ATTEMPT TO PREDICT THE STOCK MARKET
An Interactive Qualifying Project Report
submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillments of the requirement
for the
Degree of Bachelor of Science
by

Artur Wojtak

Project Number: MH - 0777

Date: 02/28/2007

Student Signature: ______________

Advisor Signature: ______________

Table of Contents

Abstract

The stock market appears in the news every day. You hear about it every time it reaches a new high or a new low. But can someone somehow actually predict short term price of an individual stock? This report will attempt to do what many investors and mathematicians have tried to do for decades, and that is to make such predications.

Executive Summary

Many have tried predicting the stock market, but very few have succeeded. It is nearly impossible to predict the market for a long period of time, but with the correct mathematical algorithms, and if other major factors that affect the stock market remain unchanged, we can predict how the stock will act from its previous behavior.

For the past six months, Professor Humi and I have been working on a design of a MatLab program, which would be able to predict the price of the stock. We constructed two programs; however one seems to provide us with a better prediction than the other. Both of these programs use many mathematical algorithms to predict the price of the stock.

The program that seems to provide us with a better prediction was Program 1. This program first took the best-fit curve of the actual price, and interpolated it for the amount of days we wish to predict the price of the stock. We then took a Fast Fourier Transformation of the actual price, cleaned out noise, and took the Inverse Fourier Transformation, in order to clean out small noises in the stock. Once we had the cleaned Inverse Fourier Transformation, we interpolated it again for the same amount of days as the interpolation of the best-fit curve, and added the two components together. These two components provided us with the prediction for a certain number of days.

The prediction we computed was fairly accurate, once small modifications were made. Certain stocks gave us a reasonable prediction without any modifications, while other predictions needed to be modified by a certain percentage (10%), after that small modification was made, the prediction was reasonable once again.

From this project we realized that predicting the stock market is very difficult due to the everyday changing of economy which we cannot predict. We however attempted to predict the stock market in ideal situations, and in some manner we have succeeded.

Introduction

The study of the stock market is a fascinating subject. It changes the lives of investors on a daily basis based on the decisions they make on what stocks to purchase. However, what if the decisions made could never be wrong? What if we could predict the price of the stock for the next day, week or even a month? Then many investors, including myself would be purchasing stocks, and becoming instant millionaires. This fact alone motivates myself, as well as many other mathematicians and investors to answer the question of how the stock market can be predicted.

But how could we always make the correct decisions on when to buy or sell? Using the knowledge I gained from certain courses such as Calculus I-IV, Linear Algebra and Probability and Statistics, many different mathematical algorithms could be used to predict the price of the stock under appropriate conditions. These courses taken at WPI have provided me with a vast range of mathematics and each of them will help me with answering the question of how to predict the stock market. Each of the courses concentrates on different areas of mathematics, and in order to answer my subject question, I will need to use many concepts from each of the areas, test the concepts, and see which works best with the stock market.

If successful in predicting the near future of some stocks, I will be able to buy the stock low, and sell high, which is the main goal on Wall Street. I would make my idea available to the public, which would be beneficial for the investors as well as the general population since everyone, whether an expert in stocks, or just a beginner, could have some sort of strategy before making their purchase.

At WPI, the Interactive Qualifying Project (IQP) is mandatory project that everyone needs to complete for their graduation requirements. An IQP topic should insure that both technological and society aspects are represented from a real-life situation. The project also needs to challenge the student to search out existing literature which explains the complexities and multiple possible solutions to the real problem. In my IQP project of how to predict the stock market, the criteria of what a good IQP project should present is met. The project does meet technological and society aspects as the stock market (ex. NASDAQ or DOW JONES) is dependent on what happens with everyday technology as well as everyday economy.

While driving towards my goal of predicting the stock market, I hope to accomplish many other goals along the way.

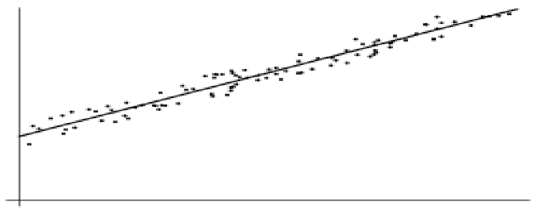
Some of the goals I want to accomplish are learning how to work with those who are interested in the same subject, proper time management skills, how to organize a project that will span over 21 weeks, and the most important skill I want to develop is how to research and present a topic in a professional atmosphere. By achieving my main goal I will be able to learn these three smaller goals as well, which will help me not only at my career at WPI, but my career after graduation. When finished with this project, there are many ways I will disseminate my work. I will start by writing a detail report of everything I've done along the way, by presenting my research in front of WPI faculty, and possibly by publishing my work over the WEB.

Report

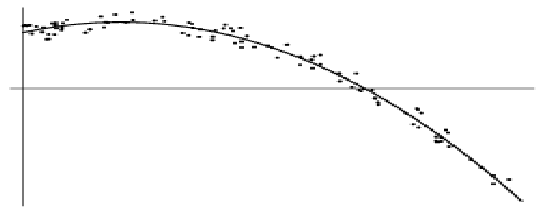
Explanation of Concepts

Predicting the stock market can be very difficult; many have tried, but very few have actually had any positive outcome. While I try to predict the price of a certain stock, I will be using various mathematical algorithms such as least squares, Fourier transformations, inverse Fourier transformations, and interpolation. These algorithms combined should provide me with an idea of how the price of the stock will behave in the following few days. The following few paragraphs will explain how each of these mathematical concepts work, and how they will be applied in this situation.

The first concept discussed will be the least square fit. The least square fit is a mathematical procedure for finding the best-fitting curve to a given set of data points. The least square concept is important for many different reasons, but the main reason is to check whether the given data points fall roughly on the best fit line. There are many different types of least square fit, but for our purpose I will use the linear least square, and the second and third degree least square. The linear least square fit is the most common of all, and this concept gives a linear fit of all the data points in the slope-intercept form $y = mx + b$. The following graphs show linear least square and second degree least square.



Linear Least Square

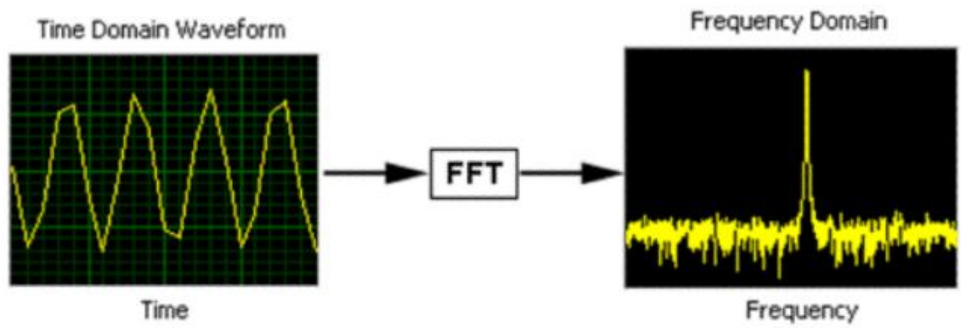


Second Degree Least Square

As we see from the graphs, the linear least square will always provide us with a straight line which is the best fit for all the points. The second and third degree least square will provide us with a parabola in the slope-intercept form of $y = ax^2 + bx + c$ for the second degree least square and $y = ax^3 + bx^2 + cx + d$ for third degree least square.

The next mathematical algorithm used was fast Fourier transformation. The fast Fourier transformation (FFT) is basically a very fast way of calculating the discrete Fourier transformation (DFT). FFT works by splitting the set of data to be transformed into smaller sets of data to be transformed. For example, if you have 16 data points, the FFT would first split the data into two sets of eight points, then split the data into four sets of four points, then split that data again into eight sets of two points; at each stage of the splitting, the FFT would calculate the

DFT and the results would all be combined in order to calculate the final FFT of the data. In general, the FFT is used in order to convert a function from amplitude as a function of time into a function of amplitude as a function of frequency. The following example shows how the transformation looks when we apply the FFT. In this example we went from time versus amplitude to time versus frequency.



In order to calculate the FFT the following equation is used:

```
a(n)=real(data(k));

b(n)=-imag(data(k));

FFT=(a(n)*cos((2*pi*(k-1)*(n-1))/L) +
b(n)*sin((2*pi*(k-1)*(n-1))/L));
```

Where L = length of data, k = number of frequencies and n = data point to be converted

The next concept used was the inverse fast Fourier transformation (IFFT). The IFFT takes the computed data of frequency versus time and converts it back to amplitude versus time. Depending of the amount of frequencies converted in the FFT, the IFFT graph will look fairly similar to the original data. For example, if we have a data set of 365 and we use k=365 in the FFT, when we use the IFFT function to convert back to time versus amplitude, we would get our original data points. However, if we used k=10, and converted back to time versus amplitude, we would only get the most important ten points instead of the original 365. Examples of how the graphs look at different k will be shown further in the report.

Program 1 Explanation

There were two programs written in MatLab, each of them attempting to predict the price of the stock for the near future. Each program will be dissected and described while showing the outcome in a graph form of each mathematical algorithm used in the described section above. Both of the complete programs can be seen on the last pages of this report. For the purpose of showing how each program differs, we will use the same stock in both programs; the stock we choose is Apple Inc. At the end of explaining each step of each program, we will show how the program needs to be altered in order to perform a similar task on other stocks.

Program 1 starts by the following code:

```
load apple.dat.txt
for i=1:365
      x(i)=i;
end
p1=polyfit(x',audibleinc(1:365),1)
xx=(1:1:458)';
pp1=polyval(p1,xx);
plot(xx,audibleinc)
hold
plot(xx,pp1)
xlabel('time(days)')
ylabel('stock price')
title('least squares vs. stock price')
print -dpsc2 -r600 apple.ps
hold
```

This section of the code loads the data for the 458 days, and then splits the data into two separate matrices; the first matrix is the data for days one through 365 and the second matrix is the data for days 366 to 458. When we run this piece of code in MatLab two things occur, we get a output for a graph of best-fit line, as well as and output from MatLab that shows us the equation of the line. The following three equations were found when we ran them through MatLab:

First Degree : $0.085x + 46.4132$

Second Degree: $-0.0004x^2 + 0.2289x + 37.6096$

Third Degree: $0.000x^3 - 0.0034x^2 + 0.6647x + 24.2266$

The following three charts will show how the data looks with best-fit polynomials of degree one, two and three.

Data points with first degree best-fit curve. $y = ax + b$.

Data points with second degree best-fit curve. $y = ax^2 + bx + c$.

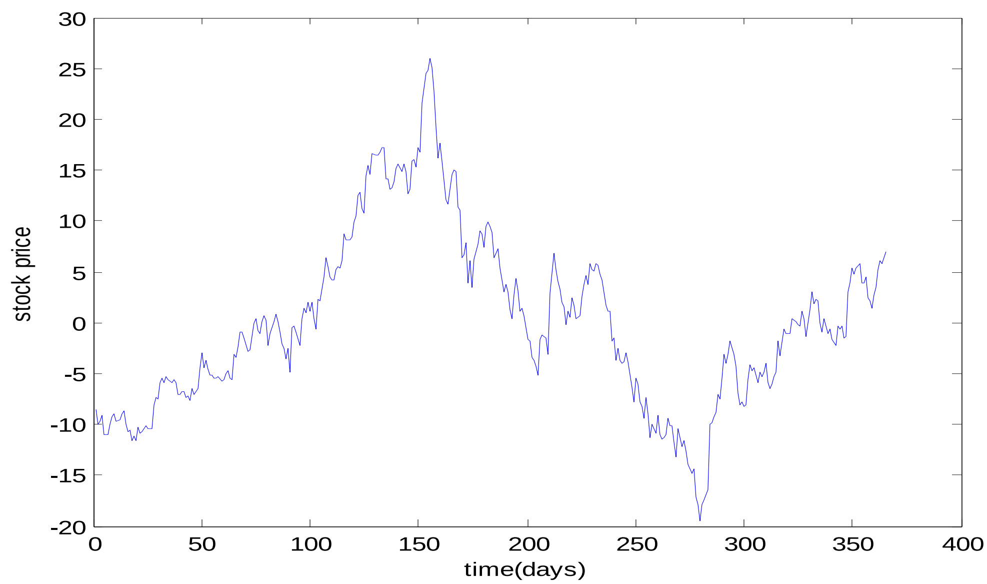
Data points with third degree best-fit curve. $y = ax^3 + bx^2 + c$.

The best-fit curve appears to be either first or third degree. We will continue with both and show how each will have a different outcome once the whole program is run. The next piece of code shows the difference between least squares and stock prices for the first 365 days. The code is run in this sequence:

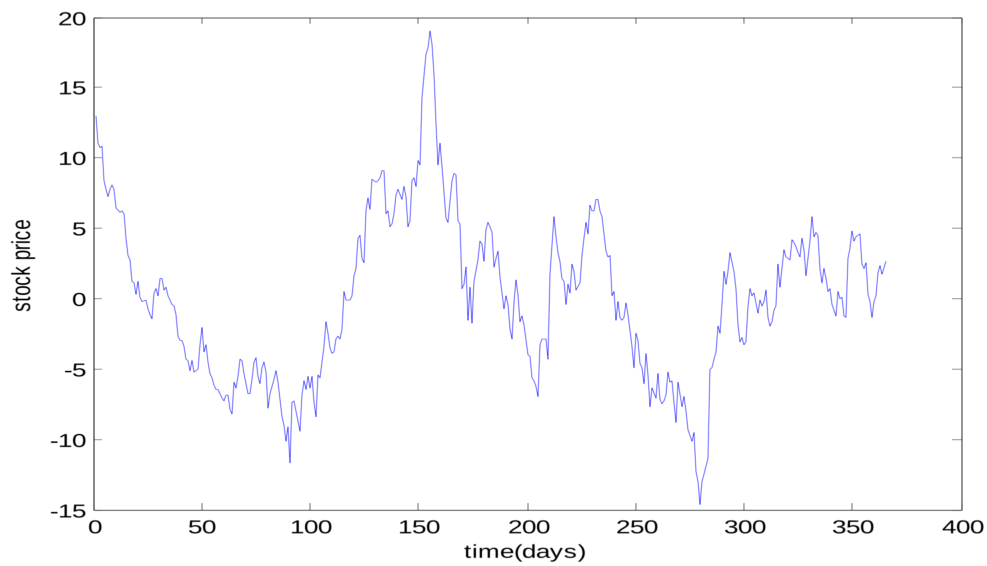
```
%% difference between data and trend for the first 365 days
diff=audibleinc(1:365)-pp1(1:365);
```

```
plot(x',diff)
xlabel('time(days)')
ylabel('stock price')
title('difference between least squares and stock price first 365 days')
print -dpsc2 -r600 -append audibleinc.ps
```

This piece of code takes the best-fit curve we found earlier, and subtracts is from the actual price of the stock. The graphs show the difference between the prices on the y-axis, and the day on the x-axis. The following two graphs will show the difference between the first degree least square and the actual stock price, and the difference between third degree least square and the actual stock price.



Difference between least square (first degree) and stock price.

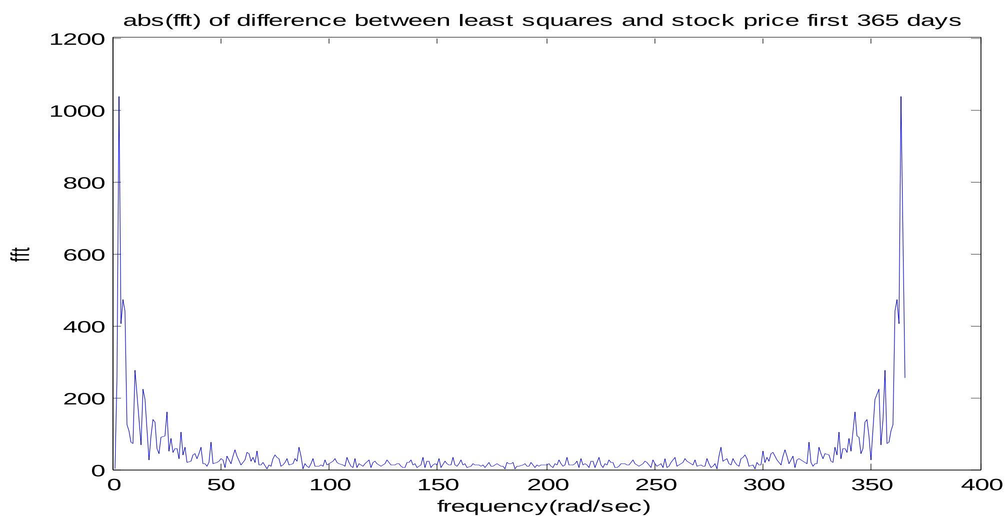


Difference between least square (second degree) and stock price.

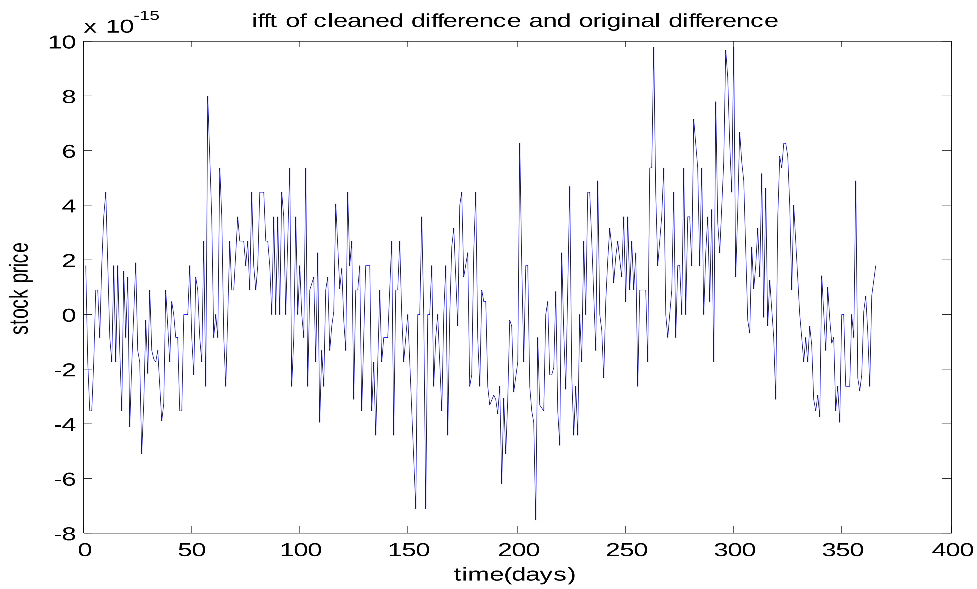
According to the graphs above, the difference between the least square and the stock price appears to be smaller in the third degree polynomial. We will continue using both degrees to show the importance of both. The next piece of code takes the FFT of the difference between least square and data points, then plots the absolute values, and finally it takes the IFFT. There could be many variation to this piece of code; in one of the variations we could set a certain number of points to zero after we take the FFT, and in other variations we could only choose a certain number of points that fall within a certain percentage of the first point of the FFT. In other words we will be cleaning the data, and showing how different cleaned data will predict the future stock prices.

```
Y=fft(diff);
plot(abs(Y))
xlabel('frequency(rad/sec)')
ylabel('fft')
title('abs(fft) of difference between least squares and stock
price first 365 days')
print -dpsc2 -r600 -append audibleinc.ps
PP=ifft(Y);
```

As earlier stated, we will first show how the graphs of the FFT and IFFT appear before any data is cleaned, then we will show how the graphs appear after different cleaning techniques. The following shows FFT and IFFT un-cleaned.

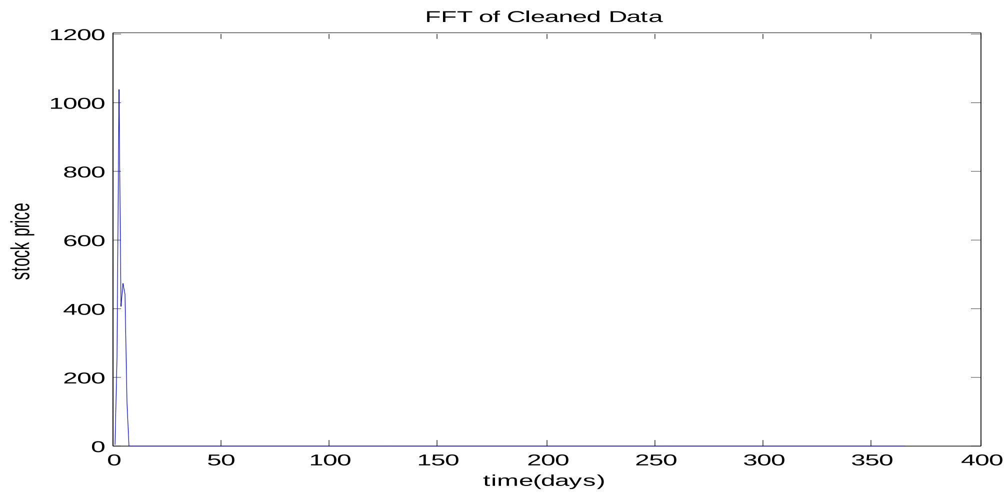


FFT of the first 365 data points.

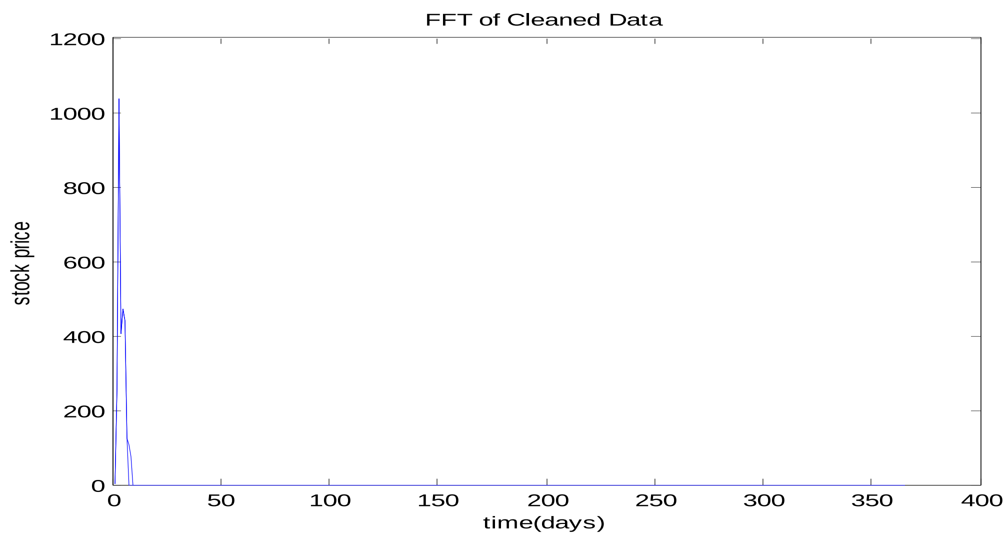


IFFT for all the data points.

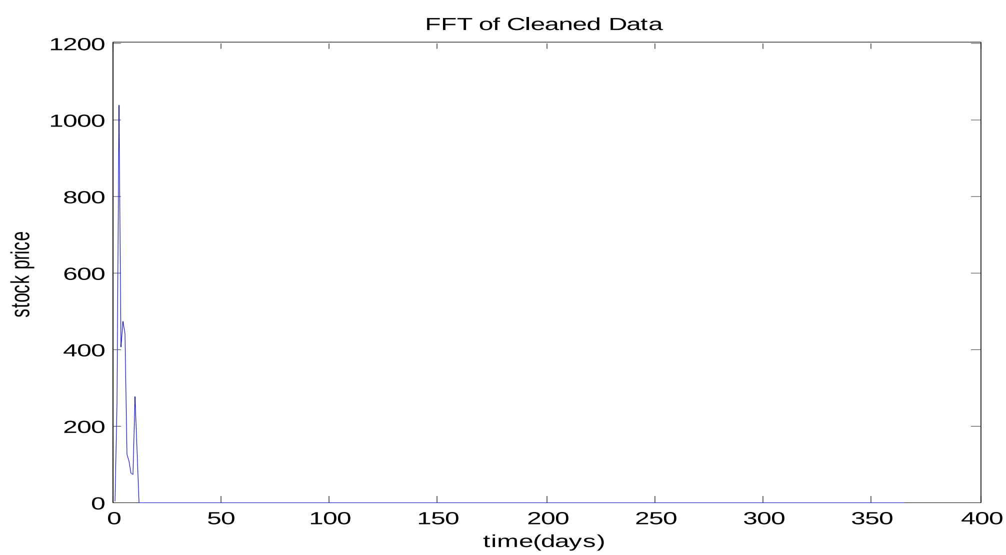
The first graph above shows the FFT for the whole data set and the second graph shows the IFFT if all the data points from the FFT are used. In the second graph we see the difference of IFFT and original data points. Since no data points were cleaned, we should have a resultant of zero as our difference, and we see from the graph that the resultant is 10^{-15}, which is caused by a rounding error. This however is so small that we can ignore this for our purpose. The following three charts will show FFT and IFFT if we only use the first 8,12, or 20 points from the FFT. The three FFT graphs will look very similar to one another. The first graphs only has the first 8 points of the FFT, and the rest are set to zero, the second graph only has the first 10 points and the rest are set to zero, and finally the third graph has 12 points and the rest are set to zero.



Cleaned FFT - using only first 8 points

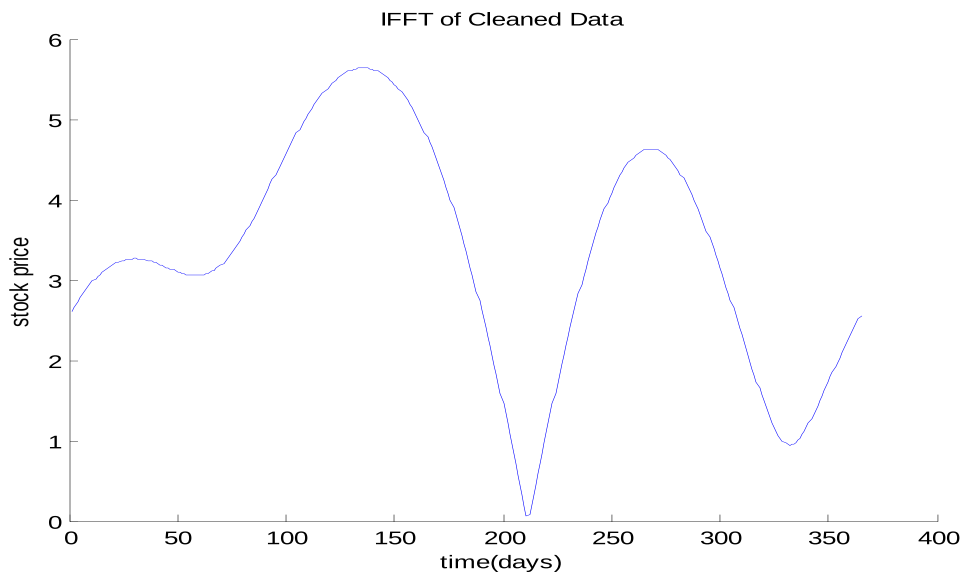


Cleaned FFT - using only first 10 points

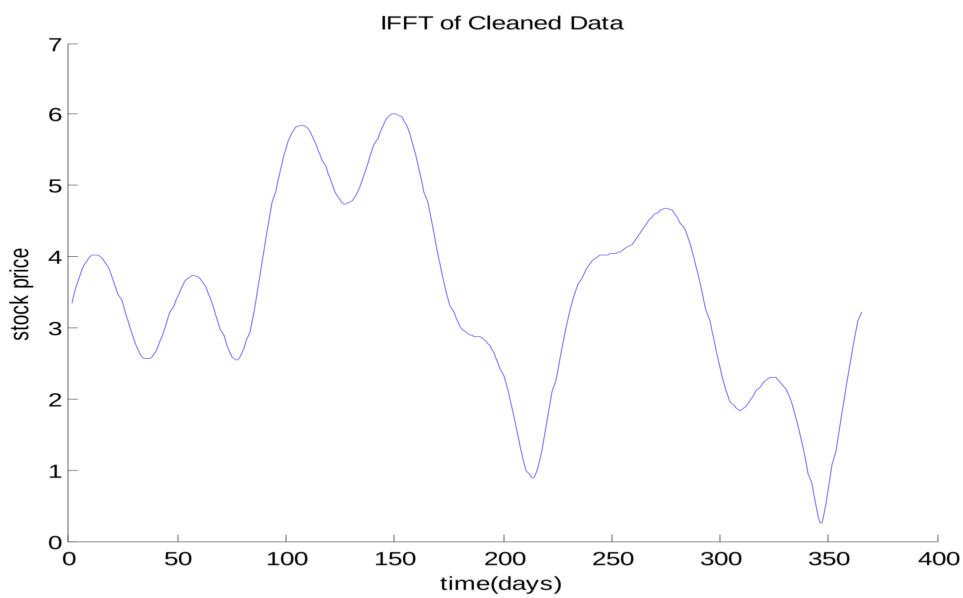


Cleaned FFT - using only first 12 points

As we could see there is a slight difference in the FFT when different amount of points are used. Since our goal is to minimize the amount of points, we will continue working with only the first 8 and 10 points to show the IFFT. The graphs looks as following:



Cleaned IFFT - using only first 8 points



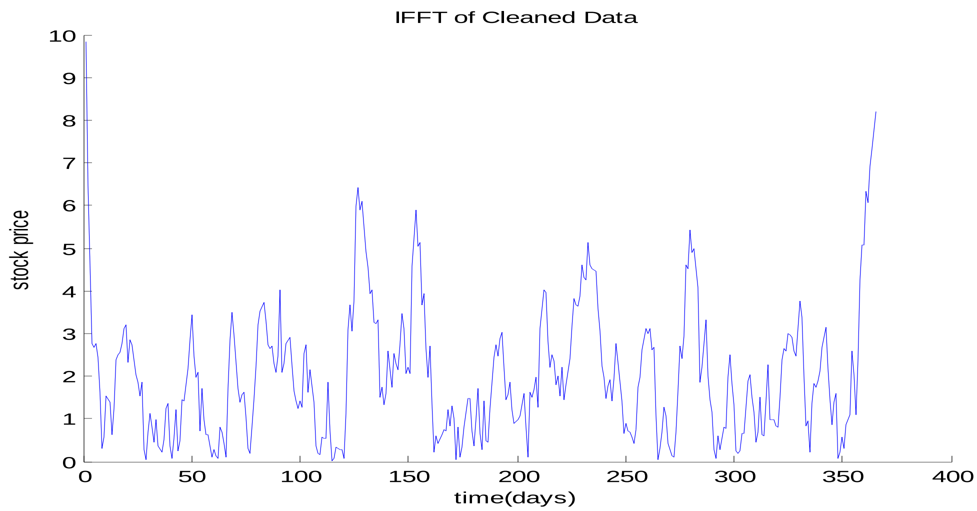
Cleaned IFFT - using only first 10 points

From the two previous graphs we see that there is a difference between how many points we take from the FFT. In our study ideally the fewer amounts of points we take the better, as we eliminate noise, and only take into consideration the major points. In the following three graphs, I will take a percentage of all the points compared to Y(1). For example in the first graph a point would need to be at least 5% of Y(1) in order to be considered, if not it will be set to zero. In graphs two and three I will use the same concept except the percentages will be set slightly higher at 10 and 15 percent.

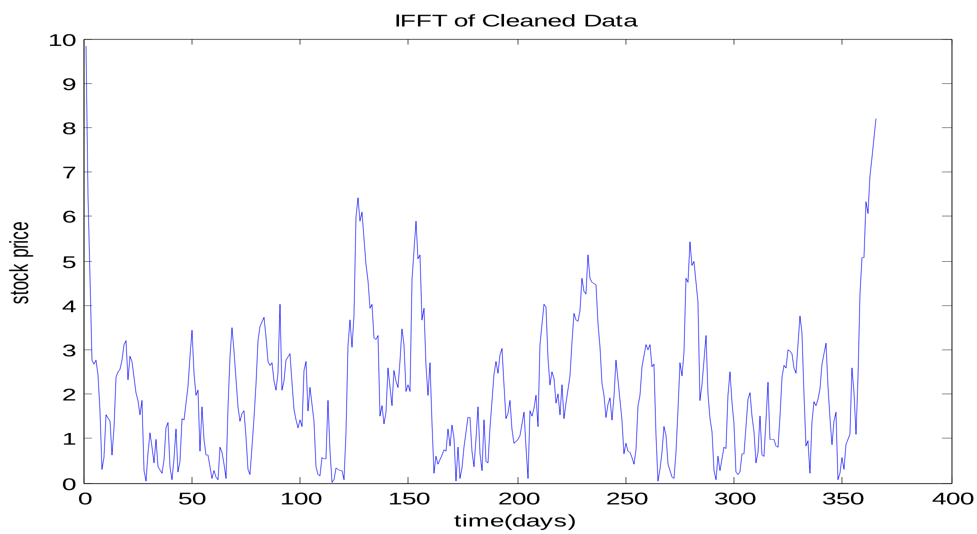
The code used to calculate the different percentages of points looks as following:

```
for i=1:365
    if Y(i) > .05*Y(1);
        Y(i)=0;
    end
end
```

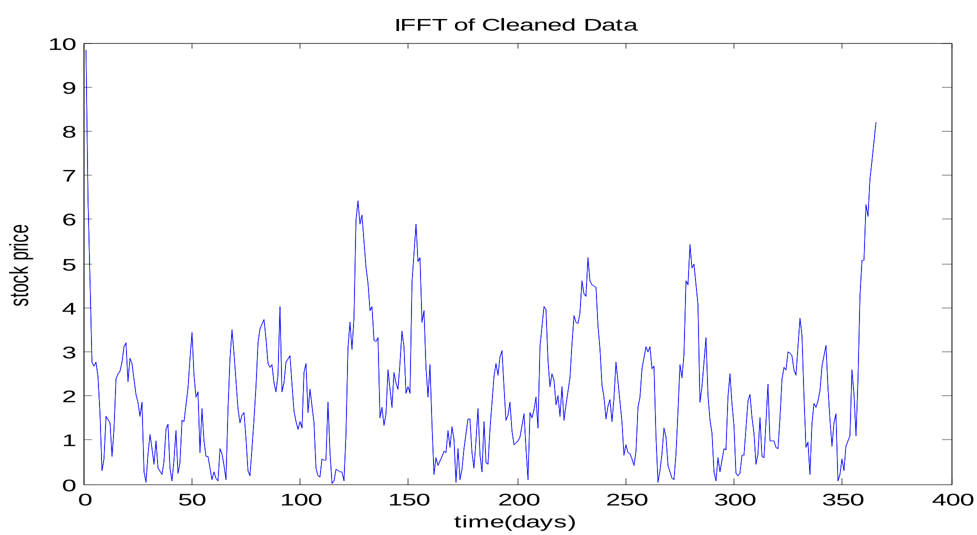
In order to change the amount of percent of points you wish to include, change the percentage which multiplies by Y(1). The following graph shows 5% of points taken.



IFFT of the data points that fall are bigger than 5% of Y(1)

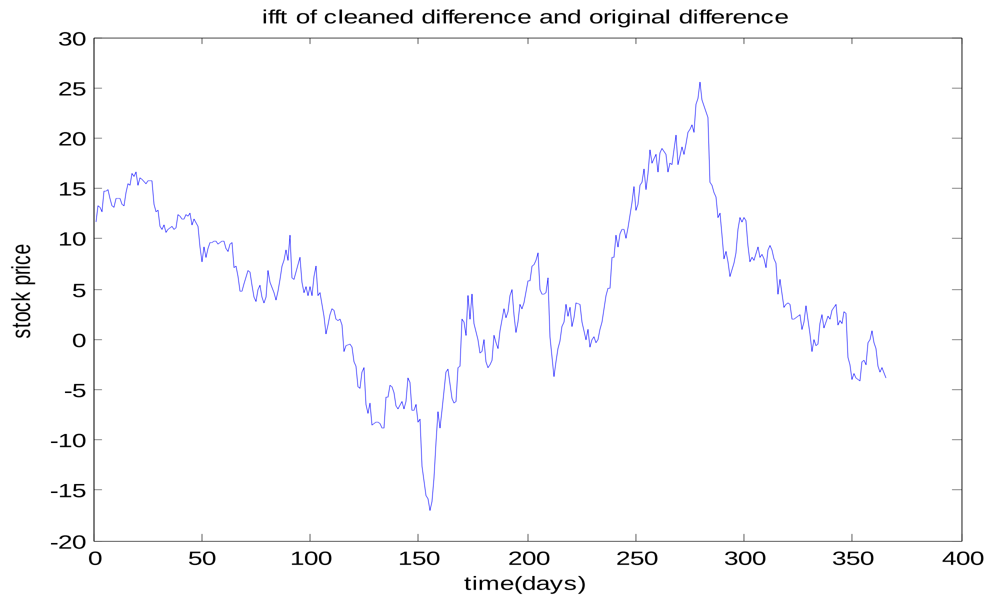


IFFT of the data points that is bigger than 10% of Y(1)

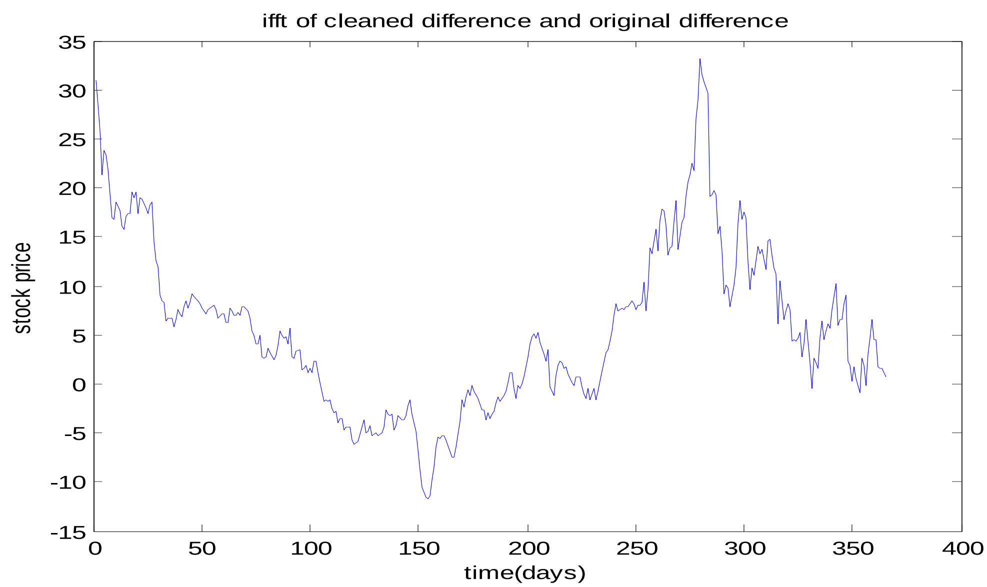


IFFT of the data points that is bigger than 15% of Y(1).

As we analyze the previous three graphs, we see that they all appear to look the same. This happens due to the fact that only first couple of points are significant, only about 10 to 15, then the rest of the points on the fft curve have a very small range compared to the first few points, therefore they are not included in the to 15%. After further analyzing, the percentage to include the rest of the points would need to be close to 100% due to their size compared to the first point of Y(1). In the next few lines of the code we take the difference and subtract it from the IFFT. The following graphs will show the difference between the best-fit subtracted from the original, and that will then be subtracted from the different types of IFFT to see which IFFT works best with each best-fit curve. We will only compute this for cleaned IFFT of 8 points, and cleaned IFFT of 10% of points.



IFFT (first 8 points) minus difference.

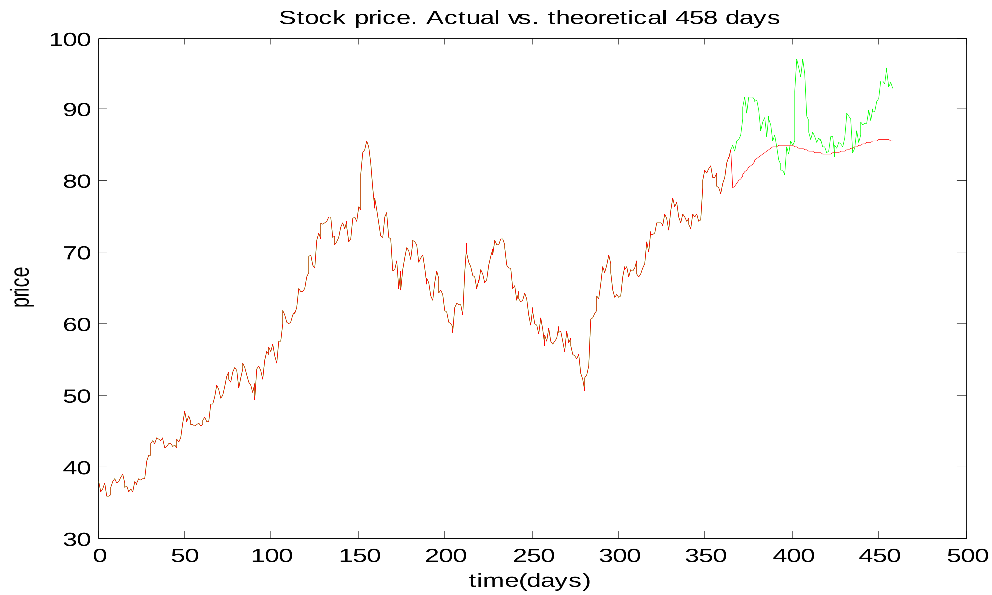


IFFT (10% of Y(1)) minus difference.

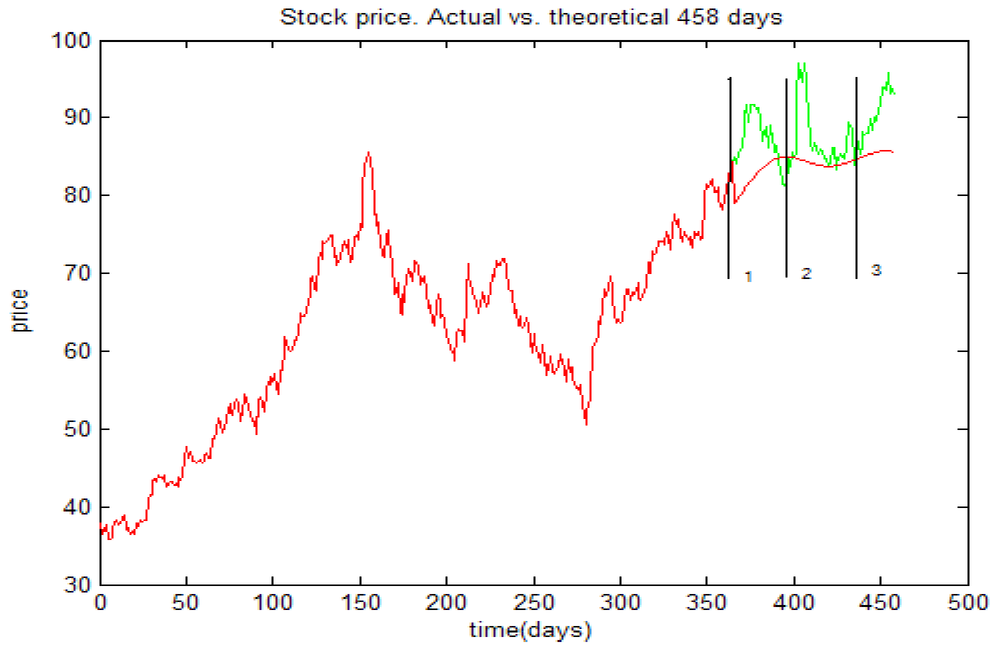
From the results above, we see that there are many ways to filter the data. We however will choose one of the best estimations and continue in our project of predicting the stock market. We will choose the FFT that only uses the first eight points. The next step that is run in the program is interpolation for the following 90 days. This is done by an equation that is often used to calculate the discrete Fourier transformation. The equation is in the form of:

```
a(k)=real(Y(k));
b(k)=-imag(Y(k));
omk=2*pi*(k-1)/365;
YY(n)=YY(n)+a(k)*cos(omk*(n-1))+b(k)*sin(omk*(n-1));
```

With this equation, we are able to interpolate the IFFT curve we obtained for the first 365 days and interpolate it for the next 90 days. Once we have the interpolation of the IFFT we add to the best fit curve line that we found in the first few lines of the program for the next 90 days. This together should yield us with a predication of how the stock will behave for the 90 days that we have predicted. The following graph shows our prediction of Apple Inc.



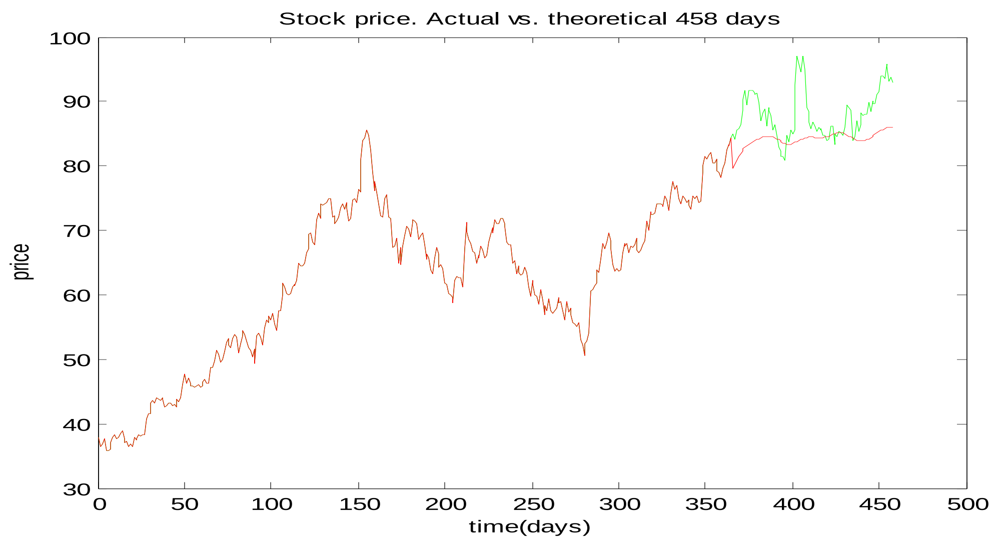
The next graph shows our prediction, however parts of the predicated stock price will be divided into 3 sections, and we will analyze each section separately.



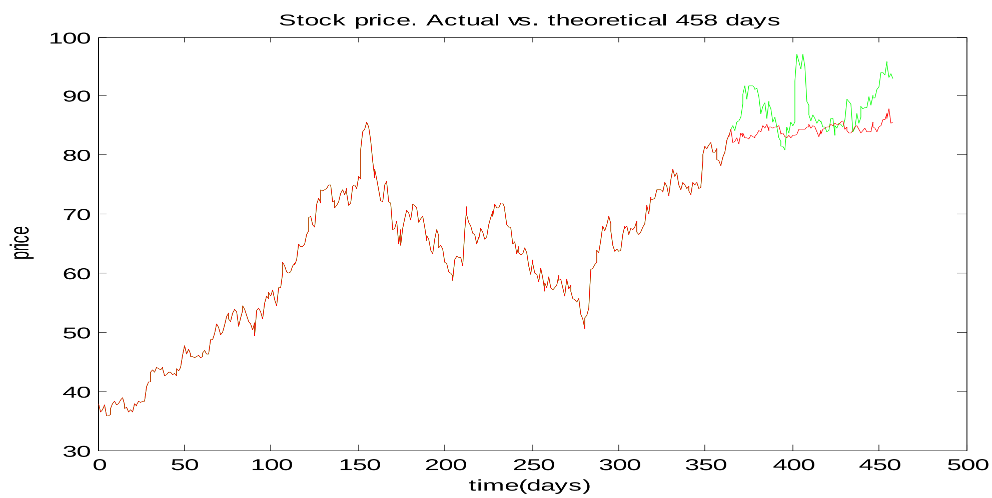
Our predicted price is color red after the first 365 days; the actual price has the color green after the first 365 days. As shown on the graph, we split predication into three sections to show advantages and disadvantages of our prediction. The first section labeled "1" predicts that the price of the stock should be rising up, however the actual price of the stock is rising only for the first couple of days; following those few days, the price of the stock begins to drop. The second section labeled "2" predicts that the stock should drop a few points, then gain a few points, but the actual price of the stock however had a major change, and the price stock rose rapidly; this could have been caused by many reasons (which will be discussed later on), which we could not have expected, therefore we could not have predicated. The third section labeled “3” predicts that the stock will be on the rise, and as we see the actual price of the stock did rise.

As expected, we could not precisely predict the price of the stock. Predicting the price of the stock can be very easily compared to predicting the weather. We could get a rough idea of how the stock or the weather will behave in the following few days or even weeks, we however cannot exactly predict the weather due to the unexpected occurrences. Some occurrences that could occur that could drastically affect the price of the stock are economic reason such as 9/11 or even the Federal Reserve lowering or raising the loan rates. There are also many reasons, not connected to our economical state why the price of the stock could sharply rise up or down, and some of those reasons include the company announcing quarterly gains or losses, as well as announcing new and improved products it will be releasing.

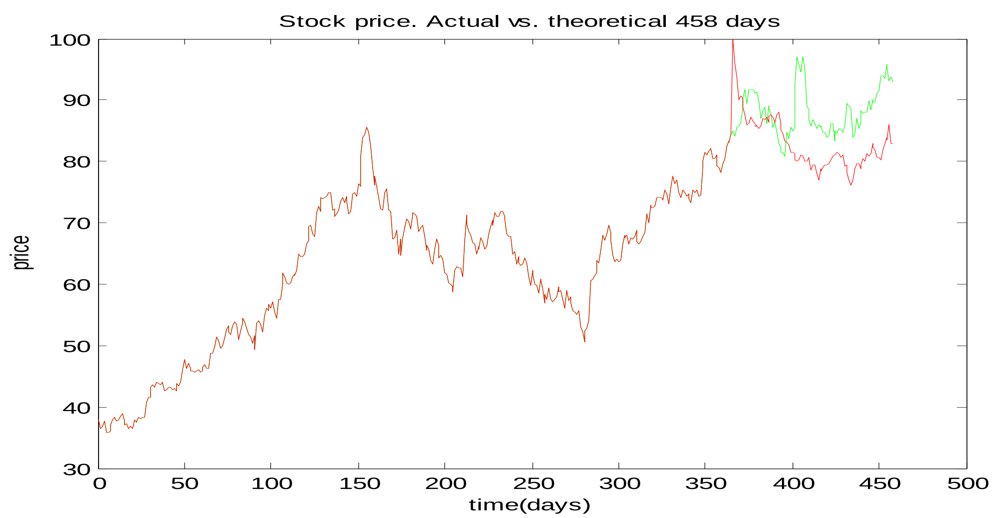
The following few graphs will show how the stock would be predicted had we used different cleaning techniques of the FFT.



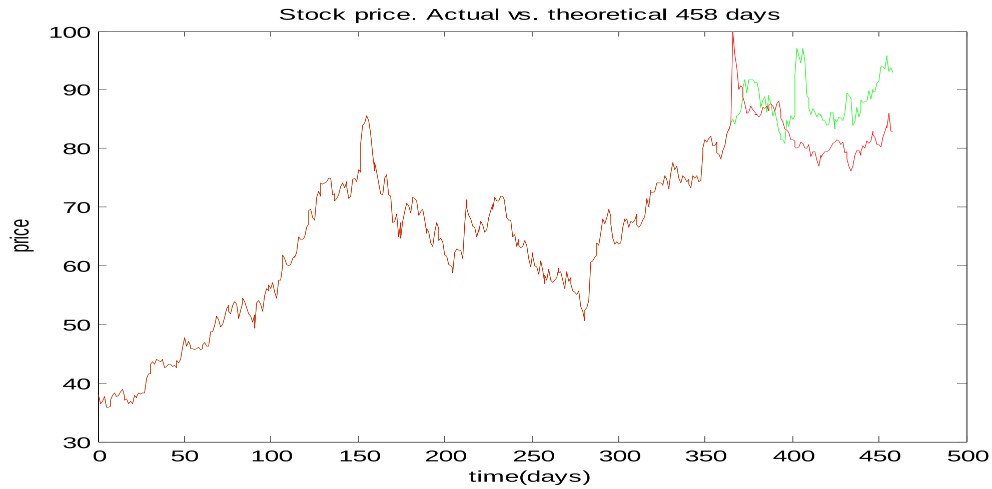
Prediction when first 20 points of FFT were used.



Prediction when first 200 points are used.

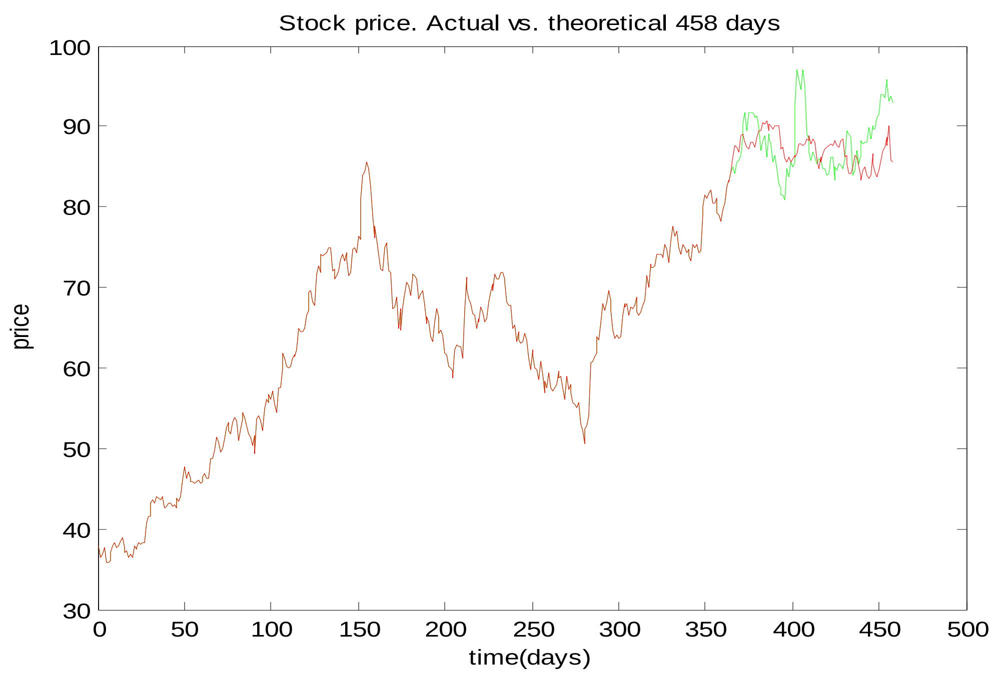


Prediction when 5% of points are used.



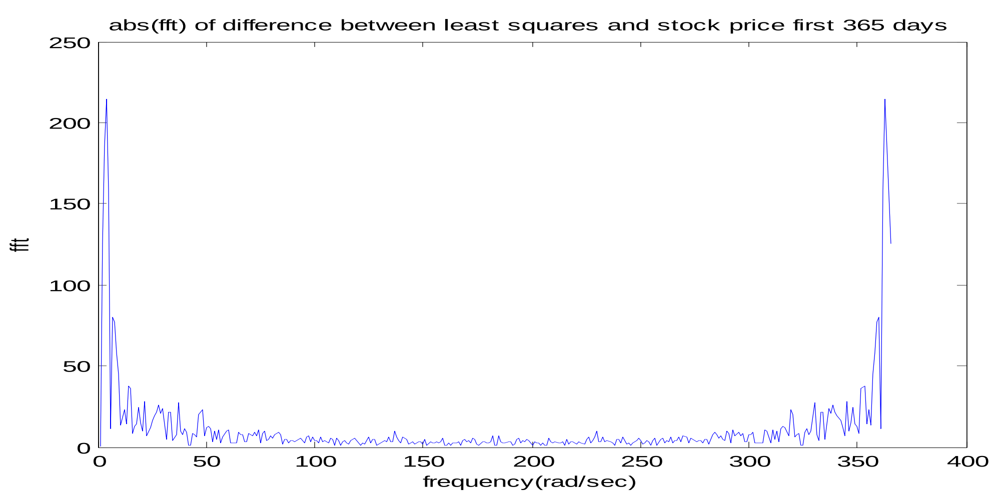
Prediction when 10% of points are used.

As we see from all the predictions we have gathered, the best predictions come when we select a certain amount of points instead of percentage. The next graph will show how our prediction will look if we don't filter out any points.

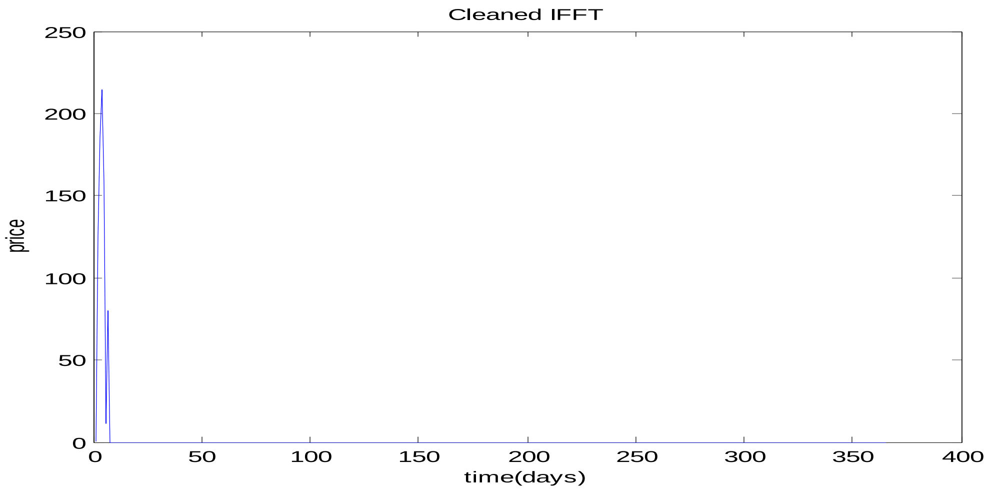


Now that we have explained the program, and the different variables that we could change in each section of the program, we will show another stock, Audible Inc., to show how the program reacts differently for different stocks. With this stock we did previous analysis and have chosen the best options for each step of the program. This is what the original stock price with the best-fit curve:

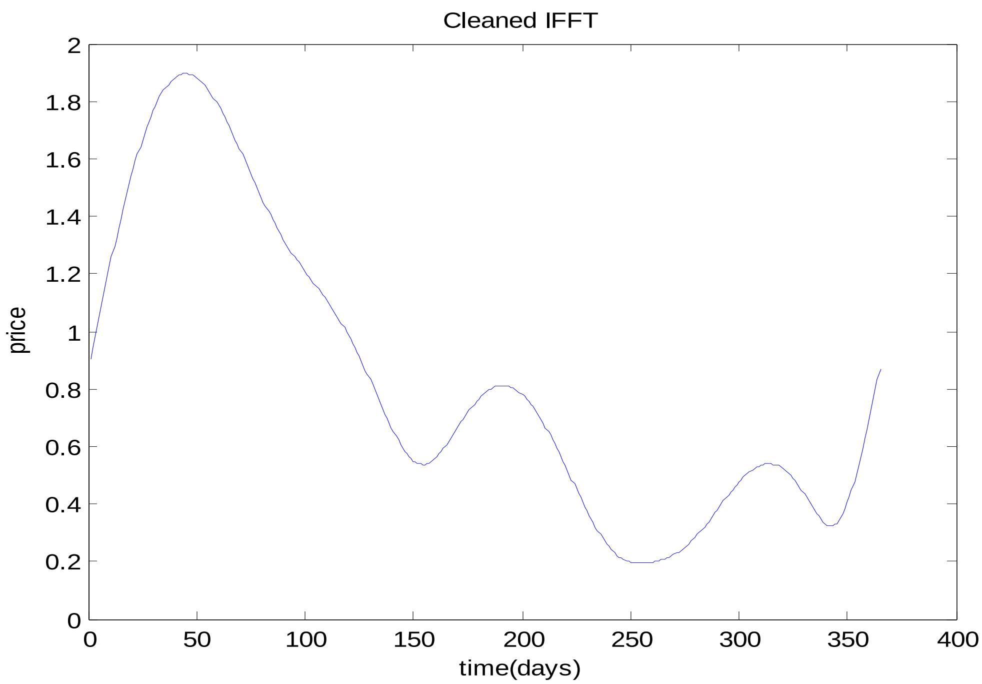
This is the FFT graph for all the points:



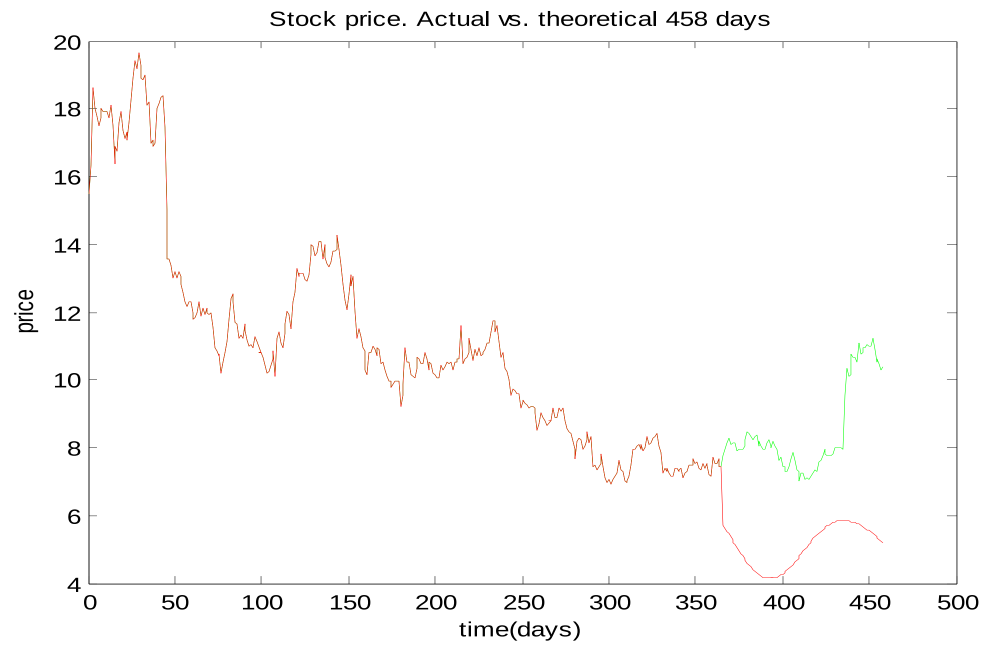
This is the cleaned FFT graph with only 8 points used:



This is the cleaned IFFT graph with only 8 points used:



This is the predicted stock price versus actual price:



As we could see from the previous graphs, our prediction for Audible Inc. is not as precise as it was for Apple Inc. We are unsure why this happens, but from the beginning we expected that one certain program would not be able to predict every stock. As we could see there is a major jump just as we start to predict our stock price. This happens for other stocks as well (other stocks that have the same problem are Bank of America Inc., Google Inc., and Exxon Mobil Inc.). For most of these stocks however, the jump is always proportional to the stock price. For example for Audible Inc. the stock dropped 2 points instantly, which is about 10% of the scale, this exact problem occurs with other stocks, once the prediction is started the stocks instantly jumps down 10% of the scale.

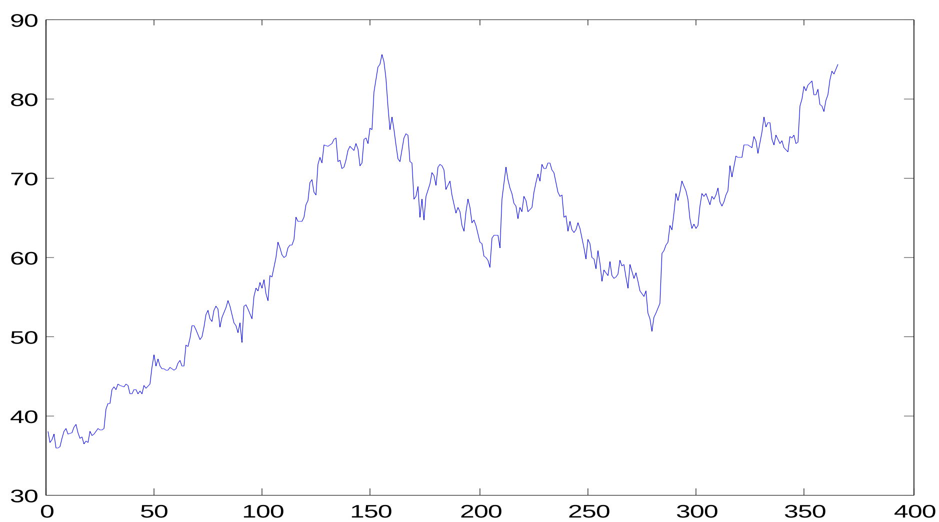
Explanation of Program 2

We will now analyze program 2 in order to see if we could get a more correct prediction. As you will see, program 2 is very similar to program 1, with few minor changes.

Program 2 starts by:

```
y=dlmread('Apple1YearData.txt');
x=dlmread('365.txt');
L=length(x);
figure(1)
plot(x,y); hold on;
```

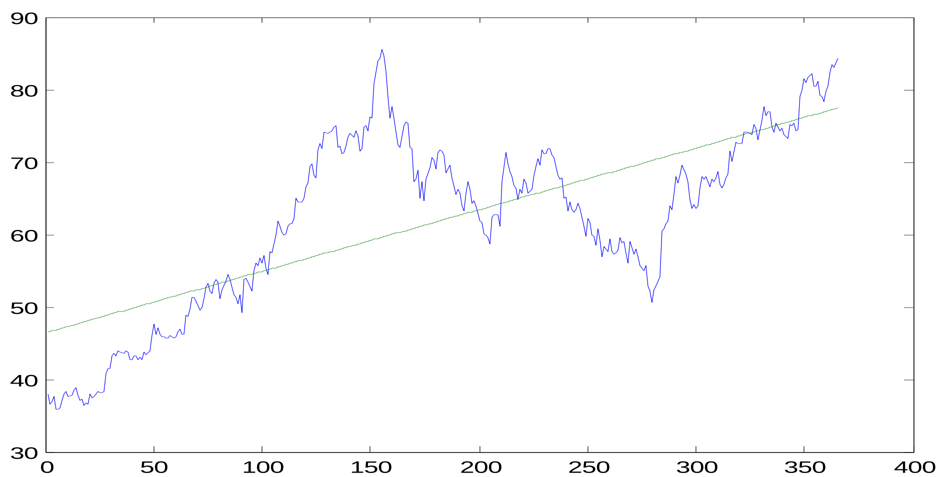
This code basically loads the price of the stock, and sets the time of days to 365. The price of the stock is on the y-axis while the time in days is on the x-axis. Once the data is loaded it plots the price versus time. This plot contains the price of the Apple stock for 365 days, the dates of this stock price is from June 1st 2006 to June 1st 2007. The plot can be seen below:



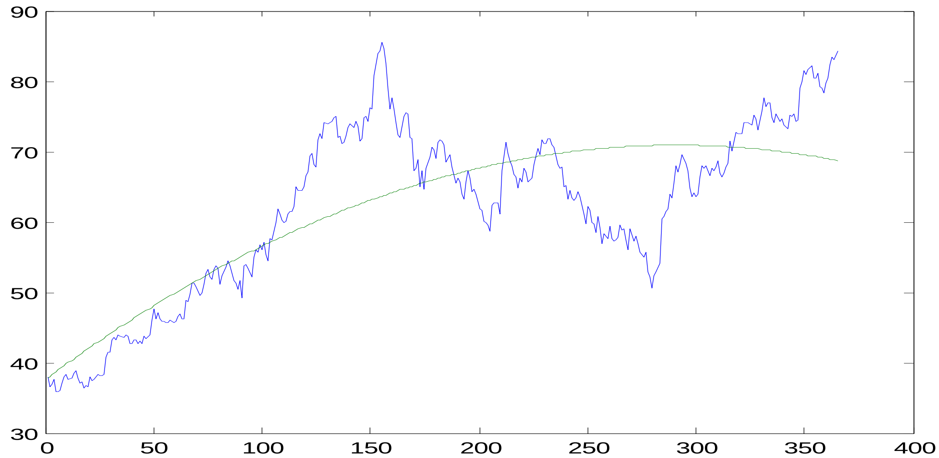
The next piece of code consists of:

```
%% finding the trend
p1=polyfit(x,y,1)
xx=(1:1:365)';
pp1=polyval(p1,xx);
plot(x,y,xx,pp1)
```

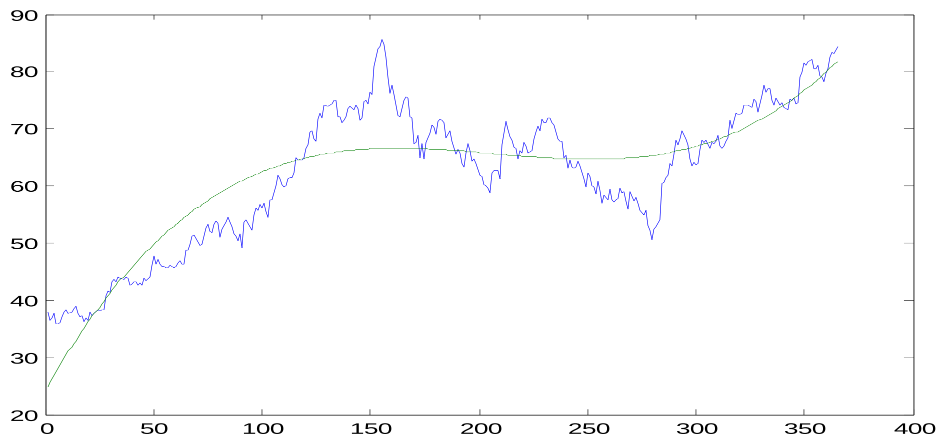
This piece of code works by taking the original data and using the least square method it finds the best-fit for the data using the first degree. The following three graphs will show the original data with first, second and third degree least squares:



Best-fit curve using first degree polynomial y=ax+b.



Best fit curve using second degree polynomial y=ax^2 + bx + c.

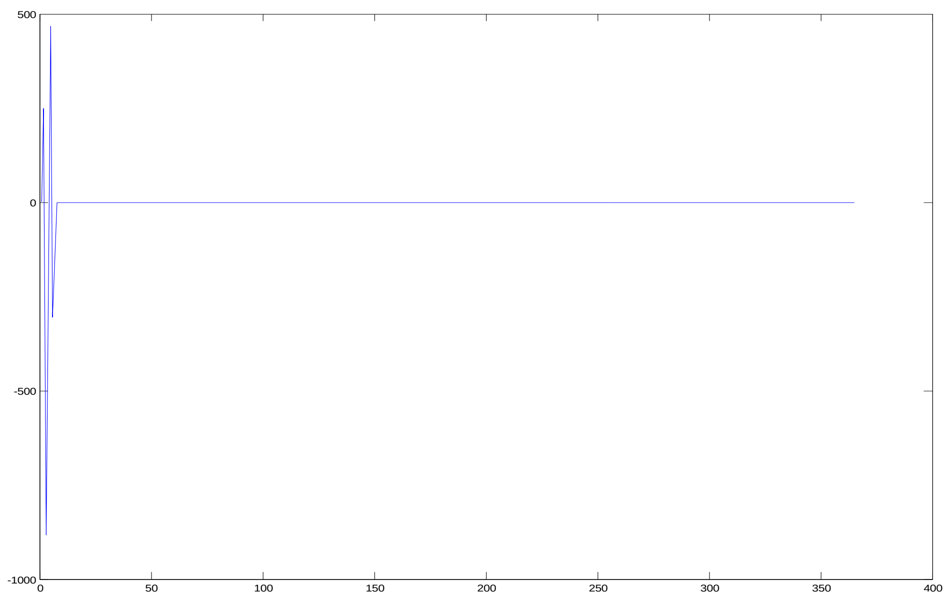


Best fit curve using third degree polynomial y=ax^3 + bx^2 +c.

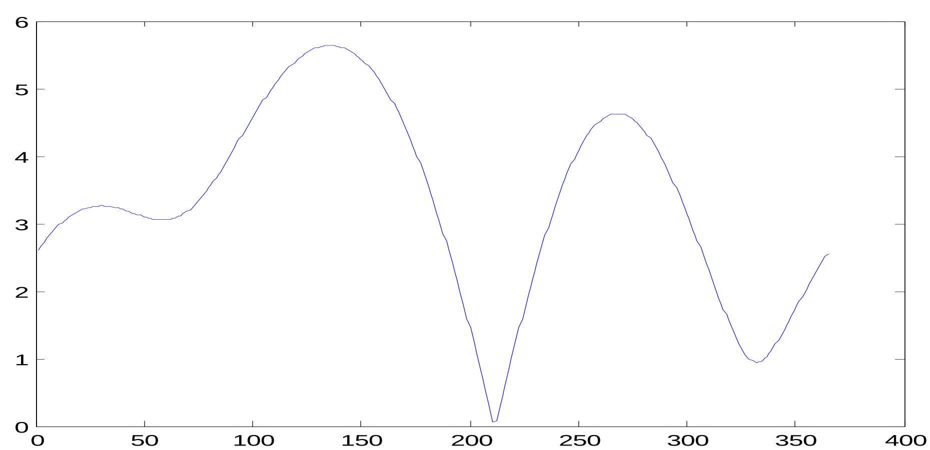
The next piece of code is:

```
difference=pp1-y;
Y=fft(difference);
for k=8:365,
      Y1=Y;
      Y(k)=0;
end
P=ifft(Y);
figure(2)
plot(abs(P))
```

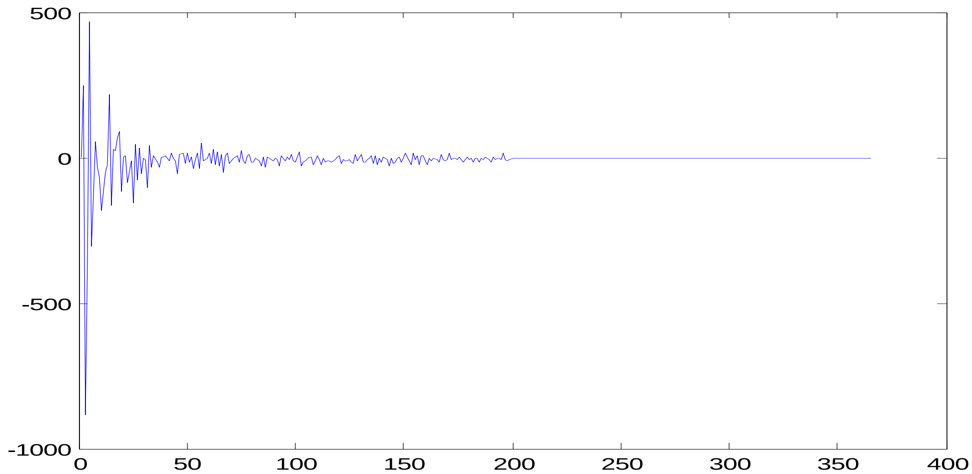
This piece of code takes the difference between the best-fit curve and the original data points. Once the difference is computed it takes a Fast Fourier transformation of the first eight points and sets the rest of the points to zero. In then takes the inverse of these eight points and graphs the plot. The following few charts will explain how setting different number of points to zero will affect the IFFT.



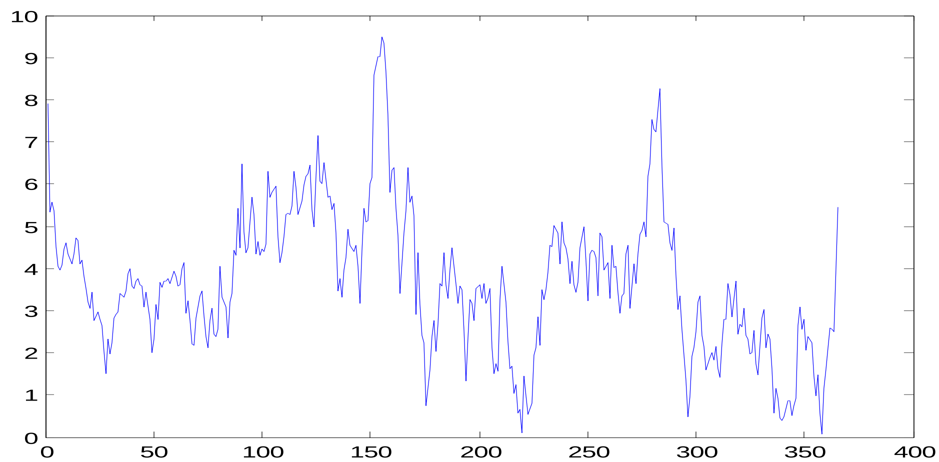
Graph shows the FFT for first 8 points. The remainders of points were set to 0;



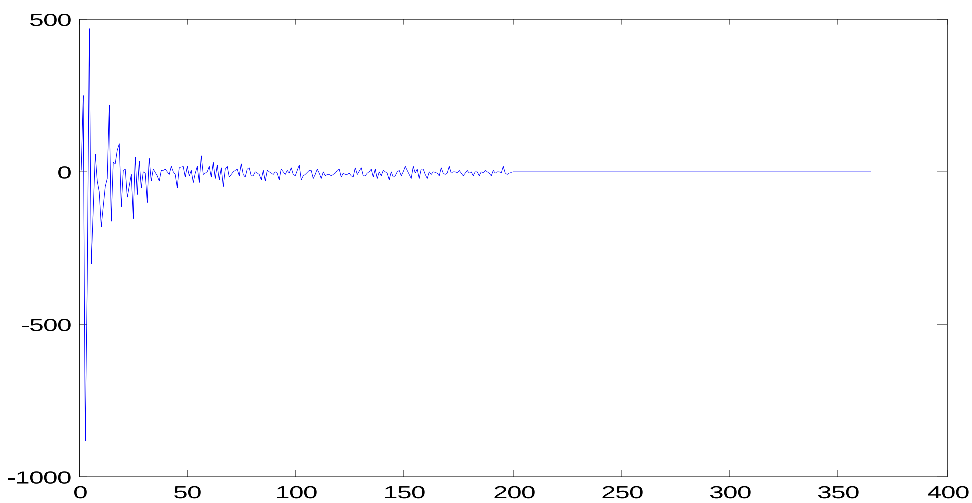
This graph shows the IFFT of the first 8 points.



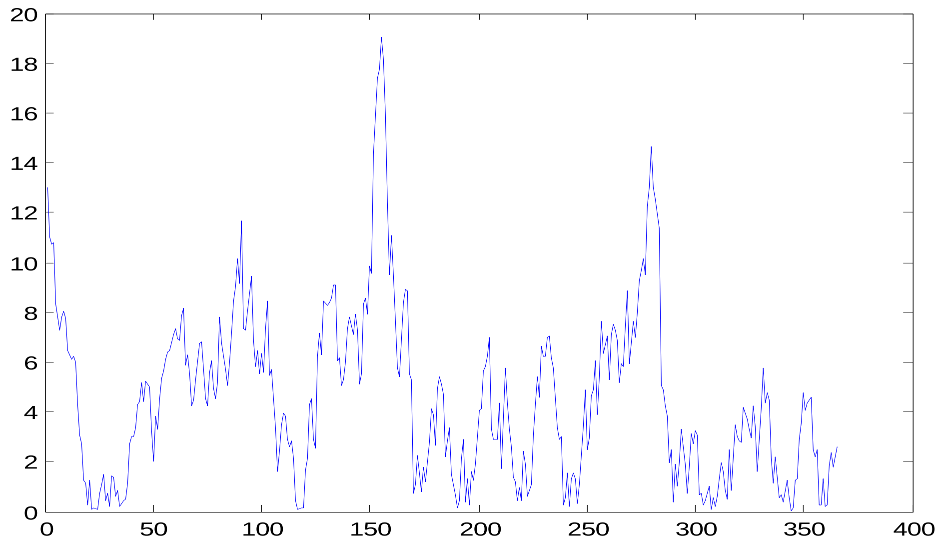
This graph shows the FFT for the first 200 points with remainder of the points set to zero.



This graph shows the IFFT for the first 200 points.

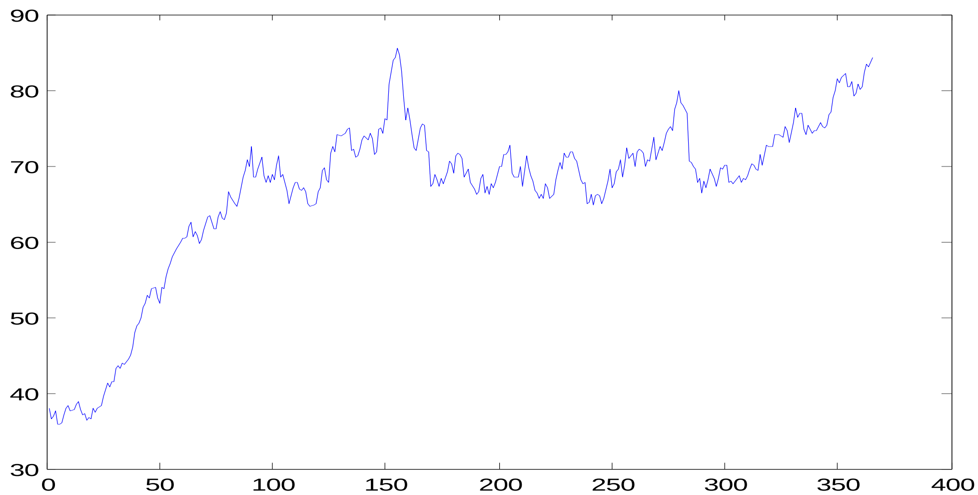


This graph shows the FFT for the first 365 points.



This graph shows the IFFT for the first 365 points.

As explained earlier, depending on how many points you decide to use for the FFT will result in a different IFFT. As we could see in the graphs above, if we only use 8 points in the FFT the IFFT will only have eight points as well. When these eight points are added to the best curve fit, we should get a line that is fairly close to the original data points. The more points we use in our FFT, the closer our IFFT line added to the best-fit curve will be to the original data points. When all 365 points are used in the FFT and they are added to the best-fit curve, we would have a line that is identical to our original data points. This can be seen in the graph below

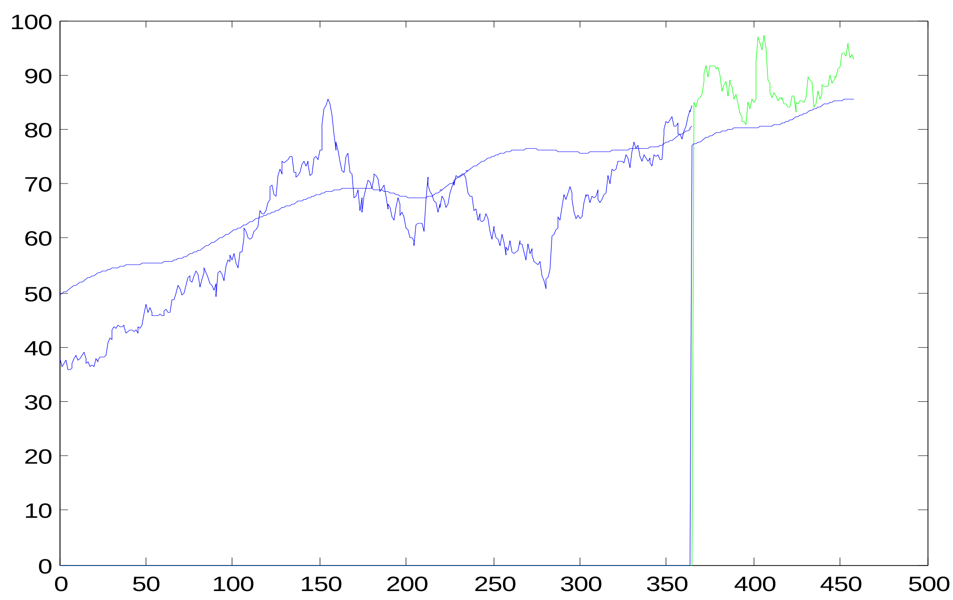


The following piece of code was used next:

```
for k=365:458,
        t(k)= p1(1)*k + p1(2);
end
for n=365:458
        sum=0;
for k=1:7;
        a(n)=real(Y1(k));
        b(n)=-imag(Y1(k));
        comp_k=(a(n)*cos((2*pi*(k-1)*(n-1))/L) +
                b(n)*sin((2*pi*(k-1)*(n-1))/L));
        sum=(sum+comp_k)/L;
end

g(n)=(sum);
figure(3)
plot(g); hold on;
end
NewTrend2 =  t + g;
monthdata=dlmread('Apple3MonthData.txt');
figure(1)
plot(monthdata, 'g');
plot(NewTrend2)
```

This piece of code interpolates both the best-fit curve and the FFT from day 365 to day 458. In order to interpolate the best-fit curve, we found the equation of the line and extended it by substituting the day number into the equation. Next we interpolated the IFFT curve; this once again was performed in a similar fashion as in the first program by using the standard equation for finding the Discrete Fourier Transformation. Once we found the interpolation of both lines, we added them together in hopes of finding a trend for the stock price for the following 90 days. The trend appears as:



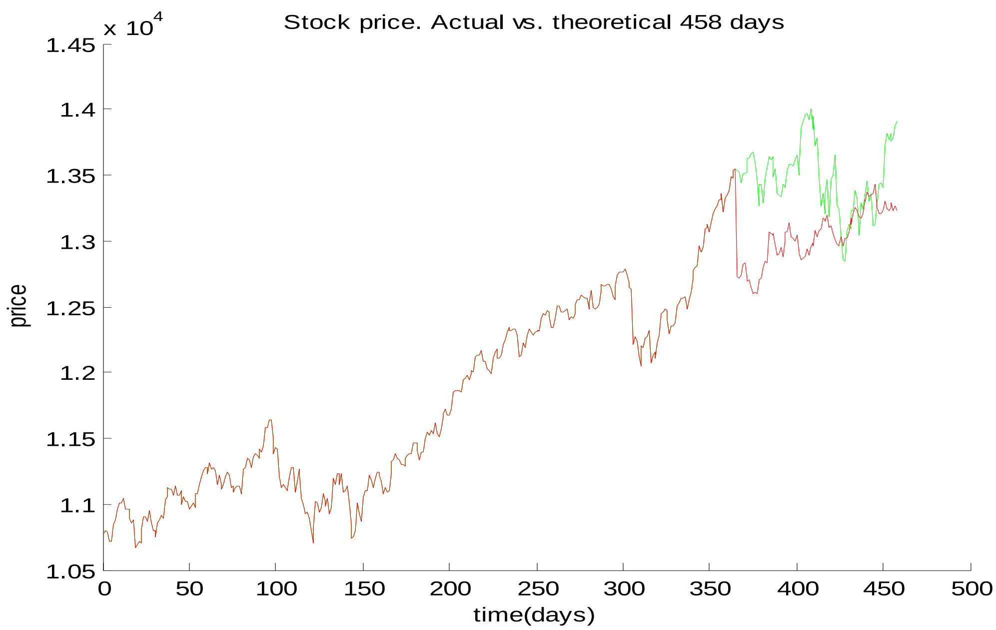
In this graph we see that we cannot make a good prediction due to the fact that the interpolated IFFT is very small. The prediction we see above is mainly the best-fit curve interpolated. From that graph we could see that this code does not give us a good prediction for this stock. Due to the reason of having such a small interpolated IFFT for other stocks as well, this piece of code will not give us an accurate prediction for any stock price.

Comparison of Programs

After careful analysis of the two programs, program 1 seems to work better with the five stocks I have chosen. The main difference in the two programs seems to be the interpolation of the FFT curve. In program 1, the interpolation of the FFT curve depends on how many points we choose from the cleaning of the FFT, despite the FFT being a small component of the interpolation it does play a major role in showing whether the stock will rise or drop. In program number 2, however when we decide to clean the program, the interpolation always seems to be in the same sequence of rising and dropping. This is very critical part of the prediction, and if there is an error of this sort, we can predict the price range, but we cannot see when the price fluctuates up or down. Another major difference between the two programs seems to occur at the very start of the prediction. Program 1 seems to always drop by the 10% mentioned earlier, program 2 however drops a different range for each of the stocks. Again, since we cannot predict how much it will drop for each stock, we cannot get an accurate prediction of how that compares to the actual price.

Further Analysis of Program 1

Since we have analyzed each of the stocks in program 1, we will now see whether the stock fluctuates according to the Dow Jones Industrial. I will run the data for Dow Jones Industrial in Program 1, and then make a correlation analysis between each of the five stocks and Dow Jones. The Jones Industrial prediction looks as following:



As we see here, there is a drop of 10% almost instantly as we start the prediction. But if we look past that drop, the prediction we calculated is fairly accurate. Now that we have the prediction of the stock, and the prediction of the Dow Jones, we ran a code, which found the correlation between the two predictions. The correlations are as follow:

1. Apple Inc.
    - Regular Program: 0.7024%
    - 15% of Points Used: 0.6755%
    - First 12 Points Used: 0.7036%
2. Bank of America Inc.
    - Regular Program: 0.8638%
    - 15% of Points Used: 0.8751%
    - First 12 Points Used: 0.8884%
3. Exxon Mobil Inc.
    - Regular Program: 0.7999%
    - 15% of Points Used: 0.7951%
    - First 12 Points Used: 0.8343%
4. Audible Inc.
    - Regular Program: 0.8400%
    - 15% of Points Used: 0.8397%
    - First 12 Points Used: 0.9759%
5. Google Inc.
    - Regular Program: 0.7428%
    - 15% of Points Used: 0.7390%
    - First 12 Points Used: 0.7410%

From the above correlation coefficients, we see that the predictions of the stocks, and Dow Jones have a very similar correlation. This shows us that the price of the stock fluctuates according to Dow Jones.

Programs 1 As Document

```
%% loading data
load audibleinc.dat.txt
%% finding the trend
for i=1:365
x(i)=i;
end
p1=polyfit(x',audibleinc(1:365),1)
xx=(1:1:458)';
pp1=polyval(p1,xx);
plot(xx,audibleinc)
hold
plot(xx,pp1)
xlabel('time(days)')
ylabel('stock price')
title('least squares vs. stock price')
print -dpsc2 -r600 audibleinc.ps
hold
%% difference between data and trend for the first 365 days
diff=audibleinc(1:365)-pp1(1:365);
plot(x',diff)
xlabel('time(days)')
ylabel('stock price')
title('difference between least squares and stock price
first 365 days')
print -dpsc2 -r600 -append audibleinc.ps
Y=fft(diff);
plot(abs(Y))
xlabel('frequency(rad/sec)')
ylabel('fft')
title('abs(fft) of difference between least squares and
stock price first 365 days')
print -dpsc2 -r600 -append audibleinc.ps
PP=ifft(Y);
plot(PP)
plot(PP-diff)
xlabel('time(days)')
ylabel('stock price')
title('ifft of cleaned difference and original difference')
print -dpsc2 -r600 -append audibleinc.ps
%interpolate the difference for 90days
for n=366:458
YY(n)=0;
for k=1:365
a(k)=real(Y(k));
b(k)=-imag(Y(k));
```

```
omk=2*pi*(k-1)/365;
YY(n)=YY(n)+a(k)*cos(omk*(n-1))+b(k)*sin(omk*(n-1));
end
YY(n)=-YY(n)/365;
end
% create one array for the original and interpolated
difference
% between the least squares curve and the stock price.
for i=1:365
YY(i)=diff(i);
end
%total price of stock from the least squares curve and
cleaned fft
%for 458 days
tot =  pp1 + YY';
%plot the stock price (green) for 458 day vs the
theoretical curve(red)
plot(xx,audibleinc,'g');
hold on
plot(xx,tot,'r')
hold off
xlabel('time(days)')
ylabel(' price')
title('Stock price. Actual vs. theoretical 458 days')
print -dpsc2 -r600 -append audibleinc.ps
%plot the difference for the extrapolated values actual-
theoretical
plot(xx(366:458),audibleinc(366:458)-tot(366:458))
xlabel('time(days)')
ylabel(' price')
title('Difference between actual stock price and
theoretical prediction.')
print -dpsc2 -r600 -append audibleinc.ps
```

Program 2 As Document

```
y=dlmread('BankofAmerica1YearData.txt');
x=dlmread('365.txt');
L=length(x);
figure(1)
plot(x,y); hold on;
%% finding the trend
p1=polyfit(x,y,2)
xx=(1:1:365)';
pp1=polyval(p1,xx);
[m,n]=size(y)
[q,w]=size(pp1)
%% difference between trend and data
difference=pp1-y;
Y=fft(difference);
[x,z]=size(Y)
%% cleaning noise out of data
for k=8:365,
    Y1=Y;
    Y(k)=0;
end
P=ifft(Y);
%% clean ifft plus trend
NewTrend=(abs(P)) + pp1;
figure(1)
plot(NewTrend); hold on;
%% interpolating data for 3 months
for k=366:458,
    t(k)= 0.0001*k^2 - 0.0212*k + 44.9922;
end
%% to interpolate FFT curve
for n=366:458
    sum=0;
    for k=1:7;
        a(k)=real(Y1(k));
        b(k)=-imag(Y1(k));
        comp_k=(a(k)*cos((2*pi*(k-1)*(n-1))/L) +
b(k)*sin((2*pi*(k-1)*(n-1))/L));
        sum=(sum+comp_k)/L;
    end

 g(n)=sum;
end
figure(3)
plot(abs(g),'r')
%% predicting 3 month data
```

```
NewTrend2 =  t + g;
monthdata=dlmread('BankofAmerica3MonthData.txt');
figure(1)
plot(monthdata, 'g');
plot(NewTrend2)
```

Privatization of Social Security Paper

Social Security was started in 1935 and begin as a funded retirement plan, meaning that its’ benefits would be paid out of a fund build over the years from contributions made by Social Security taxpayers. This has worked for many decades, but today a new problem arises. Once the baby boomers retire, many Social Security experts believe there will not be enough money in the Social Security fund to pay everyone the pension which they deserve. This causes for great deal of concern, and new measures are being put into place in order to prevent this disaster.

In the past many precautions have been put into place, in order to ensure the Social Security will not be depleted in the near future. One of the major precautions taken was raising the percentage of taxes deducted from workers’ paychecks. Another major set-back today is the low retirement age, a retiree could begin receiving benefits at an early age of sixty-two, so the other major precaution taken was to raise the retirement age to sixty-seven by the year 2022.

According to many experts, raising the retirement age will not provide us with a clear solution; many have suggested implementing higher taxes for workers, but according to Social Security experts the tax would need to be raised from 12.4% to more than 19% in 2075 in order to pay for the baby boomers pensions. This idea would not be adapted well by the individuals paying the taxes, and would cause major concern among the taxpayers. The previous solutions mentioned would offer only a temporary solution, these solutions might help to pay the pensions of the baby boomers, but there needs to be a solution adapted which would not cause these types of problems in the far future. One main idea that many have proposed is privatizing Social Security, there are many different possibilities of how this could be achieved, and a few of the popular suggestions will be discussed.

Privatization of Social Security would work on the concept that tax money would be invested in corporate stocks or mutual funds. There are many different options to privatize the Social Security; there would be options, one of them is for the government to choose where the money is invested, second option is for the tax payer to choose where their money is invested, or a third option of a combination of both. Also there could be an option of choosing how much of the tax money you wish to invest, for example a taxpayer might not want to invest their whole retirement pension on a stock market which they are not familiar with.

Another problem with privatizing the Social Security deals with the disabled. Currently, one-third of the total Social Security benefits go to the disabled. The proposed privatization would force the disabled to invest the small income they already receive, which I'm sure would cause some concern from the disabled.

Social Security privatization also has an unsure result for women. Women tend to earn less and live, on average, six years longer than men. Currently only 13 percent of women even qualify for private pensions. On the other hand, their pension would stretch over a longer time period therefore resulting in a less per month pension than the same pension held by a man.

For three years, from 1994 – 1997, the Advisory Council on Social Security worked on a solution to the upcoming problem. They however could not find one clear-cut solution, but rather a set of three solutions with each having its pros and cons.

The first option proposed, and the most conservative option of the three is to sustain Social Security as a government organized pay-as-you-go system with only minor changes. The minor changes would occur by adding cutbacks in retirees' pension rates, and a slight increase in tax. The tax increase would occur somewhere around 2045 and it would increase by about 1.6 percent. The government would also invest part of the pensions in corporate stocks chosen by their own financial advisors. This first option would only be beneficial only if the stock market continues to grow as it did in the past. The problem now arises due to the risk involved. As we look back to September 11, the stock market took a big drop following the attacks, and if a similar event occurred, many pensions would be left with very small retirement funds. This would mean that an individual's retirement account would depend on the state of the stock market, and many workers feel the risk is too great. According to a survey done by The Century Foundation, depending on when a person will retire, their pension could drastically differ. The following chart shows how much pension the person would receive at different years of retirement if they had invested in the stock market:

It is almost too risky to take that kind of a chance. For example, a worker who would invest in a stock portfolio that matched the S&P 500 in 2000 and cashed out for retirement in March of 2000 would accumulate nearly a third larger retirement pension than someone who retired a year later, with both retirees using the same investment method. Other major setbacks that could occur with this option are with the government investing our money; there would be room for corruption within Washington. As Alan Greenspan pointed out on January 20th 1999, "if the government owned lots of stock, it would sooner or later be tempted to invest with political rather than economic aims in mind. That would squander the nation's capital."

The second option proposed by the council reduces the substitution rate for upper-income workers and raises the retirement age. This option would also allow workers to create individual investing accounts, which would be overseen by the Social Security system. This option will give the individual control over their own pensions and would allow for higher returns if the correct stocks or mutual funds are chosen. This option however raises concerns; what if one individual chose the proper stocks to invest in; he would have a pension far greater than the individual who chose the stocks that did not perform well. Also what if the individual chose stocks that would ultimately go bankrupt, would that mean that the individual would not have a pension to retire on? All of these factors have to be taken into consideration.

The last option proposed was partially investing the Social Security of each individual. With this option, five percent of Social Security tax would be placed in a trust fund chosen by the government, and the other five percent would be placed in personal investment account which would be managed privately through investment companies. Once again, the problem of risk comes into play.

All of the above options give a solution to the problem we may have in the near future, but one can argue that none are good solutions. Many other aspects arise, such as the transition period. An individual who would retiree today would not have the same possibilities for a pension as a person who would retire in a year, or even thirty years. They would not have sufficient time in order to build their pensions, and since the Social Security would be invested, it would cause the retiree to retire without the proper pension. Any of the three options could however work if the government would somehow compensate the soon to be retirees'.

Social Security interest groups, as well as a current Presidential candidate Barrack Obama, believe that privatization of social security is too drastic of a measure to take, when Social Security is the most public program that we have implemented. According to analysts, the trust fund has saved funds which will last for Social Security for at least another thirty-five years. It is clear that the baby-boomers will cause a few problems, but we will have the next thirty-five years to think of a well thought of solution, and that's why analysts wonder why this sudden need of change to the Social Security is necessary.

I believe that the government will be extremely hesitant of privatization after the September 11th terrorist attacks. Nobody wants to risk their retirement pension on a market that is so unpredictable in its returns.

There are no clear-cut ways of improving Social Security; each concept has its own pros and cons; however one addition that could be made to the Social Security is raising the retirement age. This would provide that each individual works for a longer period of time, therefore saving more for their Social Security, and would decrease the amount of social security paid by a few years, which would cause a great increase in the Social Security funds.

Conclusion

Predicting the stock market is very difficult. As we saw in the report many different variables need to be considered in order to make a good prediction. Those variables range from small changes in the program, as well unpredictable occurrences in the ever day economy.

When first started we chose five random stocks to analyze. The stocks that were chosen were Apple Inc., Audible Inc., Bank of America Inc., Exxon Mobil Inc., and Google Inc., and each of these stocks behaved in a different manner when the prediction was being made. Some of the reasons for these stocks behaving differently were the vast range of the prices; for example Audible Inc. stock price ranged in the 10-20 point level, while Google Inc., ranged from 300-600 level. Another major effect on the stocks was abrupt jumps in the price. For example during a 2-3 day period Apple Inc. jumped more than ten points, while Audible Inc. dropped more than five points; which was a major shift when comparing the jump to the actual price of the stock.

The stock we spent the most time analyzing was Apple Inc. The programs that were created were actually based on the prediction of this stock. Once we analyzed this stock, we gathered many predictions that were very precise. Depending on how we cleaned the data for our prediction, we were able to predict the stock anywhere from a week to two weeks, which a very long and accurate prediction. The next few stocks had many things in common. All other four stocks mentioned earlier behaved somewhat strangely at the instant the prediction was being made. When we attempted to make our prediction, the stock would drop about 10% of the scale and then continue with a regular prediction. Once we adjusted that 10% jump, the prediction was also fairly accurate; not as accurate as Apple Inc. For these stocks the amount of time we could predict ranged drastically. When making a prediction for the 90 days, some made an accurate prediction at the beginning, but some did not until about half our predicted time. After careful analysis of each of the stocks, we compared the predicted price to the Dow Jones predicted price, and we did surprisingly have a very close correlation between the two graphs, which shows that our prediction was on target, however as stated earlier some unpredictable occurrences could have occurred that we could not had expected.

As we see, it is very difficult to predict every stock with the same algorithms. We however made that possible with small changes stated earlier in the report. For me this is a very exciting topic, as well as very interesting due to all the benefits it carries with itself. If we could accurately predict the stock market every time, we and many other would instantly take advantage of the program written and buy stocks that would make us instantly wealthy. That however is nearly impossible; with our program, we could somewhat predict the market in an ideal situation, but as the past few years in the American economy shows, nothing is impossible, and things can change drastically with every minute.

Bibliography

Aaron, Henry, Martin Feldstein and Margaret Warner. Investing for America's Future. Debate. 3 March 1999. <http://www.pbs.org/newshour/bb/health/jan-june99/ Socsec_3-3.html>

Anrig, Greg Jr. and Bernard Wason. 12 Reasons Why Privatizing is a Bad Idea. The Century Foundation. 14 December 2004. <http://www.socsec.org/publications.asp?pubid=503>

Baker, Dean. Investing the Social Security Trust Fund in the Stock Market Doesn't Help. Beat the Press Archive, The American Prospect. 1 November, 2007

Farmer, Doyne, Martin Shubik, and Eric Smith. 'Emmerging Body of Workby Physicists Addressing Questions Of Economic Organization and Function Suggests Approaches to Economic and a Broadening of the Scope of Physics.' Physics Today. September 2005: 37-42.

Francis, David. Privatized Social Security - Take 2. The Christian Science Monitor. 27 June 2005. <http://www.csmonitor.com/2005/0627/p17s01-cogn.html>

Greenspan, Alan. Greenspan Likes Social Security Private Accounts, But Urges Caution. SeniorJournal.com. 17 February 2005. <http://seniorjournal.com/NEWS/SocialSecurity/5-02-16GreenspanSays.htm>

Matthews, Merril. Some Americans Already Have Privatized Social Security. National Center for Policy Analysis. 4 November 1996. <http://www.ncpa.org/ba/ba215.html

Newman, Nathan. Social Security: Why Government Investing in Stocks Beat Personal Accounts. Online. January 12, 2005. <http://www.nathannewman.org/log/archives/002064.shtml>

Paulus, John Allen. A Mathematician Plays The Stock Market.

New York: Basic Books, 2003.

Shipman, William. Private Social Security Accounts are Best Long-Term Investment Despite Market Jitters. New York Post. 14 August, 2002.

Acknowledgements

Special Thanks to Everyone Who Contributed To This Project:

Professor Mayer Humi

Adrianna Hera

Christine Drew