BANK OF GREECE
EUROSYSTEM

Working Paper

A study of the effect of data transformation and "linearization"on time series forecasts. A practical approach

Alexandros E. Milionis
Nikolaos G. Galanopoulos

280

JUNE 2020

BANK OF GREECE
Economic Analysis and Research Department – Special Studies Division
21, E. Venizelos Avenue
GR-102 50 Athens
Tel: +30210-320 3610
Fax: +30210-320 2432

www.bankofgreece.gr

Published by the Bank of Greece, Athens, Greece

ISSN: 2654-1912 (online)

A STUDY OF THE EFFECT OF DATA TRANSFORMATION AND «LINEARIZATION» ON TIME SERIES FORECASTS. A PRACTICAL APPROACH

Alexandros E. Milionis
Bank of Greece and University of the Aegean

Nikolaos G. Galanopoulos
Bank of Greece (Trainee) and University of the Aegean

ABSTRACT

Very often in actual macroeconomic time series there are causes that disrupt the underlying stochastic process and their treatment is known as «linearization». In addition, variance non-stationarity is in many cases also present in such series and is removed by proper data transformation. The impact of either (data transformation - linearization) on the quality of forecasts has not been adequately studied to date. This work examines their effect on univariate forecasting considering each one separately, as well as in combination, using twenty of the most important time series for the Greek economy. Empirical findings show a significant improvement in forecasts' confidence intervals, but no substantial improvement in point forecasts. Furthermore, the combined transformation-linearization procedure improves substantially the non-normality problem encountered in many macroeconomic time series.

JEL Classification: C22, C51, C53, C87

Keywords: applied time series analysis, time series «linearization», time series transformation, outliers, forecasting of macroeconomic time series.

Acknowledgements: The authors are grateful to H. Gibson and the referee for helpful comments. The views expressed in the paper are those of the authors and do not necessarily reflect those of the Bank of Greece.

Correspondence:
Alexandros E. Milionis
Bank of Greece,
Directorate of Statistics
21 E. Venizelos Avenue,
Athens GR 102 50
e-mail: amilionis@bankofgreece.gr

1. Introduction

A univariate ARIMA model is a concise quantitative summary of the internal dynamics of a time series in a linear framework. As such, it is useful for several reasons, amongst others for forecasting and model-based time series decomposition in unobserved components. This work will deal with the former and, in particular, with univariate forecasts, which usually serve either as short-term, or benchmark forecasts. However, economic time series from the real world are not usually «ready» to be used for forecasting purposes and they need to undergo some statistical preparation and pre-adjustment. This is because in time series of raw data variance non-stationarity may be present. Furthermore, very often there exist causes that disrupt the underlying stochastic process (existence of outliers, calendar effects, etc.). Their treatment is known as «linearization».

Within that line of reasoning, statistical forecasts can be made after a series itself, or some variance stabilizing transformation of it, is «linearized» according to the general framework (Kaiser and Maravall, 2001):

$$y_t = w_t' b + C_t' \eta + \sum_{j=1}^{m} \alpha_j \mu_j(B) I_t(t_j) + x_t \qquad (1)$$

Where: $y_t = f(z_t), f$ is some transformation of the raw series z_t, which may be necessary to stabilize the variance.

$b = (b_1, \ldots, b_n)$ is a vector of regression coefficients.

$w_t' = (w_{1t}, \ldots, w_{nt})$ denotes n regression or intervention variables.

C_t' denotes the matrix with columns possible calendar effect variables (e.g. trading day) and η the vector of associated coefficients.

$I_t(t_j)$ is an indicator variable for the possible presence of an outlier at period t_j.

$\mu_j(B)$ captures the transmission of the j-th effect and α_j denotes the coefficient on the outlier in the multiple regression model with m outliers.

x_t follows in general a multiplicative seasonal ARIMA$(p, d, q)(P, D, Q)_s$ model:

$$\varphi(B)\Phi(B^s)\nabla^d \nabla_s^D x_t = \theta(B)\Theta(B^s)\varepsilon_t \qquad (2)$$

where:

- $\varphi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$ is the autoregressive polynomial of order p'.
- $\theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$ is the moving average polynomial of order q'.
- $\nabla^d \equiv (1 - B)^d$ is the arithmetic difference operator of order d.
- $\nabla_s^D \equiv (1 - B)_s^D \equiv (1 - B^s)^D$ is the seasonal arithmetic difference operator of order D and seasonality s.
- $\Phi(B^s) = 1 - \Phi_1 B^s - \ldots - \Phi_P B^{P \cdot s}$ is the seasonal autoregressive polynomial of order P and seasonality s.
- $\Theta(B^s) = 1 - \Theta_1 B^s - \ldots - \Theta_Q B^{Q \cdot s}$ is the moving average polynomial of order Q and seasonality s.
- ε_t is the stochastic disturbance.

As far as variance stabilization is concerned, if the variance is functionally related to the mean, it is possible to select a transformation to stabilize the variance. Widely used transformations to tackle this problem belong to the class of the power Box and Cox transformation (Box and Cox, 1964). For example, often used transformations are given by:

$$f(z_t) = \begin{cases} z_t^{\lambda} \; if \; \lambda > 0 \\ ln z_t \;\; if \; \lambda = 0 \\ -z_t^{\lambda} \; if \; \lambda < 0 \end{cases} \qquad (3)$$

Outliers are major changes in values that especially stand out in a time series. Software-wise a well-known algorithm for the detection of outliers is that of Chen and Liu (1993) and is utilized by TSW, a software product specializing in applied time series analysis, as well as by most similar products. In the TSW[1] framework, of which use will be made in this work, three types of outliers are detected according to their effect in a time series: Additive outliers (AO), Transitory Change outliers (TC), and Level shifts (LS). In an additive outlier the value of only one observation is affected. In a transitory change the value of one observation is extremely high or low and then the size of the deviation is gradually reduced. In a level shift the level of the time series is changed. As far as the detection of outliers is concerned within the TSW

[1] TSW stands for TRAMO-SEATS for Windows, a Windows version of the DOS programmes TRAMO and SEATS (see Gomez and Maravall, 1996), and is freely available by the provider (Bank of Spain).

framework, outliers are automatically detected, classified and corrected using the Chen and Liu (1993) approach (further details in section 4).

So, there are two effects with potential influence on forecasting: transformation and «linearization», each of which separately, as well as in combination, may play an important role in time series forecasting.

At the empirical level, studies which have considered the merits of mathematical transformations on forecasting have demonstrated that a data transformation generally does not have a positive effect on forecast accuracy (Nelson and Granger 1979; Makridakis and Hibon, 1979; Makridakis et. al, 1998; Meese and Geweke, 1984).

At the theoretical level, Granger and Newbold (1976) found that such forecasts are not optimal in terms of minimization of Mean Square Forecast Error (MSFE). More specifically, for instance, for the most popular transformation, namely the logarithmic one, they showed that the minimum MSFE h-step ahead forecast is not equal to $\hat{z}_{T+h} = exp(\hat{y}_{T+h})$, as implied by the previous discussion, but is given by the expression $\hat{z}_{T+h} = exp\left(\hat{y}_{T+h} + \frac{1}{2}\sigma_h^2\right)$, where σ_h^2 is the h-step ahead forecast error variance. Pankratz and Dudley (1987), building on the work of Granger and Newbold (1976), relate the bias in using simply the inversely transformed value of the forecasts on the transformed time series (as compared to the minimum MSFE forecast) amongst others to the value of the exponent λ of the power transformation. The two most frequent transformations, namely the logarithmic and the square root ones, under certain conditions may be associated with serious biases (Pankratz and Dudley, 1987).

Regarding time series linearization, such a procedure is utilized thus far mainly as a preadjustment task for seasonal adjustment (Kaiser and Maravall, 2001), so its effect on forecasting has not been examined systematically, but only indirectly and fragmentally.[2] It is also remarked that even in studies coping with forecasting with

[2] An additional advantage of "linearizing" the outliers is that such a procedure makes the original data distribution shift closer towards normality. This is important, especially for actual economic data in view of their extreme non-normality in many cases.

transformed data the attention focuses almost exclusively on point forecasts and by and large disregards interval forecasts.

Aiming at covering this research gap in the literature the objective of this work is in fact twofold: (a) to examine the effect of «linearization» and transformation separately, as well as in combination, on both point forecasts and confidence interval forecasts; (b) to use two algorithms specializing in testing, whether or not, a transformation of the original data is necessary, namely the algorithm of TSW and the algorithm recently developed by Milionis-Galanopoulos (Milionis and Galanopoulos 2018a, 2018b, 2019) and, compare the derived results from both. Hereafter the latter will be called M-G algorithm for convenience. As a further application, we rank main economic indicators of the Greek economy in terms of statistical «forecastability». The intended approach will be practical.

The structure of the paper is as follows: In section 2 the method of Milionis and Galanopoulos for detection of variance non-stationarity and identification of possible transformation of time series data is briefly reviewed; In section 3 details of the data to be used for the empirical analysis are given; section 4 presents the empirical results and relevant comments; section 5 summarizes and concludes the paper.

2. Statistical testing approach of Milionis-Galanopoulos

In this section the statistical testing approach is briefly reviewed (for further details see Milionis and Galanopoulos 2017, 2018). At first, time series are partitioned into segments (subsamples) of equal length and for each subsample the (local) mean (LM) and the (local) standard deviation (LSD) is calculated. The local standard deviation is assumed to be functionally dependent on the local mean in a non-linear fashion as follows:

$$LSD = aLM^{\beta}e^{u} \tag{4}$$

where a, β are model parameters, e is the base of natural logarithms and u the stochastic disturbance. Model parameters a, β are estimated via Ordinary Least Squares (henceforth OLS) using the corresponding log-log model. The estimated value of β ($\hat{\beta}$) provides the necessary information for the existence (or non-existence) and the type of data transformation needed to ensure variance stationarity (e.g. for the most popular transformations, namely the log-transformation and the square root one, correspond to $\beta = 1$, and $\beta = 0.5$, respectively). This is formally stated and tested by hypothesis H_a below.

To ensure robustness with respect to a particular partition and the possible existence of outliers, as this procedure should precede the detection of outliers, the procedure is repeated for different partitions. The number of different partitions is at least equal to the number of divisors of the series' length, giving a quotient (series length over divisor) ≥ 5 and restricting the size of subsamples to be ≥ 5.[3] Robustness is formally stated and tested by hypothesis H_b below.

Finally, the previous steps are repeated with the transformed data. The purpose of this last step is to test, whether or not, the suggested transformation is sufficient to stabilize the series variance. This is formally stated and tested by hypothesis H_c below.

2.1 Notation

- Index (k) indicates the ascending number of a subsample in a partition.
- Index (j) indicates the ascending number of the particular partition, $j = 1,2, \dots, j_{max}$.
- Index i_j represents the maximum value of k (number of subsamples) in partition j.
- N is the total length (size) of the initial time series.
- n_{ij} represents the size of subsamples in partition j.
- $\hat{\beta}_j$ is the estimate of the exponent β using subsamples derived from partition with ascending number j.
- $\hat{u}_{jk}$, $\hat{\varepsilon}_j$, $\hat{u}^*_{jk}$ are independent regression residuals.
- An asterisk (*) over a symbol denotes the corresponding transformed data, or the corresponding parameter estimate derived from the transformed data.

[3] Five (5) was selected as a reasonable lower limit for both the size of a subsample, as well as the number of subsamples in any partition of the original series.

2.2 Equations

- $i_j = (N/n_{ij})$, if (N/n_{ij}) is an integer; $n_{ij} \geq 5$,
- $i_j = int(N/n_{ij}) + 1$, if (N/n_{ij}) is not an integer, $n_{ij} \geq 5$ and the residual of the division is ≥ 5,
- $i_j = int(N/n_{ij})$, if (N/n_{ij}) is not an integer, $n_{ij} \geq 5$ and the residual of the division is < 5,
- $\hat{\beta}_j$ is estimated for each partition j, $j = 1,2, \dots, j_{max}$ via OLS from the model (First stage regression):

$$ln(LSD_{jk}) = ln(\alpha_j) + \hat{\beta}_j\, ln(LM_{jk}) + \hat{u}_{jk} \qquad (5)$$

- $\hat{\beta}$ is estimated via OLS as the constant term of the model (Second Stage regression):

$$\hat{\beta}_j = \hat{\beta} + \hat{d}j + \hat{\varepsilon}_j \qquad (6)$$

- Model using the transformed data (Third stage regression):

$$ln(LSD^*_{jk}) = ln(\alpha^*_j) + \hat{\beta}^*_j\, ln(LM^*_{jk}) + \hat{u}^*_{jk} \qquad (7)$$

2.3 Statistical hypotheses and comments

Briefly, the crucial statistical hypotheses are the following (more details and comments in Milionis and Galanopoulos (2018)):

1) H_a: $\beta_j = 0\ \forall j$ (or at least the majority of $\beta_j s$).

2) Robustness test. H_b: $d = 0$.

3) Under-transformation test. H_c: $\beta^*_j = 0\ \forall j$.

3. Data

The data set comprises some of the key macroeconomic time series for the Greek economy: GDP; unemployment; prices of consumer goods and services; monetary aggregates; and balance of payments statistics. In the balance of payments data, a distinction is made between imports – exports of all goods and imports - exports of goods without fuels and ships, as according to a study by the Bank of

Greece (Oikonomou et al., 2010), the dependence of the Greek economy on oil was high and was rising at the fastest pace among the euro area countries. Furthermore, in the same study it is noted that the balance of payment of sea transport is significant in the Greek balance of current transactions (4% of GPD in 2008) and will be considered separately from other BOP transactions on transport.

Of the twenty economic time series that are used, nineteen are monthly time series, one is a quarterly time series (sources: Bank of Greece (BoG) and Hellenic Statistical Authority (ELSTAT)). The list of series used is given in Table 1.

The monthly time series data cover the period from January 2004 to August 2018 and consist of one hundred and seventy-six (176) observations, except for the Industrial Production Index, where available data exists only from January 2010 to August 2018 (104 observations). The quarterly time series is GDP and covers the period from 1995 Quarter 1 to 2018 Quarter 3 (95 observations).

4. Empirical results and comments

As mention in section 1, the effect of transformation and the effect of linearization on forecasting will be examined at first separately and, subsequently, in combination.

The aforementioned effects will be studied on a comparative basis utilizing both the TSW and the M-G algorithms. In that way, together with those effects themselves, it will also be possible to evaluate the performance of each methodology.

Typical statistics to be used for the assessment of the quality of point forecasts are the following:

i) the Mean Absolute Percentage Error (MAPE) statistic given by:

$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{A_t - F_t}{A_t} \right|$,

ii) the Mean Square Forecast Error (MSFE) statistic given by:

$MSFE = \frac{1}{n} \sum_{t=1}^{n} (A_t - F_t)^2$, and

iii) the Mean Absolute Error (MAE) statistic given by:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|A_t - F_t|,$$

where A_t is the actual value and F_t is the forecast value.

Furthermore, as far as interval forecasts are concerned, the width of the forecast confidence interval (CI), or the forecast standard error, will be considered. The best forecast will be that with the minimum value of each time utilized statistic from those mentioned above.

4.1 The effect of «linearization» on forecast quality

We investigate how time series linearization affects the quality of both point forecasts and confidence interval forecasts. Here linearization is not considered in its generality, as described in section 1, but is confined to outliers' detection and adjustment[4]. Table 2 presents the number of best forecasts with data in levels[5]. Table 3 presents the number of best forecasts with log-transformed data indistinguishable for all time series, as it is often the case when using log-transformed data in econometric analyses. It is noted that in one time series with levels (that of unemployment expressed in percentages) and one time series in logs (that of industrial production index) no outliers were detected; hence, the total number of time series considered reduced to nineteen for each case.

From the results of Tables 2 and 3 it is apparent that when outliers are considered forecasts are better in every single case in terms of the width of the forecast confidence interval, as the forecast standard error is smaller in all the 19 cases. By contrast, there is no obvious improvement in point forecasts, as the number of best forecasts seems not to be affected by the treatment of outliers. One point that should be stressed is that such results are in general dependent upon the specific characteristics of each time series, especially upon whether an outlier is

[4] Calendar effects such as the trading day and leap effects were considered and indeed were found to be statistically significant on some occasions. All series were properly adjusted for calendar effects before further analysis.

[5] As the usual practice, the original data set was split up into the estimation sample, over which model estimation is performed, and the holdout (test) sample. In all cases the holdout sample for ex-post forecasts was originally set to twelve time periods for the monthly series and ten time periods for GDP. Presented results are based on one-step-ahead forecasts. Results for longer forecasting periods (not presented) are very similar and are available from the authors.

among the first, the middle or the last observations. For this reason, it would be desirable to use a large number of time series, so as to draw conclusions of indisputable confidence. Although the number of time series used in this work is relatively small (though comparable to that of other similar works, see for instance Nelson and Granger, 1976) the evidence that lead to the above conclusions, in particular regarding the width of the forecast confidence interval, is so convincing that it really stands far and beyond any concern related to micronumerosity.

4.2 The effect of Level Shifts (LS) on forecast quality

After a level shift outlier, all observations subsequent to the outlier move to a new level. In contrast to additive and transitory outliers, a level shift outlier reflects a major change in the stochastic process and affects many observations, as it has a permanent effect. For this reason, the case with only additive and transitory outliers (i.e. excluding level shifts) is considered, and their effect on forecasts is examined separately, performing the same analysis as in section 4.1. It is noted that this time only fifteen time series are considered, i.e. those including all types of outliers. The results are presented in Tables 4 and 5.

From the results below it is obvious that there is a trade-off: forecast standard errors are better with all outliers included and, conversely, point forecasts are better excluding level shifts. Given the influence of the level shift outliers it would be desirable to possibly consider stricter identification criteria for them relative to the other two types of outliers. It is noted that in existing statistical software specializing on time series analysis there is no such option and a purpose-built routine has to be created by the researcher.

4.3 The effect of a data transformation on forecast quality

As far as the effect of data transformation is concerned, first it is important to note that the effect of a transformation can take place in two ways: 1) direct, i.e. due to the transformation itself and 2) indirect, as it has been shown that a data transformation in general affects the number and the character of outliers in a time series (Milionis 2003; Milionis, 2004; Milionis and Galanopoulos, 2018).

The possible need for a data transformation of the original time series data will be examined using both the algorithms of TSW and M-G. Furthermore, each decision derived from the Milionis and Galanopoulos methodology and the corresponding one derived from the TSW routine will be compared. Once a decision about the proper data transformation is made, TSW will be used for both cases for further analysis on statistical forecasting.

Regarding the arithmetic values of the exponent λ (Equation 3) and the closely related parameter $\hat{\beta}$, as estimated by Equation (6) of section 2, for practical purposes Makridakis et al. (1998) mention that it is of no merit in using arithmetic values with several decimal points, as nearby values will produce very similar results. Simple arithmetic values of λ are easier to interpret and, hence, more meaningful.

In line with that argument, nearby arithmetic values of $\hat{\beta}$ will be grouped together, so as to create two sub-logarithmic transformations, namely the square root and cubic root ones, the logarithmic itself, and one over-logarithmic, namely the negative inverse transformation. More specifically the grouping is as follows (it is noted that no case with negative value of $\hat{\beta}$ was encountered):

(a) $\hat{\beta}$ not statistically significant, then $\lambda = 1$;
(b) $\hat{\beta}$ statistically significant and $0 \ < \hat{\beta} \leq \hat{\beta} + 1.96se(\hat{\beta})$ or 0.65, whichever is lower, then $\lambda = 1/2$;
(c) $\hat{\beta} - 1.96se(\hat{\beta})$, or 0.65, whichever is higher $< \hat{\beta} \leq \hat{\beta} + 1.96se(\hat{\beta})$ or 0.80, whichever is lower, then $\lambda = 1/3$;
(d) $\hat{\beta} - 1.96se(\hat{\beta})$ or 0.80, whichever is higher $< \hat{\beta} \leq \hat{\beta} + 1.96se(\hat{\beta})$, then $\lambda = 0$;
(e) $\hat{\beta} - 1.96se\left(\hat{\beta}\right) > 1$, then $\lambda = -1$.

Table 6 presents the results on the decision about, transforming or not, the original time series data, and the type of suggested transformation for the M-G case, based on the $\hat{\beta}$ value. From these results it is evident that, according to the M-G algorithm, no transformation of the original data is suggested in fifteen out of the twenty cases, the negative inverse transformation is suggested in four cases and the logarithmic transformation in only one case.

The same series were reanalyzed following the standard TSW procedure. It is noted that the only alternatives available with TSW are either the log-transformation, or no transformation. Using the TSW routine for these twenty cases, TSW suggested the logarithmic transformation of the original data for eighteen cases. It is remarkable that only for the two series of unemployment TSW suggests no transformation, as does the Milionis Galanopoulos method as well, for the particular two series. It should be stressed, however, that as shown by Milionis and Galanopoulos (2018a, 2018b), the TSW routine is biased towards (over)suggesting the log-transformation.

The possible effect of transforming a time series on forecasting quality is examined in Tables 7a and 7b. From the results below it can be concluded that point forecasts with either transformation method are slightly better than with no transformation in terms of MAPE and MAE, but not in terms of MSFE. As already explained, forecasts on transformed variables are not optimal in terms of MSFE. On the other hand, confidence interval forecasts are shorter in four of the five cases using transformations with the M-G approach. In contrast this happens in only eight out of the eighteen cases using the TSW approach. Though it seems that the M-G approach leads to shorter confidence interval forecasts, obviously there are very few cases available. Further empirical evidence with a larger dataset is needed so as to draw safer conclusions.

4.4 The combined effect of linearization and data transformation

The results of the examination of the forecasting performance combining both linearization and data transformation are presented in Table 8a and 8b. The conclusion that can be drawn is that, by and large, the combined effect does not lead to better point forecasts but leads to improved confidence interval forecasts with better performance for the M-G approach. The conclusion about the forecast confidence interval is reasonable and, to a large extent, expected, as with the transformation of the original time series data and the adjustment for outliers the process variance is reduced. It is possible to exploit this reduction in obtaining forecasts with increased confidence.

Appendix 1 presents the ARIMA models for the benchmark model and the combination of Milionis-Galanopoulos variance stabilizing method - linearization. It is noted that the differences in the ARIMA models for the time series where no transformation was needed should be attributed to the existence of outliers adjusted by linearization.

4.5 Sensitivity analysis - Outliers (dependence of outlier detection on the parameter $\boldsymbol{\tau}$)

Let $\hat{Y}_{T+1}/\Phi_T$ denote the optimal one-step-ahead linear forecast of Y_{T+1} given the information set Φ_T, which includes information up to time T, $e_{T+1} = Y_{T+1} - \hat{Y}_{T+1}/\Phi_T$ denote the associated forecast error, and $\sigma_{T+1}^2 = [Y_{T+1} - \hat{Y}_{T+1}/\Phi_T]^2$ denote the associated variance. The observation Y_{T+1} is considered as an outlier if the null Hypothesis: H_0: $e_{T+1} = 0$ is rejected. The appropriate statistic to test H_0 is: $\tau = \frac{e_{T+1}}{\sigma_{T+1}}$.

However, theory cannot predict the critical value of $\boldsymbol{\tau}$ above which the corresponding observation can be considered as an outlier. A usual practice is to relate the critical value of $\boldsymbol{\tau}$ to the length of a time series. The default values of TSW for $\boldsymbol{\tau}$ are presented in Table 9[6]. In the course of our experimentation it was observed that outlier detection (as well as ARIMA models for the linearized-transformed series), is sensitive to the value of parameter $\boldsymbol{\tau}$. In order to examine, whether or not, the critical $\boldsymbol{\tau}$ values could have any noticeable effect on our final conclusions, as an alternative set of critical values for $\boldsymbol{\tau}$ we use those suggested by Fischer and Planas (2000), who examined a very large number of time series. Their critical values for $\boldsymbol{\tau}$ were set at 3.5, 3.7 and 4.0 for series lengths of less than 130 observations, between 131 and 180, and more than 180 observations, respectively.

[6] In the TSW framework the subroutine TERROR is designed especially for outlier detection. Incoming data volume in institutions like EUROSTAT, ECB, OECD, NCBs, NSOs etc. may be enormous. Such data may be contaminated by errors of various types and origins. Using TERROR is a convenient, yet formal way to spot aberrant observations (outliers). It is highly possible that if erroneous data do exist, they will be included in the set of observations characterized as outliers by TERROR, hence, in a second stage, their possible identification is focused exclusively on that data set. In this work we used the first stage only.

The comparison of the results based on default critical τ values, as well as on Fischer – Planas suggestions are presented in Table 10, while the detected outliers for each time series and each set of values for the parameter τ are presented in Appendix 2. Looking at Appendix 2 it is observed that the detection of outliers is indeed sensitive even to the examined small changes in the value of τ. For instance, in the time series of unemployment (in thousands) there are no detected outliers when the TSW default values of τ are used, while there are two detected additive outliers (in May 2013 and July 2015) when using the Fischer – Planas suggested value of τ**.** On the other hand, however, from the results of Table 10, it is apparent that using the Fisher and Planas critical values for τ leads to mixed results regarding the effect on forecast quality. In short, there is only very weak evidence of improvement using the Fischer – Planas recommendations[7].

4.6 Evaluation of models' forecasting performance

The skill of a forecast can be assessed by comparing the relative proximity of both the forecast and a benchmark to the observations. The presence of a benchmark makes it easier to compare approaches and for this reason a benchmark is proposed to establish a common ground for comparison. In the present case an obvious benchmark is to use the univariate ARIMA forecasts of the twenty-time series described in section 3, non-linearized and non-transformed. These benchmark forecasts will be used together with the forecasts from the TSW and M-G approaches as three alternatives, the performances of which are to be evaluated and compared. Forecast evaluation for each model will be based on both point and interval forecasts. A simple and transparent ad-hoc approach will be used for this purpose. More specifically, for the point forecasts for each time series and for each model an arithmetic value is assigned in ascending order based on the corresponding value of the MSFE statistic (i.e. 1 for the minimum MSFE value, 2 for the mid- MSFE value, 3 for the maximum MSFE value). Then, adding up the arithmetic values for all series for a particular model their sum will represent the performance of the model.

[7] Indeed, setting the Fisher –Planas critical values instead of the default ones, the results are identical regarding those of Table 8a, while the results pertaining to those of Table 8b they are identical in terms of the standard error, and 8/18 for MAPE, MAD and MSFE with TSW, as compared to 7/18 using the default critical values).

Models will be ranked according to the value of the corresponding sum. The model with the lowest sum will be considered the best. For interval forecasts the same procedure will be followed replacing the value of the MSFE statistic with the value of the corresponding standard error around the point forecasts.

From the above, it is apparent that use will be made repetitively of the same data set. This could potentially make the whole process susceptible to the data snooping trap (White, 2000)[8]. Such a case is quite common for instance in developing trading strategies in financial markets. A well-known tool used by the developers of such strategies is the so-called reality check with its refinements and extensions (White 2000; Romano and Wolf, 2005; Hansen et. al 2011). In the present case however, the possibility that the forecasting performance of one of the three models to be used (namely the benchmark model, TSW and M-G) is superior to that of the other two simply due to chance is reduced by the fact that the number of models is much lower than the number of time series (three against twenty). Therefore it is unlikely that one and the same model would obtain superior performance in all, or at least in most of the twenty time series, just as a result of pure chance. For this reason the usage of the reality check, bearing in mind also its weaknesses (Hansen, 2005; Hansen et. al 2011), is not deemed as necessary.

The results are shown in Tables 11 and 12[9] and more detailed results are quoted in Appendix 3. It is clarified that both the TSW and M-G transformation approaches are coupled with the outlier detection-adjustment approach.

From the results of Tables 11 and 12 it is evident that the performance of neither TSW nor M-G approach for point forecasts is better than that of the benchmark model (as a matter of fact both are slightly worse). By contrast, for the forecast confidence intervals M-G is better than TSW and the benchmark model. Furthermore, TSW outperforms the benchmark model. A rather crude way to procced to an overall evaluation of the three models is to add up their performances

[8] Halbert White in his seminal paper (White, 2000) states that: "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. This problem is practically unavoidable in the analysis of time series data…"

[9] If for two models the value of MSFE or SE is exactly the same, the mid-point will be used for both.

in the two categories (i.e. point and interval forecasts). The addition gives the values of 90, 78 and 72 for the benchmark model, TSW and M-G respectively, which means that both TSW and M-G perform clearly better than the benchmark model and further the performance of M-G is better than that of TSW.

Nelson and Granger (1979) utilized the Box-Cox transformations, amongst others, for forecasting purposes (point forecasts) using twenty-one actual economic time series. As they failed to get superior forecasts, they rather pessimistically concluded that it is not worthwhile using these transformations bearing in mind the extra inconvenience, effort and cost. Their point of view was subsequently adopted by other researchers as well, as already mentioned in the introductory section. Lest one gets too disappointed, despite the fact that cost and effort are much lower nowadays than what they were at that time, we further note that Nelson and Granger did not associate forecasts on transformed time series with an outlier detection-adjustment approach. Furthermore, their conclusion was based only on point forecasts, disregarding forecast confidence intervals. The latter are of much importance especially in cases where the focus is on best-worst forecast scenarios. For instance, such is the case with actuarial time series on mortality rates, which may be used in the construction of pension plans. As shown above, the combination of transformation-linearization leads to narrower forecast confidence intervals.

It should also be stressed that neither in the existing research works thus far, nor in the present one, the treatment of the effect of data transformation on time series forecasting is complete for the simple reason that no work extends the analysis in a bivariate (in general multivariate) framework. Indeed, the existence of variance non-stationarity in time series could potentially contaminate the pre-whitening process (for details about the pre-whitening process see Box and Jenkins, 1976), and consequently the sample cross correlation function, so it will mask the true dynamic relationship between two series, one of which is supposed to be the leading indicator, thus affecting negatively the conditional (in this case) forecasts.

4.7 The shift towards normality

Another serious concern expressed by Nelson and Granger (1979) was the fact that the problem of acute non-normal distributions found in most macroeconomic time series was corrected only little by their use of data transformations. Table 13 presents the results for the Jarque-Bera statistic for normality (Jarque and Bera, 1987). This statistic is distributed as chi-square with two degrees of freedom. An asterisk next to an arithmetic value in Table 13 indicates a rejection of the null hypothesis of normality at the 5% significance level (critical value = 5.99).

The results of Table 13 allow, again, for a more optimistic view, inasmuch as it is evident that there is a general shift towards normality from the benchmark model to either TSW or M-G transformation-linearization procedure. The phenomenon on some occasions is really very pronounced indeed (e.g. in the series of M1 and Balance of Payments–transport-payments). This allows for computational algorithms such as maximum likelihood estimation, as well as standard statistical tests, to be legitimately employed with transformed-linearized data.

4.8 Statistical benchmark forecasting

Seizing the opportunity of the above analysis, it is useful to assess the relative forecastability of the twenty time series of the Greek economy. Here forecastability will be perceived in both point and confidence interval forecasts. For the former the MAPE statistic will be employed. For the latter the percentage standard error statistic will be introduced as the mean average of the ratio of the forecasts' standard error over the corresponding actual value, so as to make forecasts of the various series mutually comparable. In all cases one-step-ahead forecasts will be performed[10]. It is stressed that although these forecasts are technically perfectly acceptable, nevertheless they are purely statistical, hence, a-theoretical, and they can only serve as benchmark forecasts in order to evaluate the merit of more structural econometric forecasts. Tables 14-15 show the results in descending order in terms of statistical forecastability according to the Milionis - Galanopoulos method.

From the results of the Tables 14 – 15, it is observed that although there are many similarities in the two Tables, the ordering is not exactly the same. For this reason, the linear correlation coefficient between orderings based on MSFE and the

[10] Two-(or more)-step-ahead forecasts are available from the authors.

percentage standard error was used. In all cases there is a strong positive correlation. The method of Milionis-Galanopoulos has the highest correlation, while TSW has the lowest.

From Tables 14 and 15 it is also noticeable that the BOP series are the least forecastable in both Tables. Regarding imports-exports it is noted that the former are less forecastable than the latter. Furthermore, imports-exports excluding fuels and ships are clearly more forecastable than imports-exports including them. This justifies, here from the statistics point of view, the separate recording and usage of the imports-exports without the inclusion of fuels and ships for further economic analysis.

5. Summary and conclusions

This work dealt with the effect of data transformation for variance stabilization and linearization for outlier adjustment on the quality of univariate time series forecasts, using two methods for data transformation, those of TSW and Milionis Galanopoulos, and following a practical approach.

There is clear evidence that linearization improves the forecasts' confidence intervals and some evidence that data transformation acts likewise. However, the effect of the later needs to be reconfirmed using a larger dataset. In contrast no evidence was found that either transformation or linearization lead to better point forecasts. The combined effect of transformation-linearization improves further the forecasts confidence intervals, but worsens point forecasts. Furthermore, there is also evidence that the overall forecasting performance using the Milionis Galanopoulos data transformation procedure is somewhat better than that using the data transformation procedure of TSW.

Last, but certainly not least, the combined transformation-linearization procedure improves substantially the non-normality problem encountered in many macroeconomic time series.

References

Box, G. E. P. and Cox, D. R. (1964). 'An Analysis of Transformations', *Journal of the Royal Statistical Society*, Vol. 26, pp. 211-243.

Box, G. E. P. and Jenkins, G. M. (1976). *Time series analysis: forecasting and control*, Holden – Day.

Chen, C. and Liu, L.-M. (1993). 'Forecasting Time Series With Outliers', *Journal of Forecasting*, Vol. 12, pp. 13-35.

Clark, T.E. and West, K.D. (2007). 'Approximately normal tests for equal predictive accuracy in nested models', *Journal of Econometrics*, Vol. 138, pp. 291-311.

Fischer, B. and Planas, C. (2000). 'Large scale fitting of regression models with ARIMA errors', *Journal of Official Statistics*, Vol. 16, pp. 173-184.

Gomez, V. and Maravall, A. (1996). 'Programmes SEATS and TRAMO: instructions for the User' Bank of Spain Working Paper 9628.

Granger, C. W. J. and Newbold, P. (1976). 'Forecasting Transformed Series', *Journal of the Royal Statistical Society*, Vol. 38, pp. 189 – 203.

Hansen, P. R. (2005). 'A test for superior predictive ability', *Journal of Business & Economic Statistics*, Vol. 23, pp. 365-380.

Hansen, P. R., Lunde, A. and Nason, J. M. (2011). 'The model confidence set', *Econometrica*, Vol. 79(2), pp. 453-497.

Jarque, C. and Bera, A. (1987). 'Efficient tests for normality homoscedasticity and serial independence of regression residuals', *Econometric Letters*, Vol. 6, pp. 255-259.

Kaiser, R. and Maravall, A. (2001). *Measuring business cycles in economic time series*, Springer.

Makridakis, S. and Hibon, M. (1979). 'Accuracy of Forecasting: An Empirical Investigation', *Journal of the Royal Statistical Society,* Vol. 142.

Makridakis, S., Wheelwright, S. C. and Hyndman, R. J. (1998). *Forecasting Methods and Applications*, 3rd ed., USA, NJ, Wiley.

Meese, R. and Geweke, J. (1984). 'A comparison of autoregressive univariate forecasting procedures for macroeconomic time series', *Journal of Business & Economic Statistics,* Vol. 2, pp. 191-200.

Milionis, A. E. (2003). 'Modeling economic time series in the presence of variance non-stationarity', Bank of Greece Working Paper No 7.

Milionis, A. E. (2004). ‘The importance of Variance Stationarity in Economic Time Series Modelling; a Practical Approach’, *Applied Financial Economics*, Vol. 14, pp. 265-278.

Milionis, A. E. and Galanopoulos, N. G. (2017). ‘Time series with interdependent level and second moment: testing, consequences for modelling; preliminary results’, *Proceedings of the 30th Panhellenic Statistics conference*, pp. 109-123.

Milionis, A. E. and Galanopoulos, N. G. (2018a). ‘Time series with interdependent level and second moment: statistical testing and applications with Greek external trade and simulated data’, Bank of Greece Working Paper 246.

Milionis, A. E. and Galanopoulos, N. G. (2018b). ‘Time series with interdependent level and second moment: testing, consequences for modelling; applications (part II)’, *Proceedings of the 31th Panhellenic Statistics conference,* pp. 302-313.

Milionis, A. E. and Galanopoulos, N. G. (2019). ‘Forecasting economic time series in the presence of variance instability and outliers’, *Theoretic Economic Letters*, Vol. 9, pp. 2940-2964.

Nelson, H. L. and Granger, C. W. J. (1979). ‘Experience with using the Box-Cox transformation when forecasting economic time series’, *Journal of Econometrics*, Vol. 10, pp. 57-69.

Oikonomou, G., Sampethai, I. and Symigiannis, G. eds, (2010). *Greece’s balance of payments: reasons for imbalances and policy proposals* (in Greek).

Pankratz, A. and Dudley, U. (1987). ‘Forecasts of power-transformed series’, *Journal of Forecasting,* Vol. 6, pp. 239-248.

Romano, J. P and Wolf, M (2005). Stepwise multiple testing as formalized data snooping, *Econometrica*, Vol. 4, 1237-1282.

White, H. (2000). ‘A reality check for data snooping’, *Econometrica*, Vol. 68, pp. 1097-1126.

Tables

Table 1. *Data*

Time Series	Observation frequency	Source
Gross Domestic Product (GDP)	Quarterly	ELSTAT
Industrial Production Index (IPI)	Monthly	ELSTAT
Consumer Price Index (CPI)	Monthly	ELSTAT
Harmonised Index of Consumer Prices (HICP)	Monthly	ELSTAT
Unemployment – thousands	Monthly	ELSTAT
Unemployment – percentage	Monthly	ELSTAT
Retail sales	Monthly	ELSTAT
M1	Monthly	BoG
M2	Monthly	BoG
M3	Monthly	BoG
Balance of payments (BOP) – Transport – Payments	Monthly	BoG
Balance of payments (BOP) – Transport – Receipts	Monthly	BoG
Balance of payments (BOP) – Travelling – Payments	Monthly	BoG
Balance of payments (BOP) – Travelling – Receipts	Monthly	BoG
Balance of payments (BOP) – Sea transport – Payments	Monthly	BoG
Balance of payments (BOP) – Sea transport – Receipts	Monthly	BoG
Exports of Goods	Monthly	BoG
Exports of Goods without fuels and ships	Monthly	BoG
Imports of Goods	Monthly	BoG
Imports of Goods without fuels and ships	Monthly	BoG

Table 2. *Summary table - Number of best forecasts (levels)*

Point Forecasts	With detected Outliers	Without detection of Outliers
MAPE	10/19	9/19
MSFE	8/19	11/19
MAE	9/19	10/19
Interval Forecasts	**With detected Outliers**	**Without detection of Outliers**
Forecast Standard Error (SE)	19/19	0/19

Table 3. *Summary table – Number of best forecasts (log-data)*

Point Forecasts	With detected Outliers	Without detection of Outliers
MAPE	9/19	10/19
MSFE	11/19	8/19
MAE	10/19	9/19
Interval Forecasts	**With detected Outliers**	**Without detection of Outliers**
Forecast Standard Error (SE)	19/19	0/19

Table 4. *Summary table - Number of best forecasts (levels)*

Point Forecasts	All Outliers	Outliers without LS
MAPE	6/15	9/15
MSFE	5/15	10/15
MAE	6/15	9/15
Interval Forecasts	**All Outliers**	**Outliers without LS**
Forecast Standard Error (SE)	14/15	1/15

Table 5. *Summary table - Number of best forecasts (log-data)*

Point Forecasts	All Outliers	Outliers without LS
MAPE	6/15	9/15
MSFE	5/15	10/15
MAE	6/15	9/15
Interval Forecasts	**All Outliers**	**Outliers without LS**
Forecast Standard Error (SE)	13/15	2/15

Table 6. *Decision about data transformation*
(NSS stands for not statistically significant)

TIME SERIES	**METHOD OF TRANSFORMATION**		
	LOG-LEVEL PRETEST (Output from TSW)	**M-G**	
		Value of $\widehat{\beta}$	**TRANSFORMATION**
Gross Domestic Product (GDP)	SSlevels/(SSlog*Gmean(levels)^2)= 1.1380170 LOGS ARE SELECTED	NSS	Levels
Consumer Price Index (CPI)	SSlevels/(SSlog*Gmean(levels)^2)= 1.0781750 LOGS ARE SELECTED	NSS	Levels
Harmonised Index of Consumer Prices (HICP)	SSlevels/(SSlog*Gmean(levels)^2)= 1.0954455 LOGS ARE SELECTED	NSS	Levels
Industrial Production Index (IPI)	SSlevels/(SSlog*Gmean(levels)^2)= 1.0224433 LOGS ARE SELECTED	NSS	Levels
Unemployment – thousands	SSlevels/(SSlog*Gmean(levels)^2)= 0.87725642 LEVELS ARE SELECTED	NSS	Levels
Unemployment – percentage	SSlevels/(SSlog*Gmean(levels)^2)= 0.86356273 LEVELS ARE SELECTED	NSS	Levels
Retail sales	SSlevels/(SSlog*Gmean(levels)^2)= 1.2755206 LOGS ARE SELECTED	1.88	Negative Inverse
M1	SSlevels/(SSlog*Gmean(levels)^2)= 0.98393639 LOGS ARE SELECTED	NSS	Levels
M2	SSlevels/(SSlog*Gmean(levels)^2)= 1.0714007 LOGS ARE SELECTED	NSS	Levels
M3	SSlevels/(SSlog*Gmean(levels)^2)= 1.0422806 LOGS ARE SELECTED	NSS	Levels
Balance of payments (BOP) – Transport – Payments	SSlevels/(SSlog*Gmean(levels)^2)= 1.0351033 LOGS ARE SELECTED	NSS	Levels
Balance of payments (BOP) – Transport – Receipts	SSlevels/(SSlog*Gmean(levels)^2)= 1.1641507 LOGS ARE SELECTED	2.07	Negative Inverse
Balance of payments (BOP) – Travelling – Payments	SSlevels/(SSlog*Gmean(levels)^2)= 1.1509645 LOGS ARE SELECTED	NSS	Levels
Balance of payments (BOP) – Travelling – Receipts	SSlevels/(SSlog*Gmean(levels)^2)= 4.3996100 LOGS ARE SELECTED	0.83	Logarithmic
Balance of payments (BOP) – Sea transport – Payments	SSlevels/(SSlog*Gmean(levels)^2)= 0.98863656 LOGS ARE SELECTED	NSS	Levels
Balance of payments (BOP) – Sea transport – Receipts	SSlevels/(SSlog*Gmean(levels)^2)= 1.1948699 LOGS ARE SELECTED	2.72	Negative Inverse
Exports of Goods	SSlevels/(SSlog*Gmean(levels)^2)= 0.95751942 LOGS ARE SELECTED	NSS	Levels
Exports of Goods without fuels and ships	SSlevels/(SSlog*Gmean(levels)^2)= 0.96487436 LOGS ARE SELECTED	NSS	Levels
Imports of Goods	SSlevels/(SSlog*Gmean(levels)^2)= 1.1118244 LOGS ARE SELECTED	NSS	Levels
Imports of Goods without fuels and ships	SSlevels/(SSlog*Gmean(levels)^2)= 1.2957291 LOGS ARE SELECTED	2.17	Negative Inverse

Table 7a. *Summary table - Number of best forecasts (M-G versus benchmark)*

Point Forecasts	**M-G - no outliers**	**Levels-no outliers (Benchmark)**
MAPE	3/5	2/5
MSFE	2/5	3/5
MAE	3/5	2/5
Interval Forecasts	**M-G - no outliers**	**Levels-no outliers (Benchmark)**
Forecast Standard Error (SE)	4/5	1/5

Table 7b. *Summary table - Number of best forecasts (TSW versus benchmark)*

Point Forecasts	**TSW - no outliers**	**Levels-no outliers (Benchmark)**
MAPE	9/18	9/18
MSFE	7/18	11/18
MAE	9/18	9/18
Interval Forecasts	**TSW - no outliers**	**Levels-no outliers (Benchmark)**
Forecast Standard Error (SE)	8/18	10/18

Table 8a. *Summary table - Number of best forecasts (M-G versus benchmark)*

Point Forecasts	**M-G - All outliers**	**Levels-no outliers (Benchmark)**
MAPE	2/5	3/5
MSFE	2/5	3/5
MAE	2/5	3/5
Interval Forecasts	**M-G - All outliers**	**Levels-no outliers (Benchmark)**
Forecast Standard Error (SE)	4/5	1/5

Table 8b. *Summary table - Number of best forecasts (TSW versus benchmark)*

Point Forecasts	**TSW - All outliers**	**Levels-no outliers (Benchmark)**
MAPE	8/18	10/18
MSFE	8/18	10/18

MAE	8/18	10/18
Interval Forecasts	**TSW - All outliers**	**Levels-no outliers (Benchmark)**
Forecast Standard Error (SE)	12/18	6/18

Table 9. *Critical values for τ*

Observations	**Default values for τ in TSW**
164	0.358E+01
165 – 168	0.359E+01
169 – 172	0.360E+01
173 – 175	0.361E+01

Table 10. *Results based on Fischer – Planas recommendations*

Time series	**Improvement of forecast quality**	**Same forecast quality**	**Deterioration of forecast quality**
Gross Domestic Product (GDP)		MAPE, MSFE, MAE, SE (TSW)	MAPE, MSFE, MAE, SE (M-G)
Consumer Price Index (CPI)	MAPE, MSFE, MAE (TSW)	MAPE, MSFE, MAE, SE (M-G)	SE (TSW)
Harmonised Index of Consumer Prices (HICP)	MAPE, MSFE, MAE (M-G) MAPE, MSFE, MAE, SE (TSW),		SE (M-G)
Industrial Production Index (IPI)	MAPE, MSFE, MAE (M-G)	MAPE, MSFE, MAE, SE (TSW)	SE (M-G)
Unemployment – thousands	MSFE (M-G, and TSW)		MAPE, MAE, SE (M-G, and TSW)
Unemployment – percentage	MAPE, MSFE, MAE (M-G, and TSW)	SE (M-G, and TSW)	
Retail sales			MAPE, MSFE, MAE, SE (M-G, and TSW)
M1		MAPE, MSFE, MAE, SE (M-G, and TSW)	
M2	MAPE, MAE (M-G)		MAPE, MSFE, MAE, SE (TSW), MSFE, SE (M-G)
M3	MAPE, MSFE, MAE (M-G)		MAPE, MSFE, MAE, SE (TSW), SE (M-G)
Balance of	MAPE, MSFE, MAE		MAPE, MSFE, MAE,

payments (BOP) – Transport – Payments	(TSW)		SE (M-G) SE (TSW)
Balance of payments (BOP) – Transport – Receipts	MAPE, MSFE, MAE, SE (M-G, and TSW)		
Balance of payments (BOP) – Travelling – Payments		MAPE, MAE, SE (TSW)	MSFE (TSW) MAPE, MSFE, MAE, SE (M-G)
Balance of payments (BOP) – Travelling – Receipts		SE (M-G, and TSW)	MAPE, MSFE, MAE (M-G, and TSW)
Balance of payments (BOP) – Sea transport – Payments	MAPE, MSFE, MAE (M-G, and TSW)		SE (M-G, and TSW)
Balance of payments (BOP) – Sea transport – Receipts	MAPE, MSFE, MAE (M-G, and TSW)		SE (M-G, and TSW)
Exports of Goods	MAPE, MSFE, MAE (M-G, and TSW)		SE (M-G, and TSW)
Exports of Goods without fuels and ships			MAPE, MSFE, MAE, SE (M-G, and TSW)
Imports of Goods		MAPE, MSFE, MAE, SE (M-G)	MAPE, MSFE, MAE, SE (TSW)
Imports of Goods without fuels and ships	MAPE, MSFE, MAE (TSW)	MAPE, MSFE, MAE, SE (M-G)	SE (TSW)

Table 11. *Ranking of forecasting performance according to MSFE (point forecasts)*

Time series	Benchmark	TSW	M-G
Consumer Price Index (CPI)	1	2	3
Harmonised Index of Consumer Prices (HICP)	1	3	2
M3	1	3	2
M2	2	3	1
Gross Domestic Product (GDP)	3	1	2

M1	3	1	2
Industrial Production Index (IPI)	1	3	2
Retail sales	2	3	1
Unemployment – thousands	1	2.5	2.5
Balance of payments (BOP) – Transport – Receipts	1	2	3
Balance of payments (BOP) – Sea transport – Receipts	1	3	2
Unemployment – percentage	2	2	2
Balance of payments (BOP) – Transport – Payments	1	3	2
Imports of Goods without fuels and ships	2	1	3
Exports of Goods without fuels and ships	3	1	2
Exports of Goods	2	1	3
Balance of payments (BOP) – Sea transport – Payments	3	1	2
Imports of Goods	3	1	2
Balance of payments (BOP) – Travelling – Receipts	3	1.5	1.5
Balance of payments (BOP) – Travelling – Payments	2	1	3
SUM	**38**	**39**	**43**

Table 12. *Ranking of forecasting performance according to SE (interval forecasts)*

Time series	Benchmark	TSW	M-G
Harmonised Index of Consumer Prices (HICP)	3	2	1
Consumer Price Index (CPI)	3	2	1
M1	3	2	1
M3	3	1	2
M2	3	1	2

Gross Domestic Product (GDP)	3	1	2
Unemployment – percentage	2	2	2
Industrial Production Index (IPI)	2	3	1
Unemployment – thousands	3	1.5	1.5
Exports of Goods without fuels and ships	2	3	1
Retail sales	3	2	1
Exports of Goods	2	3	1
Balance of payments (BOP) – Transport – Receipts	3	2	1
Balance of payments (BOP) – Transport – Payments	2	3	1
Balance of payments (BOP) – Sea transport – Receipts	3	2	1
Imports of Goods without fuels and ships	3	1	2
Balance of payments (BOP) – Sea transport – Payments	2	3	1
Imports of Goods	3	1	2
Balance of payments (BOP) – Travelling – Payments	3	1	2
Balance of payments (BOP) – Travelling – Receipts	1	2.5	2.5
SUM	**52**	**39**	**29**

Table 13. *Values of the Jarque –Bera statistic (statistically significant values are indicated with an asterisk)*

Time series	Benchmark	TSW	M-G
Consumer Price Index (CPI)	2.889	0.999	0.423
Harmonised Index of Consumer Prices (HICP)	6.289*	5.850	8.263*
M3	19.78*	14.72*	12.44*
M2	16.71*	7.519*	16.31*
Gross Domestic Product (GDP)	14.17*	0.541	3.699
M1	152.6*	2.879	3.597
Industrial Production Index (IPI)	1.118	0.996	1.118
Retail sales	2.328	0.771	0.545
Unemployment – thousands	9.745*	7.613*	7.613*
Balance of payments (BOP) – Transport – Receipts	5.526	0.563	3.587
Balance of payments (BOP) – Sea transport – Receipts	7.447*	0.9231E-01	0.7904E-01
Unemployment – percentage	7.584*	7.584*	7.584*
Balance of payments (BOP) – Transport – Payments	137.5*	1.651	5.289
Imports of Goods without fuels and ships	7.938*	0.928	0.266
Exports of Goods without fuels and ships	28.26*	0.473	0.593
Exports of Goods	0.404	0.380	0.180
Balance of payments (BOP) – Sea transport – Payments	210.5*	4.633	4.598

Imports of Goods	1.589	4.115	0.924
Balance of payments (BOP) – Travelling – Receipts	15.31*	4.696	4.696
Balance of payments (BOP) – Travelling – Payments	2.286	1.978	2.013

Table 14. *Forecastability of main economic indicators. Greece. Point forecasts*

MAPE			
Time series	**Benchmark**	**TSW**	**M-G**
Harmonised Index of Consumer Prices (HICP)	0.241%	0.257%	0.252%
Consumer Price Index (CPI)	0.238%	0.289%	0.328%
M3	0.561%	0.653%	0.661%
M2	0.625%	0.650%	0.697%
M1	0.786%	0.652%	0.706%
Gross Domestic Product (GDP)	0.760%	0.729%	0.745%
Industrial Production Index (IPI)	1.011%	1.111%	1.019%
Retail sales	1.424%	1.666%	1.458%
Unemployment – thousands	2.170%	2.608%	2.608%
Balance of payments (BOP) – Sea transport – Receipts	2.789%	2.902%	2.640%
Unemployment – percentage	2.917%	2.917%	2.917%
Balance of payments (BOP) – Transport – Payments	2.929%	3.309%	3.134%
Exports of Goods without fuels and ships	3.718%	2.517%	3.208%
Imports of Goods without fuels and ships	3.032%	3.026%	3.258%
Balance of payments (BOP) – Transport – Receipts	2.748%	2.922%	3.835%
Balance of payments (BOP) – Sea transport – Payments	5.515%	3.883%	5.077%
Exports of Goods	5.021%	4.238%	5.129%
Imports of Goods	6.027%	5.750%	5.705%
Balance of payments (BOP) – Travelling – Receipts	12.194%	7.729%	7.729%
Balance of payments (BOP) – Travelling – Payments	12.553%	11.775%	13.994%

Table 15. *Forecastability of main economic indicators. Greece. Interval forecasts*

Percentage Standard Error			
Time series	**Benchmark**	**TSW**	**M-G**
Consumer Price Index (CPI)	0.454%	0.443%	0.419%
Harmonised Index of Consumer Prices (HICP)	0.439%	0.427%	0.423%
M1	1.290%	1.272%	1.142%
M3	1.451%	1.013%	1.180%
M2	1.455%	1.090%	1.219%
Gross Domestic Product (GDP)	2.145%	1.745%	1.855%
Unemployment – thousands	2.809%	2.572%	2.572%
Unemployment – percentage	2.737%	2.737%	2.737%
Retail sales	5.110%	3.636%	2.803%
Industrial Production Index (IPI)	2.808%	2.890%	2.805%
Exports of Goods without fuels and ships	4.582%	5.483%	4.008%
Balance of payments (BOP) – Transport – Payments	5.817%	6.101%	4.469%
Exports of Goods	5.495%	7.552%	5.294%
Balance of payments (BOP) – Sea transport – Receipts	6.514%	5.394%	5.332%
Imports of Goods without fuels and ships	7.151%	4.833%	5.528%
Balance of payments (BOP) – Sea transport – Payments	7.234%	7.779%	5.807%
Balance of payments (BOP) – Transport – Receipts	5.565%	5.348%	6.317%
Balance of payments (BOP) – Travelling – Receipts	24.967%	7.679%	7.679%
Imports of Goods	8.170%	7.559%	7.700%
Balance of payments (BOP) – Travelling – Payments	17.607%	14.157%	16.151%

Table 16. *Linear correlation coefficient between MSFE and percentage SE ordering*

Method	Correlation
Benchmark	95.40%
TSW	93.05%
M-G	97.23%

Appendix 1

Univariate ARIMA models with and without transformation-linearization

Time series	Benchmark	M-G
Gross Domestic Product (GDP)	ARIMA (0,1,1) $(0,1,1)_4$ $\nabla\nabla_4 Y_t = (1+0.118B)(1+0.425B^4)\varepsilon_t$	ARIMA (1,0,0) $(1,1,0)_4$ $(1+0.953B)(1-0.335B^4)\nabla_4 Y_t = \varepsilon_t$
Industrial Production Index (IPI)	ARIMA (2,0,0) $(0,1,1)_{12}$ $(1+0.379B+0.547B^2)\nabla_{12}Y_t =$ $(1+0.950B^{12})\varepsilon_t$	ARIMA (2,0,0) $(0,1,1)_{12}$ $(1+0.379B+0.547B^2)\nabla_{12}Y_t =$ $(1+0.950B^{12})\varepsilon_t$
Consumer Price Index (CPI)	ARIMA (0,1,0) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1+0.260B^{12})\varepsilon_t$	ARIMA (1,1,0) $(0,1,0)_{12}$ $(1+0.134B)\nabla\nabla_{12}Y_t = \varepsilon_t$
Harmonised Index of Consumer Prices (HICP)	ARIMA (0,1,0) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1+0.347B^{12})\varepsilon_t$	ARIMA (0,1,0) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1+0.287B^{12})\varepsilon_t$
Unemployment – thousands	ARIMA (3,2,1) $(0,1,1)_{12}$ $(1-0.681B-0.674B^2+0.062B^3)\nabla^2\nabla_{12}Y_t$ $= (1+0.758B)(1+0.938B^{12})\varepsilon_t$	ARIMA (3,2,1) $(0,1,1)_{12}$ $\begin{pmatrix}1-1.153B-1.123B^2\\ -0.340B^3\end{pmatrix}\nabla^2\nabla_{12}Y_t =$ $(1+0.614B)(1+0.907B^{12})\varepsilon_t$
Unemployment – percentage	ARIMA (2,2,1) $(0,1,1)_{12}$ $(1-0.726B-0.715B^2)\nabla^2\nabla_{12}Y_t =$ $(1+0.734B)(1+0.816B^{12})\varepsilon_t$	ARIMA (2,2,1) $(0,1,1)_{12}$ $(1-0.726B-0.715B^2)\nabla^2\nabla_{12}Y_t =$ $(1+0.734B)(1+0.816B^{12})\varepsilon_t$
Retail sales	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1+0.364B)(1+0.566B^{12})\varepsilon_t$	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}\frac{-1}{Y_t} =$ $(1+0.316B)(1+0.586B^{12})\varepsilon_t$
M1	ARIMA (0,2,1) $(0,1,1)_{12}$ $\nabla^2\nabla_{12}Y_t = (1+0.838B)(1+0.682B^{12})\varepsilon_t$	ARIMA (3,1,0) $(0,1,1)_{12}$ $\begin{pmatrix}1+0.007B+0.156B^2\\ +0.420B^3\end{pmatrix}\nabla\nabla_{12}Y_t =$ $(1+0.668B^{12})\varepsilon_t$
M2	ARIMA (3,1,0) $(1,0,1)_{12}$ $(1+0.328B+0.040B^2+0.307B^3)$ $(1+0.868B^{12})\nabla Y_t = (1+0.656B^{12})\varepsilon_t$	ARIMA (3,1,0) $(0,1,1)_{12}$ $\begin{pmatrix}1+0.375B+0.087B^2\\ +0.311B^3\end{pmatrix}\nabla\nabla_{12}Y_t$ $= (1+0.822B^{12})\varepsilon_t$
M3	ARIMA (0,2,1) $(0,1,1)_{12}$ $\nabla^2\nabla_{12}Y_t = (1+0.695B)(1+0.824B^{12})\varepsilon_t$	ARIMA (0,2,1) $(0,1,1)_{12}$ $\nabla^2\nabla_{12}Y_t =$ $(1+0.683B)(1+0.859B^{12})\varepsilon_t$
Balance of payments (BOP) – Transport – Payments	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1+0.188B)(1+0.847B^{12})\varepsilon_t$	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t =$ $(1+0.312B)(1+0.859B^{12})\varepsilon_t$
Balance of payments (BOP) – Transport – Receipts	ARIMA (3,1,1) $(0,1,1)_{12}$ $(1-0.393B-0.050B^2+0.264B^3)\nabla\nabla_{12}Y_t$ $= (1-0.288B)(1+0.950B^{12})\varepsilon_t$	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}\frac{-1}{Y_t} =$ $(1+0.180B)(1+0.829B^{12})\varepsilon_t$
Balance of payments (BOP) – Travelling – Payments	ARIMA (1,0,0) $(0,1,1)_{12}$ $(1+0.339B)\nabla_{12}Y_t = (1+0.506B^{12})\varepsilon_t$	ARIMA (1,0,0) $(1,0,0)_{12}$ $(1+0.314B)(1+0.613B^{12})Y_t = \varepsilon_t$
Balance of payments (BOP) – Travelling – Receipts	ARIMA (1,0,0) $(1,1,0)_{12}$ $(1+0.731B)(1-0.371B^{12})\nabla_{12}Y_t = \varepsilon_t$	ARIMA (1,0,0) $(1,1,0)_{12}$ $(1+0.598B)(1-0.422B^{12})\nabla_{12}\mathit{ln}Y_t$ $= \varepsilon_t$
Balance of payments (BOP) – Sea transport – Payments	ARIMA (0,1,1) $(0,0,0)_{12}$ $\nabla Y_t = (1+0202B)\varepsilon_t$	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t =$ $(1+0.290B)(1+0.846B^{12})\varepsilon_t$
Balance of payments (BOP) – Sea transport – Receipts	ARIMA (3,1,1) $(0,1,1)_{12}$ $(1-0.388B-0.020B^2+0.281B^3)\nabla\nabla_{12}Y_t$ $= (1-0.262B)(1+0.848B^{12})\varepsilon_t$	ARIMA (3,1,1) $(0,1,1)_{12}$ $\begin{pmatrix}1-0.533B-0.125B^2\\ +0.217B^3\end{pmatrix}\nabla\nabla_{12}\frac{-1}{Y_t} =$ $(1-0.414B)(1+0.826B^{12})\varepsilon_t$
Exports of Goods	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1+0.414B)(1+0.950B^{12})\varepsilon_t$	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t =$ $(1+0.387B)(1+0.950B^{12})\varepsilon_t$

Exports of Goods without fuels and ships	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1 + 0.485B)(1 + 0.922B^{12})\varepsilon_t$	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1 + 0.587B)(1 + 0.785B^{12})\varepsilon_t$
Imports of Goods	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1 + 0.495B)(1 + 0.950B^{12})\varepsilon_t$	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1 + 0.655B)(1 + 0.934B^{12})\varepsilon_t$
Imports of Goods without fuels and ships	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}Y_t = (1 + 0.434B)(1 + 0.785B^{12})\varepsilon_t$	ARIMA (0,1,1) $(0,1,1)_{12}$ $\nabla\nabla_{12}\frac{-1}{Y_t} = (1 + 0.313B)(1 + 0.799B^{12})\varepsilon_t$

Appendix 2

Detected outliers for the different values of parameter τ (the first number indicate the serial number of the corresponding observation, then follows the type of outlier and within the parentheses the corresponding month, or quarter, and year)

Time series	τ -default TSW critical values	τ -Fisher-Planas
Gross Domestic Product (GDP)	OUTLIERS: 57 AO (1 2009)	OUTLIERS: 57 AO (1 2009)
Industrial Production Index (IPI)	OUTLIERS: NO OUTLIERS DETECTED	OUTLIERS: NO OUTLIERS DETECTED
Consumer Price Index (CPI)	OUTLIERS: 93 LS (9 2011), 119 AO (11 2013)	OUTLIERS: 119 AO (11 2013)
Harmonised Index of Consumer Prices (HICP)	OUTLIERS: 119 AO (11 2013)	OUTLIERS: 119 AO (11 2013)
Unemployment – thousands	OUTLIERS: 60 LS (12 2008), 95 LS (11 2011), 98 TC (2 2012), 126 LS (6 2014), 148 TC (4 2016), 156 TC (12 2016)	OUTLIERS: 60 LS (12 2008), 95 LS (11 2011), 98 TC (2 2012), 126 LS (6 2014), 148 TC (4 2016), 156 TC (12 2016)
Unemployment – percentage	OUTLIERS: NO OUTLIERS DETECTED	OUTLIERS: NO OUTLIERS DETECTED
Unemployment – thousands	OUTLIERS: NO OUTLIERS DETECTED	OUTLIERS: 113 AO (5 2013), 139 AO (7 2015)
M1	OUTLIERS: 139 LS (7 2015)	OUTLIERS: 139 LS (7 2015)
M2	OUTLIERS: 100 AO (4 2012), 102 AO (6 2012), 133 LS (1 2015), 138 TC (6 2015)	OUTLIERS: 100 AO (4 2012), 102 AO (6 2012), 138 TC (6 2015)
M3	OUTLIERS: 100 AO (4 2012), 102 AO (6 2012), 133 LS (1 2015), 138 TC (6 2015)	OUTLIERS: 100 AO (4 2012), 102 AO (6 2012), 133 LS (1 2015), 138 TC (6 2015)
Balance of payments (BOP) – Transport – Payments	OUTLIERS: 60 LS (12 2008), 133 LS (1 2015)	OUTLIERS: 60 LS (12 2008), 133 LS (1 2015)
Balance of payments (BOP) – Transport – Receipts	OUTLIERS: 59 LS (11 2008)	OUTLIERS: 36 TC (12 2006), 59 LS (11 2008)
Balance of payments	OUTLIERS: 92 AO (8 2011)	OUTLIERS: 92 AO (8 2011)

(BOP) – Travelling – Payments		
Balance of payments (BOP) – Travelling – Receipts	OUTLIERS: 2 AO (2 2004), 113 LS (5 2013)	OUTLIERS: 2 AO (2 2004), 113 LS (5 2013)
Balance of payments (BOP) – Sea transport – Payments	OUTLIERS: 60 LS (12 2008), 113 LS (5 2013), 133 LS (1 2015)	OUTLIERS: 59 LS (11 2008), 113 LS (5 2013), 133 LS (1 2015)
Balance of payments (BOP) – Sea transport – Receipts	OUTLIERS: 36 TC (12 2006), 59 LS (11 2008), 129 AO (9 2014)	OUTLIERS: 36 TC (12 2006), 59 LS (11 2008), 129 AO (9 2014)
Exports of Goods	OUTLIERS: 81 AO (9 2010)	NO OUTLIERS DETECTED
Exports of Goods without fuels and ships	OUTLIERS: 60 LS (12 2008), 81 AO (9 2010)	OUTLIERS: 60 LS (12 2008), 81 AO (9 2010)
Imports of Goods	OUTLIERS: NO OUTLIERS DETECTED	OUTLIERS: NO OUTLIERS DETECTED
Imports of Goods without fuels and ships	OUTLIERS: 39 AO (3 2007), 59 LS (11 2008), 75 AO (3 2010), 82 AO (10 2010), 139 TC (7 2015)	OUTLIERS: 39 AO (3 2007), 59 LS (11 2008), 75 AO (3 2010), 82 AO (10 2010), 139 TC (7 2015)

Appendix 3

Detailed forecast quality statistics: MSFE, MAE and Forecast Standard Error

Time series	Benchmark	TSW	M-G
Consumer Price Index (CPI)	0.074 0.241 0.461	0.123 0.293 0.450	0.163 0.332 0.426
Harmonised Index of Consumer Prices (HICP)	0.100 0.255 0.466	0.114 0.272 0.452	0.107 0.267 0.448
M3	1,551,599 947 2,448	2,166,840 1,100 1,709	1,947,577 1,116 1,989
M2	2,410,091 1,048 2,440	2,479,304 1,094 1,831	2,224,942 1,165 2,046
Gross Domestic Product (GDP)	252,244 371 1,004	212,606 354 819	230,028 363 869
M1	1,318,053 908 1,490	849,764 752 1,470	1,138,385 815 1,319
Industrial Production Index (IPI)	1.618 0.955 2.665	1.639 1.049 2.751	1.619 0.963 2.663
Retail sales	3.159 1.423 5.111	4.389 1.671 3.646	3.048 1.480 2.821
Unemployment – thousands	546.2 20.8 26.6	819.2 24.8 24.4	819.2 24.8 24.4

Balance of payments (BOP) – Transport – Receipts	1,919 36.0 70.3	2,225 38.5 68.0	3,828 50.8 66.4
Balance of payments (BOP) – Sea transport – Receipts	1,215 31.2 70.2	1,574 32.6 58.5	1,352 30.0 58.1
Unemployment – percentage	0.399 0.584 0.544	0.399 0.584 0.544	0.399 0.584 0.544
Balance of payments (BOP) – Transport – Payments	1,002 25.4 49.7	1,217 29.0 51.8	1,106 27.3 37.9
Imports of Goods without fuels and ships	12,479 98.1 224.7	12,246 96.5 152.1	14,772 102.5 175.1
Exports of Goods without fuels and ships	6,020 67.3 81.3	2,793 45.5 97.7	4,520 58.4 71.0
Exports of Goods	20,174 130.6 138.8	16,877 108.5 192.8	20,562 133.4 133.9
Balance of payments (BOP) – Sea transport – Payments	1,276 31.0 39.8	711.8 21.8 42.9	1,095 28.7 31.8
Imports of Goods	97,620 263.4 345.1	93,330 252.6 319.3	93,509 250.3 324.0
Balance of payments (BOP) – Travelling – Receipts	19,885 87.7 96.6	13,120 78.6 98.9	13,120 78.6 98.9
Balance of payments (BOP) – Travelling – Payments	1,563 24.4 28.8	1,560 23.4 23.0	1,687 267.0 25.7

BANK OF GREECE WORKING PAPERS

265. Katseli T. L., A. Theofilakou and K-M Zekente, “Central Bank Independence and Inflation Preferences: New Empirical Evidence on the Effects on Inflation”, July 2019.
266. Panagiotou S. and H. Thanopoulou, “Tonnage Tax Revisited: The Case of Greece During a Shipping Crisis and an economic crisis period”, July 2019.
267. Economides, G. and A. Xepapadeas, “The Effects of Climate Change on a Small Economy”, September 2019.
268. Bragoudakis, Z., S. Degiannakis and G. Filis, “Oil and Pump Prices: Is There Any Asymmetry in the Greek Oil Downstream Sector?, September 2019.
269. Dellis, K., “Financial System Heterogeneity and FDI Flows: Evidence from OECD Economies”, September 2019.
270. Gibson H., S. Hall, G. Tavlas and P. Petroulas, “On the effects of the ECB’s funding policies on bank lending and the demand for the euro as an international reserve”, September 2019.
271. Karadima, M. and E. Louri-Dendrinou “Non-Performing Loans in the Euro Area: Does Market Power Matter?”, September 2019.
272. Papandreou, A. “Stranded Assets and the Financial System”, November 2019.
273. George S. Tavlas, “A Reconsideration of the Doctrinal Foundations of Monetary-Policy Rules: Fisher versus Chicago”, November 2019.
274. Bragoudakis, Z. and D. Sideris, “Asymmetric Price Adjustment and the Effects of Structural Reforms in a Low Income Environment: the Case of the Gasoline Market in Greece”, December 2019.
275. Panayiotis C. A., and S. Anyfantaki, “Financial Literacy and its Influence on Consumers’ Internet Banking Behaviour, December 2019.
276. Balfoussia, H., H. Dellas, and D. Papageorgiou, “Fiscal Distress and Banking Performance: the Role of Macroprudential Regulation,” December 2019.
277. Anastasiou, D., Z. Bragoudakis and S. Giannoulakis, “Perceived Vs Actual Financial Crisis and Bank Credit Standards: Is There Any Indication Of Self-Fulfilling Prophecy?”, January 2020.
278. Gibson, D. H., S. G. Hall, P. Petroulas, V. Spiliotopoulos and G.S. Tavlas, “The Effect of Emergency Liquidity Assistance (ELA) on Bank Lending During the Euro Area Crisis”, January 2020.
279. Tavlas, G., “On the Controversy over the Origins of the Chicago Plan for 100 Percent Reserves”, June 2020.