НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
ВЫСШАЯ ШКОЛА ЭКОНОМИКИ

Международный Институт Экономики и Финансов

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ
по образовательной программе высшего профессионального образования, направление 080100.68 Экономика

на тему:

Forecasting the Unemployment Rate in Russia with the Internet Search Volume Data

Студентка 2 курса магистратуры
Достова И.К.

Научный руководитель
Ph.D. in Economics, Fantazzini Dean

МОСКВА, 2013 год

Прогнозирование уровня безработицы в России с помощью объема поисковых запросов в Интернете

Изучение поисковых тенденций пользователей Интернета стало возможным с появлением сервиса Тренды Google, разработанного корпорацией Google. Данный сервис позволяет узнать, насколько популярен поисковый запрос по указанному ключевому слову по отношению к общему объему запросов, выполненных через Google за определенную неделю в соответствующем регионе. Безусловно, указанные данные сразу же нашли широкое применение в экономическом анализе.

В текущем исследовании данные об объеме поисковых запросов в Интернете используются с целью прогнозирования уровня безработицы в России. Интернет сегодня активно применяется для поиска работы, а данные об объеме запросов, связанных с поиском вакансий, доступны практически в реальном времени. В то же время официальные данные об уровне безработицы в России публикуются со значительной задержкой.

Данное исследование стремится проверить релевантность индексов поисковых запросов для использования в качестве основных опережающих индикаторов уровня безработицы. Это особенно актуально для России, учитывая отсутствие такого общепринятого индикатора. Для целей исследования был выбран не только сервис Тренды Google, но и статистика ключевых слов на Яндексе, принимая во внимание, что Яндекс является основной поисковой системой в России (более 60% рынка). Таким образом, в работе представлено первое сравнение индикаторов из двух источников данных поисковых запросов (а не только используемого ранее сервиса от Google). Для расчета соответствующих индикаторов были выбраны два наиболее популярных ключевых слова: “работа” и “вакансии”. Всего в исследовании тестируется широкий набор из 24 различных индикаторов поисковых запросов: нормализованных/не нормализованных, измеренных в абсолютной/относительной шкале, месячных/недельных, текущих/лаговых.

В работе проведено масштабное сравнение предсказательной способности (за пределами выборки) более 47000 моделей: стандартных ARIMA, ARIMA с соответствующим индикатором поисковых запросов и нелинейных моделей. Определение набора лучших моделей основано на передовом подходе Model Confidence Set (Hansen et al., 2011).

Среди основных результатов можно выделить следующие:

- Успех подхода, основанного на статистике поисковых запросов, выявлен для различных индикаторов Яндекс и Google, а также обоих ключевых слов.

- Использование индикаторов поисковых запросов приводит к улучшению предсказательной способности стандартных моделей временных рядов на различной дальности прогнозирования (на 1, 2 и 3 месяца вперед).
- Прогноз на 1 месяц вперед: лучшими оказались модели с индексом Google и одна модель с нормализованным объемом запросов в Яндексе: они превосходят точность прогнозов всех альтернативных моделей, включая стандартные модели временных рядов, измеренные по более широкой выборке.
- Прогноз на 3 месяца вперед: примерно 50% лучших моделей содержат индексы Google и 35% - индикаторы Яндекс.
- Качество прогноза моделей, основанных на поисковых запросах, превосходит результат моделей, основанных на классическом макроэкономическом индикаторе – Индексе Промышленного Производства (ИПП). Кроме того, дополнительное включение ИПП, не улучшает прогнозную силу лучших поисковых моделей.
- Была найдена модель, которая входит в набор лучших моделей в прогнозировании на 1, 2 и 3 месяца вперед – ARIMA модель с опережающим индексом Google для слова “вакансии”.
- Результаты прошли тест на фальсификацию.

Несмотря на то, что исследование проводилось на небольшой выборке (данные поисковых запросов Google доступны с 2004 года, Яндекс – 2008), результаты исследования являются крайне многообещающими.

Contents

1. Introduction

The availability of the internet search query data has granted the access to the invaluable information that could be incorporated into the economic analysis. Since Google has launched its Google Trends tool, it is possible to find out about the search patterns of the web users.[1] The Google index (GI) is a new tool that indicates a weekly number of searches of any given keyword, relative to the total number of searches done on Google over time.[2] There is an opportunity to specify your search query data for a region of interest, certain time period or category and get the results for a group of keywords at once.

No wonder that such a tool has been immediately appreciated in certain economic applications. In one of the first articles in this field the Google index has been adopted to monitor and track the influenza epidemics (Ginsberg et al., 2009). A simple model with the Google index for influenza search queries is developed in order to predict physician visits due to this disease. Besides that, the Google index has been successfully used to forecast sales and tourism (Choi & Varian, 2009.1), private consumption (Schmidt & Vosen, 2009), earnings-related fundamentals (Da et al., 2011) and so on.

The focus of this paper is the use of the internet search volume data in order to forecast the unemployment rate in Russia. The motivation for the use of such kind of data for unemployment prediction is the following:

- Nowadays the internet is a highly popular job-search mean. Thus, it is believed that a lot of people search for a job via the internet and their queries reflect their intentions to find a job.
- The official unemployment releases are available with a considerable lag, whereas the job-search query data is reported and updated frequently on the continuing basis. This could significantly improve the predictive power of the models for unemployment rate.

The current research tests the relevance of wide number of job-search-based indices as the first leading indicators for unemployment dynamics in Russia. The study takes advantage of the two search volume services: the Google Trends which is commonly used in literature and the Yandex Wordstat – taking into account that Yandex is the main search engine in Russia. This is the first comparison of the two search engines as sources of the search volume data. Moreover, the explored search-based indicators are quite different in their scale, frequency, and computation. Overall, there are 24 different search volume indicators for two most popular keywords for job-search in Russia. An extremely extensive forecasting "horse-race" of more than 47200 unemployment models (standard *ARIMA*, *ARIMA* augmented with job-search indicators, nonlinear models) is conducted for the purposes of the

[1] At first there used to be two tools: Google Insights for Search and Google Trends. Recently they have been merged under the name of Google Trends. Search volume data is available from 2004.
[2] GI is presented on a scale of 0 to 100, i.e. with a 100 given to the period with the highest search query level.

current research. The assessment of superior predictive ability of the chosen 1000 models is performed according to the most advanced procedure – the Model Confidence Set (Hansen et al., 2011).

The results of the research allow to identify the best unemployment forecasting models, best leading unemployment indicators and to check the performance robustness of the proposed approach as a whole. Taking into account that there is no reliable unemployment leading[3] indicator in Russia (like the Initial Claims in the US), this research is really urgent. It suggests the approach which would satisfy the need of reliable and updated unemployment forecasts in Russia.

For the correct perception of this approach several limitations should be noted. First of all, not all people search for a job via the internet, so those who do that are, probably, not randomly selected. Next, the internet search query data could be also driven by the on-the-job search. So some information incorporated in the examined data would be related to the already employed people. Finally, the increase in the job-search queries could also arise because of the increased search intensity of any individual or the fact that more and more people are becoming engaged into the internet activity. With regard to this latter point it should be noted that GI is normalized by the total number of the search quires, so that it accounts for the growing popularity of the internet. Also GI doesn't include the repeated queries done from the same IP address within the short period. The final point is the limited availability of the search volume data, since the discussed tools have been designed relatively recently. Notwithstanding all these limitations it is believed that the results of a current research are very promising.

The main body of the paper is organized as follows:

- in section 2 there is a review of the selected papers which use job-search-related query data for the unemployment forecasting in certain countries;
- section 3 describes main data sources, explored job-search-based indicators and the methodology of models estimation and their predictive ability assessment;
- section 4 discusses the main results of the out-of-sample unemployment forecasting and presents the best models and their leading indicators for short-term forecasting of the unemployment rate in Russia;
- section 5 provides the additional checks of the current research results;
- section 6 concludes.

[3] The leading indicator is defined by its ability to forecast the target variable and its earlier availability (it is released before the target variable is officially published).

2. An overview: job-search-related queries and unemployment forecasting

This section discusses several selected papers which adopt the Google index in order to improve the predictive power of the unemployment dynamics in Germany (Askitas & Zimmermann, 2009), Italy (D'Amuri, 2009) and the United States (Choi & Varian, 2009.2; D'Amuri & Marcucci, 2009, 2012).

The paper ***«Predicting Initial Claims for Unemployment Benefits»*** describes how Google Trends data could be used in order to predict the initial claims in the US (Choi & Varian, 2009.2). The indicator of the initial jobless claims shows the number of new people filed for unemployment benefits.[4] In the US it is known as a key indicator of labor market. Choi & Varian choose two Google indices from the following categories in Google Trends: "Jobs" and "Welfare & Unemployment". It should be noted that the Google indices are available a week ahead of the initial claims release date.

The standard autoregression *AR(1)* is chosen as a baseline model for the initial claims. To see how the Google data could improve the predictive power, the baseline model is augmented with the Google Trends series for the chosen categories (the alternative model). The baseline and the alternative models are estimated on the long and the short samples. The long term model is estimated for the 5-year interval from 2004 when the Google data became available for the first time. While the short term model is based on the certain recession period from December, 2007.

The analysis conducted by Choi & Varian shows that both, the long term model and the short term model, are improved significantly with an addition of the Google Trend indices. The out-of-sample *MAE*[5] for the long term model is decreased by approximately 16% and for the short term – by around 13%.

The article ***«Google Econometrics and Unemployment Forecasting»*** focuses on testing for the correlations between the search queries using certain job-related keywords and monthly unemployment rate in Germany (Askitas & Zimmermann, 2009). The authors believe that such kind of data could be used for predicting indicators measured by traditional statistics services. In Germany the unemployment rate for a certain month is based on the data of the unemployment office between the second half of the previous month and the first half of the current month. To predict the unemployment rate the authors construct the Google Insights indices separately for weeks 1 & 2 of the current month and weeks 3 & 4 of the previous month. There are 4 considered keyword groups: 1) unemployment

[4] The indicator is published in a weekly report "The Initial Jobless Claims" issued by the Department of Labor in the US.
[5] The Mean Absolute Error is computed with the rolling window for the latest 24 weeks.

office or agency, 2) unemployment rate, 3) personnel consultant and 4) the most popular job search engines in Germany. The first three indices move in the same direction whereas the last one – in the opposite.

The authors regress the change in the unemployment rate on its past level, the change in the Google indices (in different combinations) and their past levels. For modeling the change in the variables of interest and their past levels the authors use 12-month lag operators. The regressors are taken either as of the first two weeks of the current month or the last two weeks of the previous month. The statistical performance is evaluated using the *BIC* criterion.

The analysis demonstrates the presence of strong correlations between the queries using job-related keywords and the unemployment rate in Germany. The findings of the research suggest two things: models using data from weeks 3 & 4 of the previous month perform better in comparison with the rest ones and the combination of keywords from groups 1 & 4 is preferred to other specifications.

The article ***«Predicting unemployment in short samples with internet job search query data»*** tests the predictive power of job-search-related Google index for unemployment rate in Italy (D'Amuri, 2009). The unemployment rate in Italy is reported on the quarterly basis.[6] This release schedule reduces the sample that could be used for estimation and out-of-sample forecasting. The author believes that the Google index would perform well even in such short sample. The chosen Google index is based on the queries for "job offers". As soon as it is a weekly index, the author takes simple quarter averages of it. Also according to the definition of the unemployed[7] and the fact that the interview's week is unknown, the Google index is rescaled 2 weeks ahead. In addition to the Google index, there are two more standard exogenous variables: the monthly Industrial Production Index and the weighted average of the individual sector indicators[8] from the employment expectations survey.

The author choses *ARIMA(1,1,0)* as a benchmark model for the unemployment rate. This model is estimated using different exogenous variables and over samples of different lengths (the long one starts from 1885 and a short one – from 2004). The lag structure and the reference month of the dependent variable alter. The models with the Google index are compared to the otherwise identical models estimated on the same short sample or on the longer one. The total number of estimated models is nearly 40. The rolling scheme is used on 7 interval lengths: 14-20 observations for the short sample and 90-96 for the long one. The models' comparison criterion is out-of-sample *MSE*.

[6] According to the Italian Labor Force Survey.
[7] A person who is not employed and who has been looking for a job in the preceding 4 weeks.
[8] The indicators are computed by the European Commission and determine the balance of the number of professional forecasters who expect the up/down employment movements in the next 3 months.

The results of the forecasting procedure show that the Google index performs better than the more standard leading indicators of the unemployment dynamics. Moreover, the forecasts estimated with the short sample models which adopt the Google index are more accurate in comparison even with those ones estimated on a longer sample and augmented with other considered indicators.

With respect to the unemployment rate in the US, it is widely accepted to use the initial jobless claims as a leading indicator. However, the paper ***«Google it! Forecasting the US unemployment rate with a Google job search index»*** suggests the use of the Google index as an alternative leading indicator of the monthly unemployment rate (D'Amuri & Marcucci, 2009). To construct the Google index the authors use the keyword "jobs" as the most popular among the relative incidences. The cited paper is the first one using such web search indicator to predict the unemployment rate in the US. The authors believe the Google index to be the best leading indicator of the US monthly unemployment rate in the short-term forecasting, i.e. from 1- to 3-month ahead.

The paper is especially interesting as the authors run an extensive out-of-sample forecasting "horse-race" [D'Amuri & Marcucci, 2009, p.1] of more than 500 models for the first differences of the US monthly unemployment rate. The authors focus on the linear *AR(p)* and *ARMA(p,q)* models with lags up to the 2nd for *p* and *q*.[9] These models are believed to be the most appropriate to capture the short-term dynamics of the time series. The addressed models are augmented with the Google index, the initial jobless claims or combinations of both. Besides that, the authors estimate 3 most common nonlinear models which are used to capture the long-term dynamics of the US unemployment rate.[10]

The authors use a rolling forecasting scheme based on the information available at month *t*. Thus, the models are estimated over samples of different length: the short sample[11] for the models which adopt the Google index and the long one[12] for those without GI. All the models are ranged in terms of lowest out-of-sample *MSE* for all of the considered forecast horizons. Finally, the best 15[13] models were compared with best 2 models without *GI* and the nonlinear models by performing the tests for equal forecast accuracy (Diebold & Mariano, 1995), forecast encompassing (Harvey et al., 1998) and the White's reality check (2000).

As a robustness check all of the *ARMA* models are estimated for each of the 51 states of the US to assess the share for which of them the models with lowest *MSE* are those using the Google index.

[9] In addition the models are augmented with the seasonal multiplicative factors *SAR(12)* and both *SAR(12)* and *SMA(12)* for *AR* and *ARMA* models, respectively.
[10] *AAR*, *LSTAR*, *SETAR*.
[11] [2004.1-2007.2] – in-sample, [2007.3-2009.6] – out-of-sample.
[12] [1967.1-2007.2] – in-sample, [2007.3-2009.6] – out-of-sample
[13] It should be noted that in the original research all these models include GI.

Finally, the best models for the unemployment rate are used to construct the quarterly forecasts in order to compare them with the predictions gained from the Survey of Professional Forecasters[14].

The analysis of the predictive power of different models brings the following main results:

- At all forecast horizons the best 15 models always include the Google index, in some cases with the initial claims. A simple *AR(1)* model with the average monthly Google index as a leading indicator shows the best performance among 1- and 2-month-ahead forecasts. The best model for 3-month-ahead forecast is *ARMA(1,1)* augmented with the average monthly GI and the seasonal multiplicative factor. Thus, the best models outperform the similar models using the initial jobless claims, even estimated over longer samples.
- The preferred simple linear models with the Google index outperform nonlinear models, the forecasts gained from the Survey of Professional Forecasters and in state-level forecasting (the share of best models with the GI as a leading indicator ranges between 75% and 84% for 1- to 3-month-ahead forecasts, respectively).

Later on D'Amuri & Marcucci updated their previous paper and published "The Predictive power of Google searches in forecasting unemployment" (2012). It is reasonable to highlight here 2 most interesting innovations that go beyond the "Google it!" paper (D'Amuri & Marcucci, 2009):

- The best selected models (in terms of out-of-sample *MSE*) are tested for the superior predictive ability using the Model Confidence Set (MCS) proposed by Hansen et al. (2011). This procedure helps to identify the set of best models that outperform all the competing ones. The results show that about 30% of the best models in MCS include the GI as the leading indicator.
- The results survive the falsification test when the authors augment the standard *ARMA* models with a "fake" GI. This indicator shows the highest in-sample correlation with the unemployment rate, but doesn't logically relate to the job-search.[15] The results show that these alternative models are not ranked among the best 15 models according to their out-of-sample *MSE* values. Thus, only job-search indicators are relevant for the unemployment forecasting.

The analysis conducted by D'Amuri & Marcucci is the most thorough and impressive among others reviewed in this chapter and their methodology can be considered as a benchmark one.

[14] A quarterly survey conducted by the Federal Reserve Bank of Philadelphia.
[15] The Google Correlate service indicates the keyword "dos" as having the highest correlation with the US unemployment rate.

3. Data & Methodology

3.1. Main Data Sources

For the purpose of the research the two main groups of variables are used: official unemployment rate and the search volume data.

Unemployment Rate in Russia

u – monthly unemployment rate. Russian monthly unemployment rate is published in the collection "The socio-economic state of Russia" issued by the Federal State Statistics Service (FSSS)[16]. These data are available for the period [2002.1-2012.12] (the long sample). FSSS publishes the unemployment rate as of the end of each month (an instantaneous indicator). The estimation is based on the results of the labour force sample surveys. The survey is conducted by direct visiting of the households and questioning people between 15-72 years old. The examined, or reference, week is defined as the 2nd week of each month. Respondents are divided into the types by economic activity in accordance with the International Labour Organization. Thus, the person is threated as the unemployed if he does not have a job during the examined week, but is ready to work and have been looking for it during 4 prior weeks ending with the examined one. The results of the survey are then extended to the general population.[17]

As it is said the problem is that the official release of the unemployment rate is available with a 1-month lag.

Search Volume Data

The strong advantage of the use of job-related search volume data as the leading unemployment indicators is their availability almost in real time, in contrast to the lagged official unemployment releases.

The first service that is used to get such data – the Google Trends[18] made by Google. It publishes the widely used Google Index on a weekly basis. The GI is a weekly number of searches for a given keyword, relative to the total number of searches in the same week and area. As follows from the definition this index is normalized by the total number of searches so that it controls for the growing number of the internet uses. This means that the GI isn't distorted by the increased popularity

[16] http://www.gks.ru.
[17] See http://www.gks.ru.
[18] http://www.google.com/trends/explore#cmpt=q.

of the internet in the recent time. GI is presented on a scale of 0 to 100, i.e. after normalization each relative number is divided by the highest value and then multiplied by 100. GI is available since 2004 and explored in the research within the period [2004.1-2012.12].

But as opposed to the US, Google is not the main search engine in Russia. It is supposed that Google data could be less representative as the majority of the internet users in Russia (around 60%) choose Yandex as a search service. Due to this reason the current study takes advantage of the similar Yandex service as well. Yandex Wordstat[19] provides the monthly absolute search volume for a given keyword. Thus, in contrast to GI which is more like a "Black Box", Yandex enables to access the search volume explicitly. But at the same time Yandex doesn't control for the internet use growth as it is not normalized. Besides that, Yandex data is available for a shorter period as Yandex cuts its data so that to show just two recent years. The Yandex data are collected for the period [2009.11-2012.12].

Table 1 summarizes the pros and cons of both search volume services that are used in the research.

Table 1. Comparison of the search volume services: Google Trends & Yandex Wordstat

	Pros	Cons
Google Trends (relative scale; weekly indices, *w*)	• Larger time period • Normalized	Does not allow to access the search volume
Yandex Wordstat (absolute scale; monthly indices, *m*)	• Main search engine in Russia • Search volume is presented explicitly	• Limited access • Does not control for the internet use growth

3.2. Job-Search-Based Indicators

It is reasonable to divide the search volume indicators used in the research into the following types according to the computation method:

- ***eg*** – simple monthly average of the weekly Google indices, i.e. based on *w1-w4* of m_i[20] (simple monthly average GI);

[19] http://wordstat.yandex.ru.
[20] Eviews data frequency conversion.

- ***g*** – average of the weekly Google indices related to *w1-w2* of m_i & *w3-w4* of m_{i-1} (alternative average GI)[21];
- ***y*** – simple monthly Yandex search volume;
- ***q*** – monthly Yandex search volume adjusted for the number of Yandex users. The indicator is computed in the following way: ***y*** is divided by the monthly Yandex audience[22], i.e. number of people who used Yandex at least once a month.

As it is said Yandex Wordstat helps to access explicitly the popularity of the keywords (in volume terms). Besides, it displays all the related queries in the descending volume order. The analysis of the job-search queries shows that it is reasonable to calculate the chosen indicator types for the following 2 main job-related keywords[23]:

- ***1*** – "работа". According to Yandex Wordstat this is the most popular keyword used for a job-search;
- ***2*** – "вакансии". This is another popular keyword used for a job-search, but its search volume is on average about 3,6 times lower than that of the keyword ***1***. The other job-related keywords are not explored in the research as keyword ***2*** is the only significant (in volume terms) one related to keyword ***1***: the search volumes for other keywords are much lower.

It is important to note that the search volumes of the chosen keywords are cleared from the similar but unrelated search queries.[24] Another thing that should be in mind – all indicators are explored within Russia, so the sample is not distorted by the job-search queries from citizens of other countries.

3.3. Models for the Unemployment Rate

The first group of tested models consists of standard *ARIMA* models (St. *ARIMA*). To test for the stationary of the unemployment rate there is an Augmented Dickey-Fuller test (*ADF*) for the unit root process (see Description 1[25] of Appendix). This test is performed for the long sample [2002.1-2012.12] and the shorter one [2004.1-2012.12]. The results are presented in Tables 1,2 of Appendix. In the first case (the long sample) the null hypothesis of the unit root is rejected at 5%-sig. level. The results for the short sample indicate the potential presence of the unit root (failed to reject the null) at 5%-sig.

[21] It corresponds to the job-search period in the definition of the unemployed.
[22] Data for the number of Yandex users are available at http://stat.yandex.ru .
[23] The search volume services are designed to capture the effect of different keywords declension.
[24] This is done with a minus operator. For example, the search query for the keyword ***1*** looks like "работа -курсовая -диплом -контрольные -часы -дипломная -лабораторная -график -воспитательная -домашняя -по -самостоятельная -режим -время -исследовательская -ру -тему -практическая -скачать".
[25] For the details in the test see Enders (2004).

level. In order to be agnostic[26] it is decided to test the same model specifications for the level and first differences of the unemployment rate.[27] According to this agnostic approach and the in-sample correlogram analysis[28] a wide range of different *AR*- and *MA*-parts are tested in the models specifications.[29] Besides that, *ARIMA* models augmented with seasonal multiplicative factors are tested as well.[30] Overall, different combinations of the indicated specifications result in 1888 St.*ARIMA* models.

The second group of models is represented by *ARIMAX* models: standard *ARIMA* models from the previous step augmented with one of the job-search-based indicators (leading indicator, *X*). Thus, the leading indicator in each model is characterized by the chosen keyword, indicator type and the time period. Table 2 displays all possible options for *X* – their combination gives in total 24 search volume indicators (see Description 2 of Appendix).

Table 2. Options for X in ARIMAX models

Keyword	Indicator type	Time period
• ***1*** • ***2***	• ***eg*** • ***g*** • ***y*** • ***q***	• current month • 1-month lag (***_lag1***) • 2-month lag (***_lag2***)

As an example: ***eg2*** means current monthly average of weekly Google indices for a keyword “вакансии”, whereas ***y1_lag1*** **–** Yandex search volume for a keyword “работа” estimated within the previous month. In further *ARIMAX* models are named by the respective indicator name. Overall, the indicators are quite different – see descriptive statistics in Table 3 of Appendix.

In addition to these St. *ARIMA* and *ARIMAX* models, the following three nonlinear models are tested: *SETAR*, *LSTAR*, *AAR* (see Description 3[31] of Appendix). These nonlinear models are typically used for the unemployment rate in the US. The nonlinear models are believed to approximate better the long-term unemployment dynamics, while *ARIMA* models – short-term dynamics (Montgomery et al., 1998)

[26] With the term “agnostic” the author means that she does not want to restrict herself from the very beginning as it is better to estimate more potential models than less.
[27] *I(0)* – level and *I(1)* – the first differences.
[28] At first, simple time-series correlograms are studied in order to identify the most significant lags and include them into the model specifications. After that the correlogram for residuals from the identified models are explored to find other parts that need to be tested. The correlogram analysis is based only on the in-sample observations in order to avoid data mining problem in forecasting period.
[29] Tested *AR*-parts:1,2,4,7,11,12,13. Tested *MA*-parts:1,2,9.
[30] *SAR(12)*, *SMA(12)*, or both.
[31] For the details see Di Narzo (2008).

Overall, there are 47200 possible St. *ARIMA* and *ARIMAX* models along with 3 nonlinear models estimated. This means that the current research contains an extremely extensive forecasting "horse-race"!

3.4. Forecasting

The current research of the unemployment rate prediction focuses on the short-term forecasting, i.e. from 1- to 3-month ahead (*h*-step forecasts). For this purpose the rolling forecasting scheme is used. At first, the historical data is split into the in-sample model estimation and out-of-sample model evaluation parts. Then based on the initial estimation sample the *h*-step forecasts are produced for the prediction sample. "The estimation sample is then rolled ahead a given increment and the estimation and prediction exercise is repeated until it is not possible to make any more *h*-step predictions" [Zivot & Wang, 2006, p.313]. The forecasts are based only on the information available at month *t*, i.e. rolling window of the recent observations is used.

Table 3 defines the time periods and the number of observations (*E* for in-sample estimation, *P* for prediction) used for the initial sample split for the following groups of models:

- Long sample models: *ARIMA* models based on solely official unemployment rate data;
- Short sample models:
    - *ARIMA* models augmented with Google Trends indices;
    - *ARIMA* models augmented with Yandex Wordstat indicators.

Table 3. Sample split

	In-sample	Out-of-sample
Long sample	[2002.1-2011.10]; *E*=118	[2011.11-2012.12]; *P*=14
Google Trends	[2004.1-2011.10]; *E*=94	
Yandex Wordstat	[2009.11-2011.10]; *E*=24	

Because of the limited data access, the sample for Yandex Wordstat models is about 4 times shorter than that for the Google Trends one.

The ranking of the models is conducted according to the Mean Squared Error (*MSE*), where the model with the lowest *MSE* is considered as the best one (#1 rank). *Out-of-sample* $MSE = \frac{1}{P-h+1}\sum_{t=t_0+h}^{T} e_t^2$, where e_t is the difference between the real data and the forecasted value. It is important to highlight that the purpose of a current study does not consist in finding the coefficients for the in-sample models, but in the out-of-sample predictive ability assessment which is mainly dependent on the general model specification and performance of the leading indicator.

3.5. Formal Assessment of the Superior Predictive Ability

Although the out-of-sample *MSE* is used for models ranking, the comparative technique based solely on the *MSE* values is quite rough as they are usually close to each other. That is why the predictive ability of the models should be accessed in a more formal way as well. The simple Diebold & Mariano (1995) test for equal forecast accuracy or Harvey et al. (1998) test for forecast encompassing do not resolve the current research problem: the need to assess the superior predictive ability among a large number of models and control for data snooping at the same time.

That is why the current study takes advantage of the Superior Predictive Ability (SPA) Test (Hansen, 2005) and Model Confidence Set (MCS) (Hansen et al., 2011). The latter one is the most advanced procedure designed to find the best models set in terms of their out-of-sample predictive ability. The existent analytical packages have a limitation for the maximum number of models to be compared: 1000. But this is already a large number of models to be evaluated – in the benchmark research of D'Amuri and Marcucci the number of tested models is twice lower.

The target 1000 models in the current research are defined in the following way:

1. 47203 models are estimated to produce 1-step forecasts;
2. Models within each type (St.*ARIMA*, *ARIMAX* for each of the 24 search volume indicators) are ranged according to the out-of-sample *MSE* for 1-step forecast;
3. About 40 of the best models within each type and 3 nonlinear models are used to comprise exactly 1000 models at the end. Thus, the initial set for MCS is balanced.

This resulted set of models is tested for the superior predictive ability in all 3 cases: 1-, 2-, 3-step forecasting. Taking into account the computational intensity of complete models set estimation, it is not reasonable to compute *MSE* for more than 47000 models again for 2- and 3-step forecasts. Thus, 2-, 3-step forecasts are produced for the target 1000 models. It is believed that 40 best models within each type in 1-step forecasting are already enough (a very wide number of models) to capture the good models for further steps.

Superior Predictive Ability (SPA) Test

SPA test is proposed by Hansen (2005) as an alternative to White's (2000) Reality Check (RC) test. This means that this test also controls for data snooping which is important when a lot of models are estimated. “Data snooping occurs when a given set of data is used more than once for purposes of inference or model selection. When such data reuse occurs, there is always the possibility that any satisfactory results obtained may simply be due to chance rather than to any merit inherent in the method yielding the results” [White, 2000, p.1115]. But at the same time SPA test is less sensitive to the inclusion of poor models than the RC test. This is a clear advantage of SPA as there for sure might be poor ones among 1000 models.

The null hypothesis can be formulated as $H_0: E(d_{i,t}) \leq 0$, $\forall i=1,\ldots,m$[32], where $d_{i,t} = L_{b,t} - L_{i,t}$ – difference in loss functions for $t=1,\ldots,P$.[33] This means that the benchmark model is not inferior to any of the alternative m models. The following models are chosen as benchmarks: Random Walk (*RW*) and *AR(1)* – these two models are very common for economic and financial forecasting.

It is important to understand that SPA test just check whether there is an evidence of a significant model against the benchmark. To identify the whole set of the best forecasting models the more complex procedure is needed.

Model Confidence Set (MCS)

MCS was proposed by Hansen (2011) as a response to the need of multiple models comparison in terms of superior predictive ability. The procedure is designed to identify a set of models which are equivalent in terms of superior predictive ability, but outperform all the competing models. This is a great step forward in formal forecast accuracy assessment. Besides, MCS does not imply any benchmark model and allows to conduct multiple comparisons for the whole set of initial models. In addition to this MCS also allows to control for data snooping: MCS consisting of a large number of models relative to the initial set says that data is not informative and the results are not reliable.

The principal design of the MCS procedure is the following:

- First of all, it’s necessary to choose M_0 – initial set of explored models $i=1,\ldots,m$ evaluated over $t=1,\ldots,P$. In a current study $M_0 = 1000$;
- For all of the initial models the relative performance is assessed: $d_{ij,t} = L_{i,t} - L_{j,t}$, $\forall i,j = 1,\ldots,m$;
- The test sets $M \subset M_0$ to be a set of equivalent superior models M*: $H_{0,M}: E(d_{ij,t}) = 0$, $\forall i,j \in M$. The test statistics is the following: $T_R = \max_{i,j\in M}|t_{ij}|$, where $t_{ij} = \bar{d}_{ij}/$

[32] The number of tested models.

[33] The test statistics is $T_P^{sm} = \max_m \frac{\bar{d}_m\sqrt{P}}{\widehat{\sigma_m}}, where\ \bar{d}_m = \frac{1}{P}\sum_{t=1}^{P} d_{i,t}, \hat{\sigma}_m^2 = \bar{d}_m\sqrt{P}$.

$\sqrt{\widehat{var}(\bar{d}_{ij})}$ and $\bar{d}_{ij} = P^{-1}\sum_{t=1}^{P} d_{ij,t}$. To get a distribution under H_0 a stationary bootstrap[34] is used.

There are the following steps in the MCS procedure:

- Initially the following proposal is tested: $M = M_0$. This means that all our initial models are equivalent in their superior predictive ability;
- If H_0 is rejected – the forecast with the largest t_{ij} is removed, i.e. the model which has worse predictive ability;
- Then the restricted set of models (without the eliminated one at the previous step) is tested again;
- The procedure is repeated until one fails to reject the null, i.e. $(1-\alpha)$-confidence set of the best out-of-sample forecasting models is obtained.

[34] See Politis & Romano (1994).

4. Main Results

4.1. Models Ranking

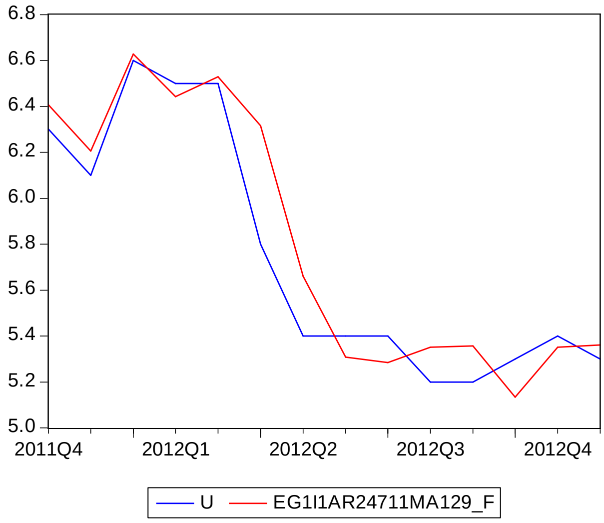
Table 4 presents the results of models ranking according to out-of-sample *MSE* for 1-step forecasts (47203 models estimated). The first column indicates the model type: St.*ARIMA* (long sample), *ARIMA* models augmented with one of the explored search-based indicators (named according to the leading indicator) and the best nonlinear model. The second and third columns show the rank of the best model within each type and the corresponding out-of-sample *MSE*, respectively.

Table 4. Models ranking, 1-step forecasting

Models	Min Rank	Min *MSE*
eg1	**1**	**0,0332**
eg2	**2**	**0,0333**
q2	**14**	**0,0423**
y2	**19**	**0,0436**
eg1_lag1	29	0,0448
eg1_lag2	139	0,0528
g2_lag2	141	0,0529
y1_lag2	178	0,0537
g2_lag1	201	0,0545
g1_lag2	221	0,0549
g2	232	0,0552
g1_lag1	243	0,0554
g1	267	0,0558
y2_lag1	271	0,0560
q1	273	0,0561
q2_lag1	287	0,0567
St. ARIMA	**313**	**0,0574**
q2_lag2	326	0,0575
eg2_lag1	444	0,0600
y2_lag2	450	0,0602
y1_lag1	571	0,0621
q1_lag1	623	0,0630
y1	747	0,0647
eg2_lag2	839	0,0661
q1_lag2	880	0,0668
Nonlinear	11505	0,1036

At the first glance, among the top there are models augmented with simple monthly average GI for both keywords (“работа”, “вакансии”) and both Yandex indicators (adjusted and simple search volumes) for keyword “вакансии”. The performance of the best standard *ARIMA* model (that is estimated over the longer sample, but is not augmented with any search volume indicator) is not impressive – 313 rank. The best nonlinear model (here it is *AAR*) shows bad performance – 11505 rank. Also the lagged search volume indicators do not work well as a whole: the best model among them has rank 29, but the others – higher than 139.

As an example, the forecast of the best model is depicted below.



Picture 1. Best model 1-step forecast

According to methodology ≈ 40 best models (1-step forecasting) within each category (24 search-based *ARIMAX* models and St.*ARIMA*) and 3 nonlinear models are selected to comprise target 1000 models. These 1000 models are used in order to produce forecasts for 2-, 3-steps as well and further formal assessment. This approached is followed for computational reasons, since estimating and forecasting with 47000 models in each case is computationally intractable, given the used IT means. The comparative models ranking within these 1000 models for 1-, 2-, 3-step forecasting is given in Table 4 of Appendix. Below there is a summary of the main results for the respective models:

- *eg1*: in top 3 models in 1-, 2-step[35] and №11 in 3-step;
- *eg2*: in top 3 models in 1-, 2-step and №7 in 3-step;
- *g1*: weak performance in 1-step, but among top 10 in 2-step and top 20 in 3-step;
- *g2*: weak performance in 1-step, but in top 10 in 2-step and №1 in 3-step;
- *y1*: bad performance in 1-step, but in top 30 in 3-step;
- *y2*: from top 20 in 1-step to top 10 in 2-, 3-step;
- *q1*: weak performance in 1-step, but in top 5 in 2-, 3-steps;
- *q2*: always in top 15;
- Lagged search indicators: bad performance in 1-step, but improvement in 2-, 3-step. The most significant results: *eg2_lag1* – top 10 (3-step), *g2_lag2* – top 5 (3-step), *q1_lag1* – top 3 (3-step), *q1_lag2* – top 5 (2-, 3-step), *q2_lag2* – top 15 (3-step), *y1_lag1* – top 20 (3-step);
- St.*ARIMA*: bad performance in 1-step, top 30 in 2-step and top 10 in 3-step;
- Nonlinear models: overall bad performance of *AAR* (the best within the nonlinear models *AAR*, *LSTAR*, *SETAR*) – rank>800 in each case.

[35] Here and further this means *h*-step forecasting.

But as it is said earlier this comparative technique is quite rough due to close *MSE* values in certain cases. Thus, a more formal forecast accuracy assessment is needed.

4.2. Forecast Accuracy Assessment: SPA test

The results of SPA test against the chosen benchmarks for the 1-step forecasting are the following:

- The null hypothesis is rejected at 5%-significance level for SPA with *RW* as a benchmark;
- The null hypothesis is rejected at 10%-sig. level for SPA with *AR(1)* as a benchmark.

This means that there is the evidence against the benchmark models at a given confidence level, i.e. at least one of the competing models does better than the benchmark.

It is reasonable to account for data-snooping problems at a 1-step forecasting as the target 1000 models are chosen based on the out-of-sample *MSE* values for this step. But as soon as SPA test does not identify all the superior models and implies a particular benchmark it is decided to conduct the MCS procedure for 2-, 3-step forecasting right away (i.e. without this additional SPA test) – it is much more advanced and controls for data-snooping by the way.

4.3. The Best Forecasting Models: MCS

The final step in the predictive ability assessment is MCS approach which helps to identify the whole set of best forecasting (out-of-sample) models. The MCS procedure for 1-step forecasting gives a set of 16 best models (out of 1000 tested). This is an impressive result: MCS contains a very little number of models which means that our dataset is informative itself. Table 5 presents the best models at 1-step forecasting.

Table 5. MCS for 1-step forecasting

ARIMA specification	*X*	*MSE*	Rank	MCS p-value
I(1) AR (2 4 7 11) MA (1 2 9)	*eg1*	0,0332	1	1*
I(1) AR (2 11) MA (2 9)	*eg2*	0,0333	2	0,9271*
I(1) AR (1 4 11) MA (1 2)	*eg1*	0,0349	3	0,9271*
I(1) AR (1 2 4 11) MA (1 2 9)	*eg1*	0,0368	4	0,9052*
I(1) AR (1 4 11) MA (1 2 9)	*eg1*	0,0392	5	0,3878*
I(1) AR (1 4 11) MA (1 9)	*eg2*	0,0398	6	0,2057*
I(1) AR (1 4 11) MA (1)	*eg2*	0,0402	7	0,2057*

I(1) AR (1 4 11) MA (1 2)	*eg2*	0,0405	8	0,2057*
I(1) AR (1 2 4 11) MA (1 2)	*eg1*	0,0407	9	0,2057*
I(1) AR (2 11) MA (9)	*eg2*	0,0415	11	0,5458*
I(1) AR (1 11) MA (1 2)	*eg1*	0,0416	12	0,1862**
I(1) AR (1 2 11) MA (9)	*eg2*	0,0422	13	0,3782*
I(1) AR (1 2 11) MA (1)	*eg2*	0,0423	15	0,2057*
I(1) AR (1 7)	*q2*	0,0424	16	0,5458*
I(1) AR (1 11) MA (1)	*eg2*	0,0424	17	0,7157*
I(1) AR (1 7) MA (9)	*eg2*	0,0436	18	0,2057*

*Note: H_0 is not rejected at *– 20%-sig.level, **– 10%-sig.level.*

The results of MCS procedure correspond to the models ranking: all the best models are search-based. Thus, models augmented with search volume indicators outperform all standard *ARIMA* models (as well as nonlinear ones) even though the latter of them are estimated over a longer sample. Among the well-performed search indicators there are simple monthly average GI for both keywords and an adjusted (normalized) Yandex search volume for keyword “вакансии”. Almost 100% of the best models are those with GI, but there is also one good model with Yandex index which is the simplest among all – $\Delta u_t = \alpha_1 \Delta u_{t-1} + \alpha_2 \Delta u_{t-7} + \gamma q2_t + \varepsilon_t$. It is important to note that all the MCS models are equivalent in their superior predictive ability. Overall, search volume data really improves the predictive ability of the standard time-series models and can be viewed as leading indicators for the unemployment rate. For this purpose both keywords can be successfully used: “работа” as the most popular one (in 37,5% of the best models) and “вакансии” (in 62,5% of the best models) which has a 3-4 times lower search volume.

The MCS procedure for 3-step forecasts gives a set of 20 best models (out of 1000 tested). This again verifies that the explored dataset is informative. Table 6 presents the best models at 3-step forecasting.

Table 6. MCS for 3-step forecasting

ARIMA specification	*X*	*MSE*	MCS p-value
I(1) AR (1 2 12 13) MA (1 9)	*g2*	0,0641	1*
I(1) AR (2 11) MA (2) SAR	*q1_lag2*	0,0652	0,8759*
I(1) AR (1 11 12 13) MA (1 2 9)	*q1_lag1*	0,0762	0,4117*
I(1) AR (2 11 12)	*g2_lag2*	0,0767	0,4117*
I(1) AR (2 11) MA (2) SAR	*q1*	0,0767	0,4117*
I(1) AR (2 11) MA (2 9)	*y2*	0,0768	0,4117*
I(1) AR (2 4 7 11) MA (2) SAR SMA	*eg2*	0,0772	0,4125*
I(1) AR (1 2 4 7 11) MA (1) SAR SMA	St. *ARIMA*	0,0774	0,4117*

I(1) AR (1 2 11) MA (1)	***eg2***	**0,0777**	**0,4117***
I(1) AR (1 4) MA (1 9)	*eg2_lag1*	0,0781	0,5197*
I(1) AR (1 2 7) MA (1 2 9)	*eg1*	0,0785	0,4117*
I(1) MA (1 2 9)	*q2_lag2*	0,0790	0,4117*
I(1) AR (2 7)	*q2*	0,0793	0,4117*
I(1) AR (1 4) MA (1 9)	*eg2*	0,0806	0,1438**
I(1) AR (1 7 11) MA (1) SAR SMA	St. *ARIMA*	0,0808	0,2469*
I(1) AR (4 12) MA (2 9)	*y1_lag1*	0,0810	0,4117*
I(1) AR (1 2 4 11) MA (1 2)	St. *ARIMA*	0,0814	0,1549**
I(1) AR (1 11 12) MA (9)	*g1*	0,0815	0,1979**
I(1) AR (1 4) MA (9)	*eg2_lag2*	0,0817	0,1438**
I(1) AR (1 2 11 12) MA (1)	*g1*	0,0828	0,2938*

*Note: H_0 is not rejected at *– 20%-sig.level, **– 10%-sig.level.*

MCS shows that 85% of best models are augmented with search volume indicators. Except for the well-performed search-based indicators from 1-step forecasting, there are also average GI based on two last weeks of the previous month and two first weeks of the current one, lagged indicators and the simple (unadjusted) Yandex search volume. Probably, the last one may evoke questions. But it's important to remember that Yandex data are available for a very short and recent period when the growth in the internet users has already slown down. As for the 1-step forecasting, both keywords perform successfully and the percentage breakdown is the same.

As for the 2-step forecast, MCS includes 92 models. This is much higher than for 1-, 3-step forecasts, but this is just 10% out of a 1000 models tested. So the data is also believed to be informative. The whole MCS is presented in Table 5 of Appendix. About 93,5% of the superior models include search volume data as a leading indicator. Among the well-performed search-based indicators there are all 4 explored types (***eg, g, q, y***) for both keywords (***1, 2***) and the lagged indicators as well.

Besides, there appeared to be the model which outperforms at all 1-, 2-, 3-step forecasts: $\Delta u_t = \alpha_1 \Delta u_{t-1} + \alpha_2 \Delta u_{t-2} + \alpha_3 \Delta u_{t-11} + \gamma eg2_t + \beta \varepsilon_{t-1} + \varepsilon_t$. This model includes simple monthly average GI for keyword “вакансии” as a leading unemployment indicator. The best models for different steps in forecasting don't have to be the same, but it's always great when such a model exists.

The general observation about the forecasting models: except for 1 model in the 3-step MCS, the MCS in all cases includes only models with *I(1)* (to remind: in order to be agnostic all specifications are tested for both *I(0)* and *I(1)*). Thus, in forecasting it's important to account for the unit root in the unemployment rate time-series.

5. Additional Checks

5.1. Classical Economic Indicator

The purpose of this section is to check the suggested search-based indicators against the classical economic one. The Industrial Production Index[36] (IPI) is tested as an alternative indicator of the unemployment rate. IPI is considered as a common proxy of GDP growth.

As it is an additional check exercise the evaluation of models is conducted for 1-step forecasts. To test IPI as a leading indicator the target 1000 models explored in the research are augmented with a 1-month lag of IPI. Thus, IPI is added to St.*ARIMA* models (IPI-based models) and St.*ARIMA* models that are already augmented with one of the search-based indicators (models containing both search volume indicator and IPI) as well. At first, the MCS procedure is applied to these 1000 models that are now augmented with 1-month lag of IPI. All the best models selected at this step represent the "new" MCS. These best models are added to our best 16 search-based models (the "previous" MCS) in order to comprise the initial set for one more MCS procedure.

Table 6 of Appendix shows all models that comprise the final MCS for this additional check exercise. As a result, MCS does not include any models augmented with IPI only (IPI-based models). There are certain models augmented with both search volume indicator and IPI, but these models were not among the well-performed search-based models in the previous section. And there is no need to care about the previously poor models. With regard to the already good search volume models, IPI does not improve them. Moreover, all our 16 best search-based models (the "previous" MCS) are still included into this new one MCS.

5.2. Falsification Test

The falsification test is proposed by D'Amuri & Marcucci (2012) as an additional check of the main results in the paper. This test would be especially interesting to those people who are skeptical about the suggested internet search-based approach for unemployment forecasting. The falsification test is conducted with the help of the Google Correlate – another interesting tool designed by Google. Google Correlate[37] helps to identify the keyword which has the highest correlation with the given time series – in our case it is the unemployment rate for the in-sample period [2004.1 – 2011.10].

---

[36] Source: "The socio-economic state of Russia", FSSS.
[37] http://www.google.com/trends/correlate

The found keyword appeared to be “6.5”. It is difficult to identify the nature of this keyword exactly, but among the related queries Google Correlate shows “Windows 6.5” and several 6.5-versions of some other software. Anyway this found keyword for sure doesn’t directly belong to the job-search activity. To conduct the falsification test the simple monthly averages GI for “6.5” are computed and added to the standard *ARIMA* models. Thus, this new index is considered as an alternative leading indicator which has the strongest correlation with the given time series, but no logical connection with the job-search. If this pseudo search-based indicator would also perform well in the out-of-sample unemployment forecasting, then the previously suggested concept (job-search queries as leading unemployment indicators in Russia) doesn’t really work well.

The best 40 models (in terms of out-of-sample *MSE* for 1-step forecasts) are chosen among the estimated simple *ARIMA* models augmented with this pseudo indicator. These models are added to our previous best 16 models in order to conduct another MCS procedure. As a result, there are no models augmented with “6.5” GI instead of our job-related search indicators: our 16 best models comprise the whole MCS. This proves again that the results of the paper are not occasional: only job-search-related indicators are relevant for forecasting the unemployment rate in Russia.

6. Conclusion

The current research tests the relevance of the search volume data for unemployment forecasting in Russia. The job-search-based indices are considered to be the first leading indicators of the unemployment rate in Russia which is released with a considerable lag. For the purpose of the research an extremely extensive forecasting "horse-race" of more than 47 000 models is conducted. Among these models there are standard time-series specifications, *ARIMA* models augmented with search volume indicators and several nonlinear models. For the first time a wide range of 24 different search-based indicators is explored: both from Yandex Wordstat & Google Trends, based on the 2 most popular keywords ("работа", "вакансии"), measured in relative or absolute scale, normalized or nonnormalized. The identification of the best forecasting models is based on the MCS approach.

Thus, the current study is a large scale comparison of Russian unemployment rate forecasting models; the first one to apply a search-based approach for unemployment forecasting in Russia; and the first worldwide comparison of search volume indicators from two different search engines: Google vs. Yandex.

The following main results regarding the out-of-sample predictive ability are highlighted:

- Success of job-search-based indicators for unemployment forecasting is robust to the choice of the main related keywords and search volume services:
  - Various Google and Yandex indicators show impressive results;
  - Both keywords are successfully used for forecasting.
- Search volume indicators are proved to be the first leading indicators for unemployment rate in Russia:
  - Models augmented with search volume indicators improve their predictive ability at 1-, 2-, 3-step forecasting;
  - In 1-step forecasting – models augmented with GI for a current month and a simple model with adjusted Yandex indicator outperform all alternatives including longer sample common models;
  - In 3-step forecasting – 50% of best models include GI and 35% - Yandex index;
  - Search-based models outperform the best analyzed macro variable models (IPI-based). And additional inclusion of IPI (with 1-month lag) does not improve the predictive ability of the best search-based models;
  - The results of the research have survived the falsification test.

- Found the model that is superior at 1-, 2-, 3-step forecasting – *ARIMA* augmented with ***eg2***.

Thus, although the sample is not large (search volume data has limited availability) the results seem very promising.

References

Askitas N., Zimmermann K. F. (2009). Google Econometrics and Unemployment Forecasting. *IZA Discussion Paper*, 4201.

Choi H., Varian H. (2009.1). Predicting the present with Google Trends. *Google technical report.*

Choi H., Varian H. (2009.2). Predicting Initial Claims for Unemployment Benefits. *Google technical report*.

Cryer, J. D., Chan K. (2008). Time series analysis with applications in R. *Springer Texts in Statistics*.

Da Z., Engelberg J., Gao P. (2011). In Search of Fundamentals. *AFA 2012 Chicago Meetings Paper*.

D'Amuri F. (2009). Predicting unemployment in short samples with internet job search query data. *MPRA working paper n. 18403*.

D'Amuri F., Marcucci J. (2009). Google it! Forecasting the US unemployment rate with a Google job search index. *Bank of Italy, mimeo*.

D'Amuri F., Marcucci J. (2012). The Predictive Power of Google Searches in Forecasting Unemployment. *Bank of Italy, mimeo*.

Diebold F. X., Mariano R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13: 253-263.

Di Narzo A. F. (2008). Nonlinear autoregressive time series models in R using tsDyn version 0.7.

Elliott G., Rothenberg T.J., Stock J.H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64: 813-836.

Enders W. (2004). Applied econometric time series. *Wiley series in probability and statistics*.

Ginsberg J., Monebbi 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.

Hansen P.R. (2005). A Test for Superior Predictive Ability. *Journal of Business & Economic Statistics*, 23(4): 365-380.

Hansen P. R, Lunde A. and Nason J. M. (2011). The Model Confidence Set, *Econometrica*, 79(2): 453–497.

Harvey D. I., Leybourne S. J., Newbold P. (1998). Tests for Forecast Encompassing. *Journal of Business & Economic Statistics*, 16: 254-259.

Montgomery, Alan L., Zarnowitz, Victor, Tsay, Ruey S. and Tiao, George C. (1998). Forecasting the U.S. Unemployment Rate, *Journal of the American Statistical Association*, 93(442): 478-493.

Politis D. N., Romano J.P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association*, 89 (428): 1303-1313.

Schmidt T., Vosen S. (2009). Forecasting Private Consumption: Survey-based Indicators vs. Google Trends. *Ruhr Economic Papers* n. 155.

White H. (2000). A Reality Check for Data Snooping. *Econometrica*, 68(5): 1097-1126.

Zivot E., Wang J. (2006). Modeling Financial Time Series with S-PLUS. *Springer-Verlag*

Data Sources

Federal State Statistics Service. http://www.gks.ru

Google Trends. http://www.google.com/trends/explore#cmpt=q

Yandex Wordstat. http://wordstat.yandex.ru

Google Correlate. http://www.google.com/trends/correlate

Yandex Statistics. http://stat.yandex.ru

Appendix: Descriptions

Description 1. ADF test

To test for the stationary of the time series there is an Augmented Dickey-Fuller test (*ADF*) for the unit root process. The widest specification of the tested model is $\Delta y_t = \alpha_0 + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + \varepsilon_{t-i}$, where $\gamma = \sum_{i=1}^{p} \alpha_i - 1$ and $\delta_i = -\sum_{j=i}^{p} \alpha_j$ ($\alpha_1, \dots, \alpha_p$ are the coefficients in the simple *p-th* order autoregressive process). The null hypothesis is that $\gamma = 0$ which means that the process is difference stationary. The alternative one states that $\gamma < 0$ and indicates the trend stationary process. The test statistic $DF = \frac{\hat{\gamma}}{\sqrt{var(\hat{\gamma})}}$ and the critical values are defined under the Dickey-Fuller distribution and for the following specifications: τ_0- no intercept and no trend, τ_μ- an intercept, τ_τ- both, an intercept and a trend.

Description 2. Explored search-based indicators

Indicator	Description
eg1	Simple monthly average of the weekly Google indices (based on *w1-w4* of m_i) for “работа”
eg2	Simple monthly average of the weekly Google indices (based on *w1-w4* of m_i) for “вакансии”
g1	Average of the weekly Google indices related to *w1-w2* of m_i & *w3-w4* of m_{i-1} for “работа”
g2	Average of the weekly Google indices related to *w1-w2* of m_i & *w3-w4* of m_{i-1} for “вакансии”
y1	Simple monthly Yandex search volume for “работа”
y2	Simple monthly Yandex search volume for “вакансии”
q1	Monthly Yandex search volume for “работа” adjusted for the number of Yandex users
q2	Monthly Yandex search volume for “вакансии” adjusted for the number of Yandex users
eg1l1	1-month lag of eg1
eg1l2	2-month lag of eg1
eg2l1	1-month lag of eg2
eg2l2	2-month lag of eg2
g1l1	1-month lag of g1
g1l2	2-month lag of g1
g2l1	1-month lag of g2
g2l2	2-month lag of g2
y1l1	1-month lag of y1
y1l2	2-month lag of y1
y2l1	1-month lag of y2
y2l2	2-month lag of y2
q1l1	1-month lag of q1
q1l2	2-month lag of q1
q2l1	1-month lag of q2
q2l2	2-month lag of q2

Note: lagged indicators are marked with “l” followed by the number of lags. Ex. eg1l1 means eg1_lag1.

Description 3: Nonlinear models

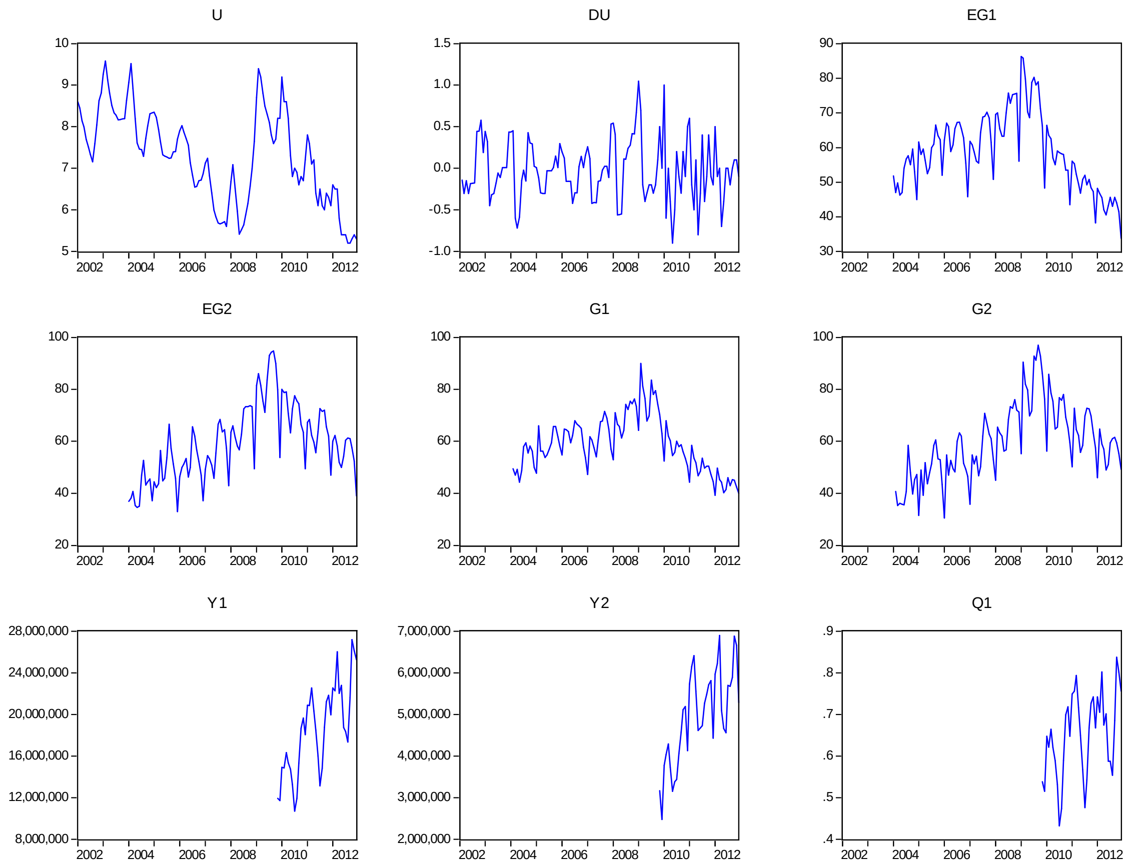
1.Self-exciting threshold autoregression (*SETAR*). $Y_t = \begin{cases} \varphi_{1,0} + \varphi_{1,1} Y_{t-1} + \dots + \varphi_{1,p_1} Y_{t-p_1} + \sigma_1 \varepsilon_t, Y_{t-d} \le c \\ \varphi_{2,0} + \varphi_{2,1} Y_{t-1} + \dots + \varphi_{2,p_2} Y_{t-p_2} + \sigma_2 \varepsilon_t, Y_{t-d} > c \end{cases}$

The 2 regimes are identified according to the value of the threshold c. Y_{t-1} is chosen as a threshold variable ($d = 1$) and two lags for each regime ($p_1 = p_2 = 2$) are adopted.

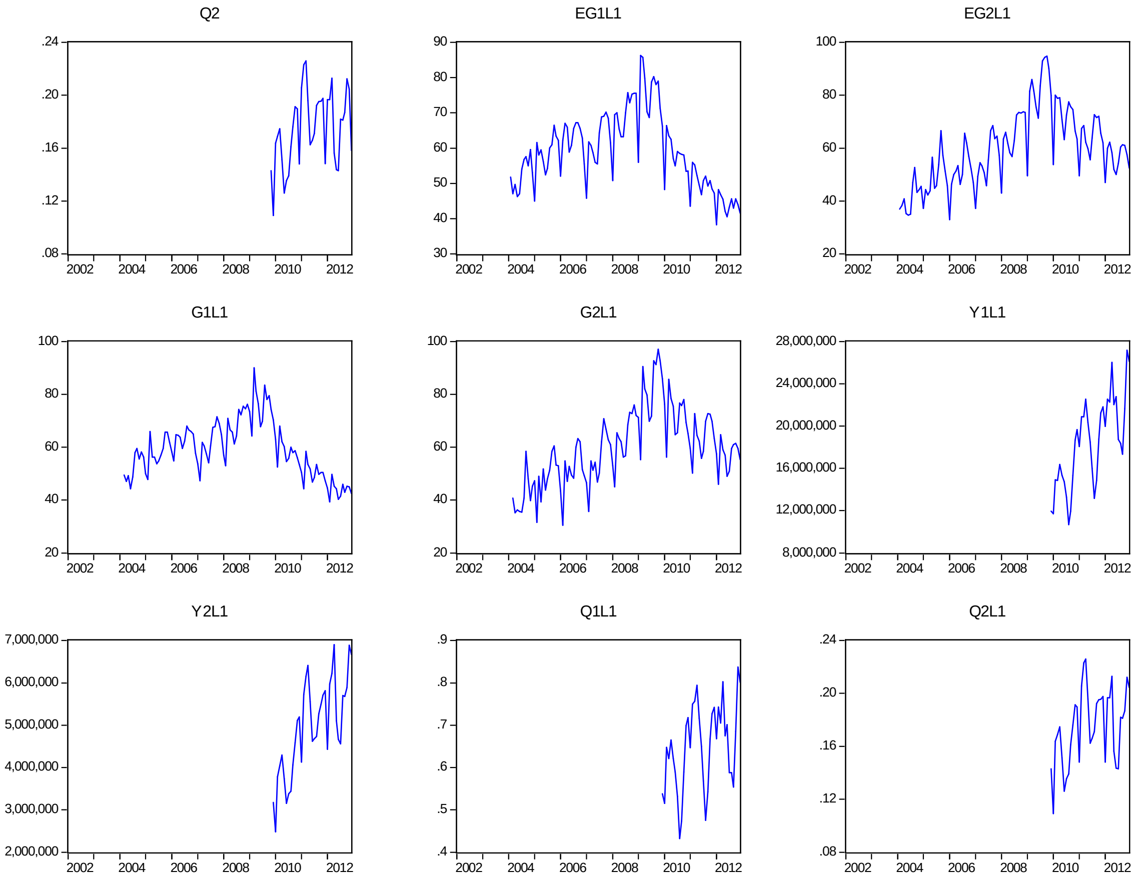
2.Logistic smooth transition autoregression (LSTAR). The adopted form is $Y_t = [\varphi_{1,0} + \varphi_{1,1} Y_{t-1} + \varphi_{1,2} Y_{t-2}][1 - G(Y_{t-1,} \gamma, c)] + [\varphi_{2,0} + \varphi_{2,1} Y_{t-1} + \varphi_{2,2} Y_{t-2}][G(Y_{t-1,} \gamma, c)] + \varepsilon_t$, where $\gamma > 0$ and $G(Y_{t-1,} \gamma, c) = \frac{1}{1+e^{-\gamma \prod_{k=1}^{K}(Y_t - c_k)}}$.

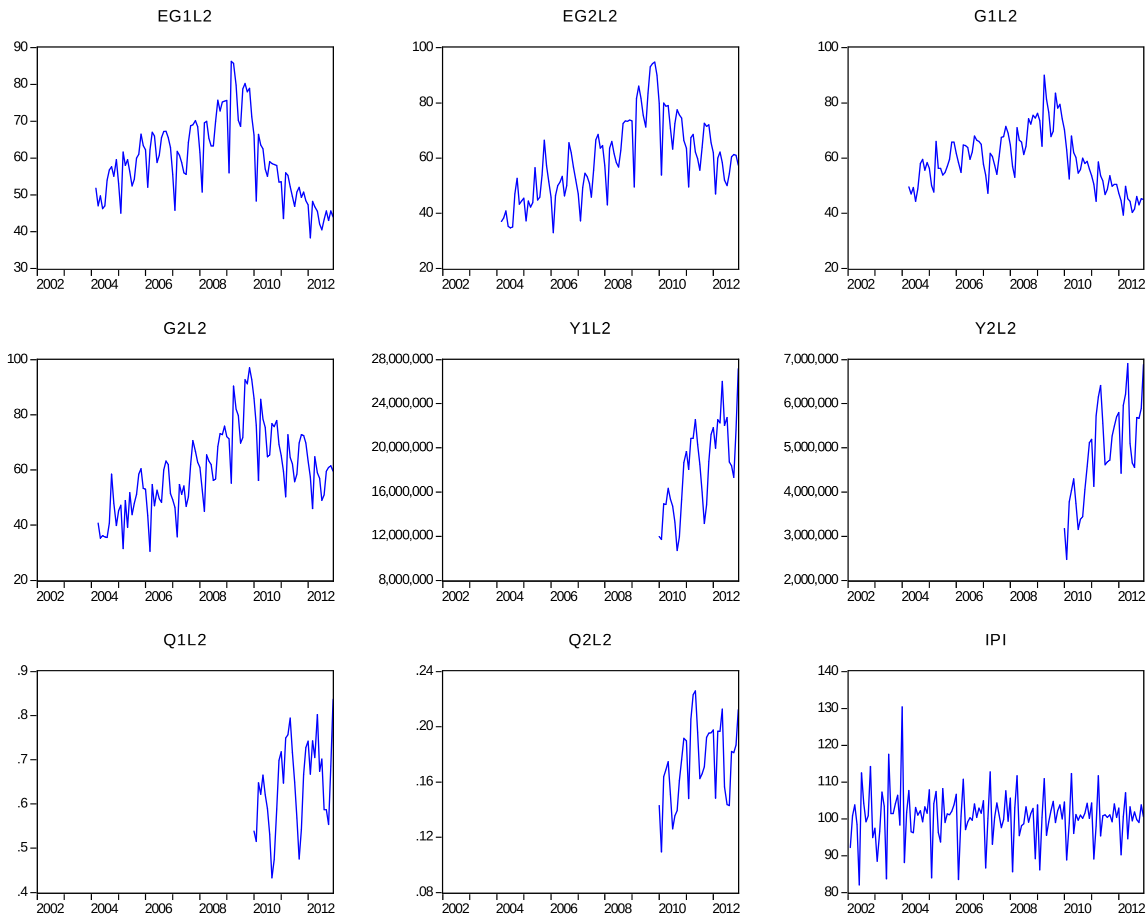
3.Additive autoregression (*AAR*) in the form $Y_t = \mu + \sum_{i=1}^{m} s_i(Y_{t-(i-1)d})$, where s_i are represented by cubic regression splines.

Appendix: Figures[38]



[38] Lagged indicators are marked with "l" followed by the number of lags. Ex. eg1l1 means eg1_lag1.


Q2
.08 .12 .16 .20 .24
2002 2004 2006 2008 2010 2012
EG1L1
30 40 50 60 70 80 90
2002 2004 2006 2008 2010 2012
EG2L1
20 40 60 80 100
2002 2004 2006 2008 2010 2012
G1L1
20 40 60 80 100
2002 2004 2006 2008 2010 2012
G2L1
20 40 60 80 100
2002 2004 2006 2008 2010 2012
Y1L1
8,000,000 12,000,000 16,000,000 20,000,000 24,000,000 28,000,000
2002 2004 2006 2008 2010 2012
Y2L1
2,000,000 3,000,000 4,000,000 5,000,000 6,000,000 7,000,000
2002 2004 2006 2008 2010 2012
Q1L1
.4 .5 .6 .7 .8 .9
2002 2004 2006 2008 2010 2012
Q2L1
.08 .12 .16 .20 .24
2002 2004 2006 2008 2010 2012


EG1L2
30 40 50 60 70 80 90
2002 2004 2006 2008 2010 2012
EG2L2
20 40 60 80 100
2002 2004 2006 2008 2010 2012
G1L2
20 40 60 80 100
2002 2004 2006 2008 2010 2012
G2L2
20 40 60 80 100
2002 2004 2006 2008 2010 2012
Y1L2
8,000,000 12,000,000 16,000,000 20,000,000 24,000,000 28,000,000
2002 2004 2006 2008 2010 2012
Y2L2
2,000,000 3,000,000 4,000,000 5,000,000 6,000,000 7,000,000
2002 2004 2006 2008 2010 2012
Q1L2
.4 .5 .6 .7 .8 .9
2002 2004 2006 2008 2010 2012
Q2L2
.08 .12 .16 .20 .24
2002 2004 2006 2008 2010 2012
IPI
80 90 100 110 120 130 140
2002 2004 2006 2008 2010 2012

Appendix: Tables

Table 1. ADF test results (long sample)

Null Hypothesis: U has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic based on SIC, MAXLAG=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.545303	0.0388
Test critical values:	1% level	-4.030157	
	5% level	-3.444756	
	10% level	-3.147221	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(U)
Method: Least Squares
Date: 05/31/13 Time: 14:08
Sample (adjusted): 2002M03 2012M12
Included observations: 130 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
U(-1)	-0.109535	0.030896	-3.545303	0.0006
D(U(-1))	0.467837	0.079071	5.916635	0.0000
C	0.915838	0.264578	3.461504	0.0007
@TREND(2002M01)	-0.001945	0.000884	-2.201092	0.0296

R-squared	0.244951	Mean dependent var	-0.024236
Adjusted R-squared	0.226974	S.D. dependent var	0.356999
S.E. of regression	0.313880	Akaike info criterion	0.550673
Sum squared resid	12.41359	Schwarz criterion	0.638905
Log likelihood	-31.79374	Hannan-Quinn criter.	0.586524
F-statistic	13.62556	Durbin-Watson stat	2.078561
Prob(F-statistic)	0.000000		

Table 2. ADF test results (short sample)

Null Hypothesis: U has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic based on SIC, MAXLAG=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.790378	0.0630
Test critical values:	1% level	-3.491928	
	5% level	-2.888411	
	10% level	-2.581176	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(U)
Method: Least Squares
Date: 05/31/13 Time: 14:11
Sample: 2004M01 2012M12
Included observations: 108

Variable	Coefficient	Std. Error	t-Statistic	Prob.
U(-1)	-0.084940	0.030440	-2.790378	0.0063
D(U(-1))	0.431788	0.087872	4.913823	0.0000
C	0.583266	0.218834	2.665332	0.0089

R-squared	0.208466	Mean dependent var	-0.030835
Adjusted R-squared	0.193389	S.D. dependent var	0.367008
S.E. of regression	0.329615	Akaike info criterion	0.645603
Sum squared resid	11.40785	Schwarz criterion	0.720107
Log likelihood	-31.86256	Hannan-Quinn criter.	0.675812
F-statistic	13.82687	Durbin-Watson stat	2.068248
Prob(F-statistic)	0.000005		

Table 3. Descriptive statistics (in-sample)

	Mean	Median	Max	Min	Std. Dev.	Skew.	Kurtosis	Jarque-Bera	Prob.	Obs.
u	7.47504	7.53389	9.57746	5.41611	1.00938	-0.1031	2.29325	2.66496	0.26382	118
eg1	61.1069	60.3750	86.2500	43.5000	9.63166	0.44799	2.71559	3.46111	0.17718	94
eg2	60.0654	59.0000	94.7500	33.0000	14.9777	0.25531	2.35340	2.65873	0.26464	94
g1	61.1586	60.0000	90.0000	44.2500	9.45676	0.5152	2.95531	4.12265	0.12728	93
g2	60.1908	59.5000	97.0000	30.5000	15.1473	0.26256	2.57115	1.78122	0.41040	93
y1	1644277	1567928	2254633	1068819	3437512	0.0823	1.87163	1.30030	0.52196	24
y2	4537200	4598329	6412968	2480783	1036329	-0.0438	2.15411	0.72322	0.69655	24
q1	0.62235	0.63437	0.79417	0.43240	0.09833	-0.1647	2.10337	0.91252	0.63364	24
q2	0.17127	0.17000	0.22589	0.10912	0.02954	-0.0829	2.49819	0.27934	0.86964	24
eg1l1	61.2435	60.7500	86.2500	43.5000	9.59184	0.44099	2.73097	3.29485	0.19254	93
eg1l2	61.3576	60.7500	86.2500	43.5000	9.58077	0.42602	2.73684	3.0483	0.21780	92
eg2l1	60.0059	58.2500	94.7500	33.000	15.0477	0.26590	2.33923	2.78783	0.24810	93
eg2l2	59.8755	57.7500	94.7500	33.0000	15.0772	0.28778	2.35183	2.88039	0.23688	92
g1l1	61.2744	60.1250	90.0000	44.2500	9.44199	0.5022	2.96315	3.87339	0.14417	92
g1l2	61.3928	60.2500	90.0000	44.2500	9.42537	0.48936	2.97275	3.63487	0.16244	91
g2l1	60.0869	59.000	97.0000	30.5000	15.1969	0.28075	2.57161	1.91204	0.38441	92
g2l2	59.9505	58.5000	97.0000	30.5000	15.2244	0.30332	2.58727	2.04131	0.36035	91
y1l1	1623416	1534897	2254633	1068819	3355756	0.14761	1.97301	1.09428	0.57860	23
y1l2	1611903	1533575	2254633	1068819	3387966	0.23443	2.00339	1.11196	0.57350	22
y2l1	4486197	4578993	6412968	2480783	1028361	0.03630	2.23850	0.56076	0.75549	23
y2l2	4440545	4436339	6412968	2480783	1028430	0.1290	2.31501	0.49118	0.78224	22
q1l1	0.61779	0.62153	0.79417	0.43240	0.09791	-0.0888	2.14805	0.72579	0.69565	23
q1l2	0.61554	0.62126	0.79417	0.43240	0.09961	-0.0250	2.09692	0.74989	0.68732	22
q2l1	0.17021	0.16890	0.22589	0.10912	0.02974	-0.0015	2.51817	0.22249	0.89471	23
q2l2	0.16908	0.16747	0.22589	0.10912	0.02992	0.08669	2.55927	0.20561	0.90230	22
"6.5	7.44680	0.00000	52.0000	0.00000	14.3268	1.82658	5.32405	36.7126	0.00000	47
IPI	100.625	101.000	130.400	82.000	7.11809	0.13720	5.58911	33.0467	0.00000	117

Note: lagged indicators are marked with “l” followed by the number of lags. Ex. eg1l1 means eg1_lag1.

Table 4. Ranking of models within target 1000

1-step			2-step			3-step		
Model[39]	min MSE	Rank	Model	min MSE	Rank	Model	min MSE	Rank
eg1	0,03324	1	eg1	0,07635	3	eg1	0,07848	11
eg1l1	0,05113	95	eg1l1	0,08955	27	eg1l1	0,08434	29
eg1l2	0,05155	109	eg1l2	0,09683	50	eg1l2	0,08858	49
eg2	0,03334	2	eg2	0,06473	1	eg2	0,07722	7
eg2l1	0,05515	213	eg2l1	0,09863	60	eg2l1	0,07811	10
eg2l2	0,05807	317	eg2l2	0,10582	117	eg2l2	0,08169	21
g1	0,05378	170	g1	0,08325	7	g1	0,08148	19
g1l1	0,05352	161	g1l1	0,10371	98	g1l1	0,08511	34
g1l2	0,0534	152	g1l2	0,10425	103	g1l2	0,08931	55
g2	0,05349	157	g2	0,08371	9	g2	0,06412	1
g2l1	0,05327	146	g2l1	0,09649	47	g2l1	0,08359	26
g2l2	0,05162	110	g2l2	0,09355	40	g2l2	0,07666	4
St. ARIMA	0,05483	204	St. ARIMA	0,08908	25	St. ARIMA	0,07744	8
q1	0,05389	173	q1	0,07523	2	q1	0,07668	5
q1l1	0,0559	249	q1l1	0,09158	32	q1l1	0,07622	3
q1l2	0,05829	327	q1l2	0,07928	5	q1l2	0,06521	2
q2	0,04228	14	q2	0,08364	8	q2	0,07931	13
q2l1	0,05407	178	q2l1	0,10731	129	q2l1	0,08905	52
q2l2	0,05455	193	q2l2	0,09998	68	q2l2	0,07897	12
y1	0,05728	284	y1	0,10313	95	y1	0,08398	27
y1l1	0,05571	238	y1l1	0,09849	58	y1l1	0,081	16
y1l2	0,06537	695	y1l2	0,17146	734	y1l2	0,23969	687
y2	0,04361	19	y2	0,07917	4	y2	0,07682	6
y2l1	0,05388	172	y2l1	0,10071	72	y2l1	0,10114	118
y2l2	0,05474	201	y2l2	0,0998	64	y2l2	0,08319	24
AAR[40]	0,1036	901	AAR	0,2923	881	AAR	0,51507	878

Table 5. MCS for 2-step forecasting

Model[41]	MSE	p-value
eg1I1AR1411MA129	0.09635	0.6184*
eg2I1AR1411MA19	0.08694	0.6184*
eg2I1AR1411MA1	0.08763	0.6184*
eg2I1AR1411MA12	0.08697	0.6184*
eg1I1AR12411MA12	0.08383	0.6184*
eg2I1AR12411MA1	0.08835	0.6184*
eg2I1AR211MA9	0.09803	0.3761*
q2I1AR211MA2	0.08629	0.6184*
eg2I1AR1211MA1	0.08749	0.6184*
eg2I1AR111MA1	0.09298	0.6184*

[39] Lagged indicators are marked with "l" followed by the number of lags. Ex. eg1l1 means eg1_lag1.
[40] The best model among nonlinear ones in all 3 cases.
[41] The models are named in the following way: eg1I1AR1411MA129 means model I(1) AR (1 4 11) MA (1 2 9) augmented with eg1.

y2I1AR111MA19	0.10285	0.5396*
eg2I1AR124711MA12	0.09043	0.6184*
eg2I1AR4711MA29	0.06473	1.0000*
y2I1AR111	0.09960	0.6184*
eg2I1AR141112MA1	0.08568	0.6184*
eg2I1AR111MA12	0.09660	0.5926*
q2I1AR14	0.10448	0.4594*
eg1I1AR1411MA1	0.09755	0.6184*
eg2I1AR24711MA2SARSMA	0.09175	0.6184*
eg2I1AR2411MA2	0.10244	0.1687**
y2I1AR17	0.09829	0.1369**
y2I1AR211	0.09849	0.3155*
eg2I1AR14MA19	0.08412	0.6383*
eg2I1AR2411MA12	0.11238	0.5396*
q2I1AR27	0.08411	0.6383*
eg1I1AR211MA19	0.09996	0.5699*
eg1I1AR124711MA19	0.09378	0.6184*
eg2I1AR1MA29	0.09693	0.3155*
y2I1AR1	0.08729	0.6184*
q2I1AR211	0.08364	0.6383*
eg2I1AR211MA12	0.11016	0.1006**
eg2I1AR27MA1	0.09632	0.5396*
eg2I1AR47MA29	0.09064	0.6184*
eg1I1AR1111213	0.09067	0.6184*
eg2I1AR1241112MA12	0.09345	0.6184*
q2I1AR11213MA29	0.08792	0.6184*
q2I1AR4MA2	0.08655	0.6184*
eg1I1AR12411MA1	0.09340	0.5536*
y2I1AR27	0.09655	0.4195*
y2I1MA1	0.10235	0.4195*
q2I1AR211MA12	0.10109	0.2363*
y2I1AR711MA2	0.09241	0.6184*
q2I1AR47MA2	0.09184	0.6184*
eg1I1AR127MA129	0.08907	0.6184*
eg1I1AR14MA9	0.10563	0.3155*
eg1l2I1AR1411MA12	0.10017	0.1638**
eg1I1AR1412MA1	0.07635	0.6383*
eg1I1AR1211MA9	0.09426	0.6184*
y2I1AR2MA9	0.08423	0.6184*
q2I1AR24	0.10024	0.5396*
eg1l1I1AR1411MA129	0.10555	0.3155*
eg1I1AR1411MA12SAR	0.10265	0.2363*
q2I1AR24MA9	0.10805	0.2388*
q1I1AR211MA2SAR	0.07523	0.6383*
y2I1AR1211MA129	0.10190	0.3807*
y2I1AR211MA29	0.07917	0.6383*
q2I1AR17MA1	0.09739	0.2662*

y2I1AR1411MA9	0.09356	0.5536*
noI1AR12411MA12	0.08908	0.6184*
noI1AR1711MA1SARSMA	0.08910	0.6184*
y1l1I1AR1111213MA129	0.09849	0.1369**
noI1AR124711MA1SARSMA	0.09302	0.5536*
g2l1I1AR124711MA12	0.09649	0.4359*
q1l1I1AR1111213MA129	0.09158	0.6184*
q1l1I1AR1211MA2	0.09862	0.4195*
eg1l1I1AR211MA19	0.08955	0.6184*
g2l2I1AR14111213	0.09995	0.5396*
eg1l2I1AR1111213	0.09683	0.5396*
noI1AR1411MA2SARSMA	0.10284	0.4626*
g2l2I1AR121112	0.10093	0.5334*
noI1AR211MA12SARSMA	0.09080	0.6184*
noI1AR14711MA2SARSMA	0.09906	0.5396*
g1l2I1AR1411MA19	0.10425	0.1064**
q1l2I1AR211MA2SAR	0.07928	0.6383*
eg1l2I1AR141112MA1	0.10137	0.3492*
g2l2I1AR1411MA1	0.10009	0.2388*
g2l2I1AR21112	0.09355	0.5396*
g2I1AR111MA12	0.08371	0.6383*
q1I1AR2MA9	0.09740	0.5536*
g1I1AR121112MA1	0.08325	0.6383*
q1l2I1AR2MA19	0.08242	0.6383*
g2l2I1AR41112MA1	0.09494	0.6184*
g2I1AR121213MA19	0.09993	0.5233*
eg1l1I1AR12411MA1	0.10157	0.1687**
g2I1AR12111213MA1	0.10112	0.3004*
eg2l1I0AR111MA19Int	0.09863	0.5536*
y2l2I1AR111MA2SAR	0.10265	0.5396*
y2l1I1MA129	0.10071	0.6184*
eg2l1I1AR14MA19	0.10205	0.5536*
q2l2I1MA129	0.09998	0.1628**
y2l2I1AR1111213MA29	0.09980	0.5396*
y1l1I1AR412MA29	0.10082	0.3062*

*Note: H_0 is not rejected at *– 20%-sig.level, **– 10%-sig.level.*

Table 6. Final MCS for IPI additional check

Model[42]	MSE	p-value
eg2I1AR2411MA9IPI	0.03960	0.7985*
eg1I1AR12411MA12IPI	0.04086	0.2093*
eg2I1AR211MA9IPI	0.04007	0.6817*
eg2I1AR2411MA29IPI	0.04179	0.2093*
eg2I1AR1111213IPI	0.04183	0.2093*
g1l2I1AR121213MA19IPI	0.04154	0.5077*

[42] The models are named in the same way as in the previous table. “IPI” at the end means that the model is also augmented with 1-month-lag IPI.

g1l2I1AR1111213MA9IPI	0.03984	0.7578*
eg2I1AR124711MA12IPI	0.04192	0.2093*
eg2I1AR24711MA12IPI	0.04270	0.2093*
eg2I1AR24711MA2SARSMAIPI	0.04383	0.1294**
eg2I1AR4711MA29IPI	0.04280	0.1294**
q2I1AR11MA1IPI	0.04260	0.1294**
q1l1I1AR2MA19IPI	0.04301	0.1294**
eg2I1AR141112MA1IPI	0.04313	0.1294**
q2I1AR211IPI	0.04418	0.1294**
g2l1I1AR1211MA19IPI	0.04319	0.1294**
eg2I1AR2411MA12IPI	0.04805	0.1294**
eg2I1AR11MA19IPI	0.04575	0.1199**
q1l1I1AR2IPI	0.04831	0.1294**
eg1I1AR24711MA129	0.03324	1.0000*
eg2I1AR211MA29	0.03334	0.9271*
eg1I1AR1411MA12	0.03490	0.9271*
eg1I1AR12411MA129	0.03683	0.9052*
eg2I1AR111MA1	0.04239	0.2697*
eg2I1AR211MA9	0.04146	0.5077*
q2I1AR17	0.04238	0.4259*
eg1I1AR1411MA129	0.03918	0.5077*
eg2I1AR1211MA9	0.04217	0.3979*
eg2I1AR1411MA19	0.03984	0.4373*
eg2I1AR1411MA1	0.04021	0.5077*
eg2I1AR1411MA12	0.04045	0.4532*
eg1I1AR12411MA12	0.04066	0.2247*
eg2I1AR1211MA1	0.04228	0.2093*
eg2I1AR17MA9	0.04360	0.2093*
eg1I1AR111MA12	0.04160	0.2093*

*Note: H_0 is not rejected at *– 20%-sig.level, **– 10%-sig.level.*