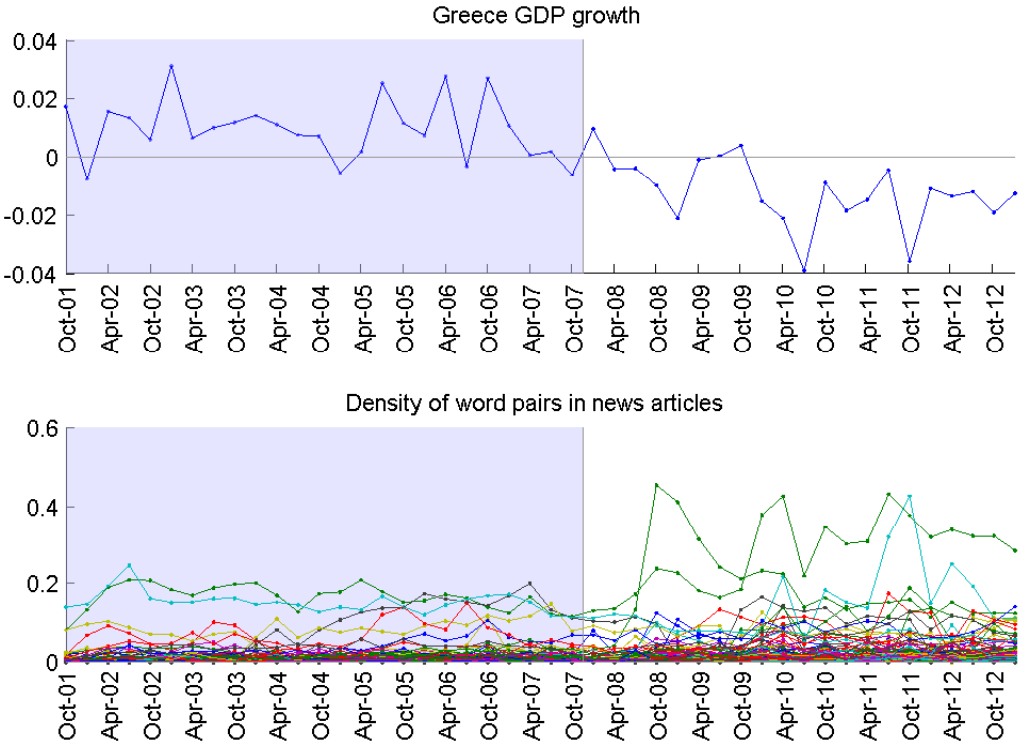

Figure 2: Plot of GDP growth and word pair variables. The fitting sample is highlighted with darker background.

Table 2 presents the five variables that are mostly correlated with GDP growth, across the whole sample. The pairs match key associations that were commonly found in the economic discourse in Greece during the financial crisis. These follow the discussions around the bad financial health of the Greek state, the so called 'Grexit' from the European Monetary Union and mentions of IMF and Troika that have been managing the most recent loans to Greece. Note that some of these variables are highly correlated with each other, so potentially not all of them are useful predictors.

3.2. Experimental setup

The aim of the experiment is to assess how accurately the GDP growth could have been nowcasted from the start of 2008 until the end of the sample

Table 2: Top 5 correlated variables with GDP growth

Keywords	Correlation
economy - treasury	-0.757
recession - deficit	-0.745
euro - stock exchange	0.730
IMF - Troika	-0.702
euro - treasury	-0.697

using the proposed inputs. Given a fitted model, for each quarter three nowcasts are produced: i) using only news from the first month of the quarter; ii) using news from the first two months of the quarter; iii) using all news in the quarter. hence, the accuracy of the nowcast can be tracked as new information becomes available.These models are named *News 1*, *News 2* and *News 3* respectively. The parameters of the model do not change when inter-quarter information is released, however when the published GDP figure becomes available, all model parameters are updated. Obviously, the same experiment could be setup with daily news variables instead of monthly, or any other sampling frequency, using the same inputs.

Lasso is used to estimate the nowcast models. The inputs of the model are all 699 non zero word pair variables and 5 autoregressive terms for GDP growth. A constant term is included in the model. The lasso regularisation parameter is identified using 20-fold cross validation. This also determines which variables are selected for the model. Note that as the model is re-estimated for every quarter to be nowcasted, the regularisation parameter and the selected variables change.

The accuracy of the nowcasts will be assessed in terms of Root Mean Squared Error (RMSE), and its median equivalent (RMdSE), against the published quarterly GDP growth figures. The performance of the proposed model is compared against nowcasts produced by an autoregressive model (AR) and the random walk (RW).

3.3. Results

Table 3 presents the results for the nowcast accuracy. For each error measure, the model with the highest accuracy is highlighted in boldface. Looking at *News 1*, *News 2* and *News 3* we can observe that as more up-to-date information becomes available, the quality of the nowcasts improve both in terms of RMSE and RMdSE. *News 3* has the best accuracy. All

models that use news inputs perform better than the benchmark *AR*. On the other hand, *RW* outperforms *News 1* that includes only limited additional information.

Table 3: Nowcast accuracy

Model	RMSE	RMdSE
News 1	0.0149	0.0103
News 2	0.0135	0.0061
News 3	**0.0133**	**0.0048**
AR	0.0150	0.0108
RW	0.0147	0.0094

The increased accuracy of the *News* models shows that the proposed variables are useful for capturing the current sentiment over the state of the economy and can partially explain the movements of GDP growth.

Although lasso regression is useful for variable selection, it does not eliminate the need for providing a useful initial pool of potential input variables. The *News* models in table 3 consider 704 variables, resulting in a difficult variable selection problem. Most of these variables are not used in any of the nowcasts for the different dates in the holdout sample. Focusing on the most frequently selected variables, we can identify a subset of variables to use as inputs, seen in table 4, together with any potential autoregressive terms. Although there are similarities with the most correlated variables in table 2, these variables are not the same, since some of the pairs presented in table 2 are highly collinear.

Table 4: Selected word pairs and percentage of nowcasts that use them.

Word pair	% of nowcasts
economy & income	100.00%
economy & treasury	100.00%
recession & deficit	100.00%
recession & unemployment	90.48%

Variable selection in the new reduced *News* models is greatly simplified. The new nowcast accuracy is presented in table 5. The best performing model is highlighted in boldface. Comparing these results with table 3 we can observe substantial reductions in the nowcast errors. Again, as more updated news information is provided, nowcasts become more accurate. With

the reduced set of variables all *News 1*, *News 2* and *News 3* outperform both *AR* and *RW* in terms of both RMSE and RMdSE.

Table 5: Nowcast accuracy - Reduced *News* models

Model	RMSE	RMdSE
News 1	0.0114	0.0083
News 2	0.0096	0.0063
News 3	**0.0090**	**0.0058**
AR	0.0151	0.0108
RW	0.0147	0.0094

Figure 3 provides the kernel density estimations of the probability density functions of the squared errors for each model, in order to better demonstrate the improvements gained using the news variables. The actual squared errors are represented in the figure with vertical ticks. As more updated news information is used the distribution of squared errors becomes more concentrated towards smaller errors and there are less outlying observations. In contrast, the distributions for both *AR* and *RW* are wider, including multiple outlying errors.

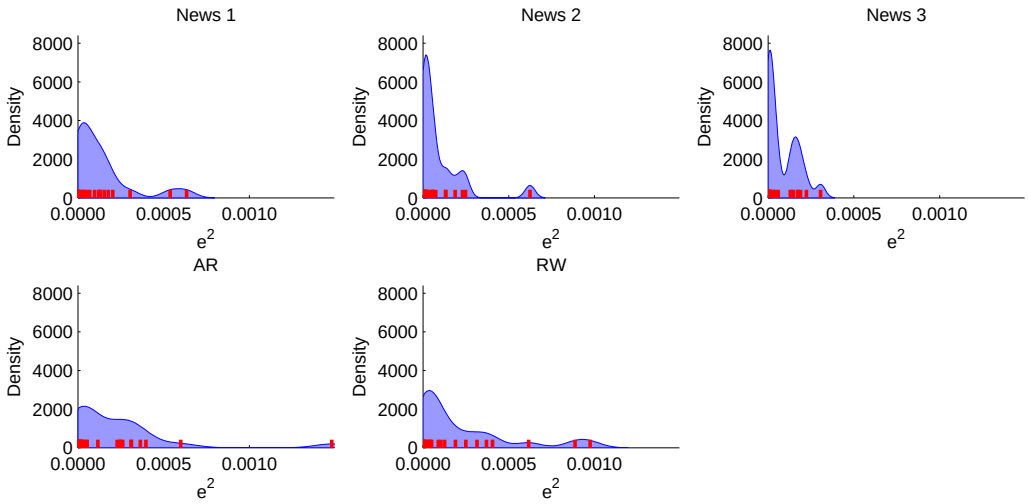

Figure 3: Kernel density estimation of density function of squared errors for the different models. Location of actual errors is noted with vertical ticks.

Nowcasting the Greek economy GDP growth is especially challenging due

to the triple dip that happened during the years 2008 and 2013 (see fig. 2). Figure 4 provides the point nowcasts for the different quarters in the holdout period. The *News* models are capable of capturing these dips. In particular, towards the latter part of the holdout sample, the estimation of the GDP growth is very accurate, as the model has assigned appropriate weights to the news variables.

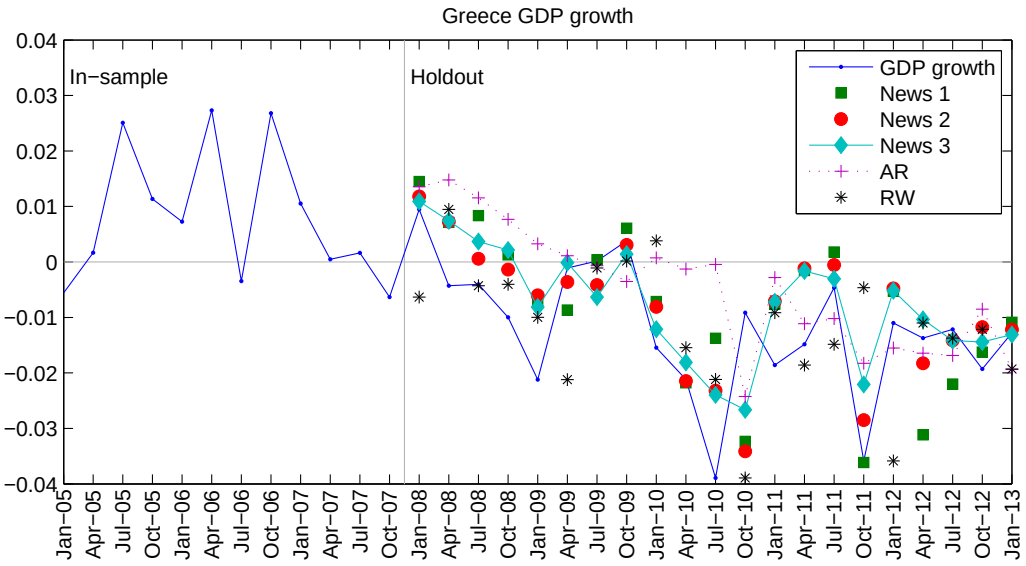

Figure 4: GDP growth nowcasts using the reduced *News* models.

3.4. Discussion

The results provided in the previous section demonstrated that there is useful information for nowcasting the GDP growth in the word pair variables, giving empirical support to the hypothesis that quantifying the sentiment about the state of the economy in the proposed way is beneficial. The empirical evaluation discussed in this paper is limited, as it does not consider nowcasting models that use other economic variables. The aim of this evaluation was to demonstrate whether this new class of input variables are useful for nowcasting and provide a proof of concept for the proposed modelling.

At this point, it is useful to discuss what kind of information is captured by such variables. News articles can have many forms and types of content. Often, they report events or statements from politicians, influential decision

makers or business leaders. Other times, they contain opinions and analysis of events or statements. Without providing an exhaustive list of the different contents that news articles may have, it is safe to say that they reflect the current sentiments of the authors and parts of the society. News capture how the market reacts to events and economic developments, and up to a certain extend even affect these developments. Therefore, they provide an invaluable source of information on the current state of the economy and how various policy and decision makers' statements and actions are reflected by the economy.

A great advantage of incorporating the current discourse about the state of the economy from news outlets is the frequency that new information becomes available. Significant events are reported with a minimal lag. This information can in turn be incorporated in the nowcasts. In the current work only monthly updates were considered, however it is trivial to model daily or even hourly updates. Once other economic variables are considered in nowcasting GDP, the high frequency news variables will be able to refine further the nowcasts, containing the most up to date sentiment about the state of the economy.

4. Conclusions

We proposed a framework to include information from the media into GDP nowcasts. To demonstrate its usefulness we used Greece as a case study. The Greek economy has been in a prolonged recession, with its GDP growth having observed a triple dip over the last years, thus providing a challenging case. Empirical evidence was provided that by incorporating such variables the accuracy of nowcasts is improved. Intuitively this is expected, as these variables capture the sentiment in the media about the state of the economy.

The key advantage of the proposed inputs is the very high frequency of updates and the absence of publication delays. The proposed framework allows incorporating any desirable rate of sampling, although results for only monthly updates were provided. The aim of the study is to provide a proof of concept of the new inputs, rather than a complete nowcast model. Future work will consider economic variables together with news variables to maximise use of available information for the nowcasts.

It is possible to create a large number of variables by mining news articles, introducing a problem of variable selection, while several of the variables may

be multicollinear. In this paper lasso regression was used to address this problem, however different alternatives should be explored.

Acknowledgements

The authors would like to acknowledge the helpful comments provided by Sigrun Valderhaug Larsen for conducting this research. The authors are also grateful to Niki Kontoe for collecting the on-line news articles, an indispensable component of this research.

References

Altheide, D. L., 1997. The news media, the problem frame, and the production of fear. The sociological quarterly 38 (4), 647–668.

Altheide, D. L., Schneider, C. J., 2012. Qualitative media analysis. Vol. 38. Sage.

Banbura, M., Giannone, D., Reichlin, L., 2010. Nowcasting. ECB working paper.

Choi, H., Varian, H., 2012. Predicting the present with google trends. Economic Record 88 (s1), 2–9.

Evans, M. D. D., 2005. Where are we now? real-time estimates of the macro economy. International Journal of Central Banking 1 (2), 127–174.

Giannone, D., Reichlin, L., Small, D., 2008. Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics 55 (4), 665–676.

Jolliffe, I. T., 2002. Principal Components in Regression Analysis. Springer, pp. 167–198.

Kourentzes, N., Petropoulos, F., Trapero, J. R., 2014. Improving forecasting by estimating time series structural components across multiple frequencies. International Journal of Forecasting 30 (2), 291–302.

Kuzin, V., Marcellino, M., Schumacher, C., 2011. Midas vs. mixed-frequency var: Nowcasting gdp in the euro area. International Journal of Forecasting 27 (2), 529–542.

Marcellino, M., Schumacher, C., 2010. Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp*. Oxford Bulletin of Economics and Statistics 72 (4), 518–550.

Mariano, R. S., Murasawa, Y., 2003. A new coincident index of business cycles based on monthly and quarterly series. Journal of Applied Econometrics 18 (4), 427–443.

OECD, Jan 2014. Main economic indicators - complete database. URL /content/data/data-00052-en

Scott, S. L., Varian, H. R., 2013. Predicting the present with bayesian structural time series. International Journal of Mathematical Modeling and Optimization.(forthcoming).

Stock, J. H., Watson, M. W., 2002. Forecasting using principal components from a large number of predictors. Journal of the American statistical association 97 (460), 1167–1179.

Stock, J. H., Watson, M. W., 2006. Forecasting with many predictors. Handbook of economic forecasting 1, 515–554.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58, 267–288.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67 (2), 301–320.