Student Project
Student Name: Joel A. Espino
VEE Course: Time Series
Session: Spring 2015
Company: The Insular Life Assurance Co., Ltd.
Mobile: +63922 802 6331
E-mail: joelespino09@yahoo.com

INTRODUCTION

Electricity plays a huge role in our world today. Virtually everything in this information age depends on it. It can be easily validated that the world's electricity consumption is increasing every year brought about by many factors like developments, business expansions, and growing population. This project will try to look closer at the annual electricity consumption of the Philippines and construct time series models for this.

SCOPE AND DATA

From the www.doe.gov.ph website, total annual power consumption (in Gwh) for years 1991-2014 were gathered, specifically for the Residential, Commercial, and Industrial sectors. Gathered data can be seen in Exhibit A. Each set of data will be analyzed.

PRE-ANALYSIS

Methodology and Initial Results

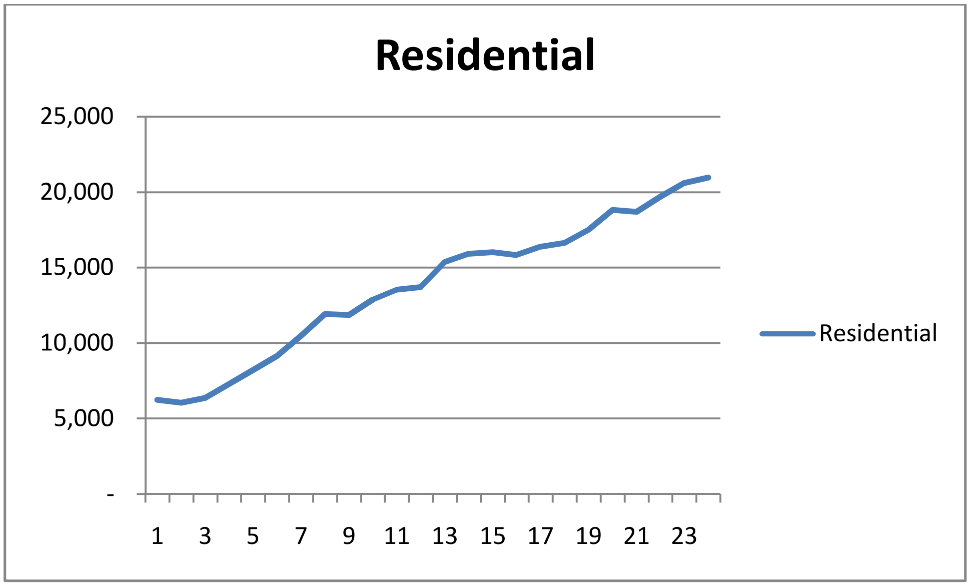
First differences, log-transformations, and first differences of the log-transformations were determined for each set of data. These new sets of data and the original sets of data were graphed and these can be seen in the succeeding pages. It is noticeable that for all sectors, the original sets of data and their log-transformations generally follow an increasing trend, while the sets of data for the first differences exhibit non-stationarity. Next, the autocorrelations for all sets of data were determined manually using an Excel spreadsheet by following the formula (for autocorrelation):

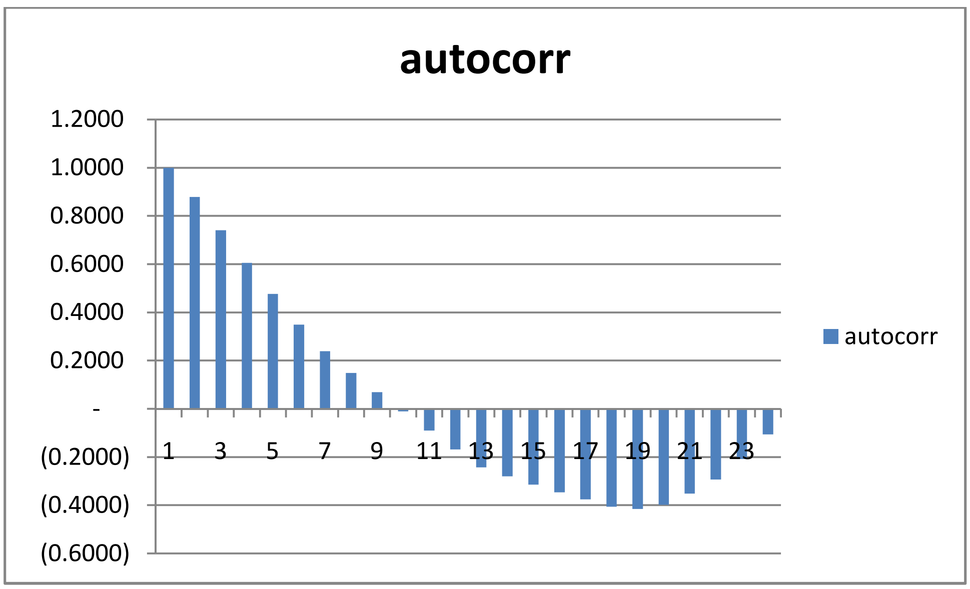
$$autocorr(k) = \frac{\Sigma_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\Sigma_{t=1}^{n}(Y_t - \bar{Y})^2}$$

After producing all the autocorrelation values, these were graphed and these can also be seen in the succeeding pages. Again, it is also noticeable that for all sectors, autocorrelations of the original sets of data and their log-transformations follow a similar trend which is positive on earlier lags and negative on later lags, while autocorrelations of the sets of data for the first differences again support non-stationarity.

Part 1: Residential Sector

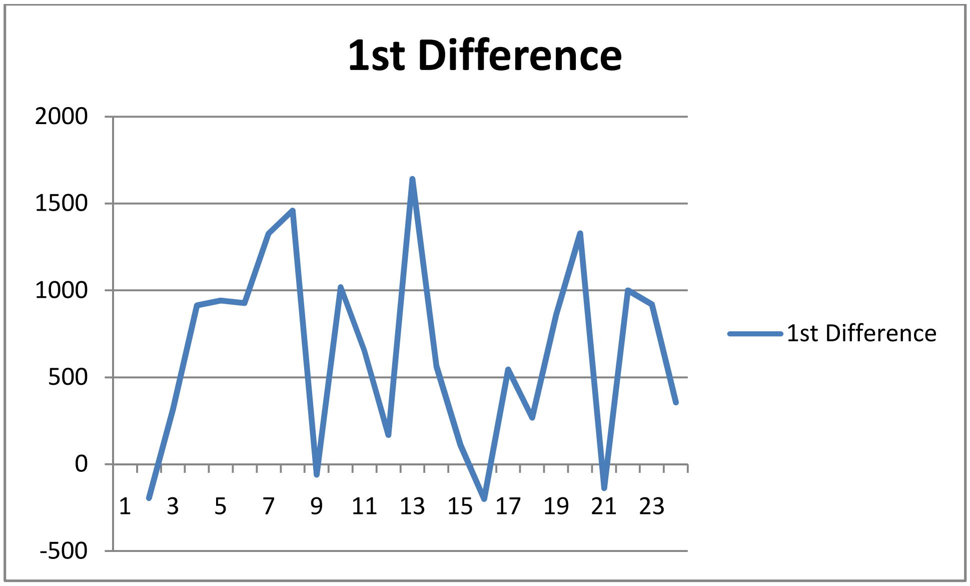
<u>Original Data</u>

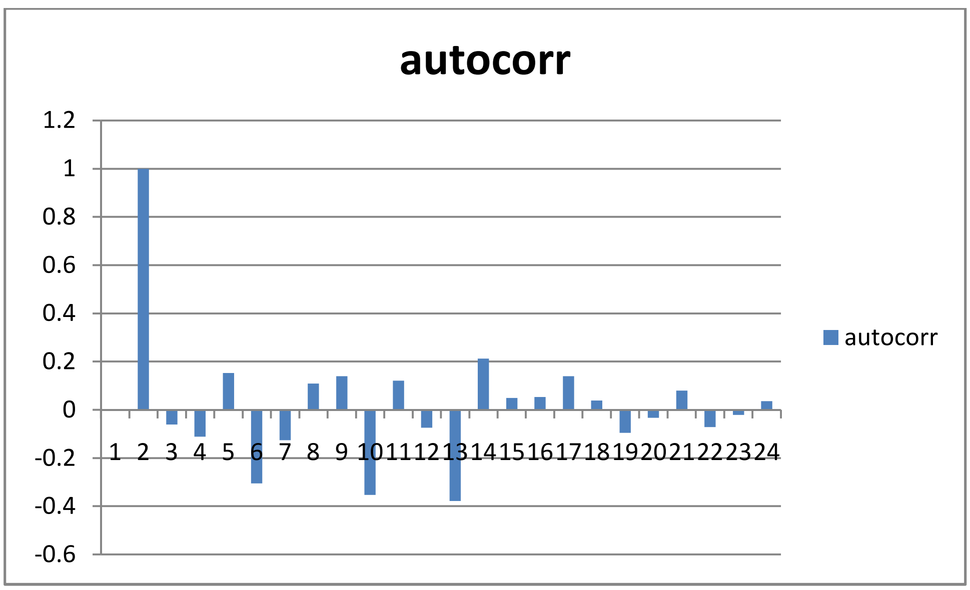




(continuation of ***Part 1: Residential Sector***)

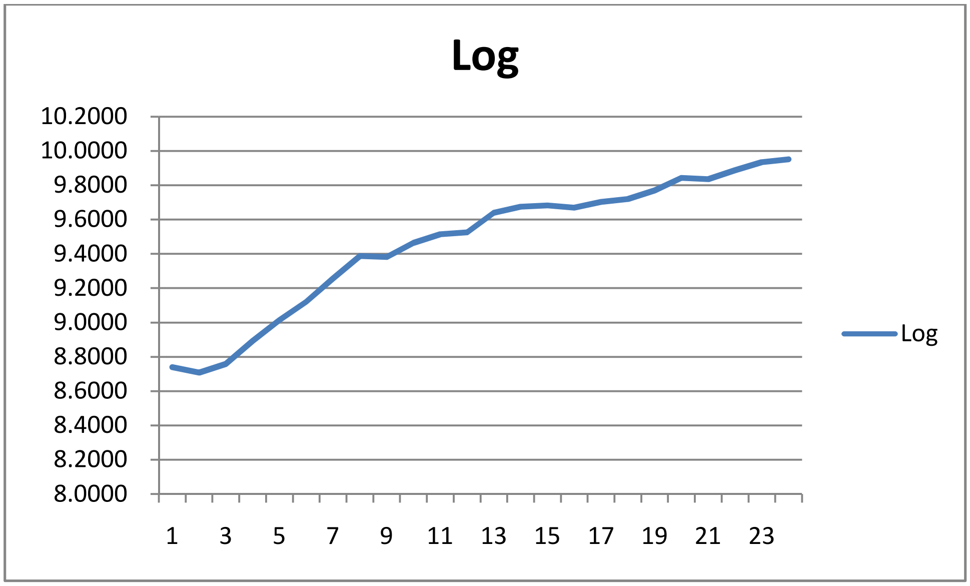
First Differences

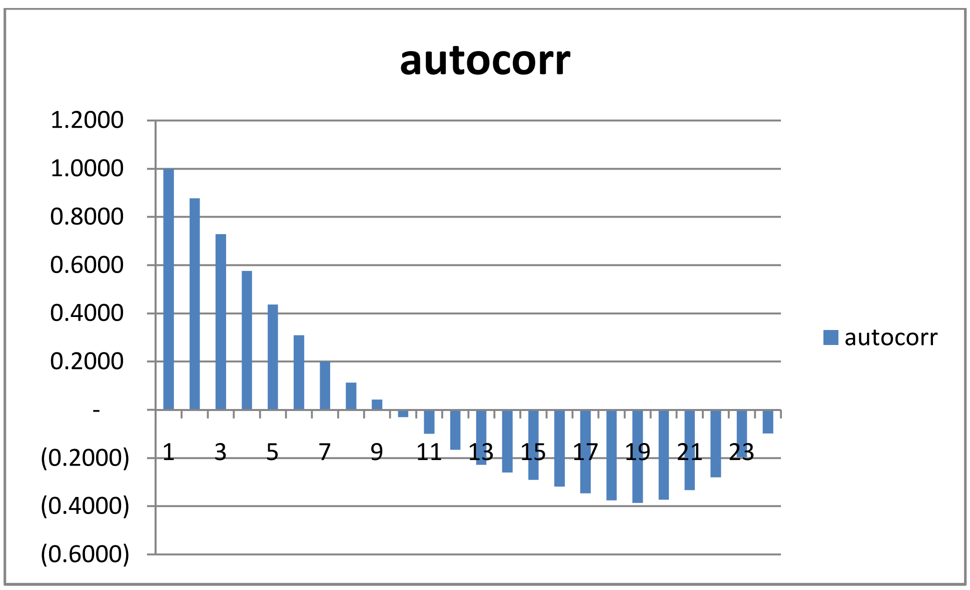




(continuation of ***Part 1: Residential Sector***)

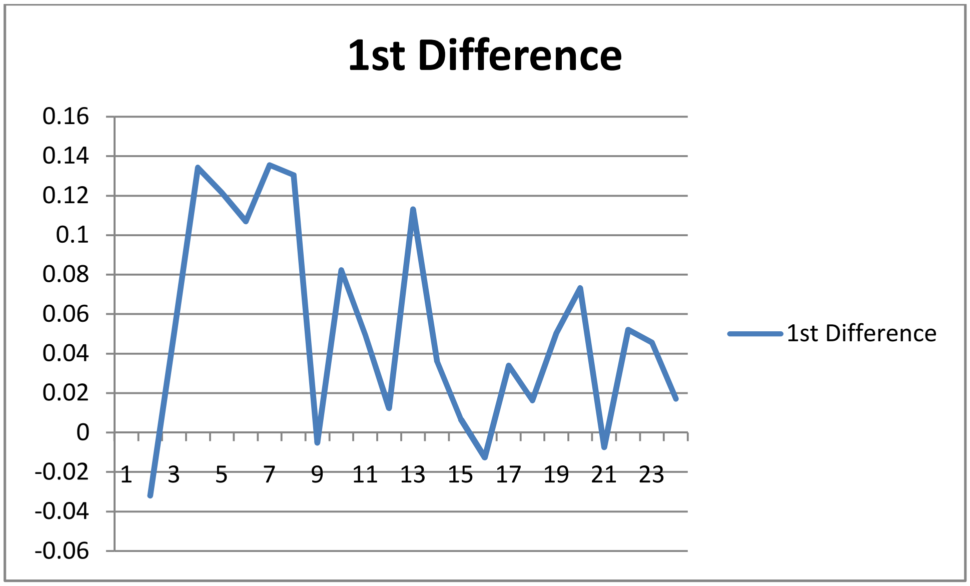
Log-Transformation

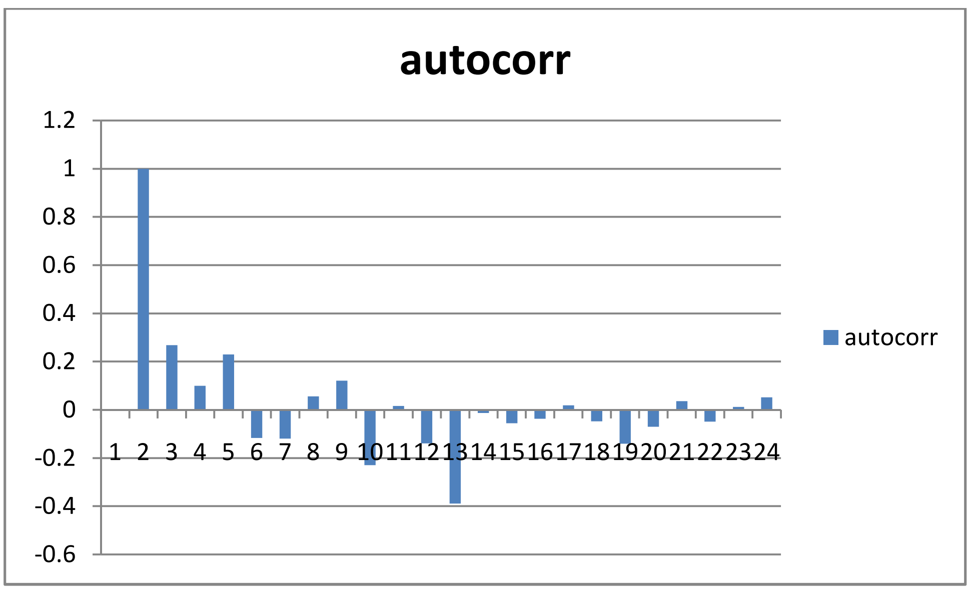




(continuation of ***Part 1: Residential Sector***)

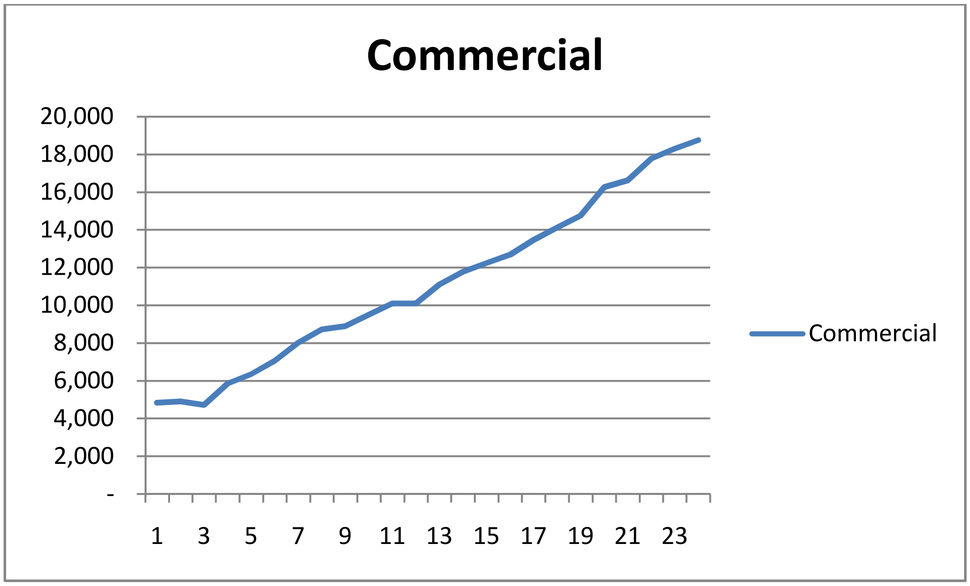
First Differences of the Log-Transformation

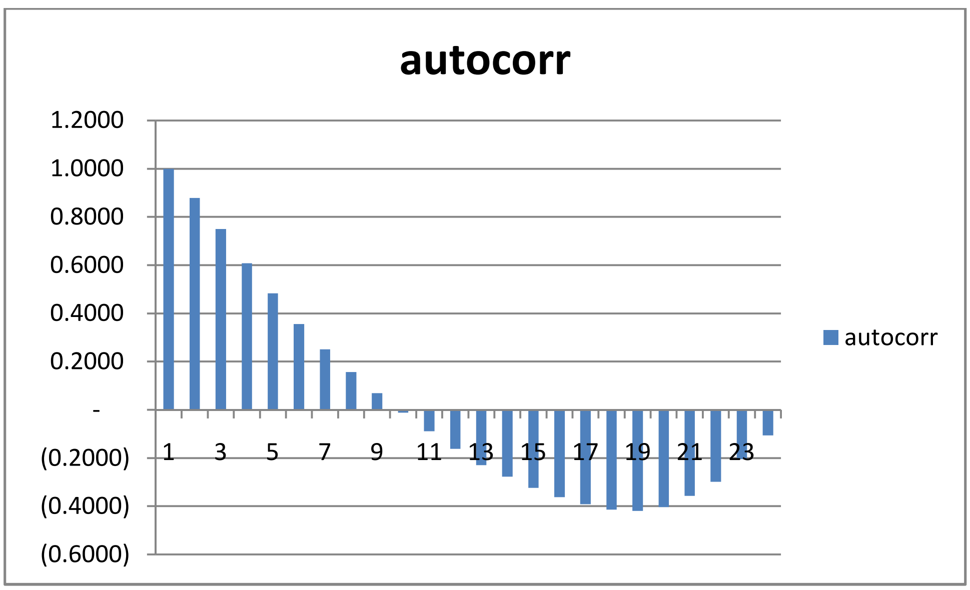




Part 2: Commercial Sector

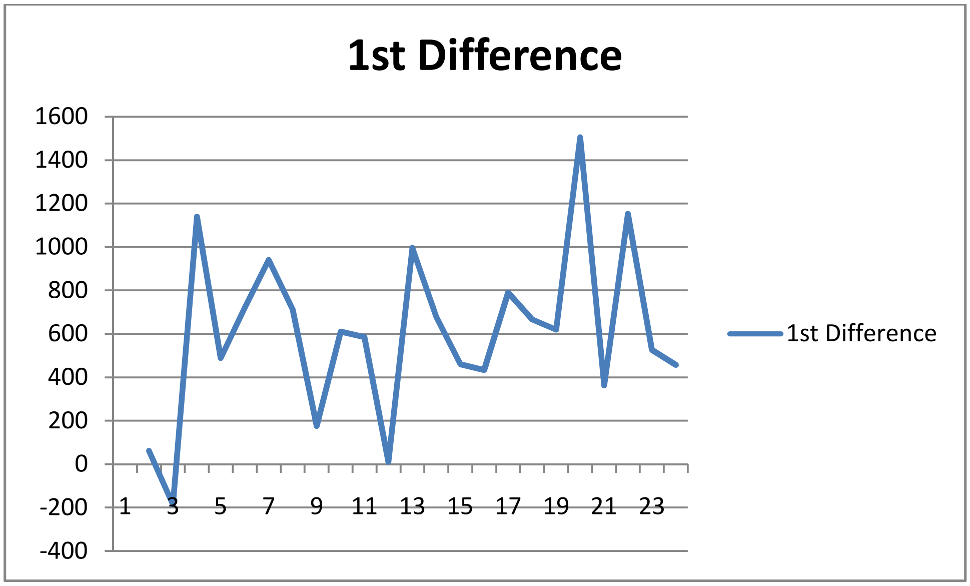
<u>Original Data</u>

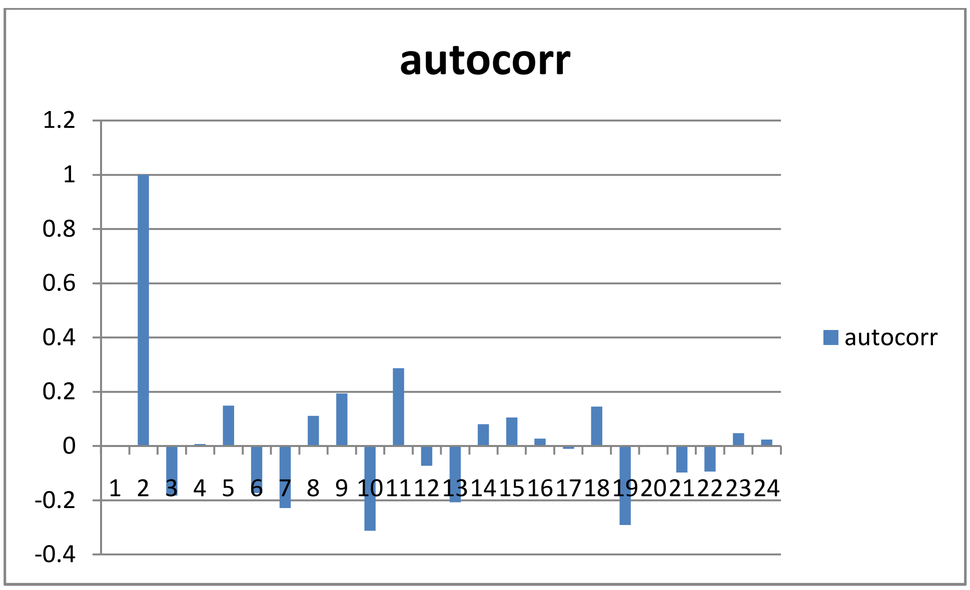




(continuation of ***Part 2: Commercial Sector***)

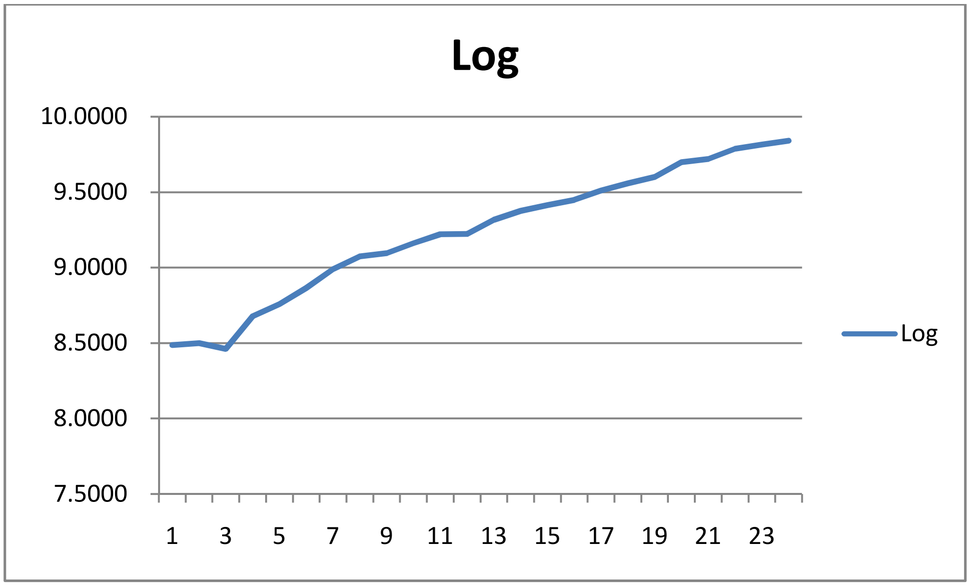
<u>First Differences</u>

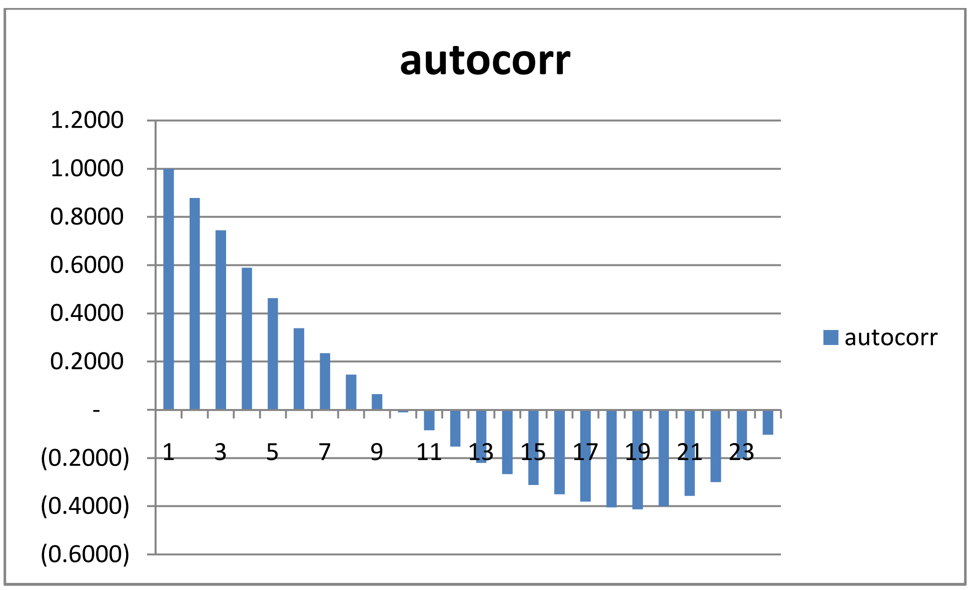




(continuation of ***Part 2: Commercial Sector***)

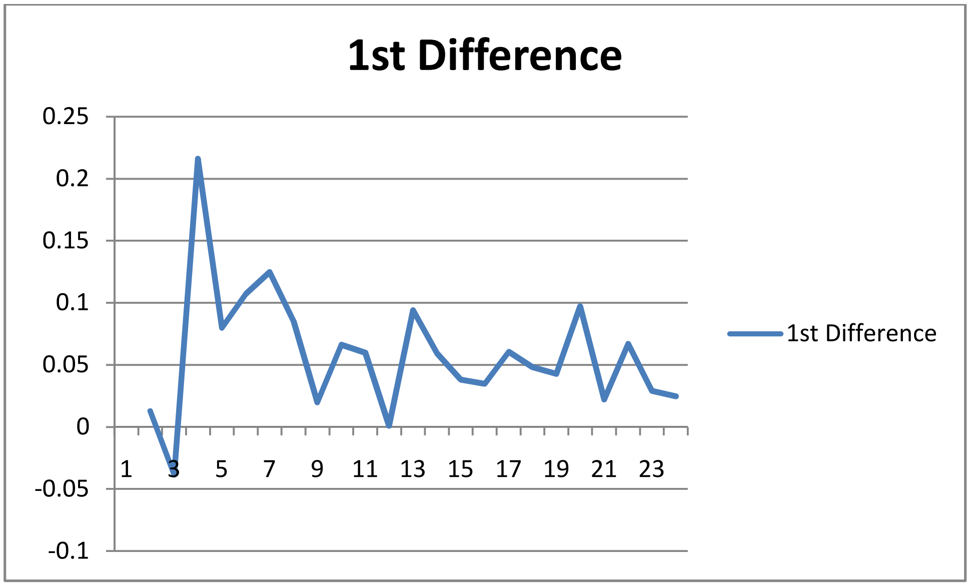
<u>Log-Transformation</u>

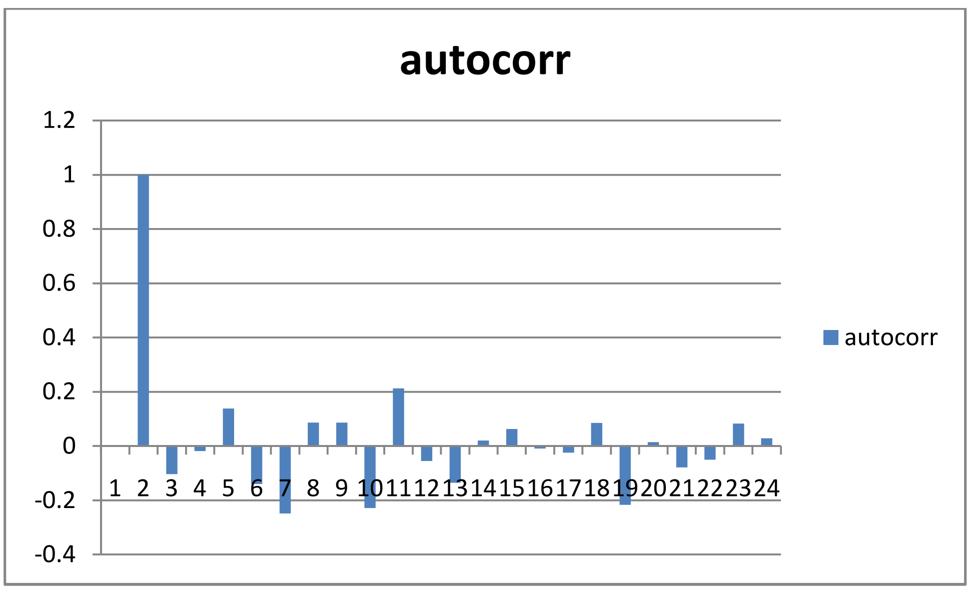




(continuation of ***Part 2: Commercial Sector***)

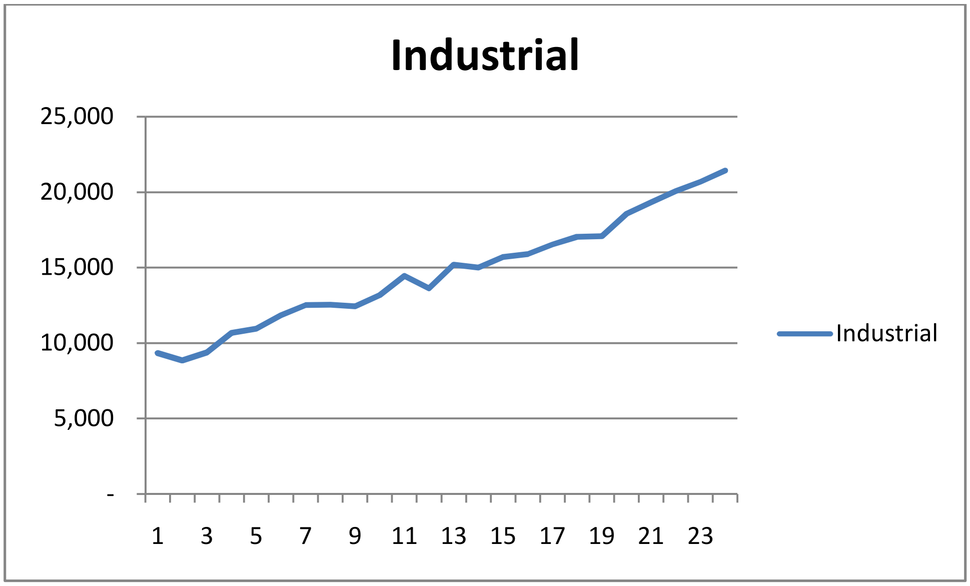
First Differences of the Log-Transformation

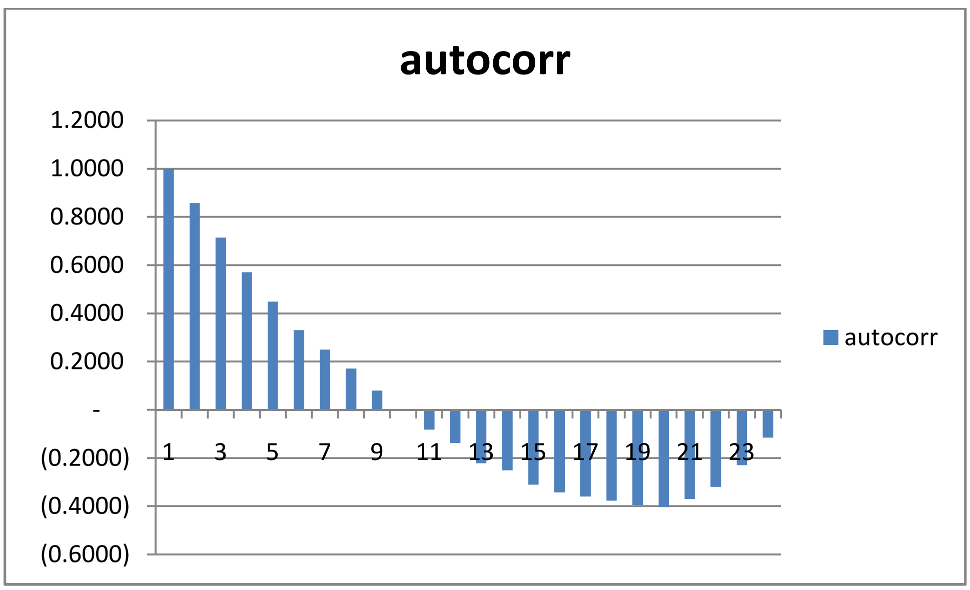




Part 3: Industrial Sector

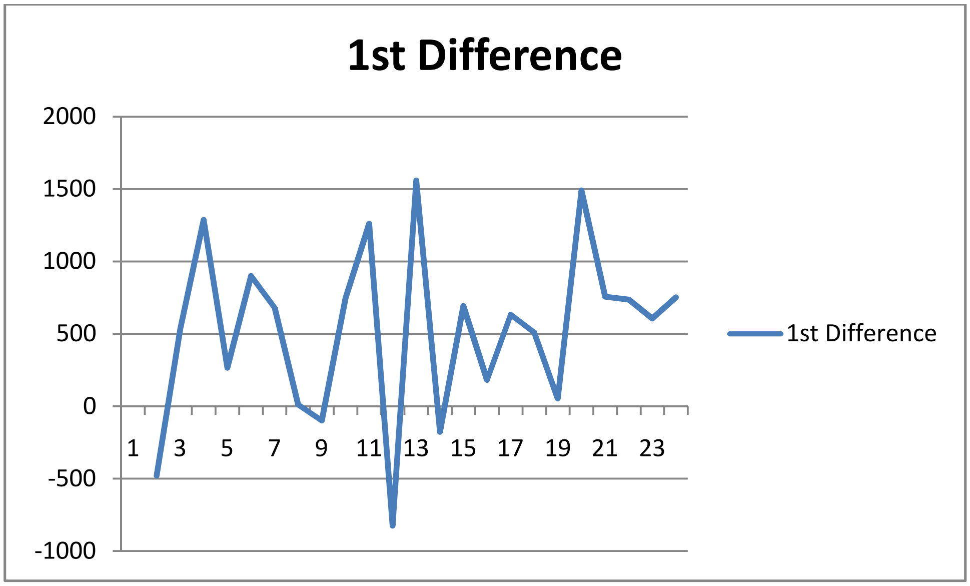
Original Data

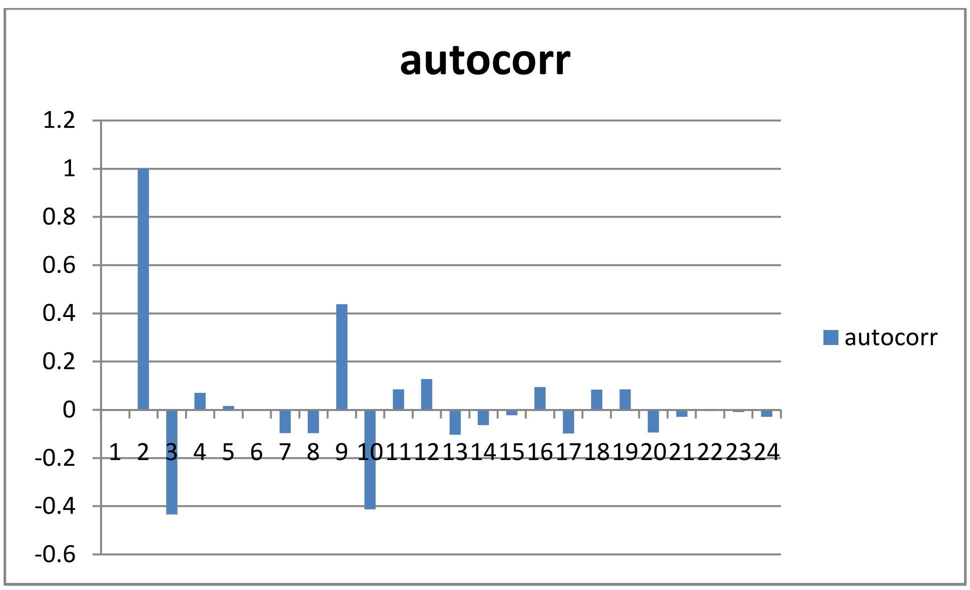




(continuation of ***Part 3: Industrial Sector***)

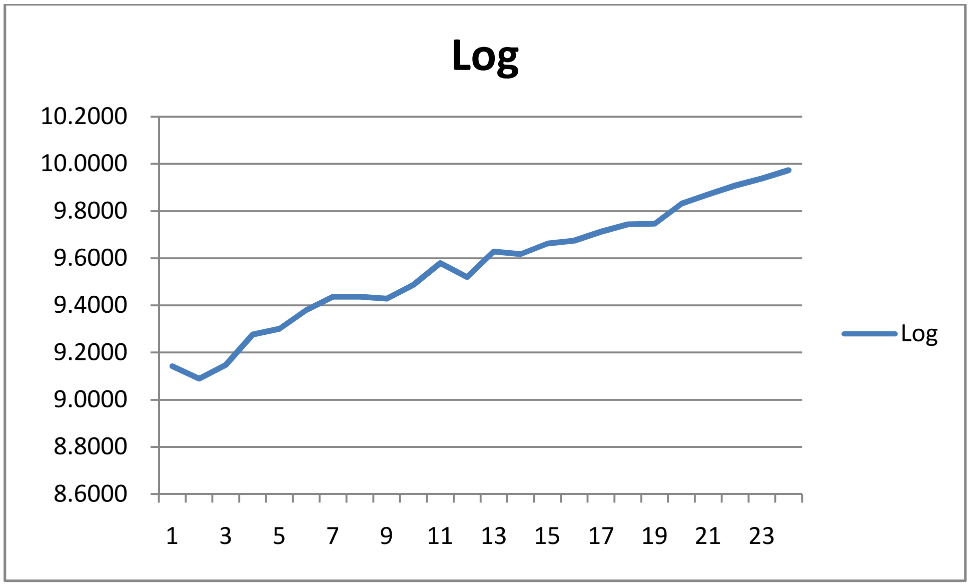
<u>First Differences</u>

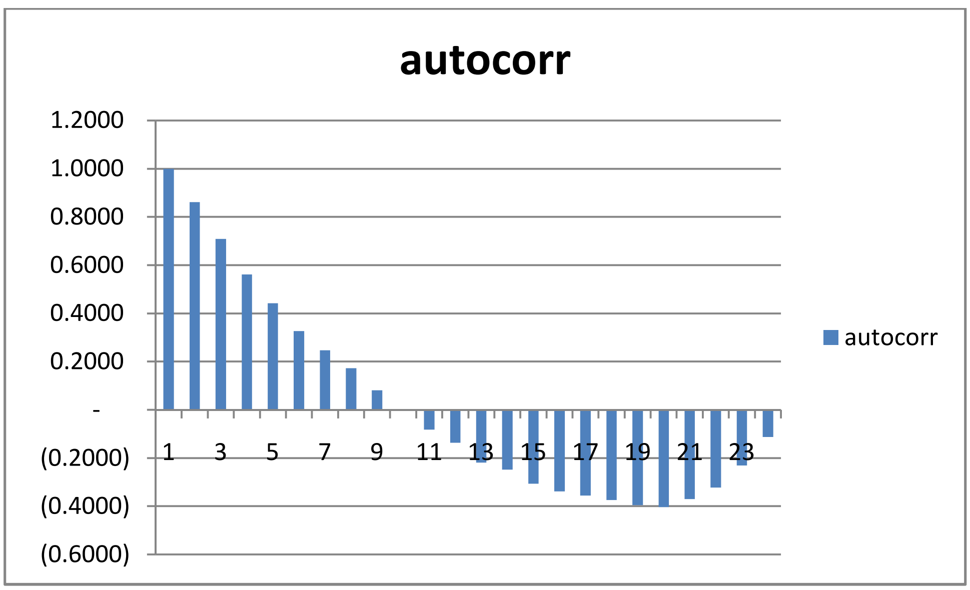




(continuation of ***Part 3: Industrial Sector***)

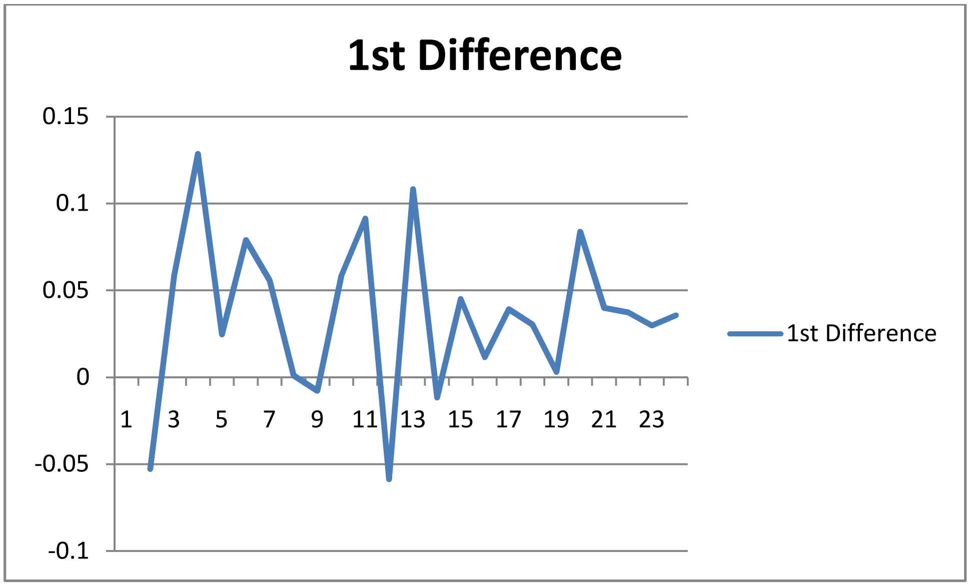
Log-Transformation

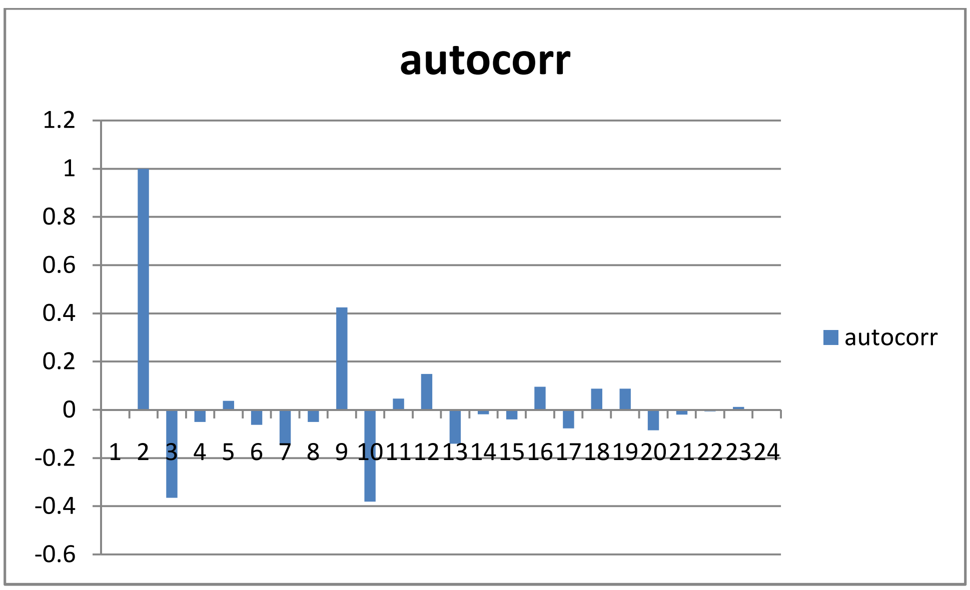




(continuation of ***Part 3: Industrial Sector***)

First Differences of the Log-Transformation





ANALYSIS

Methodology and Results

For all sectors, the following models were used:

1) ARIMA(0, 0, 0) using the original sets of data
   - formula: $Y_t = \mu + e_t$
   - mean of each set of data: μ
   - residual: e_t
2) ARIMA(1, 0, 0) using the original sets of data
   - formula: $Y_t = \mu + \phi Y_{t-1} + e_t$
   - mean of each set of data: μ
   - via least square estimation: ϕ
   - residual: e_t
3) ARIMA(0, 1, 0) using the sets of first differences of original data
   - formula: $Z_t = Y_t - Y_{t-1} = e_t$
   - residual: e_t
4) ARIMA(0, 0, 0) using the sets of log-transformed data
   - formula: $ln(Y_t) = \mu + e_t$
   - mean of each set of data: μ
   - residual: e_t
5) ARIMA(1, 0, 0) using the sets of log-transformed data
   - formula: $ln(Y_t) = \mu + \phi ln(Y_{t-1}) + e_t$
   - mean of each set of data: μ
   - via least square estimation: ϕ
   - residual: e_t
6) ARIMA(0, 1, 0) using the sets of first differences of log-transformed original data
   - formula: $Z_t = ln(Y_t) - ln(Y_{t-1}) = e_t$
   - residual: e_t

For models 2 and 5, the Regression add-in in Excel was used to determine the ϕ coefficient.

After determining the residuals, these were ranked and their respective percentiles (p) were determined. These percentiles were then used in the formula $\Phi^{-1}(p)$ to produce the z-scores. The z-scores (as the theoretical quantiles) with its corresponding sample data (as the sample quantiles) were used to create the normal q-q plots.

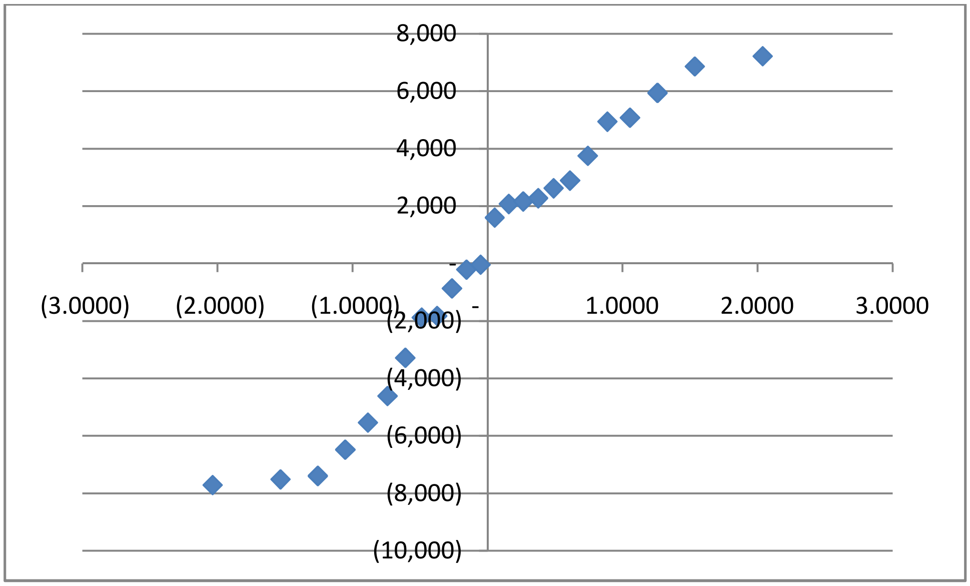
The normality of the residuals was also tested. This manually was done via Shapiro-Wilk normality test. After producing the Shapiro-Wilk statistics, the corresponding p-values were determined via interpolation. If the resulting p-value is less than 0.05, the residuals were considered not symmetric and the model was rejected, and vice versa.

The results of this procedure can be seen in the succeeding pages.

Part 1: Residential Sector

ARIMA(0, 0, 0) using the original sets of data

Normal q-q plot:



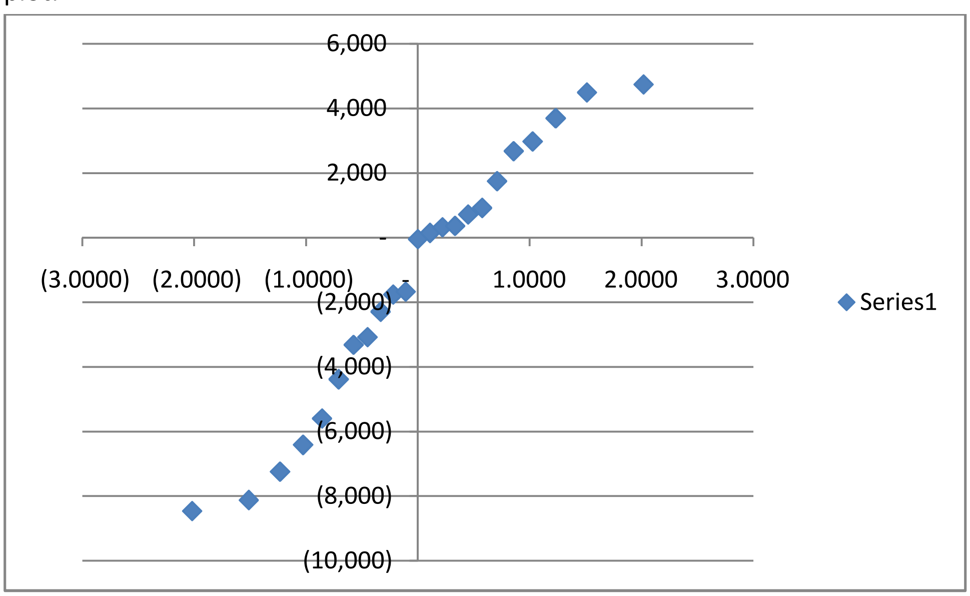
Shapiro-Wilk statistic: 0.9413

P-value: 0.2370 > 0.05

Decision: residuals are symmetric; accept model

ARIMA(1, 0, 0) using the original sets of data

Normal q-q plot:



ϕ coefficient: 0.1200

Shapiro-Wilk statistic: 0.9533

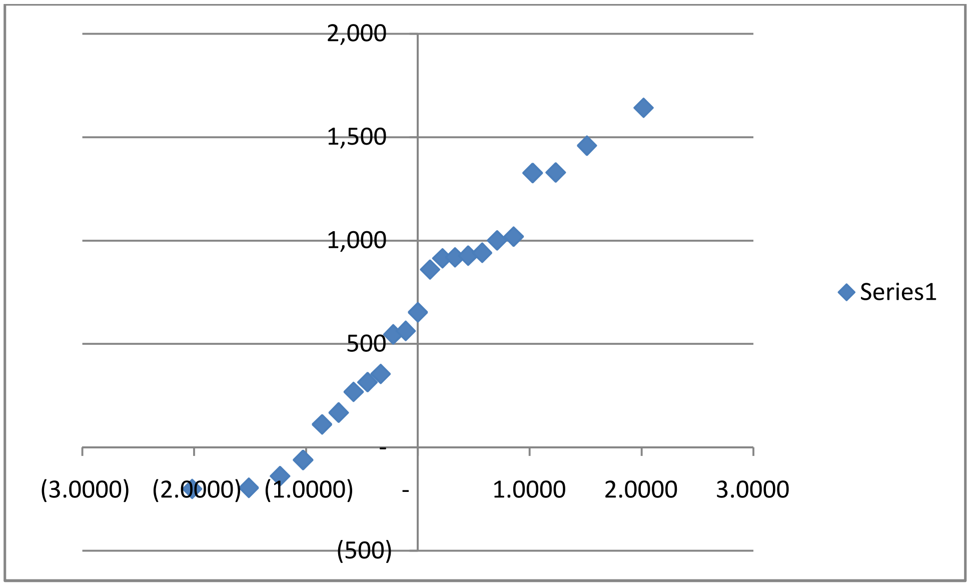
P-value: 0.3974 > 0.05

Decision: residuals are symmetric; accept model

(continuation of ***Part 1: Residential Sector***)

ARIMA(0, 1, 0) using the sets of first differences of original data

Normal q-q plot:



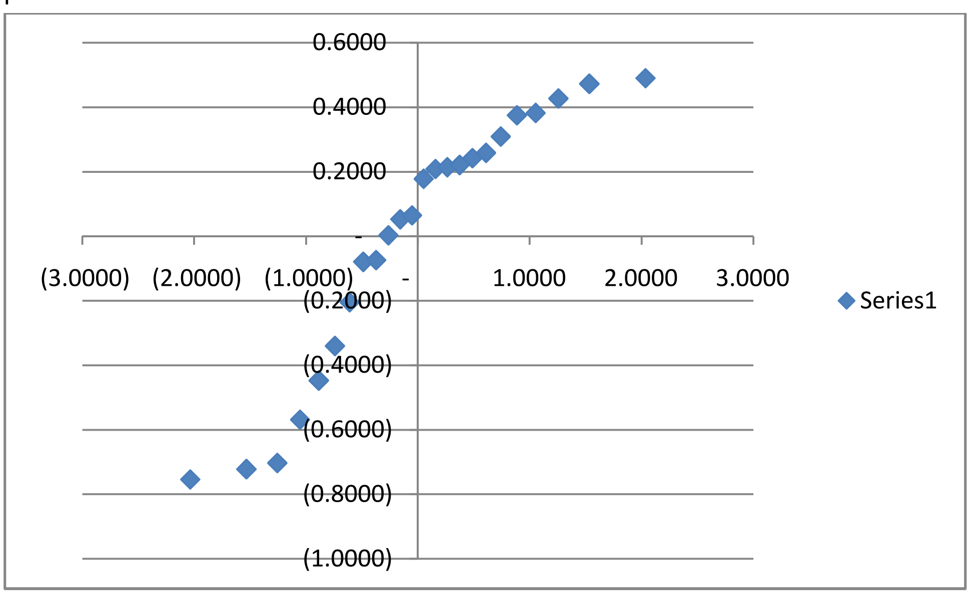
Shapiro-Wilk statistic: 0.9555

P-value: 0.4236 > 0.05

Decision: residuals are symmetric; accept model

ARIMA(0, 0, 0) using the sets of log-transformed data

Normal q-q plot:



Shapiro-Wilk statistic: 0.8964

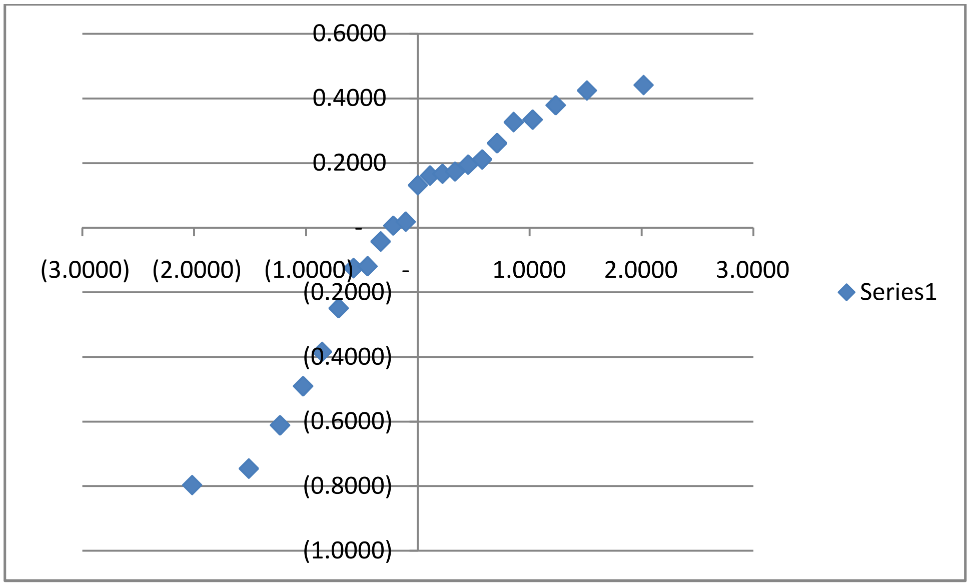
P-value: 0.0188 < 0.05

Decision: residuals are not symmetric; reject model

(continuation of ***Part 1: Residential Sector***)

ARIMA(1, 0, 0) using the sets of log-transformed data

Normal q-q plot:



ϕ coefficient: 0.0049

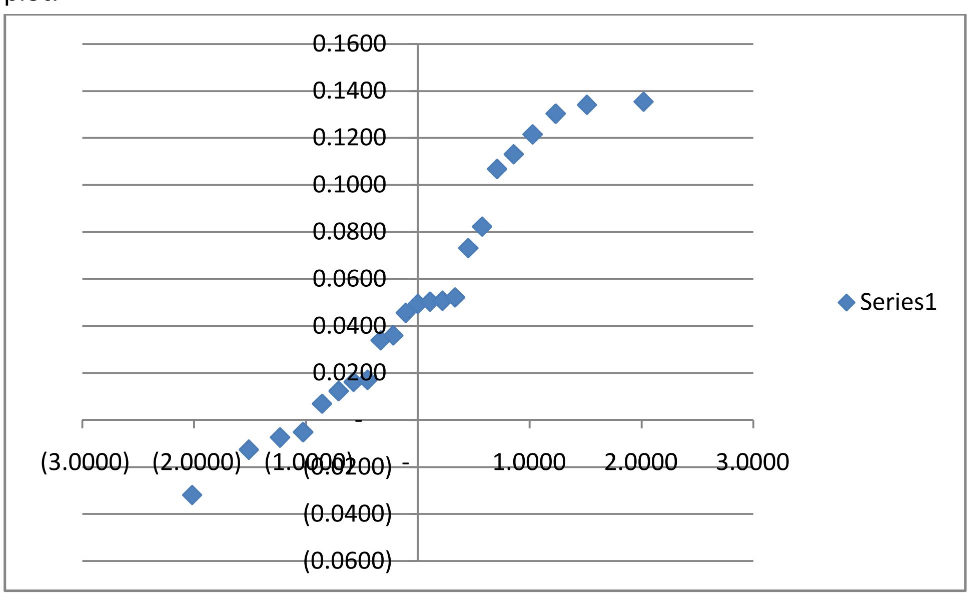
Shapiro-Wilk statistic: 0.9053

P-value: 0.0362 < 0.05

Decision: residuals are not symmetric; reject model

ARIMA(0, 1, 0) using the sets of first differences of log-transformed original data

Normal q-q plot:



Shapiro-Wilk statistic: 0.9397

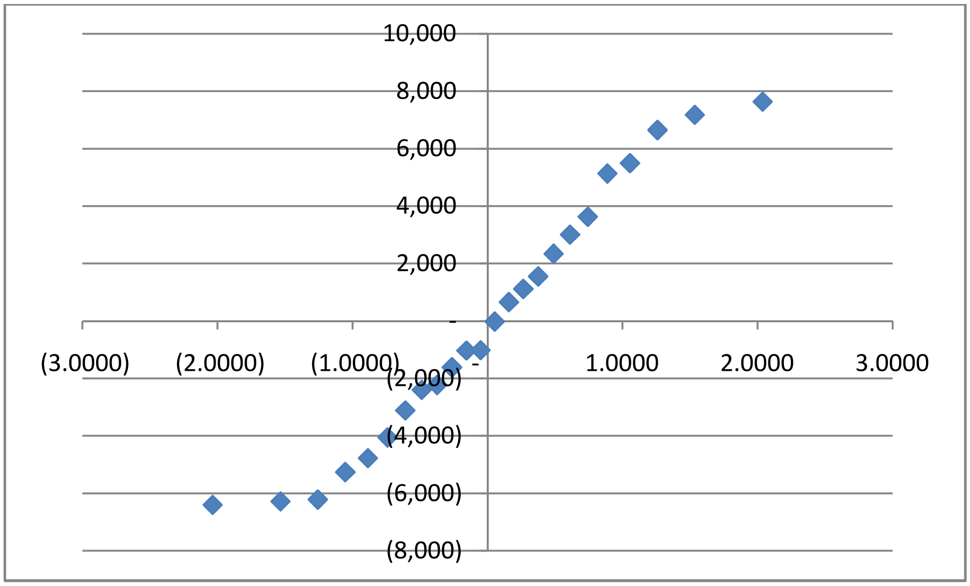
P-value: 0.2380 > 0.05

Decision: residuals are symmetric; accept model

Part 2: Commercial Sector

ARIMA(0, 0, 0) using the original sets of data

Normal q-q plot:



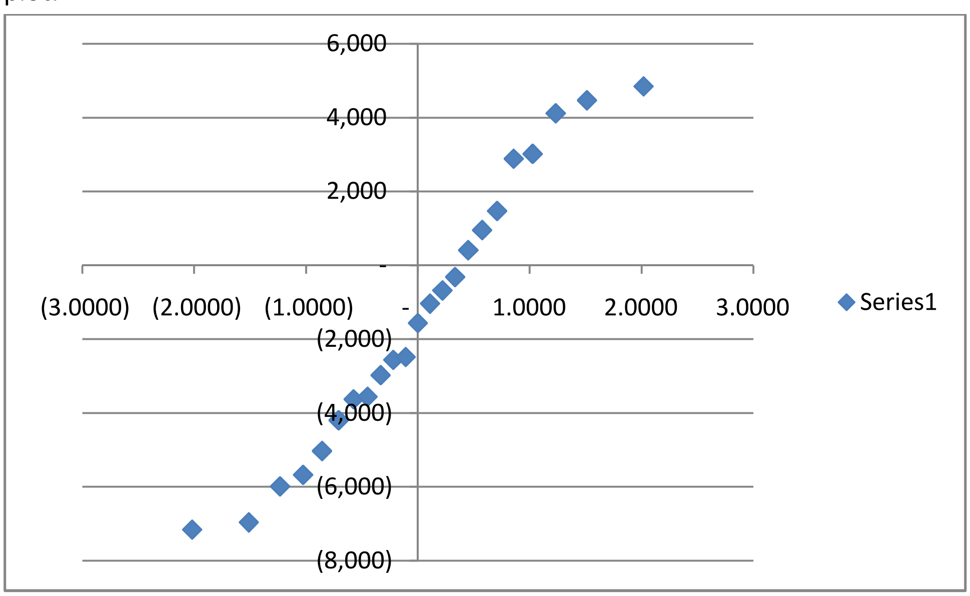
Shapiro-Wilk statistic: 0.9501

P-value: 0.3437 > 0.05

Decision: residuals are symmetric; accept model

ARIMA(1, 0, 0) using the original sets of data

Normal q-q plot:



ϕ coefficient: 0.1525

Shapiro-Wilk statistic: 0.9593

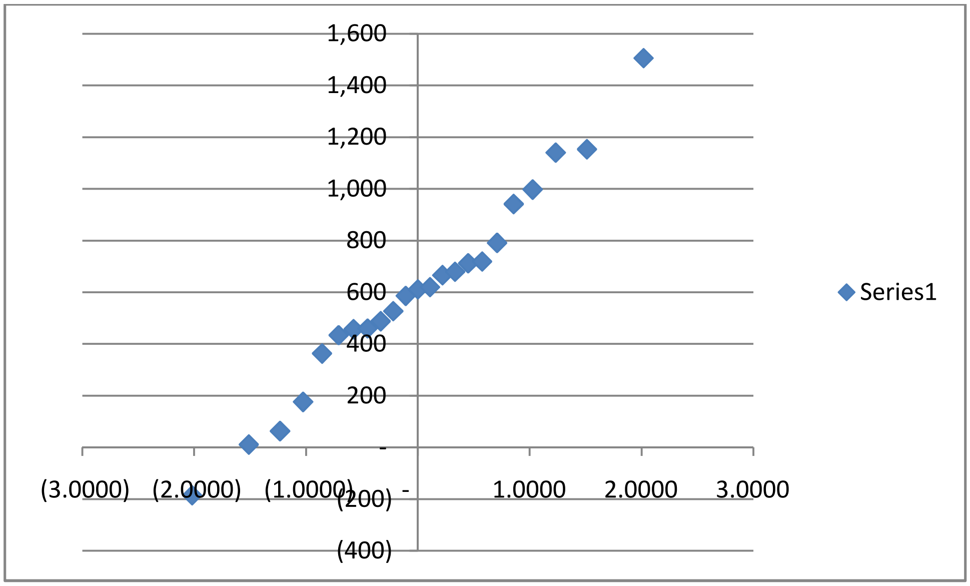
P-value: 0.4680 > 0.05

Decision: residuals are symmetric; accept model

(continuation of ***Part 2: Commercial Sector***)

ARIMA(0, 1, 0) using the sets of first differences of original data

Normal q-q plot:



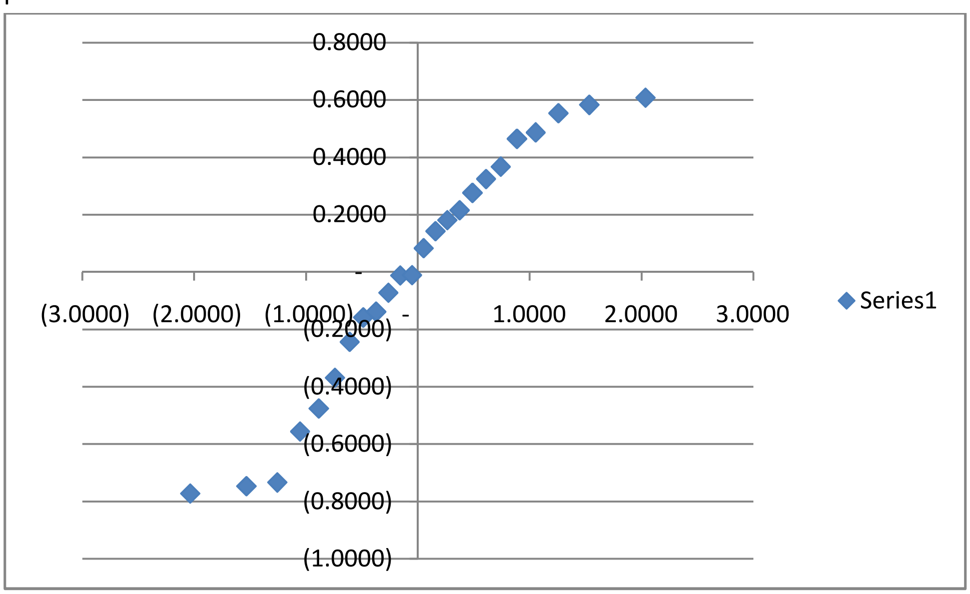
Shapiro-Wilk statistic: 0.9795

P-value: 0.8686 > 0.05

Decision: residuals are symmetric; accept model

ARIMA(0, 0, 0) using the sets of log-transformed data

Normal q-q plot:



Shapiro-Wilk statistic: 0.9409

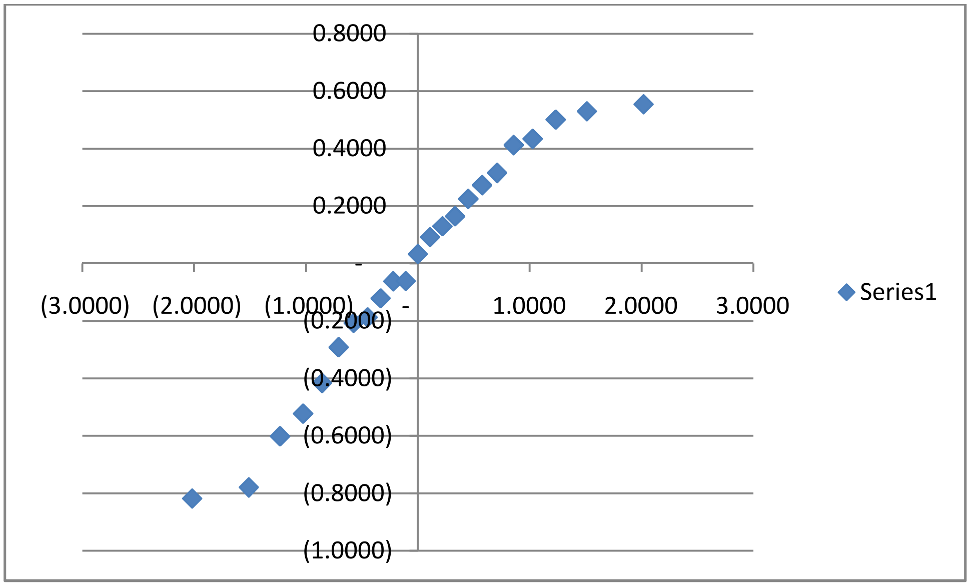
P-value: 0.2316 > 0.05

Decision: residuals are symmetric; accept model

(continuation of ***Part 2: Commercial Sector***)

ARIMA(1, 0, 0) using the sets of log-transformed data

Normal q-q plot:



ϕ coefficient: 0.0055

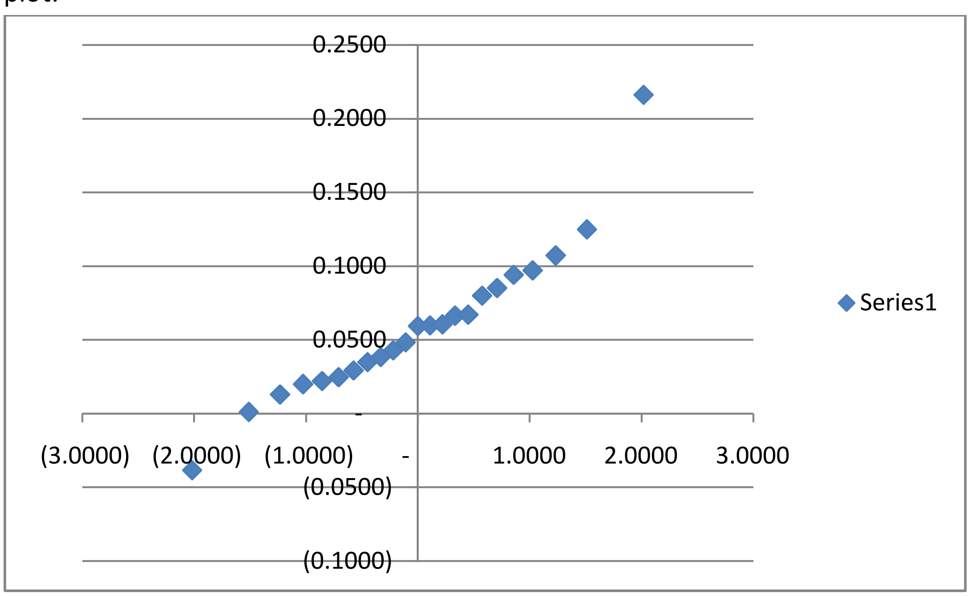
Shapiro-Wilk statistic: 0.9518

P-value: 0.3796 > 0.05

Decision: residuals are symmetric; accept model

ARIMA(0, 1, 0) using the sets of first differences of log-transformed original data

Normal q-q plot:



Shapiro-Wilk statistic: 0.9263

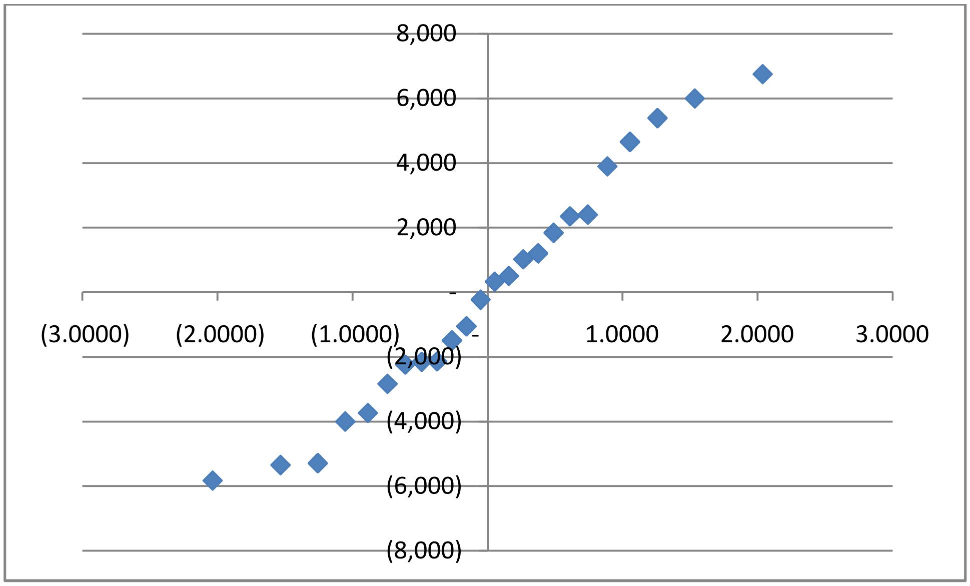
P-value: 0.0939 > 0.05

Decision: residuals are symmetric; accept model

Part 3: Industrial Sector

<u>ARIMA(0, 0, 0) using the original sets of data</u>

Normal q-q plot:



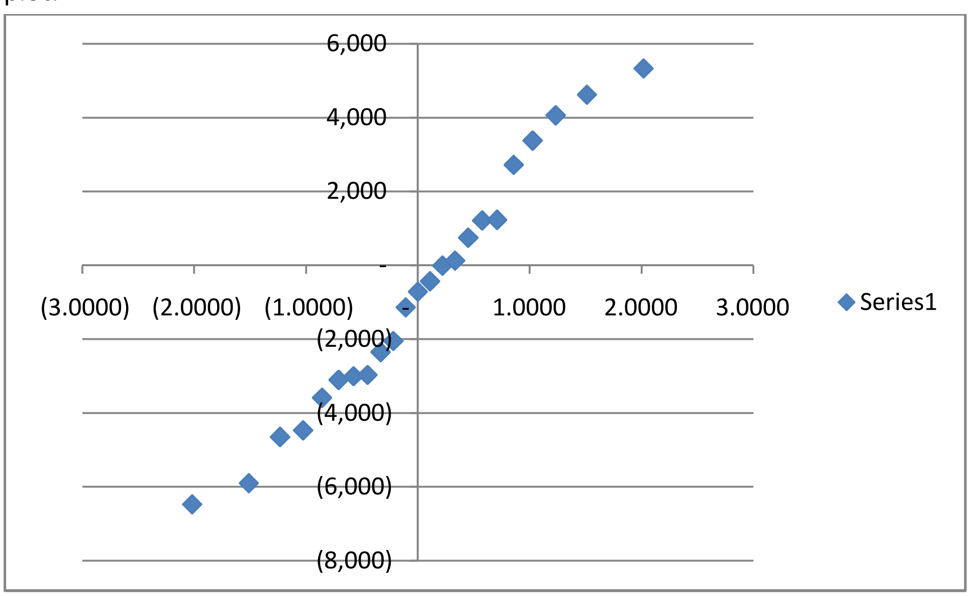
Shapiro-Wilk statistic: 0.9661

P-value: 0.5692 > 0.05

Decision: residuals are symmetric; accept model

<u>ARIMA(1, 0, 0) using the original sets of data</u>

Normal q-q plot:



ϕ coefficient: 0.0687

Shapiro-Wilk statistic: 0.9734

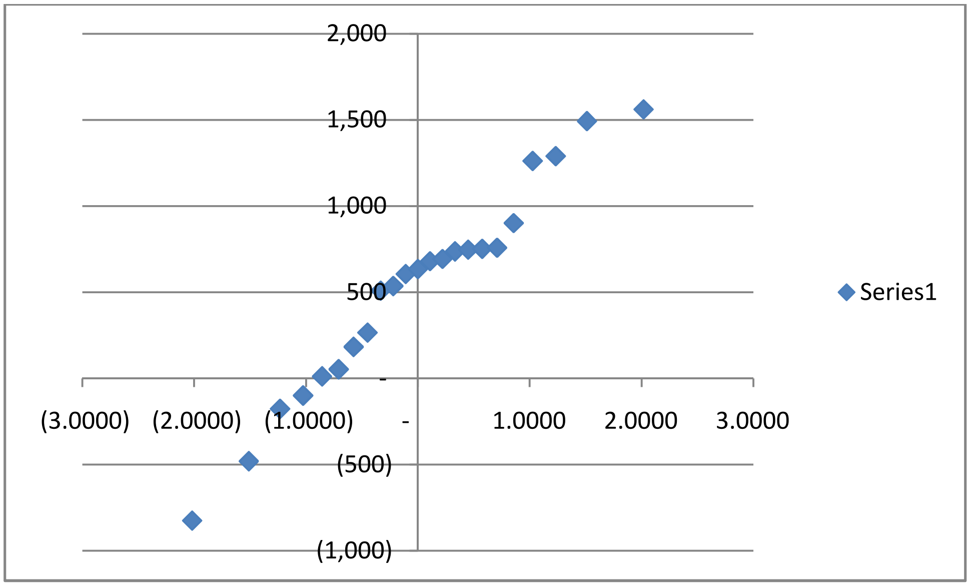
P-value: 0.7395 > 0.05

Decision: residuals are symmetric; accept model

(continuation of ***Part 3: Industrial Sector***)

ARIMA(0, 1, 0) using the sets of first differences of original data

Normal q-q plot:



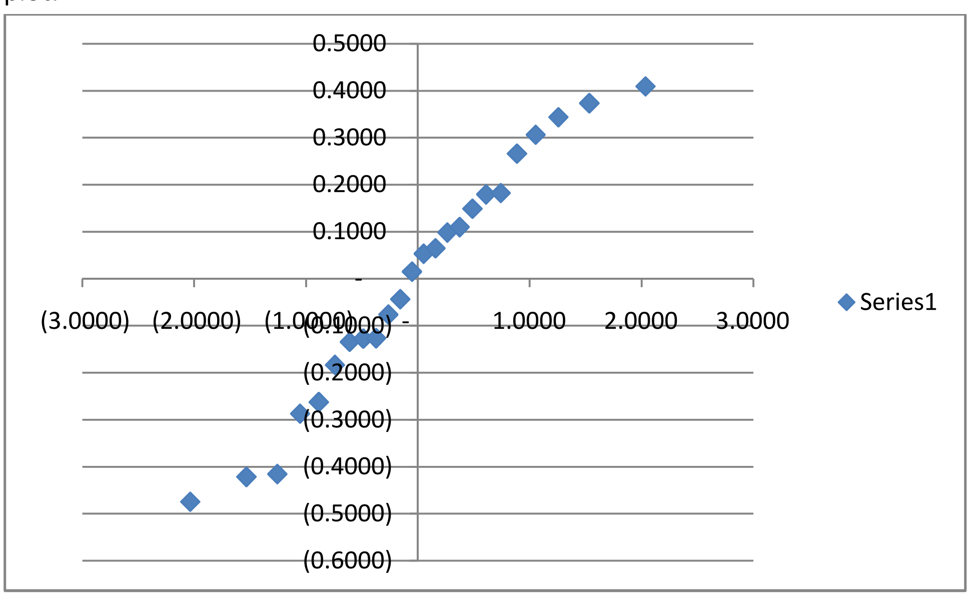
Shapiro-Wilk statistic: 0.9663

P-value: 0.5898 > 0.05

Decision: residuals are symmetric; accept model

ARIMA(0, 0, 0) using the sets of log-transformed data

Normal q-q plot:



Shapiro-Wilk statistic: 0.9654

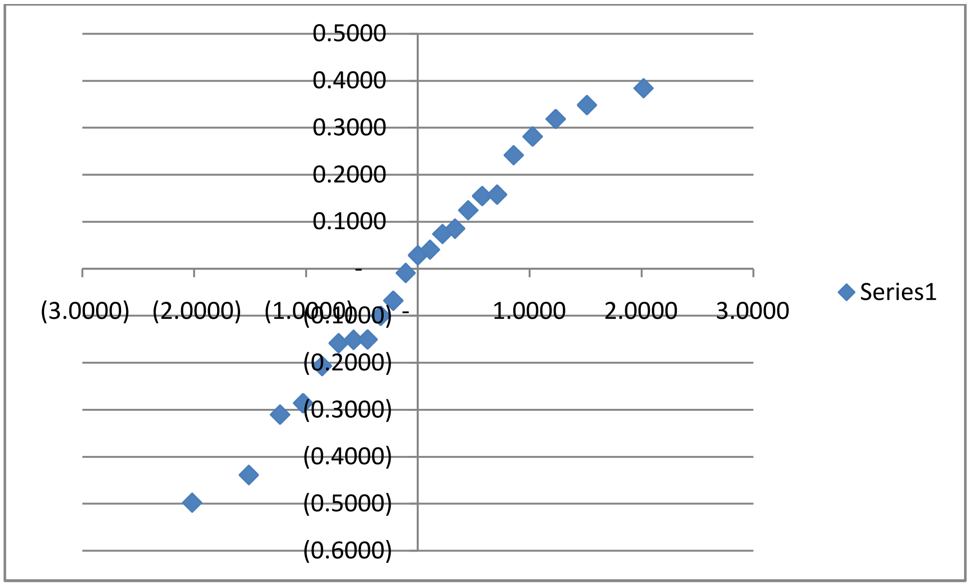
P-value: 0.5544 > 0.05

Decision: residuals are symmetric; accept model

(continuation of ***Part 3: Industrial Sector***)

ARIMA(1, 0, 0) using the sets of log-transformed data

Normal q-q plot:



ϕ coefficient: 0.0025

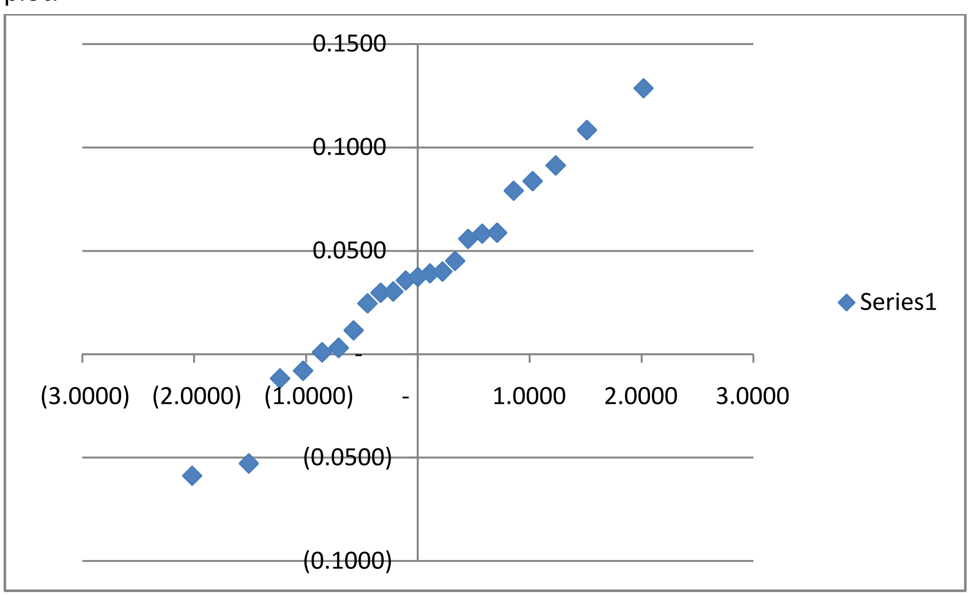
Shapiro-Wilk statistic: 0.9735

P-value: 0.7425 > 0.05

Decision: residuals are symmetric; accept model

ARIMA(0, 1, 0) using the sets of first differences of log-transformed original data

Normal q-q plot:



Shapiro-Wilk statistic: 0.9790

P-value: 0.8578 > 0.05

Decision: residuals are symmetric; accept model

SUMMARY AND CONCLUSION

Summarized below are the results for all models for each sector.

Residential Sector

Model	Formula	SW stat	p-value	residuals; accept/reject model
1	$Y_t = 13{,}760 + e_t$	0.9413	0.2370	symmetric; accept
2	$Y_t = 13{,}760 + 0.1200Y_{t-1} + e_t$	0.9533	0.3974	symmetric; accept
3	$Z_t = Y_t - Y_{t-1} = e_t$	0.9555	0.4236	symmetric; accept
4	$ln(Y_t) = 9.4617 + e_t$	0.8964	0.0188	not symmetric; reject
5	$ln(Y_t) = 9.4617 + 0.0049ln(Y_{t-1}) + e_t$	0.9053	0.0362	not symmetric; reject
6	$Z_t = ln(Y_t) - ln(Y_{t-1}) = e_t$	0.9397	0.2380	symmetric; accept

Best model: Model 3 (highest p-value)

Commercial Sector

Model	Formula	SW stat	p-value	residuals; accept/reject model
1	$Y_t = 11{,}126 + e_t$	0.9501	0.3437	symmetric; accept
2	$Y_t = 11{,}126 + 0.1525Y_{t-1} + e_t$	0.9593	0.4680	symmetric; accept
3	$Z_t = Y_t - Y_{t-1} = e_t$	0.9795	0.8686	symmetric; accept
4	$ln(Y_t) = 9.2321 + e_t$	0.9409	0.2316	symmetric; accept
5	$ln(Y_t) = 9.2321 + 0.0055ln(Y_{t-1}) + e_t$	0.9518	0.3796	symmetric; accept
6	$Z_t = ln(Y_t) - ln(Y_{t-1}) = e_t$	0.9263	0.0939	symmetric; accept

Best model: Model 3 (highest p-value)

Industrial Sector

Model	Formula	SW stat	p-value	residuals; accept/reject model
1	$Y_t = 14{,}683 + e_t$	0.9661	0.5692	symmetric; accept
2	$Y_t = 14{,}683 + 0.0687Y_{t-1} + e_t$	0.9734	0.7395	symmetric; accept
3	$Z_t = Y_t - Y_{t-1} = e_t$	0.9663	0.5898	symmetric; accept
4	$ln(Y_t) = 9.5634 + e_t$	0.9654	0.5544	symmetric; accept
5	$ln(Y_t) = 9.5634 + 0.0025ln(Y_{t-1}) + e_t$	0.9735	0.7425	symmetric; accept
6	$Z_t = ln(Y_t) - ln(Y_{t-1}) = e_t$	0.9790	0.8578	symmetric; accept

Best model: Model 6 (highest p-value)

By Model

Model	most symmetric
1	Industrial
2	Industrial
3	Commercial
4	Industrial
5	Commercial
6	Industrial

Since each of set of data used in this project only has 24 observations, it would be interesting to see if the sets of data have more observations. For example, total monthly power consumption instead of annual power consumption.

EXHIBIT A: TOTAL ANNUAL POWER CONSUMPTION (IN Gwh) OF PHILIPPINES' RESIDENTIAL, COMMERCIAL, AND INDUSTRIAL SECTORS DURING FOR YEARS 1991-2004

Year	Residential	Commercial	Industrial
1991	6,249	4,847	9,339
1992	6,053	4,910	8,859
1993	6,368	4,725	9,395
1994	7,282	5,865	10,684
1995	8,223	6,353	10,950
1996	9,150	7,072	11,851
1997	10,477	8,013	12,531
1998	11,936	8,725	12,543
1999	11,875	8,901	12,444
2000	12,894	9,512	13,191
2001	13,547	10,098	14,452
2002	13,715	10,109	13,628
2003	15,357	11,106	15,188
2004	15,920	11,785	15,012
2005	16,031	12,245	15,705
2006	15,830	12,679	15,888
2007	16,376	13,470	16,522
2008	16,644	14,136	17,031
2009	17,504	14,756	17,084
2010	18,833	16,261	18,576
2011	18,694	16,624	19,334
2012	19,695	17,777	20,071
2013	20,614	18,304	20,677
2014	20,969	18,761	21,429

EXHIBIT B: RESIDENTIAL SECTOR

Year	Residential	1st Diff	Log Trans	1st Diff of Log Trans
1991	6,249		8.7402	
1992	6,053	(196)	8.7083	(0.0319)
1993	6,368	315	8.7590	0.0507
1994	7,282	914	8.8932	0.1341
1995	8,223	941	9.0147	0.1215
1996	9,150	927	9.1215	0.1068
1997	10,477	1,327	9.2569	0.1354
1998	11,936	1,459	9.3873	0.1304
1999	11,875	(61)	9.3822	(0.0051)
2000	12,894	1,019	9.4645	0.0823
2001	13,547	653	9.5139	0.0494
2002	13,715	168	9.5262	0.0123
2003	15,357	1,642	9.6393	0.1131
2004	15,920	563	9.6753	0.0360
2005	16,031	111	9.6823	0.0069
2006	15,830	(201)	9.6697	(0.0126)
2007	16,376	546	9.7036	0.0339
2008	16,644	268	9.7198	0.0162
2009	17,504	860	9.7702	0.0504
2010	18,833	1,329	9.8434	0.0732
2011	18,694	(139)	9.8360	(0.0074)
2012	19,695	1,001	9.8881	0.0522
2013	20,614	919	9.9337	0.0456
2014	20,969	355	9.9508	0.0171

EXHIBIT C: COMMERCIAL SECTOR

Year	Commercial	1st Diff	Log Trans	1st Diff of Log Trans
1991	4,847		8.4861	
1992	4,910	63	8.4990	0.0129
1993	4,725	(185)	8.4606	(0.0384)
1994	5,865	1,140	8.6768	0.2161
1995	6,353	488	8.7567	0.0799
1996	7,072	719	8.8639	0.1072
1997	8,013	941	8.9888	0.1249
1998	8,725	712	9.0739	0.0851
1999	8,901	176	9.0939	0.0200
2000	9,512	611	9.1603	0.0664
2001	10,098	586	9.2201	0.0598
2002	10,109	11	9.2212	0.0011
2003	11,106	997	9.3152	0.0941
2004	11,785	679	9.3746	0.0593
2005	12,245	460	9.4129	0.0383
2006	12,679	434	9.4477	0.0348
2007	13,470	791	9.5082	0.0605
2008	14,136	666	9.5565	0.0483
2009	14,756	620	9.5994	0.0429
2010	16,261	1,505	9.6965	0.0971
2011	16,624	363	9.7186	0.0221
2012	17,777	1,153	9.7857	0.0671
2013	18,304	527	9.8149	0.0292
2014	18,761	457	9.8395	0.0247

EXHIBIT D: INDUSTRIAL SECTOR

Year	Industrial	1st Diff	Log Trans	1st Diff of Log Trans
1991	9,339		9.1420	
1992	8,859	(480)	9.0892	(0.0528)
1993	9,395	536	9.1479	0.0587
1994	10,684	1,289	9.2765	0.1286
1995	10,950	266	9.3011	0.0246
1996	11,851	901	9.3802	0.0791
1997	12,531	680	9.4360	0.0558
1998	12,543	12	9.4369	0.0010
1999	12,444	(99)	9.4290	(0.0079)
2000	13,191	747	9.4873	0.0583
2001	14,452	1,261	9.5786	0.0913
2002	13,628	(824)	9.5199	(0.0587)
2003	15,188	1,560	9.6283	0.1084
2004	15,012	(176)	9.6166	(0.0117)
2005	15,705	693	9.6617	0.0451
2006	15,888	183	9.6733	0.0116
2007	16,522	634	9.7124	0.0391
2008	17,031	509	9.7428	0.0303
2009	17,084	53	9.7459	0.0031
2010	18,576	1,492	9.8296	0.0837
2011	19,334	758	9.8696	0.0400
2012	20,071	737	9.9070	0.0374
2013	20,677	606	9.9368	0.0297
2014	21,429	752	9.9725	0.0357