International Journal of Economics and Financial Issues

ISSN: 2146-4138

available at http: www.econjournals.com

International Journal of Economics and Financial Issues, 2018, 8(1), 9-20.

Forecasting Unemployment Rates in USA Using Box-Jenkins Methodology

Nikolaos Dritsakis[1]*, Paraskevi Klazoglou[2]

[1]Department of Applied Informatics, University of Macedonia, Economics and Social Sciences, 156 Egnatia Street, 540 06 Thessaloniki, Greece, [2]Department of Applied Informatics, University of Macedonia, Economics and Social Sciences, 156 Egnatia Street, 540 06 Thessaloniki, Greece. *Email: drits@uom.gr

ABSTRACT

Unemployment, as a measure of market conditions, appears as an economic problem in every society and is a phenomenon with considerable negative social consequences. A low rate of unemployment is one of the main objectives for governmental macroeconomic policy. The main aim of this project is to identify the most appropriate forecasting model, i.e. the seasonal autoregressive integrated moving average (SARIMA), autoregressive conditional heteroskedasticity and the generalized autoregressive conditional heteroskedasticity (GARCH). Using one or a combination of these models could provide the best forecast for US unemployment. Applying monthly data to the US unemployment rate from January 1955 to July 2017 proved that the SARIMA$(1,1,2)(1,1,1)_{12}$ − GARCH(1,1) is the best model to project US unemployment. Finally, this project evaluates the forecasting performance of the model using forecast accuracy criteria, such as the root mean square error, mean absolute percent error and Theil's inequality coefficient.

Keywords: Unemployment Rates, Seasonal Time Series, Seasonal Autoregressive Integrated Moving Average-Generalized Autoregressive Conditional Heteroskedasticity Model

JEL Classifications: C53, E27

1. INTRODUCTION

Unemployment is an important macroeconomic variable that defines the condition of a country's economic equilibrium and is a barrier to social development. Some of its economic consequences are loss of productive forces, loss of income, as well as a burden on the state budget. High and persistent unemployment increases economic inequality, which is a term commonly used to show the differences in the distribution of wealth and income that exist between several social groups and individuals as well as amongst countries. Inequality generates redistributive pressure, which could lead to economic distortions hindering growth by causing persistent unemployment. Continued and persistent unemployment reduces growth and appears to be related to inequality.

Unemployment is a fundamental economic issue with significant negative social effects which could hinder growth not only as it is a waste of resources for the unemployed, but also because it leads people into poverty by limiting the liquidity which reduces private consumption mobility. As a consequence, repeling or avoiding the negative impact of unemployment is one of the major developmental government objectives, which apply a range of various measures so that as many people as possible find work. The unemployment phenomenon is observed not only in developing countries but also in developed countries. "American Economics Nobel prize winner, Shiller (2013), stated that income inequality is one of the main problems of unemployment."

The unemployment rate is the number of unemployed as a percentage of the labor force. To calculate the unemployment rate, the number of unemployed is divided by the total labor force, i.e. the sum of both the unemployed and employed, known as the economically active population. The labor force does not include people who cannot or do not desire to work, i.e. the economically inactive population. The significance in the amount of unemployment depends on the size of the labor force. For this reason, unemployment is measured as a percentage (%) of the labor force. This percentage is defined as follows:

Unemployment rate=(Number of unemployed)*100/(Labor force)

The US has one of the world's largest economies with a high level of productivity. A series of US historical facts may have affected its macroeconomic situation. The Great Depression of the 1930s caused an unemployment rate of 23.6%, the highest during that period. The lowest unemployment rate in the US occured in 1944, in World War II, at 1.2%. Since 1948, America experienced 11 recessions. The minimum postwar rate was 2.9% in 1953, while the maximum was 10.8% in the early 1980s and remained above 8% until September 2012. In 2008, economists faced an economic meltdown in the financial and banking sector, an international financial crisis that began in America and escalated in Europe. Between August 2005 and April 2008, unemployment was constantly below or equal to 5%, while rising in the second half of 2008 with an upward trend continuing to a maximum of 10% in October 2009. Remaining at the same levels until November 2010 (9.8%), unemployment started to decrease continuously and reached the current figure of 4.3% in July 2017.

The unemployment rate evolution is a main topic of political debate in many developed countries. The projection of the future unemployment rate is essential for economic policy makers to detect, plan and halt any persistent rise in the levels of unemployment in a country. An important question about time series forecasting is which model is the most accurate.

The main objective of this project is to identify the most appropriate model for exploring and forecasting the future unemployment rate in the US using the Box-Jenkins (1976) methodology and seasonal autoregressive integrated moving average (SARIMA) models. In particular, the ARIMA model specification is used to identify the most appropriate model and investigate possible seasonality and, by extension, the best SARIMA model. Following this, the symmetric GARCH models are estimated using the monthly unemployment rate in order to explore the best model and finally forecast future unemployment data using the most appropriate model specified.

The rest of this paper is organized as follows: Firstly, in section 2, we present a literature review about forecasting unemployment rates. Section 3 outlines a theoretical background about forecasting methodology and analysis. Section 4 outlines data and descriptive statistics and the empirical approach followed is analyzed in section 5. Section 6 proposes the forecasting methodology. Finally, the last section concludes the text with some closing remarks.

2. LITERATURE REVIEW

A number of research papers have used time series models for forecasting unemployment rates. Literature based on studies dealing with forecasting unemployment by using ARIMA models, i.e. Box-Jenkins methodology (1976) has been extensively used, to project future macroeconomic variables, such as unemployment rates. Recently, studies have also analyzed time series models with the incorporation of the autoregressive conditional heteroskedastic (ARCH) model introduced by Engle (1982). These models have been extended to the generalized autoregressive conditional heterokcedastic (GARCH) models leading to more parsimonious results rather than ARCH models.

Nkwatoh (2012) projected the unemployment rate in Nigeria using quarterly data from 1967Q1 to 2011Q4. Based on the results of the root mean square error (RMSE), mean absolute percent error (MAPE), mean absolute error (MAE) criteria and Theil's coefficient it was concluded that the ARIMA(1,1,2)-ARCH(1) model is the most appropriate to forecast unemployment in the specific period.

Rublikova and Lubyova (2013) investigated the monthly unemployment rate of Slovakia for the period from January 1999 to May 2013. The results showed that the ARIMA(0,1,2) $(0,1,1)_{12}$- GARCH(1,1) model proved to be the best to forecast the conditional mean as well as the conditional variance.

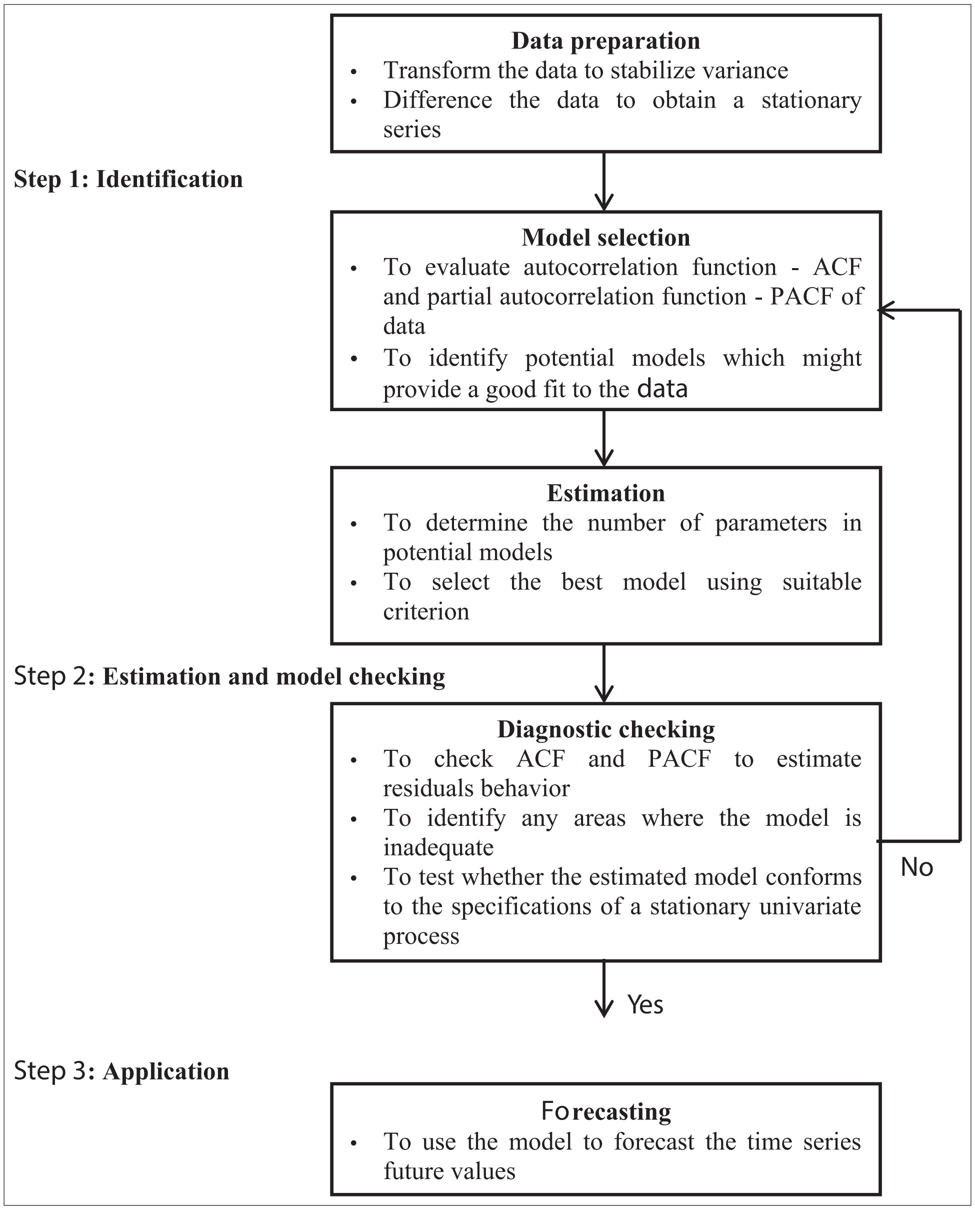
Dritsaki (2016) used monthly Greek unemployment data from 1998 to 2015 to predict the unemployment rate in both dynamic and static process. Using the Box-Jenkins methodology and SARIMA models concluded that the SARIMA$(0,2,1)(1,2,1)_{12}$ model is the best for forecasting, while the mean squared error (MSE), MAE criteria and Theil's coefficient presented more predictive accuracy in static process.

3. THEORETICAL BACKGROUND

The development and construction of ARIMA models as forecasting tools of economic variable values, is known as the Box-Jenkins (1976) method. This approach in time series analysis is a method for investigating an ARIMA(p,d,q) model or $\varphi(L)\Delta^d Y_t=\delta+\theta(L)\varepsilon$ that adequately represents the stochastic process from which the sample was derived. This method includes three steps; model identification, model estimation and diagnostic checking and finally, forecasting as shown in Figure 1.

The first step of the study is to specify the ARIMA model to determine the appropriate values of p, d and q in order to identify accurately the ARIMA (p,d,q) model. At the identification phase the appropriate value of d is estimated in order to obtain a stationary time series. If the series is not stationary the first, second or higher order of differences as well as data transformations are used to convert it into a stationary series. The Box–Jenkins (1976) methodology for ARIMA (p,d,q) model identification uses the autocorrelation-AC function and partial autocorrelation-PAC function as well as the unit root tests of augmented Dickey-Fuller (ADF) (1979, 1981) and Phillips-Perron (PP) (1998). Continuing on from this, the lag p of the AR process and the lag q of MA process is specified. This specification is based on the functions of autocorrelation and partial autocorrelation in series. The next step is to estimate the $\alpha_1,\ldots,\alpha_p$ parameters of the AR process and $\theta_1,\ldots,\theta_q$ parameters of the MA process which was identified in the previous step. If the series is an AR process, the coefficients can be estimated via the least squares method. If the series contains MA terms or a combination of both AR and MA terms (ARMA) then we can estimate the parameters using non linear estimation methods, such as the maximum likelihood (ML) method. The diagnostic checking stage then takes place to test whether the estimated model conforms to the specifications of a stationary process.

Figure 1: Schematic representation of the Box-Jenkins methodology for time series modeling



If the estimated model adequately explains the process from which the data is derived, the residuals should behave like white noise, i.e. the residuals should not exhibit autocorrelation. The residuals are tested using the Ljung–Box (1978, 1979) Q statistic. If the model is inadequate, we must return to the first phase to reconstruct a better model. After completing the above steps, the forecasting process is followed to project future time series values, based on the most appropriate model deriving from the previous stages.

3.1. Autoregressive (AR) Process

A general AR model of order p has the following form:

$$y_t=\alpha_1 y_{t-1}+\alpha_2 y_{t-2}+\ldots+\alpha_p y_{t-p}+\varepsilon_t \quad (1)$$

where,

$\varepsilon_t \sim N(0,\sigma^2)$ is white noise.

and with the lag operator L:

$$(1-\alpha_1 L-\ldots-\alpha_p L^p)y_t=\varepsilon_t \quad (2)$$

3.2. MA Process

A general MA model of order q has the following form:

$$y_t=\mu+\varepsilon_t-\vartheta_1\varepsilon_{t-1}-\vartheta_2\varepsilon_{t-2}-\ldots-\vartheta_q\varepsilon_{t-q} \quad (3)$$

and with the lag operator L:

$$y_t=\mu+(1-\vartheta_1 L^1-\vartheta_2 L^2-\ldots-\vartheta_q L^q)\varepsilon_t \quad (4)$$

3.3. Autoregressive Moving Average Process (ARMA)

A general ARMA model of orders p and q, ARMA(p,q) has the following form:

$$y_t=\alpha_1 y_{t-1}+\alpha_2 y_{t-2}+\ldots+\alpha_p y_{t-p}+\mu+\varepsilon_t-\vartheta_1\varepsilon_{t-1}-\vartheta_2\varepsilon_{t-2}-\ldots-\vartheta_q\varepsilon_{t-q} \quad (5)$$

and with the lag operator L:

$$(1-\alpha_1 L-\ldots-\alpha_p L^p)y_t=\mu+(1-\vartheta_1 L-\vartheta_2 L^2-\ldots-\vartheta_q L^q)\varepsilon_t \quad (6)$$

or

$$A(L)y_t=\mu+\Theta(L)\varepsilon_t \quad (7)$$

where:

$$A(L)=1-\alpha_1 L-\alpha_2 L^2-\ldots-\alpha_p L^p \quad (8)$$

$$\Theta(L)=1-\vartheta_1 L-\vartheta_1 L^2-\ldots-\alpha_q L^q \quad (9)$$

3.4. Autoregressive Integrated Moving Average (ARIMA) process

An ARMA(p,q) model following differences of the d order required to make the series stationary is known as ARIMA model of order (p,d,q) and symbolized as ARIMA(p,d,q).

The ARIMA(p,d,q) model with the lag operator L is as follows:

$$\Phi(L)y_t=A(L)(1-L)^d y_t=\mu+\Theta(L)\varepsilon_t \quad (10)$$

where:

$$\Phi(L)=A(L)(1-L)^d \quad (11)$$

An ARIMA(p,d,q) process may have the following forms:

- Difference equation form, as a function of the past values and the past and current values of the disturbance term. Analyzing the polynomial:

$$\Phi(L)=A(L)(1-L)^d=1-\varphi_1 L-\varphi_2 L^2-\ldots-\varphi_{p+d} L^{p+d} \quad (12)$$

The model takes the form of:

$$y_t=\mu+\varphi_1 y_{t-1}+\ldots+\varphi_{p+d} y_{t-(p+d)}+\varepsilon_t-\vartheta_1\varepsilon_{t-1}-\ldots-\vartheta_q\varepsilon_{t-q} \quad (13)$$

- Inverted form, as a function of the past values and the current value of the disturbance term. Following inversion of the polynomial we obtain:

$$\Pi(L)y_t=\left(1-\sum_{j=1}^{\infty}\pi_j L^i\right)y_t=\varepsilon_t \quad (14)$$

Therefore, the model becomes:

$$y_t=\pi_1 y_{t-1}+\pi_2 y_{t-2}+\ldots+\varepsilon_t \quad (15)$$

- Random shock form, as a function of the disturbance term, current and past values:

$$(y_t-\mu)=\Phi(L)-\Theta(L)\varepsilon_t=\varepsilon_t+\psi^1\varepsilon_{t-1}+\psi^2\varepsilon_{t-2}+\ldots=\psi^1 L\varepsilon_t+\psi^2 L_2\varepsilon_t+\ldots$$
$$=(1+\psi^1 L+\psi^2 L_2+\ldots)\varepsilon_t=\psi(L)\varepsilon_t \quad (16)$$

where:
$\psi(L)=1+\psi_1 L+\psi_2 L^2+\ldots$ is the random shock and
ψ_i is the i^{th} parameter of $\psi(L)$.

3.4.1. SARIMA process

In seasonal ARIMA models, seasonal differencing is required to turn the series into stationary, as is the case with general ARMA models. In non-stationary seasonal data with periodicity s the seasonal first order difference is defined as:

$$\Delta_s y_t=(1-L^s)y_t=y_t-y_{t-s} \quad (17)$$

While the seasonal Δ order difference is:

$$\Delta_s^D y_t=(1-L^s)^D y_t \quad (18)$$

where the lag operator L^s shows that observations present a seasonal periodic behavior.

The seasonal ARMA(p,q) model for every s is defined as:

$$\Phi(L^s)y_t=\theta(L^s)u_t \quad (19)$$

Where the random error u_t is white noise and θ symbolizes the seasonal lag parameter u_{t-12}

We consider that u_t in (19) follows an ARMA(p,q) model in the form of:

$$A(L)u_t=\Theta(L)\varepsilon_t \quad (20)$$

where ε_t is white noise. The polynomial A(L), Θ(L) orders are p and q, respectively. By replacing (20) with (19) the multiplicative seasonal ARMA(p,q)(P,Q)$_S$ model results in:

$$A(L)\Phi(L^s)y_t=\theta(L)\Theta(L^s)\varepsilon_t \quad (21)$$

If we consider that u_t in (19) are a form of ARIMA(p,d,q) model the multiplicative seasonal model has the following form:

$$A(L)\Phi(L^s)(1-L)^d(1-L^s)^D y_t=\theta(L)\Theta(L^s)\varepsilon_t \quad (22)$$

And is symbolized as ARIMA(p,d,q)(P,D,Q)$_s$. This model is determined by the constants p,d,q,P,D,Q,s which are calculated using a method similar to Box-Jenkins (1976).

3.4.2. Estimation of SARIMA models

To estimate SARIMA models the ML method is used. Under the assumption of independent and distributed standardized z_t, the log-likelihood (LL) function of $\{y_t(\theta)\}$, for a T observations sample, is given by:

$$\ln L[(y_t),\theta]=\sum_{t=1}^{T}\left[\ln[D(z_t(\theta),v)]-\frac{1}{2}\ln[\sigma_t^2(\theta)]\right] \quad (23)$$

where,
θ is the vector of the parameters that have to be estimated for the conditional mean, conditional variance and density function.
z_t is a sequence of independent and distributed random variables with mean as zero and variance as one.

3.5. Conditional Heteroskedasticity Models: ARCH and GARCH

The term "conditional" mean and "conditional" variance of a series Y_t is defined by the total available information I_{t-1} until the period t−1:

$$\mu_t = E\left(\frac{Y_t}{I_{t-1}}\right) \tag{24}$$

and

$$\sigma_t^2 = E\left(\frac{(Y_t-\mu_t)^2}{I_{t-1}}\right) \tag{25}$$

Where μ_t expresses the "conditional" mean and σ_t^2 the "conditional" variance of the series Y_t. The total information I_{t-1} includes all finite values of the time series under investigation.

If we consider the model:

$$Y_t=\mu_t+\varepsilon_t \tag{26}$$

where,

$$\mu_t = \delta_0 + \theta' X_t + \sum_{j=1}^{p}\alpha_j Y_{t-j} + \sum_{j=1}^{q}\beta_j \varepsilon_{t-j} \tag{27}$$

where, μ_t is the mean equation. Based on the definition (25) the "conditional" variance equals to:

$$\sigma_t^2 = \left(\frac{(Y_t-\mu_t)^2}{I_{t-1}}\right) = E\left(\frac{\varepsilon_t^2}{I_{t-1}}\right) \tag{28}$$

The random variable ε_t is linearly uncorrelated, but not independent (iid), due to variance volatility. Vector X_t contains determinants and θ′ symbolizes the vector of their coefficients. Equation (27) is known as the mean equation and (28) as the variance equation of the time series.

3.5.1. ARCH models

The ARCH model was developed for the first time by Engle (1982) and is based on the idea that the random error ε_t is linearly uncorrelated, but not independent over time. The general form of an ARCH(q) model is given by:

$$\varepsilon_t=u_t\sigma_t \text{ (mean equation)} \tag{29}$$

where,
$u_t\sim$ iid(0,1) and σ_t is the volatility that evolves over time.

The volatility σ_t^2 in the basic ARCH (q) model is defined as:

$$\sigma_t^2 = \delta_0 + \beta_1\varepsilon_{t-1}^2 + \ldots + \beta_q\varepsilon_{t-q}^2 = \delta_0 \sum_{i=1}^{q}\beta_i\varepsilon_{t-1}^2 = \delta_0 + A(L)\varepsilon_{t-1}^2$$

(variance equation) (30)

where
σ_t^2 is the conditional variance,
$\delta_0>0$ and $\beta_i\geq 0$ for σ_t^2 must be positive.
A(L) is the lag polynomial of squared residuals.

3.5.2. GARH models

Bollerslev (1986) extended the ARCH model into a model in which the conditional variance depends not only on the previous squared error values but also on the previous values of the variance itself. The proposed model is known as Generalized Autoregressive Conditional Heteroskedastic – the GARCH model. The general form of the GARCH(p,q) model is as follows:

$$R_t=\mu+\varepsilon_t \text{ (mean equation)} \tag{31}$$

$$\sigma_t^2 = \delta_0 + \sum_{j=1}^{q}\beta_j\varepsilon_{t-j}^2 + \sum_{j=1}^{p}\alpha_j\sigma_{t-j}^2 \text{ (variance equation)} \tag{32}$$

or

$$\sigma_t^2 = \delta_0 + B(L)\varepsilon_{t-j}^2 + D(L)\sigma_{t-j}^2 \tag{33}$$

where,
μ is the mean value.
ε_t is the error term at time t, which is assumed to be normally distributed with zero mean and conditional variance σ_t^2.
p is the order of GARCH and *q* is the order of ARCH process.
μ, δ_0, α_j and β_j are parameters for estimation. All parameters in the variance equation must be positive ($\mu>0$, $\delta_0>0$, $\alpha_j\geq 0$, and $\beta_j\geq 0$ for σ_t^2 must be positive**)**.
D(L) is the lag polynomial of the conditional variance σ_t^2 and
B(L) is the lag polynomial of the mean equation squared residuals.

3.5.3. Estimation of GARCH models

To estimate GARCH models (as in SARIMA models) the maximum likelihood (ML) method is used. The parametres of CARCH models are estimated by maximizing the LL function. Parameter estimation in the maximum LL function is obtained via non linear least squares using the Marquardt (1963) algorithm. The LL function is as follows:

$$\ln L[(y_t)\theta] = \sum_{t=1}^{T}\left[\ln[D(z_t(\theta),v)] - \frac{1}{2}\ln\left[\sigma_t^2(\theta)\right]\right] \tag{34}$$

where,
θ is the vector of the parameters that have to be estimated for the conditional mean, conditional variance and density function.
z_t is a sequence of independent and distributed random variables with mean as zero and variance as one.

In our project we estimate the maximum LL function using the distributions; normal, t-student, and generalized error distribution (GED).

In the case of standard normal distribution for similar distributed iid(0,1) random variables $\{z_t\}$, the following LL function has to be maximized.

$$\ln L[(y_t)\theta] = -\frac{1}{2}\left[T\ln(2\pi) + \sum_{t=1}^{T} z_t^2 + \sum_{t=1}^{T}\ln(\sigma_t^2)\right] \tag{35}$$

where,
θ is the vector of the parameters that have to be estimated for the conditional mean, conditional variance and density function,
T symbolizes the observations.

Bollerslev (1987) recommended the standardized t-student distribution with v>2 degrees of freedom. The LL function is defined as follows:

$$\begin{aligned}\ln L\left[(y_t,\theta)\right] = T\left[\ln\Gamma\left(\frac{v+1}{2}\right) - \ln\Gamma\left(\frac{v}{2}\right) - \frac{1}{2}\ln\left[\pi(v-2)\right]\right] \\ -\frac{1}{2}\sum_{t=1}^{T}\left[\ln(\sigma_t^2) + (1+v)\ln\left(1+\frac{z_t^2}{v-2}\right)\right]\end{aligned} \tag{36}$$

where,

$\Gamma(v) = \int_0^\infty e^{-x}x^{v-1}dx$ is the gamma function and

v is the degree of freedom.

For v→∞, the density function of standardized t-student distribution converges to the density function of normal distribution.

Nelson (1991) in constrast, proposed the use of the GED and the LL function is defined as follows:

$$\ln L\left[(y_t),\theta\right] = \sum_{t=1}^{T}\left[\begin{array}{l}\ln\left(\frac{v}{\lambda}\right) - \frac{1}{2}\left|\frac{z_t}{\lambda}\right|^{v} - (1+v^{-1})\ln(2) \\ -\ln\Gamma\left(\frac{1}{v}\right) - \frac{1}{2}\ln(\sigma_t^2)\end{array}\right] \tag{37}$$

where,

$$\lambda = \left[2^{-2/v}\frac{\Gamma\left(\frac{1}{v}\right)}{\Gamma\left(\frac{3}{v}\right)}\right]^{1/2}$$

4. DATA AND DESCRIPTIVE STATISTICS

The monthly data examined in our study is obtained from the OECD database and covers the period from January 1955 to July 2017. In Figure 2 the US unemployment rate in levels is presented for the specific period.

From the diagram in Figure 2 we notice that the US unemployment rate presents fluctuations throughout the period under analysis. The highest unemployment rise (10.8%) is presented in the months of November and December 1982 and the lowest (3.4%) from September 1962 to May 1969. In Table 1 the variable estimation in relation to time is given to determine whether a trend exists.

In Table 1 it is observed that a trend exists in the estimated model. As a consequence, we can assume that the investigated series is non-stationary. Onwards, the autocorrelation plot is also used to assess the series stationarity.

From the autocorrelation plot in Figure 3 we observe that the autocorrelation coefficients decline slowly, which indicates that the series is non-stationary. The next step is therefore, to reapply the above tests to identify if the series is stationary in first differences. In Figure 4 the unemployment rates in first differences are displayed.

The diagram results in Figure 4 indicate that possible stationarity exists in first differences. Table 2 is used to assess if a trend exists.

Stationarity is confirmed by the absence of trend in Table 2. Figure 5 also verifies the stationarity of the series under investigation.

From the correlogram in Figure 5 it is observed that the series appears to show stationarity in first differences, as well as seasonality. The series stationarity in first differences is validated by the unit root tests of Dickey–Fuller (1979, 1981) and Phillips Perron (1998).

Figure 2: The US unemployment rate from January 1955 to July 2017

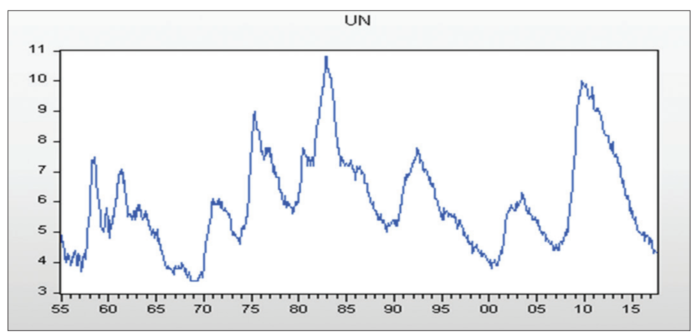


Table 1: Estimation of the US unemployment rate from January 1955 to July 2017

Dependent variable: UN				
Method: Least squares				
Sample: 1955M01 2017M07				
Included observations: 751				
Variable	**Coefficient**	**Std. Error**	**t‑statistic**	**Prob.**
C	5.376399	0.112016	47.99667	0.0000
@TREND	0.001586	0.000259	6.132478	0.0000
R^2	0.047809	Mean dependent var		5.971105
Adjusted R^2	0.046538	S.D. dependent var		1.573448
S.E. of regression	1.536399	Akaike info criterion		3.699419
Sum squared resid	1768.030	Schwarz criterion		3.711726
Log likelihood	−1387.132	Hannan‑Ouinn criter.		3.704161
F‑statistic	37.60729	Durbin‑Watson stat		0.014945
Prob (F‑statistic)	0.000000			

Figure 3: The autocorrelation and partial autocorrelation plots for the US unemployment rate from January 1955 to July 2017

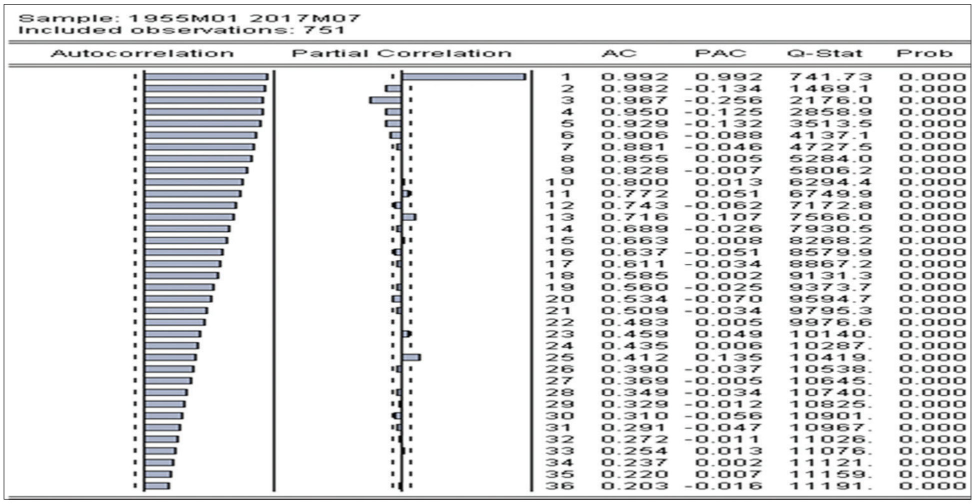


Figure 4: The US unemployment rate in first differences from January 1955 to July 2017

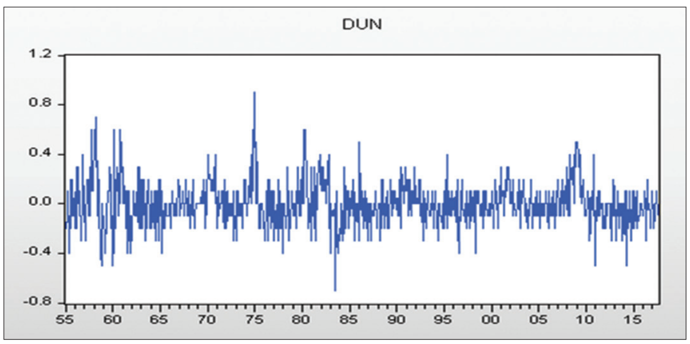


Table 2: Estimation of the US unemployment rate in first differences from January 1955 to July 2017

Dependent variable: DUN				
Method: Least squares				
Sample: 1955M01 2017M07				
Included observations: 750				
Variable	**Coefficient**	**Std. Error**	**t-statistic**	**Prob.**
C	0.010198	0.013731	0.742717	0.4579
@TREND	−2.93E-0.5	3.17E-0.5	−0.924584	0.3555
R-squared	0.001142	Mean dependent var		−0.000800
Adjusted R^2	−0.000194	S.D. dependent var		0.187811
S.E. of regression	0.187829	Akaike info criterion		−0.503902
Sum squared resid	26.38936	Schwarz criterion		−0.491582
Log likelihood	190.9632	Hannan-Ouinn criter.		−0.499155
F-statistic	0.854855	Durbin-Watson stat		1.707128
Prob (F-statistic)	0.355481			

The results in Table 3 confirm that the series is stationary in first differences. In Table 4 the descriptive statistics of the US unemployment rate in levels and first differences are shown.

From the Table 4 we can see that unemployment in levels and first differences does not constitute normal distribution.

Also, the unemployment distribution in levels and first differences is leptokurtic and positively skewed.

5. EMPIRICAL RESULTS

Following this, we define the SARIMA $(p,d,q)(P,D,Q)_s$ model form from the diagram results in Figure 5. Figure 5 and Table 3 exhibit stationary behavior in first differences and thus, the value of the d parameter equals to 1. The parameters p and q of the ARMA model can be identified by the partial autocorrelation and autocorrelation coefficients respectively, comparing them with the critical value

$$\pm\frac{2}{\sqrt{n}} = \pm\frac{2}{\sqrt{750}} = \pm 0.073$$

From the parameter values of partial autocorrelation and autocorrelation plots in Figure 5 the p value will be between 0<p<5 and respectively, the q value will be between 0<q<7. Figure 6 and Table 5 then present an unemployment graph and a trend estimation into seasonal differences of the data in first differences, respectively.

The results in Table 5 indicate that stationarity exists in seasonal differences of the data in first differences. In the correlogram of Figure 7 seasonal differences are presented, which correspond to the unemployment series in first differences.

From the correlogram in Figure 7 we observe that the seasonal differences for the autocorrelation function are significant for lag 12, while for the partial autocorrelation function, the seasonal differences are significant for 11, 24 and 36. Consequently, P and Q values will be: 0<P<3, 0<Q<1.

By using the above values of p, q, P and Q, we select the best SARIMA $(p,1,q)(P,1,Q)_{12}$ model from the lowest values of the AIC, SC και HQ criteria, estimating the model using the numerical optimization of Berndt-Hall-Hall-Hausman (1974) algorithm. In the following Table 6 the values of p, q, P και Q are shown.

The results in Table 6 indicate that according to Akaike (AIC) criterion, the SARIMA $(4,1,3)(2,1,1)_{12}$ model is the most appropriate, while according to the Schwartz (SIC) and Hannan-

Table 3: Summary table of ADF and PP unit root tests

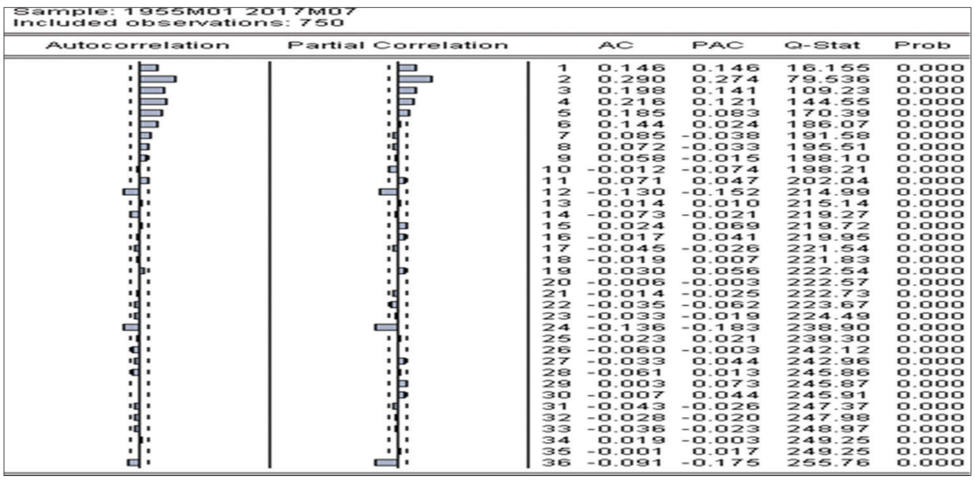
Variable	ADF		P-P	
	C	C, T	C	C, T
un	−2.4785 (5)	−3.0154 (5)	−2.3420 [18]	−2.7867 [18]
dun	−8.6906 (3)***	−8.7105 (3)***	−27.2330 [18]***	−27.204 [18]***

*,** and *** show significant at 1%, 5% and 10% levels respectively. The numbers within parentheses followed by ADF statistics represent the lag length of the dependent variable used to obtain white noise residuals. The lag lengths for ADF equation were selected using Schwarz Information Criterion (SIC). Mackinnon (1996) critical value for rejection of hypothesis of unit root applied. The numbers within brackets followed by PP statistics represent the bandwidth selected based on Newey and West (1994) method using Bartlett Kernel. C: Constant, T: Trend. d: First differences, ADF: Augmented Dickey–Fuller, PP: Phillips Perron

Table 4: Descriptive statistics of unemployment in levels and first differences

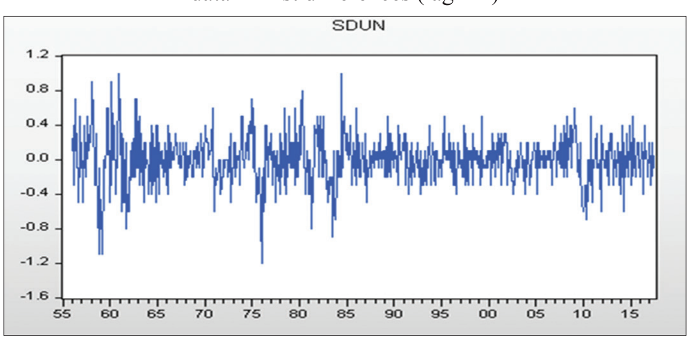
Variables	Mean	Median	Maximum	Minimum	Std. Dev	Skewness	Kurtosis	Jarque-Berra	Probability	Observations
un	5.97	5.70	10.80	3.40	1.57	0.73	3.07	67.39	0.00	751
dun	−0.0008	0.00	0.90	−0.70	0.19	0.55	4.70	128.40	0.00	750

Figure 5: The autocorrelation and partial autocorrelation plots for the US unemployment rate in first differences from January 1955 to July 2017


Sample: 1955M01 2017M07
Included observations: 750

	AC	PAC	Q-Stat	Prob
1	0.146	0.146	16.155	0.000
2	0.290	0.274	79.536	0.000
3	0.198	0.141	109.23	0.000
4	0.216	0.121	144.55	0.000
5	0.185	0.083	170.39	0.000
6	0.144	0.024	186.07	0.000
7	0.085	-0.038	191.58	0.000
8	0.072	-0.033	195.51	0.000
9	0.058	-0.015	198.10	0.000
10	-0.012	-0.074	198.21	0.000
11	0.071	0.047	202.04	0.000
12	-0.130	-0.152	214.99	0.000
13	0.014	0.010	215.14	0.000
14	-0.073	-0.021	219.27	0.000
15	0.024	0.069	219.72	0.000
16	-0.017	0.041	219.95	0.000
17	-0.045	-0.026	221.54	0.000
18	-0.019	0.007	221.83	0.000
19	0.030	0.056	222.54	0.000
20	-0.006	-0.003	222.57	0.000
21	-0.014	-0.025	222.73	0.000
22	-0.035	-0.062	223.67	0.000
23	-0.033	-0.019	224.49	0.000
24	-0.136	-0.183	238.90	0.000
25	-0.023	0.021	239.30	0.000
26	-0.060	-0.003	242.12	0.000
27	-0.033	0.044	242.96	0.000
28	-0.061	0.013	245.86	0.000
29	0.003	0.073	245.87	0.000
30	-0.007	0.044	245.91	0.000
31	-0.043	-0.026	247.37	0.000
32	-0.028	-0.020	247.98	0.000
33	-0.036	-0.023	248.97	0.000
34	0.019	-0.003	249.25	0.000
35	-0.001	0.017	249.25	0.000
36	-0.091	-0.175	255.76	0.000

Figure 6: The US unemployment rate in seasonal differences of the data in first differences (lag=12)



Quinn (HQ) criteria the best model is SARIMA $(1,1,2)(1,1,1)_{12}$. In Tables 7 and 8 estimations of the above models are given.

The results in Table 7 show that for the SARIMA $(4,1,3)(2,1,1)_{12}$ model there is a problem with coefficient significance.

The results in Table 8 show that all of the coefficients are statistically significant and therefore, the SARIMA $(1,1,2)(1,1,1)_{12}$ model is the best for forecasting. Figure 8 then tests for the existence of heteroskedasticity (ARCH(q) process test), from residual squares of the above model.

The results in Figure 8 show that all autocorrelation and partial autocorrelation coefficients are statistically significant. As a consequence of this, we can then argue that the ARCH-GARCH process exists in the model being studied. Table 9 contains the results of conditional heteroskedasticity ARCH(1) which confirm the ARCH model existence.

Through analysis of all the estimated models, the GARCH (1,1) model proved to be the most sutitable to test this process. We then estimate the symmetric SARIMA$(1,1,2)(1,1,1)_{12}$ – GARCH(1,1) model using the three following different distributions:

1. Conditional normal distribution of residuals,
2. T-student distribution and
3. GED.

Parameter estimation is performed by the LL method using the numerical optimization of Broyden-Fletcher-Goldfarb-Shanno (BFGS/Marquardt) algorithm. The results of this model using these three distributions are presented in Table 10.

From the Table 10 we notice that for all of the distributions, the coefficients are statistically significant (except for t-student). Furthermore, no problems exist in either autocorrelation or conditional autocorrelation, apart from normality. In addition, the

Table 5: Estimation of the US unemployment rate in seasonal differences of the data in first differences (lag=12)

Dependent variable: SDUN				
Method: Least squares				
Sample: 1956M02 2017M07				
Included observastions: 738 after adjustments				
Variable	**Coefficient**	**Std. Error**	**t-statistic**	**Prob.**
C	0.006752	0.021402	0.315500	0.7525
@TREND	−1.66E-05	4.90E-05	−0.339605	0.7343
R^2	0.000157	Mean dependent var		0.000407
Adjusted R^2	−0.001202	S.D. dependent var		0.283298
S.E. of regression	0.283468	Akaike info criterion		0.319271
Sum squared resid	59.14061	Schwarz criterion		0.331748
Log likelihood	−115.8110	Hannan-Ouinn criter.		0.324082
F-statistic	0.115332	Durbin-Watson stat		1.811953
Prob (F-statistic)	0.734251			

Table 6: Comparison of models with the AIC, SC και HQ criteria

SARIMA model	**AIC**	**SC**	**HQ**
dun			
$(4,1,3)(2,1,1)_{12}$	−0.726160	−0.652316	−0.697708
$(1,1,2)(1,1,1)_{12}$	−0.721747	−0.679919	−0.706397

model has the maximum LL value in GED. Thus, we can use this model for forecasting.

In Figure 9 the SARIMA$(1,1,2)(1,1,1)_{12}$ – GARCH(1,1) model distribution of standardized residuals is presented to confirm the hypothesis of normality.

From the results in Figure 9 we can see that in accordance with the skewness (0.22) and kurtosis (3.46) coefficients, the distribution presents positive skewness and leptokurtosis. Moreover, Figure 9 shows that the distribution of standardized residuals diverges from normality.

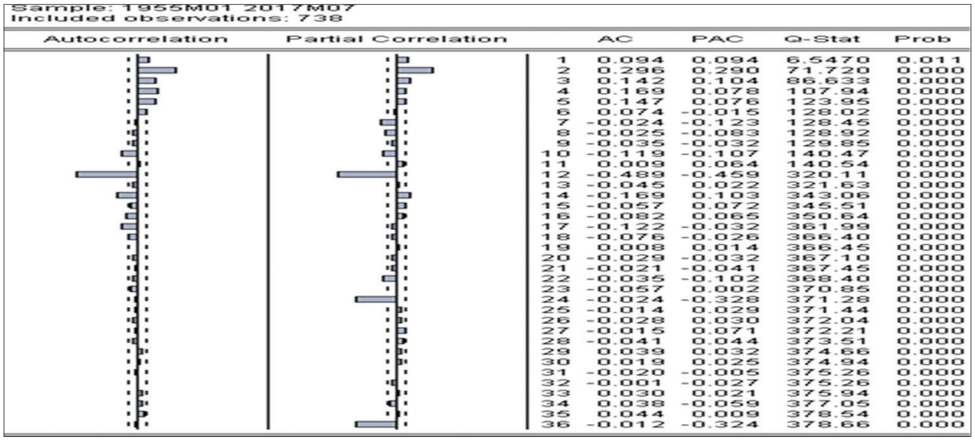
6. FORECASTING METHODOLOGY

In this section we present the forecasting outcomes for the SARIMA $(1,1,2)(1,1,1)_{12}$ – GARCH(1,1) model. To project unemployment we use both dynamic and static forecasting (one-step ahead). Static forecasting extends the recursion forwards from the end of the sample estimation, allowing one-step ahead projection both in structural samples and innovations. In the literature, a variety of statistics are used to evaluate forecasting. The optimum value of forecasting is assessed by the MSE. Other indexes are the MAE, RMSE, MAPE and Theil's inequality coefficient (1961). Having selected the form of the SARIMA

Table 7: Estimation of the SARIMA $(4,1,3)(2,1,1)_{12}$ model

Dependent variable: DUN				
Method: ARMA maximum likelihood (OPG–BHHH)				
Sample: 1955M01 2017M06				
Included observations: 750				
Convergence achieved after 37 iterations				
Coefficient covariance computed using outer product of gradients.				
Variable	**Coefficient**	**Std. Error**	**t-statistic**	**Prob.**
AR (1)	2.269052	0.176319	12.86905	0.0000
AR (2)	−1.745685	0.311379	−5.606304	0.0000
AR (3)	0.297328	0.191911	1.549297	0.1217
AR (4)	0.135082	0.048535	2.783183	0.0055
SAR (12)	0.522516	0.082200	6.356614	0.0000
SAR (24)	−0.072128	0.043923	−1.642148	0.1010
MA (1)	−2.244906	0.175973	−12.75709	0.0000
MA (2)	1.874065	0.285611	6.561602	0.0000
MA (3)	−0.521351	0.149318	−3.491547	0.0005
SMA (12)	−0.743396	0.071624	−10.37910	0.0000
SIGMASQ	0.027261	0.001171	23.28837	0.0000
R^2	0.226115	Mean dependent var		−0.000800
Adjusted R^2	0.215643	S.D. dependent var		0.187811
S.E. of regression	0.166333	Akaike info criterion		−0.729793
Sum squared resid	20.44567	Schwarz criterion		−0.662032
Log likelihood	284.6724	Hannan-Ouinn criter.		−0.703683
Durbin-Watson stat	1.997028			
Inverted AR roots	0.90−0.02i	0.90+0.02i	0.84	0.81+0.44i
	0.81−0.44i	0.78+0.43i	0.78−0.43i	0.77−0.46i
	0.77+46i	0.46−0.77i	0.46+0.77i	0.43+0.78i
	0.43−0.78i	0.02−0.90i	0.02+.90i	−0.02+0.90i
	−0.02−0.90i	−0.19	−0.43+.78i	−0.43−0.78i
	−0.46+.77i	−0.46−0.77i	−0.77+0.46i	−0.77−0.46i
	−0.78+0.43i	−0.78−0.43i	−0.90+0.02i	−0.90−0.02i
Inverted MA roots	0.98	0.84_0.49i	0.84−0.49i	0.84−0.47i
	0.84+0.47i	0.56	0.49−0.84i	0.49+.84i
	0.00−0.98i	−0.00+0.98i	−0.49−0.84i	−0.49+0.84i
	0.84+0.49i	−0.84−0.49i	−0.98	

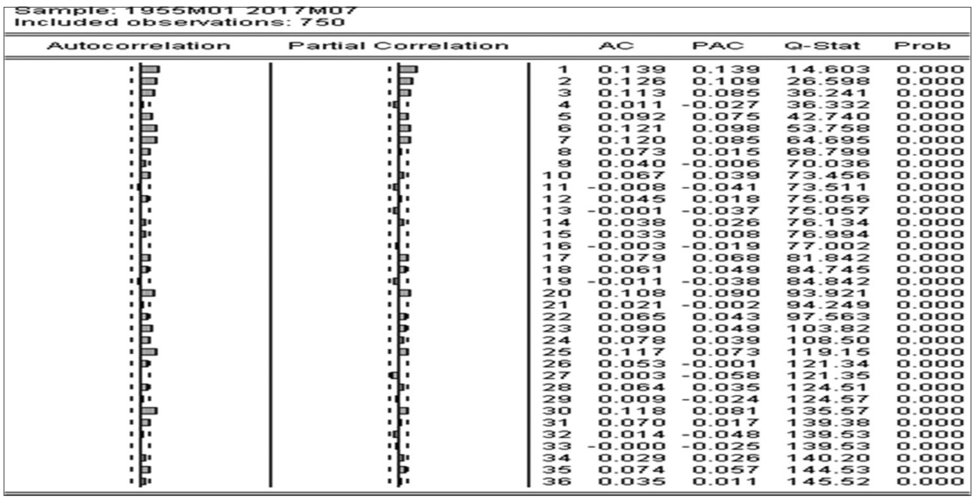
Figure 7: The autocorrelation and partial autocorrelation plots for the US unemployment rate in seasonal differences of the data in first differences (lag=12)



Sample: 1955M01 2017M07
Included observations: 738

	AC	PAC	Q-Stat	Prob
1	0.094	0.094	6.5470	0.011
2	0.296	0.290	71.720	0.000
3	0.142	0.104	86.633	0.000
4	0.169	0.078	107.94	0.000
5	0.147	0.076	123.95	0.000
6	0.074	-0.015	128.02	0.000
7	-0.024	-0.123	128.45	0.000
8	-0.025	-0.083	128.92	0.000
9	-0.035	-0.032	129.85	0.000
10	-0.119	-0.107	140.47	0.000
11	0.009	0.064	140.54	0.000
12	-0.489	-0.459	320.11	0.000
13	-0.045	0.022	321.63	0.000
14	-0.169	0.103	343.06	0.000
15	-0.057	0.072	345.51	0.000
16	-0.082	0.065	350.64	0.000
17	-0.122	-0.032	361.99	0.000
18	-0.076	-0.026	366.40	0.000
19	0.008	0.014	366.45	0.000
20	-0.029	-0.032	367.10	0.000
21	-0.021	-0.041	367.45	0.000
22	-0.035	-0.102	368.40	0.000
23	-0.057	0.002	370.85	0.000
24	-0.024	-0.328	371.28	0.000
25	-0.014	0.029	371.44	0.000
26	-0.028	0.030	372.04	0.000
27	-0.015	0.071	372.21	0.000
28	-0.041	0.044	373.51	0.000
29	0.039	0.032	374.66	0.000
30	0.019	0.025	374.94	0.000
31	-0.020	-0.005	375.26	0.000
32	-0.001	-0.027	375.26	0.000
33	0.030	0.021	375.94	0.000
34	0.038	-0.059	377.05	0.000
35	0.044	0.009	378.54	0.000
36	-0.012	-0.324	378.66	0.000

Figure 8: ARCH(q) process test



Sample: 1955M01 2017M07
Included observations: 750

	AC	PAC	Q-Stat	Prob
1	0.139	0.139	14.603	0.000
2	0.126	0.109	26.598	0.000
3	0.113	0.085	36.241	0.000
4	0.011	-0.027	36.332	0.000
5	0.092	0.075	42.740	0.000
6	0.121	0.098	53.758	0.000
7	0.120	0.085	64.695	0.000
8	0.073	0.015	68.799	0.000
9	0.040	-0.006	70.036	0.000
10	0.067	0.039	73.456	0.000
11	-0.008	-0.041	73.511	0.000
12	0.045	0.018	75.056	0.000
13	-0.001	-0.037	75.057	0.000
14	0.038	0.026	76.134	0.000
15	0.033	0.008	76.994	0.000
16	-0.003	-0.019	77.002	0.000
17	0.079	0.068	81.842	0.000
18	0.061	0.049	84.745	0.000
19	-0.011	-0.038	84.842	0.000
20	0.108	0.090	93.921	0.000
21	0.021	-0.002	94.249	0.000
22	0.065	0.043	97.563	0.000
23	0.090	0.049	103.82	0.000
24	0.078	0.039	108.50	0.000
25	0.117	0.073	119.15	0.000
26	0.053	-0.001	121.34	0.000
27	0.003	-0.058	121.35	0.000
28	0.064	0.035	124.51	0.000
29	0.009	-0.024	124.57	0.000
30	0.118	0.081	135.57	0.000
31	0.070	0.017	139.38	0.000
32	0.014	-0.048	139.53	0.000
33	-0.000	-0.025	139.53	0.000
34	0.029	0.026	140.20	0.000
35	0.074	0.057	144.53	0.000
36	0.035	0.011	145.52	0.000

Figure 9: The SARIMA(1,1,2)$(1,1,1)_{12}$ - GARCH(1,1) model distribution of frequencies and standardized residuals (GED distribution)

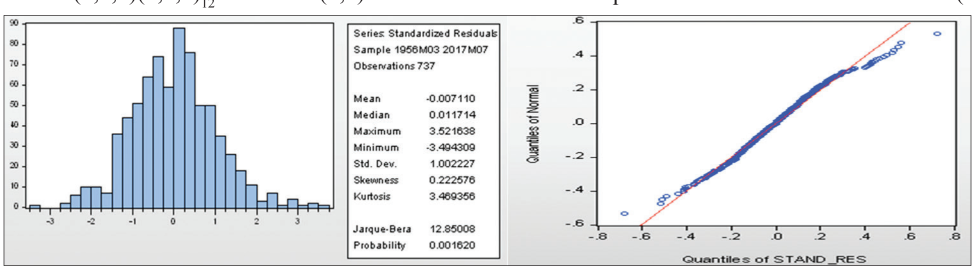


Figure 10: Static forecasting for unemployment

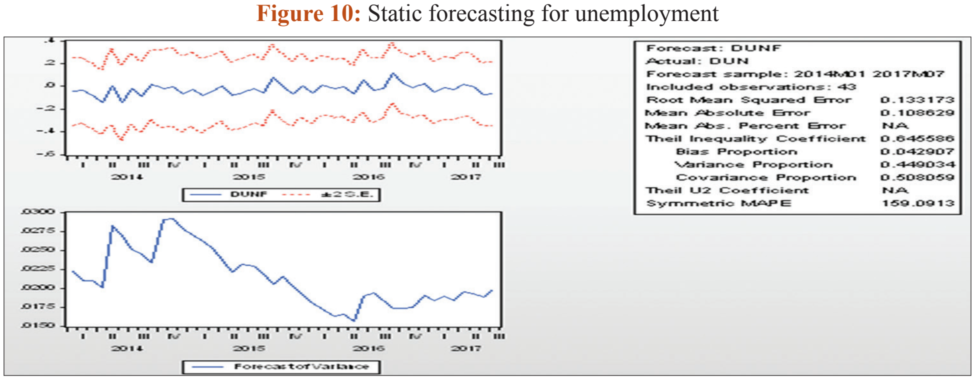


Figure 11: Dynamic forecasting for unemployment

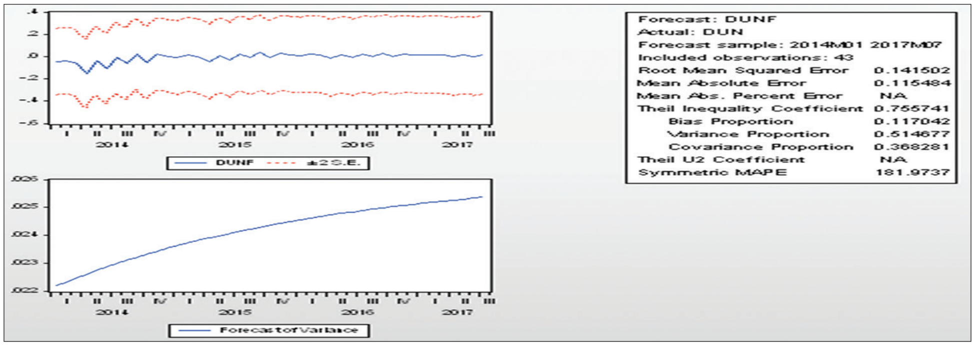


Table 8: Estimation of the SARIMA (1,1,2)(1,1,1) $_{12}$ model

Dependent variable: DUN				
Method: ARMA maximum likelihood (OPG–BHHH)				
Sample: 1955M01 2017M07				
Included observations: 750				
Convergence achieved after 22 iterations				
Coefficient covariance computed using outer product of gradients.				
Variable	**Coefficient**	**Std. Error**	**t-statistic**	**Prob.**
AR (1)	0.885978	0.030532	29.01754	0.0000
SAR (12)	0.512227	0.069498	7.370417	0.0000
MA (1)	−0.848613	0.040025	−21.20182	0.0000
MA (2)	0.157021	0.037166	4.224858	0.0000
SMA (12)	−0.796621	0.049251	−16.17456	0.0000
SIGMASQ	0.027737	0.001135	24.44790	0.0000
R^2	0.212605	Mean dependent var		−0.000800
Adjusted R^2	0.207313	S.D. dependent var		0.187811
S.E. of regression	0.167214	Akaike info criterion		−0.726624
Sum squared resid	20.80260	Schwarz criterion		−0.689664
Log likelihood	278.4840	Hannan-Ouinn criter.		−0.712382
Durbin-Watson stat	1.992193			
Inverted AR roots	0.95	0.89	0.82+0.47i	0.82−0.47i
	0.47+0.82i	0.47−0.82i	0.00+0.95i	−0.00−0.95i
	−0.47+0.82i	−0.47−0.82i	−0.82−0.47i	−0.82+0.47i
	−0.95			
Inverted MA roots	0.98	0.85+0.49i	0.85−0.49i	0.58
	0.49+0.85i	0.49−0.85i	0.27	0.00−0.98i
	−0.00+0.98i	−0.49−0.85i	−0.49+0.85i	−0.85−0.49i
	−0.85+0.49i	−0.98		

Table 9: Estimation of the ARCH (1) model

Heteroskedasticity test: ARCH			
F-statistic	14.77849	Prob. F (1,747)	0.0001
Obs*R^2	14.53058	Prob. Chi-square (1)	0.0001

$(1,1,2)(1,1,1)_{12}$ − GARCH(1,1) model with GED distribution we present the graphs of actual and projected values for static and dynamic forecasting in Figures 10 and 11, respectively, as well as statistic indicators of the model for static and dynamic forecasting and their innovations.

From the results in Figures 10 and 11 we discern that depending on indicators of the MAE, RMSE and Theil's inequality coefficient, static forecasting provides more accurate predictions for US unemployment in contrast to dynamic forecasting for the model in a wide confidence interval ±2SE.

7. CONCLUSION

All countries regard unemployment as one of the most economic and social hardships. On one hand, it stimulates researchers scientific interest while on the other, it encourages measures and policies to be taken by governments to confront it. The upward trend of unemployment plagues many countries and hence, it is important for researchers to investigate this trend and propose solutions. The use of ARIMA models is an exceptional tool to forecast the unemployment rate and deal with this rising trend. The main purpose of this study is to investigate the most appropriate model to project US unemployment. The results of this project proved that the US unemployment rate could be constructed and better predicted using the SARIMA$(1,1,2)(1,1,1)_{12}$ − GARCH(1,1) model. By applying the Box-Jenkins methodology the form of the SARIMA$(1,1,2)(1,1,1)_{12}$ − GARCH(1,1) model was estimated via non-linear optimization of maximum likelihood, using the

Table 10: Estimation of the SARIMA (1,1,2) $(1,1,1)_{12}$ - GARCH (1,1) model using the three distributions

Parameter	Normal	t-student	GED
δ_0	0.001 (0.018)	0.001 (0.064)	0.001 (0.048)
β_1	0.080 (0.000)	0.065 (0.001)	0.073 (0.001)
α_1	0.881 (0.000)	0.902 (0.000)	0.891 (0.000)
		D.O.F=14.38 (0.07)	PAR=1.706 (0.000)
Persistence	0.961	0.967	0.964
LL	311.8103	314.2802	313.8456
Jarque-Bera	11.146 (0.003)	13.850 (0.000)	12.850 (0.001)
ARCH (1)	1.104 (0.293)	2.563 (0.109)	1.796 (0.180)
Q^2 (12)	10.427 (0.579)	13.723 (0.394)	11.176 (0.514)
Q^2 (24)	19.501 (0.725)	21.452 (0.612)	20.242 (0.683)
Q^2 (36)	34.132 (0.558)	35.399 (0.497)	34.643 (0.533)

The persistence is calculated as $\beta_1+\alpha_1$ for SARIMA $(1,1,2)(1,1,1)_{12}$ - GARCH (1,1) model. Values in parentheses denote the P values. LL is the value of the log-likelihood

numerical optimization of the BFGS algorithm. To potentially forecast the model, both dynamic and static processes were used. The forecasting outcome showed that the projected value of unemployment is close to the real value. This result showed that the suitability of the $SARIMA(1,1,2)(1,1,1)_{12} - GARCH(1,1)$ model could be used to project US unemployment in the years to come with static forecasting. Undoubtedly, this result may be affected by changes in both time horizon and sample data size.

REFERENCES

Berndt, E., Hall, B., Hall, R., Hausman, J. (1974), Estimation and inference in nonlinear structural models. Annals of Economic and Social Measurement, 3(4), 653-665.

Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31, 307-327.

Bollerslev, T. (1987), Conditionally heteroskedastic time series model for speculative prices and rates of return. Review of Economics and Statistics, 69(3), 542-547.

Box, G.E.P., Jenkins, G.M. (1976), Time Series Analysis. Forecasting and Control. San Francisco: Holden-Day.

Broyden, C.G. (1970), The convergence of a class of double-rank minimization algorithms. Journal of the Institute of Mathematics and its Applications, 6, 76-90.

Dickey, D.A., Fuller, W.A. (1979), Distributions of the estimators for autoregressive time series with a unit root. Journal of American Statistical Association, 74(366), 427-431.

Dickey, D.A., Fuller, W.A. (1981), Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, 49(4), 1057-1072.

Dritsaki, C. (2016), Forecast of SARIMA models: An application to unemployment rates of Greece. American Journal of Applied Mathematics and Statistics, 4(5), 136-148.

Engle, R.F. (1982), Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. Econometrica, 50(4), 987-1008.

Fletcher, R. (1970), A new approach to variable metric algorithms. Computer Journal, 13(3), 317-322.

Goldfarb, D. (1970), A family of variable metric updates derived by variational means. Mathematics of Computation, 24(109), 23-26.

Ljung, G., Box, G. (1978), On a measure of lack of fit in time series models. Biometrika, 65(2), 297-303.

Ljung, G., Box, G. (1979), The likelihood function of stationary autoregressive-moving average models. Biometrika, 66(2), 265-270.

MacKinnon, J.G. (1996), Numerical distribution functions for unit root and cointegration tests. Journal of Applied Econometrics, 11(6), 601-618.

Marquardt, D.W. (1963), An algorithm for least squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics, 11, 431-441.

Nelson, D.B. (1991), Conditional heteroskedasticity in asset returns: A new approach. Econometrica, 59, 347-370.

Newey, W.K., West, K.D. (1994), Automatic lag selection in covariance matrix estimation. Review of Economic Studies, 61(4), 631-654.

Nkwatoh, L.S. (2012), Forecasting unemployment rates in Nigeria using univariate time series models. International Journal of Business and Commerce, 1(12), 33-46.

Phillips, P.C.B., Perron, P. (1998), Testing for a unit root in time series regression. Biometrika, 75, 335-346.

Rublikova, E., Lubyova M. (2013), Estimating ARIMA-ARCH model rate of unemployment in Slovakia. Forecasting Papers/Prognosticke Prace, 5(3), 275.

Shanno, D.F. (1970), Conditioning of quasi-newton methods for function minimization. Mathematics of Computation, 24(111), 647-656.

Shiller, R.J. (2013), Nobel Price Winner in Economics 2013 Say: Housing Not a Great Investment.

Theil, H. (1961), Economic Forecasts and Policy. Amsterdam: North-Holland Publishing Company.