Chapter I

Introduction, Background and Objectives of the Study

I am not interested in developing a powerful brain
All I'm after is just a mediocre brain...

1.0 Introduction

The question "what is the need of forecasts at all?" has engrossed many economists and managers for many years. Forecasts are essential for two basic reasons: the future is uncertain and the full impact of many decisions taken now is not felt until later. If a person is told that he or she is to relocate to a new city, he or she might very well be interested in knowing what life will be like there. It is for the same reason that every one should be highly interested in the future - for it is where they are going to spend the rest of their lives and it is likely to be quite different from the past or present. However, for economists and managers, knowing the future is not merely a matter of academic interest. Economics and Management are decision sciences: they are concerned with sensible decision making and the effects of decisions. Any decision has to take a viewpoint about what will occur in the future and thus comes the idea of forecasts. But forecasting is generally not easy for the reason that the methods that can be used to forecast can vary greatly and will depend upon data availability, the quality of models available, and the kinds of assumptions made, amongst other things. This is one reason why the topic is so interesting.

Two aspects of forecasting that are of particular importance are the degree of predictability, and forecasting accuracy. The degree of predictability of variables is one important aspect of forecasting. An astronomer can predict just where every star will be at half past eleven tonight; he can make no such prediction about his daughter. Even this simple avowal brings out an important point about forecasting. It states that variables to be forecast vary greatly in their degree of predictability. An astronomer can have much greater confidence in forecasting the movement of a star than the
The movement of his own daughter. The reason is that the movement of the star is more predictable than the movement of his own daughter. To generalize, it can be said that some variables can be predicted to a considerable degree of certainty, and others are almost entirely unpredictable.

The other aspect, forecasting accuracy is of major interest to anyone concerned with the future. On one side, increasing forecasting accuracy could facilitate the saving of millions of dollars and becomes a major motivation for using formal forecasting methods. On the other, forecasting errors are inevitable, because forecasting is not a substitute for prophecy. Forecasts unfortunately do not possess crystal balls enabling them to look into the future. What is extremely important here, from a practical point of view, is to be able realistically to assess the advantages and limitations of forecasting methods and to take them into account while utilizing prediction for some purposes. A forecasting method with less error and greater accuracy should be adopted and appreciated.

Keeping in view the importance of forecasting, a study dealing with the prediction of exchange rates and stock returns is of obvious interest as both practitioners and policy makers always keep a close watch on the movement of exchange rates and stock prices. However, forecasting exchange rates and stock returns is rather a hazardous operation as these variables are notorious for their unpredictability, and are often characterized by high volatility, complexity, noise, nonstationarity, non-linearity and chaos. The view of their unpredictability has been further reinforced by the efficient market hypothesis (EMH), which found broad acceptance in the financial community.

According to efficient market hypothesis, a market is said to be efficient if price at any point of time contains all available information about the market. In its strong form, it says the current price of an asset reflects all publicly available information, while in its weak form it says the current price reflects all of the information gleaned from past prices of asset. Thus there is no arbitrage possibility, and the movement of a price is unpredictable. The EMH is based on the assumption
that all news is promptly incorporated in prices; since news is unpredictable by
definition, prices are unpredictable. The best prediction for a price is the current price
and the actual prices follow what is called a random walk. In other words, it claims
that exchange rates and stock prices follow a random walk and hence can not be
predicted from their past prices. If a market is efficient, price will show a lower
degree of volatility.

However, whether the financial markets are truly efficient and follow a
random walk has always been a matter of empirical investigation. Much effort has
been expended in trying to prove or disprove the EMH. Current opinion is that the
theory has been disproved (Taylor, 1994 and Ingber, 1996) and much evidence
suggests that financial markets are not efficient (Malkiel, 1996 and Lo et al. 1999).
Many financial theorists believe that the death-knell for the random walk model (for
financial markets) occurred in 1987, October 19, when the collapse (in tandem) of the
world equity markets occurred. Now, all the academics who were grounded in an
efficient market model felt that there must be something drastically wrong with the
theory if there could be a 500-point move in the Dow in one day. These markets could
not be efficient if they reflected all available information on Friday, October 16. This
ignited enthusiasm among financial economists to make an attempt to forecast
exchange rates and stock prices. Various alternative linear models such as linear
regression, vector autoregression (VAR), and ARIMA models were then used for
forecasting these variables. But unfortunately these linear models failed to improve
upon the simple random walk model in out-of-sample forecasting of both stock prices
and exchange rate prices. The reason could be that all the models investigated were
linear and it was natural to conjecture that exchange rate and stock price data contain
non-linearities, which may not be accounted for or approximated well by linear
models. This inevitably called for the emergence of a non-linear model, which could
capture the non-linearities in financial variables.

Following this need, a variety of parametric non-linear models such as
autoregressive conditional heteroskedasticity (ARCH), general autoregressive
conditional heteroskedasticity (GARCH), and self-exciting threshold autoregression
models have been proposed and applied to financial forecasting. While these models may be good for a specific data series, they do not have general appeal for other applications. Because there are too many possible nonlinear patterns, the pre-specification of the model form restricts the usefulness of these parametric nonlinear models.

Attempts to do away with these deficiencies gave birth to the use of a powerful nonparametric non-linear model called artificial neural network (ANN). Neural networks for forecasting exchange rates and stock prices have been proposed and examined in recent years. Broadly speaking, artificial neural network is a nonlinear, nonparametric and data driven modeling approach. It allows one to fully utilize the data and let the data determine the structure and parameters of a model without any restrictive parametric modeling assumptions. ANN is appealing in financial area because of the abundance of high quality financial data and the paucity of testable financial models. As the speed of computers increases and the cost of computing declines exponentially, this computer intensive method becomes attractive.

Neural network has some advantages over other linear and nonlinear models which make it attractive in financial modeling. First, neural network has flexible nonlinear function mapping capability which can approximate any continuous measurable function with arbitrarily desired accuracy (Hornik et al, 1989; and Cybenko, 1989), whereas most of the commonly used non-linear time series models do not have this property. Second, being a nonparametric and data-driven model, neural network imposes few prior assumptions on the underlying process from which data are generated. Because of this property, neural network is less susceptible to model misspecification problem than most parametric nonlinear methods. This is an important advantage since exchange rates and stock returns do not exhibit a specific nonlinear pattern. Third, neural network is adaptive in nature. The adaptivity implies that the network's generalization capabilities remain accurate and robust in a nonstationary environment whose characteristics may change over time. Fourth, neural network model uses only linearly many parameters, whereas traditional
polynomial, spline, and trigonometric expansions use exponentially many parameters to achieve the same approximation rate (Barron, 1991).

1.1 Some Interesting Aspects of Exchange Rate and Stock Return

Since 1973, with the abandonment of the fixed exchange rates and the implementation of the floating exchange rate system by industrialized countries, researchers have been striving to explain the movement of exchange rates. Foreign exchange rates are affected by many highly correlated factors. These factors could be economic, political and even psychological factors. The interaction of these factors is in a very complex form. Consequently, to forecast the changes of foreign exchange rates is generally very difficult. Again, the interaction between economic systems of different countries and the consequences of globalization phenomenon in economies, increase the dimensionality of factors affecting exchange rates, which makes the accurate prediction of the exchange rate difficult.

Before going to the prediction of foreign exchange rate one should be aware of the general properties it possesses. The first property is nonlinearity. A number of empirical studies (e.g. Hsieh, 1989; De Grauwe et al, 1993; Brooks, 1996; and Drunat et al, 1996) have uncovered significant nonlinearities in nominal bilateral exchange rates and some authors have estimated nonlinear models for these nominal rates. Economic theory offers a number of potential explanations for the presence of nonlinearities and cycles in exchange rates. Heterogeneity of participants in the foreign exchange market is often cited as the major source of nonlinearities in the exchange rate process. Market imperfections such as taxes, transaction costs and the timing of the information reaction introduce nonlinearities in the foreign exchange market. This means market participants do not trade every time news arrives in the market, rather they trade whenever it is economically feasible, leading to clustering of price changes. Furthermore, nonlinearities are observed when announcement of important factors are made less often than the sampling frequency. For example, weekly money supply announcements will cause nonlinearities in daily but not in monthly data. The other
sources of nonlinearities are discrete regime shifts, time varying coefficients and data generating process that is inherently nonlinear.

Second, the exchange rates are found to be unconditionally leptokurtic. Leptokurtosis in exchange rates has been appreciated since Mandelbrot (1963). Similarly, Westerfield (1977), Boothe et al (1987), Hsieh (1988), and Diebold and Nerlove (1989) all show that exchange rates are leptokurtic. A number of explanations have been advanced for the leptokurtosis and volatility clustering. To cite one, Clark (1973) in his seminal work shows that subordination to an i.i.d. information-arrival process produces leptokurtic returns.

Third, it is widely agreed that exchange rate returns are conditionally heteroskedastic. Diebold (1988), for example, examines seven nominal dollar spot rates and finds little linear predictability but conditional heteroscedasticity implied by strong ARCH effects in all of them. Cumby and Obstfeld (1984), Domowitz and Hakkio (1985), Diebold (1988), Hsieh (1989) and Engel et al (1990) also find conditional heteroscedasticity in the residuals of exchange rates. The evidence of nonlinear dependencies, leptokurtosis and heteroscedasticity goes against the random walk hypothesis and further indicates the predictability of exchange rates.

Given these general properties, obviously benchmark models such as linear regression, ARIMA and random walk models are unlikely to give good predictions of the exchange rate. However, using a nonlinear model or introducing nonlinearities in exchange rate models does not necessarily improve the forecasting of exchange rates as these effects operate through even ordered moments. For example, using non-parametric kernel regression, Diebold and Nason (1990) and Meese and Rose (1990) were not able to improve upon a simple random walk in the out-of-sample prediction of 10 major dollar spot rates in the post-1973 float period. Meese and Rose (1991) examine five structural exchange rate models to account for potential nonlinearities. Their results indicate that although nonlinear effects may be important in the even moments of the exchange rate dynamics, the incorporation of nonlinearities into the
structural models of exchange rates does not improve our ability to understand exchange rate movements.

A number of explanations have been offered for this happening. One, of course, is that the nonlinearity may be present in even ordered conditional moments and therefore are not useful for point prediction. Only nonlinearity in conditional mean can improve the point prediction. Second, in-sample nonlinearities such as outliers and structural shifts may be present, and may cause rejection in various linearity tests while nevertheless being of no use for Out-of-sample forecasting. Third, very slight conditional-mean nonlinearities might be truly present and be detectable with large data sets, while nevertheless guiding negligible ex ante forecast improvement. Finally, even if conditional nonlinearities are present and are important, the overwhelming variety of plausible candidate nonlinear models makes determination of a good approximation to data generating process (DGP) a difficult task. The seemingly large variety of parametric nonlinear models that have received attention lately (e.g. bilinear, threshold and exponential autoregressive) is in fact a very small subset of the class of plausible nonlinear DGPs.

Similarly, stock returns are also well described as nonlinear, heteroscedastic and leptokurtic. Forecasting stock returns is also a difficult job for the same reasons cited in the case of exchange rates. A great deal of effort has been devoted to developing systems for predicting stock returns in the capital markets. Limited success has been achieved. It is believed that the main reason for this is that the structural relationship between an asset price and its determinants changes over time. However, using linear models Chen et al (1986), Campbell (1987), Fama and French (1988, 1989), Whitelaw (1994) find that stock returns are predictable by publicly available information such as time series data on financial economic variables. Their findings go against the efficient market hypothesis and random walk model, which advocate for the unpredictability of asset returns. More importantly, Hinich and Patterson (1985), Abhyankar et al (1997) have indicated the presence of structural nonlinear dynamics. Pesaran and Timmermann (1994) find significant nonlinear effects in quarterly and monthly regression of excess stock returns on economic
variables and some non-linear terms such as the lagged values of the square of returns.

1.2 Need for the Present Study

We can summarize the above discussion to say that random walk hypothesis in financial community has been refuted but there also appears to be little linear predictability of exchange rates and stock returns. Both exchange rates and stock returns are seen to contain nonlinearities, which may not be accounted for or approximated well by linear models. However, it is difficult to zero in on a particular data generating process, given the large class of plausible nonlinear data generating processes. It is in this context that the use of artificial neural network, which does not presume any specific form of nonlinearity, gains importance. It would be of interest to see whether the predictability of exchange rates and stock returns can be improved upon by using artificial neural network, a nonlinear model. A number of studies have been carried out on expediency of neural network in predicting exchange rate and stock returns overseas, but very few of them are done in the context of India. The only studies that come to mind are Nag and Mitra (2002), Bordolai (2001) and Kamath (2001). There is a paucity in the number of studies, and even those studies that exist are not extensive enough to prove the credibility of neural network as an alternative model of forecasting financial variables in India. It is for this reason that a study exploring the possibility of artificial neural network in predicting exchange rates and stock returns assumes immense importance. The present work proposes to do just this.

1.3 Objectives of the Study

In the light of the above discussion, the chief objectives of the study are set as follows

To use artificial neural network in forecasting daily and weekly exchange rate returns and to compare its predictive ability with benchmark models such as linear autoregressive and random walk models.
To use artificial neural network in forecasting daily and weekly stock returns and to compare its predictive ability with linear autoregressive and random walk models.

To improve the generalization or Out-of-sample performance of neural network through Bayesian regularization and early stopping technique in forecasting exchange rate and stock returns.

1.4 Nature and Sources of Data

The study has made use of daily and weekly data on closing spot rates of Indian rupee/US dollar. We have used Indian rupee/US dollar exchange rate for our study keeping in view the fact that US is the major trading partner of India. And of course, the US dollar is regarded as the best hedging currency in the world. The data are collected from Pacific FX database. The daily data consisting of 2396 observations covers the period from January 4, 1993 to July 12, 2002. The weekly data consisting of 497 observations covers the period of January 6, 1993 to July 10, 2002. The study period has been chosen purely on the basis of availability of data, and keeping in mind that neural network estimation requires a long time series data. The exchange rate returns are calculated as the log first difference of the levels.

As far as stock market data is concerned, the study has taken BSE Sensitive Index (Sensex) as the representative. This comprises 30 companies from specified group in Bombay Stock Exchange constituting 150 shares. The selection is made on the basis of liquidity, depth, floating stock adjusted depth and industry representation. The compilation of the index is based on the weighted aggregate method. The data on daily and weekly closing values of BSE Sensitive Index are collected from web pages of Bombay stock exchange (BSE). The daily stock price data consisting of 2553 observations covers the period from January 2, 1991 to December 31, 2001. The weekly stock price data consisting of 558 observations covers the period from January 3, 1992 to November 8, 2002. Here too, the study period has been chosen on the basis of availability of a long time series data. The stock returns are calculated by taking the logarithmic differences between successive trading days.
1.5 Organization of the Thesis

The remainder of the thesis is organized as follows. In chapter II, alternative forecasting techniques are discussed. This includes a detailed discussion on artificial neural network and a brief discussion on linear autoregressive and random walk models. In this chapter, we have also reviewed some select works on the expediency of neural network in forecasting exchange rates and stock returns.

Chapter III presents our empirical findings on forecasting daily and weekly exchange rates by neural network, linear autoregressive and random walk models. The in-sample and Out-of-sample performances of three studied models are compared. Besides, the effects of forecast horizon on the in-sample and Out-of-sample performances of neural network, linear autoregressive and random walk models are carried out in this chapter. This chapter also tests the statistical significance of the forecasting results of all three studied models by using forecast encompassing test.

In chapter IV, we have compared in-sample and Out-of-sample forecasting of daily and weekly stock returns by neural network, linear autoregressive and random walk models. Forecast horizon effects and forecast encompassing tests are also performed in this chapter.

Chapter V is devoted to increase the Out-of-sample performance of neural network in forecasting exchange rate and stock returns by using Bayesian regularization and early stopping technique. Finally, chapter VI concludes the thesis with a summary of major findings and the implications of the study. The scope for further research is also discussed, together with the limitations of the present study.