FORECASTING THE S&P 500 INDEX USING TIME SERIES ANALYSIS AND SIMULATION METHODS

By

Eric Glenn Chan

Submitted to the MIT Sloan School of Management and the School of Engineering
In partial fulfillment of the requirements for the Degree of

Master of Science in Engineering and Management
In conjunction with the SDM Fellows Program at the

Massachusetts Institute of Technology
June 2009

Signature of Author____

Eric Glenn Chan, SDM Fellow
MIT Sloan School of Management
MIT School of Engineering
June 2009

Certified by__________________________

Roy E. Welsch, Thesis Advisor
Professor of Statistics and Management Science and Engineering Systems
MIT Sloan School of Management

Accepted by____

Patrick Hale
Director, SDM Program
MIT School of Engineering

FORECASTING THE S&P 500 INDEX USING TIME SERIES ANALYSIS AND SIMULATION METHODS

By

Eric Glenn Chan

Submitted to the MIT Sloan School of Management and the School of Engineering
In partial fulfillment of the requirements for the Degree of

Master of Science in Engineering and Management

ABSTRACT

The S&P 500 represents a diverse pool of securities in addition to Large Caps. A range of audiences are interested in the S&P 500 forecasts including investors, speculators, economists, government and researchers. The primary objective is to attempt to provide an accurate 3 month and 12 month forecast using the recent credit crisis data, specifically during the time range of 10/2008 – 09/2009. Several methods were used for prediction fit including: Linear Regression, Time Series Models: Autoregressive Integrated Moving Averages (ARIMA), Double Exponential Smoothing, Neural Networks, GARCH, and Bootstrapping Simulations.

The criteria to evaluate forecasts were the following metrics for the evaluation range: Root Mean Square Error (RMSE), Absolute Error (MAE), Akaike information criterion (AIC) and Schwartz Bayesian criterion (SBC). But most importantly, the primary forecasting measure includes MAE and Mean Absolute Percentage Error (MAPE), which uses the forecasted value and the actual S&P 500 level as input parameters.

S&P 500 empirical results indicate that the Hybrid Linear Regression outperformed all other models for 3 month forecasts with the explanatory variables: GDP, credit default rates, and VIX volatility conditioned on credit crisis data ranges, but performed poorly during speculation periods such as the Tech Bubble. The Average of Averages Bootstrapping Simulation had the most consistent historical forecasts for12 month levels, and by using log returns from the Great Depression, Tech Bubble, and Oil Crisis the simulation indicates an expected value -2%, valid up to 12 months. ARIMA and Double Exponential smoothing models underperformed in comparison. ARIMA model does not adjust well in the "beginning" of a downward/upward pattern, and should be used when a clear trend is shown. However, the Double Exponential Smoothing is a good model if a steep incline/decline is expected. ARMAX + ARCH/EGARCH performed below average and is best used for volatility forecasts instead of mean returns. Lastly, Neural Network residual models indicate mixed results, but on average outperformed traditional time series models (ARIMA/Double Exponential Smoothing).

Additional research includes forecasting the S&P 500 with other nontraditional time series methods such as VARFIMA (vector autoregressive fractionally integrated moving averages) and ARFIMA models. Other Neural Network techniques include Higher Order Neural Networks (HONN), Psi Sigma network (PSN), and a Recurrent Neural Network (RNN) for additional forecasting comparisons.

Thesis Advisor:
Roy E. Welsch
Professor of Statistics and Management Science and Engineering Systems
MIT Sloan School of Management

ACKNOWLEDGEMENTS

Many thanks to Professor Roy Welsch, for challenging me to think about different dimensions in a well conversed and tried topic. I was fortunate to find an advisor whom was patient enough to teach me a few things about statistics, give me guidance on the endless questions, and share my enthusiasm about statistics, finance and life.

Special thanks to Philip Melville for inspiring me to think about the markets from a macro perspective, and for his guidance on drafting this thesis topic. Additional thanks to Joseph Demetrick, Mei-Feng Tau, Gregory Tell, and the MetLife Investments team for their mentorship in Finance.

I appreciate Patrick Hale, my SDM colleagues, MIT School of Engineering, and the Engineering Systems Division (ESD) for encouraging and teaching holistic thinking and the use of advanced engineering methods. I thank you for your support to transfer and apply this knowledge in the field of Finance. I would also like to thank Professor Mort Webster for his mentorship in simulation methods.

To my parents and grandparents, thank you for your lifelong devotions to your children. Thank you for all your love, support and sacrifice so that we (your children and all generations thereafter) could have the opportunity to pursue our dreams.

Lastly, I would like to thank my wife Jessica for all her unconditional love, support and encouragement. I fully appreciate her selfless sacrifices so that I could experience the opportunities available to me at MIT. I hope that I can extend the same support, encouragement, and kindness not only to our future children, but also to those who are less fortunate.

TABLE OF CONTENTS

TABLE OF FIGURES

1 INTRODUCTION AND MOTIVATION

"All models are wrong, some are useful."

George E. P. Box, British Statistician

1.1 PROBLEM STATEMENT

This thesis explores different methodologies used to forecast the S&P 500 Index during recessionary periods, potential factors and indicators that indicate possible recessions, and bull periods. Forecasting the S&P 500 Index has been of interest since its inception to several stakeholders within and outside of the financial domain. From an investor's perspective the Standard & Poor's 500 Index is one of the most influential indices in the world, several passive investment funds hold billions of dollars that track this index. From an economist's view, the index is a broad indicator of the U.S. economy, in some instances a global economic indicator as well. Forecasting the S&P 500 level is an important analysis tool for investors. Forecasting the index allows traders and portfolio managers to hedge their portfolio for the future or to consider longing or shorting U.S. Equity. For economists the S&P 500 Index forecasts help determine the next step for fiscal or monetary policy. For example, if index forecasts indicate a pessimistic 12 month level or return, then cutting interest rates or lowering banking reserves is a consideration. Finally, forecasting the S&P 500 is of interest to quantitative researchers and econometricians to understand the factors that drive the value of the future markets. Specifically, the objective of this thesis is to attempt forecasting the S&P 500 level in a 3 month and 12 month period using Linear Regression, Exponential Smoothing, ARIMA, GARCH, Neural Network, and Bootstrapping Simulation models. **Long range data forecasting** (blue) will be defined as prediction fit with data ranges of 10 years or greater with a forecast of 12 months. **Short range data forecasting** (red) is with a prediction fit range of 7 years or less, and forecasting of 1 to 3 months ahead. Moreover, the context of these forecasts will be to determine a 3 and 12 month forecast during the volatile months of the credit crisis (10/2008 to 09/2009). In addition to a specific period forecast, this thesis will advocate and recommend the best performing forecasting time series and statistical methodologies.

1.2 S&P 500 OVERVIEW

On March 4, 1957 the S&P 500 came into existence. Standard & Poor's issued many similar indices prior to the S&P 500 as early as 1926, but few indices have claimed the tracking success or the fame of the index with the exception of the Dow Jones which is an older index established in 1896. The S&P 500 Index holds constituents that are primarily U.S. based, in fact in the original S&P 500 this accounted for 90% of the index[i]. Today it has shrunk to 75%, which indicates that the index is moving towards an international list. Surprisingly, there are no restrictions against foreign stocks to include in the index. Therefore, a growing exposure to have more foreign based securities in the index is to be seen in the future. The common myth is that the S&P 500 represents the 500 largest companies in the United States. In reality, the S&P 500 is chosen by a select committee (think of a private club) that uses the criteria of "market size, liquidity, and group representation". The majority of the S&P 500 constituents are Large Cap, however the minimum capitalization needed to get into the index is $3 billion. Moreover, there are certainly at least 20 companies included in the S&P 500 that are small and mid capitalization in nature. The constituents are chosen from these major industry categories: Energy, Materials, Industrials, Consumer Discretionary, Consumer Staples, Health Care, Financials, Information Technology, Telecommunications Services, and Utilities[ii]. In addition, there are over 100 subgroups within these industries that determine the criteria.

2 ARIMA, EXPONENTIAL SMOOTHING, AND LINEAR TIME SERIES FORECASTING METHODOLOGY

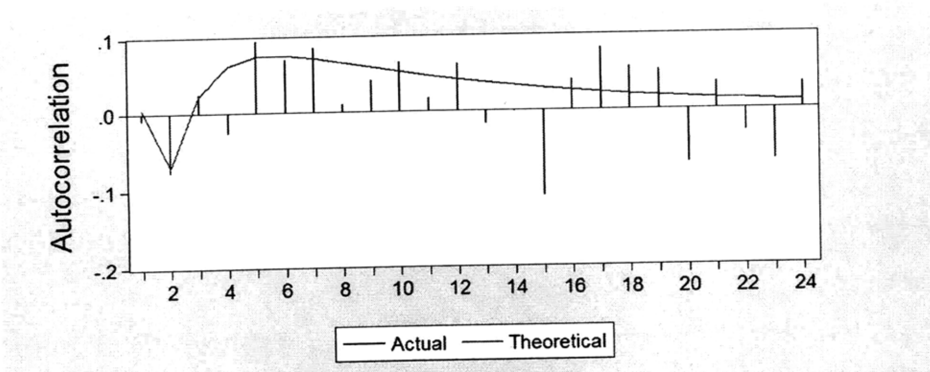
Time series forecasting is an important area which uses past observations to describe an underlying relationship. The most frequently used model during the research is the Autoregressive Integrated Moving Average (ARIMA) model. The following equation denotes a linear relationship of past observation values (autoregressive terms) and random errors (moving averages). p is the number of autoregressive terms [AR(p)] and q as the number of moving average terms [MA(q)]. p and q are referred as integers and are considered orders of the model. ø is the parameter estimation of the autoregressive terms and θ is the estimation for the random errors.

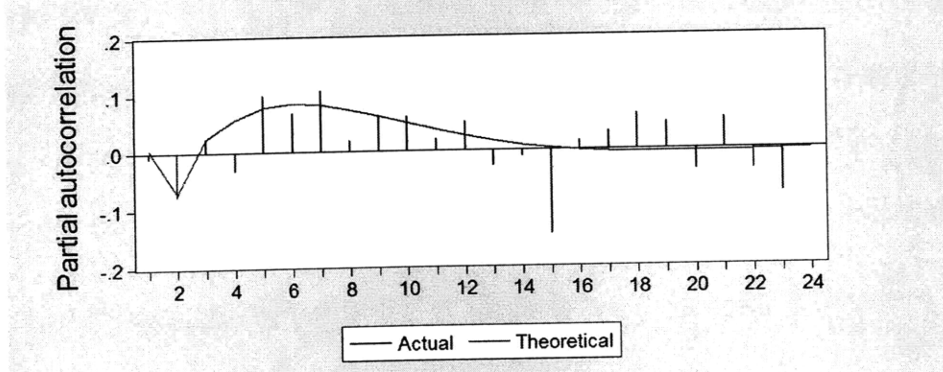
2.1 The Autoregressive Integrated Moving Average (ARIMA) Model[iii]

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q},$$

The popular Box Jenkins methodology[iv] is a three step stage to assist in accurate model building:

- Identification – First step is to identify a time series to be stationary. Inducing transformations and differencing for the time series achieve stationarity. [v] Second step is to determine the number of AR (p) and MA (q) by matching the theoretical autocorrelation patterns with the empirical ones. The following graphical example of the theoretical and empirical matching for an ARIMA (2, 1, 3) Model (2 AR, 1 differencing, 3 Moving Average) for the S&P 500 monthly data range of 1990 to 2008.





- Estimation – Estimation of parameters (ø , θ) using nonlinear optimization techniques. Models are compared where the smallest values of Akaike information criterion (AIC) and Schwartz Bayesian criterion (SBC) and parsimonious models are preferred.

- Diagnostic Checking – Examining for goodness of fit. Standard practice is to plot residuals and check for outliers as well as using the Q-statistic and Box-Pierce statistic to test for autocorrelations of the residuals.

For empirical results purposes ARIMA modeling will be used for 3 and 12 month forecasts.

2.2 Double (Brown) Exponential Smoothing Model[vi]

This equation is quoted directly from the SAS User Guide version 8.

The model equation for Double Exponential Smoothing is:

$$Y_t = \mu_t + \beta_t t + \epsilon_t$$

The smoothing equations are:

$$L_t = \alpha Y_t + (1 - \alpha) L_{t-1}$$

$$T_t = \alpha(L_t - L_{t-1}) + (1 - \alpha) T_{t-1}$$

This method may be equivalently described in terms of two successive applications of simple exponential smoothing:

$$S_t^{[1]} = \alpha Y_t + (1 - \alpha) S_{t-1}^{[1]}$$

$$S_t^{[2]} = \alpha S_t^{[1]} + (1 - \alpha) S_{t-1}^{[2]}$$

where $S_t^{[1]}$ are the smoothed values of Y_t and $S_t^{[2]}$ are the smoothed values of $S_t^{[1]}$. The prediction equation then takes the form:

$$\hat{Y}_t(k) = (2 + \alpha k/(1 - \alpha)) S_t^{[1]} - (1 + \alpha k/(1 - \alpha)) S_t^{[2]}$$

The ARIMA model equivalency to Double Exponential Smoothing is the ARIMA (0, 2, 2) model:

$$(1 - B)^2 Y_t = (1 - \theta B)^2 \epsilon_t$$

$$\theta = 1 - \alpha$$

The Double (Brown) Exponential Smoothing Model will be used for 3 and 12 month forecasts.

2.3 Damped-Trend Linear Exponential Smoothing Model (SAS V8, Section 11)

This equation is quoted directly from the SAS User Guide version 8.

The model equation for damped-trend linear exponential smoothing is:

$$Y_t = \mu_t + \beta_t t + \epsilon_t$$

The smoothing equations are:

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + \phi T_{t-1})$$

$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)\phi T_{t-1}$$

The k-step prediction equation is:

$$\hat{Y}_t(k) - L_t + \sum_{i=1}^{k} \phi^i T_t$$

The ARIMA model equivalency to damped-trend linear exponential smoothing is the ARIMA (1, 1, 2) model:

$$(1-\phi B)(1-B)Y_t = (1-\theta_1 B - \theta_2 B^2)\epsilon_t$$

$$\theta_1 = 1 + \phi - \alpha - \alpha\gamma\phi$$

$$\theta_2 = (\alpha - 1)\phi$$

Damped-Trend Linear Exponential Smoothing Model will be used for 3 and 12 month forecasts.

2.4 Log Linear Trend Model[vii]

$$x_t = b_0 + b_1 t + \varepsilon_t$$

The log linear trend is similar to the regression model; the key difference is that the explanatory variable uses 't' as a time period, and the natural log is taken on the left side of the equation: (ln x_t).

The Log Linear Trend Model is used for 3 and 12 month forecasts.

2.5 Log Random Walk w/ Drift Model[viii]

$$y_t = \mu + y_{t-1} + u_t$$

The random walk model is equivalent to a pure integrated ARIMA (0, 1, 0), taking the difference of its past observation. The drift is equivalent to the constant μ, so every observation is influenced by this constant. The Log Random Walk w/ Drift Model will be used for 3 and 12 month forecasts.

3 FORECASTING CRITERIA AND DATASOURCE

3.1 Criteria and Evaluation of Statistical Fit and Forecasting

To evaluate the goodness of fit for the evaluation range/hold-out sample (10/2007 - 9/2008) the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), AIC, SBC, and Adjusted R-Square are used as metrics. Specifically, the criteria used are to choose a model that has the lowest RMSE, and MAE. Secondary criteria include a low AIC, SBC score. The adjusted R-Square is for reference, and is not actively used in any criterion. The evaluation of statistical fit is a guideline in determining potential forecasting models. However, since a majority of the models perform within a similar range, i.e. RMSE is between 50-65, and *the fact these statistics are calculated solely within the evaluation range and not the forecasting range, a supplemental measure is needed.* These metrics will be shown in the statistical fit table. The metrics above (RMSE, AIC, SBC, etc.) can be calculated for every method presented in this thesis. However, the *focus of the thesis is evaluating forecasting metrics.*

The metric to evaluate the forecasting errors will be the Mean Absolute Error (MAE), this metric calculation will be separate from the one calculated on the evaluation range. The MAE will take the forecasted value and the actual value (not in the SAS datasets, but located in another source) as input parameters. This calculated measure will be shown on the forecasting tables.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |f_i - y_i| = \frac{1}{n}\sum_{i=1}^{n} |e_i| .$$ [ix]

In addition the Mean Absolute Percentage Error will be evaluating available forecasting periods, which calculates the absolute value of the actual observation minus the forecasted value divided by the actual observation. This result is then divided by the number of observations, which represents the average percentage error.

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{A_t - F_t}{A_t}\right|$$ [x]

3.2 Datasource and Statistical Tools

The source of the monthly S&P 500 return and index data can be found at the Wharton Research Data Services (WRDS) using The Center for Research in Security Prices (CRSP) data set. However, monthly data only extends to the end of December 2007. Therefore, the 2008 data was sourced directly from the official Standard & Poor's website in the index data section. SAS version 9.1 was used as the statistical tool to calculate the statistical fit and parameter estimations for the ARIMA, Double Exponential Smoothing, and other Linear Models. Moreover, EVIEWS version 6.0 was used to determine the best ARIMA model in the Box-Jenkins identification stage (determination of the order number of AR and MA).

The Hybrid Linear Regression model used Bloomberg Professional to gather monthly level pricing data on the VIX, GDP, and Corporate Default rates which ranged from 1/1990 to 1/2008. EVIEWS was also used for parameter estimation and statistics of fit for the Hybrid Linear Regression model.

Lastly, the Neural Network models used SAS Enterprise Miner 4.3, for weighting and parameter estimation. MATLAB R2008b was used to perform Bootstrapping Simulations and to estimate parameters and mean/volatility forecasts for ARCH/GARCH models.

4 ARIMA, EXPONENTIAL SMOOTHING, AND LINEAR APPROACH AND EMPIRICAL RESULTS

The approach was to take different time ranges of the S&P 500 monthly index data to forecast a 3 month and 12 periods ahead of the recent credit crisis (forecasts beginning on October 2008). Prediction fit vary from 12/1925 to 09/2007. The concept was to try different models defined in the methodology section with different monthly ranges. Each data set was optimized with the appropriate model. The prediction fit range is generally organized into five categories:

- Entire dataset, with a holdout sample of 8 to 12 months (1925 to 2007). This was done to capture all historical economic boom and recession data.
- 30 Year dataset with a holdout sample of 12 months (1975 to 2007). This was to include the Oil Crisis in 1973 and 1974, the '92 and '93 recession, and the Tech Bubble in 2000-2003.
- 17 Year dataset with a holdout sample of 8 to 12 months (1990 to 2007). This was to include the '92 and '93 recession, and the Tech Bubble data.
- 7 Year dataset with a holdout sample of 12 months (2000 to 2007). This was to include data of the Tech Bubble.
- 1 and 2 year dataset, no holdout sample. This data is to observe specifically the recent events of the credit crisis.

The ARIMA (2, 1, 3) model is the base model which is used in multiple data set categories, the 17 and 30 year data sets. The model was determined using the Box Jenkins approach mentioned earlier. Other models such as the Damped and Double Exponential Smoothing are used in multiple data set categories as well. See the fit statistics table for a comprehensive overview of the type of model and period range.

4.1 The Credit Crisis Forecasting Results with ARIMA and Exponential Smoothing

For the models highlighted in blue in table 1, ARIMA(2,1,3), Log Damped Exponential Smoothing models showed strong metrics for 12 month forecasting with a 12 month evaluation range. The Log Linear Trend had the lowest RMSE and MAE, but also had the shortest fit of data of 1 year and no hold-out sample. The primary idea with the short range models (in red) was to fit a model that put a lot of weight on the recent credit crisis observations. The range of the forecast is a specific 3 month and 12 month forecast: 10/2008 – 12/2008, 10/2008 - 09/2009. These dates were chosen given that the latest data at this time only extends to 12/2008. Furthermore, only 3 month metrics can be calculated at this time. Validating the 12 month forecasting will be addressed in the Tech Bubble forecasting section.

So does the models that fit well in the evaluation range typically forecast well? Not necessarily. The forecasting table 2 displays the results. The Log Linear Model, which initially seemed to be a good short run forecast, had the second worst MAE ranking (The Log Random Walk w/Drift had the worst). The Log Double (Brown) Exponential Smoothing, 7 year dataset had the lowest forecasting MAE, but in the evaluation range ranked in the middle (Parameter estimation can be found in the appendix). The Double Brown Exponential Model outperformed the other models, because of the factoring of the strong non-linear differences that was present in the past observations and errors in the index. *Since, the credit crisis during the months of Oct – Dec of 2008 saw unprecedented drops in the level, the Double Brown Exponential Model adjusts by predicting a steeper descent.* However, even though the MAE is lower the Brown model still displayed a MAE of 205, meaning that each forecasting month differed by 205 points per observation on average. At the end of December, the 95% confidence interval for the Brown model was between 949 and 1265, with the actual of 903.06 (see figure 3). *Clearly, the fall in the index was beyond a 2 sigma event and this leads to a discovery of the limitation of all the linear and exponential smoothing models: the point estimates are poor predictors of unprecedented events.*

For the 12 month forecasting the ARIMA(2,3,1) model or the equally weighted forecasting combination ARIMA(2,3,1) and Log Damped Exponential Smoothing is recommended (Parameter estimation can be found in the appendix). Since the full 12 month of forecasting data is not yet available for the credit crisis, the validation of these models will be shown in the *12 month Tech Bubble prediction section. Furthermore, after analysis a key insight can be revealed about these models: the S&P 500 level is potentially undervalued when considering long term data. In figure 4, shows a 12 month level combination prediction of 961 and a 95% confidence interval of 481 to 1442.* December 2008 shows a value of 903.25. The S&P 500 Index may very well continue to fall, and the 12 month prediction may be inaccurate. However, the long range models are using several decades of data as input, and through historical observation, the current level has declined to the level of October 2002, which was the bottom of the Tech Bubble. The S&P 500 has a history of adjusting towards the long run mean. Figure 5 shows the long run S&P 500 monthly returns mean of .58%, and generally shows negative returns being offset by positive runs having properties of a normal distribution. Furthermore, the monthly mean represents the S&P 500 price returns and does *not include dividend yields.* Figure 6 shows the S&P 500 returns during the Credit Crisis, the short run mean is -1.99%. *If the assumption of long run mean converge holds true, then the S&P 500 level will be higher to reflect positive returns that offsets the recent negatively skewed data.* See appendix for detailed distribution data.

Table 1 – Fit Statistics Overview for All S&P 500 Datasets, Credit Crisis Evaluation Range (10/2007-09/2008)

	Fit Range	Eval Range	RMSE	MAE	MAPE	AIC	SBC	Adj R-Square
Log Damped Exponential Smoothing	12/1925 - 09/2007	10/2007 - 9/2008	63.578	49.24	3.727	105.65	107.11	0.528
ARIMA(2,1,3) + Log Damped Exp. Forecast Combo	12/1925 - 09/2007	10/2007 - 9/2008	69.589	50.07	3.785	103.63	104.6	0.576
Log Random Walk W/ Drift	1/1975 - 09/2007	10/2007 - 9/2008	69.247	52.94	4.02	103.7	104.19	0.542
Log Damped Exponential Smoothing	01/1970 - 01/1985	10/2007 - 9/2008	63.404	48.73	3.691	105.59	107.04	0.531
ARIMA (2,1,3)	01/1990 - 01/2007	02/2007 - 09/2008	63.85	50.89	3.844	109.756	112.18	.388
Linear Regression	1/1990 to 9/2007	10/2007 - 9/2008	55.397	39.75	3.053	-3.906	-3.844	0.995
Log Double (Brown) Exponential Smoothing	1/2000 - 9/2007	10/2007 - 9/2008	65.911	58.05	4.323	102.52	103	0.585
Log Random Walk W/ Drift	10/2006 - 9/2008	10/2006 - 9/2008	51.577	42.45	3.07	191.27	192.45	0.691
Log Linear Trend	10/2007 - 9/2008	10/2007 - 9/2008	44.388	37.08	2.754	95.031	96.001	0.793

Table 2 - S&P 500 Index 3 Month and 12 Month Forecast: 10/2008 – 12/2008, 10/2008 - 09/2009

			Actual	Actual	Actual			
			968.8	896.24	903.25			
	Fit Range	Eval Range	Predict Month 1	Predict Month 2	Predict Month 3	3 Mon Forecast Error	3 Mon Forecast MAE	12 Mon Forecast
Log Damped Exponential Smoothing	12/1925 - 09/2007	10/2007 - 9/2008	1162	1158	1154	706	235	1134
ARIMA(2,1,3) + Log Damped Exp. Forecast Combo	12/1925 - 09/2007	10/2007 - 9/2008	1156	1136	1115	639	213	**961**
Log Random Walk W/ Drift	1/1975 - 09/2007	10/2007 - 9/2008	1175	1184	1193	784	261	1278
Log Damped Exponential Smoothing	01/1970 - 01/1985	10/2007 - 9/2008	1162	1159	1157	710	237	1153
ARIMA (2,1,3)	12/1925 - 09/2007	10/2007 - 9/2008	1160	1157	1147	696	232	1017
Log Double (Brown) Exponential Smoothing	1/2000 - 9/2007	10/2007 - 9/2008	1160	1128	1096	616	**205**	850
Log Random Walk W/ Drift	10/2006 - 9/2008	10/2006 - 9/2008	1150	1141	1133	656	219	1069
Log Linear Trend	10/2007 - 9/2008	10/2007 - 9/2008	1193	1170	1147	742	247	961

Figure 3 - Log Double (Brown) Exponential Smoothing, Seven Year Data Set (2000 – 2007)

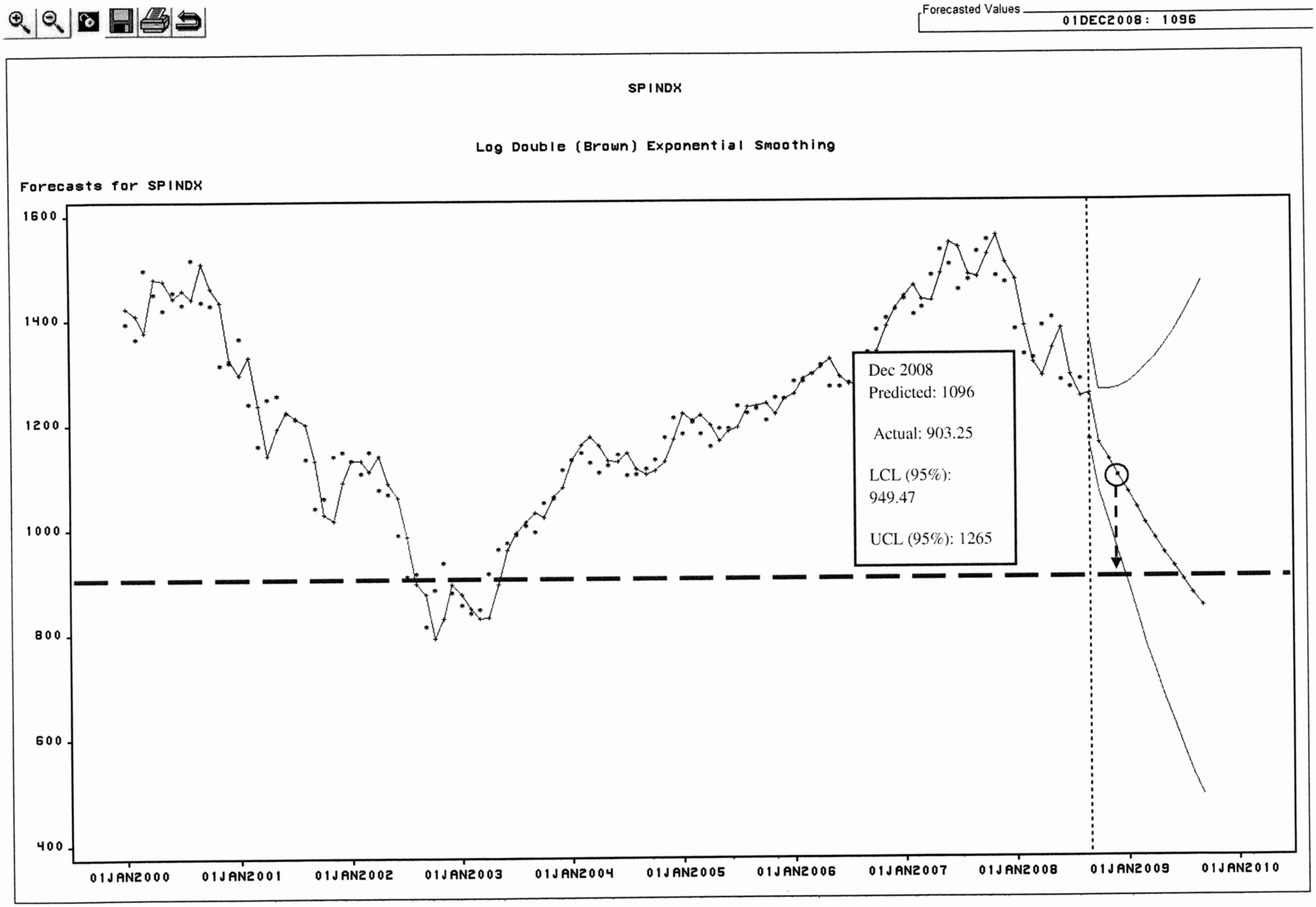


Figure 4- Combined Forecast ARIMA(2,1,3) + Log Damped Trend Exponential, Entire Data Set (1925 – 2008)

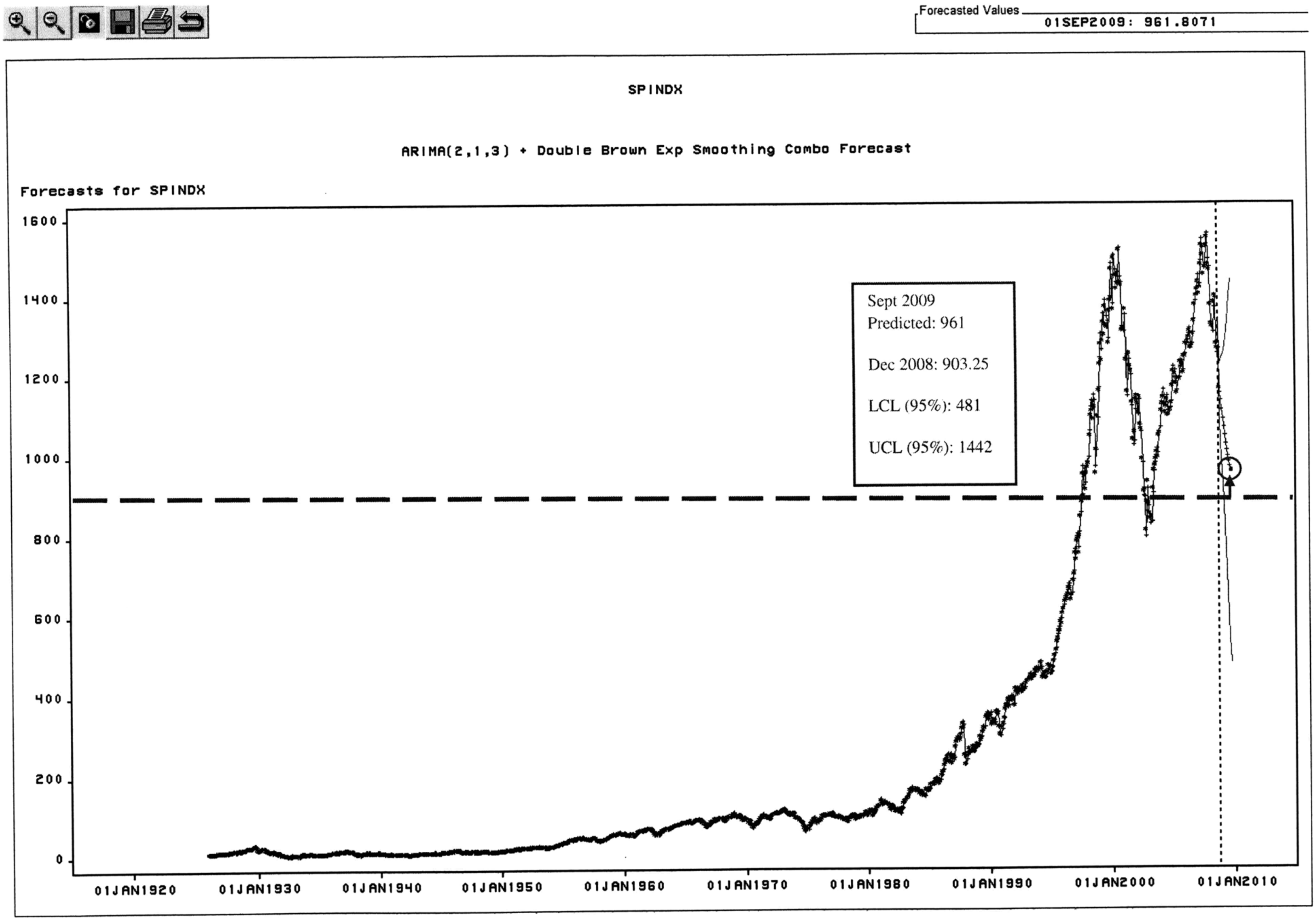


Figure 5 - S&P 500 Monthly Returns Long Run Mean (No Dividend Yield)

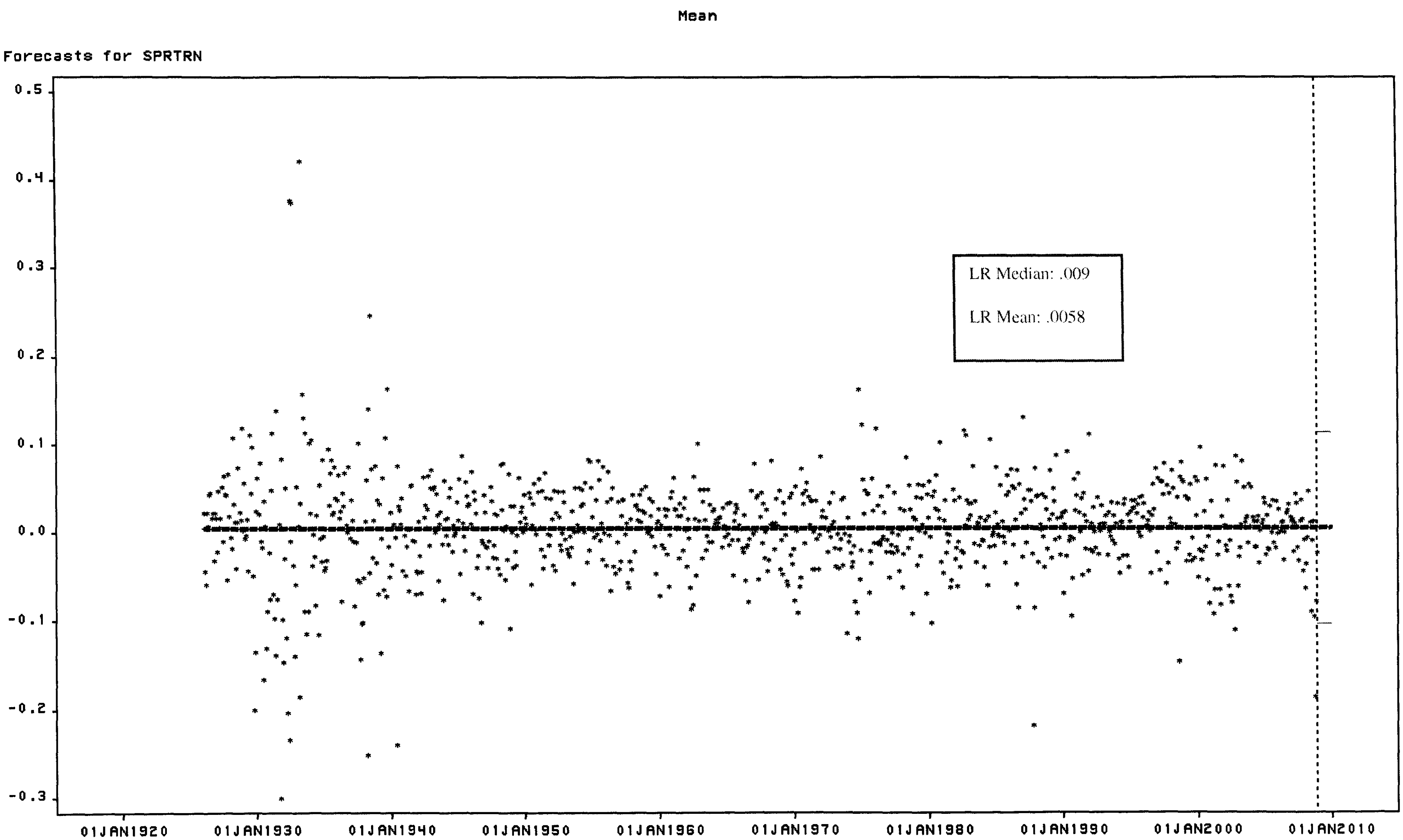
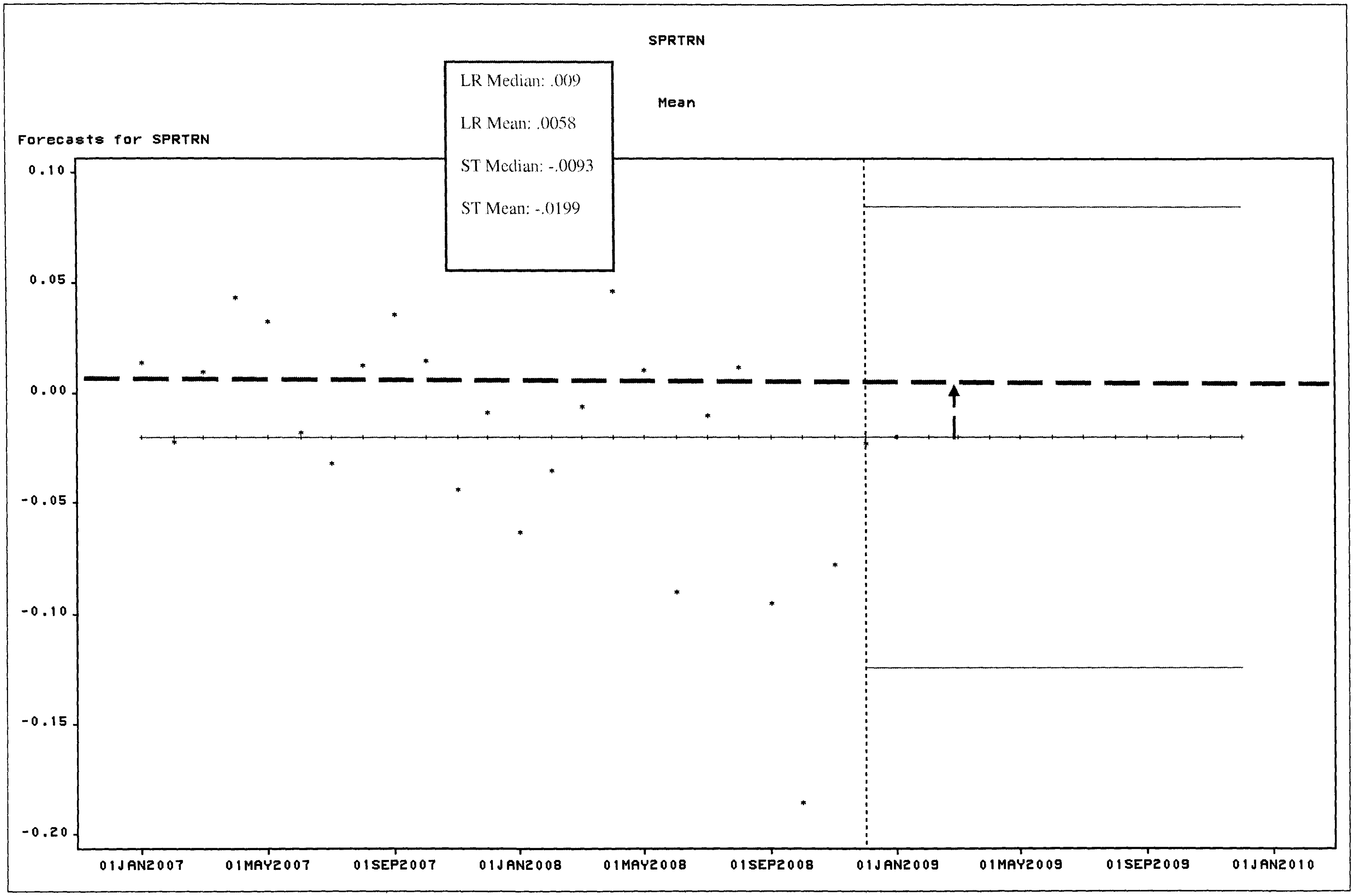


Figure 6 - S&P 500 Returns Credit Crisis Mean, Long Run Mean Convergence



4.2 Tech Bubble (2000-2003) Prediction, Validating the 12 Month Forecasting Models

The Credit Crisis has been hailed as unprecedented in its time, but according to the index history, the decrease in levels has not proven to be more severe than previous recessions[xi] (The Great Depression of 1929-1932, The Oil Crisis of 1973-1974). The Tech Bubble time range was chosen because it shows similar level patterns of the boom and bust (1998-2002) vs. the rise and fall of the mortgage market (2005-2008). Again, the current levels of the credit crisis is close to the bottom of the Tech Bubble (793 vs. 876) and similar peak levels (1518 vs. 1558), all within a 3 to 4 year period. Therefore modeling for the Tech Bubble seemed appropriate considering the similar patterns.

Historically, the Tech Bubble or "Dot-Com Bubble" is an event of rapidly increasing stock prices due to speculative valuation. John Cassidy illustrated the example of this era with the IPO of Priceline.com. After three weeks of Priceline's IPO, the stock reached $150 at which this "tiny" company was *worth more than the entire Airline Industry*. After the crash, Priceline.com was trading at less than $2 a share, which would not have covered the cost of two Boeing 747s[xii]. Many entrepreneurs relied on venture capitalists and IPOs of stocks to pay their expenses. With low interest rates (1995 to 1998) and lower capital requirements than a traditional brick and mortar businesses, entrepreneurs attracted the investment capital needed to fund several startups. The dot com business model relied on the network effect or "get big fast" to collect market share first with the hopes to be profitable in the future (Cassidy 2003). Many investors naively expected the higher valuation to match future returns of the tech companies. The burst of the bubble occurred on March 10, 2000 and was attributed to several reasons:

- To balance the speculative rise, the Federal Reserve increased interest rates six times over the period of1999 to 2000.
- Microsoft was declared a monopoly in the federal court ruling of The United States vs. Microsoft.
- The NASDAQ peaked at 5048.62 which were more than double the previous year.
- After the posting of poor year end earning announcements in 1999 and 2000, investors realized that tech stocks were overvalued.

- Decreased IT budget spending after the Y2K system overlay.

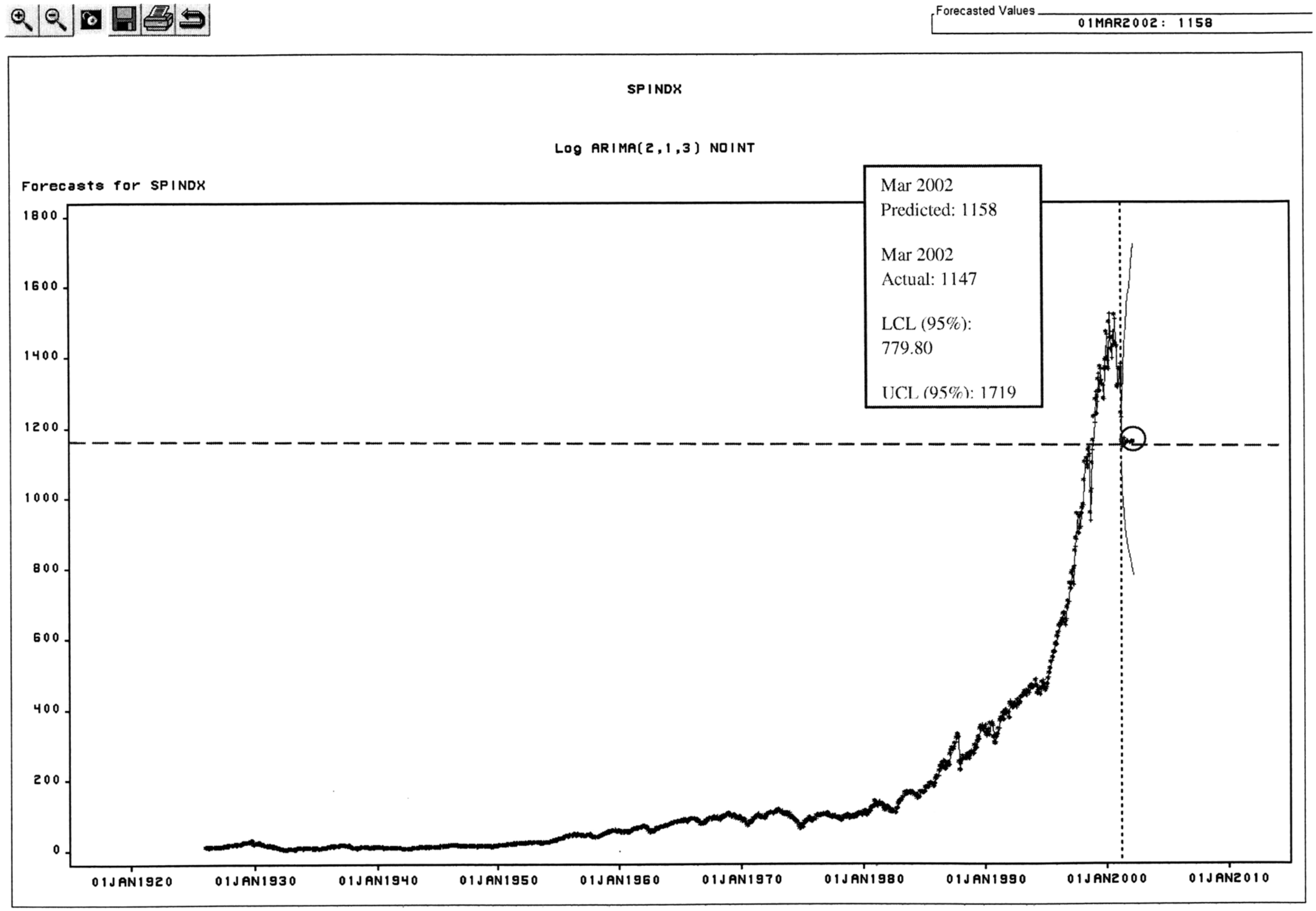
Lastly, the S&P 500 Index aggregates companies of market capitalization. With tech companies representing inflated market capitalization, the S&P 500 had increased exposure to the tech sector, thus posting increased declines. In fact the S&P 500 peaked at 1527 on March 24, 2000. And by the end of December 2002, the S&P 500 fell to 875, a decline of 55.7% from its peak.

The primary purpose of the Tech Bubble model is to validate the Log ARIMA (2, 1, 3) model and the Log Double Exponential Smoothing model for an accurate 12 month forecast. The 2 SAS datasets used was within 12/1925 to 12/2000 (12 month forecast starting on 1/2001) and 12/1925 to 3/2001 (12 month forecast starting on 4/2001). Each model performed differently between these two datasets. Looking at the forecasting results shown in table 7, the Log Double Exponential Smoothing had the lowest MAE of 48.99 and the best 12 month average fit for the 1/2002 forecast, but the ARIMA model performed poorly with a MAE of 159.49. At the end of the data set the last 6 observations had a subtle decline in movement (2 observations that decreased the level from 81 to 100). *ARIMA model does not adjust well in the "beginning" of a downward/upward pattern, and should be used when a clear trend is shown. However, the Double Exponential Smoothing is a good model if a steep decline is expected.* In the second dataset which is 3 months later from the first (forecast starting on 4/2001), the exact opposite effect is observed. This dataset shows that the Log ARIMA (2, 1, 3) model outperforms the Double Exponential Smoothing model. This is because the Double Exponential Smoothing predicts an exponential increasing or decreasing value. The results show a large discrepancy to undervalue the prediction and had a sizable MAE of 205. The Log ARIMA model performed better due to a clear downward pattern in the past 5 observation lags, which led to having the best 12 month point forecast (but not the best overall MAE) between the 2 data sets. See Figure 8 for the graphical fit. To summarize, if a steep decline or rise in the index is expected with only 1 or 2 observations showing subtle hints of decline, and then the Double Exponential Smoothing Model will outperform linear models. *However, if there is an established trend at the end of the dataset (4 -5 observations showing a clear trend), then the ARIMA (2, 1, 3) model may perform and forecast better.*

Table 7 – 12 Month Forecast, Predicting the Tech Bubble

				Start Jan 2001 (Actual)	1366	1240	1160	1249	1256	1224	1211	1134	1041	1060	1139	1148
				Start April 2001 (Actual)	1249	1256	1224	1211	1134	1041	1060	1139	1148	1130	1107	1147
	Fit Range	**Eval Range**	**12 Mon Forecast Error**	**12 Mon Forecast MAE**	**Mo 1**	**Mo 2**	**Mo 3**	**Mo 4**	**Mo 5**	**Mo 6**	**Mo 7**	**Mo 8**	**Mo 9**	**Mo 10**	**Mo 11**	**Mo 12**
Log Double (Brown) [Start Jan]	12/1925 – 12/1999	01/2000 – 12/2000	587.9	48.99167	1288	1260	1234	1210	1189	1171	1156	1144	1135	1130	1129	1132
Log ARIMA (2,1,3) [Start Jan]	12/1925 – 12/1999	01/2000 – 12/2000	1914	159.4983	1319	1331	1333	1329	1333	1338	1339	1340	1343	1346	1348	1350
Log Double (Brown) [7 year prev data]	12/1992 – 12/1999	01/2000 – 12/2000	608.9	50.74167	1299	1275	1251	1229	1207	1186	1166	1147	1129	1112	1096	1082
Log ARIMA (2,1,3) [Start April]	12/1925 – 03/2000	04/2000 – 03/2001	676.8	56.39833	1142	1158	1169	1157	1153	1161	1162	1157	1157	1161	1159	1158
Log Double (Brown) [Start April]	12/1925 – 03/2000	04/2000 – 03/2001	2569	214.1083	1129	1084	1043	1004	968.4	936	907	881	858	838.1	821.6	808.4

Figure 8 - Good Forecast: Log ARIMA (2,1,3), Entire Data Set (1925 – 2001)



5 LINEAR REGRESSION METHODOLOGY AND EMPIRICAL RESULTS

5.1 Linear Regression Methodology

Linear Regression Model[xiii]

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i, \qquad i = 1, \ldots, n$$

The linear regression is fundamentally different from ARIMA. Instead of having past observations and errors define the underlying relationship, the response variable Y_i is a factor of X (i) number of explanatory variables. The linear regression model uses an ordinary least squares (OLS) method for β parameter estimation. The baseline linear regression model will be used for long range data forecasts.

Multivariate Representation

In matrix from the multivariate linear regression can be represented as follows:

$$Y = X\beta + \varepsilon$$

$$X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix}$$

Where *X* represents a matrix of repressors, *Y* is a column vector that represents the dependent variables and $\varepsilon_1, \ldots, \varepsilon_n$ represents the unobserved vector errors. $\beta_0 \ldots \beta_n$ can be represented as the parameter estimates for matrix *X*, but if there are any linear dependency within columns of *X* then creating parameter estimates with $\beta_0 \ldots \beta_n$ is problematic, therefore choosing explanatory variables that reveal additional insight versus overlapping information is key for an accurate prediction model. Lastly, $\beta_0 \ldots \beta_n$ can be estimated via the least squares method and the solving of normal equations:

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

I refer you to Montgomery (2008) for a full derivation of the least square normal equation estimate.

5.2 The Baseline Linear Regression Model

The number of explanatory variables used for the linear regression model was kept to a minimum. The chosen factors are GDP expressed in billions of chained 2000 dollars, CAY (consumption, asset wealth, and labor income), and Default Spread using the Lehman Brothers Investment Grade Spread Index w/ Option Adjusted Spread (IG OAS), and finally a AR(1) term, which has a relationship with past errors. The consumption-wealth ratio is from Lettau and Ludvigson (2001, 2004) and is calculated as CAY = c – 0.2985a – 0.597y, where c is real per-capita consumption of nondurables and services, a is financial wealth, and y is labor income, all measured in natural logarithms.[xiv] Statistically it showed a very good fit. This model shows promise as a fair value model, evaluating the current S&P 500 level to determine relatively if the price is fair, under, or overpriced. However, as a forecasting model there are several challenges. Each explanatory variable will need to have 3 to 12 month forecast projections. Since the CAY data is not updated frequently this poses challenges to this model, and we will not explore any further in using the baseline linear regression model for 12 month forecasting. However, each explanatory variable would use either a projection from a reputed analyst or another black box forecasting method like ARIMA to project, which may or may not provide an accurate forecast. This next section will attempt to introduce a hybrid linear regression with an ARIMA and Exponential smoothing forecasting model.

5.3 Linear Regression Hybrid Model with ARIMA or Double Exponential Smoothing Forecasts

In this hybrid model a revisitation of the explanatory variables was warranted. Lettau and Ludvigson implied that the CAY data point are cointegrated and thus lead to important predictive components for future returns. [xv] After a second evaluation, CAY is a difficult measure to forecast effectively with a black box statistical method. Moreover, the forecasted the values that were constant and slightly above 0 in the context of mean returns. Instead of using CAY, the

VIX indicator is a more appropriate substitute for the CAY. VIX is a ticker symbol used in the Chicago Board Options Exchange Volatility Index. The VIX is often a measure of the implied volatility of the S&P 500 Index options. This volatility measure is more specific to the S&P 500 and in the past has proved to be a good indicator for forecasting. However, there have been concerns about the VIX's ability to predict market decline. Kearns and Tsang state the VIX failed to predict the October 2008 17 percent drop, but also noted that the VIX in the past has only failed five times in its 18 year history to predict the S&P 500 volatility within a 10% margin[xvi] Bill Luby describes the VIX as an annualized volatility measure, and can use a simplistic Gaussian distribution to predict monthly pricing by converting the annualized measure to monthly and using the confidence intervals of a logarithmic or Gaussian distribution as a heuristic estimate.[xvii] Nonetheless, this thesis will test the viability of using black box models to forecast potential volatility measures. For the hybrid approach, the ARIMA (2, 1, 3) will be used to forecast VIX volatility. The IG OAS and GDP level will remain unchained from the baseline regression. The table below shows the forecasting model chosen for each explanatory variable and the respective forecasted values. The appendix contains the parameter estimates to these forecasts.

Table 9 – Forecasted Results of Explanatory Variables

Forecasting Model	ARIMA(2,1,3)	ARIMA(2,1,3)	Double Exponential Smoothing
Date	**VIX**	**IG_OAS**	**GDP_LVL**
10/30/2008	36.73	355.75	11707
11/30/2008	33.903	349.47	11702
12/30/2008	32.27	349.74	11697
1/30/2009	31.4891	354.24	11692
2/30/2009	31.2881	353.64	11687
3/30/2009	31.4558	350.45	11682
4/30/2009	31.8466	351.17	11677
5/30/2009	32.34	353.4	11672
6/30/2009	32.86	352.68	11667
7/30/2009	33.37	351.14	11662
8/30/2009	33.38	351.8	11657
9/30/2009	34.23	352.85	11652

5.4 Linear Regression Hybrid Models and Credit Crisis Empirical Results

Overall, two linear regressions were estimated with the Hybrid model, one with a dependent variable of S&P 500 Index returns (SPINDX Return), and the other using the first differences of the S&P 500 Index D(SPINDX). The SPINDX Return uses a static forecast, which essentially will perform parameter estimation based upon the explanatory variable data in addition to the forecasted ARIMA/Exponential Smoothing data referenced in the previous section. The first differences model: D (SPINDX) uses a dynamic forecasting model. The equation below represents the initialization of the dynamic forecast model for the *first observation.*

$$\hat{y}_S = \hat{c}(1) + \hat{c}(2)x_S + \hat{c}(3)z_S + \hat{c}(4)y_{S-1},$$

where c represents the beta coefficients and x_S and z_S represent explanatory variables within the X matrix. And Y_{S-1} represents the *actual* value of the lagged endogenous variable: D(SPINDX), in the period prior to the start of the forecast sample.

$$\hat{y}_{S+k} = \hat{c}(1) + \hat{c}(2)x_{S+k} + \hat{c}(3)z_{S+k} + \hat{c}(4)\hat{y}_{S+k-1}.$$

Subsequent observations uses the equation above, the key difference is that the variable Y_{S+k-1} uses the *predicted value Y_hat* in the period prior instead of the actual endogenous variable to perform its forecast.

Overall, the hybrid linear regression models provide superior short term forecasts with both regressions outperforming all stand alone models which include the ARIMA and Double Exponential Smoothing models, and the damped exponential combination. The D(SPINDX) had an impressive 3 month short term forecast with only a *20 Mean Absolute Error* per month. The SPINDX returns model had the second lowest MAE with 153 per month. Although the D(SPINDX) predicted the 3 month volatility with close precision, it is *unlikely* that its 12 month forecast of 439, will be accurate unless *persistent increase* of volatility will continue in 2009. In addition, this estimate has a highly variable 95% confidence interval between 0 and 1647. With a stabilization of volatility the SPINDX Return model will be a *better* 12 month forecast, which accounts for the actual endogenous values versus a predicted one. However, both linear regressions indicate declining 12 month forecasts, which a possible scenario is given the assumption of increasing volatility and corporate defaults. Lastly, there are issues with heteroskedasticity with the assumption of constant variance violated in the classical linear regression model (Asteriou and Hall 2007). After performing a series of ARCH (10) and White test the null hypothesis of homoskedasticity is rejected. The consequences may include underestimation of the SPINDX return model, thus leading to an underweighted forecast. The GARCH models detailed in the next section will adjust for this issue and bring a different forecasting perspective.

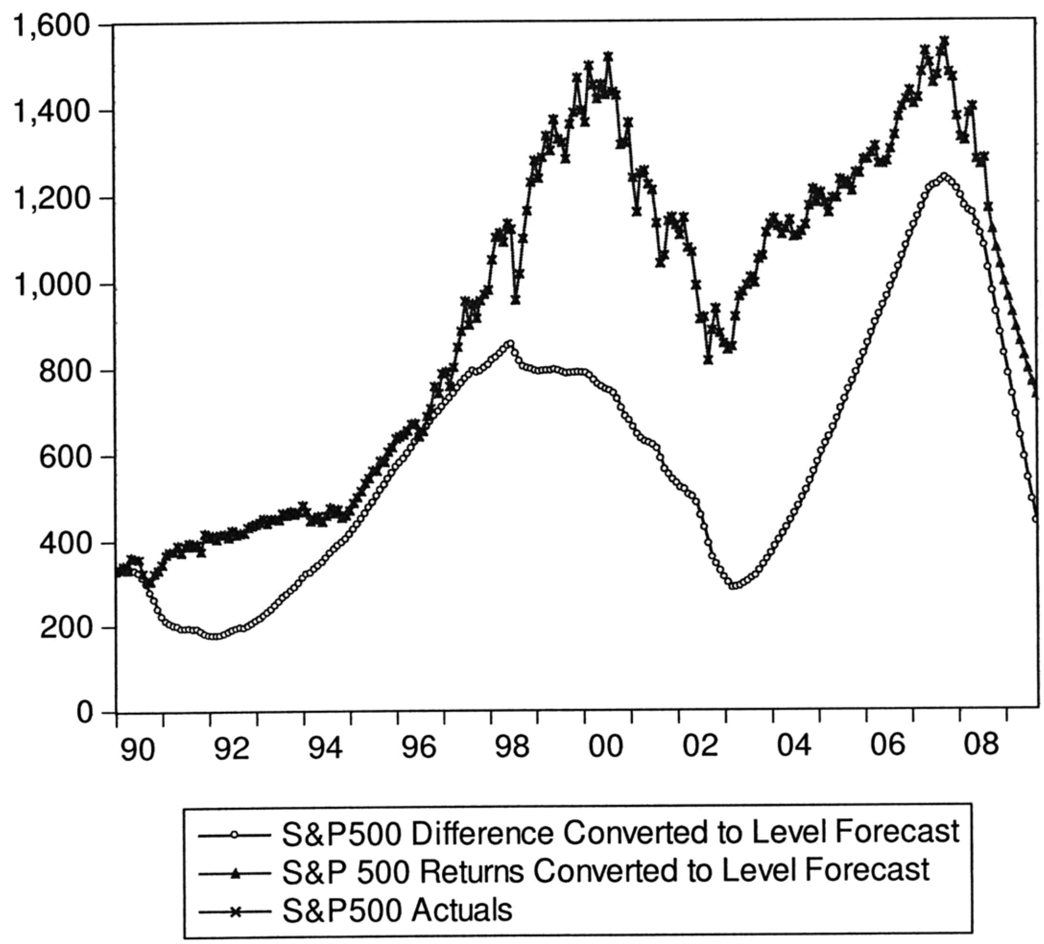
Table 10 – Linear Regression Comparison to Credit Crisis Forecasted Results

			Actual	Actual	Actual			
			968.75	896.24	903.25			
	Fit Range	Eval Range	Predict Month 1	Predict Month 2	Predict Month 3	3 Mon Forecast Error	3 Mon Forecast MAE	12 Mon Forecast
ARIMA(2,1,3) + Log Damped Exp. Forecast Combo	12/1925 - 09/2007	10/2007 - 9/2008	1156	1136	1115	639	213	961
ARIMA (2,1,3)	01/1990 - 01/2008	02/2008 - 09/2008	1156	1153	1141	682	227	1027
Log Double (Brown) Exponential Smoothing	1/2000 - 9/2007	10/2007 - 9/2008	1160	1128	1096	616	205	850
Baseline Linear Regression	01/1990 - 01/2008	02/2008 - 09/2008	1167	1160	1130	689	230	N/A

Hybrid Linear Regression Y=(SPINDX RET)	01/1990 - 01/2008	02/2008 - 09/2008	1118	1075	1035	460	153	**733**
Hybrid Linear Regression Y=D(SPINDX)	01/1990 - 01/2008	02/2008 - 09/2008	973	924	877	59	**20**	439

Figure 11 combines the actual values and the two hybrid linear regression model forecasts converted into level amounts into a single graph.

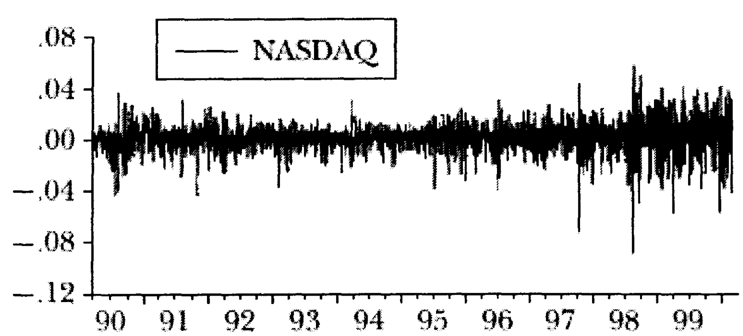
Figure 11 – Forecasted D(SPINDX), SPINDX Returns with S&P 500 Actuals



6 ARCH/GARCH TIME SERIES FORECASTING METHODOLOGY

Linear regression models that use the least squares model assume that the variance is constant at any given point time, that is the expected value of all the squared error teams is constant. The assumption of constant variance is called homoskedasticity. Financial time series data often do not follow this assumption. The S&P 500 Index and returns can be extremely volatile during recessionary and growth periods. With careful inspection, the S&P 500 and many other financial equities display increasing variances during certain periods. This increasing variance, also called heteroskedasticity is quite problematic causing standard errors and confidence intervals estimated by the ordinary least squares regression to be too narrow in precision.[xviii] The coefficients of the OLS are preserved and are unbiased, but this false precision typically leads to erroneous conclusions. There are different methodologies to correct for heteroskedasticity. Douglas Montgomery suggests that if the changes in the variance at certain time intervals are known then a weighed least squares regression can be employed.[xix] Although, these particular time intervals are often not known, and there is value in knowing when and why variance increases. ARCH (Autoregressive Conditional Heteroskedasticity) and GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models the variance of each error term and corrects the deficiencies of heteroskedasticity for least squares. Moreover, these models deal with the issue of "volatility clustering", when the amplitude of the time series level varies over time. For example, the NASDAQ Index has shown in figure 12, wide amplitude of returns. This suggests that certain NASDAQ returns are riskier/less riskier than others. ARCH/GARCH models are methods to deal with volatility clustering and to provide a measure of volatility.

Figure 12 - NASDAQ Daily Returns (Engle 2001)



Forecasting and modeling for variance is extremely useful in option pricing and value at risk calculations. This thesis will not concentrate deeply on these subjects, but focus on the conditional mean prediction and forecasting of the ARCH/GARCH models. However, the forecast of these variances will still be very useful for in calculating the confidence intervals and forecasted standard errors.

6.1 Conditional Mean Equation: ARMAX(R,M,Nx)

The following equation below is considered the conditional mean, where y_t is considered the returns and $\varepsilon_t = \sigma_t z_t$, where z_t are i.i.d random variables with mean 0 and variance 1, independent of past realizations of εt_{-j} and $\sigma^2_{t.}$ This conditional mean model is also known as the ARMAX(R, M, Nx) which stands for Autoregressive moving average model with exogenous inputs model. R indicates the order of the autoregressive terms, M specifies the order of the moving average terms for the innovations, and N_X represents the number of exogenous or time series and explanatory factors that is to be included in the conditional mean.

$$y_t = C + \sum_{i=1}^{R} \phi_i y_{t-1} + \varepsilon_t \sum_{j=1}^{M} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{Nx} \beta_k x(t,k)$$

6.2 The Conditional Variance Equation: GARCH(1,1)

This thesis will employ several conditional variance methods: The GARCH (P, Q) model invented by Bollerslev (1986), which is essentially an extension of the ARCH model proposed by Engle (1982). h_t is defined as the variance, and the weights are $(1 - \alpha - \beta, \beta, \alpha)$, and the long run variance is calculated as $\sqrt{\omega/(1 - \alpha - \beta)}$. This works if all the weights are positive and $\alpha + \beta < 1$.

$$h_{t+1} = \omega + \alpha(r_t - m_t)^2 + \beta h_t = \omega + \alpha h_t \varepsilon_t^2 + \beta h_t.$$

This equation is an example of a GARCH (1, 1). P indicates the order of past (t – P) squared return residuals and Q is the order of (t – Q) past variances.

6.3 The Conditional Variance Equation: Exponential GARCH

The Exponential GARCH also known as EGARCH was introduced by Nelson (1991):

$$\log \sigma_t^2 = \omega_t + \sum_{k=1}^{\infty} \beta_k g(Z_{t-k}) + \sum_{k=1}^{\infty} \alpha_k \log \sigma_{t-k}^2$$

Where $g(Z_t) = \theta Z_t + \lambda(|Z_t| - E(|Z_t|))$, σ_t^2 is the conditional variance and ω, β, α, θ and λ are coefficients, and Z_t is a standard normal variable.[xx] ARCH/GARCH models have a symmetric effect on the errors for the conditional variance, i.e. a positive error has the same effect as a negative error of the same magnitude.[xxi] The EGARCH model adjusts for asymmetry by using function $g(Z_t)$. When $0 < Z_t < \infty$, $g(Z_t)$ is linear in Z_t with slope $(\theta + \lambda)$; when $-\infty < Z_t < 0$, $g(Z_t)$ is linear with slope $(\theta - \lambda)$. This allows for the conditional variance to respond asymmetrically to the rises and falls of the process.

6.4 The Conditional Variance Equation: Glosten-Jagannathan-Runkle GARCH

Lastly, the Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) model by Glosten, Jagannathan and Runkle (1993) will be used, which is another way to model asymmetry.

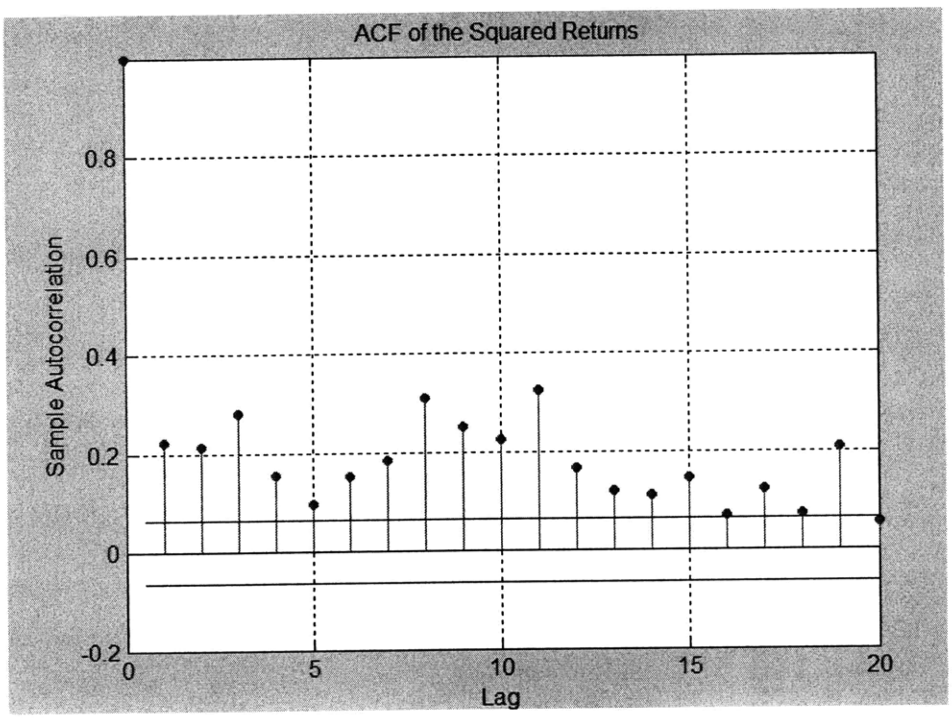
$$\sigma_t^2 = K + \delta\, \sigma_{t-1}^2 + \alpha\, \epsilon_{t-1}^2 + \phi\, \epsilon_{t-1}^2 I_{t-1}$$

Where $\epsilon_t = \sigma_t z_t$, z_t is i.i.d and $I_{t-1} = 0$ if $\epsilon_{t-1} \geq 0$, and $I_{t-1} = 1$ if $\epsilon_{t-1} < 0$.

7 GARCH EMPIRICAL RESULTS

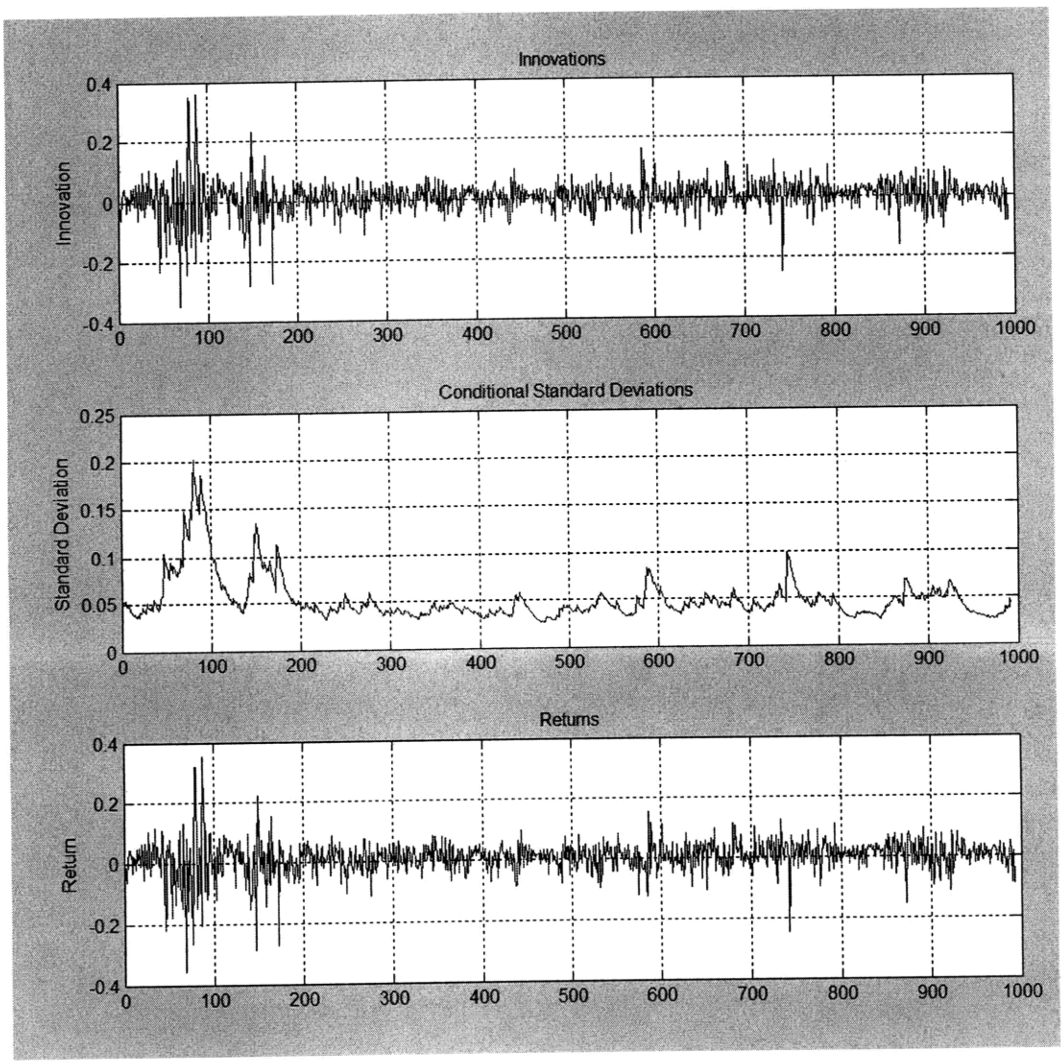
The GARCH approach will also forecast a 3 month and 12 periods ahead of the recent credit crisis (forecasts beginning on October 2008 and December 2008). This is to provide a baseline period to every method tested. The GARCH approach uses the methodology stated in the previous section. Instead of using different time ranges, 2 different orders were tested for each GARCH, E-GARCH, and GJR GARCH models. The results indicate that the GARCH (1, 1) model was the most significant. Furthermore the credit crisis forecasting was tested with the order P=1, Q=1, with both a mean constant and without for each model. The first step of analysis was to check for correlation of the return series, by looking at the ACF (autocorrelation) and PACF (partial autocorrelation). The monthly returns did not exhibit any serial correlation. However, when the returns were squared, a slowly decaying pattern was revealed. The variance suffers from heteroskedasticity and is outside of the standard error. This can be validated with the ARCH and Q-test. Both results showed that a high order of lags was significant and thus the null hypothesis of no heteroskedasticity is rejected.

Figure 13 - ACF of the Squared Returns



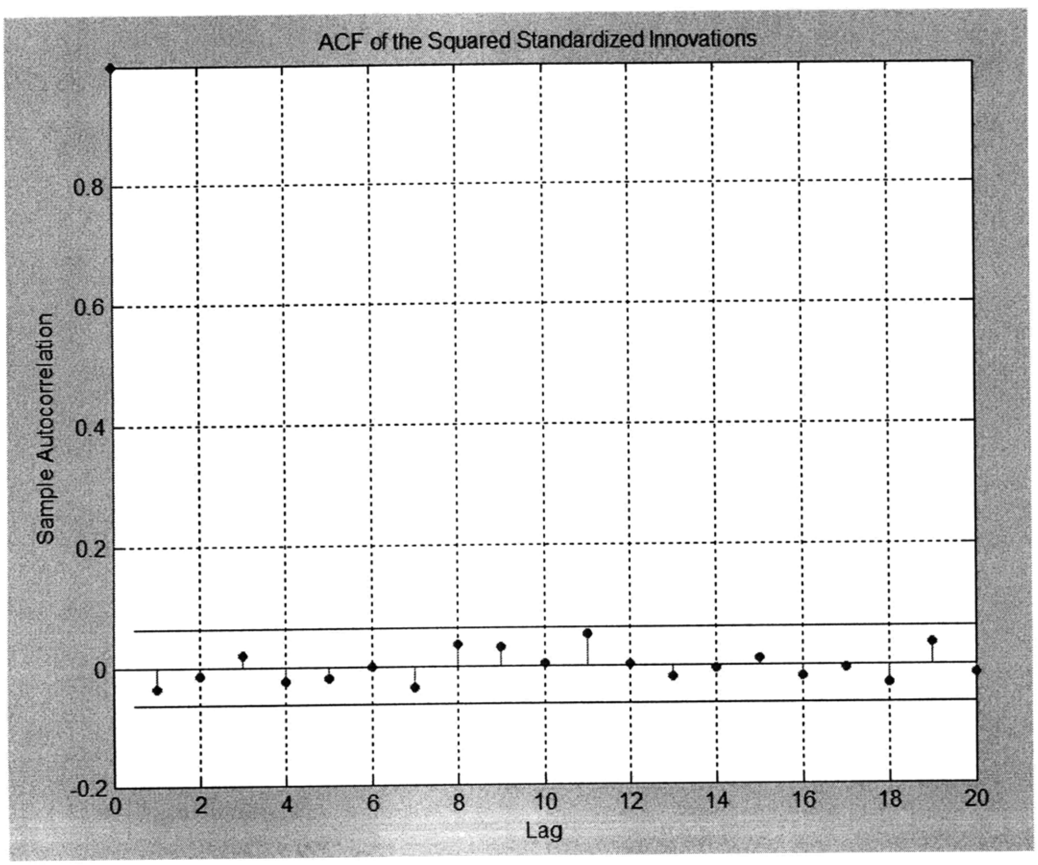
Because a high order of the ARCH (q) test was found in several lags over 10, the selection of a GARCH (1, 1) model becomes appropriate which acts as a parsimonious alternative to an infinite ARCH (q) process. After the order has been identified, the fitting of the models and parameter estimation was performed. MATLAB performed all of the estimations. These estimates are in the appendix section. Figure 14 displays a GARCH plot of the innovations (errors), a sigma plot derived from the conditional variance equations, and the S&P 500 returns. Notice the high volatility that occurred during the Great Depression with monthly returns as high as 40% and as low as 40%. The credit crisis has often been compared to have volatility "near" depression levels. But close examination show that the credit crisis monthly returns (near the end of the standard deviation/return graph) is not nearly as volatile as the Depression era. However, the graph data ends at September 30th 2008, and in October, -18% returns were posted increasing the volatility. Also, if we graphed daily volatility, we would see much more activity.

Figure 14 - A GARCH plot of the Innovations, Conditional STD Dev, S&P 500 Returns dated 12/1925 to 09/2008.



After the models have been fitted and forecasted, a plot of the squared standardized innovations was created. A standardized innovation plot is composed of each error at a given time and divided by its respective standard deviation derived from the GARCH conditional variance equation. The plot below shows that the GARCH (1, 1) model indeed adjusted for heteroskedasticity.

Figure 15 - ACF of the Squared Standardized Innovations



7.1 GARCH Forecasting Results of the Credit Crisis

Table 16 shows the GARCH forecasting results, similar to the ARIMA and Exponential smoothing results. GARCH models do not predict well for the short term forecast, but show promise as a long term model (see the next section on the Black Monday modeling). However, the GARCH model did adjust for signals of increased volatility. The EGARCH(1,1) has the best 3 month forecast with the lowest MAE of 229. Overall, the forecasts between all three models were fairly close to one another. The EGARCH(1,1) model shows an optimistic 12 month forecast of 1095.

The GARCH models depend on previous volatility to adjust their forecasts. When the 3 month data of September, October, and December 2008 were included in the dataset, the forecasts changed dramatically. 12 month forecasts ending in December 2009 are substantially bearish showing the index price to be at 779 for the EGARCH model. Also note that these models were not fitted with a constant in the mean. So the ARMA terms have a downward pressure on the forecasted returns and price index.

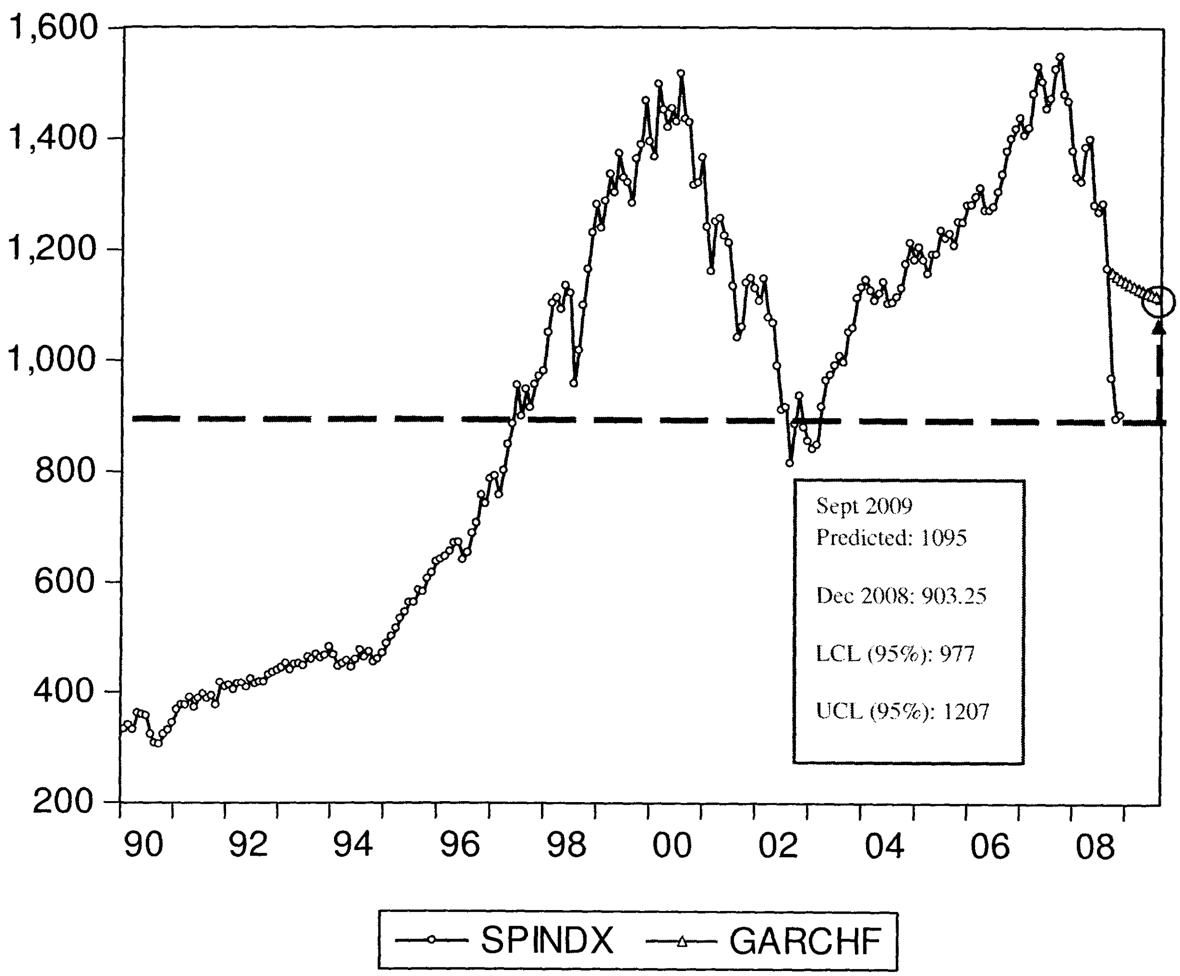
When a constant was included in the model testing, the results were positive and the parameter estimation was significant as well. See the appendix for these estimations. The mean incorporates a positive drift that is inherited from the long run mean of the data set. Whether or not the constant model is a good forecasting fit is undetermined with the credit crisis data, but forecast well in the Black Monday scenarios. The next section will validate and back test these GARCH models with the Black Monday stock market crash to further evaluate the forecasting results.

Table 16 – GARCH Model of Credit Crisis, S&P 500 Index 3 Month and 12 Month Forecast

		Actual	Actual	Actual			
		968.75	**896.24**	**903.25**			
	Fit Range	**Predicted Month 1**	**Predicted Month 2**	**Predicted Month 3**	**3 Mon Forecast Error**	**3 Mon Forecast MAE**	**12 Mon Forecast**
ARMAX(1,1,0) + GARCH(1,1)	12/1925 - 09/2008	1159.68	1153.58	1147.98	693.00	231.00	1114.57
ARMAX(1,1,0) +EGARCH(1,1)	12/1925 - 09/2008	1158.62	1151.28	1144.30	685.96	228.65	**1095.00**
ARMAX(1,1,0) +GJR(1,1)	12/1925 - 09/2008	1159.26	1152.59	1146.32	689.92	229.97	1104.00
ARMAX(1,1,0) + GARCH(1,1)	12/1925 - 12/2008	891.89	881.68	872.51	N/A	N/A	821.52
ARMAX(1,1,0) +EGARCH(1,1)	12/1925 - 12/2008	888.17	874.21	861.25	N/A	N/A	778.72
ARMAX(1,1,0) +GJR(1,1)	12/1925 - 12/2008	890.12	878.07	867.00	N/A	N/A	799.33
C + ARMAX(1,1,0) + GARCH(1,1)	12/1925 - 12/2008	908.62	914.56	920.08	N/A	N/A	973.41
C +ARMAX(1,1,0) +EGARCH(1,1)	12/1925 - 12/2008	908.47	913.63	918.81	N/A	N/A	966.79
C + ARMAX(1,1,0) +GJR(1,1)	12/1925 - 12/2008	908.15	913.62	918.66	N/A	N/A	967.43

The following graph shows the 3 month difference between the actual and the EGARCH forecast. The magnitude of the monthly returns volatility during September and the previous months was not high enough to capture October's -18% drop. *The 12 month point estimate is 1095 with a 95 % lower confidence interval of 977 and an upper confidence interval of 1207.*

Figure 17 – GARCH Model of Credit Crisis, S&P 500 Index 3 Month and 12 Month Forecast



7.2 Modeling Black Monday and Backtesting GARCH Models

Historically, Black Monday has been one of the single largest daily drops in history which occurred on October 19, 1987.[xxii] The Dow Jones Index had the largest drop of 22.6%, and the S&P 500 fell by 20.4%. This steep drop attributed to further decline and the S&P 500 posted a 24.53% decline at the end of the month. In November 5, 1987, President Ronald Regan issued an executive order to examine the events surrounding Black Monday in order to understand what happened, why it happened, and how to prevent such an event from occurring in the future. The report indicated that selling pressure during October 14 to 16 was the cause of 2 major events[xxiii]:

- Disappointing poor merchandise trade figures, which put pressure on the dollar in currency markets and upward pressure on long term interest rates.

- Filing of anti-takeover tax legislation, which caused risk arbitrageurs to sell stocks of takeover candidates resulting in a downward ripple effect throughout the market.

Moreover, the reactive selling primarily involved a small group of participants, portfolio insurers and a few mutual fund groups. The portfolio insurers typically used computerized insurance models that calculated sales in excess of 20 percent in response to a 10 percent decline in the market (Brady 1988). For one particular insurance client the models indicated that 70% of the portfolio's remaining equity should be sold by the close of Friday (October 16th). Overall, insurers had $60 to $90 Billion of equity assets that were under administration during the market break. Insurers sold $530 million on Wednesday, $965 Million on Thursday, and $2.1 billion on Friday primarily in the futures market. Mutual Funds faced customer redemptions, and by Friday, customer redemptions at these funds exceeded fund sales by $750 million. On Black Monday alone, total equity sales for the day totaled $19.1 billion in the equity sales and $19.8 billion in futures sales. In the midst of the frenzy selling amongst the handful of participants, there were little buyers. Index arbitrageurs pulled away from purchasing discounted futures when the overwhelming selling created a complete disconnect in pricing information between futures and equity markets. This disconnect created a freefall in markets, allowing historical discounts and substantial downward price movements in both stock and futures markets. For more detailed and historical information on Black Monday, please refer to "Report of The Presidential Task Force of Market Mechanisms" (Brady 1988).

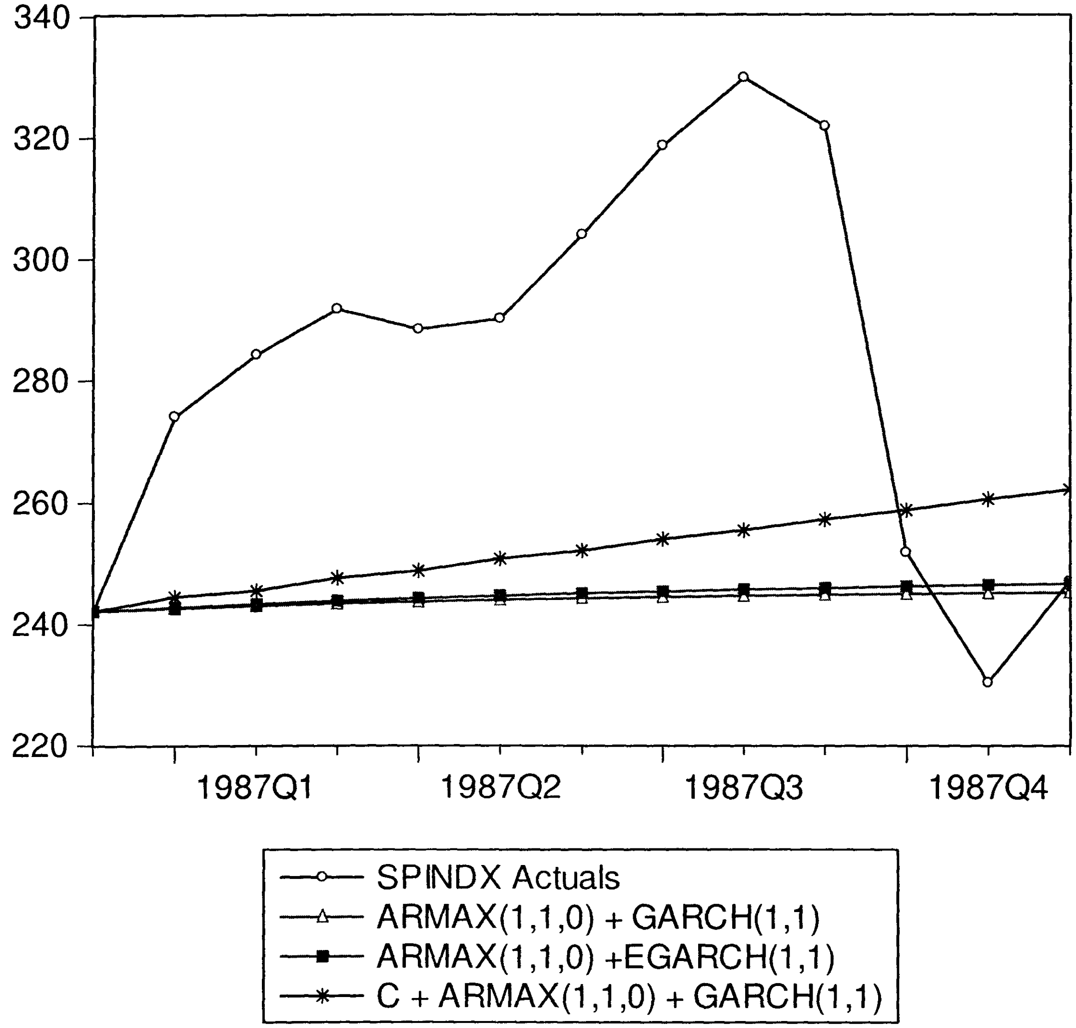
Epistemologist and author Dr. Nassim Nicholas Taleb considers this drop a "black swan" event, which is a large impactful event that is rare and in many instances improbable, and an event that is taken completely by surprise. Similar to the credit crisis, this event happened on a six sigma (if not greater) scale. Black Monday spurred a recession afterwards and dampened the S&P 500 price. After the crash, the index did not recover until May of 1989, almost 2 years later. Dr. Taleb's black swan theory does hold some truth in attempting to use statistical models to fit data prior to the October 19^{th} to predict the actual crash date and month. From a short term perspective, it was extremely difficult to account for the crash with the GARCH models. No previous time data could give a volatility magnitude high enough for the GARCH models to come within a significant interval (5-10%) of the return and level drop. However, forecasting a 12 month and longer term estimates has shown more success. The model that provided the best point estimate was the ARMAX (1, 1, 0) no constant for the conditional mean and EGARCH (1, 1) for the variance. This model did not have the lowest MAE, but it had a 12 month point estimate residual of .4, which was very close to the actual. When a constant was used for the conditional mean, the GARCH (1, 1) provided the lowest MAE, and served overall as the best fitted model. Note that all the GARCH models did not account for the accelerated levels shown in between months 2-9. These months account for the highest residual error, and the crash October readjusted the price level to the forecast, with monthly residuals less than 10. Overall, GARCH models are best used to model heteroskedasticity, but the conditional mean forecasting is most effective in 12 month or greater forecasts around steep recessionary and growth periods. GARCH simulation methods were also used, but overall the estimates converged to be very close to the minimum mean square error forecast used for our models. See the appendix for the GARCH model estimates.

Table 18 - Predicting Black Monday, 12 Month Forecast

	Fit Range	12 Mon Forecast Error	12 Mon Forecast MAE	Mo 1	Mo 2	Mo 3	Mo 4	Mo 5	Mo 6	Mo 7	Mo 8	Mo 9	Mo 10	Mo 11	Mo 12
			Start Jan 1987 (Actual)	274.1	284.2	291.7	288.4	290.1	304.0	318.7	329.8	321.8	251.8	230.3	247.1
ARMAX(1,1,0) + GARCH(1,1)	12/1925 - 12/1986	531.4	44.3	242.6	243.0	243.4	243.7	244.0	244.2	244.4	244.6	244.8	245.0	245.1	245.2
ARMAX(1,1,0) +EGARCH(1,1)	12/1925 - 12/1986	523.9	43.7	242.8	243.3	243.8	244.3	244.7	245.1	245.4	245.7	246.0	246.2	246.5	246.7
ARMAX(1,1,0) +GJR(1,1)	12/1925 - 12/1986	529.4	44.1	242.7	243.1	243.5	243.8	244.2	244.4	244.7	244.9	245.1	245.3	245.5	245.6
C + ARMAX(1,1,0) +GARCH(1,1)	12/1925 - 12/1986	499.4	41.6	244.5	245.5	247.6	248.8	250.7	252.1	253.9	255.4	257.2	258.8	260.6	262.2
C + ARMAX(1,1,0) +EGARCH(1,1)	12/1925 - 12/1986	506.3	42.2	243.2	244.6	246.1	247.5	249.0	250.4	251.9	253.4	254.9	256.4	257.9	259.4
C + ARMAX(1,1,0) +GJR(1,1)	12/1925 - 12/1986	501.6	41.8	244.3	245.1	247.0	247.9	249.7	250.8	252.5	253.7	255.3	256.6	258.2	259.5

Figure 19- GARCH Forecast of Black Monday, 12/1985 – 12/1987

The GARCH models provided an *acceptable* 12 month point estimate largely due to *mean reversion.* However, it did *not perform well for short term and interim months, missing the accelerated peaks.* The constant model adjusted for a better short term forecast, but not by a wide margin compared to the non-constant model.



8 NEURAL NETWORK FORECASTING AND HYBRID METHODOLOGY

8.1 What is a Neural Network?

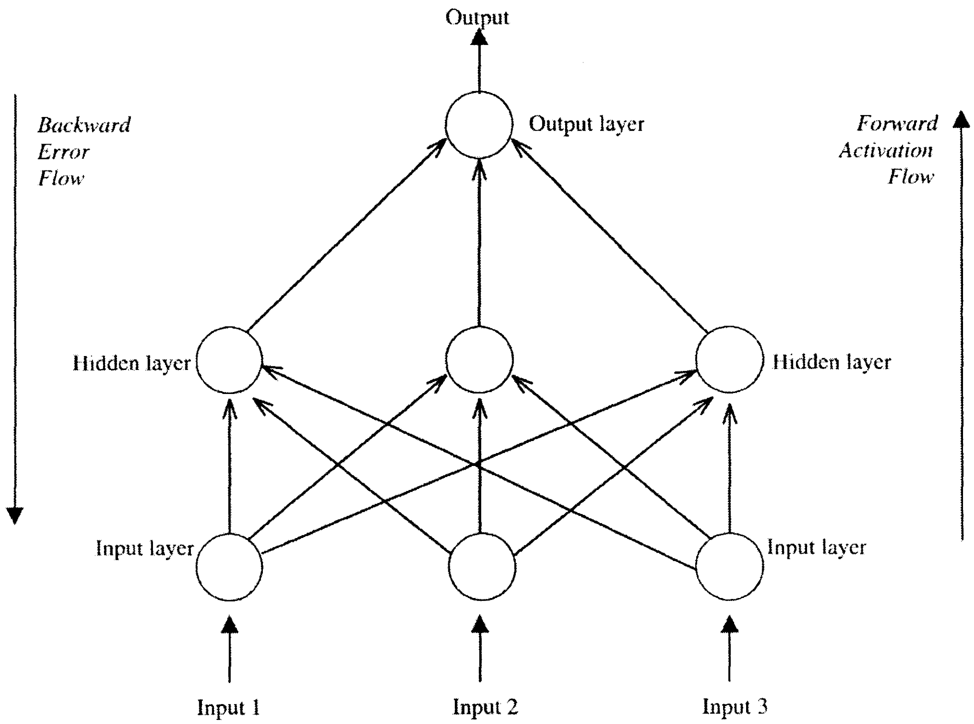
A Neural Network is a computational method that mimics how the human brain processes information and data. Hamid and Iqbal cite that the network imitates the structure and operations of a 3 or N dimensional lattice of networks among brain cells, referred to as nodes or neurons.[xxiv] There have been variations of Neural Network architectures proposed, but every network has the following characteristics:

- A network is composed of certain *number of layers* indicating which input variables to be processed. Take for example, that a typical Neural Network would have at least three layers: an input layer, a hidden layer, and an output layer.
- Within a layer, there can be 1 to many nodes or neurons within a layer. Each neuron is responsible for learning the relationship between input and output nodes. Also, each node is considered a variable.
- Each neuron has *connections between neurons*, often forming a many to many relationship between nodes in a downstream layer. For example if an input layer has 3 nodes, and if the hidden layer has 3 nodes as well, then 3! connections are created.
- A network or layers can have a *transfer function* which seeks to determine the relationship between the input and output data.

Think of a layer as an aggregation of nodes. The first layer is the input layer which is where the input data is defined and processed. Each neuron or node can be considered to have an input variable. You can also look at the input neurons as explanatory variables in a regression. The hidden layer can be considered as a constructed set of variables (Montgomery 2008). This process explores patterns and correlations to make generalizations that will influence the output. This layered architecture is typically called a feed-forward multi-layered network, also known as *backpropagation*. Figure 20 is an example of this architecture.

Figure 20 - A 3x3x1 Feed-Forward Back Propogation Neural Network

Image Source "Using neural networks for forecasting volatility of the S&P 500 Index Future Prices", S.Hamid, Z. Ibqal

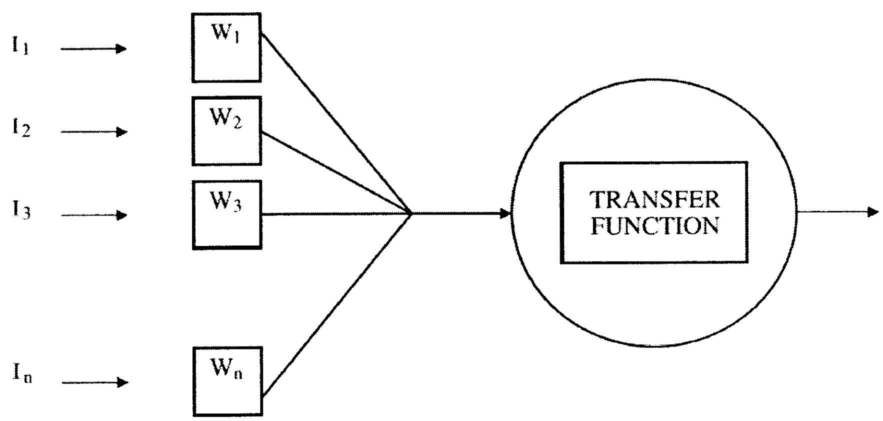


The neurons can either be fully connected or partially connected. Figure 20 shows an example of a fully connected 3x3x1 network. All nodes in the input layer are connected to all hidden neurons, and all hidden neurons are connected to the output. Each connection has an *initial weight*, which are generated by small random initial seed values. These weights are "trained" during subsequent runs to create the optimal value. Moreover, algorithm of the backpropagation (a.k.a. generalized delta rule) provides the "learning rule" through which the network changes its connection weights during training until the optimized value is found. Training is synonymous to parameter estimation, and the usual approach is to estimate the parameters by minimizing the residual sum of squares taken over all responses and all observations. This procedure is a nonlinear least squares problem, and there are several algorithms to do this. I refer you to Tang and Fishwick, they give more detail in the backpropagation algorithm as an example.[xxv]

The transfer function, a.k.a. activation function is responsible for transforming the product of the input and weights into an output. Input data ($I_1, I_2, \ldots, I_n$) are multiplied by the weights ($W_1, W_2, \ldots, W_n$) associated with the connection of the weights. Then this product is used as an input to the transfer function, which in turn converts the sum into an interval between 1 and -1 or 1 and 0. This process is recursively repeated until the next node becomes the output layer. At the last output node this transfer function output becomes the predicted response. Figure 21 is an example of the product and transformation flow.

Figure 21 – The Transfer Function in relation to input variables

Image Source "Using neural networks for forecasting volatility of the S&P 500 Index Future Prices", S.Hamid, Z. Ibqal



The typical transformation function for the hidden layer is the logistic function:

$$g(x) = \frac{1}{1 + \exp(-x)}.$$

Other functions can be chosen such as Gaussian, Hyperbolic tangent, cosine, and sine transfer are all valid options. In the case of a single hidden layer feedforward network that is used for time series forecasting the relationship between the output (y_t) and the inputs ($y_{t-1}, y_{t-2 \ldots} y_{t-p}$) can be represented with the equation below. Where α_j ($j = 0,1,2,\ldots,q$) and β_{ij} ($i = 0,1,2,\ldots,p; j = 1,2,\ldots,q$) are the connection weights, (p) is the number of input nodes and (q) is the number of hidden nodes.

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j g \left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} y_{t-i} \right) + \varepsilon_t,$$

The Neural Network model does have limitations. Deciding the number of layers and nodes is still a trial and error process. Unlike the ARIMA process which uses the Box-Jenkins approach in determining the appropriate model, Neural Networks is an immature technology and lacks formal theory in model determination. Moreover, there is no analytical formula to determine the best networking model. The actual number of layers and nodes will depend on the complex problem being solved. This implies that users will need to be sophisticated to choose the appropriate network architecture and input variable selection. Perhaps, the largest issue with Neural Network is *overfitting*. Take for example, if too many nodes and layers are chosen, the prediction will resemble a near perfect fit to the response, making the prediction ineffective. Overfitting can be addressed by stopping the parameter estimation process before convergence, reducing unknown parameters, and using cross validation to determine the number of iterations to use (Montgomery 2008). Lastly, Neural Networks are a closed architecture, and debugging and decomposing the output is not possible.

8.2 Hybrid ARIMA/Exponential Smoothing Models with Neural Network Residual Modeling

The hybrid approach will adopt Peter Zhang's method of combining ARIMA Time Series forecasting with Neural Networks. Zhang asserts that ARIMA models do not handle complex nonlinear problems adequately. In addition, Artificial Neural Nets (ANN) do not consistently fit linear problems well, yielding mixed results. He cites an example of ANN outperforming linear regressions when outliers and multicollinearity exists in the data, but there are dependencies on sample size and noise level [Zhang (iii)]. *An approach is recommended to combine both ARIMA and Neural networks, where ARIMA models will capture the linear autocorrelations of a time series and Neural Networks will fit the nonlinear aspects.* Mathematically, it can simply be represented as the following:

$$y_t = L_t + N_t,$$

Where L_t denotes the linear component and N_t denotes the nonlinear component. The combination of the model will work as follows:

- Model and fit the data with the ARIMA model first.
- The residuals from the ARIMA model represent the nonlinear relationship.

$$e_t = y_t - \hat{L}_t,$$

Where e_t denotes the residual at time t from the ARIMA model.

- Take the residuals and model the ANN with n input nodes.

$$e_t = f(e_{t-1}, e_{t-2}, \ldots, e_{t-n}) + \varepsilon_t,$$

Where f is a nonlinear function determined by the ANN and ε_t is the random error. Note that if f is not appropriate, then the error may not necessarily be random.

- The combined forecast of the ARIMA and ANN prediction will simply be:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t.$$

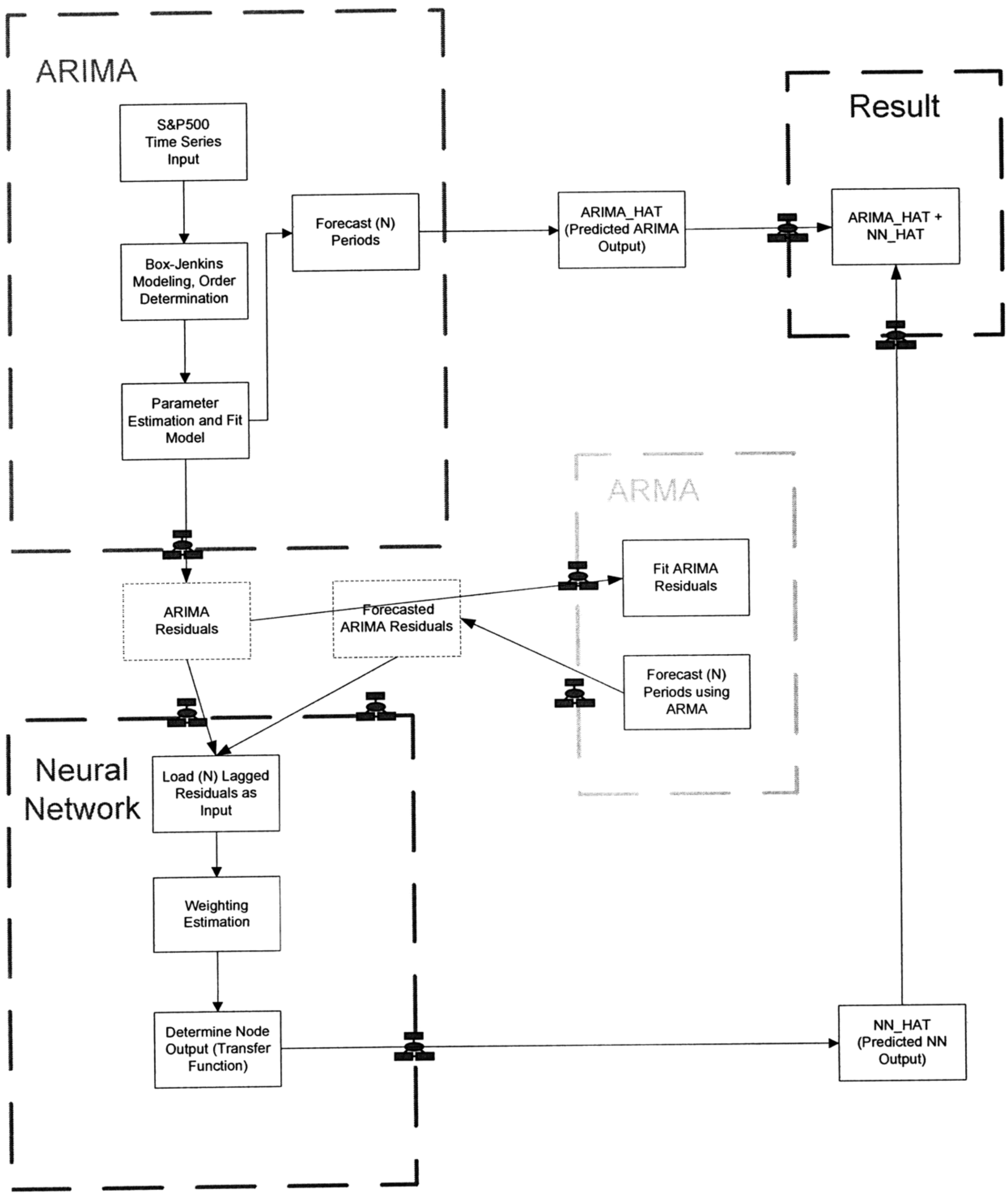
Zhang does not really clarify on how to determine the number of input nodes for the ANN residual modeling. As a rule of thumb, from the empirical research, the number of residual lags should equal to the forecast horizon. For example, forecasting a 12 period horizon should have 12 residual lags with each lagged time series being (t – 1, t – 2, t – 3, ...) used as an input. Note that this heuristic rule isn't strict and is an initial state to start the model. Also, the results from the ANN can be used as predictions of the error terms for the ARIMA model. Overall the hybrid model aims to capture the unique characteristics of both linear and nonlinear components through ARIMA and ANN respectively to produce a better time series forecast.

Lastly, in the next section we will apply this hybrid method to the Exponential Smoothing models, replacing the linear ARIMA component with the Double Exponential Smoothing. The intuition comes from the fact that the exponential smoothing models outperformed the ARIMA model during the 3 month forecast in the steep price drop resulting from the credit crisis. The anticipation is that the exponential smoothing model combined with the ANN residual modeling will yield a superior forecast than the exponential smoothing model by itself, even with the possibility that it may capture some of the nonlinear characteristics of the time series.

9 NEURAL NETWORK FORECASTING AND HYBRID EMPIRICAL RESULTS

Forecasting the credit crisis using Artificial Neural Networks is fundamentally different from the other methods used in the previous sections. Similar to a regression, all input variables have to be defined for all time periods. Also, forecasting multi period steps ahead without any input data for that given time is often a constant value and is an ineffective forecast. In Figure 22, an architectural process and component map lays out the high level model components and processes that realized the hybrid methodology discussed in the previous section. Input and Output interfaces are represented with this symbol: . For example, ARIMA residuals are an output from the ARIMA component, and an input to the Neural Network and ARMA component. The boxes within the components (inside the dotted line) reflect primary internal processes used to produce a statistical output. Since, the Neural Network provides marginal forecasting without the forecasting data, the ARMA model was used to provide a forecast for the ARIMA residuals. This straight forward model forecasted 12 month periods within the range of (10/2008 – 09/2009). *These forecasted residuals in addition to the ARIMA residuals* served as the input for the Neural Network model. The predicted and forecasted ARIMA output is denoted by ARIMA_HAT, and the Neural Network predicted and forecasted residuals are denoted by NN_HAT. Finally, in following Zhang's methodology the final prediction result is ARIMA_HAT + NN_HAT. As a validation of the prediction, ARIMA_HAT (t) + NN_HAT (t) + NN_Residuals (t) = S&P500 (t). Overall, the fit was good with insignificant differences between the actual and the predicted values.

Figure 22 - The process and component architecture of the Hybrid ARIMA and Neural Network Forecast



9.1 Neural Network and Hybrid Forecasting Results of the Credit Crisis

Several Neural Network models were used to forecast the credit crisis. As a baseline of comparison the ARIMA (2,1,3) and the Log Double Exponential Smoothing model results from the previous empirical time series section was included. The 3 (input) x 7 (hidden) x 1(output) Neural Network model is a pure S&P 500 time series fit. Moreover, it used all the observations from the fit range from the S&P 500 and used the forecasted values from the ARIMA (2, 1, 3) model as the input. This Neural Network model did not incorporate residual lags in the input layer and is also considered as a baseline comparison to the other Neural Network models. The baseline Neural Network model (3x7x1) overall underperformed against all other models. Different data sample partitions were used: one partition contained 100% training data, another partition contained 70% training, 15% validation, and 15% test data sample which also underperformed. Overall, the best performing model was the ARIMA + Neural Nets Residual (7x7x1). This model had the best 3 month forecast with a MAE of 147.87, which outperformed all the other baseline models. The 12 month point estimate is 1017.54 with a *95% confidence interval between 1147 and 886.54*. The third best model was the Log Double Brown Exp + NN Residuals (12x3x1). This model had a 3 month MAE of 163.45, which was surprising in this case. Typically, through the previous the Log Double Brown performs better in the short run forecast than the ARIMA. However, the Neural Network residuals used to model this did not have the same negative magnitude as the ARIMA + NN model. *As a standalone model the Double Exponential Smoothing model will outperform in a short term forecast, but the empirical results show that as a hybrid Neural Network model the ARIMA + NN model outperforms.* In addition, the exponential smoothing + NN model has a 12 month forecast of 803.81. This forecast could potentially be underpriced due to the exponential behavior of the model, with the assumption of a positive outlook and return for the S&P 500 on 09/2009. To adjust for this view, a combination forecast can be used between ARIMA+NN and Double Exponential Smoothing + NN with equal weighting on the forecasts (.5). This combination forecast will give a weighted forecast of 910.68, which is essentially the mean between the two models. One final observation is that the ARIMA (2, 1, 3) model from the previous empirical study of the Tech Bubble revealed to have the best 12 month point estimate. However, the hybrid forecast gives another point estimate to consider when evaluating forecasting estimates. The appendix will have the ARIMA parameter estimations and as well as the ARIMA+NN weighting estimates. The next section will validate these models in the 73 to 74 Oil Crises.

Table 23 – Neural Network Residual Credit Crisis Forecast

			Actual 968.75	Actual 896.24	Actual 903.25			
	Fit Range	**Eval Range**	**Predict Month 1**	**Predict Month 2**	**Predict Month 3**	**3 Mon Forecast Error**	**3 Mon Forecast MAE**	**12 Mon Forecast**
ARIMA (2,1,3)	12/1925 - 09/2007	10/2007 - 9/2008	1160.00	1157.00	1147.00	695.76	231.92	1017.00
Log Double (Brown) Exponential Smoothing	1/2000 - 9/2007	10/2007 - 9/2008	1160.00	1128.00	1096.00	615.76	205.25	849.51
ARIMA + NN (7x7x1) Residuals	12/1925 - 09/2008	12/1925 - 09/2008	1037.25	1079.40	1095.21	443.62	147.87	1017.54
(3x7x1) NN 100% Training + ARIMA(2,1,3) Forecast	12/1925 - 09/2008	12/1925 - 09/2008	1148.65	1183.61	1149.04	713.06	237.69	999.37
(3x7x1) NN, 70% Training + ARIMA(2,1,3) Forecast	12/1925 - 09/2008	12/1925 - 09/2008	1164.38	1173.48	1165.53	735.15	245.05	1038.10
(12x3x1) Log Double Brown Exp + NN Residuals	12/1925 - 09/2008	12/1925 - 09/2008	1139.20	1101.55	1017.85	490.36	163.45	803.81
Combined Forecast Log Double Exp (12x3x1) + ARIMA (7x7x1)	12/1925 - 09/2008	12/1925 - 09/2008	1088.22	1090.48	1056.53	466.99	155.66	910.68

9.2 Modeling the 73-74 Oil Crisis using Neural Network Models

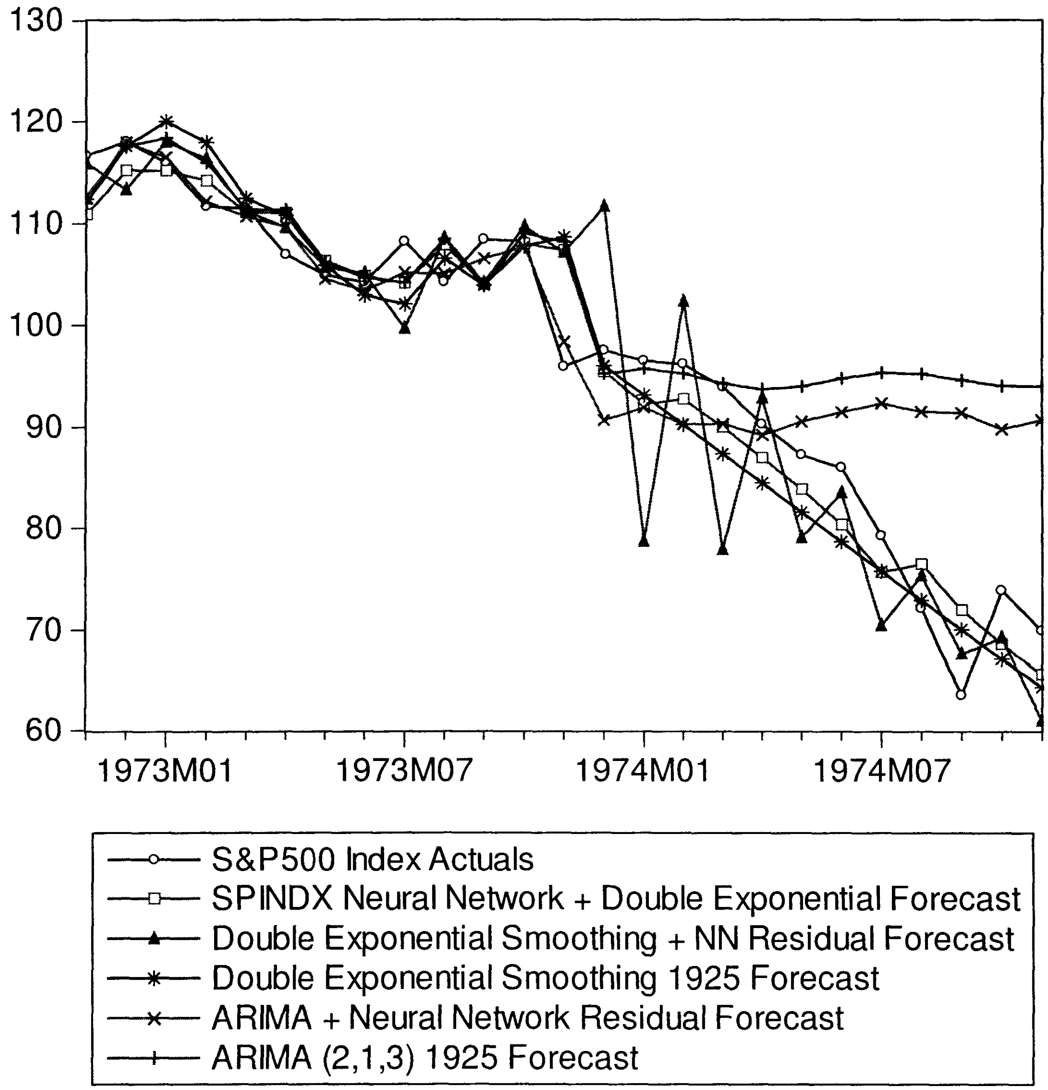
As a brief introduction to the oil crisis, Arab members of the OPEC (Organization of Petroleum Exporting Countries) placed an oil embargo on the United States in response to their decision to resupply the Israeli military during the Yom Kippur War.[xxvi] The act of this embargo and the changing of nature of oil contracts caused an upward spiral in oil prices that had a global influence. The prices first doubled and then later quadrupled, leading to increased costs worldwide (U.S. Department of State 1974). The United States, facing growing oil consumption and increased reliance on imported oil was a prime factor in the dramatic price inflation. The oil crises eventually spurred the 1973-1974 stock market crash, which will be modeled with the S&P 500 Index. This market crash is often regarded as the first event since the Great Depression to have a persistent recessionary effect.[xxvii] From January 1973 to December 1974 the Dow Jones Index had lost over 45% of its value, which is similar to the credit crisis which lost the same amount within a much shorter time frame (3-5 months). Forecasting the S&P 500 Index during the Oil Crisis was an interesting exercise. The model parameters were first estimated starting at 1925 (baseline ARIMA and Exponential Smoothing data sets), but 20 year data sets starting in 11/1954 proved to be a better statistical fit. The double exponential smoothing model overall yielded excellent results that were very close to the 12 month actual value (See table 24). Unfortunately, using a Double Exponential + Residual model made the forecast worse with a MAE of 8.09 versus 5.03. The best performing model was the (3x7x1) Neural Network + Double Exponential Forecast model. Combining the forecast and using Neural Networks to reweight and refine the smoothing estimate proved to be a better prediction model with the lowest MAE of 4.34. The hybrid ARIMA + NN residual and related models underperformed for the 12 month forecast. However, the ARIMA model which had the highest MAPE *still outperformed* the lowest MAPE of all the models compared to the 3 month forecast previously in table 23. Furthermore, the Oil Crisis *highest* 12 month MAPE outperformed: 15.72% versus 16.25% of the *lowest* 3 month MAPE of the credit crisis. Also note that the ARIMA + NN residual model outperformed the standalone ARIMA model. In conclusion, the hybrid NN residual models yielded mixed forecast results, but overall Neural Network models provide acceptable forecasting results and should not be overlooked. See appendix for parameter estimates.

Table 24 – Predicting the 1973 Oil Crisis with Neural Networks Model, 12 Month Forecast

	Fit Range	12 Mon Forecast Error	12 Mon Forecast MAE	Mo 1	Mo 2	Mo 3	Mo 4	Mo 5	Mo 6	Mo 7	Mo 8	Mo 9	Mo 10	Mo 11	Mo 12
			Start Dec 1973 (Actual)	97.55	96.57	96.22	93.98	90.31	87.28	86	79.31	72.15	63.54	73.9	69.97
ARIMA (2,1,3)	12/1925 - 11/1973	138.07	11.51	95.26	95.78	95.29	94.33	93.76	94.03	94.82	95.40	95.30	94.67	94.11	94.09
Log Double (Brown) Exponential Smoothing	12/1925 - 11/1973	59.42	4.95	96.04	93.16	90.27	87.38	84.50	81.61	78.73	75.84	72.95	70.07	67.18	64.29
ARIMA (2,1,3)	11/1954 - 11/1973	135.97	11.33	95.43	96.07	95.31	93.99	93.33	93.81	94.87	95.52	95.24	94.40	93.79	93.92
Log Double (Brown) Exponential Smoothing	11/1954 - 11/1973	60.37	5.03	96.00	93.10	90.19	87.28	84.38	81.47	78.57	75.66	72.75	69.85	66.94	64.03
ARIMA + NN (7x7x1) Residuals	11/1954 - 11/1973	127.98	10.66	90.68	91.93	90.28	90.35	89.24	90.56	91.50	92.36	91.56	91.42	89.82	90.74
(3x7x1) NN 100% Training + Double Forecast	11/1954 - 11/1973	52.08	4.34	95.47	92.28	92.78	90.04	87.01	83.89	80.41	75.71	76.53	72.00	68.67	65.60
(12x3x1) Log Double Brown Exp + NN Residuals	11/1954 - 11/1973	97.13	8.09	111.66	78.74	102.39	77.98	92.94	79.13	83.56	70.48	75.40	67.68	69.28	61.01

The graph shows an overall good fit for most of the models. During the prediction months (12/73 – 11/74) a divergence is seen with the ARIMA (2, 1, 3) models. Moreover, the Double Exponential smoothing + NN residual prediction becomes more volatile. SPINDX NN + Double Forecast outperform all other prediction models, showing a slight improvement over the standalone Double Exponential Smoothing model. See appendix for parameter estimates.

Figure 25 - **Oil Crisis, Prediction Time series 11/1972 – 11/1974**



10 BOOTSTRAPING SIMULATION FORECASTING METHODOLOGY

10.1 Bootstrapping Simulation Methodology

Bootstrapping simulation is a nonparametric method that differs from other methodologies used in the previous sections. To generate a series of random movements, bootstrapping uses real asset returns from historical data. Paul Wilmott (2007) introduces a technique to bootstrap financial time series[xxviii]:

- Determine N assets with X periods of data. To illustrate a simple example we choose N=3 securities: Google, Microsoft, and Yahoo and X=13 periods of monthly data ranging from 12/31/2007 to 12/31/2008.

DATE	GOOG	MSFT	YHOO
31-Dec-08	307.65	19.44	12.2
28-Nov-08	292.96	20.22	11.51
31-Oct-08	359.36	22.33	12.82
30-Sep-08	400.52	26.69	17.3
29-Aug-08	463.29	27.29	19.38
31-Jul-08	473.75	25.72	19.89
30-Jun-08	526.42	27.51	20.66
30-May-08	585.8	28.32	26.76
30-Apr-08	574.29	28.52	27.41
31-Mar-08	440.47	28.38	28.93
29-Feb-08	471.18	27.2	27.78
31-Jan-08	564.3	32.6	19.18
31-Dec-07	691.48	35.6	23.26

- Convert the X daily data into M returns ($\ln X_t / \ln X_{t-1}$). The table below shows the conversion of the price levels to M=12 returns.

DATE	GOOG	GOOG_RET	MSFT	MSFT_RET	YHOO	YHOO_RET
31-Dec-08	307.65	0.049	19.44	-0.039	12.2	0.058
28-Nov-08	292.96	-0.204	20.22	-0.099	11.51	-0.108
31-Oct-08	359.36	-0.108	22.33	-0.178	12.82	-0.300
30-Sep-08	400.52	-0.146	26.69	-0.022	17.3	-0.114
29-Aug-08	463.29	-0.022	27.29	0.059	19.38	-0.026
31-Jul-08	473.75	-0.105	25.72	-0.067	19.89	-0.038
30-Jun-08	526.42	-0.107	27.51	-0.029	20.66	-0.259

30-May-08	585.8	0.020	28.32	-0.007	26.76	-0.024
30-Apr-08	574.29	0.265	28.52	0.005	27.41	-0.054
31-Mar-08	440.47	-0.067	28.38	0.042	28.93	0.041
29-Feb-08	471.18	-0.180	27.2	-0.181	27.78	0.370
31-Jan-08	564.3	-0.203	32.6	-0.088	19.18	-0.193
31-Dec-07	691.48		35.6		23.26	

- Assign an index for each M return. The index will refer to the cross section period of all N returns. In the example below an M x N (excluding index) matrix is constructed. Note that in index 1 refers to a row vector of cross section returns of Google, Microsoft, and Yahoo. This will maintain the correlation of returns between the assets for that specific index.

INDEX	GOOG_RET	MSFT_RET	YHOO_RET
1	0.049	-0.039	0.058
2	-0.204	-0.099	-0.108
3	-0.108	-0.178	-0.300
4	-0.146	-0.022	-0.114
5	-0.022	0.059	-0.026
6	-0.105	-0.067	-0.038
7	-0.107	-0.029	-0.259
8	0.020	-0.007	-0.024
9	0.265	0.005	-0.054
10	-0.067	0.042	0.041
11	-0.180	-0.181	0.370
12	-0.203	-0.088	-0.193

- Determine how many P periods to forecast/simulate forward, then draw a random number U from a uniformly distribution from 1 to M (total number of return periods) for each forecasting period. Map U to the index, and use the index number to reference the cross section return for N assets for each forecasting period. Following our example we determine that P=6 number of forecast periods. Next we randomly draw U for each period, which can range from 1 to 12. In this case the result is recorded in the Random Index row. The cross sectional returns referenced to that index is realized by that period column. For example, in period 2 the random index drawn was 6, this index references the cross sectional returns of -.105, -.067, -.038 for Google, Microsoft, and Yahoo respectively.

6 Month Forecast

	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6
Random Index	1	6	7	2	10	1
GOOG	0.049	-0.105	-0.107	-0.204	-0.067	0.049
MSFT	-0.039	-0.067	-0.029	-0.099	0.042	-0.039
YHOO	0.058	-0.038	-0.259	-0.108	0.041	0.058

- Repeat N Times, and save each result. In this example, if I assume that my portfolio is an equally weighted average of Google, Microsoft, and Yahoo, then I can take a simple arithmetic average for each trial. After 3 trials, I forecasted an Average of Averages return of 1% in period 1, -3.1% in period 2, -4.1% in period3, etc. Note that the Average of Averages is not part of Wilmott's original method, and it is introduced in the conditional bootstrapping method in the next section. This example is simplistic, a large number of trials (10,000+) and historical data (36+) are normally observed.

Trial 1

	Period 1	Period 2	Period 3	Period 4	Period 5	Period 6
GOOG	0.049	-0.105	-0.107	-0.204	-0.067	0.049
MSFT	-0.039	-0.067	-0.029	-0.099	0.042	-0.039
YHOO	0.058	-0.038	-0.259	-0.108	0.041	0.058
Average Trial #1	*0.023*	*-0.070*	*-0.132*	*-0.137*	*0.005*	*0.023*

Trial 2

GOOG	-0.180	-0.146	-0.067	-0.146	-0.022	-0.204
MSFT	-0.181	-0.022	0.042	-0.022	0.059	-0.099
YHOO	0.370	-0.114	0.041	-0.114	-0.026	-0.108
Average Trial #2	*0.003*	*-0.094*	*0.005*	*-0.094*	*0.004*	*-0.137*

Trial 3

GOOG	-0.180	0.265	-0.180	0.049	-0.180	-0.107
MSFT	-0.181	0.005	-0.181	-0.039	-0.181	-0.029
YHOO	0.370	-0.054	0.370	0.058	0.370	-0.259
Average Trial #3	*0.003*	*0.072*	*0.003*	*0.023*	*0.003*	*-0.132*

Average of Averages	**0.010**	**-0.031**	**-0.041**	**-0.069**	**0.004**	**-0.082**

10.2 Conditional Bootstrapping Simulation Methodology

To adjust for simulations in recessionary periods for the S&P 500, the following modification to Wilmott's bootstrapping is proposed:

- Determine a fixed time period of N time series that exhibit similar correlation structures or similar trend patterns. For example, S&P 500 36 month return snapshots of the Great Depression and the Credit Crisis. Moreover, this is equivalent to taking a conditional time range snapshot of the S&P 500, and classifying it as an "asset".

- Take the average or median of the N cross section return at the end of each simulation loop. In addition, you can take the Average of Averages by storing the mean of *each period* from the NxP matrix, and calculating the Average of all Averages after all iterations are complete. Please see the empirical results section for an example.

Closed form methods like the linear regression often violate the normality assumption for financial time series resulting in inaccurate estimation. Alternatively, bootstrapping makes no assumptions about the distribution and can accurately account for the joint distribution of several time series involved. Another advantage is that bootstrapping incorporates cross section correlations between chosen financial time series. Lastly, to implement the bootstrapping technique is fairly easy. One disadvantage is that the historical returns drawn may not accurately reflect the current economic environment. However, by using a conditional data set this disadvantage is less warranted. For example, modeling recessionary periods using time series from historical recessionary periods like the Great Depression will provide more accurate forecasts compared to using a random selection from the entire data set. This modified method assumes that a decision maker has a prior belief of the economy (bull versus bear), so the results would be inaccurate if the prior is incorrect, otherwise more accurate than unconditional simulation.

11 CONDITIONAL BOOTSTRAPING SIMULATION EMPIRICAL RESULTS

The approach was to simulate with different recessionary time periods in the S&P 500. Choosing the length of monthly periods and the appropriate time series is essential. As a guideline the recessionary period tables provided by the National Bureau of Economic Research was used as an initial evaluation.

Table 26 – U.S. Recessionary Periods

Peak month	Year		Trough month	Year
November	1948		October	1949
July	1953		May	1954
August	1957		April	1958
April	1960		February	1961
December	1969		November	1970
November	1973		March	1975
January	1980		July	1980
July	1981		November	1982
July	1990		March	1991
March	2001		November	2001

After evaluating each period, not every recessionary event had influence on the S&P 500. The analysis showed that the following time series was best suited for recessionary simulation:

Table 27 – Selected Empirical Recessionary Time Series

Event	Period Range	Length
The Great Depression	07/1929 - 06/1932	36 Months
The Oil Crisis	10/1972 - 09/1975	36 Months
Tech Bubble	01/2001 - 12/2002	36 Months

Other time series was chosen such as the recession in 1948 and 1957, but the economic recession seemed to have small impact on the S&P 500 returns, thus those time series were not used. The 36 month period range was chosen, because that was average length an economy to recover from a downturn. The Great Depression duration was above the mean, and the Tech Bubble and Oil Crisis was equivalent to the mean. Also, the data was constructed so that the first 3 observations would be prior to a "crash" return. This approach was adopted so that it could be comparable to the time series period length in the previous empirical sections. Notice that the cross section correlations co-moved together on the 4th and 5th observations.

Table 28 – First five cross section returns

Index #	Great Depression	Oil Crisis	Tech Bubble
1	0.045619	0.009317	-0.050904
2	0.097992	0.045617	-0.020108
3	-0.04888	0.011828	0.09672
4	**-0.199271**	**-0.017111**	**-0.030796**
5	-0.133747	-0.03749	-0.021915

Using this simulation approach assumes that the analyst or decision maker has a prior belief that the S&P 500 will be entering a bear market. Moreover, this approach is not appropriate if such a prior does not exist. In this scenario, using a random sample over a large sample of the recent observations is a better approach. To validate whether these time series co-moved together, the time series was tested for cointegration (Johansen), as an alternative to correlation testing:

Table 29 – Johansen Conintegration Test of Great Depression, Oil Crisis, Tech Bubble

Sample (adjusted): 3 36
Included observations: 34 after adjustments
Trend assumption: Linear deterministic trend
Series: DEP OIL TB
Lags interval (in first differences): 1 to 1

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.646648	68.83888	29.79707	0.0000
At most 1 *	0.503452	33.46897	15.49471	0.0000
At most 2 *	0.247463	9.666380	3.841466	0.0019

Trace test indicates 3 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

This test indicated that all time series involved were cointegrated, and that these time series are an appropriate choice for recessionary simulation.

11.1 Empirical Results of the Credit Crisis Bootstraping Simulation

The time series was used to simulate the Credit Crisis, the objective of this thesis. In addition, the simulation is evaluated in the following statistics:

- Averages of Averages – The average of all 3 time series for each random sample, and a final average taken of *all* the aggregated averages at the end of the simulation.
- Median of Medians – The median of all 3 time series for each random sample, and a final median taken of *all* the aggregated medians at the end of the simulation.
- Mean of 3 Time Series – At the end of 100,000 simulations the mean of the last random sample is taken.
- Median of 3 Time Series – At the end of 100,000 simulations the median of the last random sample is taken.

- Random Draw - At the end of 100,000 simulations a random 12 period cross section period is selected.

Table 30 – Table of Empirical Results of the Credit Crisis Bootstraping Simulation

	Actual 968.75	Actual 896.24	Actual 903.25			
	Predict Month 1	Predict Month 2	Predict Month 3	3 Mon Forecast Error	3 Mon Forecast MAE	12 Mon Forecast
Average of Averages	1142.35	1118.82	1095.81	588.74	196.25	908.50
Median of Medians	1147.31	1128.58	1110.28	617.93	205.98	957.84
Mean of 3 Time Series	1099.27	1033.98	952.21	317.22	105.74	734.36
Median of 3 Time Series	1133.79	1015.67	999.98	381.19	127.06	903.70
Random Draw #1	1155.97	1199.13	1225.19	812.05	270.68	630.74
Random Draw #2	1209.91	1046.32	918.69	406.68	135.56	962.74

Overall, the simulation results forecasted better results than other time series methods. The Double Exponential Smoothing for the Credit Crisis had the best MAE of 205. Compare that result to the mean of 3 time series which provided the best 3 month simulation of 105.74. However, there is uncertainty whether the mean of 3 time series will provide an accurate 12 month forecast, due to the fact that the statistic is derived only through a single average of 3 time series for a random draw. As a preference of long term forecasting, the Average of Averages will most likely be a better indicator, since it takes into account a longer simulation history data. *In fact, the simulation returns showed that after an averaging of over 1,000,000 samples, the results show that each recessionary period can expect on average a -2.08% return during a recessionary period with a 95% lower confidence interval of -14.54% and upper confidence interval of 9.67%* . The next section on recessionary analysis will compare the previous time series methods to bootstrapping, in order to evaluate the 12 month period forecasting accuracy. See the appendix for attached simulation implementation written in MATLAB.

Figure 31 – Simuation of Credit Crisis 12 Month Forecasted S&P 500 Returns

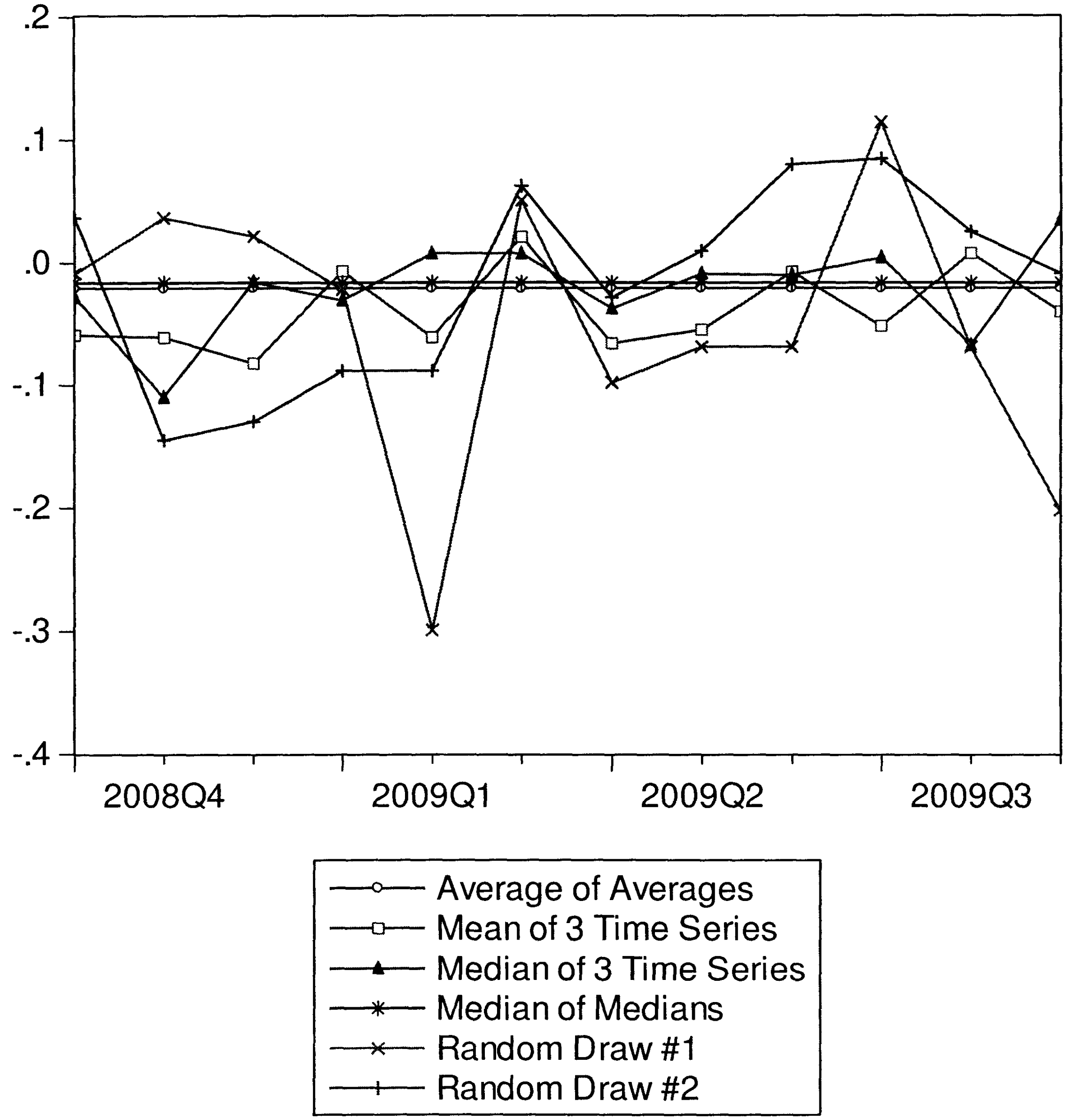
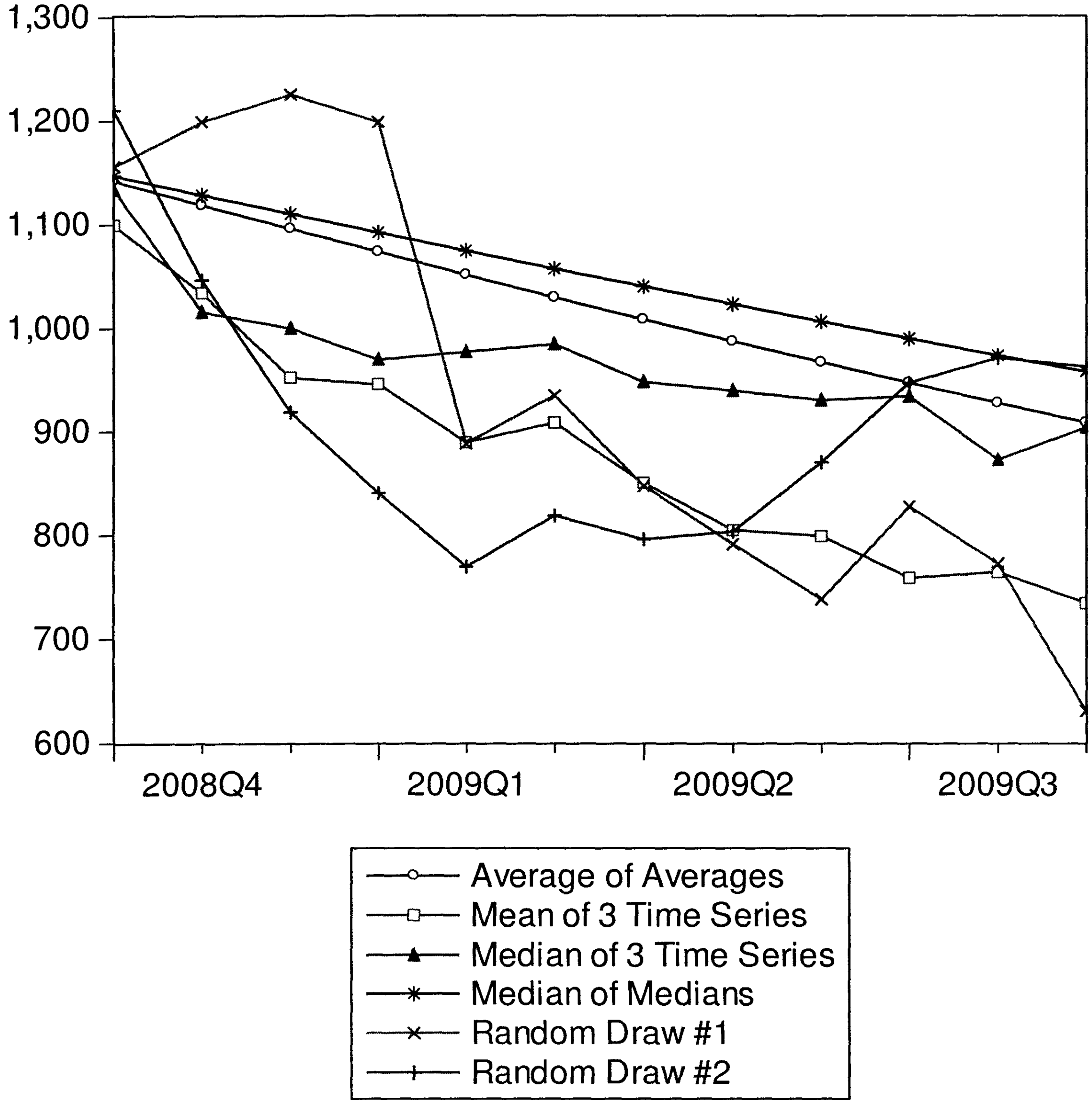


Figure 32 – Simulation of Credit Crisis, 12 Month Forecasted S&P 500 Level



12 COMPARISON OF COMBINED METHODOLOGIES DURING RECESSION/DEPRESSION PERIODS

The section will evaluate the best performing empirical models in its category. The "best performing " criteria to select the models in each category that had the lowest MAE in 3 month and 12 month forecasting errors. The following "best performing" models were identified:

- Log ARIMA
- Log Double (Brown) Exponential Smoothing
- Hybrid Linear Regression, Y=D(SPINDX)
- ARMAX(1,1,0) Mean + EGARCH(1,1) Volatility
- (3x7x1) Neural Network, with 100% Training, using a 12 month Double Exponential Forecast
- ARIMA Forecast + (7x7x1) Neural Network Residuals
- Average of Averages Bootstrapping Simulation
- Mean of 3 Time Series Bootstrapping Simulation

Previously, the analysis only involved evaluating models under different major event scenarios within its own category or method. For example, GARCH 12 month forecasts were evaluated under Black Monday. A baseline of comparison is needed to evaluate all the models under the same time series and metrics. This section on comparison analysis evaluates which model is most effective during recessions. Specifically, each of the models/simulation above will forecast 12 month periods for the Great Depression and the Tech Bubble, and will summarize which models were most effective in 3 month and 12 month forecasts.

12.1 Tech Bubble Empirical Results of Best Performing Models

The Tech Bubble era was chosen, for a couple of reasons. The first was that the time period had available VIX data so that the Linear Regression model could be compared. Secondly, the Tech Bubble fell within the 36 month recession period, which would allow the use of simulations. Previous time periods such as Black Monday in the 1980's wasn't selected due to the recessionary period was too short. Third, the Tech Bubble was a speculative period primarily concentrated in the technology sector and was different from other recessions. Overall, the Log Double Exponential Smoothing model outperformed all other forecasting models in the 3 month and 12 month benchmark. With a 12 month MAPE of 3.77%, and a MAE error of 44. In addition, the Double Exponential Smoothing had the most accurate 12 month point estimate of 1132. The (3x7x1) Neural Network Model had the second best 12 month results with a MAPE of 4.29% and a MAE of 51. Another model that performed well was the Average of Averages bootstrapping simulation, which had the second best 3 month forecast with a MAPE of 4.80%, and the third best 12 month forecast with a MAPE of 4.43%. The most surprising result was the linear regression, which had the worst 3 month forecast, and the bottom 3rd 12month forecast with a MAPE of 8.90%, 13.04% respectively. The analysis showed that the regression had consistently produced level values higher than the actual. This is due to the fact that the Tech Bubble had small influence on the explanatory values.

Table 33 - Linear Regression Forecasted Explanatory Values

Forecasting Model	Log Double Exponential Smoothing	ARIMA(2,1,3)	ARIMA(2,1,2)
Date	**VIX**	**IG_OAS**	**GDP_LVL**
2001M01	24.8696	201.8488	9906
2001M02	25.0635	194.5873	9920
2001M03	25.268	188.4603	9932
2001M04	25.4842	185.8845	9941
2001M05	25.7132	186.2597	9949
2001M06	25.9561	187.8497	9955
2001M07	26.2141	189.2339	9960

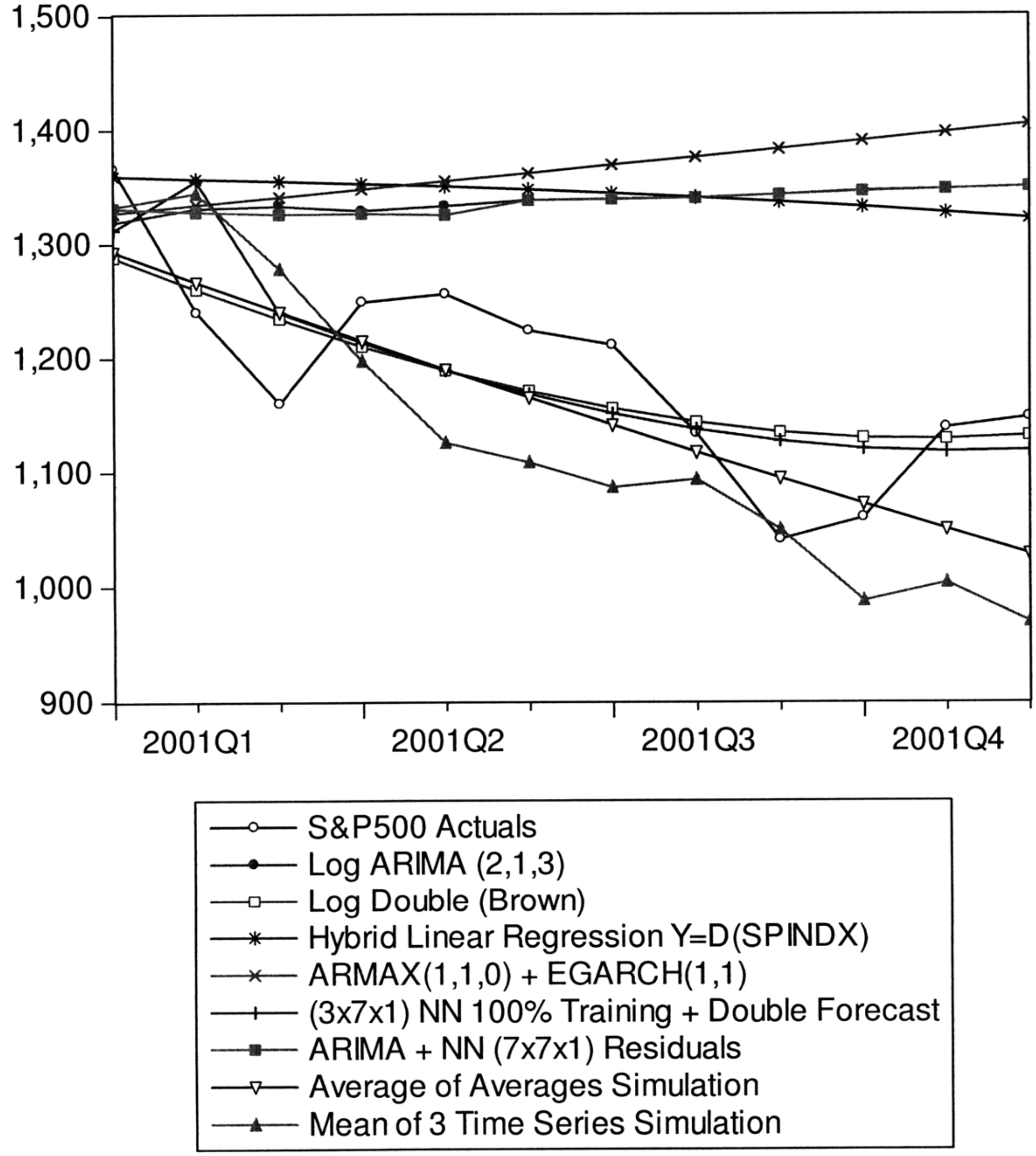
2001M08	26.4884	189.8396	9965
2001M09	26.7804	189.7783	9968
2001M10	27.0913	189.4309	9971
2001M11	27.4227	189.1186	9973
2001M12	27.7761	188.9766	9975

Notice that the forecasted explanatory values during these months did not fluctuate greatly. However, the VIX did increase in volatility and the corporate default rates (IG_OAS) actually fell during the Tech Crisis instead of rise, which in turn produced a lower forecasted result in the regression model. The same issue is seen with the GDP, where levels forecasted are increasing instead of decreasing. *This analysis draws a particular insight that not all recessions and sharp market declines are equated with higher default rates and lower GDP.* The linear regression model performed well for the first 3 months of the Credit Crisis for those circumstances, but failed to appropriately model the tech speculation bubble. Another surprising result was the EGARCH model that had the worst 12 month forecast, and the second worst 3 month forecast with a MAPE of 15.30%, 8.68%, respectively. This is due to insignificant parameter estimations in the ARMAX portion. A constant is not defined for this model, and since the historical observations had no recent decline, the model assumes a positive drift. This model estimation performs poorly during recessionary periods. Moreover, this positive drift issue is exactly the same with the Log ARIMA model, which also performed poorly. Lastly, the ARIMA+ NN Residual model indicated that the NN residual did improve the forecast slightly and outperformed the Log ARIMA model in comparison. Compare the 12 month ARIMA+ NN MAPE of 12.92% versus Log ARIMA 13.14%.

Table 34 - 3 and 12 Month Forecast Comparison Table, Tech Bubble

							Start Jan 2001 (Actual)	1366	1240	1160	1148
	Fit Range	3 Mon Forecast Error	3 Mon MAE	3 Mon MAPE	12 Mon Forecast Error	12 Mon Forecast MAE	12 Mon Forecast MAPE	Mo 1	Mo 2	Mo 3	Mo 12
Log ARIMA (2,1,3)	12/1925 – 12/2000	311	104	8.56%	1787	149	13.14%	1319	1331	1333	1350
Log Double (Brown)	12/1925 – 12/2000	172	57	4.57%	531	44	3.77%	1288	1260	1234	1132
Hybrid Linear Regression Y=D(SPINDX)	1/1990 - 12/2000	318	106	8.90%	1776	148	13.04%	1360	1357	1355	1322
ARMAX(1,1,0) +EGARCH(1,1)	12/1925 – 12/2000	314	105	8.68%	2080	173	15.30%	1327	1334	1341	1405
(3x7x1) NN 100% Training + Double Forecast	12/1925 – 12/2000	249	83	6.70%	607	51	4.29%	1312	1355	1239	1119
ARIMA + NN (7x7x1) Residuals	12/1925 – 12/2000	289	96	7.98%	1754	146	12.92%	1331	1328	1326	1350
Average of Averages Simulation	N/A	180	60	4.80%	629	52	4.43%	1293	1266	1240	1028
Mean of 3 Time Series Simulation	N/A	256	85	7.02%	990	82	6.98%	1332	1344	1278	969

Figure 35 – Tech Bubble, 12 Month Forecast Comparison Graph



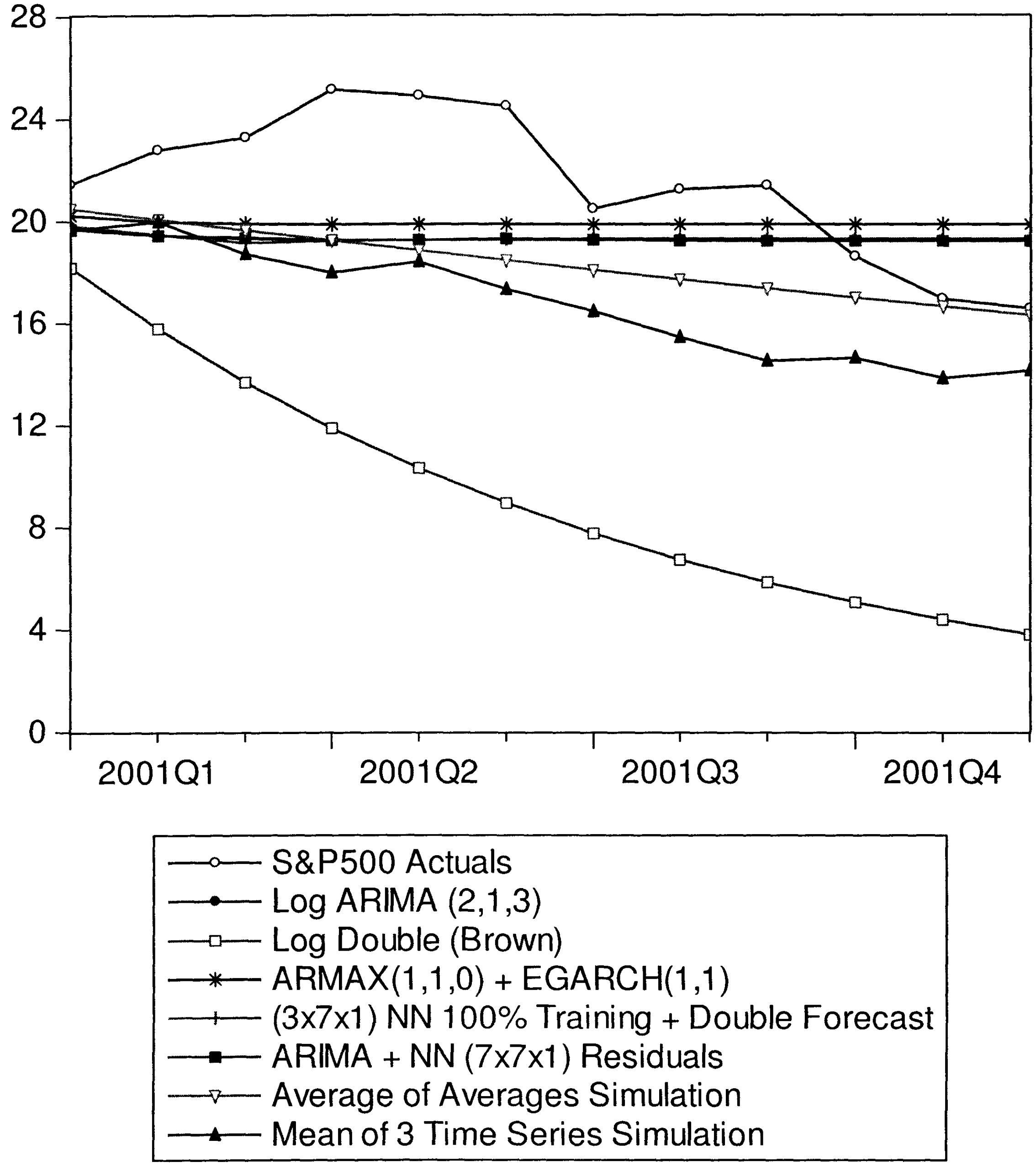
12.2 Great Depression Empirical Results of Best Performing Models

The Great Depression was chosen due to its historical impact to the stock market and to the S&P 500. It is a historical model that can be considered on the magnitude and persistence of negative returns. For example, in October 1929, the Great Crash occurred within a span of 3 days, which is also known as "Black Thursday, Black Monday, and Black Tuesday"[xxix], which caused the S&P 500 to fall by 19.92% in that month. This steep decline started a movement in sporadic monthly negative returns extending up until February of 1933. No recessionary event can compare to this event with the exception of the Credit Crisis. Therefore, the Great Depression is an appropriate time era to compare our results with the Credit Crisis forecasts. Overall, the Average of Averages bootstrapping simulation outperformed all models for the 3 month period with a MAPE of 10.67%. In addition, the Average of Averages had the best 12 month point estimate. The EGARCH model which had the worst Tech Bubble results had the best overall 12 month forecast with a MAPE of 12.50%, and a MAE of 2.72. EGARCH also had the second best 3 month forecast with a MAPE of 10.84%. EGARCH outperformed due to similar reasons stated previously. EGARCH and ARIMA models perform better when there are more recent observations that include negative returns, and significant parameter estimation. *The time range in periods were arranged purposely to test differences in time periods ending with and without negative returns prior to forecasting.* To no surprise the Log ARIMA (2, 1, 1) had the third best 12 and 3 month forecast with a MAPE of 13.25%, 13.03% respectively. Also, consistent with our previous evaluation, the Log Double Brown had the worst 3 and 12 month forecast with a MAPE of 52.27%, 29.05% respectively. Again, this is due to the last 3 observations ended with negative returns prior to forecasting, resulting in an extreme forecast. The neural network models ranked average in comparison, but the ARIMA + NN residual had the worst 3 month forecast in comparison with the ARIMA(2,1,1), and tied for the 12 month model with a MAPE of 13.23%, 13.25% respectively. This is an indication that the NN residuals did not provide much value to the ARIMA (2, 1, 1) forecast, and overall produces mixed results. There are in previous cases the NN residuals improved the forecast, but in this scenario it underperformed. In addition, the 3 out of 4 events (Credit Crisis, Tech Bubble, Oil Crisis, Great Depression) tested, modeling the residuals with NN indicate to improve the ARIMA forecast.

Table 36 - Great Depression, 3 and 12 Month Forecast Comparison Table

	Fit Range	3 Mon Forecast Error	3 Mon MAE	3 Mon MAPE	12 Mon Forecast Error	12 Mon Forecast MAE	12 Mon Forecast MAPE	Mo 1	Mo 2	Mo 3	Mo 12
							Start Dec 1929 (Actual)	21.45	22.79	23.28	16.57
Log ARIMA (2,1,1)	12/1925 – 11/1929	8.89	2.96	13.03%	35.37	2.95	13.25%	19.84	19.46	19.32	19.24
Log Double (Brown)	12/1925 – 11/1929	19.86	6.62	29.05%	132.20	11.02	52.27%	18.19	15.78	13.69	3.80
Hybrid Linear Regression Y=D(SPINDX)	12/1925 – 11/1929	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ARMAX(1,1,0) +EGARCH(1,1)	12/1925 – 11/1929	7.41	2.47	10.84%	32.69	2.72	12.50%	20.23	19.99	19.90	19.85
(3x7x1) NN 100% Training + ARIMA(2,1,1) Forecast	12/1925 – 11/1929	9.04	3.01	13.24%	35.57	2.96	13.33%	19.84	19.48	19.16	19.29
ARIMA + NN (7x7x1) Residuals	12/1925 – 11/1929	9.02	3.01	13.23%	35.41	2.95	13.25%	19.70	19.44	19.36	19.21
Average of Averages Simulation	N/A	7.31	2.44	10.67%	34.97	2.91	12.64%	20.49	20.07	19.65	16.30
Mean of 3 Time Series Simulation	N/A	9.13	3.04	13.38%	52.08	4.34	19.80%	19.68	19.99	18.72	14.14

Figure 37 - Great Depression, 12 Month Forecast Comparison Graph



13 SUMMARY OF MODEL PERFORMANCE

To summarize the data, the MAPE was aggregated for each major event, and arithmetic average was taken for all the 3 months MAPE. The results indicate that the Hybrid Linear Regression model was the best short term forecasting model. Note that the linear regression only accounted for two historical periods (Credit Crisis and the Tech Bubble) due to data constraints of the VIX, which its earliest data started in 01/1990. The mean simulation model was ranked second, but could very well be a more consistent performing model than the regression. The Double Exponential Smoothing model was ranked last with the indication that this model can be extreme, varying from getting very accurate results to inaccurate. Another interesting observation is that the Neural Networking residual model on average outperforms the standalone ARIMA time series model, which shows promise for using Neural Networks as a forecasting method.

Table 38 - Performance Summary For 3 Month Forecast Models

	3 Month MAPE Credit Crisis	3 Month MAPE Tech Bubble	3 Month MAPE Great Depression	3 Month MAPE Avg	3 Month Rank
Hybrid Linear Regression Y=D(SPINDX)	2.15%	8.90%	N/A	5.52%	1
Mean of 3 Time Series Simulation	11.42%	7.02%	13.38%	10.61%	2
Average of Averages Simulation	21.36%	4.80%	10.67%	12.28%	3
ARIMA + NN (7x7x1) Residuals	16.25%	7.98%	13.23%	12.49%	4
ARMAX(1,1,0) +EGARCH(1,1)	24.91%	8.68%	10.84%	14.81%	5
(3x7x1) NN 100% Training + ARIMA(2,1,3) Forecast	25.95%	6.70%	13.24%	15.30%	6
ARIMA (2,1,3)	25.27%	8.56%	13.03%	15.62%	7
Double (Brown) Exponential Smoothing	22.31%	4.57%	29.05%	18.64%	8

For the 12 month performance summary, the same approach was taken where the 12 month MAPE of all the major events were aggregated, and a final arithmetic average was taken. The ranking indicated that the Average of Averages simulation ranked first. The "2%" rule of thumb indicates to have some predictive results over a longer range. The mean of 3 time series simulation ranked 6th, indicating that it is best to be considered for 3 month forecasts. Also, the mean of 3 time series has a tendency to forecast extreme negative returns and does not forecast well for 12 months. ARMAX + EGARCH have also shown to be ranked 2nd to last due to the "constant" inaccuracy of hit and miss, and the Double Exponential Model placed last due to the extreme results. Lastly, the Neural Network models have generally performed well over their standalone ARIMA/Double Exponential counterparts with the NN (SPINDX) + ARIMA forecast ranking second.

Table 39 - Performance Summary for 12 Month Forecast Models

	12 Month MAPE Tech Bubble	12 Month MAPE Great Depression	12 Month MAPE Avg.	12 Month Rank
Average of Averages Simulation	4.43%	12.64%	8.54%	1
(3x7x1) NN 100% Training + ARIMA(2,1,3) Forecast	4.29%	13.33%	8.81%	2
Hybrid Linear Regression Y=D(SPINDX)	13.04%	N/A	13.04%	3
ARIMA + NN (7x7x1) Residuals	12.92%	13.25%	13.09%	4
ARIMA (2,1,3)	13.14%	13.25%	13.19%	5
Mean of 3 Time Series Simulation	6.98%	19.80%	13.39%	6
ARMAX(1,1,0) + EGARCH(1,1)	15.30%	12.50%	13.90%	7
Double (Brown) Exponential Smoothing	3.77%	52.27%	28.02%	8

The overall summary takes the arithmetic average of the 3 month MAPE average and the 12 month MAPE average. This approach indicates that the Hybrid Linear Regression is ranked first originally. However, since the linear regression had only one category to evaluate, a N/A ranking was applied. Instead, the Average of Averages simulation was ranked first and consistently outperformed most models on 3 month and 12 month forecasts. It is also easy to implement, assumes no distribution of the data and is a dependable and powerful method for forecasting.

Table 40 - Overall Performance Summary, Weighted 3 and 12 Months

	3 Month MAPE Avg	12 Month MAPE Avg	Weighted Avg	Weighted Rank
Hybrid Linear Regression Y=D(SPINDX)	5.52%	13.04%	9.28%	1
Average of Averages Simulation	12.28%	8.54%	10.41%	2
Mean of 3 Time Series Simulation	10.61%	13.39%	12.00%	3
(3x7x1) NN 100% Training + ARIMA(2,1,3) Forecast	15.30%	8.81%	12.05%	4
ARIMA + NN (7x7x1) Residuals	12.49%	13.09%	12.79%	5
ARMAX(1,1,0) +EGARCH(1,1)	14.81%	13.90%	14.36%	6
ARIMA (2,1,3)	15.62%	13.19%	14.41%	7
Double (Brown) Exponential Smoothing	18.64%	28.02%	23.33%	8

14 CONCLUSIONS

Overall, forecasting the S&P 500 with accurate point estimates can be difficult. Models that were shown to have a good fit in the evaluation range may not necessarily perform well in actual forecasts due to overfitting. However, certain models are more effective than others depending on the time range or on the trend. The Log Double Brown exponential smoothing is a good short range (3 months) and long range (12 months) forecasting model to use if you have a prior belief that a step recession or boom is in the early stages. However, the Double Exponential Smoothing model can produce extreme forecasts, so this method is best modeled prior to a large reduction in returns. The ARIMA (2, 1, 3) model and the Damped Exponential models are possible long range forecasting models if a clear trend is shown, or variation within 2 sigma is expected. These 12 month forecasting models were validated using the Tech Bubble data, because of the similarities in trends and data patterns. The EGARCH models are best used for forecasting volatility, and indicated below average performance results in comparison. In addition EGARCH models have similarities to the ARIMA model in misforecasting steep declines and increases in level. However, EGARCH and ARIMA models provide good 12 month forecasts when there are more recent observations that include correlated and subsequent negative/positive returns. Neural networks performed average in comparison to the other methods, but combination and residual Neural Networking models on outperformed results of traditional ARIMA and Double Exponential Smoothing models on average. Neural network research shows potential for improved forecasting accuracy in the future. Bootstrapping Simulation is the favored method to forecast the S&P 500 during recessionary periods due to its consistent MAPE performance in 3 and 12 month forecast range. *In conclusion, the S&P 500 index is considered undervalued if the assumptions of long run mean convergence holds true.* Reflecting upon the last table below, the 3 month forecasts of the Hybrid Linear regression show an excellent forecast with a low 3 month MAPE of 2.15%. *Since actual data is not available for 09/2009, the Average of Averages simulation, the favored method, forecasts a 12 month point estimate of 909 with a 95% lower confidence interval of 802 and upper confidence interval of 1021.*

Table 41 – Summary of 3 and 12 Month Level Forecasts of the Credit Crisis 10/2008 – 09/2009

	Predict Mon 1	Predict Mon 2	Predict Mon 3	3 Mon Forecast Error	3 Mon Forecast MAD	3 Mon MAPE	12 Mon Forecast	Lower 95% CI	Upper 95% CI
	Actual	Actual	Actual						
	968.75	896.24	903.25						
Hybrid Linear Regression Y=D(SPINDX)	973	924	877	58	19	2.15%	439	0	1647
Mean of 3 Time Series Simulation	1099	1034	952	317	106	13.69%	734	661	842
ARMAX(1,1,0) +EGARCH(1,1)	1159	1151	1144	686	229	24.91%	1095	977	1207
Log Double (Brown) Exponential Smoothing	1160	1128	1096	616	205	22.31%	850	491	1470
Average of Averages Simulation	1142	1119	1096	589	196	21.36%	909	802	1021
ARIMA (2,1,3)	1160	1157	1147	696	232	24.25%	1017	886	1157
(3x7x1) NN 100% Training + ARIMA(2,1,3) Forecast	1149	1184	1149	713	238	25.95%	1018	887	1158
ARIMA + NN (7x7x1) Residuals	1037	1079	1095	444	148	16.25%	1017	886	1157

15 FURTHER RESEARCH

In addition to validating long term models through the tech boom, there are other unprecedented events that can be used for back testing such as the first (1914-1918) and second world wars (1939-1945), the recession in the 1957-1958, 1960-1961, and prior to the oil crisis in 1969-1970. The models used in this thesis have been constrained to time series and simulation methods. Although introductory research in Neural Networks has been conducted, further research is needed to better understand the use of Neural Networks. This research has shown that Neural Networks hold forecasting potential and on average Neural Network models outperformed time series models. However, mixed results can occur and at times they do not outperform the ARIMA or Exponential model when combined. There is also potential in using VARFIMA (vector autoregressive fractionally integrated moving averages) and ARFIMA models to model long term memory. A challenge with using VARFIMA and ARFIMA models is that high frequency data is usually needed, and a strong data pattern may be required to yield effective results. One VARFIMA approach is to produce forecasts and level pricing by capturing both short term correlation structures and long-range dependence characteristics and feedback relationships between series. [xxx] Moreover, Ravishanker (2009) offers an approach to determine maximum likelihood estimation of parameters from vector ARFIMA models with Gaussian errors by using multivariate preconditioned conjugate gradient (MPCG) algorithm. Comparing VARFIMA with existing methods and the accuracy of VARFIMA producing 3 and 12 month forecasts is an avenue of new research. Lastly, another area of research is researching different Neural Network techniques such as Higher Order Neural Networks (HONN), Psi Sigma network (PSN), and a Recurrent Neural Network (RNN) for additional forecasting comparisons. Dunis et al demonstrate that these Neural Network techniques outperform the following traditional models: ARMA model, moving average convergence divergence technical model (MACD), and logistic regression model (LOGIT).[xxxi] Dunis's research compared these methods under the context of trading simulations of the EUR/USD Exchange rates. Researching and comparing these Neural Network methods against traditional methods for 12 month forecasts for the S&P 500 is an application of interest.

16 APPENDIX

16.1 Parameter Estimates of Preferred 12 Month Forecast: Combined Forecast ARIMA(2,1,3) + Log Damped Trend Exponential, Entire Data Set (1925 – 2008)

Parameter Estimates
SPINDX
Forecast combination 1: (Combination of 2 models)

Model Parameter	Estimate	Std. Error	T	Prob>|T|
ARIMA(2,1,3) NOINT	0.50000	.	.	.
Log Damped Trend Exponential Smoothing	0.50000	.	.	.
Combined Model Variance	0.00306	.	.	.

Parameter Estimates
SPINDX
ARIMA(2,1,3) NOINT

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Moving Average, Lag 1	1.56100	0.1464	10.6606	<.0001
Moving Average, Lag 2	-0.53854	0.1722	-3.1265	0.0167
Moving Average, Lag 3	-0.07952	0.0433	-1.8353	0.1091
Autoregressive, Lag 1	1.51684	0.1450	10.4607	<.0001
Autoregressive, Lag 2	-0.54658	0.1416	-3.8592	0.0062
Model Variance (sigma squared)	343.95079	.	.	.

Parameter Estimates
SPINDX
Log Damped Trend Exponential Smoothing

Model Parameter	Estimate	Std. Error	T	Prob>|T|
LEVEL Smoothing Weight	0.99900	0.0275	36.3026	<.0001
TREND Smoothing Weight	0.02712	0.0206	1.3150	0.2210
DAMPING Smoothing Weight	0.90578	0.0865	10.4745	<.0001
Residual Variance (sigma squared)	0.00306	.	.	.
Smoothed Level	7.33086	.	.	.
Smoothed Trend	0.00222	.	.	.

16.2 Parameter Estimates of Preferred 3 Month Forecast: Log Double (Brown) Exponential Smoothing, 7 Year Data Set (2000 – 2007)

Parameter Estimates

SPINDX

Log Double (Brown) Exponential Smoothing

Model Parameter	Estimate	Std. Error	T	Prob>|T|
LEVEL/TREND Smoothing Weight	0.43961	0.0305	14.4001	<.0001
Residual Variance (sigma squared)	0.00184	.	.	.
Smoothed Level	7.31144	.	.	.
Smoothed Trend	0.00748	.	.	.

16.3 Parameter Estimates of 12 Month Forecast, Tech Bubble: ARIMA (2, 1, 3), Entire Data Set (1925 – 2001)

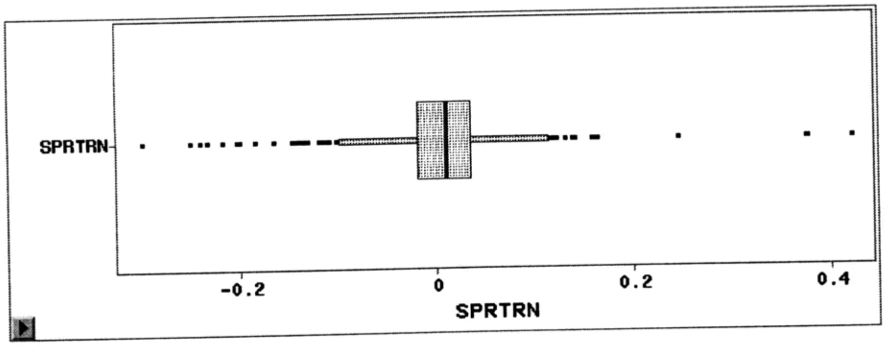
Parameter Estimates

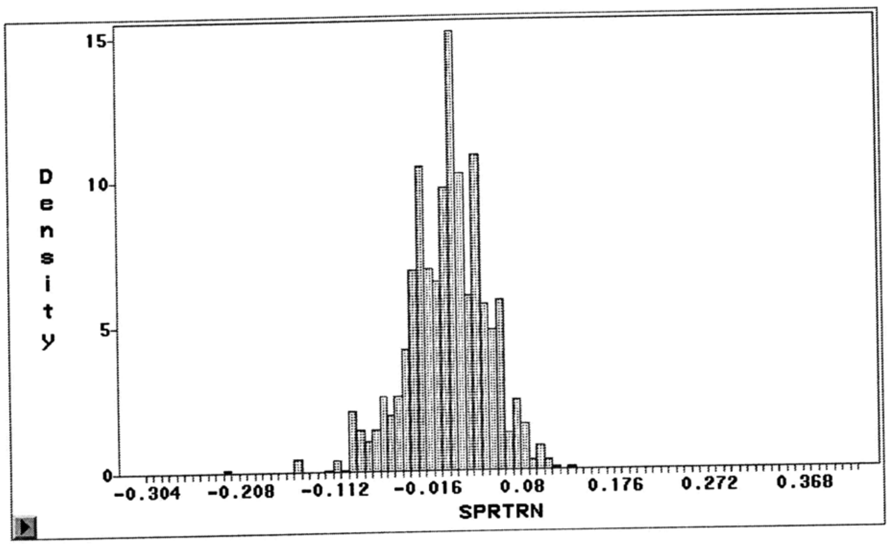
SPINDX

Log ARIMA(2,1,3) NOINT

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Moving Average, Lag 1	-0.31753	0.1463	-2.1701	0.0666
Moving Average, Lag 2	-0.64814	0.1394	-4.6497	0.0023
Moving Average, Lag 3	0.06029	0.0460	1.3108	0.2313
Autoregressive, Lag 1	-0.22981	0.1439	-1.5974	0.1542
Autoregressive, Lag 2	-0.60951	0.1340	-4.5476	0.0026
Model Variance (sigma squared)	0.00314	.	.	.

16.4 Box Plot and Distribution of S&P 500 Data, Entire Data Set (1925-2008)





Moments

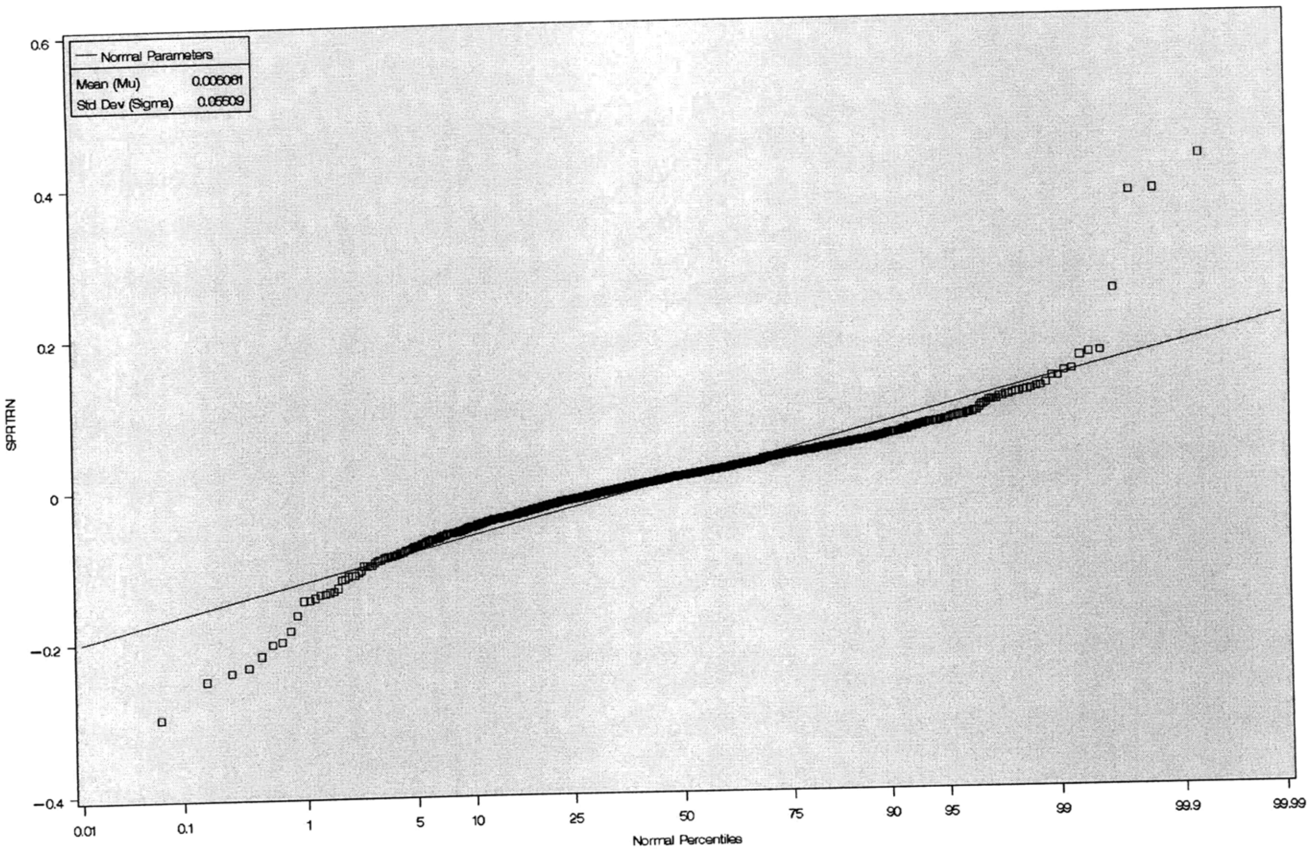
N	259833.000	Sum Wgts	259833.000
Mean	0.0062	Sum	1605.4380
Std Dev	0.0419	Variance	0.0018
Skewness	-0.4548	Kurtosis	2.0947
USS	465.9911	CSS	456.0715
CV	678.0644	Std Mean	8.219E-05

Quantiles

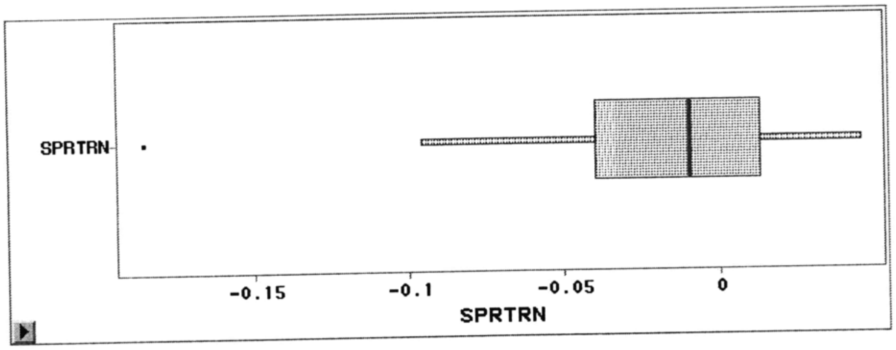
100%	Max	0.4222	99.0%	0.0967
75%	Q3	0.0346	97.5%	0.0803
50%	Med	0.0101	95.0%	0.0711
25%	Q1	-0.0188	90.0%	0.0568
0%	Min	-0.2994	10.0%	-0.0457
	Range	0.7216	5.0%	-0.0642
	Q3-Q1	0.0535	2.5%	-0.0899
	Mode	0.0148	1.0%	-0.0952

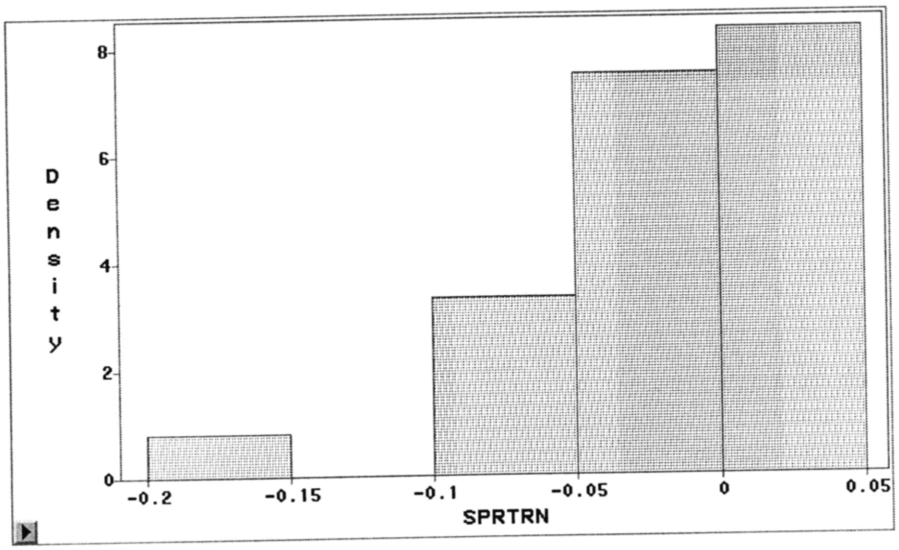
\

16.5 Normality Plot of S&P 500 Data, Entire Data Set (1925-2008)



16.6 Box Plot and Distribution of S&P 500 Data, Credit Crisis Data Set (01/2007- 12/2008)





Moments

N	24.0000	Sum Wgts	24.0000
Mean	-0.0199	Sum	-0.4774
Std Dev	0.0532	Variance	0.0028
Skewness	-1.4293	Kurtosis	2.7866
USS	0.0747	CSS	0.0652
CV	-267.6331	Std Mean	0.0109

Quantiles

100% Max	0.0465	99.0%	0.0465
75% Q3	0.0135	97.5%	0.0465
50% Med	-0.0093	95.0%	0.0433
25% Q1	-0.0397	90.0%	0.0358
0% Min	-0.1856	10.0%	-0.0899
Range	0.2321	5.0%	-0.0952
Q3-Q1	0.0532	2.5%	-0.1856
Mode	.	1.0%	-0.1856

16.7 Parameter Estimation, Linear Regression, VIX Data point, ARIMA(2,1,3)

Parameter Estimates
VIX: VIX
ARIMA(2,1,3) NOINT

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Moving Average, Lag 1	1.78200	0.1432	12.4458	<.0001
Moving Average, Lag 2	-0.75903	0.2426	-3.1284	0.0020
Moving Average, Lag 3	-0.06449	0.1080	-0.5970	0.5511
Autoregressive, Lag 1	1.54618	0.1280	12.0792	<.0001
Autoregressive, Lag 2	-0.61630	0.1278	-4.8227	<.0001
Model Variance (sigma squared)	12.40973	.	.	.

Fit Range: JAN1990 to SEP2008

16.8 Parameter Estimation, Linear Regression, IG OAS Data point, ARIMA(2,1,3)

Parameter Estimates
IG_OAS: IG_OAS
ARIMA(2,1,3) NOINT

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Moving Average, Lag 1	-0.17084	0.0997	-1.7130	0.0881
Moving Average, Lag 2	-0.83083	0.0631	-13.1692	<.0001
Moving Average, Lag 3	-0.31492	0.0760	-4.1422	<.0001
Autoregressive, Lag 1	-0.09030	0.0876	-1.0312	0.3036
Autoregressive, Lag 2	-0.72022	0.0809	-8.8987	<.0001
Model Variance (sigma squared)	142.62082	.	.	.

Fit Range: JAN1990 to SEP2008

16.9 Parameter Estimation, Linear Regression, GDP Level Data point, Double Exponential Smoothing

Parameter Estimates
GDP_LVL: GDP_LVL
Double (Brown) Exponential Smoothing

Model Parameter	Estimate	Std. Error	T	Prob>|T|
LEVEL/TREND Smoothing Weight	0.99900	0.0237	42.1547	<.0001
Residual Variance (sigma squared)	128.95299	.	.	.
Smoothed Level	11712	.	.	.
Smoothed Trend	-5.00000	.	.	.

Fit Range: MAR1990 to SEP2008

16.10 Parameter Estimation, SPINDX Return Linear Regression, Credit Crisis

Dependent Variable: SPINDX_RET
Method: Least Squares
Date: 02/07/09 Time: 00:01
Sample (adjusted): 2005M04 2008M09
Included observations: 42 after adjustments

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.026321	0.009250	2.845503	0.0070
HOMEDEF	4.10E-07	2.13E-07	1.926538	0.0613
IG_OAS	-0.000383	9.80E-05	-3.908391	0.0004

R-squared	0.333917	Mean dependent var	-0.000289
Adjusted R-squared	0.299759	S.D. dependent var	0.032075
S.E. of regression	0.026840	Akaike info criterion	-4.329071
Sum squared resid	0.028096	Schwarz criterion	-4.204952
Log likelihood	93.91049	Hannan-Quinn criter.	-4.283576
F-statistic	9.775621	Durbin-Watson stat	1.914902
Prob(F-statistic)	0.000362		

16.11 Parameter Estimation, D(SPINDX) Linear Regression, Credit Crisis

Dependent Variable: D(SPINDX)
Method: Least Squares
Date: 02/07/09 Time: 00:07
Sample (adjusted): 1990M03 2008M09
Included observations: 223 after adjustments

	Coefficient	Std. Error	t-Statistic	Prob.
VIX	-0.834168	0.460197	-1.812633	0.0713
GDP_LVL	0.004651	0.000897	5.186794	0.0000
IG_OAS	-0.215100	0.067178	-3.201952	0.0016

R-squared	0.109340	Mean dependent var	3.742018
Adjusted R-squared	0.101243	S.D. dependent var	41.20658
S.E. of regression	39.06499	Akaike info criterion	10.18169
Sum squared resid	335736.2	Schwarz criterion	10.22753
Log likelihood	-1132.259	Hannan-Quinn criter.	10.20020
Durbin-Watson stat	2.063864		

16.12 Parameter Estimation, Credit Crisis Data, ARMAX(2,3,0) + EGARCH(1,1)

```
Mean: ARMAX(1,1,0); Variance: EGARCH(1,1)

Conditional Probability Distribution: Gaussian
Number of Model Parameters Estimated: 6
```

Parameter	Value	Standard Error	T Statistic
AR(1)	0.94178	0.037496	25.1170
MA(1)	-0.88899	0.050912	-17.4614
K	-0.18814	0.040946	-4.5948
GARCH(1)	0.96726	0.0069873	138.4316
ARCH(1)	0.22529	0.024525	9.1862
Leverage(1)	-0.10197	0.022123	-4.6093

16.13 Parameter Estimation, Credit Crisis Data, Constant + ARMAX(2,3,0) + GARCH(1,1) Model

```
Mean: ARMAX(1,1,0); Variance: GARCH(1,1)

Conditional Probability Distribution: Gaussian
Number of Model Parameters Estimated: 6
```

Parameter	Value	Standard Error	T Statistic
C	0.011491	0.0025628	4.4838
AR(1)	-0.83974	0.12099	-6.9404
MA(1)	0.87152	0.11143	7.8214
K	5.5752e-005	1.6939e-005	3.2914
GARCH(1)	0.85585	0.01388	61.6590
ARCH(1)	0.13405	0.017048	7.8634

16.14 Parameter Estimation, Black Monday Data, ARMAX(2,3,0) + EGARCH(1,1)

```
Mean: ARMAX(1,1,0); Variance: EGARCH(1,1)

Conditional Probability Distribution: Gaussian
Number of Model Parameters Estimated: 6
```

Parameter	Value	Standard Error	T Statistic
AR(1)	0.90579	0.063172	14.3384
MA(1)	-0.84897	0.080959	-10.4864
K	-0.17056	0.046044	-3.7043
GARCH(1)	0.96991	0.0080176	120.9721
ARCH(1)	0.21039	0.024848	8.4673
Leverage(1)	-0.10819	0.02565	-4.2178

16.15 Parameter Estimation, Black Monday Data, Constant + ARMAX(2,3,0) + GARCH(1,1) Model

```
Mean: ARMAX(1,1,0); Variance: GARCH(1,1)

Conditional Probability Distribution: Gaussian
Number of Model Parameters Estimated: 6
```

Parameter	Value	Standard Error	T Statistic
C	0.011797	0.0031436	3.7525
AR(1)	-0.81706	0.18744	-4.3590
MA(1)	0.84814	0.1754	4.8354
K	6.898e-005	2.0583e-005	3.3512
GARCH(1)	0.85412	0.01575	54.2300
ARCH(1)	0.1274	0.018214	6.9943

16.16 ARIMA(2,1,3) Neural Network Residual Fit Statistics, and Weighting Estimation, Credit Crisis

	Fit Statistic	Training
1	[TARGET=ERROR]	.
2	Average Error	237.32218868
3	Average Squared Error	237.32218868
4	Sum of Squared Errors	238508.79963
5	Root Average Squared Error	15.405264966
6	Root Final Prediction Error	16.41962421
7	Root Mean Squared Error	15.920525241
8	Error Function	238508.79963
9	Mean Squared Error	253.46312394
10	Maximum Absolute Error	128.48703954
11	Final Prediction Error	269.6040592
12	Divisor for ASE	1005
13	Model Degrees of Freedom	64
14	Degrees of Freedom for Error	941
15	Total Degrees of Freedom	1005
16	Sum of Frequencies	1005
17	Sum Case Weights * Frequencies	1005
18	Akaike's Information Criterion	5624.7657572
19	Schwarz's Baysian Criterion	5939.1812977

	From	To	Weight
1	ERROR1	H11	-0.568339481
2	ERROR2	H11	-0.813869681
3	ERROR3	H11	1.1231864768
4	ERROR4	H11	-1.49964644
5	ERROR5	H11	0.7982100526
6	ERROR6	H11	0.1886645684
7	ERROR7	H11	-1.165239711
8	ERROR1	H12	0.0529356329
9	ERROR2	H12	-0.111725463
10	ERROR3	H12	0.4100516329
11	ERROR4	H12	-1.56075493
12	ERROR5	H12	0.1105915503
13	ERROR6	H12	0.4308771758
14	ERROR7	H12	1.4506655996
15	ERROR1	H13	-0.355260731
16	ERROR2	H13	0.2789253631
17	ERROR3	H13	0.2775489495
18	ERROR4	H13	-1.001263813
19	ERROR5	H13	0.8412174359
20	ERROR6	H13	-0.09472083
21	ERROR7	H13	-0.192825135
22	ERROR1	H14	1.0715936615
23	ERROR2	H14	0.5417945492
24	ERROR3	H14	-1.100098257
25	ERROR4	H14	1.3184865278
26	ERROR5	H14	1.6765060752
27	ERROR6	H14	-0.79608386
28	ERROR7	H14	-1.959002734

29	ERROR1	H15	-0.599294236
30	ERROR2	H15	-0.821964239
31	ERROR3	H15	1.1440428323
32	ERROR4	H15	-1.534245712
33	ERROR5	H15	0.8156417968
34	ERROR6	H15	0.1894496044
35	ERROR7	H15	-1.17556977
36	ERROR1	H16	-0.572121654
37	ERROR2	H16	-0.249605909
38	ERROR3	H16	0.0808344677
39	ERROR4	H16	0.498707449
40	ERROR5	H16	-0.321091195
41	ERROR6	H16	-0.431459559
42	ERROR7	H16	0.2300927811
43	ERROR1	H17	-0.499794872
44	ERROR2	H17	0.5395390222
45	ERROR3	H17	-0.430869589
46	ERROR4	H17	-0.579171337
47	ERROR5	H17	1.2915406385
48	ERROR6	H17	0.6226783505
49	ERROR7	H17	-0.402987594
50	BIAS	H11	-1.569678657
51	BIAS	H12	-1.805767057
52	BIAS	H13	0.660205842
53	BIAS	H14	-0.471388933
54	BIAS	H15	-2.502587373
55	BIAS	H16	0.3031878657
56	BIAS	H17	0.1971611848
57	H11	ERROR	0.3610953982
58	H12	ERROR	-0.07612045
59	H13	ERROR	0.7316271669
60	H14	ERROR	-8.536855E-7
61	H15	ERROR	-0.724542904
62	H16	ERROR	-0.06496634
63	H17	ERROR	-0.066264445
64	BIAS	ERROR	-0.944353753

16.17 SPINDX Neural Network + Double Exponential Smoothing Fit Statistics, Oil Crisis

	Fit Statistic	Training
1	[TARGET=SPINDX]	.
2	Average Error	26.292295419
3	Average Squared Error	26.292295419
4	Sum of Squared Errors	6336.443196
5	Root Average Squared Error	5.1276013319
6	Root Final Prediction Error	5.9604246989
7	Root Mean Squared Error	5.5596293946
8	Error Function	6336.443196
9	Mean Squared Error	30.909479005
10	Maximum Absolute Error	41.407149622
11	Final Prediction Error	35.526662591
12	Divisor for ASE	241
13	Model Degrees of Freedom	36
14	Degrees of Freedom for Error	205
15	Total Degrees of Freedom	241
16	Sum of Frequencies	241
17	Sum Case Weights * Frequencies	241
18	Akaike's Information Criterion	859.89550309
19	Schwarz's Baysian Criterion	985.34819269

16.18 Parameter Estimates, Double Exponential Smoothing, Oil Crisis

Parameter Estimates

SPINDX: SPINDX

Double (Brown) Exponential Smoothing

Model Parameter	Estimate	Std. Error	T	Prob>|T|
LEVEL/TREND Smoothing Weight	0.51636	0.0132	38.9872	<.0001
Residual Variance (sigma squared)	4.39889	.	.	.
Smoothed Level	101.63378	.	.	.
Smoothed Trend	-2.88645	.	.	.

Fit Range: DEC1925 to NOV1973

16.19 MATLAB Implementation of Bootstraping Simulation, Mean of Time Series

```
load recession;

%define total number of iterations
iter = 100000;

%Generate random number from 1-36, for a vector of 12
col = 3;
horizon_ret = zeros(col,12);
avg = zeros(col*iter,12);

row_counter = 0;
col_counter = 0;

for i=1:iter
     u = unidrnd(36,12,1);

    for j=1:12

       horizon_ret(:,j) = [D(u(j)); O(u(j)); TB(u(j))];
       row_counter = row_counter + 1;
    end
    col_counter = col_counter + col;


    if(i==1)
            avg(1:col,:) = horizon_ret;
    else
            avg(col_counter-col+1:col_counter,:) = horizon_ret;
    end


end

final_avg = mean(avg)
```

17 REFERENCES

[i] Kathman, David. "The S&P 500 affects every investor whether they care". http://news.morningstar.com/articlenet/article.aspx?id=1054. 09/25/1998.

[ii] *Standard and Poor's.* (2009, January 1). Retrieved March 15, 2009, from Standard and Poor's: http://www.standardandpoors.com

[iii] Zhang, P. (2001). Time Series Forecasting using a hybrid ARIMA and Neural Network Model. *Neurocomputing* .

[iv] G.E.P. Box, G. Jenkins (1970), *Time Series Analysis, Forecasting and Control*, San Francisco, CA, 1970.

[v] Asteriou, D. (2007). *Applied Econometrics a Modern Approach.* NY, NY: Palgrave Macmillian.

[vi] SAS Institute Inc. (1999). SAS Documentation Version 8. Section 11. Cary, N.C.

[vii] SAS Institute Inc. (1999) SAS Documentation Version 9.2. Section 11. Cary, NC.

[viii] SAS Institute Inc. (1999) SAS Documentation Version 8. Section 25. Cary, NC.

[ix] *Mean Absolute Error.* (2005). Retrieved Februrary 10, 2009, from Wikipedia: http://en.wikipedia.org/wiki/Mean_absolute_error

[x] *Mean Absolute Percentage Error.* (2005). Retrieved Februrary 10, 2009, from Wikipedia: http://en.wikipedia.org/wiki/Mean_absolute_percentage_error

[xi] Craig, P (2008). Credit Crisis is Not Worst in History. *Wealth Bulletin.*

[xii] Cassidy, John (2003). *How America Lost Its Mind and Money in the Internet Era.* NY,NY: Harper Perennial.

[xiii] *Linear Regression.* (2003). Retrieved Februrary 10, 2009, from Wikipedia: http://en.wikipedia.org/wiki/Linear_regression

[xiv] Rapach, D. W. (2006). Structural Breaks and Predictive Regression Models of Aggregate U.S. Stock Returns. *Journal of Financial Econometrics* , Vol 4, No.2, 238-274.

[xv] Lettau, M. L. (2001). Consumption, Aggregate Wealth and Expected Stock Returns. *The Journal of Finance* , Vol. LVI, No.3.

[xvi] Jeff Kearns, M. T. (2008, December 22).. *VIX fails to forecast S&P 500 drop, Loses Followers*. Retrieved January 25 2009, 2009, from Bloomberg: http://www.bloomberg.com/apps/news?pid=20601110&sid=a6bJKg1wZ8Vc

[xvii] Luby, B. (2008, December 22). *The VIX, Lehman Brothers, and Forecasting*. Retrieved January 25, 2009, from vixandmore: http://vixandmore.blogspot.com/2008/12/vix-lehman-brothers-and-forecasting.html

[xviii] Engle, R. (2001). GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives* , Volume 15, Number 4, 157-168.

[xix] Montgomery, D. C. (2008). *Introduction to Time Series Analysis and Forecasting*. Hoboken, New Jersey: Wiley.

[xx] *Autoregressive Conditional Heteroskedasticity*. (2001). Retrieved Februrary 15, 2009, from Wikipedia: http://en.wikipedia.org/wiki/Autoregressive_conditional_heteroskedasticity

[xxi] Wei, W. S. (2005). *Time Series Analysis Univariate and Multivariate Methods, Second edition.* NY, NY: Addison Wesley.

[xxii] *Black Monday*. (2002). Retrieved Februrary 11, 2009, from Wikipedia: http://en.wikipedia.org/wiki/Black_Monday_(1987)

[xxiii] Brady, N. e. (January 1988.). *Report of the Presidential Task Force on Market Mechanisms.* Washington, D.C.: The President of the United States, The Secretary of the Treasury and The Chairman of the Federal Reserve Board.

[xxiv] S.A. Hamid, Z. I. (2002). Using Neural Networks For Forecasting SP500 Volatility. *Journal of Business Research* , 1116-1125.

[xxv] Tang, Z. F. (1993). Feedforward Neural Nets as Models for Time Series Forecasting. *ORSA Journal on Computing* , Vol 5, No 4.

[xxvi] The Office of Electronic Information, Bureau of Public Affairs. (1974). *Second Arab Oil Embargo, 1973-1974*. Retrieved February 15, 2009, from U.S. Department of State: http://www.state.gov/r/pa/ho/time/dr/96057.htm

[xxvii] Perron, P.; University, Princeton; Program, Econometric Research (1988) (PDF), *The Great Crash, the Oil Price Shock and the Unit Root Hypothesis*, Econometric Research Program, Princeton University Princeton, NJ

[xxviii] Wilmott, P. (2007). Paul Wilmott Introduces Quantitative Finance, 2nd Edition . Hoboken, NJ: John Wiley & Sons.

[xxix] Brooks, John. (1969). *Once in Golconda: A True Drama of Wall Street 1920-1938*. New York: Harper & Row.

[xxx] Ravishanker, N. (2009). A multivariate preconditioned conjugate gradient approach for maximum likelihood estimation in vector long memory processes. *Statistics and Probability Letters* , 1282-1289.

[xxxi] Dunis, C. L. (February 2008). Higher Order and Recurrent Neural Architectures for Trading the EUR/USD Exchange Rate. *Liverpool Business School, CIBEF* .

Hull, J. C. (2005). *Options, Futures and Other Derivatives, 6th Edition.* Upper Saddle River, NJ: Prentice Hall.

Pfaff, B. (2006). *Analysis of integrated and cointegrated time series with R.* NY, NY: Springer Science.