ARIMA with Intervention Model for Predicting Consumer Price Index (CPI) in Malang City, Indonesia

By: Rauzan Sumara

1. Background

The central bank of Indonesia has the objective of achieving and maintaining the stability of the Indonesian currency (Rupiah). The stability of the Rupiah contains two aspects, the stability of the Rupiah to goods and services, and the stability of the Rupiah to the currency of other countries. The first aspect is reflected in the inflation rate while the second aspect is reflected in exchange rates.

As the monetary authority, The central bank of Indonesia sets and directs policies based on the inflation rate target to be achieved by taking into account various other macroeconomic targets, in the short, medium, and long term. The indicator that is often used to measure the inflation rate is the Consumer Price Index (CPI). Changes in the CPI from time to time indicate the price movements of goods and services consumed by the public.

East Java Province has quite a large economic activity in Java Island, Indonesia. This can be seen from the value of the Gross Regional Domestic Product (PDRB) per capital in 2015 which reached 43.5 million Rupiah. According to the 2015 East Java Province Regional Development Analysis Series, one of the biggest economic drivers in East Java is the Malang city with the category of regions with high economic growth above the average. Also, Malang is one of eight cities conducted by SBH (Cost of Living Survey) in East Java. The SBH survey results will serve as a reference for calculating the national Consumer Price Index (CPI).

Therefore, this report seeks to provide information for the CPI control program in Malang city in the coming periods by considering the CPI series data in previous periods. By using this prediction data, it is hoped that it can be taken into consideration in the formulation of policies to improve people's welfare.

2. Methodology

Time series data is a series of observational data arranged according to time, where the observational data are random (Cryer and Chan, 2008). Time series analysis is an analysis of a series of observational data that occurs based on a sequential time index with fixed time intervals and a statistical procedure to predict the probabilistic structure of future conditions to make decisions regarding related problems. The aim is to model the stochastic mechanism and to predict the value in the future period based on the previous period and other related factors.

2.4. Autoregressive Integrated Moving Average (ARIMA)

According to Wei (2006), the ARIMA model is a combination of AR and MA models with differencing order d. ARIMA model (p, d, q) in general is:

$$\phi_p(B)(1-B)^d Y_t = \theta_0 + \theta_q(B)e_t \tag{1}$$

with $\phi_p(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$, $\theta_q(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$, and θ_0 is *intercept* model.

Cryer and Chan (2008) in their book state that the time series Y_t follows the ARIMA model if the differencing d of $W_t = \Delta^d Y_t$ follows the ARMA model (p, q) then Y_t is ARIMA (p, d, q). They also formulated some general ARIMA models as follows:

a. AR(p) Model

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t \tag{2}$$

b. MA(q) Model

$$Y_t = \theta_0 + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \tag{3}$$

c. ARMA(p,q) Model

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \tag{4}$$

d. ARIMA(p,d,q) Model

$$W_t = \phi_1 W_{t-1} + \cdots + \phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \cdots - \theta_q e_{t-q} \tag{5}$$

which is $W_t = Y_t - Y_{t-1}$ where ϕ is parameter belongs to AR model, θ is parameter belong to MA model, p is parameter degree of AR model, d is the *difference*, q is parameter degree of MA model, and e_t is error model.

2.5. ARIMA Model with Intervention

Intervention is a time-series data analysis model that is used to explore the impact of unexpected external events on the variables that are observed. The general form of the intervention model is as follows (Wei, 2006):

$$Y_t = \sum_{j=1}^{k} \frac{\omega_{sj}(B) B^{bj}}{\delta_{rj}(B)} I_{j,t} + N_t \tag{6}$$

where

$$N_t = \frac{\theta_q(B)}{\phi_p(B)} a_t$$

with,

Y_t : time-series data,

$I_{j,t}$: intervention variables j-th and the intervention at time t,

$\omega_{sj}(B)$: $\omega_{0j} - \omega_{1j} B^1 - \omega_{2j} B^2 - \cdots - \omega_{sj} B^s$ is the numerator of the intervention,

$\delta_{rj}(B)$: $1 - \delta_{1j} B^1 - \delta_{2j} B^2 - \cdots - \delta_{rj} B^r$ is the denominator of the intervention,

b : *delay* time of the intervention effect, and

N_t : ARIMA(p,d,q) model,

In general, there are two kinds of intervention effects, **the step function model** and **the pulse function model**. A step function is a form of intervention that occurs from time T onwards in the long term. The form of step function intervention for these case examples is starting from a new policy being established until the policy is no longer valid. The form of step function intervention can be written as follows:

$$I_t = S_t = \begin{cases} 0, & t < T \\ 1, & t \geq T \end{cases} \tag{7}$$

where T is a time of intervention exist.

While the pulse function is a form of intervention that occurs only at a certain time (temporary), for example, the case of the bombing of the New WTC building. The form of pulse function intervention can be written as follows:

$$I_t = P_t = \begin{cases} 0, & t \neq T \\ 1, & t = T \end{cases} \tag{8}$$

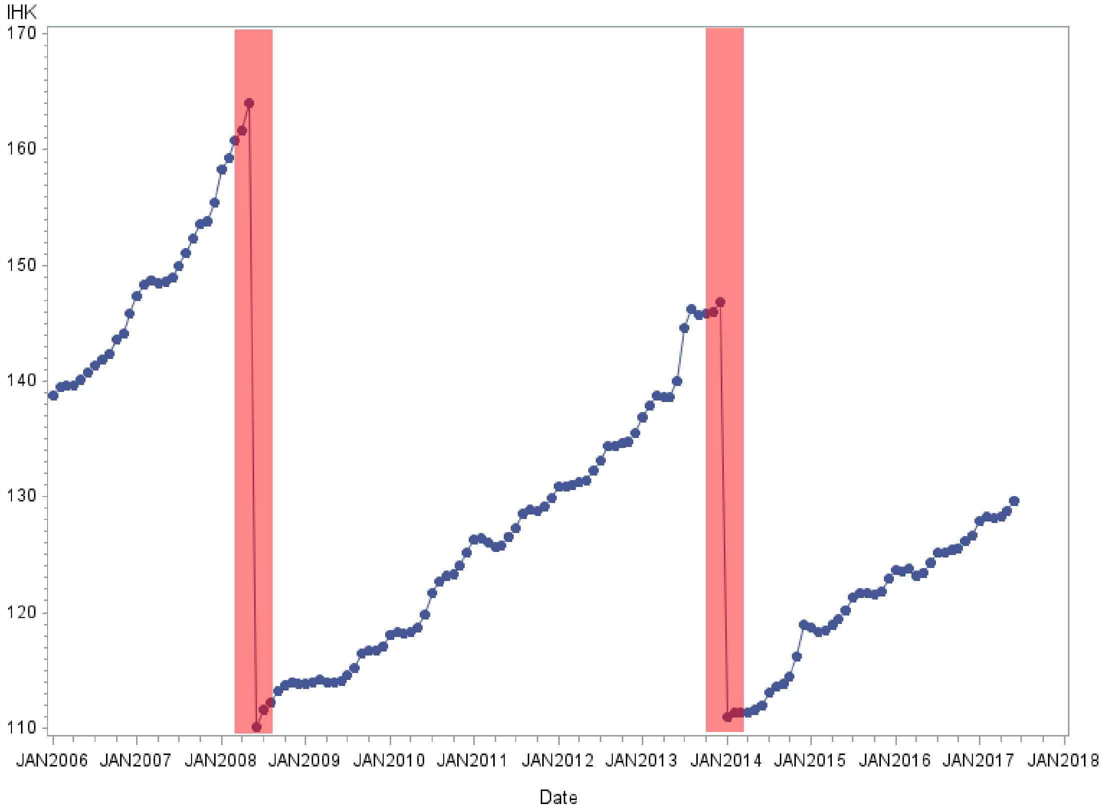
The stages of forming an intervention model are almost the same as the stages of forming a transfer function model (Wei, 2006).

2.6. Dataset

In this project, we will use Consumer Price Index (CPI) in Malang City, Indonesia from January 2006 to June 2017. The data can be downloaded in https://malangkab.bps.go.id (Complete URL in the references). The Autoregressive Integrated Moving Average (ARIMA) model with Intervention will be used, then the best model is for predicting the CPI for the next 12 months.

3. Result

Descriptive analysis in this study is used to see an overview of the data. The variable used is the Consumer Price Index (CPI) January 2006 - June 2017. The following is a plot descriptive analysis of the CPI between January 2006 - June 2017,



The results of the descriptive analysis indicate that the CPI from January 2006 - June 2017 reached the highest index in May 2008 which is 164.01. One month after that, in June 2008, it decreased drastically to 110.08. After that, there was an increase periodically in the following months and reached an index of 146.84 in December 2013. In January 2014, it decreased again to 110.99. This symptom is caused by a change in the reference basis for calculating the CPI. Since June 2008, the CPI calculation has been based on the consumption of the living cost survey in 66 cities in 2007. Whereas since

January 2014, the CPI calculation has been based on the consumption of the living cost survey in 82 cities in 2012.

Through the explanation above, it can be informed that there are effects of the interventions for the CPI from January 2006 - June 2017. The interventions were determined to occur in June 2008 (Intervention I) and January 2014 (Intervention II). The intervention represented policy factors in influencing the January 2006 - June 2017 CPI situation.

3.1. ARIMA with Intervention Modeling

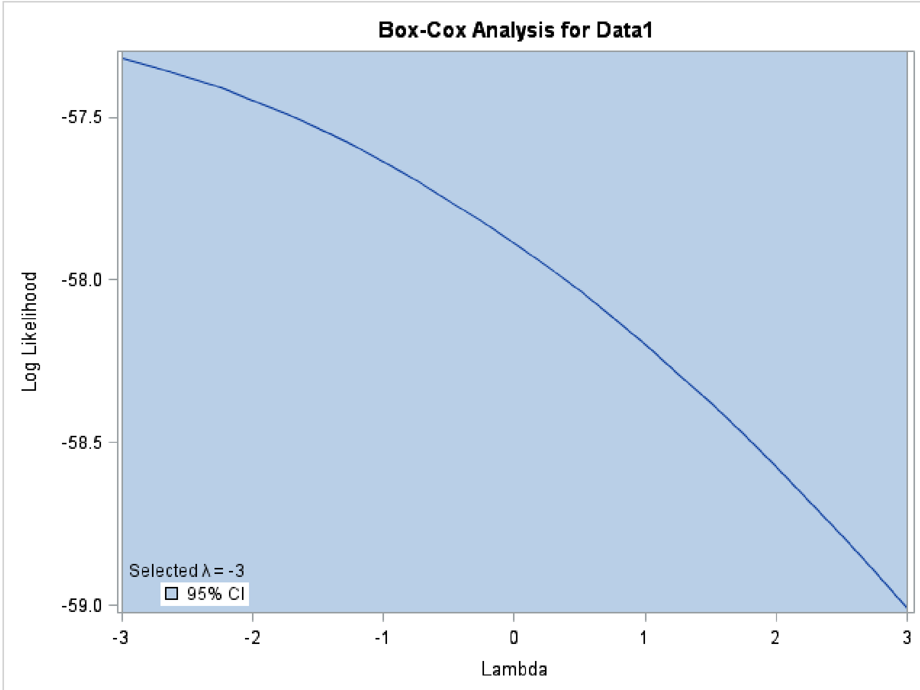
Based on the previous explanation, the appropriate intervention function was determined, namely the step function for intervention I and intervention II. The formation of the intervention model begins by dividing the data series into three parts which are data from the first to before intervention I, data from the first to before Intervention II, and data from the first to the last observation.

3.1.1. ARIMA Model (Using the First Data)

a. Stationarity Test

Stationary testing is intended to obtain data that has a stable mean and variety that does not contain a unit root. Data that has been declared stationary means that the data is stable for the forecasting process because the model that can produce accurate forecasting comes from data that is stable in both means and variety.

Variance stationery is carried out by the Box-Cox transformation, with the criteria if the confidence interval contains the value of lambda = 1, it can be stated that the data is stationary to the variance.



Based on the Box-Cox transformation plot in the Figure above, it is known that the 95% confidence interval in the CPI contains a lambda value = 1. This shows that the data is stationary concerning variance.

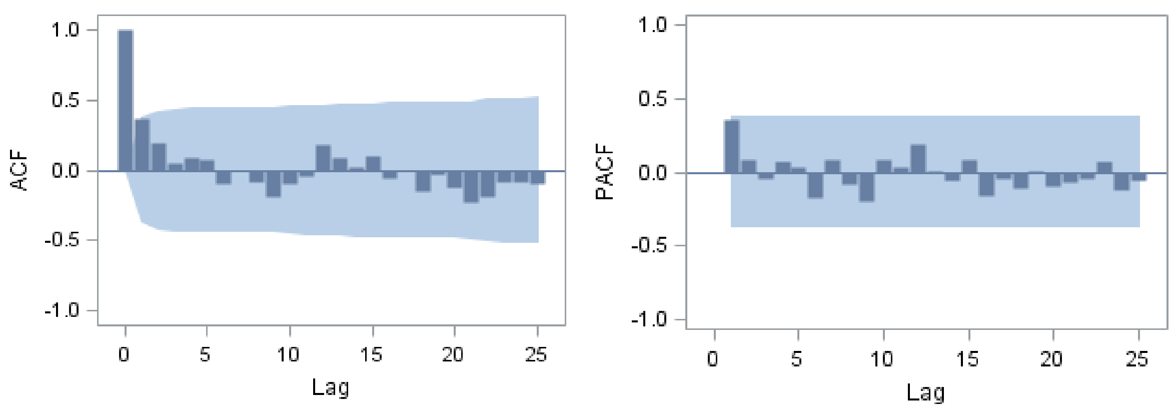
The stationarity to the mean was done by using the Augmented Dickey-Fuller Test (ADF Test). The test criteria state that if the probability ≤ level of significance (alpha (α) = 5%) then the data is stationary in mean, and if the probability > level of significance (alpha (α) = 5%) then the data is declared non-stationary, so the CPI data must be done by differencing. The results of stationary testing can be seen through the summary in the following table:

Level of *d*	**Rho**	**Probability**
Without Difference	1.434	0.992
1st Difference	-15.572	0.016

Based on the test results in the table above, it is known that at the level (without a difference), the rho value is 1.434 with the probability of 0.992. These results indicate a probability > level of significance (alpha (α) = 5%) so that the data is non-stationary at the level. Thus, the CPI must be done by 1st differencing, and the rho obtained after differencing is -15,572 with the probability of 0.016. These results indicate the probability < level of significance (alpha (α) = 5%) so that the data is stationary at the 1st differencing.

b. Identification of ARIMA Model

Identification of the ARIMA model is intended to obtain the ARIMA model (*p, d, q*) from data that has been stationary. ARIMA model (*p, d, q*) can be identified through a significant lag from ACF and PACF. The significant lag from the ACF will form the MA (*q*) model, while the significant lag from the PACF will form the AR (*p*) model, and the differencing is expressed as order *d*. The plot of the ACF and PACF can be seen as follows:



Based on the ACF plot, the lag that crosses the confident interval is the 1st lag, so it can be identified that the *q* order formed is MA (1). Then based on the PACF plot, the lag that crosses the confident interval is the 1st lag, so it can be identified that the *p* order formed is AR (1). Thus, the tentative ARIMA models formed are as follows:

1. ARIMA (1,1,0) 2. ARIMA (0,1,1) 3. ARIMA (1,1,1)

c. Estimation of Parameter Model

No.	Tentative	Parameter	Probability	Information	AIC	SBC
1.	ARIMA (1,1,0)	constant	0.0001	Significant	57.46736	60.13177
		ϕ_1	0.0382	Significant		
2.	ARIMA (0,1,1)	constant	0.0001	Significant	58.46591	61.13032
		θ_1	0.1199	Insignificant		
3.	ARIMA (1,1,1)	constant	0.0001	Significant	59.3568	63.35341
		ϕ_1	0.2209	Insignificant		
		θ_1	0.7455	Insignificant		

Based on the table above, it can be seen from the three tentative models, showing that all parameters in the ARIMA model (1,1,0) are significant and it has the smallest AIC and SBC, while the ARIMA (0,1,1) and ARIMA (1,1,1) models have insignificant parameters. Therefore, the ARIMA model (1,1,0) is chosen as the basic model for the formation of an intervention model on the CPI data.

d. Diagnostics of ARIMA (1,1,0) Model

The diagnostics test of the model is to determine whether the residual of the model is normally distributed and white noise or not. The following are the results of testing the normality assumption and white noise test:

White Noise Test			**Normality Test**	
Lag	**Chi-Square**	**Probability**	**Kolmogorov Smirnov**	**Probability**
6	1.92	0.8597	0.148	0.109
12	6.20	0.8599		
18	10.76	0.8688		
24	17.21	0.7987		

To determine whether the remainder is normally distributed or not, it can be seen through the Kolmogorov Smirnov test. These results indicate that the probability > level of significance (α = 5%). This means that the residual of the ARIMA model is normally distributed, then the normality assumption is met.

Testing of the white noise assumption uses the Ljung-Box (Q) test with the testing criteria stating that if the probability > level of significance (alpha (α) = 5%) then the residual is white noise (assumption fulfilled), on the other hand, if the probability ≤ level of significance (alpha (α) = 5%) then the residual is not white noise (the assumption is not fulfilled). Based on the results of the white noise test, the residual of the ARIMA (1,1,0) model, the test statistics on all lags are insignificant. It means that the residuals in the ARIMA (1,1,0) model are white noise.

e. ARIMA(1,1,0) Model

The ARIMA (1,1,0) model is as follows:

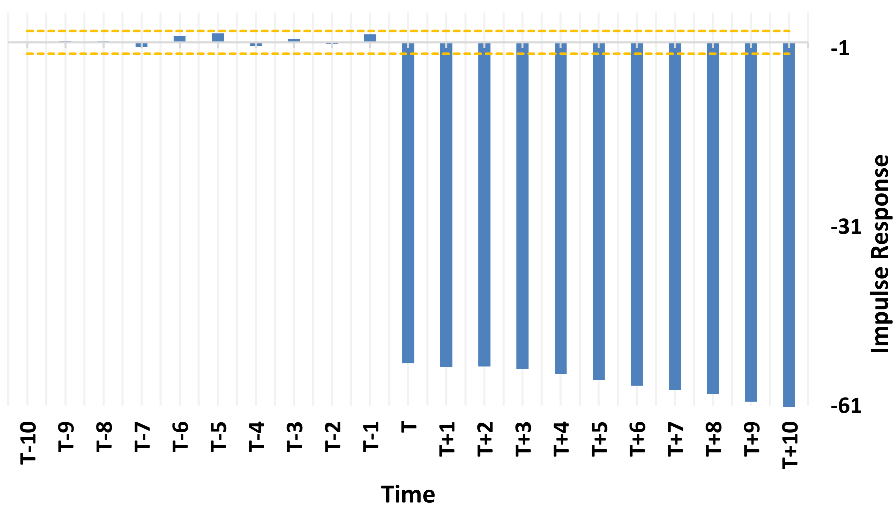
$$Y_t = 0.932 + Y_{t-1} + 0.405Y_{t-1} - 0.405Y_{t-2} + e_t$$

This model can be interpreted that the current state or level of the CPI is influenced by the state of the CPI in the two previous periods. The model coefficient of 0.405 means that the CPI in the previous period had a positive and significant effect on the CPI in the current period. This means that an increase in the CPI of 1 index in the previous period will increase the current period's CPI by 0.405. The model coefficient of -0.405 means that the CPI of the two previous periods has a negative and significant effect on the CPI in the current period. This means that an increase in the CPI of 1 index in the previous two periods will reduce the CPI of the current period by 0.405.

3.1.2. Intervention I Modeling (Using the Second Data)

a. Order Determination of b, r, and s

The first intervention is a change in the basis for calculating the CPI based on the consumption of living cost survey in 66 cities in 2007, since June 2008 or since T = 30 is a form of the step function. The first thing in modeling is determining the order of b, r, and s for the first intervention model. Determine the order of the first intervention can be seen through the diagram of the residuals in the following Figure.



From the Figure above, the impulse response is significant from the T period, it can be presumed that the tentative order of the impulse response is

Order	b	r	s
Respon Impuls (b,r,s)	0	1	1
	0	0	1
	0	1	0

Because there are three tentative models, it is necessary to choose the best order that will represent the impulse response to intervention I.

b. Parameter Estimation of Intervention I Model

From the tentative impulse response orders, here are the results of parameter estimation, and the significant parameter test of the Intervention I model can be seen in the following table:

Order b,r,s	**Parameter**	**Estimate**	**Probability**	**Information**	**AIC**	**BIC**
(0,1,1)	constant	0.932	0.0001	Significant	198.8545	211.6239
	ϕ_1	0.394	0.0001	Significant		
	ω_{01}	-55.385	0.0001	Significant		
	ω_{11}	-54.989	0.0001	Significant		
	δ_{11}	-0.014	0.2429	Insignificant		
(0,0,1)	constant	0.923	0.0001	Significant	198.2729	208.4885
	ϕ_1	0.405	0.0001	Significant		
	ω_{01}	-55.655	0.0001	Significant		
	ω_{11}	-55.294	0.0001	Significant		
(0,1,0)	constant	2.885	0.0145	Significant	597.7264	607.9419
	ϕ_1	0.112	0.3116	Insignificant		
	ω_{01}	-7.376	0.0032	Significant		
	δ_{11}	-0.870	0.0001	Significant		

Based on the table above, it can be seen from the three tentative models based on the order of the Intervention I showing that all parameters in the ARIMA (1,1,0) with intervention order b = 0, r = 0, and s = 1 are significant and it has the smallest AIC and SBC, while the ARIMA (1,1,0) with intervention order b = 0, r = 1, and s = 1 and the ARIMA (1,1,0) with intervention order b = 0, r = 1, and s = 0 have insignificant parameters. Therefore, the ARIMA (1,1,0) with intervention order b = 0, r = 0, and s = 1 is chosen.

c. Intervention I Model

ARIMA (1,1,0) with intervention I model is as follows:

$$Y_t = \theta_0 + (\omega_{01} - \omega_{11}B)\, S_t^{(T)} + \frac{1}{(1-\phi_1 B)(1-B)} e_t$$

$$= 0.923 + (-55.655 + 55.294B)\, S_t^{(30)} + \frac{1}{(1-0.405B)(1-B)} e_t$$

$$= 0.923 - 55.655\, S_t^{(30)} + 55.294\, S_{t-1}^{(30)} + \frac{1}{(1-0.405B)(1-B)} e_t$$

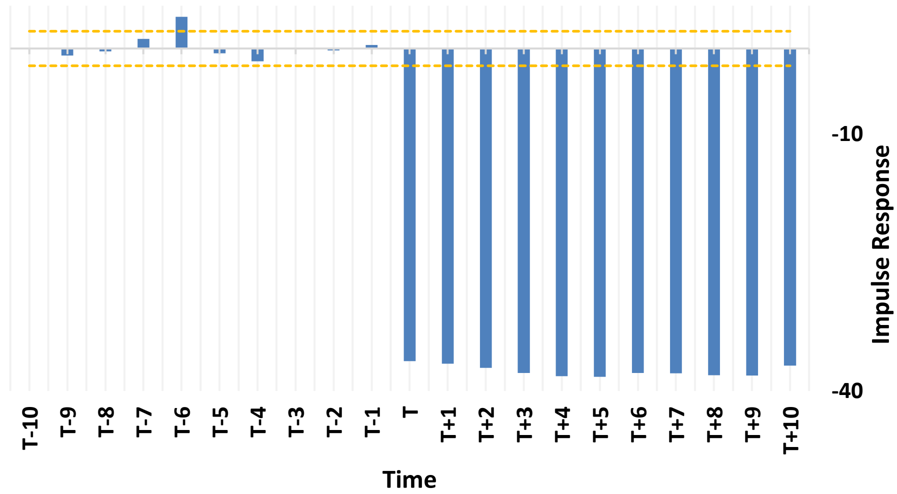
The $S_t^{(30)}$ is the intervention of changing the basis for CPI calculation based on the consumption of living cost survey in 66 cities in 2007 which was initialized through a dummy variable with the value of 1 from June 2008 onwards. Changing of reference basis for calculating the CPI cause a decline in the CPI of 53.89 since June 2008.

3.1.3. Intervention II Modeling (Using the Third Data)

a. Order Determination of b, r, and s

Intervensi kedua yaitu perubahan dasar acuan perhitungan IHK didasakan pada pola konsumsi survei biaya hidup di 82 kota tahun 2012, sejak Januari 2014 atau sejak T = 97 merupakan bentuk fungsi *step*. Langkah selanjutnya adalah menentukan orde dugaan b, r, dan s untuk model intervensi kedua. Untuk menentukan orde intervensi kedua dapat dilihat melalui diagram residual pada Gambar berikut.

The second intervention is the change in the basis for calculating the CPI based on the consumption of living cost survey in 82 cities in 2012, since January 2014 or since T = 97 is a form of the step function. The next step is to determine the order of b, r and s for the second intervention model. To determine it can be seen through the residual diagram in the following Figure



From the Figure above, the impulse response is significant from the T period, it can be presumed that the tentative order of the impulse response is

Orde	b	r	s
Respon Impuls (b,r,s)	0	1	1
	0	0	1
	0	1	0

Because there are three tentative models, it is necessary to choose the best order that will represent the impulse response to intervention II.

b. Parameter Estimation of Intervention II Model

From the tentative impulse response orders, here are the results of parameter estimation, and the significant parameter test of the Intervention II model can be seen in the following table:

Order b,r,s	Parameter	Estimate	Probability	Information	AIC	BIC
(0,1,1)	constant	0.921	0.0001	Significant	277.6871	298.1269
	ϕ_1	0.371	0.0001	Significant		
	ω_{01}	-55.603	0.0001	Significant		
	ω_{11}	-55.245	0.0001	Significant		

	ω_{02}	-36.455	0.0001	Significant		
	ω_{12}	-36.354	0.0001	Significant		
	δ_{12}	0.002	0.9209	Insignificant		
(0,0,1)	constant	0.921	0.0001	Significant	275.6973	293.2171
	ϕ_1	0.371	0.0001	Significant		
	ω_{01}	-55.604	0.0001	Significant		
	ω_{11}	-55.245	0.0001	Significant		
	ω_{02}	-36.435	0.0001	Significant		
	ω_{12}	-36.331	0.0001	Significant		
(0,1,0)	constant	0.910	0.1803	Insignificant	703.6984	721.2183
	ϕ_1	0.146	0.1110	Insignificant		
	ω_{01}	-55.082	0.0001	Significant		
	ω_{11}	-55.354	0.0001	Significant		
	ω_{02}	-4.737	0.0002	Significant		
	δ_{12}	-0.867	0.0001	Significant		

Based on the table above, it can be seen from the three tentative models based on the order of the Intervention II showing that all parameters in the ARIMA (1,1,0) with intervention order b = 0, r = 0, and s = 1 are significant and it has the smallest AIC and SBC, while the ARIMA (1,1,0) with intervention order b = 0, r = 1, and s = 1 and the ARIMA (1,1,0) with intervention order b = 0, r = 1, and s = 0 have insignificant parameters. Therefore, the ARIMA (1,1,0) with intervention order b = 0, r = 0, and s = 1 is chosen as the best intervention II model.

c. Intervention II Model

ARIMA (1,1,0) with intervention II model is as follows:

$$Y_t = \theta_0 + (\omega_{01} - \omega_{11}B)\, S_t^{(T)} + (\omega_{02} - \omega_{12}B)\, S_t^{(T)} + \frac{1}{(1-\phi_1 B)(1-B)} e_t$$

$$= 0.92 + (-55.60 + 55.24B)\, S_t^{(30)} + (-36.43 + 36.33B)\, S_t^{(97)} + \frac{1}{(1-0.37\,B)(1-B)} e_t$$

$$= 0.92 - 55.60\, S_t^{(30)} + 55.24\, S_{t-1}^{(30)} - 36.43\, S_t^{(97)} + 36.33\, S_{t-1}^{(97)} + \frac{1}{(1-0.37\,B)(1-B)} e_t$$

The $S_t^{(97)}$ is an intervention of changing the basis for CPI calculation to the consumption of the living cost survey in 82 cities in 2012 which has been in effect since January 2014. From the model above, it can be explained that the effect caused by changes in the basic calculation of CPI is a decrease in the CPI. Here it can be seen also the fact that changing the basis for CPI calculation to the consumption of the living cost survey in 66 cities in 2008 tends to be greater than changing the basis for CPI calculation to the consumption of the living cost survey in 82 cities in 2012.

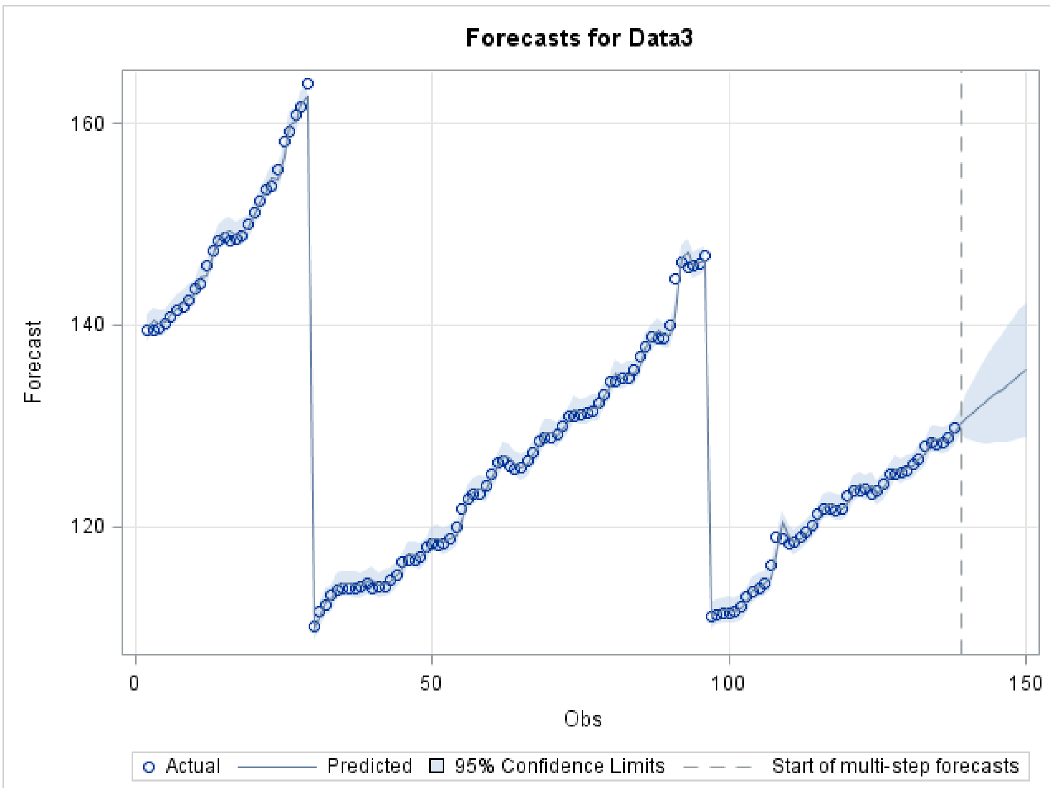
3.1.1. Forecasting

The ARIMA (1,1,0) with Intervention II model is used to forecast the next 12 periods from the CPI which is forecasting results as follows:

Periods	Forecasting	Lower Limit	Upper Limit
Jul-17	130.339	129.070	131.608
Aug-17	130.857	128.703	133.011
Sep-17	131.338	128.455	134.221
Oct-17	131.805	128.307	135.303
Nov-17	132.267	128.235	136.299
Dec-17	132.726	128.219	137.234
Jan-18	133.186	128.247	138.124
Feb-18	133.644	128.309	138.979
Mar-18	134.103	128.399	139.807
Apr-18	134.562	128.510	140.613
May-18	135.020	128.641	141.400

Jun-18	135.479	128.788	142.171

Based on the table above, it can be seen that the CPI forecast tends to increase from the previous period. For example, CPI of 131,608 in July 2017, 133,011 in August 2017, 134,221 in September 2017, and up to 142,171 in June 2018. Visually the forecast results are stated in the following graph:



Based on the graph, it is known that in the next 12 periods the CPI tends to increase over time. Likewise, the confidence interval for the forecasting results will widen as the forecast period is high. This means that the longer the period to be predicted, the higher the error will be.

References

Cryer, D.J. dan Chan, K.S. 2008. *Time Series Analysis With Applications in R*. Second Edition. Iowa City: Springer.

Wei, W.W.S. 2006. *Time Series Analysis Univariate & Multivariate*. Second Edition. Boston: Addison Wesley.

Malang City, 2021. https://malangkab.bps.go.id/statictable/2016/09/20/584/indeks-harga-konsumen-dan-inflasi-bulanan-malang-2005-2006-2007-jan-mei-2008-2002-100-juni-2008---desember-2013-2007-100-januari-2014---desember-2019-2012-100-.html

Appendices

Appendix 1. Dataset

Date	IHK	Date	IHK	Date	IHK
Jan-06	138.72	Feb-10	118.36	Mar-14	111.37
Feb-06	139.53	Mar-10	118.19	Apr-14	111.35
Mar-06	139.57	Apr-10	118.37	May-14	111.53
Apr-06	139.64	May-10	118.71	Jun-14	112.01
May-06	140.16	Jun-10	119.86	Jul-14	113.05
Jun-06	140.79	Jul-10	121.74	Aug-14	113.58
Jul-06	141.42	Aug-10	122.67	Sep-14	113.89
Aug-06	141.88	Sep-10	123.21	Oct-14	114.42
Sep-06	142.42	Oct-10	123.29	Nov-14	116.14
Oct-06	143.65	Nov-10	124.03	Dec-14	119
Nov-06	144.14	Dec-10	125.17	Jan-15	118.71
Dec-06	145.89	Jan-11	126.29	Feb-15	118.28
Jan-07	147.41	Feb-11	126.46	Mar-15	118.48
Feb-07	148.32	Mar-11	126.05	Apr-15	118.91
Mar-07	148.67	Apr-11	125.66	May-15	119.50
Apr-07	148.43	May-11	125.81	Jun-15	120.14
May-07	148.58	Jun-11	126.5	Jul-15	121.26
Jun-07	148.92	Jul-11	127.35	Aug-15	121.73
Jul-07	149.99	Aug-11	128.54	Sep-15	121.67
Aug-07	151.11	Sep-11	128.89	Oct-15	121.57
Sep-07	152.32	Oct-11	128.74	Nov-15	121.82
Oct-07	153.53	Nov-11	129.18	Dec-15	122.99
Nov-07	153.81	Dec-11	129.91	Jan-16	123.62
Dec-07	155.5	Jan-12	130.9	Feb-16	123.51
Jan-08	158.26	Feb-12	130.96	Mar-16	123.75
Feb-08	159.29	Mar-12	131.05	Apr-16	123.19
Mar-08	160.81	Apr-12	131.32	May-16	123.48
Apr-08	161.73	May-12	131.41	Jun-16	124.29
May-08	164.01	Jun-12	132.23	Jul-16	125.15
Jun-08	110.08	Jul-12	133.16	Aug-16	125.13
Jul-08	111.59	Aug-12	134.43	Sep-16	125.41
Aug-08	112.16	Sep-12	134.45	Oct-16	125.59
Sep-08	113.25	Oct-12	134.67	Nov-16	126.18
Oct-08	113.76	Nov-12	134.76	Dec-16	126.71
Nov-08	113.9	Dec-12	135.49	Jan-17	127.94
Dec-08	113.86	Jan-13	136.88	Feb-17	128.24
Jan-09	113.78	Feb-13	137.91	Mar-17	128.22
Feb-09	114.02	Mar-13	138.78	Apr-17	128.33

Mar-09	114.27	Apr-13	138.64	May-17	128.83
Apr-09	113.92	May-13	138.6	Jun-17	129.72
May-09	113.97	Jun-13	140.03		
Jun-09	114.1	Jul-13	144.63		
Jul-09	114.61	Aug-13	146.25		
Aug-09	115.25	Sep-13	145.74		
Sep-09	116.46	Oct-13	145.87		
Oct-09	116.68	Nov-13	146.04		
Nov-09	116.65	Dec-13	146.84		
Dec-09	117.03	Jan-14	110.99		
Jan-10	118.01	Feb-14	111.28		

Appendix 2. Code

/*First Code*/

```
/*Import data*/
proc import out=work.Intervensi datafile='D:\DATA IHK.xlsx'
dbms=xlsx replace;
sheet='sheet1';
getnames=yes;
run;

/*show data*/
proc print data=Intervensi;
run;

/*plot time series data */
proc gplot data=Intervensi;
symbol1 i=splines v=dot;
plot IHK*Date;
run;

/*manipulating data*/
data Intervensi;
set Intervensi;
z=0;
*t_data1=log(data1);
run;

/*Stationarity in variance*/
proc transreg maxiter=0 nozeroconstant;
    model BoxCox(data1) = identity(z);
    output;
run;

/* Stationarity in Mean*/
proc arima data=Intervensi;
identify var=data1(1) nlag=40 stationarity=(adf);
run;
```

```
/*Estimate ARIMA model + Forecasting*/
proc arima data=Intervensi;
identify var=data1(1) nlags=25;
run;
estimate p=1 method=ml;
run;
forecast out=ramalan lead=67 printall;
run;

/*Normality Test*/
proc univariate normaltest;
var residual;
run;
```

/*Second Code*/

```
/*Import data*/
proc import out=work.Intervensi datafile='D:\DATA IHK.xlsx'
dbms=xlsx replace;
sheet='sheet1';
getnames=yes;
run;

/*Print data*/
proc print data=Intervensi;
run;

/* ARIMA model + Intervention I*/
proc arima data=Intervensi;
identify var=data2(1) crosscorr=(inter_1);
estimate p=1 input=(0 $ (1) / (0) inter_1) method=ml;
forecast out=ramalan2 lead=42;
run;
```

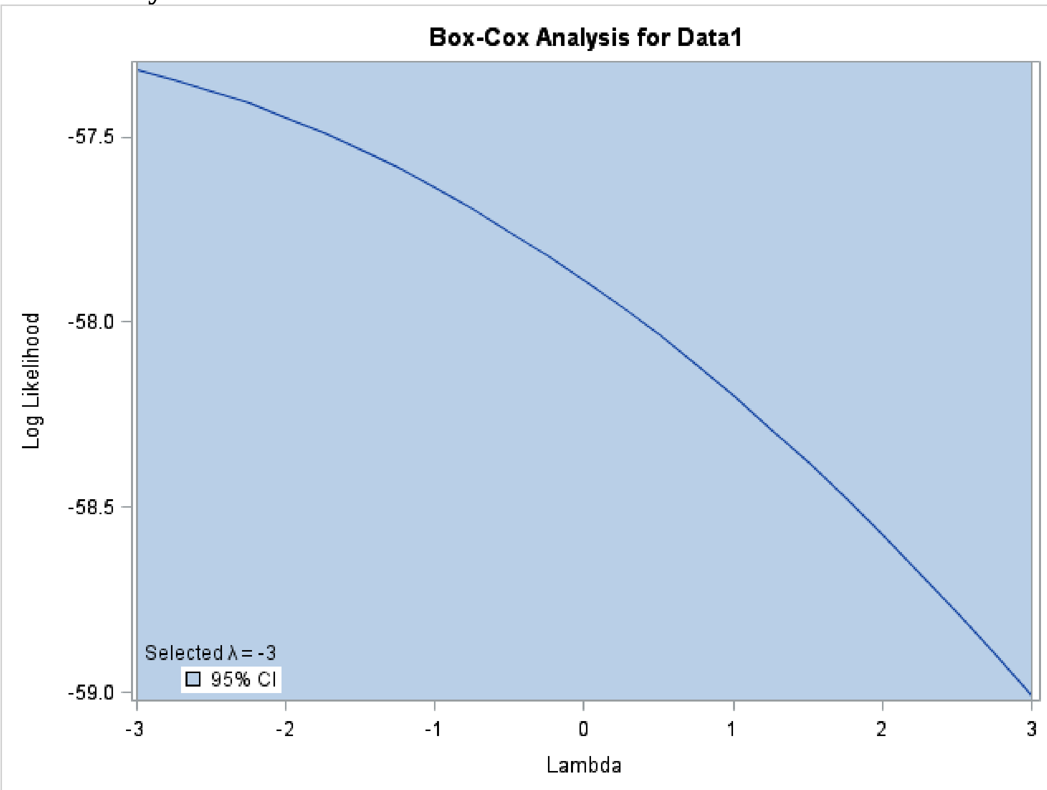
```
/*Normality Test */
proc univariate normaltest;
var residual;
run;

/*Print data*/
proc print data=ramalan2;
run;
```

/*Third Code*/

```
/*Import data*/
proc import out=work.Intervensi datafile='D:\DATA IHK.xlsx'
dbms=xlsx replace;
sheet='sheet1';
getnames=yes;
run;

/*Print data*/
proc print data=Intervensi;
run;

/*ARIMA model + Intervention II*/
proc arima data=Intervensi plots = all;
identify var=data3(1) crosscorr=(inter_1 inter_2);
estimate p=1 input=(0 $ (1) / (0) inter_1 0 $ (1) / (0) inter_2) method=ml;
forecast out=ramalan3 lead=12;
run;

/*Normality Test */
proc univariate normaltest;
var residual;
run;

/*Print data*/
proc print data=ramalan3;
run;
```

Appendix 3. ARIMA Model (Using the First Data)

Stationarity in Variance



Stationarity in Mean (Level)

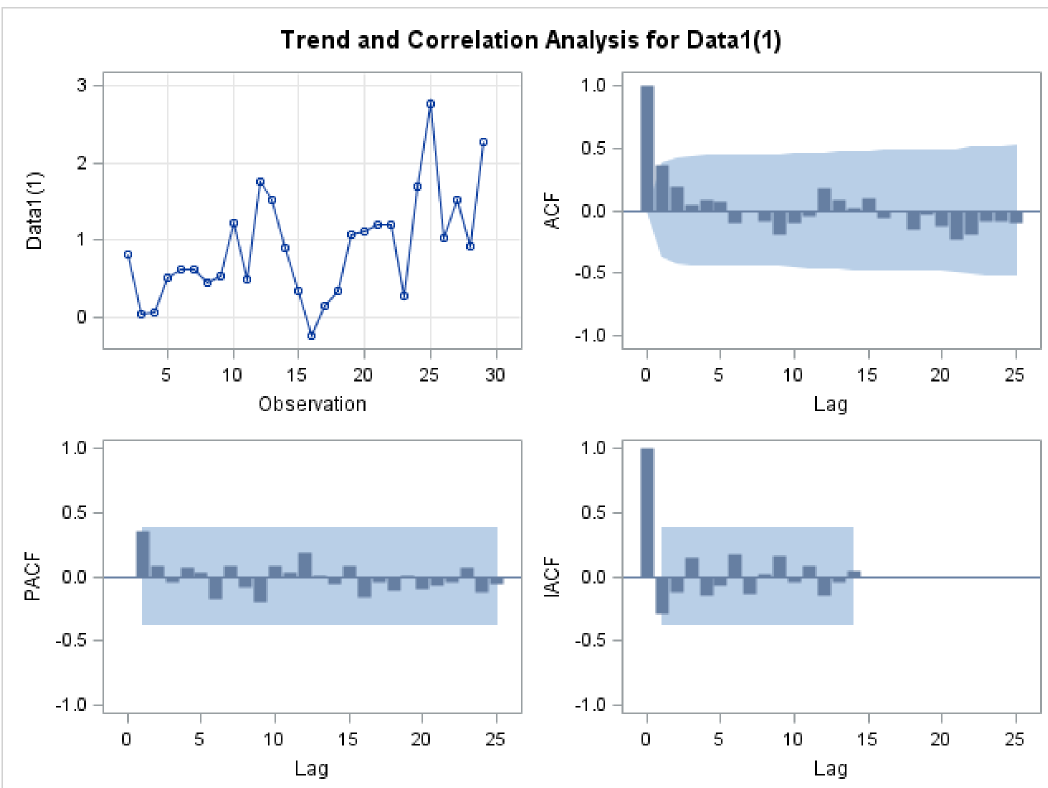
Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	**0**	0.1735	0.7169	7.23	0.9999		
	1	0.1740	0.7154	2.74	0.9976		
	2	0.1773	0.7163	2.39	0.9944		
Single Mean	**0**	1.4347	0.9920	3.05	0.9999	35.81	0.0010
	1	1.5038	0.9926	2.23	0.9999	6.29	0.0157
	2	1.3419	0.9905	1.88	0.9996	4.57	0.0656
Trend	**0**	-0.3060	0.9920	-0.12	0.9917	4.82	0.2495
	1	-2.5754	0.9451	-0.73	0.9602	3.26	0.5429
	2	-1.8083	0.9687	-0.43	0.9804	2.07	0.7687

Stationarity in Mean (First difference)

Augmented Dickey-Fuller Unit Root Tests							
Type	**Lags**	**Rho**	**Pr < Rho**	**Tau**	**Pr < Tau**	**F**	**Pr > F**
Zero Mean	**0**	-4.4875	0.1358	-1.26	0.1853		
	1	-1.6618	0.3662	-0.60	0.4477		
	2	-0.6703	0.5258	-0.28	0.5749		
Single Mean	**0**	-15.5726	0.0168	-2.89	0.0599	4.25	0.0829
	1	-12.9075	0.0396	-2.24	0.1975	2.74	0.3932
	2	-13.5643	0.0308	-2.03	0.2725	2.38	0.5002
Trend	**0**	-21.8828	0.0120	-3.95	0.0237	8.07	0.0208
	1	-24.0720	0.0047	-3.12	0.1234	4.90	0.2340
	2	-36.7404	<.0001	-2.86	0.1912	4.10	0.4158

ACF and PACF



Tentative of ARIMA Model

ARIMA (1,1,1)

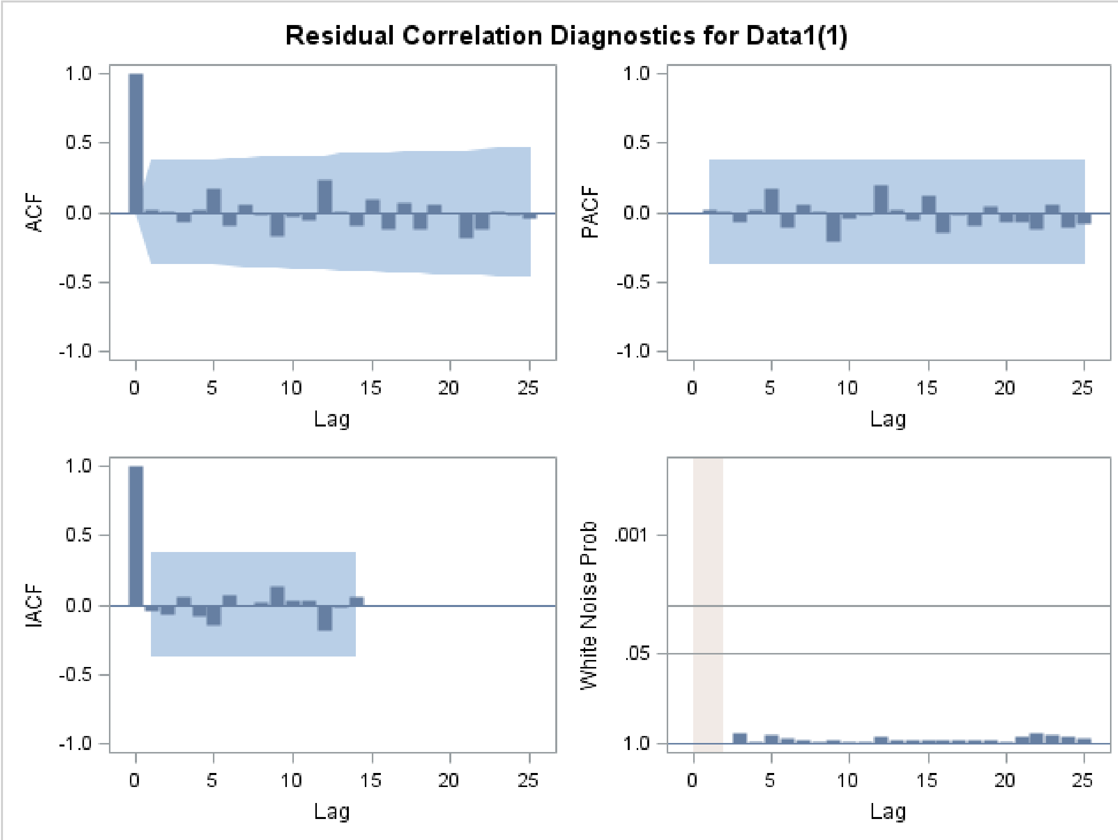
Maximum Likelihood Estimation

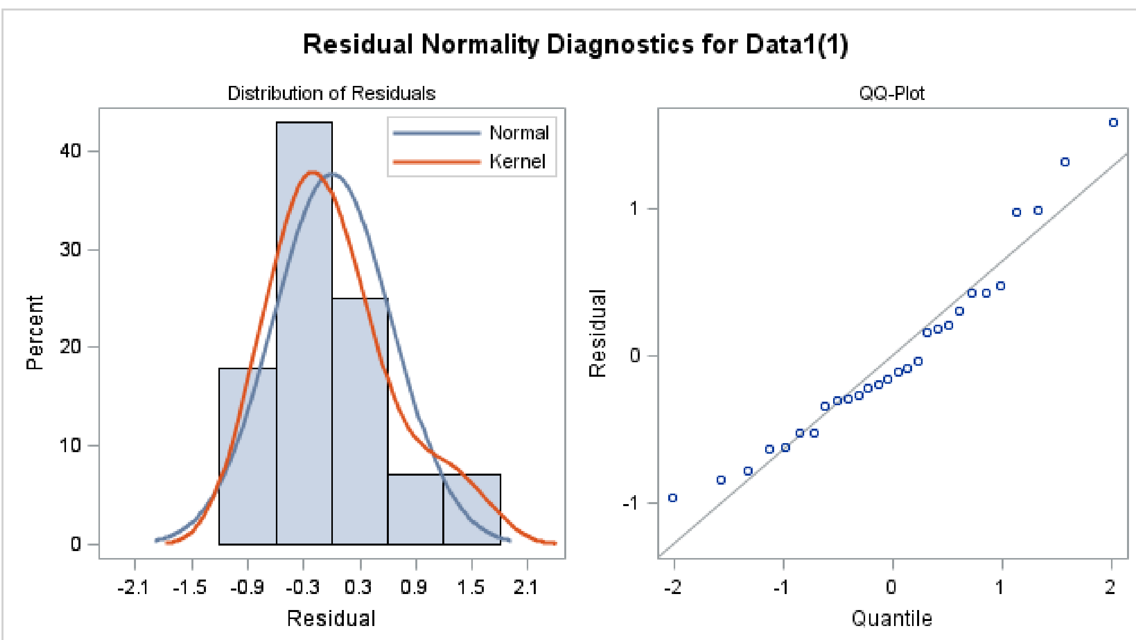
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MU	0.93447	0.22363	4.18	<.0001	0
MA1,1	0.16822	0.51821	0.32	0.7455	1
AR1,1	0.54438	0.44471	1.22	0.2209	1

Constant Estimate	0.425766
Variance Estimate	0.437943
Std Error Estimate	0.661773
AIC	59.3568
SBC	63.35341
Number of Residuals	28

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	1.63	4	0.8043	0.015	0.013	-0.067	0.020	0.173	-0.096
12	6.27	10	0.7924	0.056	-0.015	-0.171	-0.033	-0.062	0.238
18	10.02	16	0.8656	0.010	-0.097	0.097	-0.122	0.074	-0.115
24	16.37	22	0.7972	0.065	-0.009	-0.180	-0.119	0.003	-0.018


Residual Correlation Diagnostics for Data1(1)
ACF
PACF
IACF
White Noise Prob
Lag
1.0
0.5
0.0
-0.5
-1.0
.001
.05
1.0
0
5
10
15
20
25


Residual Normality Diagnostics for Data1(1)
Distribution of Residuals
Normal
Kernel
Percent
Residual
-2.1
-1.5
-0.9
-0.3
0.3
0.9
1.5
2.1
0
10
20
30
40
QQ-Plot
Residual
Quantile
-2
-1
0
1
2

ARIMA(0,1,1)

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MU	0.92139	0.16540	5.57	<.0001	0
MA1,1	-0.33099	0.21280	-1.56	0.1199	1

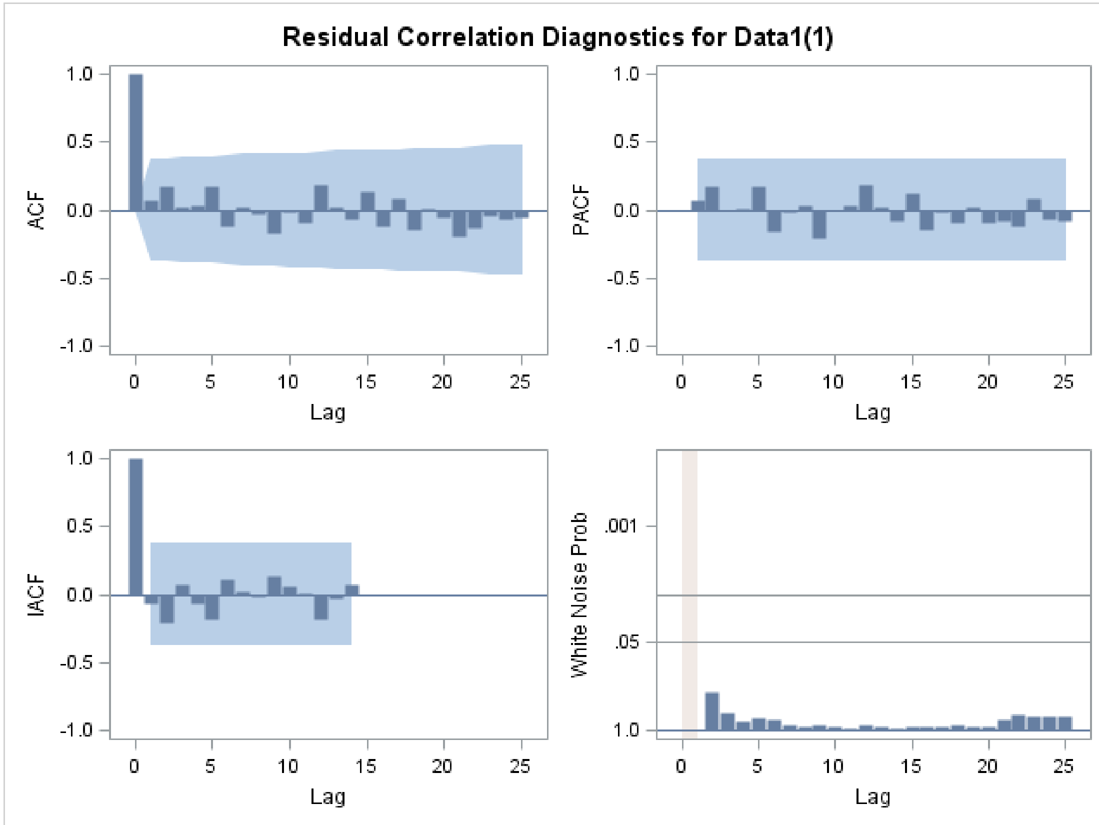
Constant Estimate	0.921393
Variance Estimate	0.439243
Std Error Estimate	0.662754
AIC	58.46591
SBC	61.13032
Number of Residuals	28

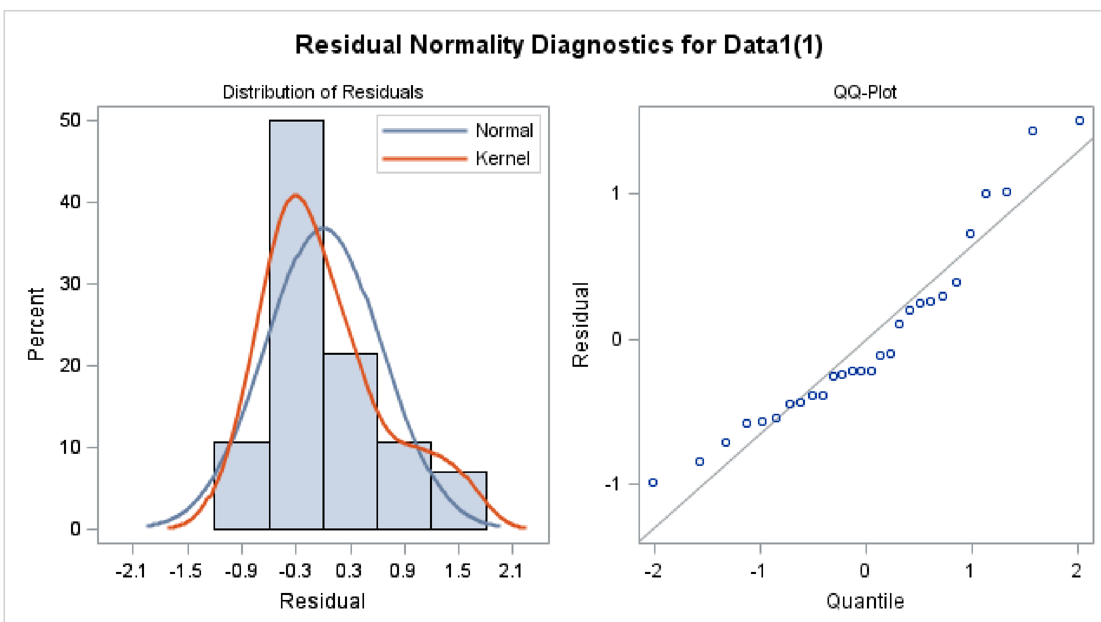
Correlations of Parameter Estimates

Parameter	MU	MA1,1
MU	1.000	-0.078
MA1,1	-0.078	1.000

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	2.93	5	0.7105	0.068	0.178	0.015	0.035	0.176	-0.123
12	6.60	11	0.8302	0.024	-0.027	-0.171	-0.014	-0.089	0.191
18	11.45	17	0.8324	0.020	-0.063	0.130	-0.123	0.090	-0.147
24	20.20	23	0.6300	0.008	-0.058	-0.194	-0.137	-0.045	-0.065


Residual Correlation Diagnostics for Data1(1)
ACF
PACF
IACF
White Noise Prob
Lag
1.0
0.5
0.0
-0.5
-1.0
.001
.05
0
5
10
15
20
25


Residual Normality Diagnostics for Data1(1)
Distribution of Residuals
QQ-Plot
Normal
Kernel
Percent
Residual
Quantile
50
40
30
20
10
0
-2.1
-1.5
-0.9
-0.3
0.3
0.9
1.5
2.1
1
-1
-2
2

ARIMA(1,1,0)

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MU	0.93299	0.20273	4.60	<.0001	0
AR1,1	0.40517	0.19552	2.07	0.0382	1

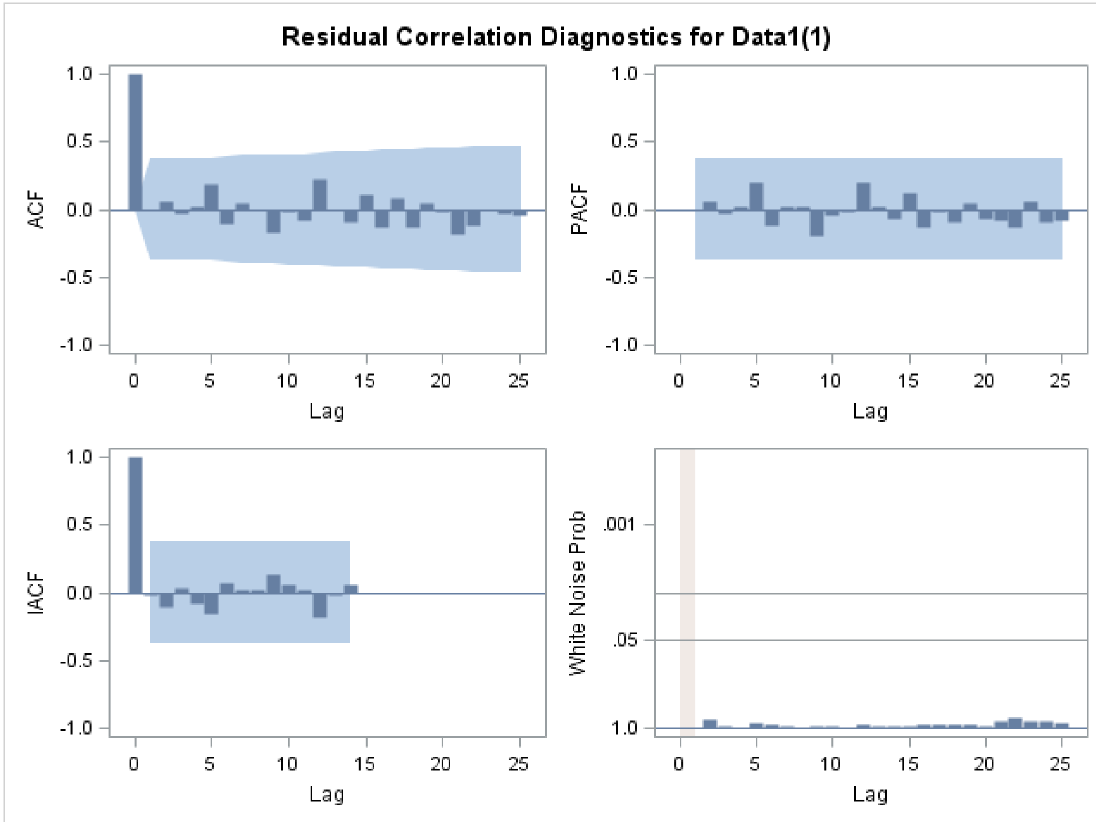
Constant Estimate	0.554971
Variance Estimate	0.422898
Std Error Estimate	0.650306
AIC	57.46736
SBC	60.13177
Number of Residuals	28

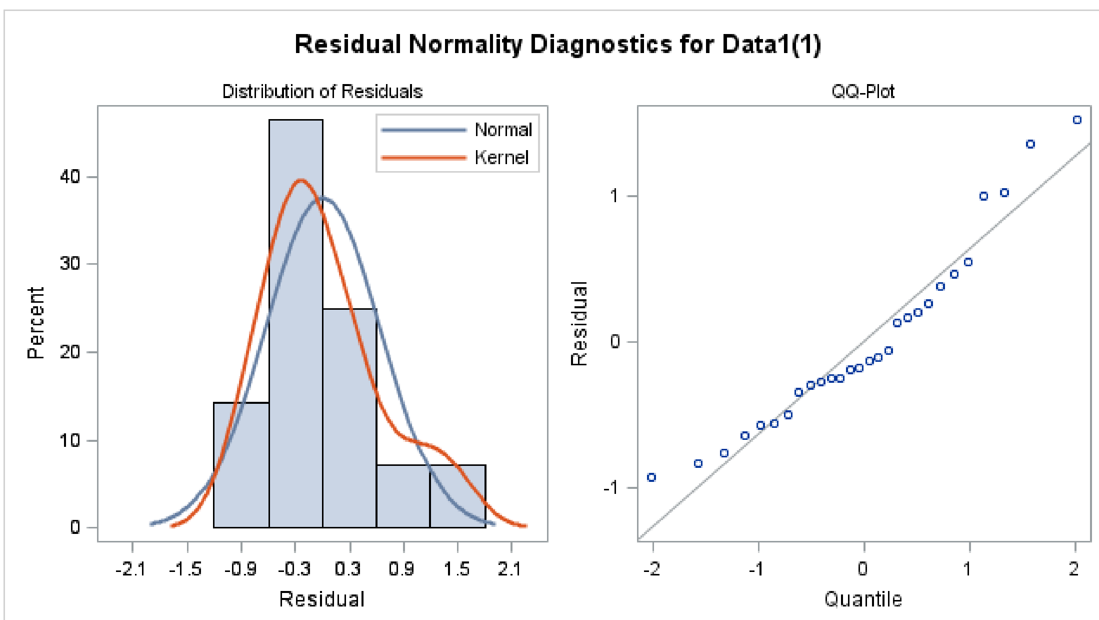
Correlations of Parameter Estimates

Parameter	MU	AR1,1
MU	1.000	0.125
AR1,1	0.125	1.000

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	1.92	5	0.8597	-0.009	0.062	-0.033	0.022	0.187	-0.111
12	6.20	11	0.8599	0.047	-0.011	-0.165	-0.015	-0.079	0.225
18	10.76	17	0.8688	0.001	-0.098	0.115	-0.129	0.090	-0.127
24	17.21	23	0.7987	0.049	-0.016	-0.182	-0.120	-0.006	-0.032


Residual Correlation Diagnostics for Data1(1)
ACF
PACF
IACF
White Noise Prob
Lag
1.0
0.5
0.0
-0.5
-1.0
.001
.05
0
5
10
15
20
25


Residual Normality Diagnostics for Data1(1)
Distribution of Residuals
Normal
Kernel
Percent
Residual
QQ-Plot
Quantile
40
30
20
10
0
-2.1
-1.5
-0.9
-0.3
0.3
0.9
1.5
2.1
1
-1
-2
2

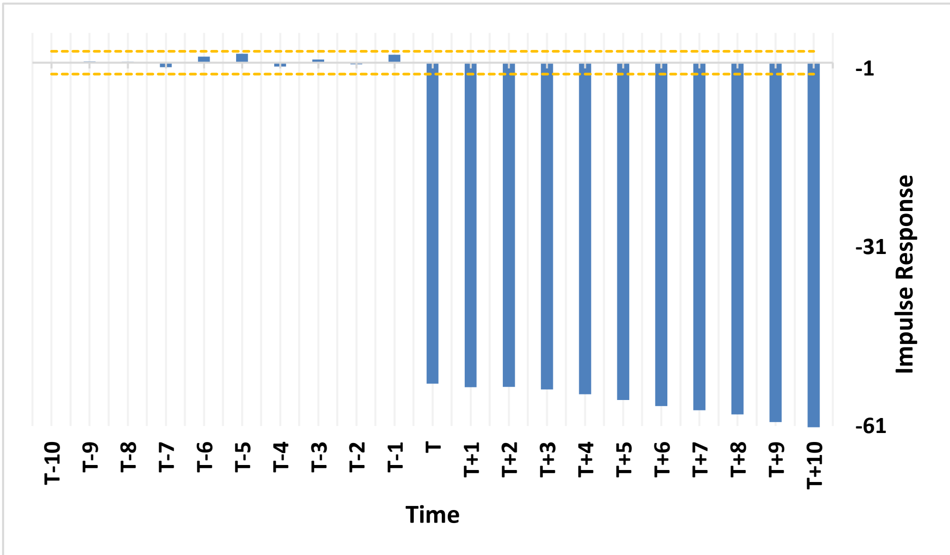
Diagnostics of ARIMA(1,1,0) Model

Tests for Normality

Test	Statistic		p Value	
Shapiro-Wilk	**W**	0.933811	**Pr < W**	0.0770
Kolmogorov-Smirnov	**D**	0.148927	**Pr > D**	0.1097
Cramer-von Mises	**W-Sq**	0.103506	**Pr > W-Sq**	0.0969
Anderson-Darling	**A-Sq**	0.639716	**Pr > A-Sq**	0.0886

Appendix 4. Intervention I Modeling

Order Determination of b, r, and s



Tentative of Intervention I Model

(0,1,1)

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	0.93201	0.20480	4.55	<.0001	0	Data2	0
AR1,1	0.39410	0.09874	3.99	<.0001	1	Data2	0
NUM1	-55.38506	0.68774	-80.53	<.0001	0	Inter_1	0
NUM1,1	-54.98985	0.69024	-79.67	<.0001	1	Inter_1	0
DEN1,1	-0.01422	0.01218	-1.17	0.2429	1	Inter_1	0

Constant Estimate	0.564701
Variance Estimate	0.450389
Std Error Estimate	0.67111
AIC	198.8545
SBC	211.6239

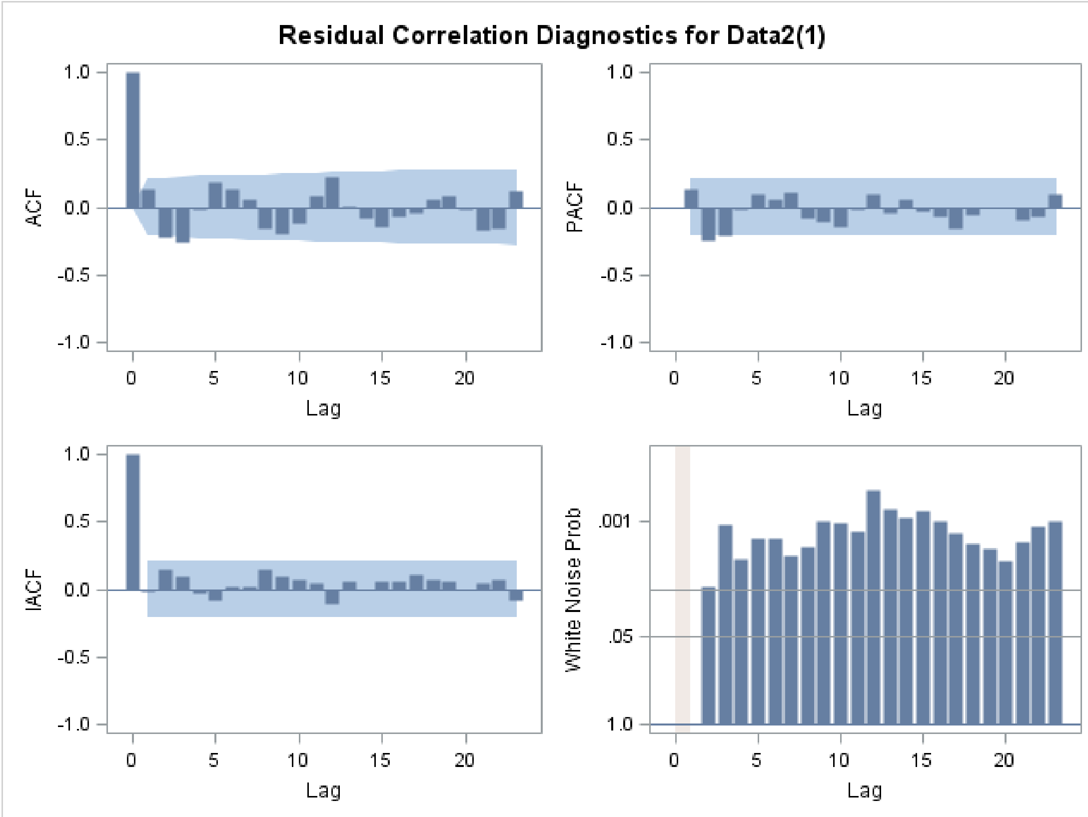
Number of Residuals	95

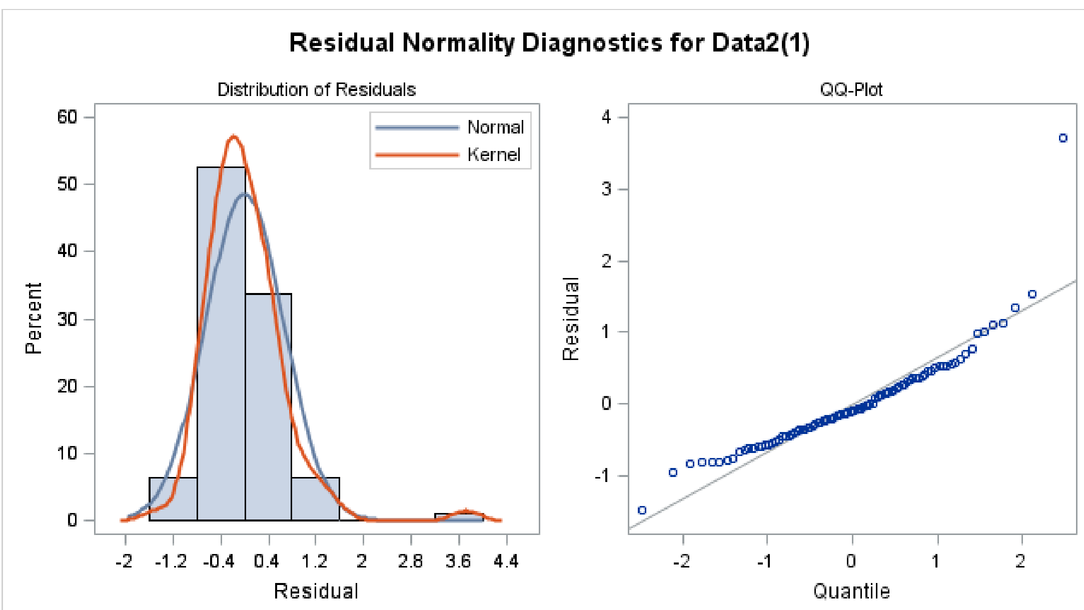
Correlations of Parameter Estimates

Variable Parameter	Data2 MU	Data2 AR1,1	Inter_1 NUM1	Inter_1 NUM1,1	Inter_1 DEN1,1
Data2 MU	1.000	0.062	-0.194	0.107	-0.033
Data2 AR1,1	0.062	1.000	-0.193	-0.172	0.117
Inter_1 NUM1	-0.194	-0.193	1.000	0.934	-0.344
Inter_1 NUM1,1	0.107	-0.172	0.934	1.000	-0.384
Inter_1 DEN1,1	-0.033	0.117	-0.344	-0.384	1.000

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	19.26	5	0.0017	0.130	-0.227	-0.260	-0.018	0.189	0.139
12	34.12	11	0.0003	0.064	-0.154	-0.191	-0.125	0.088	0.222
18	38.46	17	0.0021	0.006	-0.081	-0.143	-0.072	-0.044	0.058
24	50.21	23	0.0009	0.088	-0.018	-0.172	-0.157	0.126	0.122


Residual Correlation Diagnostics for Data2(1)
ACF
PACF
IACF
White Noise Prob
Lag
1.0
0.5
0.0
-0.5
-1.0
.001
.05
0
5
10
15
20


Residual Normality Diagnostics for Data2(1)
Distribution of Residuals
Normal
Kernel
Percent
Residual
QQ-Plot
Quantile
60
50
40
30
20
10
0
-2
-1.2
-0.4
0.4
1.2
2
2.8
3.6
4.4
4
3
2
1
-1

(0,0,1)

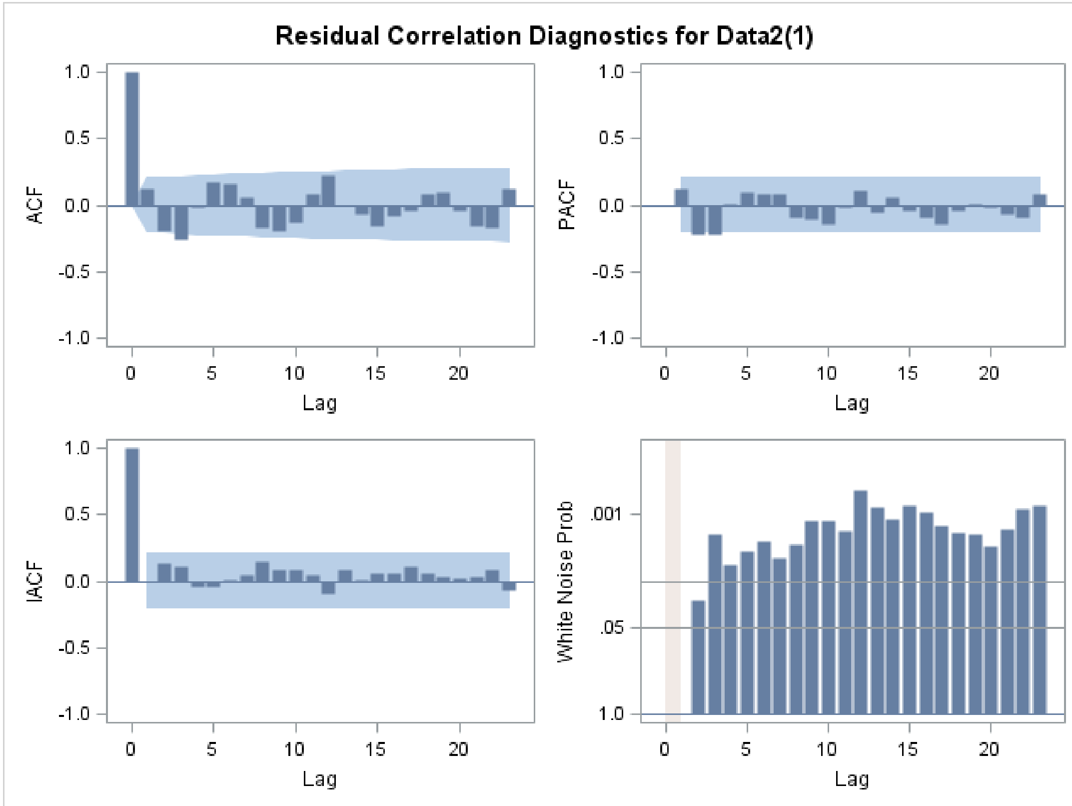
Maximum Likelihood Estimation

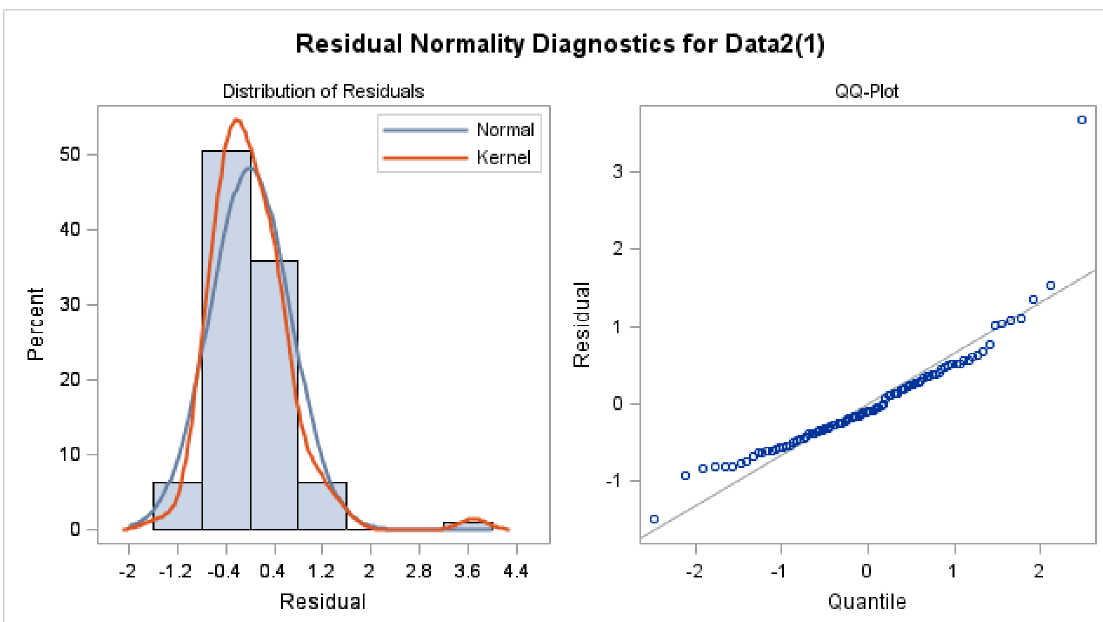
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	0.92314	0.20855	4.43	<.0001	0	Data2	0
AR1,1	0.40511	0.09672	4.19	<.0001	1	Data2	0
NUM1	-55.65524	0.64596	-86.16	<.0001	0	Inter_1	0
NUM1,1	-55.29406	0.63801	-86.67	<.0001	1	Inter_1	0

Constant Estimate	0.549169
Variance Estimate	0.45209
Std Error Estimate	0.672377
AIC	198.2729
SBC	208.4885
Number of Residuals	95

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	18.41	5	0.0025	0.124	-0.199	-0.266	-0.016	0.175	0.163
12	33.53	11	0.0004	0.060	-0.167	-0.190	-0.127	0.083	0.221
18	38.91	17	0.0018	-0.008	-0.073	-0.158	-0.084	-0.047	0.082
24	51.89	23	0.0005	0.100	-0.043	-0.164	-0.172	0.121	0.141





Tests for Normality

Test		Statistic		p Value
Shapiro-Wilk	**W**	0.873279	**Pr < W**	<0.0001

Tests for Normality

Test		Statistic	p Value	
Kolmogorov-Smirnov	**D**	0.092831	**Pr > D**	0.0429
Cramer-von Mises	**W-Sq**	0.230295	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	1.603153	**Pr > A-Sq**	<0.0050

(0,1,0)

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	2.88598	1.18039	2.44	0.0145	0	Data2	0
AR1,1	0.11287	0.11154	1.01	0.3116	1	Data2	0
NUM1	-7.37658	2.49951	-2.95	0.0032	0	Inter_1	0
DEN1,1	-0.87084	0.12040	-7.23	<.0001	1	Inter_1	0

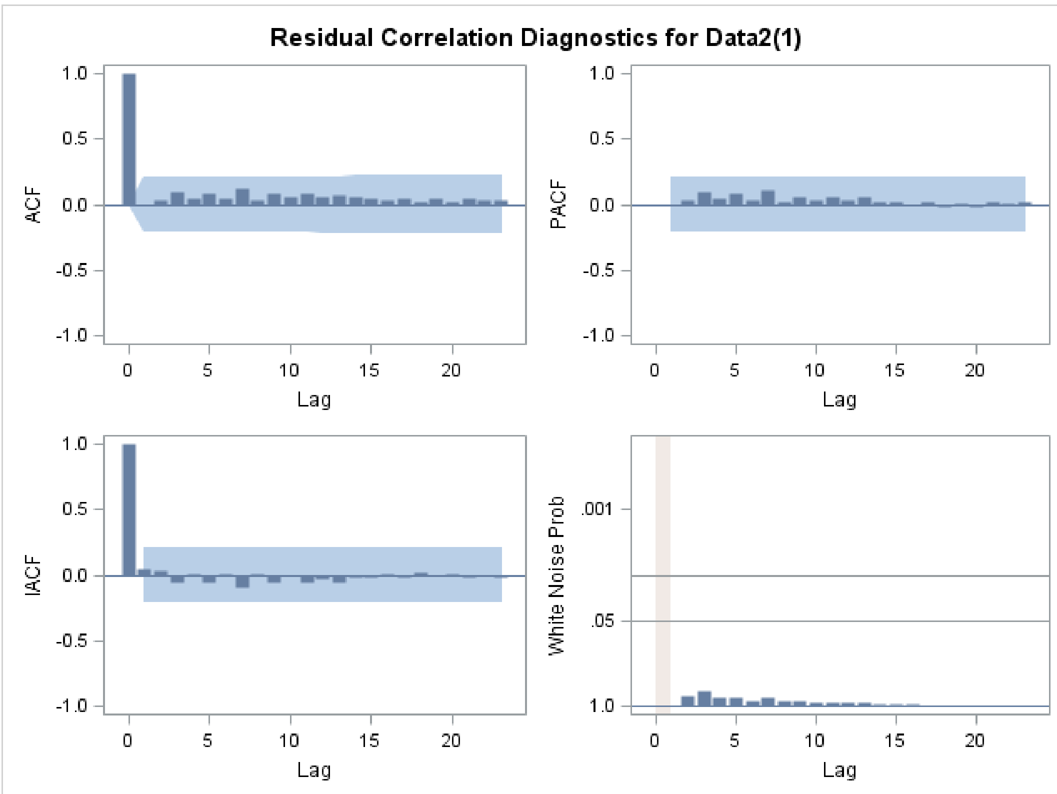
Constant Estimate	2.560245
Variance Estimate	30.34562
Std Error Estimate	5.508686
AIC	597.7264
SBC	607.9419
Number of Residuals	95

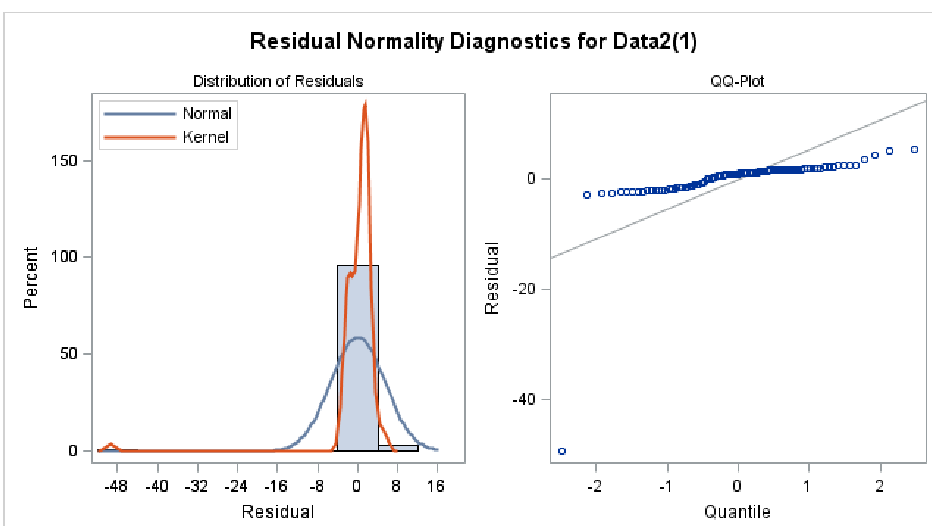
Correlations of Parameter Estimates

Variable Parameter	Data2 MU	Data2 AR1,1	Inter_1 NUM1	Inter_1 DEN1,1
Data2 MU	1.000	0.299	-0.829	0.284
Data2 AR1,1	0.299	1.000	-0.356	0.071
Inter_1 NUM1	-0.829	-0.356	1.000	-0.165
Inter_1 DEN1,1	0.284	0.071	-0.165	1.000

Autocorrelation Check of Residuals

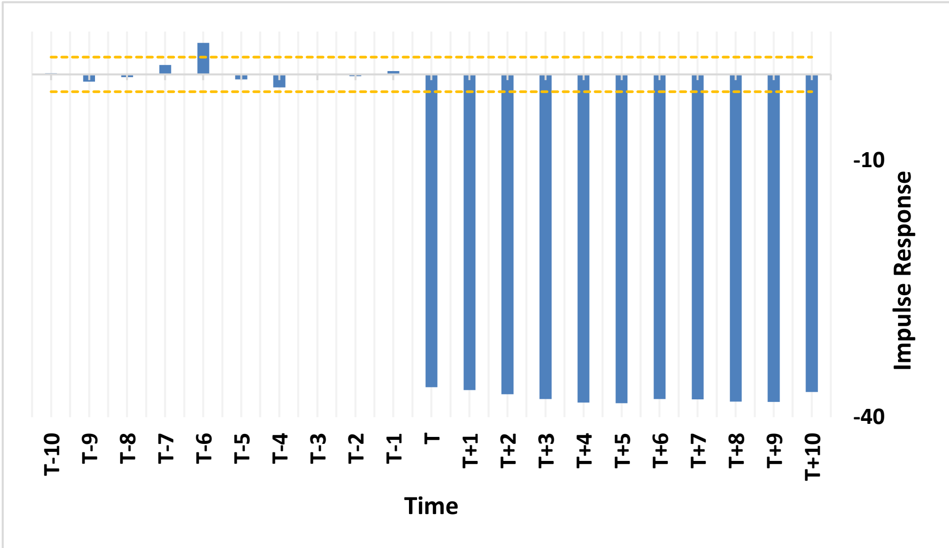
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	2.13	5	0.8315	-0.004	0.037	0.096	0.041	0.084	0.040
12	5.88	11	0.8814	0.119	0.037	0.078	0.054	0.079	0.064
18	7.65	17	0.9735	0.077	0.056	0.047	0.039	0.046	0.020
24	8.64	23	0.9970	0.043	0.024	0.050	0.035	0.038	0.018





Appendix 5. Intervention II Modeling

Order Determination of b, r, and s



Tentative of Intervention II Model

(0,1,1)

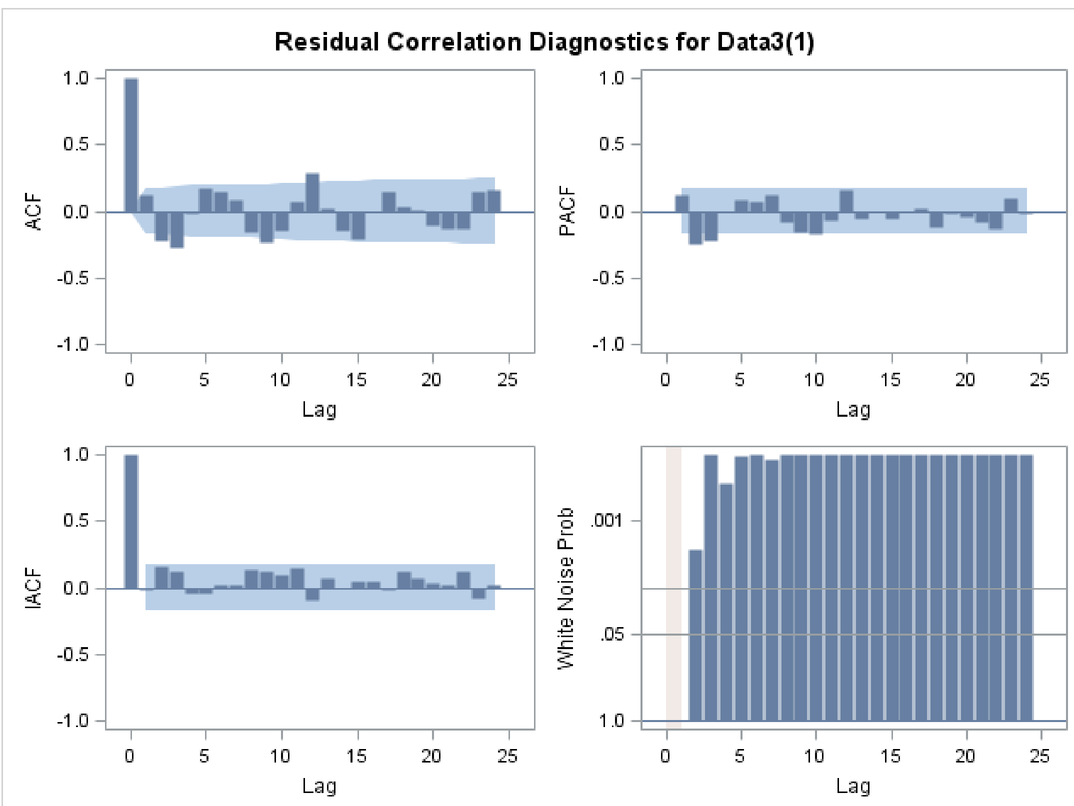
Maximum Likelihood Estimation

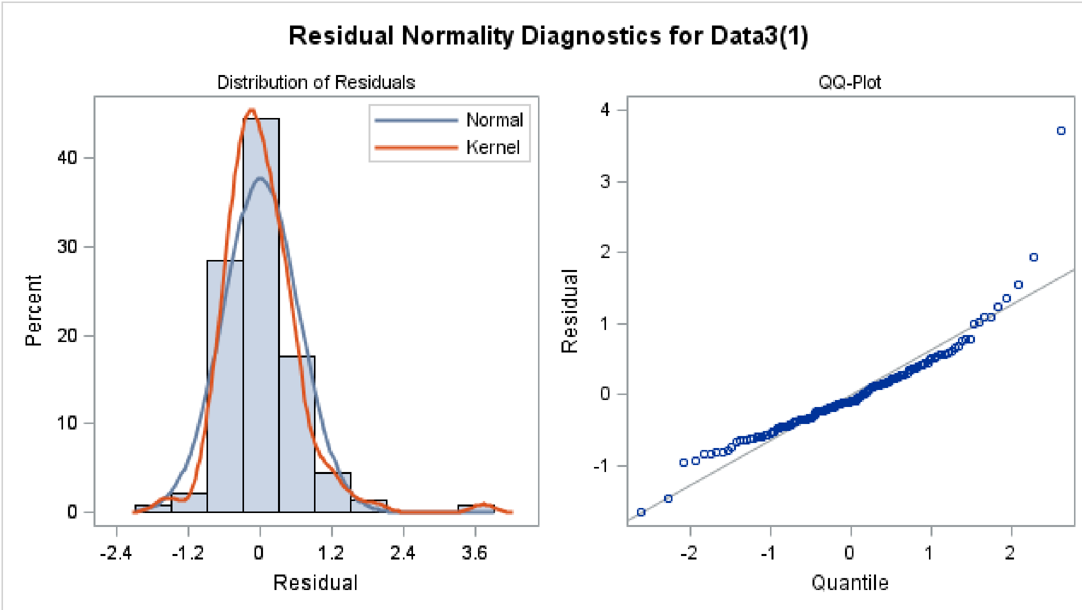
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	0.92122	0.19140	4.81	<.0001	0	Data3	0
AR1,1	0.37143	0.08226	4.52	<.0001	1	Data3	0
NUM1	-55.60395	0.62950	-88.33	<.0001	0	Inter_1	0
NUM1,1	-55.24525	0.62226	-88.78	<.0001	1	Inter_1	0
NUM2	-36.45534	0.64981	-56.10	<.0001	0	Inter_2	0
NUM1,1	-36.35418	0.66232	-54.89	<.0001	1	Inter_2	0
DEN1,1	0.0017846	0.01797	0.10	0.9209	1	Inter_2	0

Constant Estimate	0.579048
Variance Estimate	0.422405
Std Error Estimate	0.649927
AIC	277.6871
SBC	298.1269
Number of Residuals	137

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	26.55	5	<.0001	0.119	-0.223	-0.266	-0.013	0.172	0.148
12	55.90	11	<.0001	0.079	-0.154	-0.233	-0.151	0.067	0.289
18	69.70	17	<.0001	0.015	-0.148	-0.206	-0.008	0.148	0.038
24	84.90	23	<.0001	0.007	-0.101	-0.128	-0.129	0.152	0.159





(0,0,1)

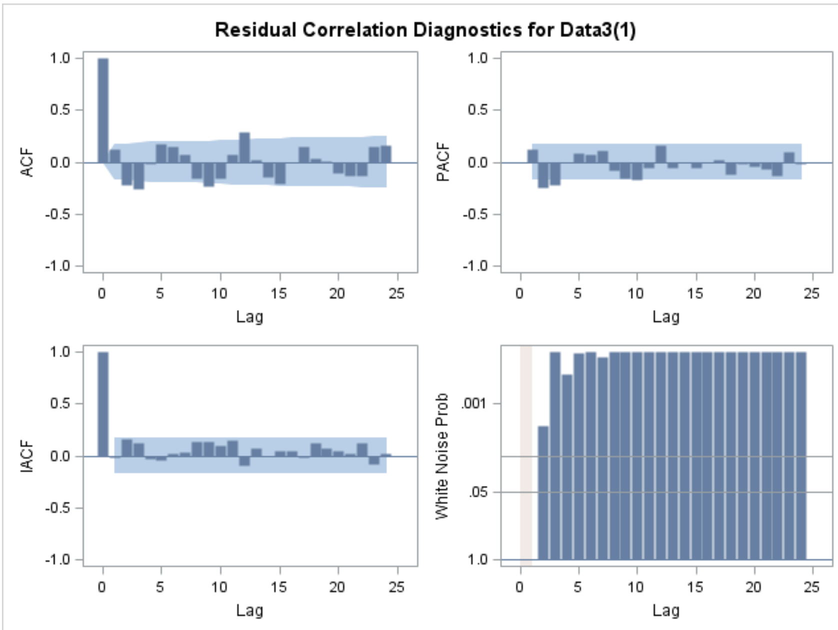
Maximum Likelihood Estimation

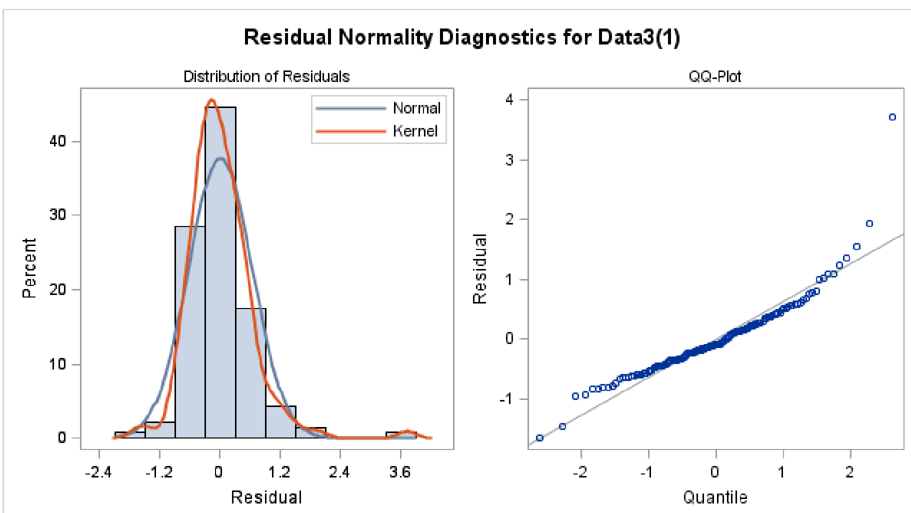
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	0.92123	0.19076	4.83	<.0001	0	Data3	0
AR1,1	0.37167	0.08192	4.54	<.0001	1	Data3	0
NUM1	-55.60426	0.62705	-88.68	<.0001	0	Inter_1	0
NUM1,1	-55.24584	0.61985	-89.13	<.0001	1	Inter_1	0
NUM2	-36.43518	0.61476	-59.27	<.0001	0	Inter_2	0
NUM1,1	-36.33104	0.61747	-58.84	<.0001	1	Inter_2	0

Constant Estimate	0.578839
Variance Estimate	0.419211
Std Error Estimate	0.647465
AIC	275.6973
SBC	293.2171
Number of Residuals	137

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	26.46	5	<.0001	0.119	-0.223	-0.265	-0.013	0.173	0.146
12	55.79	11	<.0001	0.075	-0.152	-0.233	-0.154	0.067	0.290
18	69.52	17	<.0001	0.015	-0.148	-0.206	-0.007	0.148	0.037
24	84.76	23	<.0001	0.007	-0.101	-0.127	-0.129	0.152	0.160





Tests for Normality

Test		Statistic		p Value
Shapiro-Wilk	**W**	0.89473	**Pr < W**	<0.0001
Kolmogorov-Smirnov	**D**	0.082939	**Pr > D**	0.0211
Cramer-von Mises	**W-Sq**	0.328746	**Pr > W-Sq**	<0.0050
Anderson-Darling	**A-Sq**	2.134719	**Pr > A-Sq**	<0.0050

(0,1,0)

Maximum Likelihood Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	0.91052	0.67958	1.34	0.1803	0	Data3	0
AR1,1	0.14628	0.09179	1.59	0.1110	1	Data3	0
NUM1	-55.08215	3.11375	-17.69	<.0001	0	Inter_1	0
NUM1,1	-55.35476	3.08692	-17.93	<.0001	1	Inter_1	0
NUM2	-4.73765	1.26880	-3.73	0.0002	0	Inter_2	0
DEN1,1	-0.86756	0.10481	-8.28	<.0001	1	Inter_2	0

Constant Estimate	0.777336
Variance Estimate	9.541416
Std Error Estimate	3.088918
AIC	703.6984
SBC	721.2183
Number of Residuals	137

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	4.26	5	0.5125	-0.006	0.029	0.108	0.061	0.081	-0.082
12	6.56	11	0.8334	0.070	0.048	0.075	-0.012	-0.016	0.048
18	7.79	17	0.9708	0.061	0.037	0.034	0.037	0.016	-0.002

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
24	8.26	23	0.9979	0.017	0.033	0.029	0.022	0.012	-0.003