Time Series Study of an ETF

NEAS Time Series Final Project

A. Wil Edie

September 8, 2013

Contents

1 Introduction

In 1860 the first Standard & Poor index was created by Henry Poor and contained financials for railroad and canal companies. In the first half of the 20^{th} century Standard Statistics and Moody's Manual company merged with the company Poor founded.With the addition of an IBM computer in 1946 the Standard & Poor index expanded to be based upon 500 companies [2]. Since then, the S&P500 has been one of the most relied upon indexes for determining the value of the market and its relative value.

For various reasons, there have been attempts to build portfolios based upon some of the most common indices, like the S&P500, but it was not until the first index based exchange-traded fund (ETF) was created in 1993 to follow the S&P500 that it was a simple process to accomplish [1]. Since then, many more ETFs have been created and one that I have found most interesting is the inverse leveraged ETF (LETF) created by ProShares called the UltraPro Short S&P500 (symbol SPXU) [4]. What is intriguing about this LETF is that it is uses leverage to inversely magnify gains and losses by a magnification of 3 times, meaning that an increase of 1% in the S&P500 translates to a decrease in SPXU of 3%, and vice versa. There are obvious benefits to having such a product on the market, since an investor who believes that the market is heading for a plunge can not only

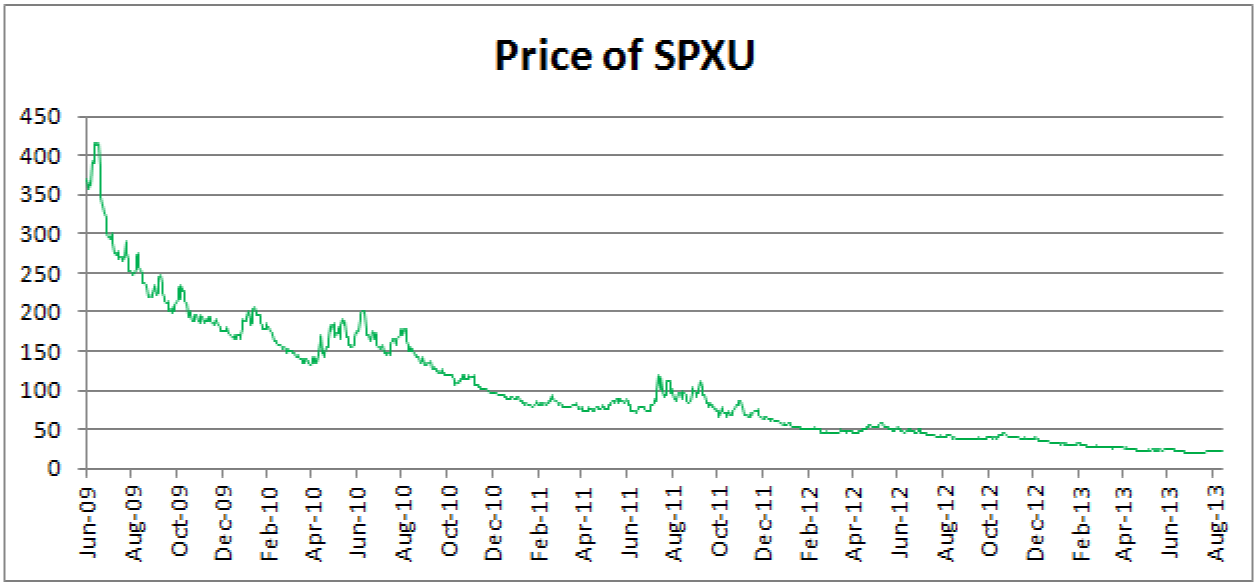

Figure 1: Graph of SPXU Daily Close Price

gain instead of lose value if that happens, but can get three times the gains that he would have had losses.

This report will use time series processes to test whether or not SPXU can be predicted with any measure of certainty. Because of the nature of SPXU and its connection to the S&P500 this is the same question of whether the index can be forecasted.

2 The Data

SPXU has a fairly short history, having been first listed on June 23, 2009 which means we will not have to worry about the 2007 and 2008 price movements in the data. The market has been more volatile since then than it was before, but we will only be comparing post-recession values so this will not need to be accounted for. There was a 1:5 split on May 11, 2012, but the share price was continuous directly after and because of the nature ETFs the price appears to have change little as a direct result.

The data that is used for analysis is the daily close price from Google Finance [3]. The data is from June 25, 2011 to September 6, 2013 can been seen in Table 1 on the following page from most recent to oldest; there are 1067 values. The graph of this data can be seen in Figure 1.

3 Initial Analysis

The first thing that is apparent is that this is not a stationary process, but the sample autocorrelation (ACF) was run anyway, just to be sure. According to equation 3.6.2 in Cryer and Chan's textbook [5] the sample ACF is calculated for lag k by:

$$r_k = \frac{\sum_{t=k+1}^{n} \left(Y_t - \bar{Y}\right)\left(Y_{t-k} - \bar{Y}\right)}{\sum_{t=1}^{n} \left(Y_t - \bar{Y}\right)^2}$$

21.89	24.51	28.54	36.5	36.97	52.1	45.3	62.45	94.5	78.95	83.25	91.1	119.45	148.9	149.45	185.95	188.7	254.9
21.93	23.9	28.06	37.75	36.93	52.97	45.85	62.65	92.85	77.4	80.5	91.75	119.65	147	148.2	195.2	189.1	248.95
22	23.23	27.96	39.21	37.01	50.49	45.6	65.65	98	78.25	81.4	91.4	120.5	157.26	170.55	195.2	196.35	247.75
22.55	23.22	28.42	38.73	36.89	51.68	44.9	64.95	104.3	77.3	81.65	88.75	121.3	157.5	162.85	194.9	194.85	250.15
22.88	24.15	28.53	38.03	36.54	48.44	44.5	66.75	111.1	73.25	81.55	89.2	122	155.4	148.6	200.8	187.45	250.7
22.63	24.81	28.34	38.14	36.98	48.17	46.4	64.5	102.05	71.7	79.85	91.16	125.8	152.45	145.65	199.9	187.35	251.8
22.78	23.79	28.67	38.87	38.9	49.61	46.9	64.7	95.4	74.1	79.1	90.75	126	152.2	135.85	207.5	187.85	251.75
22.97	23.47	29.01	39.02	39.3	50	45.9	66.3	97.65	74.45	78.3	93.35	122.85	157.3	141.85	202.8	196.4	267.3
21.94	23.83	29.18	39.36	39.59	51.45	45.65	68.25	92.1	74.2	77.85	94.3	123.5	161.5	135.05	204.3	200.05	275.45
21.69	22.86	29.27	39.78	38.91	53.17	45.25	68.7	95.35	77.7	78.5	94	122.35	173	140.35	186.85	193.75	282.85
21.92	23.15	30.13	39.97	39.44	52.2	45.75	75.35	102.55	79.9	78.45	93.45	125.1	166.6	143.45	184.2	196.85	291.05
22.52	22.68	30.6	39.75	41.96	54.06	46.1	73.15	104.3	82.1	78.65	92.9	126.35	172.3	133.95	191.65	196.6	270.85
22.12	23.11	30.94	39.3	41.9	52.1	46.8	73.45	95.45	85.45	79.7	94.35	126.75	175.3	132.5	200.7	211.4	264.5
22.43	23.03	30.73	39.19	41.8	53.37	46.65	74.3	87.35	87.75	79.4	94.25	128.8	161.75	135.1	194.4	213.5	270.65
22	22.84	31.93	39.81	42.34	53.47	49.4	71.9	87.15	84.75	81	97.05	128.35	162.15	136.3	187.8	226.25	269.7
21.83	22.37	32.61	40.7	41.45	57.46	49.4	69.95	84.7	84.05	82.7	97.25	128.4	162.2	135.65	190.65	227.3	268
20.92	22.43	30.86	40.06	41.55	58.66	49.95	67.1	86.25	82.6	82.15	96.9	136.7	169.8	139.2	188.3	229.75	278.85
20.6	22.44	31.8	39.93	41.42	58.6	51.5	70.55	91	86.15	82.85	97.15	133.68	170.4	140.94	190.95	234	275.35
20.83	23.1	31.25	41.53	41.45	54.5	52.65	66.35	95	87.4	85.3	97.35	135.45	174.15	134.5	179.15	216.15	272.5
20.75	22.83	30.08	41.8	42.22	54.12	50.3	66.9	97.4	88.3	86.05	97.55	134.4	179.3	134.8	169.3	230.75	275.1
20.57	23.16	30.77	41.89	41.24	51.85	49.8	67.15	99.55	88.8	85.25	97.1	133.4	198.35	139.35	164.4	218.55	288.75
20.78	23.91	30.64	44.65	41.27	53.82	49.35	69.3	92.2	84.3	89.2	98.15	134.85	201.75	139.75	170.55	215.9	289.2
20.57	23.99	30.75	45.29	40.9	53.31	50.25	69.2	89.45	87.7	90.45	100	133.05	201.75	140.75	170.55	208.8	299.85
20.21	24.21	30.85	44.97	40.85	53.56	49.6	69.1	97.6	88	94	100.65	141.75	198.4	143.35	165.05	201.85	296.2
20.18	24	31	43.28	41.08	53.63	50	79.05	95.6	84.4	88.95	101.05	138.4	196.75	144.75	166.45	208.3	293.3
20.26	24.36	31	42.84	42.06	53.93	50.3	79.65	88.8	86.35	86.08	102.8	136.5	190.7	142.45	170.85	202.6	296.2
21	24.74	31.49	43.04	42.17	56.78	50.65	87.5	86.3	85.35	84.65	101.4	135.7	174.85	143.6	166.3	199.55	299.9
21.03	24.94	31.36	43.24	42.2	55.5	51.35	86.8	87.55	85.15	86.35	101.66	142.5	172.7	146.95	167	204.45	320.75
21.04	25.72	31.46	41.53	42.11	53.04	50.8	81.45	88	82.5	81.9	101.85	142.35	175.05	146.95	168.65	201	321.05
20.86	26.46	32.4	38.94	42.31	52.47	50.85	80.65	96.2	80.1	81.55	103.7	142.3	167	150.25	170.65	202.6	326
20.93	25.76	31.35	39.81	42.41	51.56	51.35	76.25	100.6	79.9	83.75	105.05	143.85	165.3	148.74	171.15	213.5	336.35
21.03	25.96	32.35	40.15	42.66	49.99	53	75.9	96.4	74.8	81.85	106.1	143.7	157.75	148.9	172.65	211.85	336.5
20.83	26.49	32.13	39.02	43.32	49.55	52.25	72.35	100.6	77.2	80.05	106.2	148.5	156.35	151.4	181.45	215.2	345.6
20.71	26.31	31.76	40.49	43.5	49.8	52.15	69.1	111.6	78.15	84.4	106	151.1	156.84	151.55	181.45	218.95	378.8
20.8	26.67	32.16	40.38	46.17	48.95	53.25	70.2	111.6	79.25	84.95	106.8	153.05	157.5	150.85	176.5	224.05	385.55
20.94	26.69	32.03	40.25	45.25	48.35	52.2	68.2	106.3	80.05	80.95	111.05	156.1	157.45	148.6	176.45	225.9	416.5
21.27	27.56	32.56	40.63	45.16	48.45	52.35	72.45	94.45	79.8	82.4	119.2	151.25	168.6	151.95	175.6	235.8	413.75
21.44	27.94	32.65	40.31	44.27	46.2	52.85	74.4	94.7	77.05	85.15	116.75	151.25	168.08	154.5	176.4	246.6	414.65
21.21	28.67	32.81	38.65	44.18	45.2	53.2	67.05	92.5	75.45	84.92	117.4	157.5	170.55	152.25	176.4	243.6	414.7
21.45	28.21	33.33	38.73	46.9	44.8	53.1	69.75	98.45	75.95	83.45	113.9	162	187.3	152	179.25	227.2	391.1
21.5	26.98	33.55	36.86	49.36	45.75	55.4	71.1	100.95	77.95	78.7	119.25	177.45	184.6	154.7	180.5	224.25	392
22.4	28.22	34.2	36.59	49.29	45.1	55.8	69.75	116.15	77.9	79.15	114.3	177.85	190.3	158.5	182.8	222.5	392
22.47	26.42	34.17	37.07	48.07	45.4	57.3	73.8	102.25	76.52	79.95	114.05	170.64	182.65	158.6	188	234.9	363.05
22.93	26.23	34.26	38.26	46.66	46.4	57.15	77.55	119.65	74.75	81.45	114.9	179.1	166.1	158.7	191.5	231.3	366.9
23.34	26.44	34.14	39.22	45.43	48.25	56.65	71.75	100.1	75.8	80.75	120.2	175.46	167.6	160.75	184.7	223.7	358.15
24.08	27.45	34.19	38.84	45.82	48.9	56.5	66.8	99.6	73.5	81.3	120.35	177.15	181.9	163	185.6	218.4	369.15
24.16	27.79	34.96	38.9	46.83	47.7	55.7	68.75	87.3	75.5	82.75	115.05	169.75	172.8	163.7	182.95	222.3	366.45
24.06	28.3	35.27	38.2	47.88	47.9	57.1	74.6	88.75	76.4	82.85	114.55	167.95	172.8	163.85	186.4	220.15	
24.72	27.97	34.94	37.1	47.53	47.05	56.9	76.9	82.65	77.4	82.25	110.55	166	166.84	171.25	189.35	221.2	
24.21	28.34	34.66	36.72	49.96	46.55	57	72.75	81.55	75.35	83.4	109.45	158	185.75	172.85	192.45	220.2	
24.64	27.48	35.12	36.65	49.35	48.85	57.15	75.7	79.9	73.9	85	110.7	159	182.28	173.45	194.65	230.9	
25.36	27.86	34.93	37.55	49.25	48.65	58.05	80	79.05	73.1	85.75	108.3	164.9	182.8	174.75	188.75	233.3	
26.14	27.57	37.75	38	48.05	46.9	60.1	81.15	74.65	72.65	86.4	107.65	164.95	176.1	180.2	187.6	237.2	
25.22	27.86	39.99	38.12	47.8	49	60.45	78.3	73.7	73.3	85.95	109	163.5	184	180.45	190.65	237.05	
25.45	27.84	38.63	38.38	46.54	50.1	59.8	83.25	72.6	74.1	90.25	115.55	160.25	165.1	180	186.68	244.48	
23.65	28.43	38.52	37.85	45.97	47.7	60.2	78.75	72.7	75.5	92.3	117.1	148.05	162.4	185.2	187	250.4	
22.78	28.15	37.97	39.01	46.9	46.15	60.35	83.1	75.9	77.58	87.7	120.05	145.8	155.96	178.6	193.6	257	
23.27	28.86	37.64	38.25	47.21	46.1	62.05	82.6	75.7	77.3	88.25	120.05	148.15	156.35	178.6	195.7	268.85	
23.83	28.12	36.74	37.18	51.14	44.7	62.25	85	79.6	78.46	89.45	120.2	146.5	148.35	179.7	186.45	275.2	
23.42	28.68	37.28	37	50.67	44.3	61.9	85.25	77.6	81.95	89.5	120.4	145.95	143.2	183.25	186.45	272.6	

Table 1: Data Values for SPXU from June 25, 2011 to Sept 6, 2013

Algorithm 1 Sample Autocorrelation

```
SUMPRODUCT(OFFSET(data,0,0,COUNT(data)-G2,1)-AVERAGE(data),OFFSET(data,G2,0,COUNT(data)-G2,1)-
AVERAGE(data))/DEVSQ(data)
```

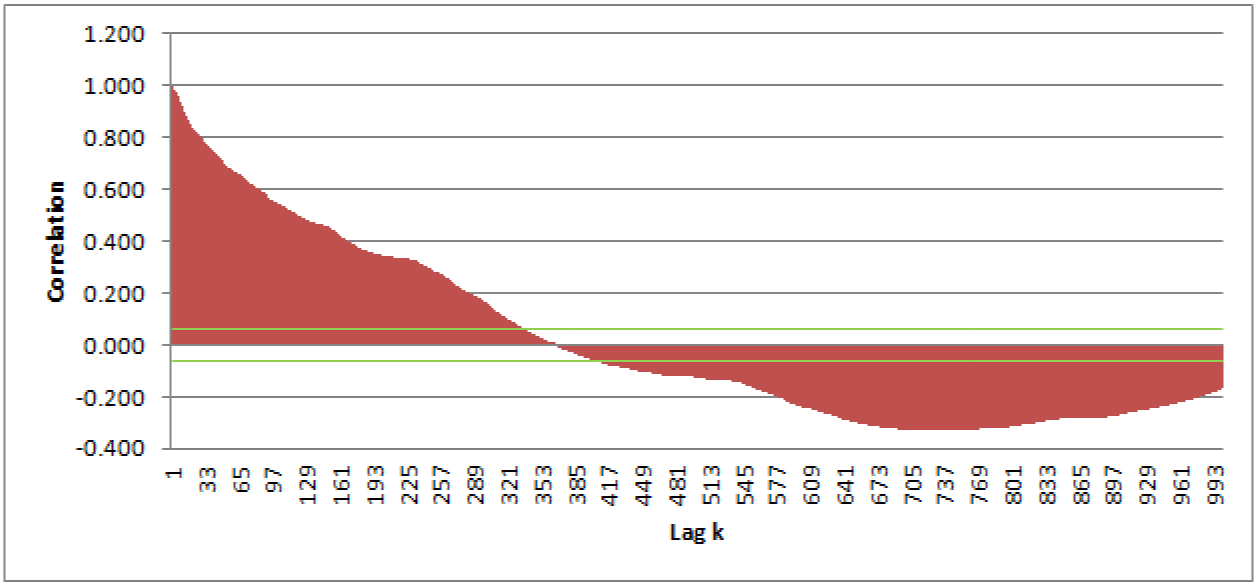

Figure 2: Sample ACF for Original Data

Using the series of Excel functions shown in Algorithm 1, using the named range "data" for the closing prices, we get the values that generates Graph 2 On this graph the correlation between lags is in red and the approximate standard deviation of $\pm 2/\sqrt{n}$ is in green. Clearly, this will need some work to become stationary.

4 Taking Differences

According to chapter 5 of the Cryer and Chan text if a process is not stationary, differences can be taken to see if a stationary process can be made from them [5]. Graphs 3 on the following page , 4 on the next page , and 5 on page 6 show the first, second, and third differences for the series and their associated ACFs are shown in graphs 6 on page 6 , 7 on page 7 , and 8 on page 7 .

From these graphs it appears that of the three differences only the second has relatively low correlations, with the first correlation being the greatest by far. There are some values that extend beyond the approximate standard error lines at several lags. Taking logarithms and then differences was attempted for first, second, and third differences but their sample ACFs were quite worse than the non-log differences.

We will use the second difference to build an autoregression model on, and the ACF graph indicates that an AR(1) should be appropriate.

5 Estimating the Parameters

Using the Regression add-in for Excel a regression was run using the original series for the Y_t values and the Y_{t-2} values for the X values. The output is shown in Table 2 on page 8 where the p-value for both the intercept and X Variable 1 are low enough to suggest a good fitting model.

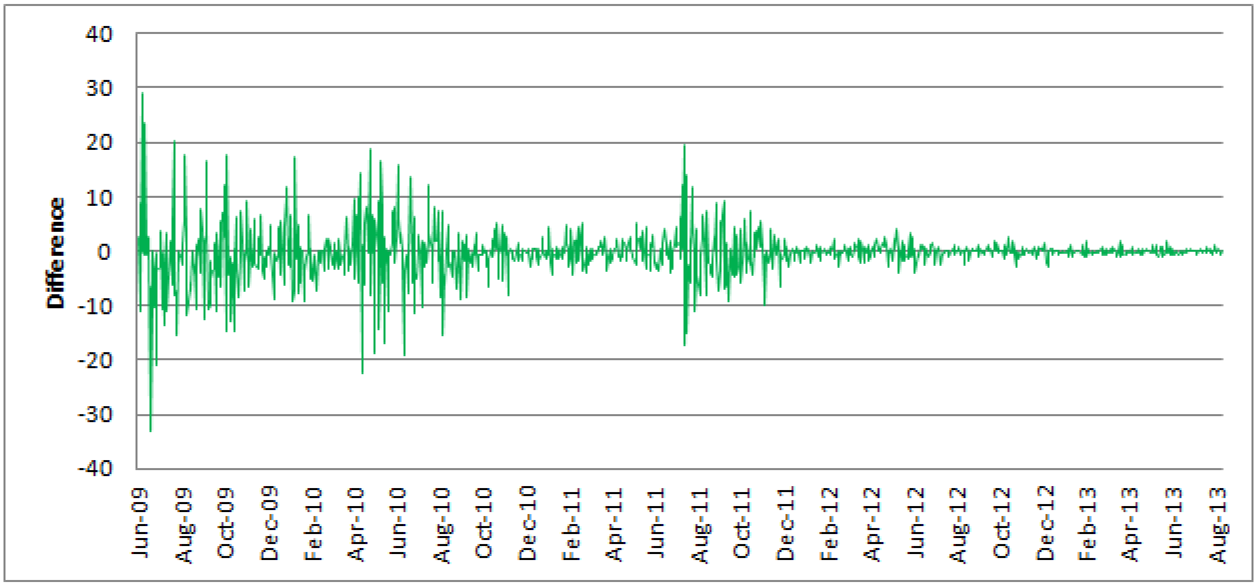

Figure 3: Values of First Difference

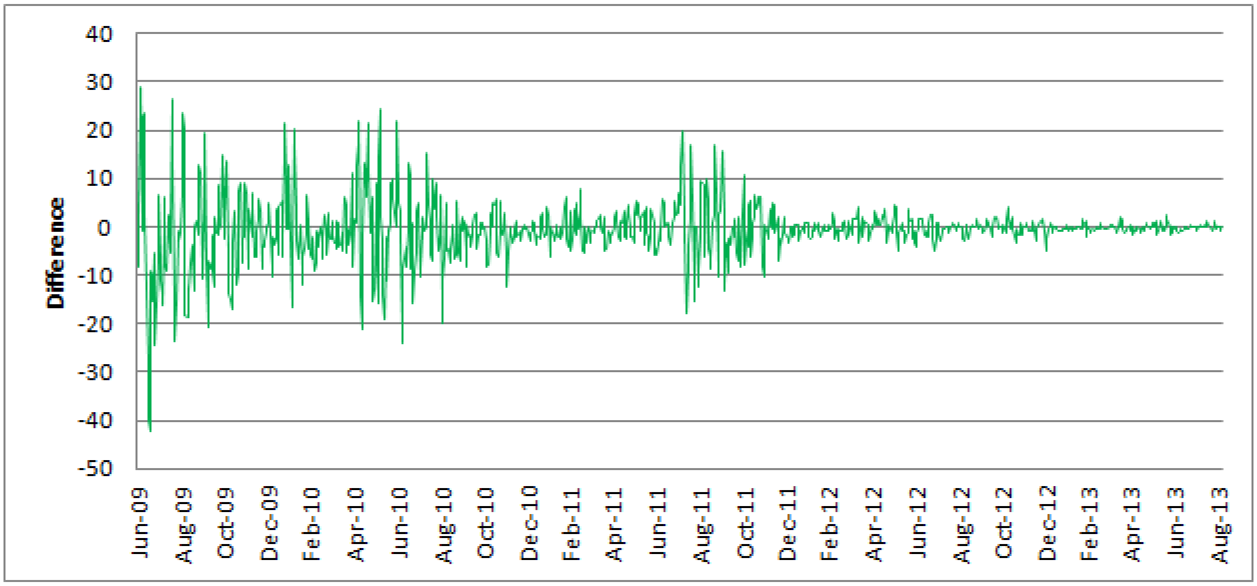

Figure 4: Values of Second Difference

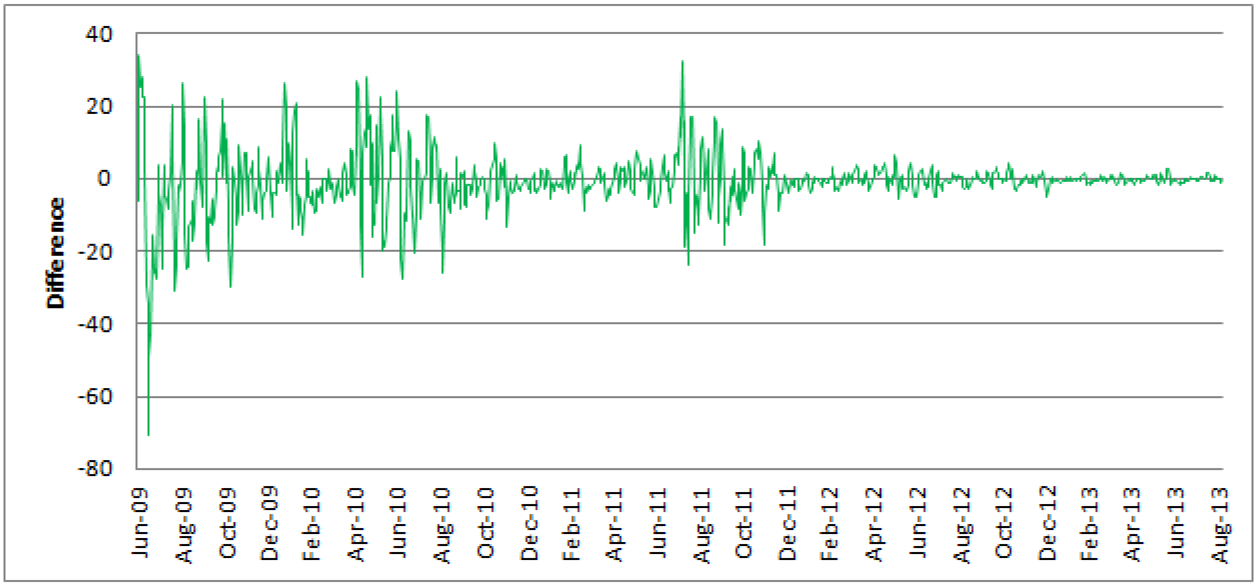

Figure 5: Values of Third Difference

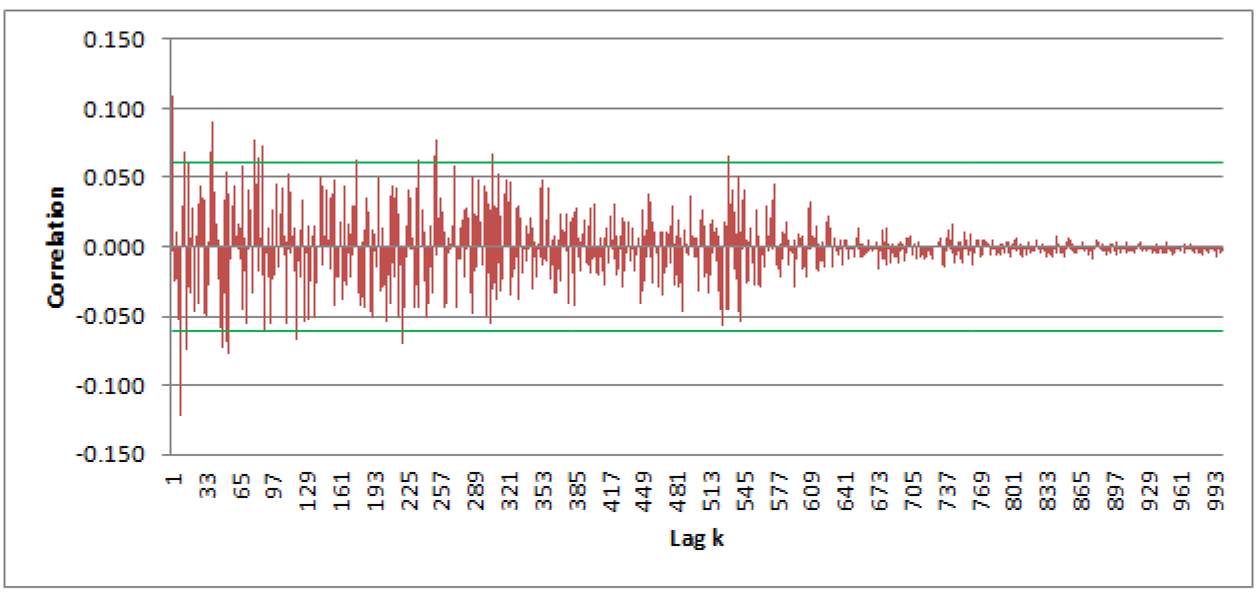

Figure 6: Sample ACF of First Differences

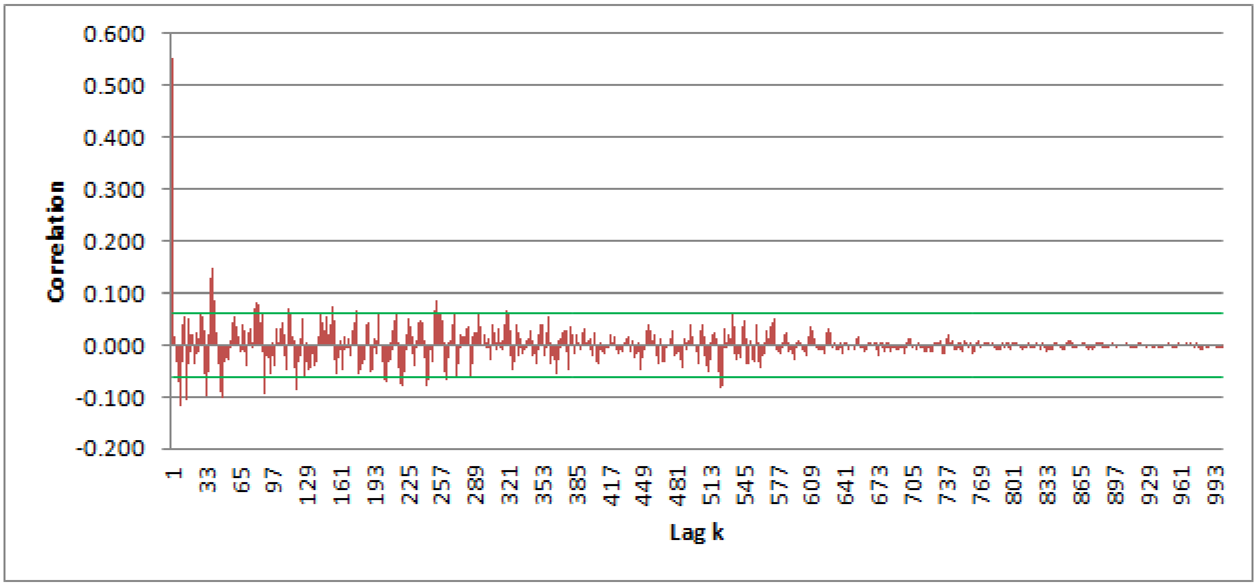

Figure 7: Sample ACF of Second Differences

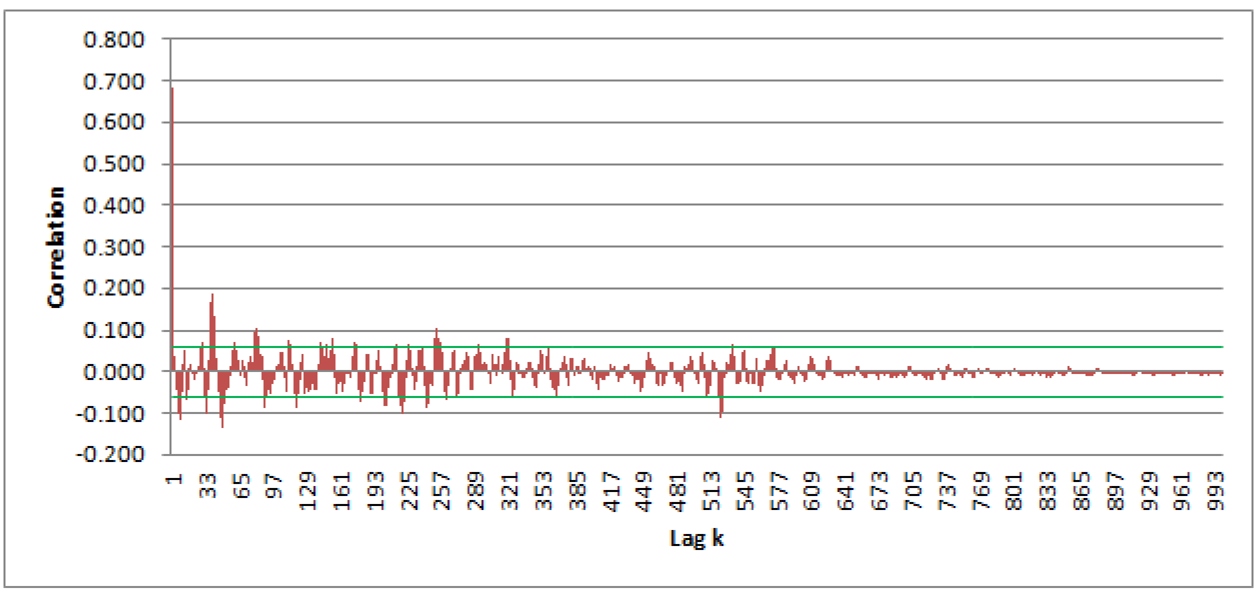

Figure 8: Sample ACF of Third Differences

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.996
R Square	0.993
Adjusted R Square	0.993
Standard Error	6.233
Observations	1065

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	5838313	5838313	150262	0
Residual	1063	41302	39		
Total	1064	5879615			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.807	0.323	2.498	0.013	0.173	1.441	0.173	1.441
X Variable 1	0.986	0.003	387.636	0.000	0.981	0.991	0.981	0.991

Table 2: Regression Output for ARI(1,2)

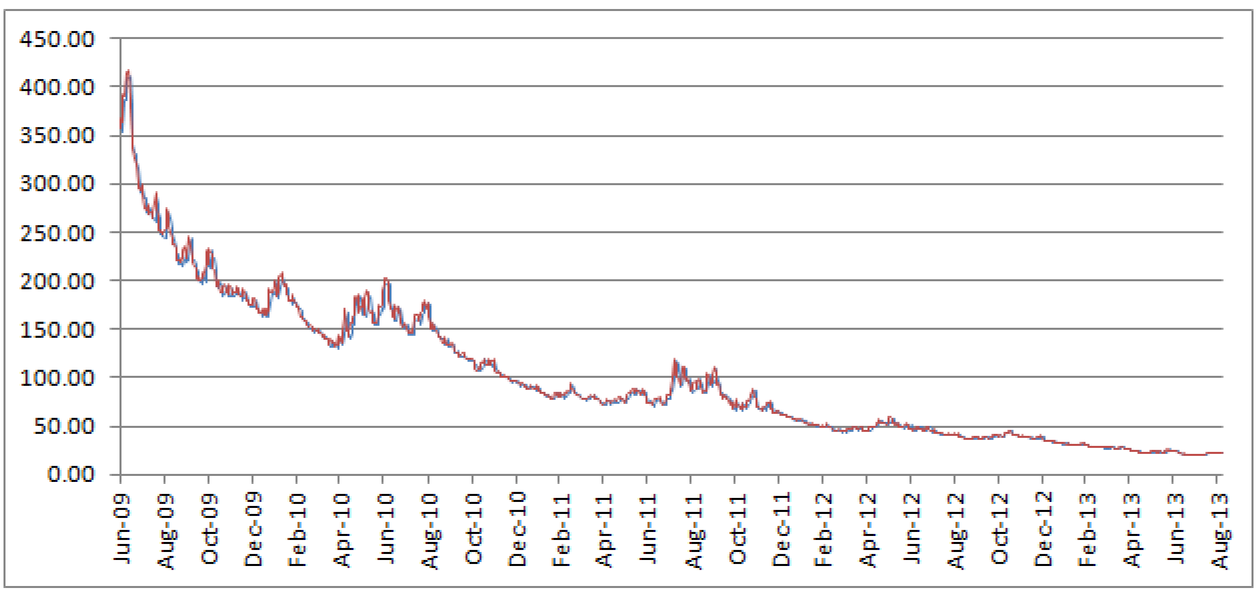

Figure 9: Graph of ARI(1,2) (blue) Versus Actual (red) Values

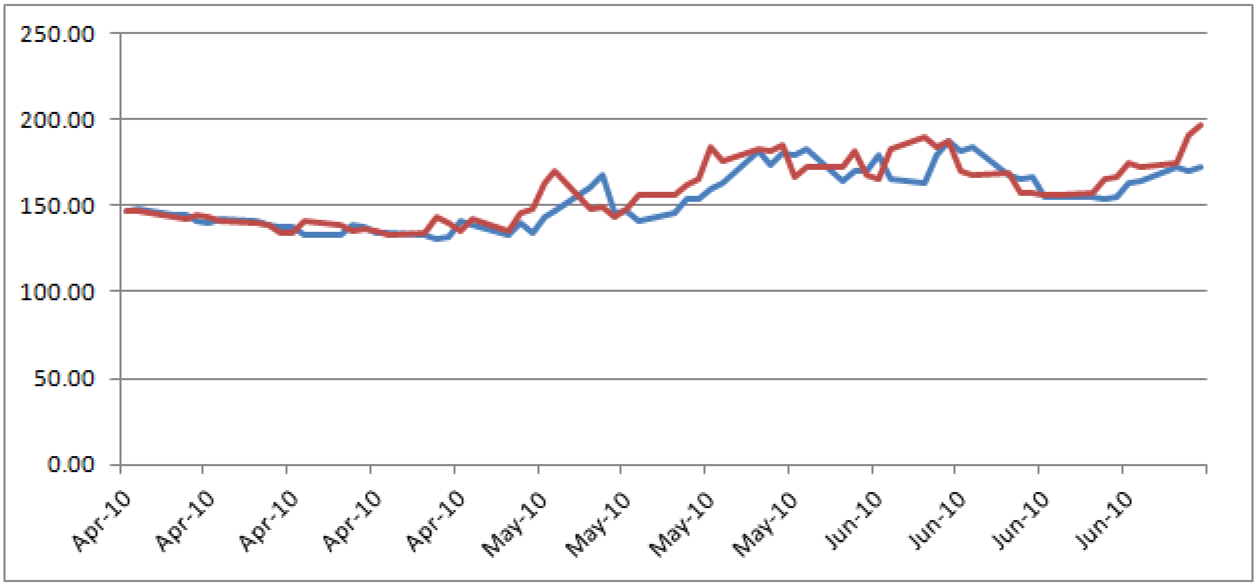

Figure 10: Graph of April - June 2010 of Predicted Values (blue) Versus Actual (red) Values

The regression estimates the Intercept value as 0.807 and the X Variable 1 as 0.986. Using these values to create the equation $\hat{Y}_t = 0.986 \cdot Y_{t-2} + 0.807$ which is the predicted value of Y_t. Graphing these predicted values against the actual values of Y_t we can see that this model is a good approximate to the actual values on Graph 9 on the previous page. A three month period is displayed on Graph 10 to show how closely the model is to the actual values. Unfortunately the volatility of stock prices and their quasi-random nature makes the model very reactive and the predicted values are not as helpful as one would hope.

A q-q plot of the differenced data is shown in Graph 11 on the next page, which helps show why the model is not as helpful as one would hope. Instead of a linear relationship there is convexity below the origin and concavity above. A study of other differences' q-q plots shows that the second differences is indeed the closest to a stationary series as one would hope.

6 Forecasting

To make a forecast and to see how helpful it is, the same model was used, but the values were removed for August and September of 2013, and only values from January through July 2013 were used to build the model. Once the values were trimmed the parameters were re-estimated for an ARI(1,2) model with the regression output shown in Table 3 on the following page where we can see that the estimators are similar to the last regression, albeit with a slight decrease in Intercept estimate. Using this model we forecast the values for August and September of 2013 using only the last July data point. This is shown along with the actual values in Graph 12 on page 11. The forecast follows the general shape of the actual values; however the forecast overestimates all future values.

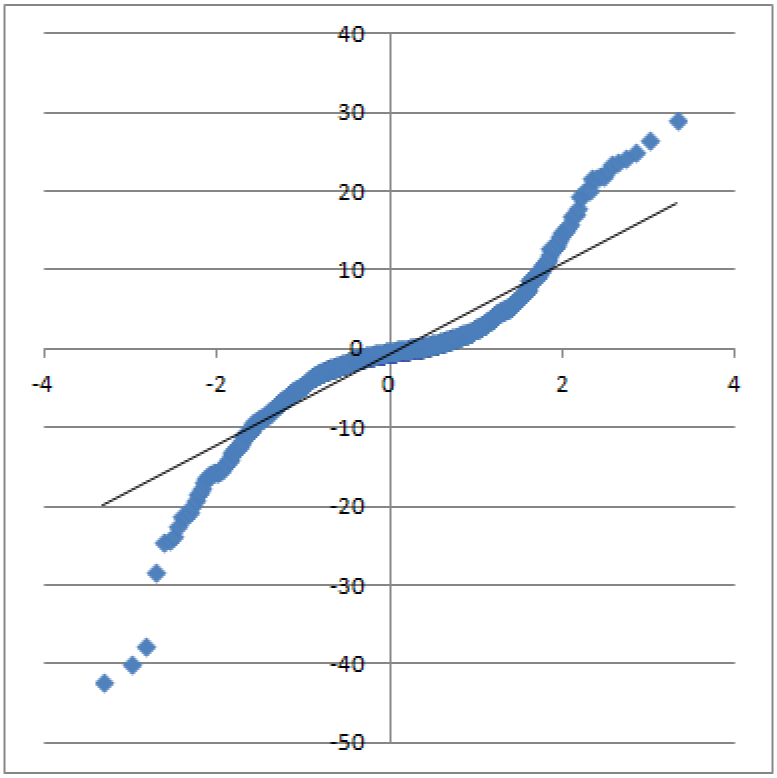

Figure 11: q-q Plot For Second Differences

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.985
R Square	0.970
Adjusted R Square	0.969
Standard Error	0.705
Observations	144

ANOVA

	df	*SS*	*MS*	F	*Significance F*
Regression	1	2249	2249	4521	0
Residual	142	71	0		
Total	143	2320			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.523	0.398	1.314	0.191	-0.264	1.309	-0.264	1.309
X Variable 1	0.974	0.014	67.235	0.000	0.945	1.002	0.945	1.002

Table 3: Regression Output for Prediction Model

Figure 12: Graph of Forecast Versus Actual Values

7 Conclusion

In conclusion, the model that was constructed as shown in this paper helps emphasize why the random walk theory for asset prices is so resilient against detractors. The model that was constructed, although fairly simple, had no real use for forecasting future prices. The usefulness of a model is not necessarily whether it can model past values, but whether any kind of useful prediction can be made. The first difference model using an autoregression parameter simply was not complex enough to forecast very accurately. Having said that, a person who would have used the model on July 31^{st} to invest in SPXU, based on the forecast that it would go up, and sold on Sept 6^{th} would have pocketed a profit of 4.25% over the course of 37 days. All-in-all that proves that even though the model may not be perfect, it would have at least been profitable.

List of Figures

List of Tables

References

[1] A focus on etfs. *Wall Street Journal*, 2006. http://online.wsj.com/ad/focusonetfs/history.html.

[2] Investopedia - how stock market indexes changed investing, April 2011. URL `http://www.investopedia.com/articles/07/history-indexes.asp`.

[3] Google finance - spxu, July 2013. URL `https://www.google.com/finance/historical?cid=12136779&startdate=Dec+1%2C+1999&enddate=Sep+8%2C+2013&num=30&ei=hI8sUpDuH8KQrgGXDg`.

[4] Yahoo! finance, August 2013. URL `http://finance.yahoo.com/q/pr?s=SPXU`.

[5] Jonathan Cryer and Kung-Sik Chan. *Time Series Analysis*. Springer, second edition, 2008.