

NSE RESEARCH INITIATIVE

PAPER NO.: 10

Empirical Investigation of Multifactor Asset Pricing Model Using Artificial Neural Networks

M. V. Kamath

Empirical Investigation of Multifactor Asset Pricing Model Using Artificial Neural Networks

M. V. KAMATH*

In recent times there have been many advances in quantitative modeling of financial markets and in this paper, I attempt to use one such technique, Artificial Neural Networks for the modelling of Asset prices. The Capital Asset Pricing model (CAPM) has been in use for over two decades but of late it is becoming increasingly clear that there are sources of priced risk other than the market portfolio. Artificial Neural Networks allows us to examine these other factors of risk without any strong assumptions about the model. Further Neural Nets can approximate any non-linear and complex relationship between asset returns and various factors of risk thereby allowing us to understand the sensitivity of asset returns to various macro-economic factors even under structural or policy changes.



National Stock Exchange of India Limited

* M. V. Kamath is a doctoral student at the S.J.M.School of Management, Indian Institute of Technology, Bombay. Prior to joining IIT Bombay he has done his Master of Business Management in finance from IIT Kharagpur and B.Tech in Civil Engineering from IIT Madras. The views expressed in this paper are of the author and not necessarily of NSE.

I Introduction

This study investigates the behavior of multifactor asset pricing model in the Indian context using a non-parametric tool – Artificial Neural Networks (ANN).

Artificial Neural Networks are widely used in various branches of engineering and science and their property to approximate complex and nonlinear equations makes it a useful tool in econometric analysis. A number of econometric models have been developed for pricing financial assets and forecasting financial asset returns. An empirical econometric analysis starts with the specification of the econometric model and subsequently we have to specify a priori the functional form of the equation, the variables to be included in the equation and the assumptions about the independent, dependent variables and error terms. Neural Nets however make no a priori assumptions about the problem and let the data speak for themselves without the pre-specification of a model or equation. Although the CAPM has worked well for a long time, however, it is now clear that there are sources of priced risk other than market portfolio. The evaluation of the sensitivities of these other factors to changes in expected return will broaden our understanding of capital markets.

II. Asset pricing

The price of an asset is its *fundamental value* which for reasons of ease and practicality is measured by the asset's *market price* assuming that both will converge over a long horizon period, that the efficient market hypothesis holds and market price follows a Random Walk. The market price under equilibrium should be equal to the present value of the future cash flow accruing due to holding such asset. This therefore means that we have to determine with a certain degree of accuracy the future cash flows as well as the discount rate.

Cochrane [2001] defines Asset pricing as:

“Asset pricing theory tries to understand the prices or values of claims to uncertain payments. A low price implies a high rate of return, so one can also think of the theory as explaining why some assets pay higher average returns than others.”

Asset prices can be explained by two types of models, absolute pricing models and relative pricing models. CAPM and multifactor models are examples of the former and the Black-Scholes option-pricing model is an example of the latter. I quote Cochrane [2001]:

“The central and unfinished task of absolute asset pricing is to understand and measure the sources of aggregate or macroeconomic risk that drive asset prices. Of course, this is also the central question of macroeconomics, and this is a particularly exciting time for researchers who want to answer these fundamental questions in macroeconomics and finance.

A lot of empirical work has documented tantalizing stylized facts and links between macroeconomics and finance. For example, expected returns vary across time and across assets in ways that are linked to macroeconomic variables, or variables that also forecast macroeconomic events; a wide class of models suggests that a *recession* or *financial distress* factor lies behind many asset prices. Yet theory lags behind; we do not yet have a well-described model that explains these interesting correlations.”

The fundamental question is to arrive at the expected return for a particular security so as to price that security. The CAPM, which is an equilibrium model, arrives at this using market index using the concept of mean-variance efficiency. However, other than the problem of determining the composition of the market index, beta may vary over time. APT arrives at this by determining the factor sensitivities to various factors under the conditions that price levels will adjust to eliminate any arbitrage opportunities. In APT, given various factors and security's sensitivity to these factors the expected return can be calculated. Neural Net is a generalized approach that attempts to capture the relationship of the particular security with various macro-economic factors to arrive at the expected return for the security. In NN approach, given past values and future expected values of macro economic factors, the future expected return from a security can be calculated without the sensitivities being explicitly calculated. Further, even a modest ability to forecast financial asset returns can give handsome returns. One US dollar invested in US treasury Bills in January 1926 and the proceeds reinvested every month would give a total value of \$14 by December 1996. The same investment made in S&P500 over the same period would fetch \$1,370. However, if an individual were to forecast in advance for each month, which of these two investments would give a higher return for the next month and switch to that investment then by December 1996 the investment would fetch a total of more than two billion dollars¹

Cochrane [2001] states:

“The absolute approach is most common in academic settings, in which we use asset pricing theory positively to give an economic explanation for why prices are what they are, or in order to predict how prices might change if policy or economic structure changed.”

Kent D. Daniel *et al* [March 2000] indicates:

“Many empirical studies attempt to predict security stock returns using not just risk measures like CAPM beta, but also variables such as book / market that are open to multiple interpretations as either proxies for factor risk or for market misvaluation. The debate over empirical results has been pursued in the absence of an explicit model of the ability of different

¹ Source: Farmer J.D, Lo A.W. [1999] in “Frontiers of Finance: Evolution and Efficient Markets”

proxies to predict returns when there are both misvaluation effects and risk effects among a cross-section of securities.

Furthermore, it has often been argued that mispricing effects will tend to be ‘arbitraged away’ by smart traders. Such arbitrage strategies may include diversification by trading portfolios of mispriced securities, and hedging away of factor risk. The risk and profitability of multi-security arbitrage strategies, and the extent to which these do indeed eliminate mispricing are issues that have yet to be explored in the literature.”

Financial asset pricing can be determined by two types of models: parametric or non-parametric. Parametric methods include the beta method and the stochastic discount factor method. The beta method uses a linear regression to estimate the parameters or betas of the various macroeconomic factors.

$$E_i = r + \sum_{j=1}^k \lambda_j \beta_{ij}$$

r = a constant (which equals the riskless rate of return if there is a riskless asset)

λ = is a vector of risk premia.

The stochastic discount factor (SDF) method defines the price of the financial asset in period ‘t’ as the expected value of the product of the stochastic discount factor for period ‘t +1’ and the payoff on the asset for period ‘t +1’ where the stochastic discount factor is a function of the historical data and model parameters.

$$P_t = E (m_{t+1} x_{t+1}),$$

$$m_{t+1} = f(\text{data, model parameters})$$

Where P_t = asset price,
 x_{t+1} = value of asset at time t + 1
 m_{t+1} = stochastic discount factor

Non-parametric methods include Artificial Neural Networks, which are essentially data driven and statistical methods like Generalised Adaptive Models (GAM) and Projection Pursuit Regression (PPR). GAMs and PPR are non-parametric methods of lagged regression analysis.

III. Capital Asset Pricing Model:

Assumptions:

- The market is made up of risk averse investors who measure risk in terms of standard deviation of portfolio returns.

This assumption is based on the use of standard Normal distribution² However there is now evidence that returns may not follow a Normal distribution. I quote Farmer J.D. [1999]:

“the growing awareness of fat tails is changing the way people characterize risk. Ten years ago, sophisticated quantitative trading firms characterized risk in terms of standard deviation. Increasingly, this is changing to Value-at-Risk (VaR), the size of the loss that would be experienced with an unfavorable move of a given probability”.

- All investors have a common time horizon. Even as a generalization this assumption may be flawed as investors would have different time horizons. I refer to Fractal Market Hypothesis³:
 “If an event occurs that makes the validity of fundamental information questionable, long-term investors either stop participating in the market or begin trading based on the short term information set. When the over-all investment horizon of the market shrinks to a uniform level, the market becomes unstable. There are no long term investors to stabilize the market by offering liquidity to short term investors”
- Homogeneous expectations i.e. all investors are expected to have same expectations about future security returns and risks. It seems to me that, in at least part of the trading activity, the buyer and seller would be having different expectations about future security returns, for otherwise the trade would not take place at all.
- Capital markets are perfect i.e. assets are completely divisible, there are no transactions costs or differential taxes and borrowing and lending rates are equal.

$$E (R_p) = R_f + \beta (E (R_m) - R_f)$$

where $E (R_p)$ = Expected return on portfolio
 R_m = Return on market portfolio
 R_f = Risk free rate of return

I quote Andreas Krause [2001]:

“Models with time varying betas and risk premia have attracted increased attention in recent years. The reason on one hand is the empirical evidence that

²Normal distribution has the following moments:

$$\mu = E(x) = \text{mean}; \quad \sigma^2 = E(x - \mu)^2 = \text{variance};$$

$$\text{Third Moment} = E(x - \mu)^3 = 0$$

$$\text{Fourth Moment} = E(x - \mu)^4 = 3\sigma^4$$

³Peters E.E. (1994) : “Fractal Market analysis – Applying chaos theory to Investment and economics”.

covariances, variances and risk premia are not constant over time. On the other hand the poor performance of the traditional CAPM gives rise to modification of this model.”

In an attempt to correct the deficiencies in CAPM other models have been developed. A variation of the CAPM is the conditional CAPM.

Engle [1982] proposed the ARCH [Autoregressive conditional Heteroskedasticity] model where returns are given by a linear combination of explanatory variables and an error term with mean zero and the variance of the error term is assumed to follow a qth order Auto regressive process.

Bollerslev [1986] proposed a GARCH [generalized Autoregressive conditional Heteroskedasticity] model wherein the variance of the error term follows an (p, q) ARMA [Autoregressive moving average] process.

Engle /Kroner [1995] provided a multivariate version of the GARCH process where a larger number of assets are being investigated.

Nelson [1991] proposed the exponential GARCH (EGARCH) model.

However all these models are difficult to estimate from data and most investigations used GARCH (1,1) for computational simplicity.

IV. Arbitrage Pricing Theory

APT states that the single period expected return on any risky asset is approximately linearly related to its associated factor loadings (i.e. systematic risks) given the following assumptions:

- All investors exhibit homogeneous expectations that the stochastic properties of capital asset returns are consistent with a linear structure of K factors.
- Capital markets are in competitive equilibrium i.e. there are no arbitrage opportunities in the capital markets.
- The number of securities in the economy is either infinite or very large.
- The APT does not specify what the factors are but it is assumed that the number of factors can be correctly estimated by the investigator. APT model asserts that the random rate of return on security i is given by the relationship.

The problem of determination of market portfolio in CAPM is replaced with the uncertainty over the choice of and measurement of the underlying factors in APT. Further non-linear and linear feedback between stock returns and macro-economic variables indicate a complex system. Shanken [1982] has questioned whether an exact risk-return linear relationship is implied in APT. Ross *et al* have identified the following factors: growth rate in industrial production, rate of inflation (both expected and unexpected), spread between long term and short term interest rate, spread between low-grade and high grade bonds.

In a non-equilibrium factor model such as the BARRA⁴ model about 68 fundamental and industry factors are used.

⁴ Developed by Barr Rosenberg and Vinay Marathe at the University of California, Berkeley in the 1970's.

APT model asserts that the random rate of return on security ‘i’ is given by the relationship

$$\tilde{R}_i = E(R_i) + \sum_{j=1}^k \beta_{ij} \delta_j + \tilde{\epsilon}_i$$

where \tilde{R}_i = the random rate of return

$E(R_i)$ = the expected rate of return

β_j = vector of factor loadings on asset i

$\tilde{\epsilon}_i$ = idiosyncratic noise term associated with asset i

δ_j = vector of systematic factors which commonly influence asset returns

Under assumptions that ϵ_i is independent across assets, independent of the factors and has zero mean, and also that δ_j have zero mean and variances exist and further that firm specific risk represents a diversifiable risk which should have a zero price in a market with no arbitrage opportunities, APT reduces to:

$$E_i = r + \sum_{j=1}^k \lambda_j \beta_{ij}$$

where E_i = expected return of ith security

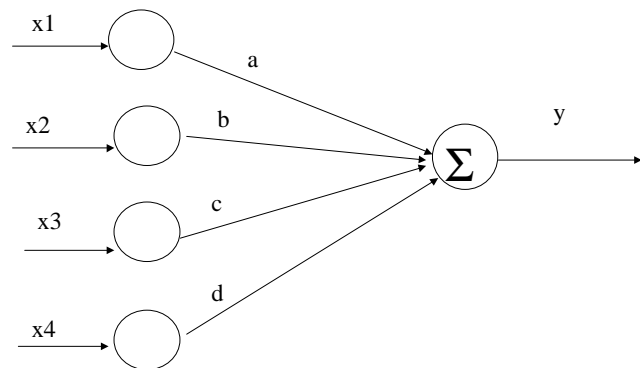
r = a constant (which equals the riskless rate of return if there is riskless asset)

λ_j = vector of risk premia of dimension k

V. Artificial Neural Networks:

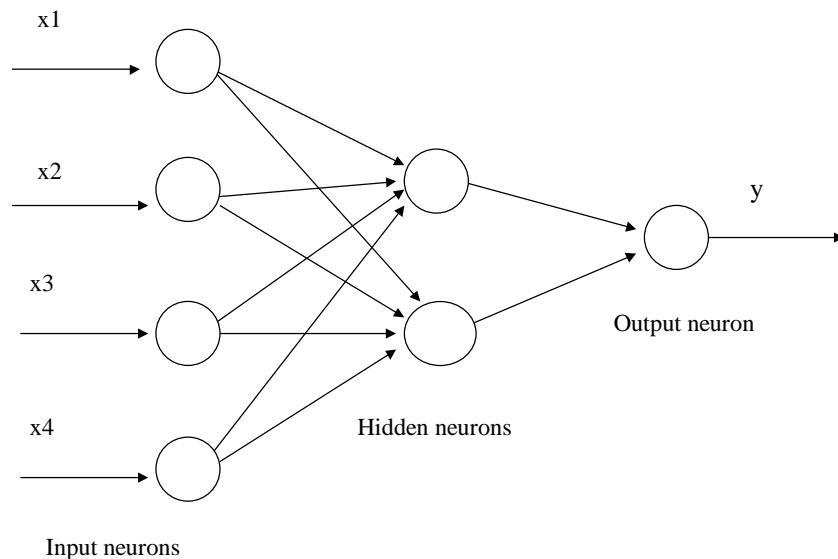
Although the original inspiration for ANN came from biological Neural Networks, ANN are mathematical models that have very little resemblance to biological neural systems. The primary parallel between biological nervous systems and Artificial Neural Networks is that each typically consists of a large number of simple elements that learn and are able to collectively solve complicated and ambiguous problems. The interest in ANN from the viewpoint of econometric forecasting is that Artificial Neural Networks are able to represent highly complex and non-linear equations. To get an intuitive idea about the operation of a simple connected model let us assume a prior linear equation.

$$Y = ax_1 + bx_2 + cx_3 + dx_4$$



This equation can be described by the above model (shown in figure) and the output neuron performs the summation of all inputs multiplied by their respective weights. However, in reality, ANN does not need any prior equation as in statistical regression but are able to map the relationship between input and output vectors by changing the values of the network's weights. If instead of the above linear equation, the output is a non-linear function of the inputs, and assuming that this function is not a-priori known to us, then for the sake of brevity we can write.

$$Y = f(x_1, x_2, x_3, x_4)$$



To model this using ANN we require a set of data, which gives the values of y for each set of (x_1, x_2, x_3, x_4) . This set of data is then divided into a learning set (about 80% of the data) and a test set. Each set of (x_1, x_2, x_3, x_4) is presented to the ANN. Initial random weight parameters are chosen and the output is compared to the value of y . The difference is called the *error term*. The task then is to reduce this error term to a very small pre-specified level which is called *convergence*.

An *epoch* consists of one cycle of presenting each set (x_1, x_2, x_3, x_4) to the ANN. The error term is then minimised by changing the weight parameters. There are many methods by which this is done but the back-propagation method is the most popular. After a large number of epoch's ranging from a few hundred to ten's of thousands, the error term is reduced to a prespecified level. We then say that the ANN has learned the input-output relationship between (x_1, x_2, x_3, x_4) and y . However, it should be understood that the initial layer through which the values of (x_1, x_2, x_3, x_4) are input do not do any processing. The intermediate layers (there can be more than one) are called *hidden layers* and have an *activation function* or *squashing function*. The number of neurons in each hidden layer and the type of activation function to be used depends on the judgement of the modeler. The output layer can also have an activation function and the number of neurons depends on the number of output variables. After the ANN has been trained (after many epochs) the weights are frozen and the test set data are then presented to the ANN to examine how well it is able to predict using the weight values frozen after the training of the ANN.

Statistical tools like multiple regression analysis require the prespecification of an equation either linear or non-linear and can therefore be misspecified if the modeler has no idea of the input-output relationship of the data set being examined. Studies in ANN gained momentum after 1986 following the publication by *Rumelhart et al* of the general framework for an ANN model. However, it is still a developing subject and has some grey areas where our understanding is not complete and is therefore a subject of much research. However, the capacity of ANN to model any complex non-linear relationship makes it a valuable tool in econometric analysis. ANN has been used in such areas as stock price prediction, option trading, bond rating, forecasting financial distress and also portfolio optimization.

Refenes et al [1995] in "*Modelling stock returns in the framework of APT: a comparative study with regression model*" have indicated

"Neural network architectures have drawn considerable attention in recent years because of their interesting learning abilities. They are capable of dealing with the problem of structural instability. Several researchers have reported exceptional results with the use of neural networks. Neural networks are generally believed to be an effective modeling procedure when the mapping from input to the output vector space contains both regularities and exceptions.

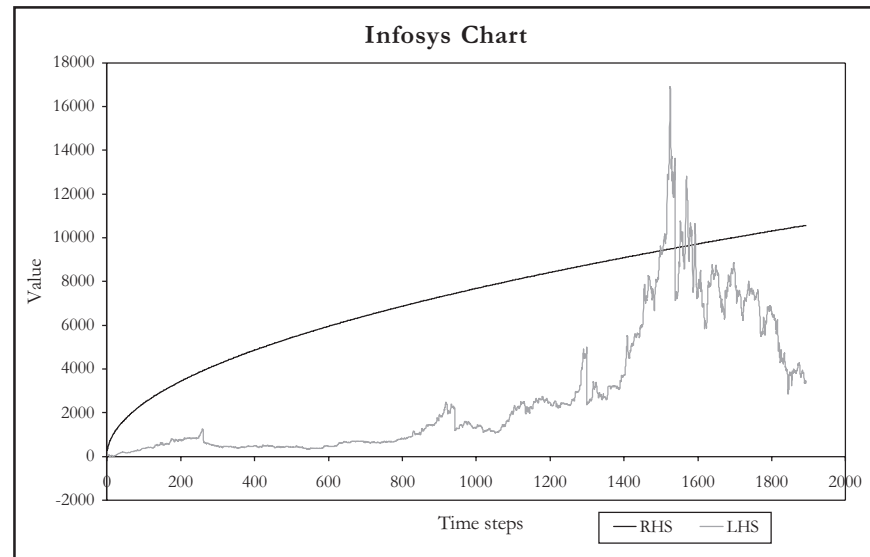
They are, in principle, capable of solving any nonlinear classification problem, provided that the network contains a sufficiently large number of free parameters (hidden units and/or connections)”

Baestaens D.E. et al [1995] have carried out a study of Amsterdam stock-market return using ANN. They conclude their paper with the following remarks:

“We have shown that Multi Layer Backpropagation network analysis is able to unravel dynamic functional relationships between stock-market returns and contextual variables and may be helpful in increasing our understanding of the working of financial markets.”

VI. Random Walk Hypothesis:

A Random Walk model assumes that successive price changes are independent of each other. If market returns are determined by the normal distribution then volatility should increase with the square root of time. However, a study on US data shows that US stocks and bonds were bounded at about 4 years. If the fundamental value of the asset is to be measured by the market value of the asset then the assumptions under which this holds is that the markets follow a random walk and that efficient Markets hypothesis holds. While a rigorous examination of the random Walk Hypothesis in the Indian context is beyond the scope of this paper, a preliminary analysis has been made. The specification test for examination of the Infosys stock for a period from 1993 to 2001 is given here below.



The following specification was used to determine if the series was random.

$p_t = p_{t-1} + u_t$
 u_t is a random variable and i.i.d. having Gaussian or Normal distribution with mean zero and constant variance.

$$E(p_t)^2 = E(p_{t-1} + u_t)^2$$

$$\bar{P}_t^2 = \bar{P}_{t-1}^2 + E(u_t)^2$$

$$\bar{P}_t^2 = \bar{P}_{t-1}^2 + \sigma^2$$

$$\sqrt{(P_t^2 - P_0^2)} = \sigma\sqrt{t} \dots\dots\dots (1)$$

In equation (1) the Left hand side of the equation refers to the square root of the difference of the square of the price of the stock at various times and the Right hand side refers to the standard deviation of the error term. Both the right hand side and left hand side of the above equation are shown separately in the above chart indicating the deviations. No attempt has been made to examine for heteroskedasticity.

VII. Empirical study using India Data

The empirical study assumes that asset returns are related to certain systematic economy wide factors plus a unique risk or noise. No presumption of linearity of relationship is made.

$$R_i = f(x_1, x_2, x_3, \dots, x_n) + u_i$$

The empirical study examines the relationship of macro economic factors to the returns of individual assets using Artificial Neural Networks. The BSE sensx index as well as LT stock has been examined. The results are given in the appendix.

VIII. Conclusions

The study has been constrained due the lack of availability of high quality data in the Indian context. This would invariably lead to some errors as a complete set of factors could not be examined. The analysis of the results of the preliminary analysis for Random walk of Infosys stock shows that the period after December 1998 has deviation from a theoretical Random Walk. This deviation could be due to the process not following a Random Walk or due to heteroskedasticity or non-normality.

The results of the BSE study show that the functional relationship of the factors to P/E of the BSE sensx is non-linear and complicated. It should be noted that the resulting sensitivity graphs are generated only from the data without the prespecification of any model.

References:

1. Apostolos-Paul Refenes, 1995, "Neural Networks in the Capital Markets". John Wiley and sons, England
2. Baestaens D.E., Willem Max van den Bergh, "tracking the Amsterdam Stock Index using Neural Networks" in "Neural Networks in the Capital Markets" edited Refenes A.P. John Wiley & sons.
3. Campbell J.Y., Lo Andrew W., MacKinlay A.C. 1997 "The econometrics of Financial Markets"
4. Chen, N., 1983, "Some Empirical Tests of Arbitrage Pricing." Journal of Finance, Vol 38, 1393-1414.
5. Chen, N., R.Roll, S.A.Ross, "Economic forces and the stock market". Journal of Business, Vol 59, 383-403.
6. Cochrane, J.H., [2001] "Asset Pricing". pp3 - 48, Princeton University Press.
7. Daniel, K.D, Hirshleifer, D., Subrahmanyam, A., [2000] "Covariance risk, mispricing, and the cross section of security returns". Working Paper 7615, National Bureau of Economic Research.
8. Ferson Wayne E., Jagannathan Ravi [1996]: "Econometric evaluation of asset pricing models". Research report, Federal Reserve Bank of Minneapolis.
9. Hodrick Robert J., Zhang Xiaoyan [2000]: "Evaluating the specification errors of asset pricing models". Working paper, National Bureau of Economic Research.
10. John Wei, K.C., 1988, "An Asset-Pricing Theory Unifying the CAPM and APT". Journal of Finance, Vol 43, 881-892.
11. Kaastra Ibeling , Boyd Milton [1995]: "Designing a Neural Network for forecasting financial and economic time series". *Neurocomputing* 10(1996).
12. Kamath M. V. [1999]: "An Introductory Survey of Artificial Neural Networks for Application in Finance and Economics". M.B.M. Thesis, IIT Kharagpur.
13. Krause Andreas [2001] "An overview of Asset pricing models" Working paper, School of Management, University of Bath.
14. Kutsurelis Jason E. [1998]: "Forecasting financial markets using Neural networks: an analysis of methods and accuracy" MS thesis, US Naval postgraduate School.
15. Peters E.E., 1994, "Fractal Market Analysis – Applying Chaos Theory to Investment and economics". John Wiley & Sons.
16. Peters E.E., 1996, "Chaos and order in the Capital markets- A new view of Cycles, prices and Market volatility" (second edition), John Wiley & Sons.
17. Philip H.Dybvig , Stephen Ross, 1985 : "Yes , The APT Is Testable". The Journal of finance, Vol 40, 1173-1188.
18. Shanken, J., 1982, "The Arbitrage Pricing Theory: Is it Testable?". Journal of Finance, Vol 37, 1129-1140.
19. Trzcinka, C., 1986, "On the numbers of Factors in the Arbitrage Pricing Model". Journal of Finance, vol 41, 347-368.

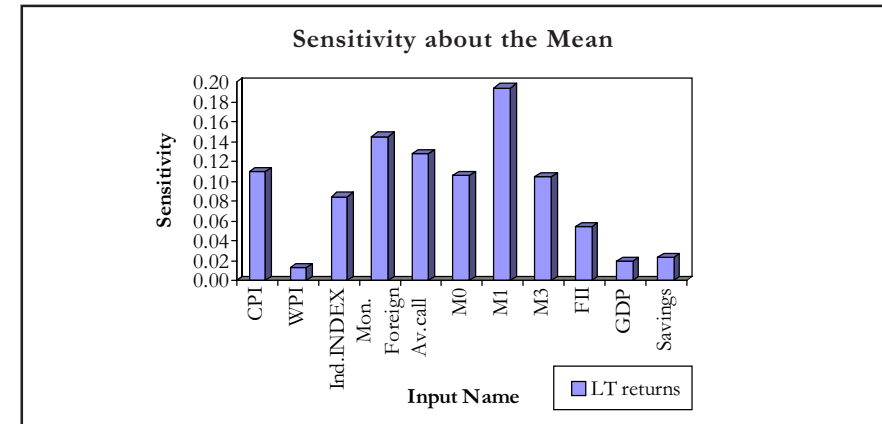
Appendix

In examining the P/E ratio of the BSE Sensex the following factors have been used:

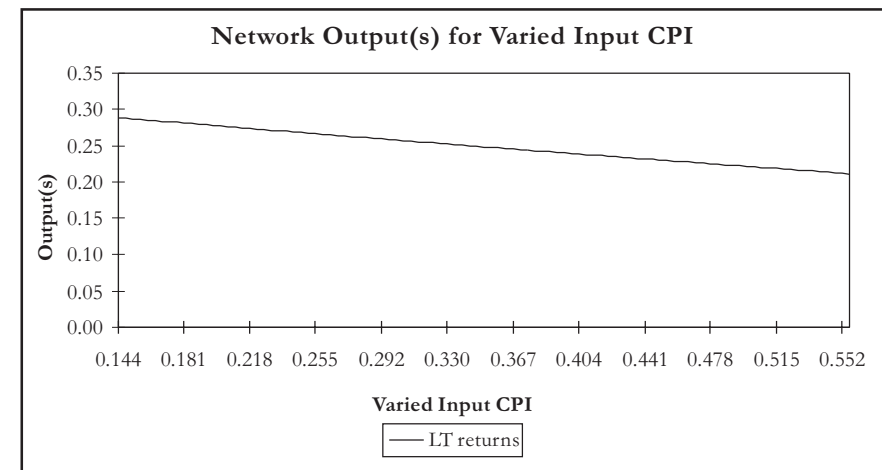
1. Consumer price index
2. Whole sale price index
3. Industrial index
4. Monthly foreign exchange reserves
5. Monthly average call money rate
6. M0 = currency in circulation + Bankers and other deposits with RBI
7. M1= currency with public + demand deposits a + other deposits with RBI.
8. M3 = M1 + time deposits
9. Foreign institutional investments
10. Gross Domestic product
11. Domestic savings.

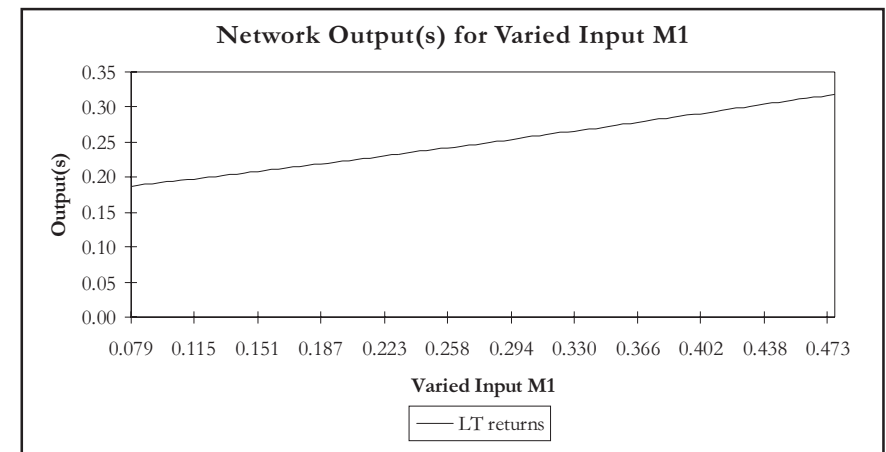
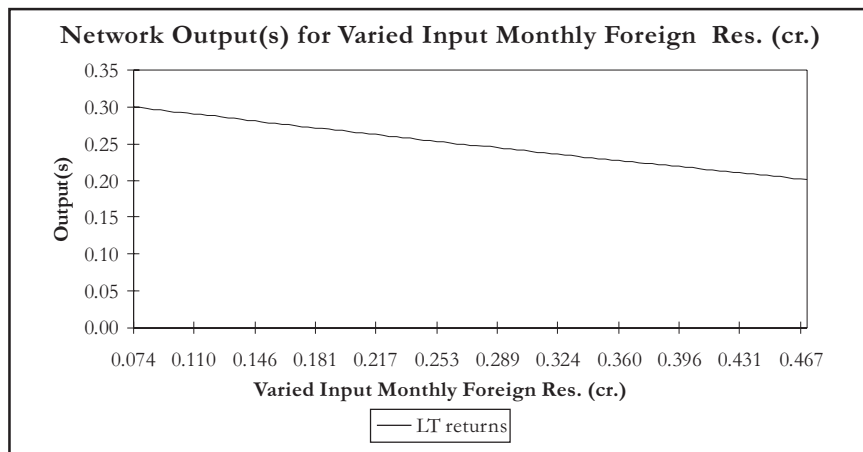
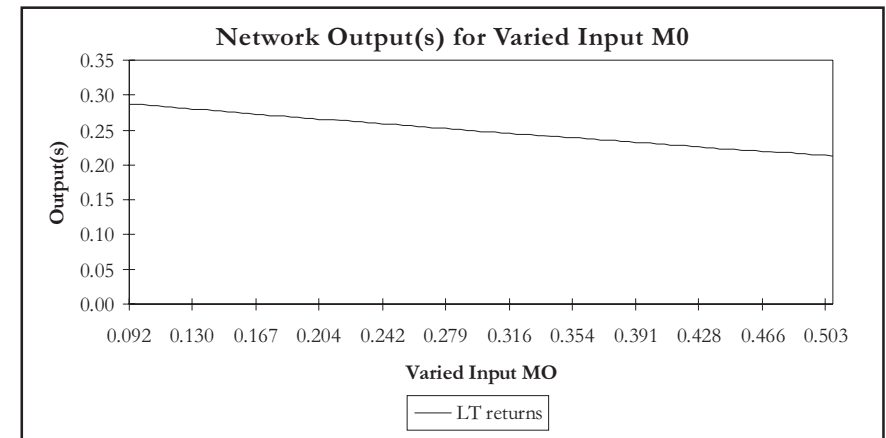
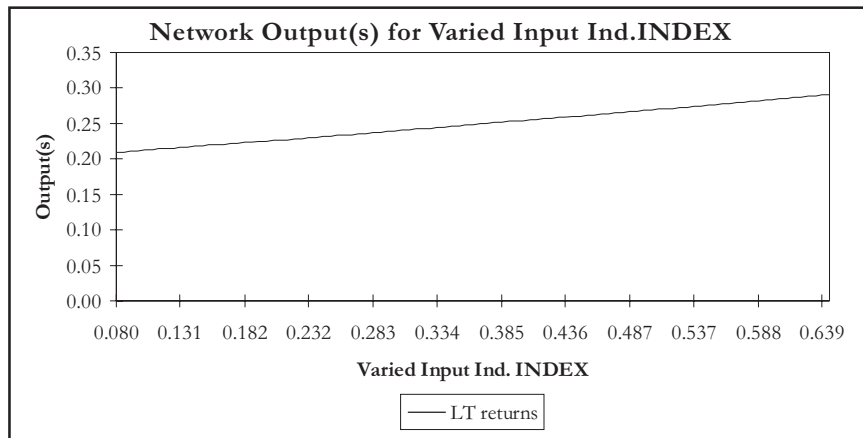
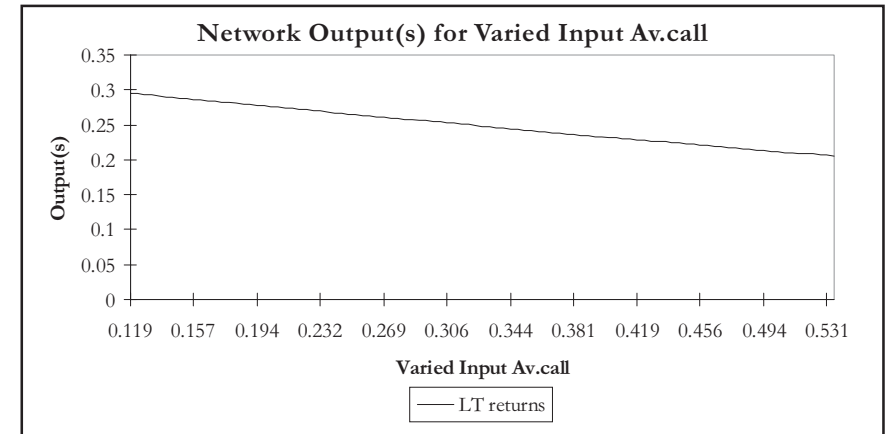
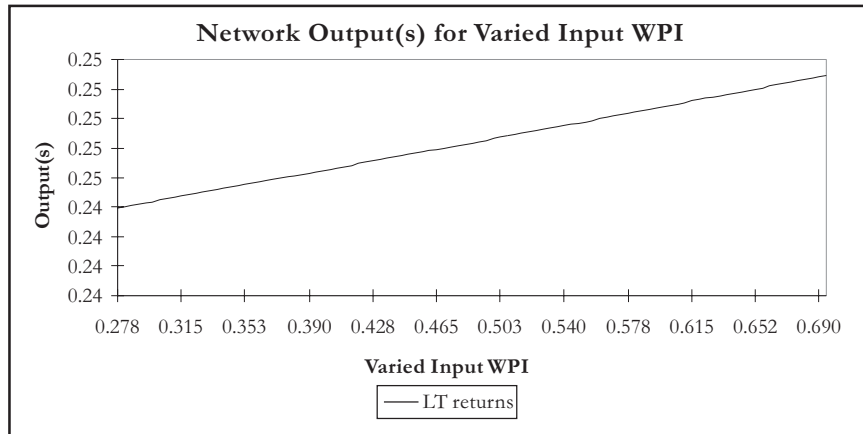
A multi layer perceptron network was used with the following configuration: 11 input neurons, 4 hidden neurons with sigmoid activation function, one output neuron with sigmoid activation function. Pre processing of the data included normalizing the data to the range [0,1]. Training data (monthly) was from April 1990 to September 1997 and test data was from October 1997 to March 1998.

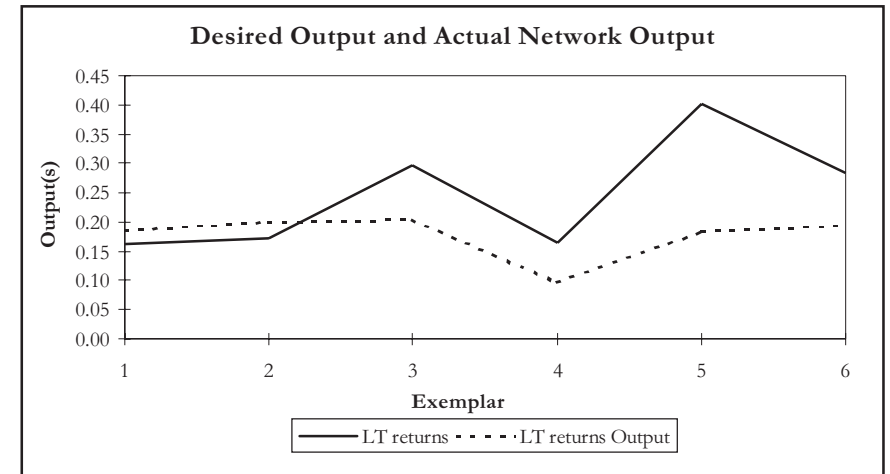
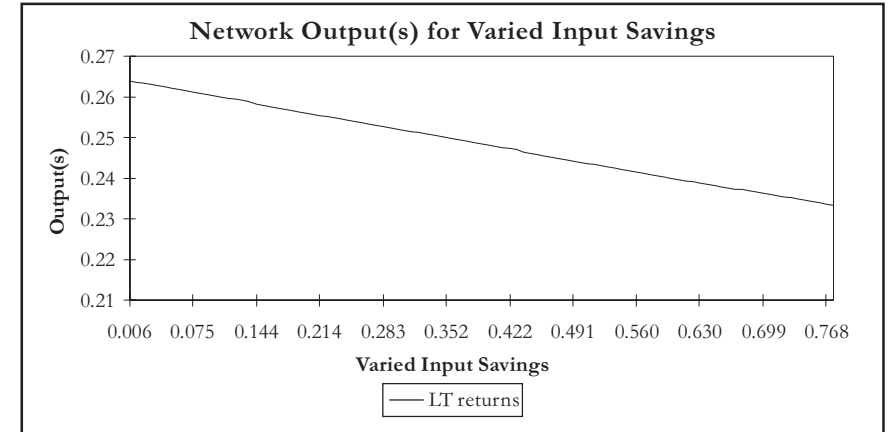
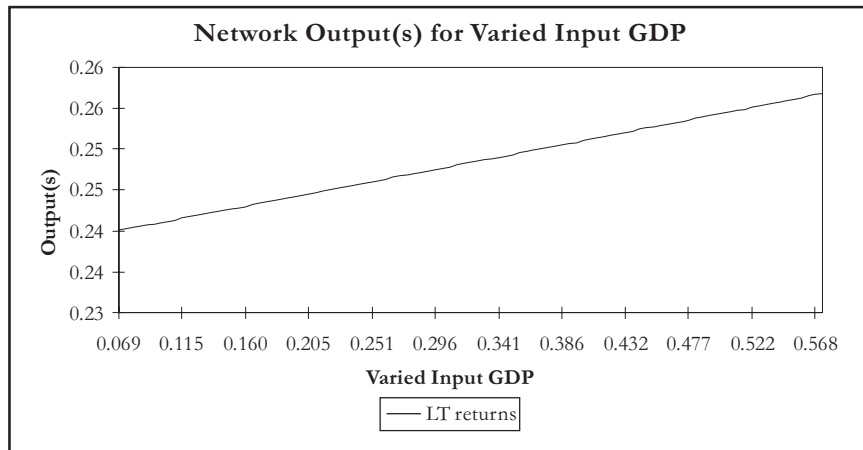
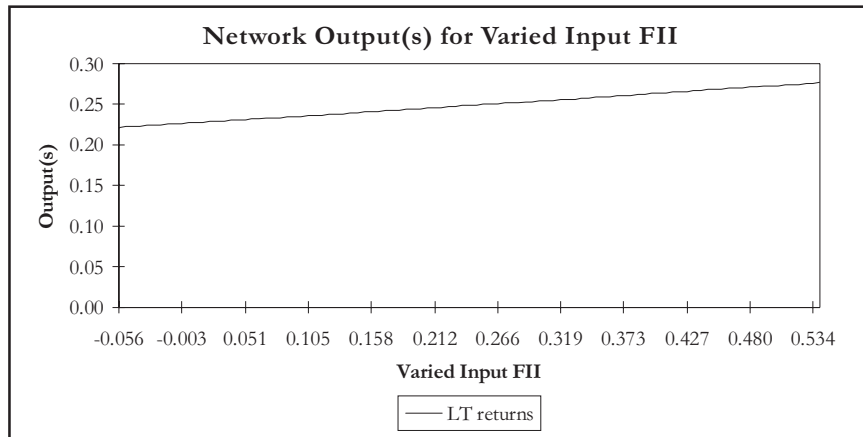
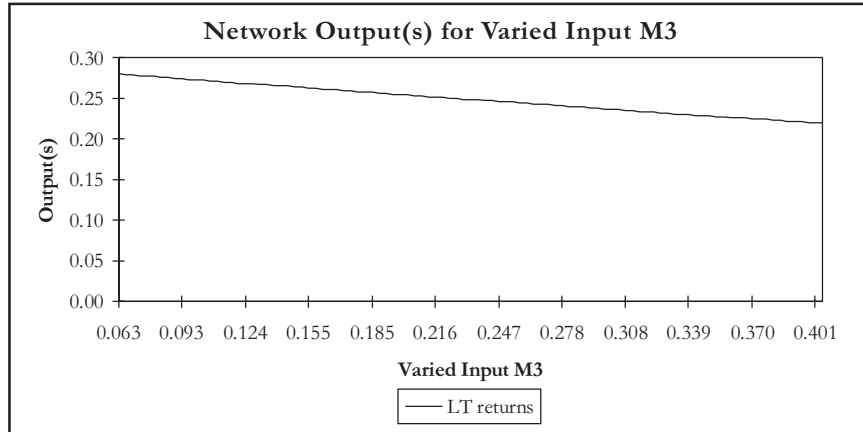
In examining the returns on LT during above period the above 11 factors have been used. However a generalized feed forward neural network with two hidden layers and jumping connections has been used. The configuration is as follows: 11- input neurons, 5 sigmoid layer one



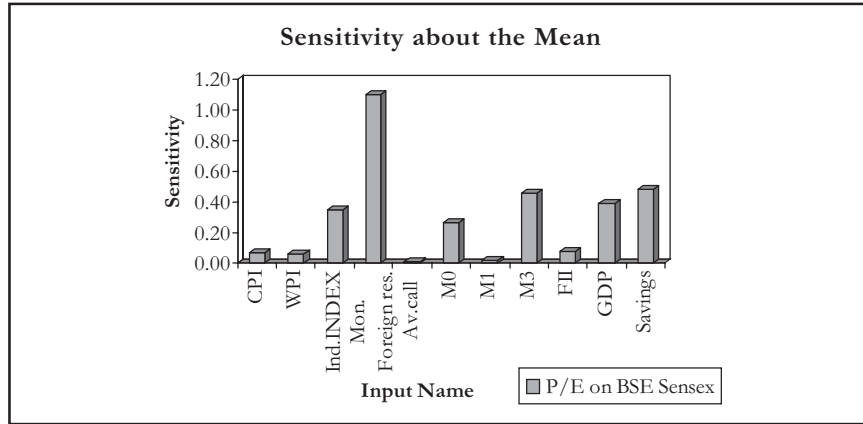
<i>Sensitivity</i>	<i>LT returns</i>
CPI	0.109953642
WPI	0.012613669
Ind.INDEX	0.084565975
Mon. Foreign Res. (cr.)	0.145320013
Av.call	0.12737447
M0	0.105685815
M1	0.194216639
M3	0.104476102
FII	0.054202203
GDP	0.019500101
Savings	0.023411894



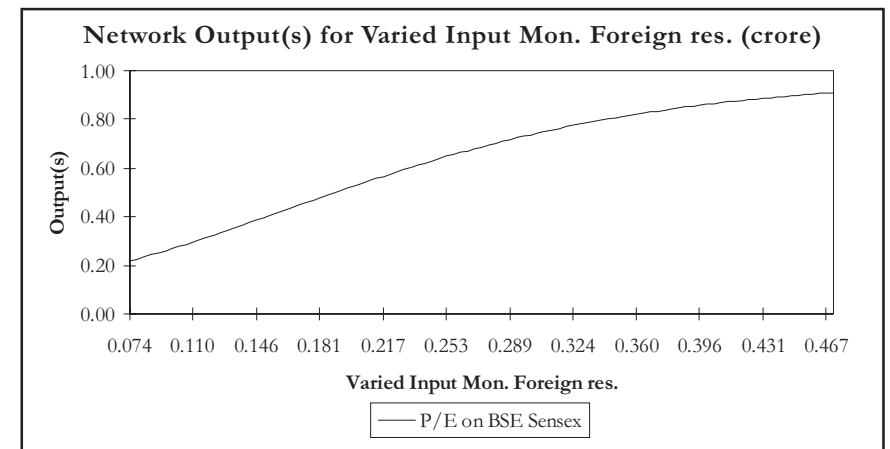
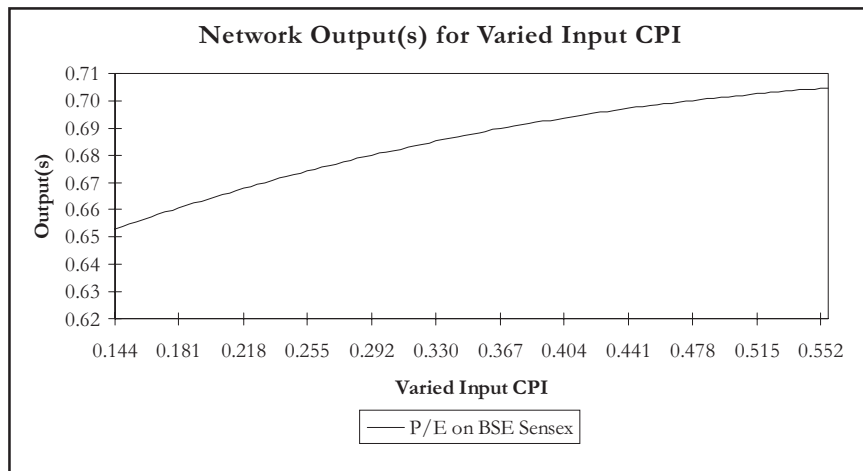
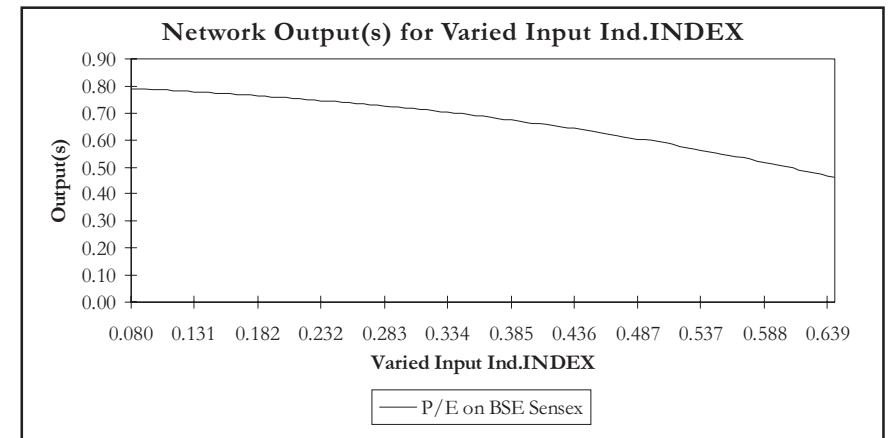
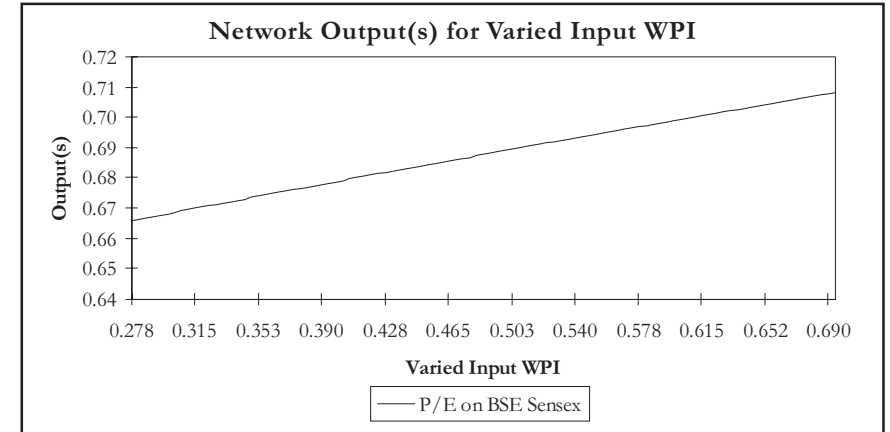


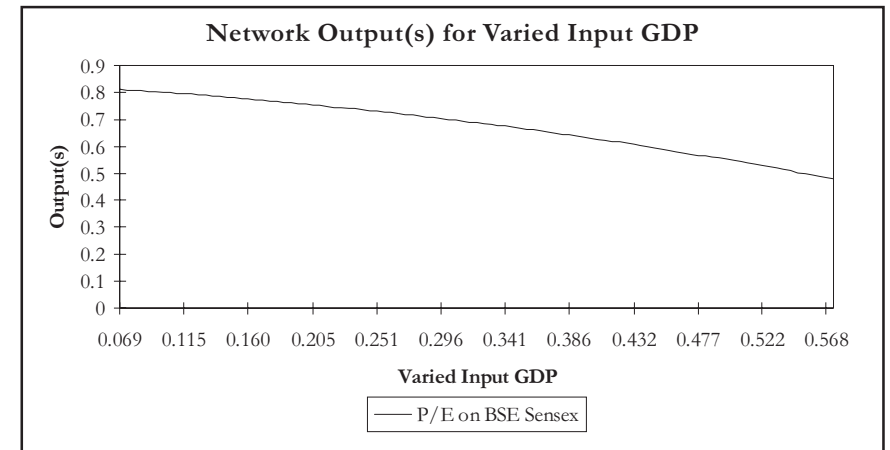
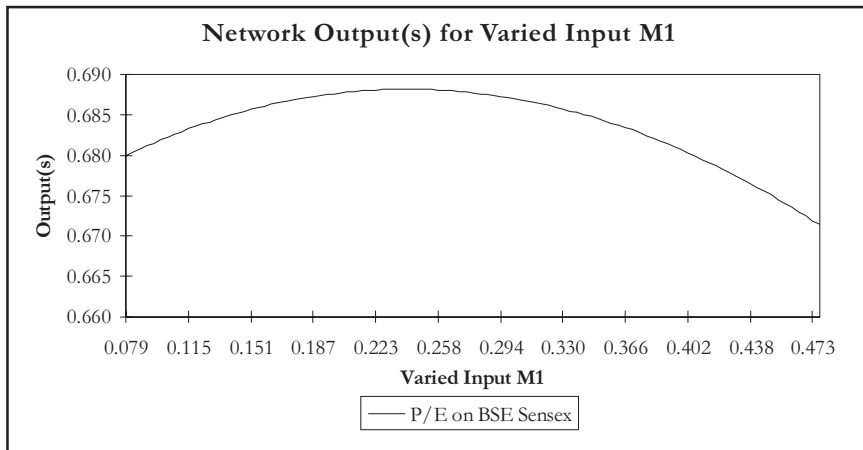
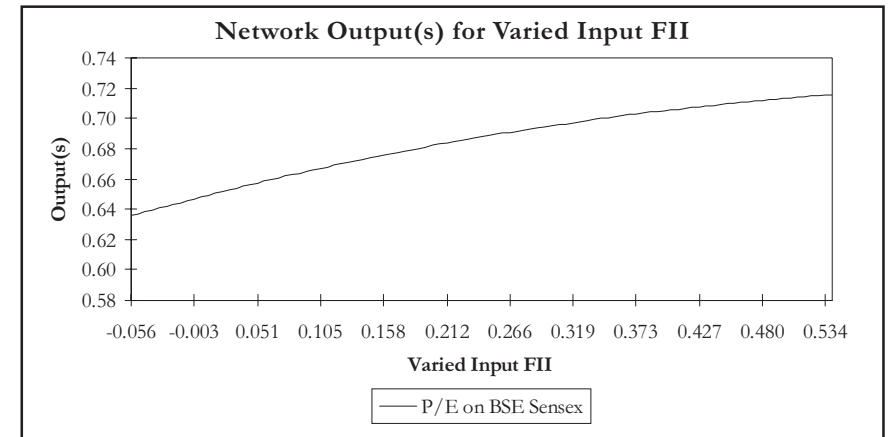
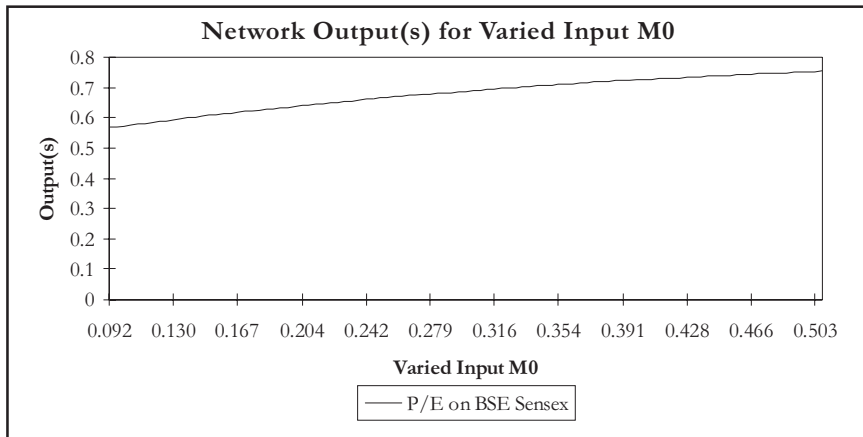
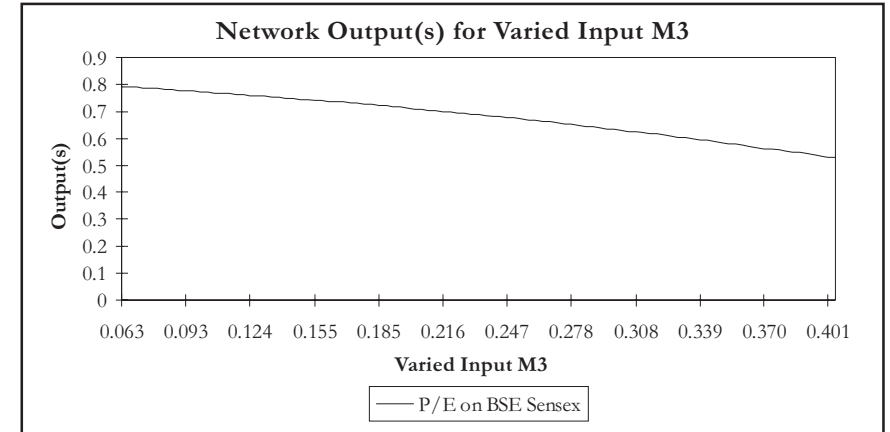
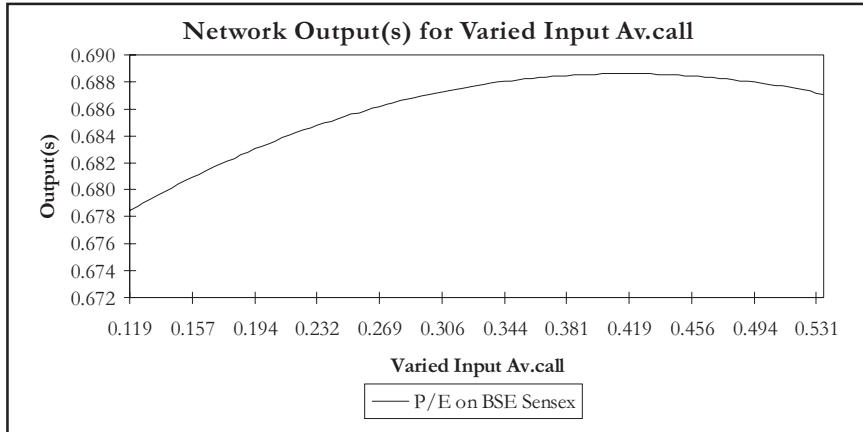


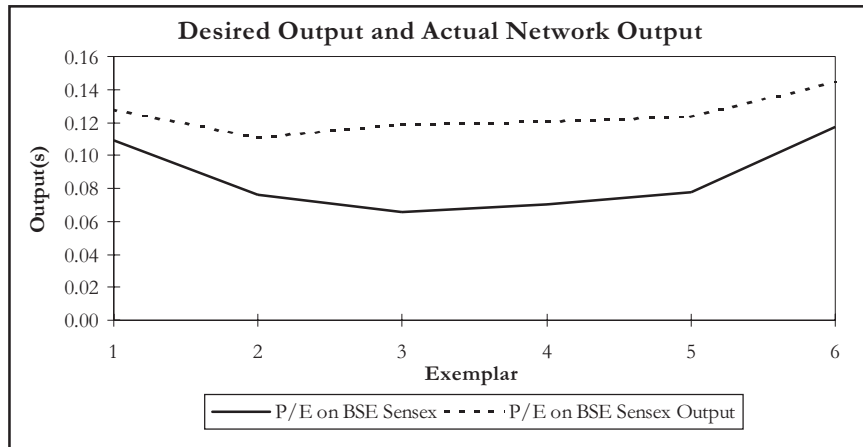
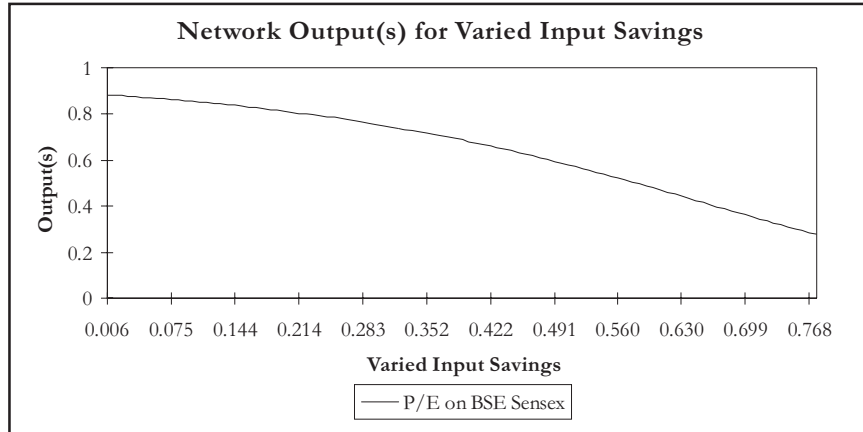
<i>Performance</i>	<i>LT returns</i>
MSE	0.011799372
NMSE	1.500996872
MAE	0.086656998
Min Abs Error	0.021930292
Max Abs Error	0.219578258
r	0.359098544



<i>Sensitivity</i>	<i>P/E on BSE Sensex</i>
CPI	0.073848091
WPI	0.059373919
Ind.INDEX	0.342368484
Mon. Foreign res. (crore)	1.094350576
Av.call	0.013990178
M0	0.263882577
M1	0.022682317
M3	0.455477446
FII	0.079190604
GDP	0.388252527
Savings	0.481378436







<i>Performance</i>	<i>P/E on BSE Sensex</i>
MSE	0.001614439
NMSE	4.100545055
MAE	0.038173809
Minimum Absolute Error	0.018939063
Maximum Absolute Error	0.05308755
r - Linear correlation coefficient	0.836792494