

Modeling & Forecasting

Consumer Confidence Index

Yip Nam Kim

Jonathan Kizer

Arshiya Sabzevari

The University of Texas at Austin

May 2018

Table of Contents

Table of Contents	1
Executive Summary	2
Analysis	2
Findings	2
Consumer Confidence Index & Elections	3
Consumer Confidence Index: 1960 - Present	4
Modeling the Consumer Confidence Index	5
ARIMA Model	6
ARMA-GARCH Model	7
Forecasting CCI with ARMA-GARCH Model	9
Conclusion	10
Appendix A: ARIMA Model Plots	11
Appendix B: ARMA-GARCH Model Plots	13
Appendix C: Project Output	16

Executive Summary

Analysis

This report includes an analysis and forecast of the Consumer Confidence Index in the United States from January, 1960 to February, 2018, in order to determine whether or not the Index will remain above 100 for the November, 2018 midterm elections. When Consumer Confidence is above 100, political scientists have observed that the incumbent political party typically stays in power.

Findings

This report's findings, created through the use of an ARMA-GARCH model, indicate that the Consumer Confidence Index will almost certainly remain above 100 in November, 2019. Therefore, this indicator would suggest that the Republican Party is likely to continue to hold its electoral advantage during the midterm elections.

Consumer Confidence Index & Elections

The Consumer Confidence Index is described by the Organisation for Economic Co-operation and Development as a reflection of a household's plans for future major purchases, as well as their current and immediate future economic situation. Opinions are collected from survey respondents; these answers are then compared to a "normal" state, with the difference between positive and negative statements providing a qualitative index on economic conditions.¹

Beyond this, the Consumer Confidence Index in the United States has been positively correlated to the odds that an incumbent retains his or her seat in elected office.² Generally, the higher the elected office, the more this effect is seen. This has lead political scientists and journalists to develop a rule of thumb that a President up for reelection is very likely to win if the Consumer Confidence Index in the United States is above 100.

While the trend is not quite as pronounced for predicting United States Senate and House elections, the general rule is still useful for beginning to forecast future election results. While other macroeconomic and social trends have a large impact on the outcome of these more local elections, the general sentiment of voters towards the economy sets a decent baseline for electoral expectations.

Therefore, in addition to attempting to build a model that sufficiently explains the in-sample data and trends, this report also aims to project nine months into the future, to the

¹ OECD.org Consumer confidence index (CCI). (n.d.). Retrieved from <http://www.oecd.org/>

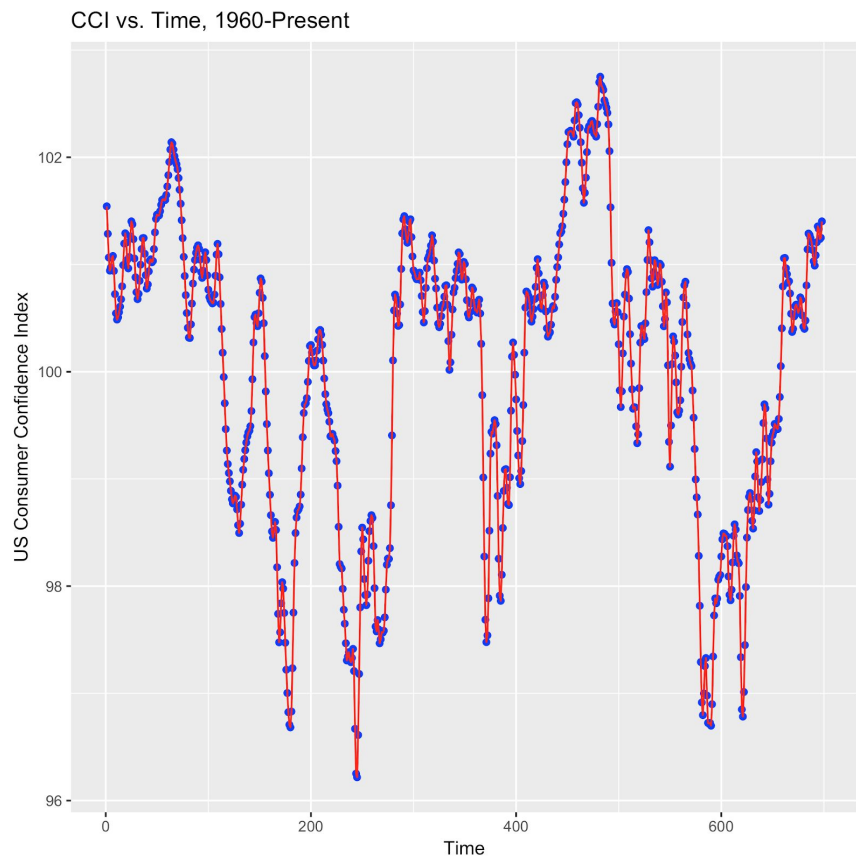
² Erikson, R. S., MacKuen, M., & Stimson, J. A. (2006). *The Macro Polity*. New York: Cambridge University Press.

upcoming November elections, in order to predict whether or not the Consumer Confidence Index will still be above this 100 threshold that seems to be influential for elections.

Modeling, and therefore predicting, Consumer Confidence is difficult. There are many factors that influence the index, including social, political, and financial trends. Additionally, it is difficult to gauge which of these types of factors typically matters most to consumers. In some cases, Presidential elections seem to swing the Consumer Confidence Index one way or another, even as other financial metrics remain relatively stable from one President to the next. At other times, it is clear that financial indicators like the stock market are heavily influential, such as during the economic crisis of 2008, when the Consumer Confidence Index fell to 96.9.

Consumer Confidence Index: 1960 - Present

The image to the left shows Consumer Confidence Index from January, 1960 to February, 2018. While the actual variation in US Consumer Confidence Index is small -- all values are between 96 and 103 -- there appears to be large discrepancies in



correlation structure and variation over time.

This points towards the dataset being non-stationary. There also appears to be a large amount of difference in variation; there are periods where the Consumer Confidence Index is very volatile, dropping or rising sharply month to month. At other times, the Index is relatively stable, staying within a small upper and lower range for possibly years. While the data appears to be heteroskedastic, there is also a minute amount of seasonality, which is taken into account with the `stl()` R function. The seasonality was so small as to be essentially meaningless, and so the dataset was analyzed without it.

Modeling the Consumer Confidence Index

Various methods were used to attempt to model the Consumer Confidence Index. The full printout of every model tested -- a total of five -- may be found in Appendix C: Project Output, while the following three models will be covered in-depth in this report:

1. ARIMA Model
2. ARMA-GARCH Model

The ARIMA model (created using the `auto.Arima()` function) offered a result that initially appeared useful, but upon further examination failed to have Normally-distributed errors. Because of this, the model was useless for forecasting purposes, and proved that a more general framework was needed to take into account all of the trends within the Consumer Confidence Index dataset.

From that conclusion, an ARMA-GARCH model was fitted to the data, and ultimately provided the best results as it took into account the heteroskedastic nature of the dataset.

ARIMA Model

```

ARIMA(3,1,2)

Coefficients:
      ar1      ar2      ar3      ma1      ma2
    0.1847  0.4349 -0.4429  1.3208  0.4222
s.e.  0.0853  0.0947  0.0621  0.0841  0.0473

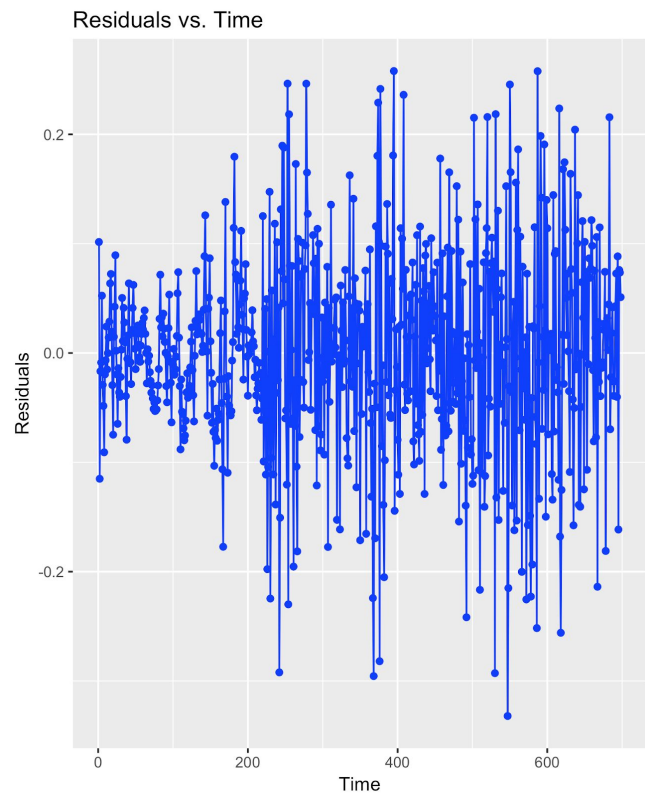
sigma^2 estimated as 0.008244:  log likelihood=685.09
AIC=-1358.18  AICc=-1358.06  BIC=-1330.9

```

The results of the `auto.arima()` function was an ARIMA(3, 1, 2) model. This model results in a standard deviation of: 0.091.

Graphing the results of this model and the actual Consumer Confidence Index appears promising (Appendix B, Figure I). The model seems to track well with the trends within the actual data.

Additionally, the ACF and PACF (Appendix B, Figures II & III) appear to show that the data is uncorrelated; the model therefore does a good job of extracting data and underlying trends within the Index dataset.



Issues begin to arise when observing the residuals from this model (see above graph). There appears to still be trends within the residuals. Upon further work and after applying the

Anderson-Darling test and creating a QQPlot (Appendix B, Figure IV) of the errors, it becomes clear that the errors are not Normally distributed. Therefore, despite a good AIC and low sigma value, one of the underlying assumptions of the model is violated, and it is therefore useless for creating meaningful forecasts.

This supports the original observation that the data may not be heteroskedastic, and that therefore an ARMA model with GARCH errors is almost certainly necessary.

ARMA-GARCH Model

With each model displaying the same violation of Normality to a significant degree, it became clear that the Consumer Confidence Index required the use of a model that could track changing errors over time. Additionally, the correct model still had to incorporate trend and level throughout the large Index dataset.

The final model tested was therefore an ARMA-GARCH model. This model made use of the input parameters previously found with the `auto.arima()` command for the $p=3$ and $d=2$ components. The GARCH component was set to (1, 1), after running various iterations and choosing the result with the lowest AIC value, as well as inspecting the ACF and PACF plots. This model was created using the `rugarch` R package, which allows for the simultaneous creation of a model that incorporates an ARMA model for the mean, and a GARCH model for the variance.³

³ Ghalanos, A. (2018, January 28). `Rugarch`. Retrieved from <http://www.unstarched.net/r/rugarch/>


```

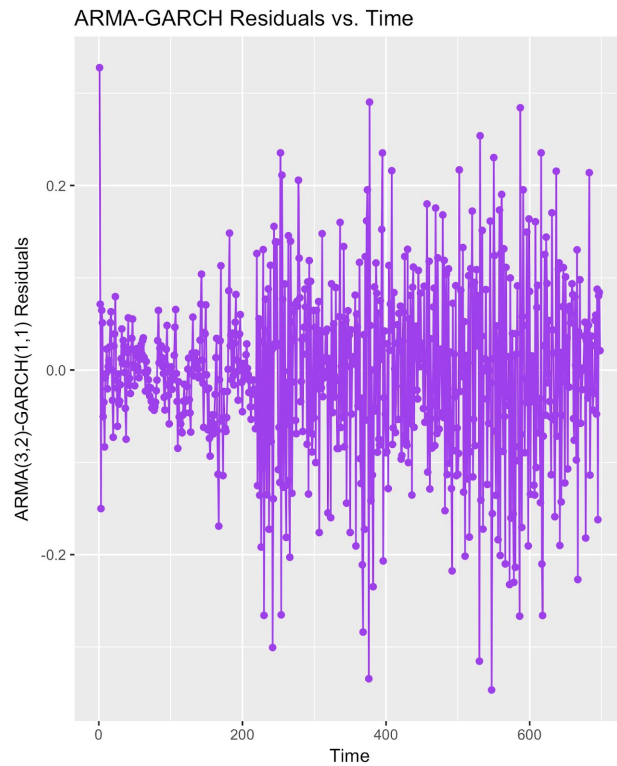
*-----*
*           GARCH Model Fit           *
*-----*

Conditional Variance Dynamics
-----

GARCH Model : sGARCH(1,1)
Mean Model   : ARFIMA(3,0,2)
Distribution : norm
Optimal Parameters
-----

```

	Estimate	Std. Error	t value	Pr(> t)
mu	101.215054	0.059340	1705.6772	0.0000
ar1	1.962917	0.009355	209.8334	0.0000
ar2	-1.398994	0.015342	-91.1857	0.0000
ar3	0.426647	0.020589	20.7221	0.0000
ma1	0.683141	0.043430	15.7296	0.0000
ma2	0.285677	0.040023	7.1377	0.0000
omega	0.000037	0.000024	1.5726	0.1158
alpha1	0.113514	0.020929	5.4238	0.0000
beta1	0.885486	0.019457	45.5098	0.0000



This model results in an AIC, after being multiplied by the number of data points in the Consumer Confidence Index dataset, of -1517.173. Additionally, the ACF and PACF show that the η and η^2 are uncorrelated (Appendix C, Figures I & II). The residuals, plotted to the left, display fewer signs of clear trends and correlation. Finally, the Ljung-Box test of both the η and the η^2 terms indicates constant variance and that the eta terms are

uncorrelated.

The model may be seen compared against the original CCI data in Appendix C, Figure IV.

While the ARMA-GARCH method does seem to better explain the Index dataset than the other attempted models, there are still issues. This likely is because Consumer Confidence Index itself depends on a number of factors.

While some of the trends could be accounted for with explanatory variables, it would take a subject matter expert to identify those variables and know that they are a cause of whatever trend is seen in the Index dataset. A future report, built upon this ARMA-GARCH model, may try to account for this.

Forecasting CCI with ARMA-GARCH Model

To answer the actual question -- whether or not the Consumer Confidence Index will remain above 100 in nine months from the last data point included in this dataset (i.e., November, 2019's midterm elections), the ARMA-GARCH model was used.

```
*-----*
*          GARCH Model Forecast          *
*-----*
```

```
Model: sGARCH
Horizon: 9
```

```
Forecast:
Series  Sigma
T+1    101.6  0.08264
T+2    101.7  0.08283
T+3    101.7  0.08301
T+4    101.7  0.08320
T+5    101.7  0.08338
T+6    101.6  0.08356
T+7    101.6  0.08374
T+8    101.6  0.08393
T+9    101.6  0.08411
```

Therefore, if Erikson's *The Macro Polity* and a general rule of thumb for political scientists holds true, it is unlikely that Republicans will experience a major loss in the number of seats in the House and the Senate, since the Consumer Confidence Index will remain above 100 during the time period between now and the election. It is possible, of course, that external variables could change this forecast, or else that this election breaks the trend that Erikson described and codified.

Conclusion

This report finds an ARMA-GARCH model to be best suited to forecast Consumer Confidence Index. This model incorporates an ARMA model to predict the mean, and a GARCH model to account for unstable variance throughout time.

While this model was able to account for both the changes in mean and the changing errors over time, further analysis could be applied to the Index dataset in the form of external variables. However, due to the wide array of possible variables that may be influential on consumer answers that ultimately create the Index -- and the issue of consumers weighting different factors higher or lower over time depending on still other variables -- an expert in the area of Consumer Confidence Index collection and creation would be required.

The ARMA-GARCH analysis and forecast finds it exceedingly unlikely that the Consumer Confidence Index will fall below 100, a threshold used by political scientists to judge whether or not a change in the dominant political party is likely, between now and the elections of November, 2019.

Appendix A: ARIMA Model Plots

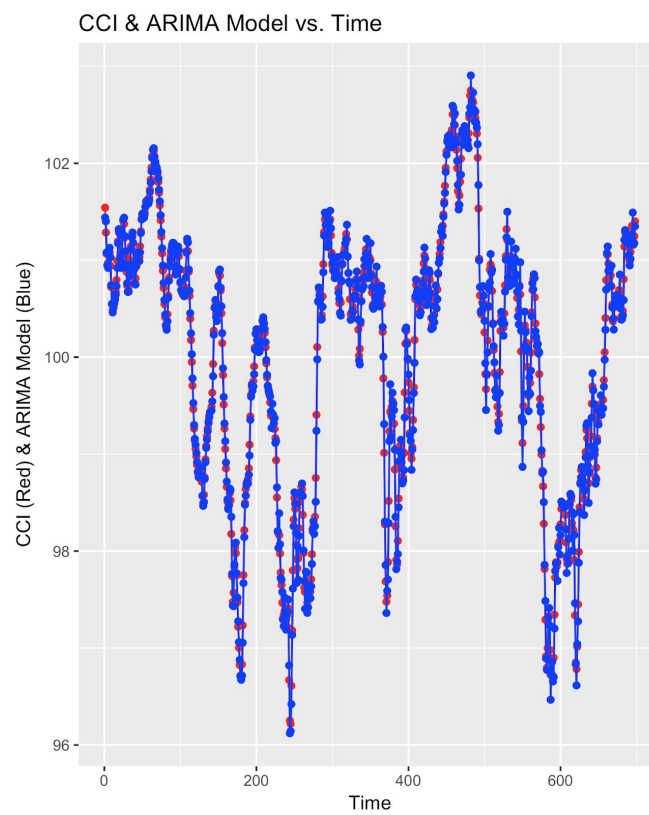


Figure I

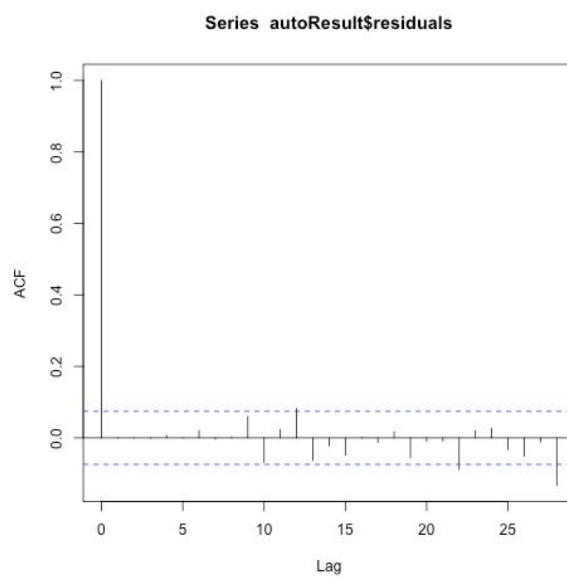


Figure B.II: ARIMA Residuals ACF

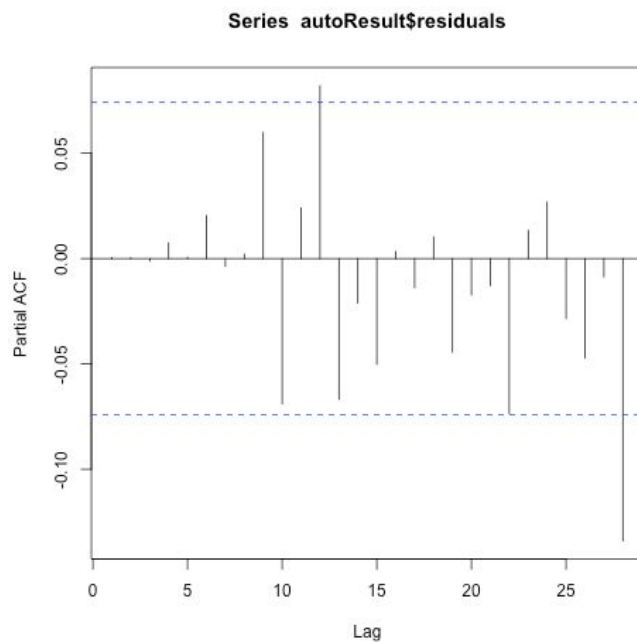


Figure III: ARIMA Residuals PACF

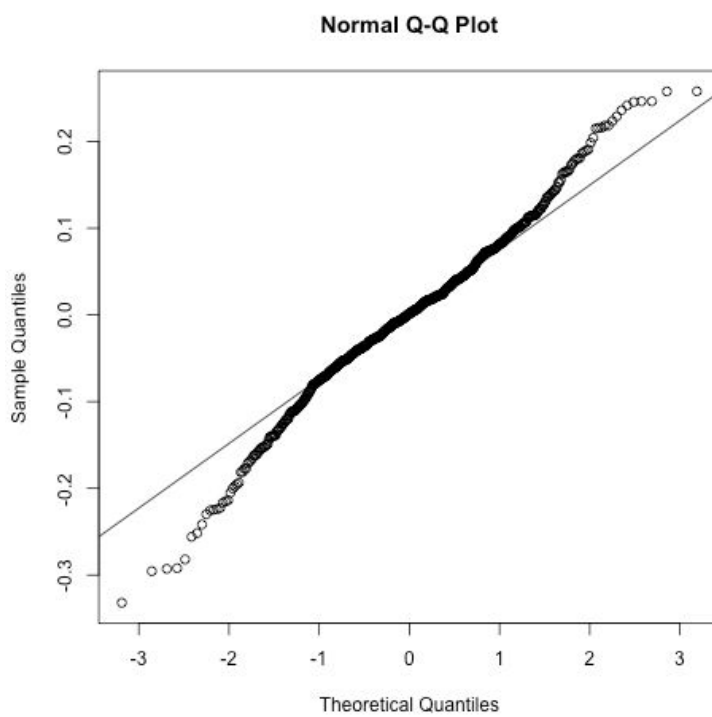


Figure IV: ARIMA Residuals QQPlot

Appendix B: ARMA-GARCH Model Plots

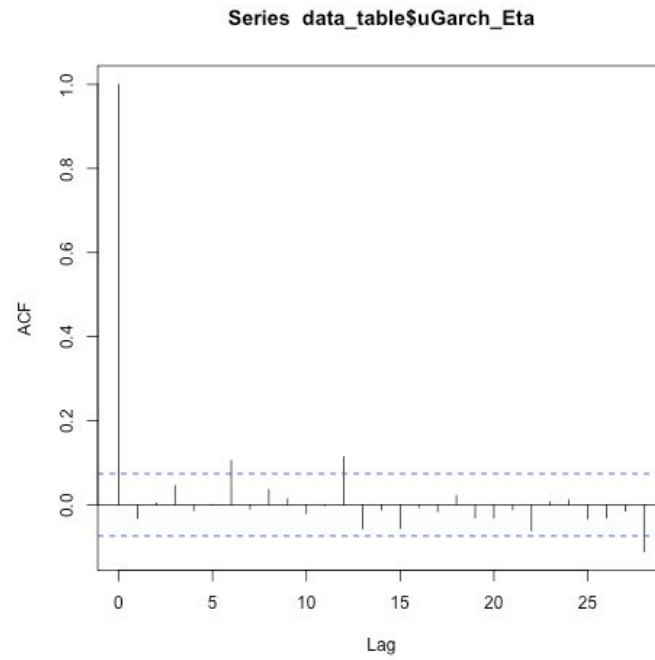


Figure 1: ARMA-GARCH η ACF

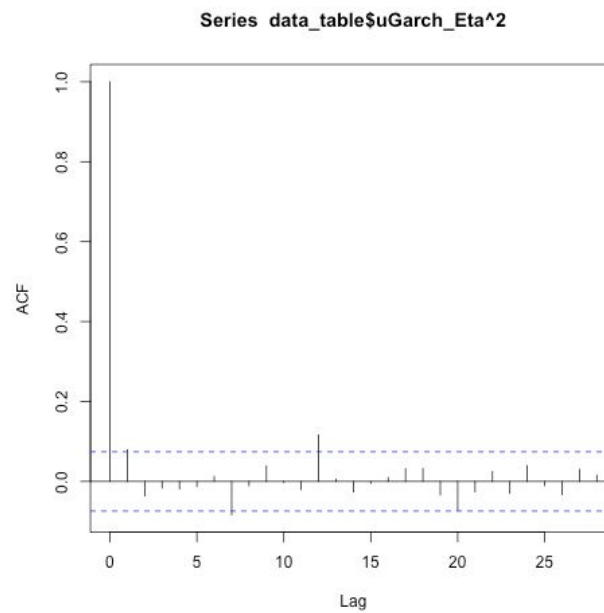


Figure 2: ARMA-GARCH η^2 ACF

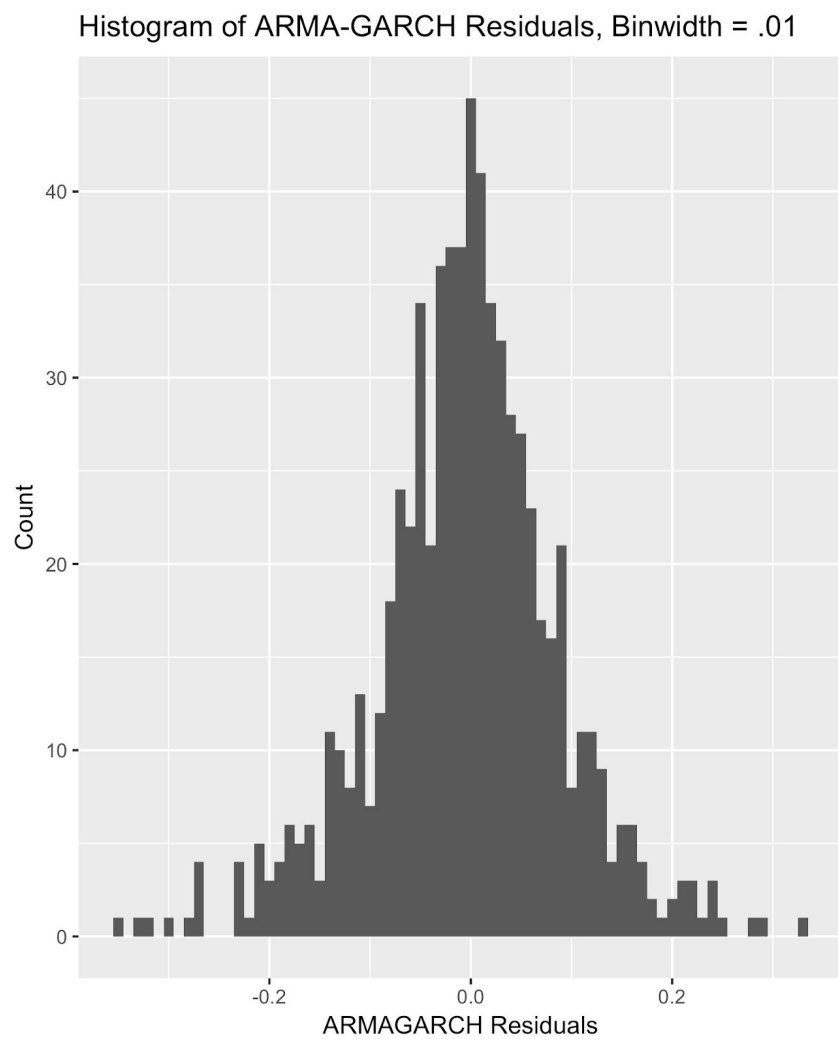


Figure III: Histogram of Residuals

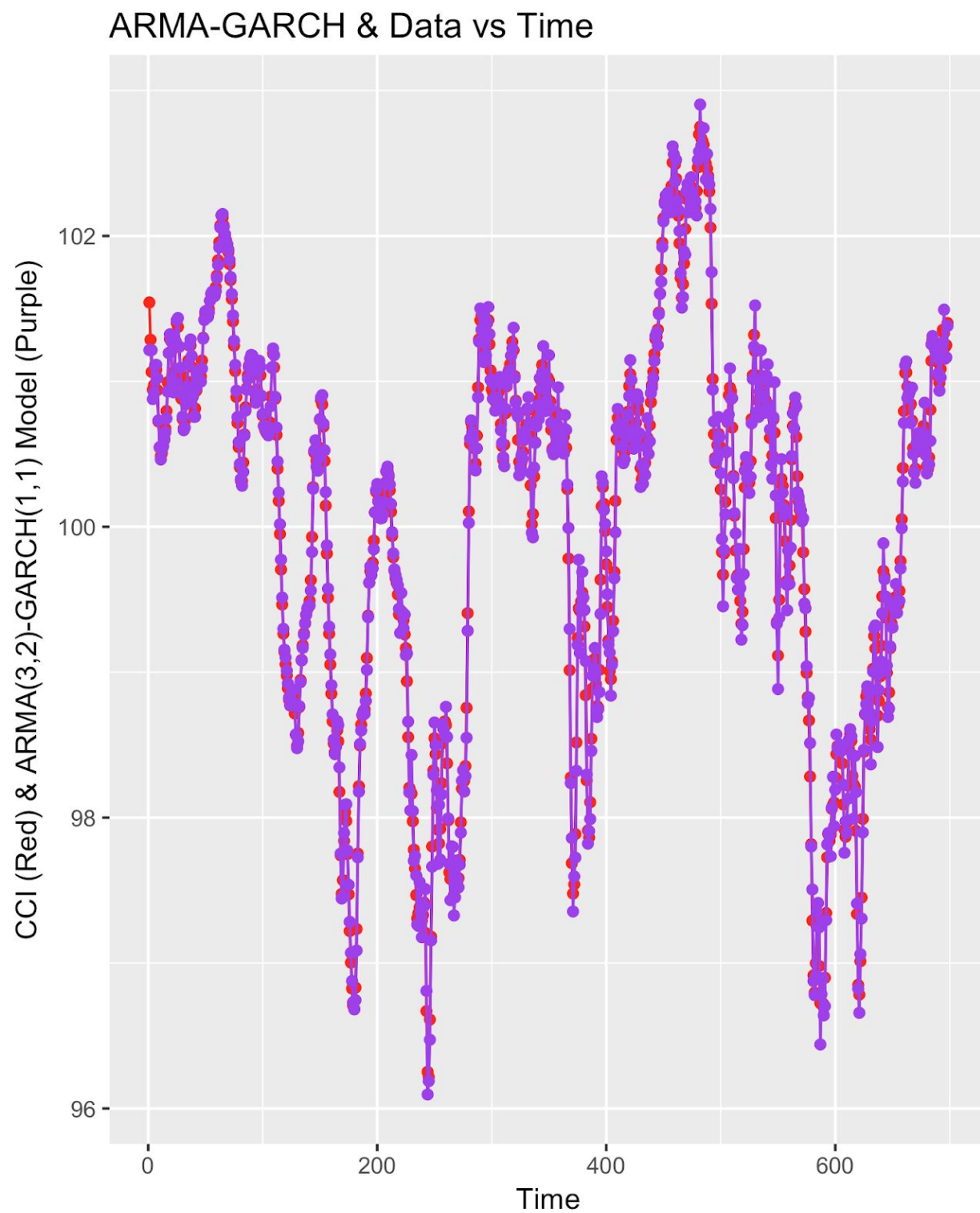


Figure IV: ARMA-GARCH Model & CCI Data vs. Time

Appendix C: Project Output

-----BEGIN ANALYSIS-----

Data Table First and Last Six Columns:

	Time	YearMonth	CCI
1	1	1960-01	101.5427
2	2	1960-02	101.2864
3	3	1960-03	101.0649
4	4	1960-04	100.9428
5	5	1960-05	100.9736
6	6	1960-06	101.0609
	Time	YearMonth	CCI
693	693	2017-09	101.2071
694	694	2017-10	101.3554
695	695	2017-11	101.3308
696	696	2017-12	101.2372
697	697	2018-01	101.2492
698	698	2018-02	101.4019

Augmented Dickey-Fuller Test

```
data:  ts(data_table[, 3])
Dickey-Fuller = -3.1288, Lag order = 8, p-value = 0.1005
alternative hypothesis: stationary
```

-----ARIMA Model Analysis-----

ARIMA(0,1,0)	: -251.4833
ARIMA(0,1,0) with drift	: -249.4724
ARIMA(0,1,1)	: -910.1783
ARIMA(0,1,1) with drift	: -908.1616
ARIMA(0,1,2)	: -1266.423
ARIMA(0,1,2) with drift	: -1264.401
ARIMA(0,1,3)	: -1323.583
ARIMA(0,1,3) with drift	: -1321.554
ARIMA(0,1,4)	: -1344.371
ARIMA(0,1,4) with drift	: -1342.336
ARIMA(0,1,5)	: -1344.57
ARIMA(0,1,5) with drift	: -1342.529
ARIMA(1,1,0)	: -877.4608
ARIMA(1,1,0) with drift	: -875.4441
ARIMA(1,1,1)	: -1225.638
ARIMA(1,1,1) with drift	: -1223.615

```

ARIMA(1,1,2) : -1329.168
ARIMA(1,1,2) with drift : -1327.14
ARIMA(1,1,3) : -1333.358
ARIMA(1,1,3) with drift : -1331.324
ARIMA(1,1,4) : -1343.199
ARIMA(1,1,4) with drift : -1341.158
ARIMA(2,1,0) : -1299.607
ARIMA(2,1,0) with drift : -1297.584
ARIMA(2,1,1) : -1353.973
ARIMA(2,1,1) with drift : -1351.944
ARIMA(2,1,2) : -1354.6
ARIMA(2,1,2) with drift : -1352.566
ARIMA(2,1,3) : -1353.795
ARIMA(2,1,3) with drift : -1351.754
ARIMA(3,1,0) : -1351.422
ARIMA(3,1,0) with drift : -1349.393
ARIMA(3,1,1) : -1353.867
ARIMA(3,1,1) with drift : -1351.832
ARIMA(3,1,2) : -1358.058
ARIMA(3,1,2) with drift : -1356.017
ARIMA(4,1,0) : -1354.146
ARIMA(4,1,0) with drift : -1352.111
ARIMA(4,1,1) : -1352.197
ARIMA(4,1,1) with drift : -1350.156
ARIMA(5,1,0) : -1352.283
ARIMA(5,1,0) with drift : -1350.242

```

Series: CCITimeSeries

ARIMA(3,1,2)

Coefficients:

	ar1	ar2	ar3	ma1	ma2
	0.1847	0.4349	-0.4429	1.3208	0.4222
s.e.	0.0853	0.0947	0.0621	0.0841	0.0473

sigma^2 estimated as 0.008244: log likelihood=685.09

AIC=-1358.18 AICc=-1358.06 BIC=-1330.9

Augmented Dickey-Fuller Test

data: autoResult\$residuals

Dickey-Fuller = -8.0567, Lag order = 8, p-value = 0.01

alternative hypothesis: stationary

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
699	101.5753	101.4589	101.6917	101.39734	101.7533
700	101.6900	101.3761	102.0039	101.20990	102.1701
701	101.7190	101.1917	102.2462	100.91256	102.5253

```

702      101.6974 100.9892 102.4055 100.61438 102.7804
703      101.6552 100.8099 102.5005 100.36240 102.9480
704      101.6252 100.6802 102.5701 100.17999 103.0704
705      101.6108 100.5891 102.6326 100.04823 103.1735
706      101.6138 100.5271 102.7006  99.95182 103.2758
707      101.6214 100.4736 102.7693  99.86600 103.3769

```

Anderson-Darling normality test

```

data:  autoResult$residuals
A = 2.8841, p-value = 2.957e-07

```

```

-----Holt's Model Analysis-----
Holt's method

```

```

Call:
holt(y = logCCIITS, h = 9)

```

```

Smoothing parameters:
  alpha = 0.9999
  beta  = 0.9999

```

```

Initial states:
  l = 4.6383
  b = -0.028

```

```

sigma:  0.0015

```

	AIC	AICc	BIC			
	-4459.656	-4459.569	-4436.915			
Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95	
699	4.620599	4.618625	4.622572	4.617580	4.623617	
700	4.622106	4.617692	4.626519	4.615356	4.628855	
701	4.623613	4.616228	4.630997	4.612319	4.634906	
702	4.625120	4.614310	4.635930	4.608587	4.641652	
703	4.626627	4.611990	4.641263	4.604242	4.649011	
704	4.628134	4.609307	4.646961	4.599340	4.656927	
705	4.629641	4.606289	4.652993	4.593927	4.665354	
706	4.631148	4.602959	4.659336	4.588037	4.674258	
707	4.632655	4.599336	4.665973	4.581699	4.683610	

Anderson-Darling normality test

```

data:  holtResult$residuals
A = 7.9629, p-value < 2.2e-16

```

-----SES Model Analysis-----

Simple exponential smoothing

Call:

```
ses(y = SESTS, h = 9)
```

Smoothing parameters:

```
alpha = 0.9999
```

Initial states:

```
l = 101.5423
```

```
sigma: 0.2016
```

	AIC	AICc	BIC			
	2341.073	2341.107	2354.717			
Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95	
699	101.4019	101.1435	101.6603	101.0067	101.7970	
700	101.4019	101.0365	101.7673	100.8431	101.9607	
701	101.4019	100.9544	101.8494	100.7175	102.0863	
702	101.4019	100.8852	101.9186	100.6116	102.1921	
703	101.4019	100.8242	101.9796	100.5184	102.2854	
704	101.4019	100.7691	102.0347	100.4341	102.3697	
705	101.4019	100.7184	102.0854	100.3565	102.4473	
706	101.4019	100.6712	102.1326	100.2843	102.5194	
707	101.4019	100.6268	102.1769	100.2165	102.5872	

Augmented Dickey-Fuller Test

data: SESResult\$residuals

Dickey-Fuller = -8.1181, Lag order = 8, p-value = 0.01

alternative hypothesis: stationary

Anderson-Darling normality test

data: SESResult\$residuals

A = 3.5122, p-value = 8.796e-09

-----GARCH Model Analysis-----

Title:

GARCH Modelling

Call:

```
garchFit(formula = data_table$CCI ~ garch(1, 0), data = data_table$CCI,
  trace = F, include.constant = T)
```

Mean and Variance Equation:

```
data ~ garch(1, 0)
<environment: 0x112e031b8>
[data = data_table$CCI]
```

Conditional Distribution:

norm

Coefficient(s):

	mu	omega	alpha1
	100.640560	0.023246	0.975175

Std. Errors:

based on Hessian

Error Analysis:

	Estimate	Std. Error	t value	Pr(> t)
mu	1.006e+02	3.137e-02	3207.812	< 2e-16 ***
omega	2.325e-02	3.766e-03	6.173	6.69e-10 ***
alpha1	9.752e-01	6.253e-02	15.596	< 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood:

-783.8057 normalized: -1.122931

Description:

Mon Apr 30 11:25:07 2018 by user:

Standardised Residuals Tests:

			Statistic	p-Value
Jarque-Bera Test	R	Chi^2	44.98769	1.702339e-10
Shapiro-Wilk Test	R	W	0.8868257	0
Ljung-Box Test	R	Q(10)	2733.066	0
Ljung-Box Test	R	Q(15)	3439.065	0
Ljung-Box Test	R	Q(20)	3943.882	0
Ljung-Box Test	R^2	Q(10)	100.7589	0
Ljung-Box Test	R^2	Q(15)	107.5815	4.440892e-16
Ljung-Box Test	R^2	Q(20)	120.3588	2.220446e-16
LM Arch Test	R	TR^2	105.1672	0

Information Criterion Statistics:

AIC	BIC	SIC	HQIC
2.254458	2.274006	2.254421	2.262015

Autocorrelations of series 'data_table\$GARCH1_eta', by lag

	0	1	2	3	4	5	6	7	8	9	10	11	12
13	14	15	16	17	18	19	20	21					
1.000	0.908	0.766	0.659	0.600	0.572	0.559	0.538	0.512	0.494	0.490	0.488	0.473	
0.449	0.417	0.391	0.378	0.378	0.386	0.379	0.352	0.341					
	22	23	24	25	26	27	28						
0.348	0.364	0.372	0.359	0.330	0.300	0.283							

-----ARMA-GARCH Model Analysis-----

```

*-----*
*           GARCH Model Fit           *
*-----*

```

Conditional Variance Dynamics

```

-----
GARCH Model  : sGARCH(1,1)
Mean Model   : ARFIMA(3,0,2)
Distribution: norm

```

Optimal Parameters

```

-----
      Estimate  Std. Error  t value Pr(>|t|)
mu      101.215054    0.059340 1705.6772  0.0000
ar1       1.962917    0.009355  209.8334  0.0000
ar2      -1.398994    0.015342  -91.1857  0.0000
ar3       0.426647    0.020589   20.7221  0.0000
ma1       0.683141    0.043430   15.7296  0.0000
ma2       0.285677    0.040023    7.1377  0.0000
omega     0.000037    0.000024    1.5726  0.1158
alpha1    0.113514    0.020929    5.4238  0.0000
beta1     0.885486    0.019457   45.5098  0.0000

```

Robust Standard Errors:

```

      Estimate  Std. Error  t value Pr(>|t|)
mu      101.215054    0.109387  925.2927 0.000000
ar1       1.962917    0.012191  161.0077 0.000000
ar2      -1.398994    0.019261  -72.6319 0.000000
ar3       0.426647    0.027800   15.3468 0.000000
ma1       0.683141    0.079303    8.6143 0.000000
ma2       0.285677    0.066655    4.2859 0.000018
omega     0.000037    0.000030    1.2448 0.213203
alpha1    0.113514    0.026285    4.3186 0.000016
beta1     0.885486    0.025028   35.3804 0.000000

```

LogLikelihood : 767.5859

Information Criteria

```

-----
Akaike      -2.1736
Bayes       -2.1150
Shibata     -2.1739
Hannan-Quinn -2.1509

```

Weighted Ljung-Box Test on Standardized Residuals

```

-----
                        statistic  p-value
Lag[1]                  0.7254 3.944e-01
Lag[2*(p+q)+(p+q)-1][14] 10.1811 2.379e-05
Lag[4*(p+q)+(p+q)-1][24] 17.5235 3.936e-02
d.o.f=5
H0 : No serial correlation

```

Weighted Ljung-Box Test on Standardized Squared Residuals

```

-----
                        statistic p-value
Lag[1]                  4.443 0.03505
Lag[2*(p+q)+(p+q)-1][5] 5.425 0.12247
Lag[4*(p+q)+(p+q)-1][9] 7.487 0.16171
d.o.f=2

```

Weighted ARCH LM Tests

```

-----
Statistic Shape Scale P-Value
ARCH Lag[3]    0.1946 0.500 2.000 0.6591
ARCH Lag[5]    0.4709 1.440 1.667 0.8922
ARCH Lag[7]    2.6542 2.315 1.543 0.5811

```

Nyblom stability test

```

-----
Joint Statistic: 8.8357
Individual Statistics:
mu      0.006757
ar1     0.072521
ar2     0.025178
ar3     0.026779
ma1     1.094480
ma2     1.624457
omega   0.507932
alpha1  0.199209
beta1   0.351447

```

Asymptotic Critical Values (10% 5% 1%)

```

Joint Statistic:      2.1 2.32 2.82
Individual Statistic: 0.35 0.47 0.75

```

Sign Bias Test

```
-----
                t-value   prob sig
Sign Bias      0.3059 0.7598
Negative Sign Bias 0.3229 0.7468
Positive Sign Bias 1.1654 0.2443
Joint Effect    1.5163 0.6785
```

Adjusted Pearson Goodness-of-Fit Test:

```
-----
group statistic p-value(g-1)
1    20      13.06      0.8355
2    30      37.47      0.1345
3    40      47.67      0.1607
4    50      52.86      0.3274
```

Elapsed time : 0.3263631

Box-Ljung test

```
data: data_table$uGarch_Eta
X-squared = 11.669, df = 10, p-value = 0.3078
```

Box-Ljung test

```
data: data_table$uGarch_EtaSq
X-squared = 12.147, df = 10, p-value = 0.2753
```

```
*-----*
*          GARCH Model Forecast          *
*-----*
```

```
Model: sGARCH
Horizon: 9
Roll Steps: 0
Out of Sample: 0
```

0-roll forecast [T0=1971-11-30]:

```
Series   Sigma
T+1  101.6 0.08264
T+2  101.7 0.08283
T+3  101.7 0.08301
T+4  101.7 0.08320
T+5  101.7 0.08338
T+6  101.6 0.08356
T+7  101.6 0.08374
T+8  101.6 0.08393
T+9  101.6 0.08411
```


Corrected AIC (AIC * n):

-1517.173

-----END ANALYSIS-----