ESTIMATION OF TIME VARYING RISK PREMIA
FOR THE NIKKEI 225 STOCK INDEX FUTURES CONTRACTS

by

JUNGMANN LEE

A dissertation submitted to the Graduate Faculty in Economics
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy, The City University of New York

1997

UMI Number: 9732941

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

This manuscript has been read and accepted for the Graduate Faculty in Economics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

April 29, 1997

Date

Chair of Examining Committee

Apr 30, 1997

Date

Executive Officer

Thom Thurston

T. D. Agbeyegbe

Salih N. Neftci

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

ESTIMATION OF TIME VARYING RISK PREMIA FOR THE NIKKEI 225 STOCK INDEX FUTURES CONTRACTS

by

Jungmann Lee

Advisor: Professor Thom Thurston

The purpose of this research is to examine the intertemporal relation between risk and expected equity returns. In particular, I investigate whether the excess holding period return on the Nikkei 225 stock index, defined as the expected return on a stock index minus the risk-free interest rate, is positively related to risk as measured by the conditional volatility of the Nikkei stock index returns.

If the degree of uncertainty in asset returns varies over time, the compensation (or risk premium) required by risk averse investors for holding risky assets must be varying. This is called time varying risk premia (TVRP). To derive the TVRP, a simple index arbitrage in the stock index futures market is employed for the equilibrium value of the intertemporal asset pricing model which was modified to allow for risk aversion. The time varying risk premia are applied to the hedging model. Generalized Autoregressive Conditional Heteroscedasticity in mean (GARCH-M) models of excess returns on the Nikkei 225 stock index are used to estimate the time varying conditional variances and a mean-variance ratio that represents the risk-return trade off. In the ARCH model, the

nature of heteroskedastic error terms over time is captured by the conditional variance, which depend upon past available information. Thus, the conditional variance is an useful variable for measuring uncertainty and provides a model of the TVRP.

The empirical results show that there was strong evidence of both ARCH and GARCH effects and the time varying risk premia, as reflected in heteroscedastic error terms over time, using data on spot and futures prices of the Nikkei 225. This suggests that an investor's risk premium changes over time. Another finding is the persistence of shocks to the Nikkei 225 stock index returns volatility. My results also show a negative relationship between excess returns on the Nikkei 225 stock index returns and conditional volatility using the GARCH-M model. The findings of this paper support that an increase in riskiness can either increase or decrease the stock index returns.

The specification tests demonstrate that the expected returns on the Nikkei 225 stock index can be improved by adding its own lagged endogenous variable and the first order of moving average which contain important information.

ACKNOWLEDGEMENTS

I would like to thank many faculty members and students while attending the graduate school at the City University of New York. They gave me valuable advice and encouragement. First and foremost, I wish to thank my dissertation committee members, Dr. Thom Thurston, Dr. T.D. Agbeyegbe, and Dr. Salih N. Neftci for their valuable comments and discussions. I would especially like to thank Dr. Thom Thurston, whose encouragement, guidance, academic support, and elaborate advice not only helped me understand my subject better, but helped me to complete this research throughout the entire period of my research. He always cheered me up when I had hard times. I also thank Dr. T.D. Agbeyegbe, and Dr. Salih N. Neftci for their valuable suggestions and ideas to develop my dissertation. Second, I am deeply grateful to Dr. Michael Grossman for giving me the opportunity to teach at York College, City University of New York, for my teaching career. I hope to show my gratitude by always doing my best. Third, my thanks go to Tong Yang Futures Trading Co., Ltd. and Samsung Insurance Economics Research Institute which provide the data for my research. Fourth, I would like to thank my friend, Dini Chandra, who took on the task of proof reading my dissertation.

Fifth, my special thanks go to my parents, Sookwan Lee and Seungim Lee, for their endless love and supports. This dissertation I dedicate to my father, who passed away in the middle of this research and encouraged me to accomplish my dream. I am also so indebted to my mother because she has always prayed for everything about me. I would like to acknowledge the help of my father-in-law, Bongsoo Shin, my mother-in-law, Woohee Jeon, my elder sister, Gayoung, and my elder brother, Sungman. Sixth, I would like to thank my brother-in-law, Dr. Changi Nam, who inspired the curiosity in me that led me to pursue my doctorate and advised me to study finance. Finally, I thank my wife, Sunmee, for her true love, patience, and devotional support from the bottom of my heart. Without her, this dissertation would not be possible. One more thing I want to say is to share this big joy with my expecting daughter, Joan, who will give us the second new life to cherish. Now, I can feel my Buddha smiles at me.

TABLE OF CONTENTS

LIST OF TABLES

Chapter I

INTRODUCTION

1.1 Overview

The introduction of stock index futures in April 1982 has reduced the cost of transactions, resulting in increased inter-market trading activity. Stock index futures contracts are attractive to investors due to low transactions costs and high liquidity in comparison to trading in the underlying market. Moreover, stock index futures has played an important role in enabling stock market participants to reduce the risks which they face due to stock price fluctuations. However, the investors and participants still bear the risks from the unknown movements of stock prices and stock index futures prices. Thus, the valuation of risk is very important in these markets. In general, the risk premium can be defined as a higher expected rate of return of an asset as it becomes more risky. If the degree of uncertainty in asset returns varies over time, the compensation (or risk premium) required by risk averse investors for holding risky assets must be varying. This is called time varying risk premia (TVRP). TVRP, as reflected in an error term which is heteroskedastic through time, is an important measure of the degree of riskiness in the financial markets. The changing conditional variances over time can be one of the factors which contribute to the volatility of financial markets.

The importance of such risk premia in foreign exchange markets, the term structure of interest rates, and stock returns have been studied extensively. To date, not much work has been done on the existence of TVRP in stock index futures. This dissertation concerns with examining the intertemporal relation between excess returns on the Nikkei 225 stock index and conditional volatility with the inclusion of futures prices. Most asset pricing models postulate a positive relation between a stock portfolio's expected returns and risk. In general, the risk premium is expected to go up when the market's volatility rises since risk averse investors are required to pay a larger amount of compensation to get into the increasingly volatile markets. For example, French, Schwert, and Stambaugh (1987), and Chou (1988) found statistically significant positive relation between conditional variance and excess returns on stock market indices using GARCH-M models. However, some studies question the existence of a positive mean-variance ratio. Naka (1989) showed that the returns on the spot currency exchange markets using five major currencies can be positive or negative when the TVRP increases. Abel (1988) claims that in general equilibrium the mean-variance relationship is not necessarily positive when the investor's preference has not logarithmic (which have a unitary elasticity of marginal utility). Glosten, Jagannathan, and Runkle (1989) also describes the asymmetry in the conditional volatility of Japanese stock returns and obtains a negative ARCH-M parameter when they include the nominal risk-free rate in the conditioning information set.

In this dissertation, to derive TVRP based on the rational expectation hypothesis, a simple index arbitrage is employed for the equilibrium value of intertemporal asset pricing model which was developed by Lucas (1978). The time varying risk premia are applied to the hedging model. Generalized Autoregressive Conditional Heteroskedasticity in mean (GARCH-M) models of excess returns on the Nikkei 225 stock index are utilized to estimate the time varying conditional variances and a mean-variance ratio that represents the risk-return trade off.

This dissertation's objectives are : (a) to find whether or not an investor's risk attitude (or risk premium) changes over time, (b) to investigate the relationship between excess returns on the Nikkei stock index and conditional volatility and the existence of a negative GARCH-M parameter or mean-variance ratio.

The paper consists of five chapters. The first consists of the introduction and literature review. The second chapter explains theories based on the rational expectations hypothesis (which is most widely used models in pricing futures) and risk premia. The third chapter shows how to apply the TVRP to the hedging model. A simple index arbitrage theory is employed for the equilibrium value of intertemporal asset pricing model to derive the TVRP. The GARCH-M model of excess returns on the Nikkei 225 stock index is utilized to estimate the time varying conditional variances and a mean-variance ratio. The fourth chapter explains the econometric ARCH, GARCH, and ARCH-M models, which explain heteroscedasticity (or conditional volatility) and the time varying risk premia in the stock index futures market, to empirically test the hypotheses. A procedure by which to estimate ARCH models is presented.

The fifth chapter describes data and is the implications the empirical results. The Nikkei 225 stock index and futures contract are used to estimate the time varying conditional variances and a mean-variance ratio. Using data on spot and futures prices of the Nikkei 225, the empirical results found significant heteroscedasticity and non-normality of the error term of the OLS regression. Moreover, there was strong evidence of both ARCH and GARCH effects and the time varying risk premia, as reflected in heteroskedastic error terms over time in the hedging model. This suggests that an investor's risk premium changes over time. Another finding is the persistence of shocks to the Nikkei 225 stock index returns volatility. My results also show negative relationship between excess returns on the Nikkei stock index and conditional volatility. The findings of this paper support that an increase in riskiness can either increase or decrease the stock index returns. The specification tests demonstrate that the expected returns on the Nikkei 225 stock index can be improved by adding the lagged dependent variables and the first order of moving average [or MA(1)] which contain important information. The GARCH (1,1) process is rejected in favor of the GARCH(1,2) in the hedging model.

The conclusion in Chapter 6 presents a summary of this study and comments on the potential for future research in this area.

1.2 Review of the Literature

Mandelbrot (1963) observed that large price changes tend to be followed by large changes - of either sign - and small changes tend to be followed by small changes and that in the speculative markets the prices follow non-normal stable distributions (leptokurtic). The ARCH model of Engle (1982) explains well Mandelbrot's observation by providing possible parameterization of heteroscedasticity. The ARCH model estimates the unobservable second moments by parameterizing the conditional variance to capture much of the leptokurtosis. The ARCH process recognizes the difference between the unconditional and the conditional variance allowing the latter to change over time as a linear function of past squared errors leaving the unconditional variance constant. Bollerslev (1986) proposed the generalized ARCH (or GARCH), which provides a much more flexible lag structure. GARCH allows for past conditional variances in the current conditional variance equation. Engle, Lilien, and Robins (1987) introduced the ARCH in mean model, which allows the conditional variance to affect the mean. In other words, changing conditional variances directly affect the expected return on a portfolio. The ARCH, GARCH, and ARCH-M model and the estimation and test of GARCH (and ARCH) are explained in Chapter IV in more detail.

Lucas (1978) asset pricing model is examined why stock price is volatile. The first order condition arising from the intertemporal utility maximization model shows that changes in expectations regarding futures dividends, in the discount factor, or the marginal rate of substitution of consumption will cause volatility in stock prices. Since expectations are conditional upon available information, new information can change expectations about these variables. If investors' utility functions display more aversion, then an increase in the variance of these variables is likely to increase the volatility in stock prices. This intertemporal asset pricing model based on the rational expectations hypothesis allows for risk aversion. Cox, Ingersoll, and Ross (1981) show the difference between futures prices and forward prices under stochastic interest rates. A major difference between a futures contract and a forward contract results from variation margin, the daily "marking to market" of an open future position. If an investor is long a futures contracts and its settlement price on a given day is below that of the previous day, he is required to pay the price difference in cash. If the price rises, he will receive the price difference in cash. However, with forward contracts, no cash changes hands until expiration of the contracts. Richard and Sundarsen (1981) show the difference between futures prices and forward prices under a continuous time equilibrium model.

For the hedging model, Figlewski (1986) used the ordinary least squares regression to estimate the optimal hedge ratio relating changes in cash prices to changes in futures prices. They run a regression of the returns on the spot market and futures contract so that the optimal hedge ratio (the sample covariance between both returns divided by the sample variance of the futures returns) is obtained [see Figlewski (1984)]. Bera, Bubnys, and Park (1993) found that significant heteroskedasticity (conditional and unconditional) and nonnormality of the disturbance term of the OLS regression, using data on spot and futures prices of S&P 500, NYSE, and Kansas City Value Line indices from the first trading date of each contract to June 1985. They also found that there is the improved efficiency of the hedge ratio through the ARCH model.

Work has been done regarding the existence for time varying risk premia (TVRP) in the foreign currency markets. Domowitz and Hakkio (1985) define the risk premium as the difference between the expectation of the logarithm of the expected spot price and the logarithm of the forward price. The existence of a risk premium for five foreign exchange rates based on the conditional variance of market forecast errors by using the ARCH-M model was examined. Their results show that there is some evidence of a nonzero constant risk premium for Japanese Yen and UK pound and that the estimated risk premium switches from positive to negative for different periods. Hodrick and Sryivastava(1987) [thereafter H&S] are the first to derive the existence of TVRP in futures markets for foreign exchange by using Generalized Method of Moments. They also define the TVRP as the difference between today's futures prices and expected spot prices (or expected futures prices). In other words, the TVRP is expressed as the conditional covariance of the marginal rate of substitution of consumption between time t and time t+1 with known return of futures contracts and expected futures prices at time t+1. Thus, the current futures price is equal to expected futures price plus a risk premium. The null hypothesis is unbiasedness of daily futures prices as predictors of the following day's futures price. The null hypothesis is rejected for all five major currencies. Hence, the daily futures premium is useful in predicting each of the following day's futures premium. However, to allow the variability of the risk premium to become larger, the highly positive autocorrelation in daily risk premium is required. They do not estimate TVRP directly and ignore the problems of convergence (the futures prices will be equal to the spot prices at the maturity) and the presence of thin markets at the maturity of the contracts (there will be very little active trading at the maturity date), because the GMM method does not explicitly specify the risk term. Naka (1989) investigate the existence of the TVRP in the foreign currency futures markets (five major foreign currencies). He defines the TVRP as the covariance between marginal rates of substitution [the definition of H&S]. He also introduced the TVRP into a hedging model and a basis risk model which is based on the intertemporal asset pricing between futures and spot currency exchange markets by Richard and Sundaresan (1981). The conditional variance and covariance matrices of the returns on the spot currency exchange markets are utilized to test for the presence of the TVRP. ARCH-M (GARCH-M) model and one-factor ARCH in Mean model (for the multivariate cases) are employed to test for the presence of the TVRP. The results show that there is strong evidence for varying risk premia, as reflected in heteroscedastic error terms through time in the foreign exchange markets and that the returns on the spot currency exchange markets using five major currencies can be positive or negative when the TVRP increases.

Models for the term structure of interest using an estimate of the conditional variance as a proxy for the risk premium are given in Engle, Lilien, and Robins (1987). They extended the simple ARCH of measuring conditional variances to the ARCH-M model where the conditional variance is a determinant of the current risk premium, and thus enters the forecasting equation of expected financial returns. Their results show that the ARCH effect was clearly present in the forecast errors of bond holding yields from 1960 through 1984 II on 3 and 6 month Treasury bills indicating variation in the degree of uncertainty over time. They conclude that the time varying risk premium and yield spread are a significant and positive determinant of the excess holding yield, indicating that risk premia vary systematically with agent's perceptions of underlying uncertainty . Under an assumption of conditional t-distribution, Bollerslev (1987) finds the conditional standard deviations [GARCH (1,1)-t model] help to explain variations in the expected return of the S&P 500 index and foreign exchange rates. The standard t-distribution and ARCH or GARCH model fails to capture the leptokurtosis. Instead, the GARCH (1,1)-t model fits the return series of the S&P 500 index and foreign exchange rates. Bollerslev, Engle, and Wooldridge (1988) examine whether the expected return is proportion to the conditional covariance (which measure the non-diversified risk) of each return with the market portfolio return. They extended GARCH(1,1) to a multivariate GARCH process to test a conditional CAPM with time varying covariances of the returns of bills (6-month Treasury bills), bonds (20-year Treasury bonds) and stocks. They found that the conditional covariances vary over time and are significant determinant of the TVRP.

Their results also show that bills and bonds have rising (or positive) risk premia during the volatile post-October 1979 period. However, the negative premia were obtained for bonds and stocks in some periods because the preferential tax treatments provide incentives to hold these assets even at otherwise unfavorable rates of return. The estimated beta for stocks and bills is close to zero, while the beta for bonds is slightly above one. There is, however, substantial movement over the sample period. Their diagnostic tests indicate that lagged excess holding yields and innovations in consumption appear to have some explanatory power for the asset returns.

Most asset pricing models postulate a positive relationship between a stock portfolio's expected returns and risk. French, Schwert and Stambaugh (1987) and Chou (1988) found statistically significant positive relation between conditional variance and excess returns on stock market indices using GARCH-M models. French, Schwert, and Stambaugh (1987) studied the relation between risk and stock market behavior, but focusing instead on stock prices. To estimate the ex ante relation between risk premiums and volatility, the GARCH-M model with a conditional normality is used. The expected risk premium (or the daily excess holding period return on the S&P composite portfolio over the 1928-1984 period) is defined as the expected return on a stock market portfolio minus the risk-free interest rate and the volatility is the variance of the unexpected excess holding period return. They found reliable evidence of a positive relation between the expected risk premium and the predicted volatility. Chou (1988) found the existence of changing risk premia in returns on the NYSE value-weighted index with dividends. The persistence of shocks to the stock return volatility is so high that a non-stationary volatility process specification can not be rejected. He demonstrates that the GARCH-M model is a more suitable tool than the two-stage method and gives much more reliable results. He also found a statistically significant positive relationship between conditional variance and returns on stock market indices using GARCH-M models. Baillie and DeGennaro (1990) use the GARCH-M models to examine the relationship between mean returns on a stock portfolio and its conditional variance or standard deviation. The GARCH-M model with a conditional student t density is found to provide a better description of daily and monthly returns on the CRSP value weighted index with dividends than the GARCH-M model with a conditional normality. Their results show that the relationship between mean returns and own variance or standard deviation is weak. Their results imply that traditional two-parameter models relating portfolio returns to risks are inappropriate and that investors consider some other risk measure to be more important than the variance of portfolio returns.

A number of studies question the existence of a positive mean-variance ratio. Abel (1988) examine the relation between this risk premium on stocks and dividend volatility. In other words, an increase in the variance of dividends is likely to increase the risk premium on stocks. Under logarithmic utility, which has a unitary elasticity of marginal utility, the volatility of the rate of return on stocks is equal to the volatility of the dividend. However, an increase in the conditional risk reduces the stock price if the elasticity of marginal utility is less than one. Abel (1988) claims that the mean-variance relationship is not necessarily positive when the investor's preference is not logarithmic and that it depends on the curvature of the utility function. He also derived exact solutions for asset prices [the prices of the aggregate stock and riskless one-period bonds] in a rational expectations model [the Lucas (1978) model] with a time-varying distribution of dividends and provided some theoretical support for aspects of previous studies by French, Schwert, and Stambaugh (1987). Glosten, Jagannathan, and Runkle (1989) also describes the asymmetry in the conditional volatility of Japanese stock returns and obtains a negative ARCH-M parameter when they include the nominal risk-free rate in the conditioning information set. My paper focuses on the relation between conditional volatility and the returns on the Nikkei stock index, instead of focusing on stock prices and dividend.

Chapter II

EXPECTATIONS and RISK PREMIA

2.1 *Futures Prices and Rational Expectations*

To incorporate the TVRP into the hedging model which is based on the rational expectations hypothesis, I shall explain the relationship between rational expectations and risk premia. The most widely used models in pricing futures are the cost of carry model and the expectations hypothesis. The former implies that the futures price depends on the current cash price for the underlying security and the net cost of carrying it until the futures delivery date. The market's expectation also plays a large role in establishing futures prices. The expectations model implies that the current price of a futures contract is equal to the market expectations of the spot price on the delivery date. The rational expectations hypothesis is not inconsistent with the cost of carry model [see section 2.3].

If the expectations theory did not hold, there would be attractive speculative opportunities. For instance, assume that market participants expect the price of the fruit in the nest harvest to be \$10. If the futures price were \$12, speculators would sell the futures contract and then plan to buy the fruit for \$10 on the harvest date. They would be able to deliver the fruit and collect \$12, for a \$2 profit. By contrast, if the futures price were \$8, speculators would buy the futures contract and then plan to sell the fruit for \$10 on the harvest date. They also take a profit of \$2. Therefore, speculative strategies are available when the futures price does not equal the expected spot price. In short, profit-seeking speculators will trade as long as the futures price is sufficiently far away from the expected spot price.

The futures price can diverge from the expected spot price due to transaction costs or due to risk aversion on the part of traders. First, transaction costs can keep the futures price from being exactly the expected spot price. Assume that the fruit has a futures price of $9 and an expected spot price of $10, and assume that the cost of transaction to take advantage of this discrepancy is $1.50. With these prices, a trader can not buy the futures for $9 and plan to make a profit by selling the fruit at its expected spot price. This opportunity is not profitable with the transaction costs, because the total cost of acquiring the fruit would be the $9 delivering on the futures plus the $1.50 transaction cost. Second, the futures price can differ from the expected spot price if traders are risk averse. Traders in futures markets can be classified into hedgers and speculators. Hedgers are traders who attempt to reduce a pre-existing risk (or unwanted risk), while speculators are traders who require compensation for the risks they are bearing. The speculators accept the risk that the hedgers are unwilling to bear. From the definitions of hedgers and speculators, the hedgers must be more risk averse than the speculators. Most participants in financial markets are risk averse, so they take risk only if the expected profit from bearing the risk will compensate them for the risk exposure. Assume that the fruit has a futures price of $10.05 and an expected spot price of $10.00. Assume also that there is tremendous uncertainty about what the actual price of the fruit will be. The market expects a cash price of $10 upon harvest, but the fruit is very susceptible to weather conditions. For a speculator, there appears to be a $.05 profit available from the strategy of selling the futures, buying fruit for $10.00 at harvest, and delivering against the futures contract.

However, this strategy subjects the speculator to considerable risk if the weather is bad. Speculators may decide that the expected profit of $.05 is not worth the risk exposure. If they are more risk averse, they ask more compensation from the risk exposure. Thus, the expected profit will be higher. So, the difference between the futures price and the expected spot price will be bigger. Thus, risk aversion among speculators can allow the futures price to diverge from the expected spot price. According to the theory of normal backwardation, this divergence occurs in a systematic way. Normal backwardation is discussed in the next section.

2.2 Futures Prices and Risk Aversion

We explore two theories of the relationship between futures price and the expected spot price. Whether the futures price is equal to the expected spot price is still an open question. Assume that speculators are rational; that is, they make assessments of expected future prices based on available information. In assessing this information, rational speculators occasionally make mistakes, but on the whole, they process the information efficiently. The expectational errors are randomly distributed around the true price that the commodity will have in the future. Assuming that hedgers are net short in most futures market to reduce the risk they face in their businesses and the number of short positions should be same as that of long positions, Keynes (1930) explained the normal backwardation. The normal backwardation can be explained well in the market which has a stable demand and supply of commodity. For example, a wheat farmer has a long position in cash wheat because he or she grows wheat. The farmer(or hedger) can reduce risk by selling wheat futures. If hedgers will be net short, speculators must be net long. Consider a single speculator who is considering to take a long position. In this case, the rational speculator takes a long futures position only if the expected spot price exceeds the current futures price. Otherwise, the speculator do not make any profit.

The hedger must be willing to sell the futures contracts at a price below the expected spot price. Otherwise, he cannot induce the speculator to accept the long side of the contract.

In this case, the expected futures spot price should be higher than the current futures price (the futures price will rise over time) due to the hedger's general desire to be net short. This view is known as normal backwardation. In effect, the hedger transfers his unwanted risk to the speculator and pays expected profit to the speculator for the bearing the risk. The payment to the speculator is the difference between the futures price and the expected spot price. It is called risk premia. Even so, the speculator does not receive any sure payment. The speculator must still wait for the expected spot price to receive the expected profit. Conversely, when the expected spot price should be lower than the current futures price (the futures price will decline over time), the traders will have the opposite positions, short positions for speculator and long positions for hedger. This view is known as contango. If the speculator is net short, then he can hope to earn a return for his risk-bearing services only if the futures price lies above the expected spot price. The fall in futures prices, as the contract approaches maturity, gives the short speculator the compensation that induces him to enter the market.

Thus far, the discussion has focused on a single hedger and a single speculator. Let's extend it to many hedgers and many speculators as two groups with different needs, different levels of risk aversion, and different expectations about expected spot price. By assuming that hedgers will be net short, no matter what the futures price, their pre-existing risk requires a short position. Thus, hedgers are defined as those who enter the futures market to reduce a pre-existing risk. Figure 1 depicts the situation that might prevail in the futures market. Line HH shows the hedgers' supply curve for various futures prices. At higher futures prices, hedgers want to sell more futures contracts, as the downward slope for line HH indicates. For convenience, lines HH and SS are drawn as straight lines. Note that the hedgers hedge different amounts depending on the futures price. With low prices, they sell fewer contracts, hedging less of their pre-existing risk than they would if futures price were high. Speculators are willing to hold either long or short net positions as the situation demands. Line SS shows the speculators' demand curve as a function of the futures price. If the futures price exceeds the expected spot price (the futures price lies above point C), the speculators will desire to be net short as well as the hedgers. If the futures price lies below the expected spot price (normal backwardation case), speculators will desire to be net long, holding some position between C and S. If the futures price equals the expected spot price, the speculator will be neither long nor short. In such a situation, some speculators would be long, others short, reflecting their divergent opinions. But, in the aggregate, they hold a zero net position in the futures market. With a price of E, the net short position desired by the hedgers exactly offsets the net long position desired by the speculators(market clearing condition: AE = BE). The futures market must reach an equilibrium price at E when the futures price lies below the expected spot price.

2.3 Risk Premia

Market efficiency in this research means that the prices of the futures market contain all the information for determining the underlying spot prices when the contracts mature. The relationship between futures price and the expected futures spot price is assumed:

(2-1) $F(t,T) = E[S(t+T)| \Psi(t)] - RP$

where $F(t,T)$ = futures price at current time t for a futures contract that expires at time T, $F(t+n,T)$ = futures price at time t+n with the maturity date at t+T $[T \geq n]$. $S(t+T)$ = spot price at time t+T, $\Psi(t)$ = all the set of available information for market participants at time t, and RP is risk premium. For simplicity, the notation F(t) is used for F(t,T) and F(t+n) is used for F(t+n,T). Samuelson (1965) utilizes the property of conditional expectations to explain the theory of market efficiency. He hypothesizes that today's expectation of tomorrow's forecast equals today's forecast, that is,

$$E\{E[S(T)| \Psi(t+1)] | \Psi(t)\} = E[S(T) | \Psi(t)]$$

where $S(T)$ = spot price at time T, today's forecast = $E[S(T) | \Psi(t)]$, for $T > t$, and tomorrow's forecast = $E[S(T) | \Psi(t+1)]$, for $T > t+1$. Samuelson also assumes that futures prices are conditional expectations of future spot prices. $F(t) = E[S(T) | \Psi(t)]$, $F(t+1) = E[S(T)| \Psi(t+1)]$, and so on.

The derivation of (2-1) indicates that the rational expectations is consistent with the cost of carry model and is as follows: Suppose there is speculative demand for the commodity and this commodity can be storable and seasonal. Carrying costs (storage costs) happen to speculators to reserve the commodity. The profit of speculators in no futures market will be

(2-2) $R(t,T) = S(T) - S(t) - C(t,T)$

where $R(t,T)$ = Profit from carrying the good from the present to the selling date, $S(T)$ = the spot price at selling date, $S(t)$ = the spot price at buying date, $C(t,T)$ = the carrying cost per unit to carry the good from the present to the selling date. $R(t,T)$ is a random variable because $S(T)$ is not known at time t. Thus, the speculators will consider the expected profit from this transaction. Assuming that the speculators know the cost of carry, the expected profit is

(2-3) $E[R(t,T)] = E[S(T) \mid \Psi(t)] - S(t) - C(t,T)$.

If the speculator is risk-neutral, the increasing speculative demand for a commodity will continue until all arbitrage profits disappear. Thus, the increasing demand for the spot or the decreasing demand for $E[S(T) \mid \Psi(t)]$ will drive the expected profit to zero. Thus, the expected spot price at time T will be

(2-4) $E[S(T) \mid \Psi(t)] = S(t) + C(t,T)$.

Now, when we apply (2-4) to the futures markets, we can see that both rational expectations and cost of carry model are not inconsistent. According to the rational expectations, the futures price, $F(t)$, is equal to $E[S(T) \mid \Psi(t)]$. Substituting $F(t)$ into (2-4) gives $F(t) = S(t) + C(t,T)$ [the cost of carry model(see ch.3.2 for more details)]. Hence, $F(t) = E[S(T) \mid \Psi(t)] = S(t) + C(t,T)$. Thus, we can say that both the rational expectations and the cost of carry model are consistent. Now, suppose that the speculators are risk averse. They will ask for compensations for the price risk they are bearing. In this case, the expected spot price at time T will be

(2-5) $E[S(T) \mid \Psi(t)] = S(t) + C(t,T) + RP$.

When we apply (2-5) to the futures market again, we get a necessary condition for the decision of the equilibrium price for the futures contract. Substitute $E[S(T) \mid \Psi(t)]$ - RP from (2-5) into the cost of carry model. Finally, we have

$F(t) = E[S(T) \mid \Psi(t)] - RP$. QED

The above futures price is the equilibrium price for the futures contract in perfect markets. This equation also means that based on the market efficiency, the futures price converges to the expected future spot price, i.e. the futures price is unbiased predictor of the expected futures spot price $\{F(t) = E[S(T) \mid \Psi(t)]\}$, if there is no risk premium. However, if the normal backwardation hypothesis hold, then $F(t) = E[S(T) \mid \Psi(t)] - RP$ or $F(t) = S(t) - RP$. [since $E[S(T) \mid \Psi(t)] = S(t)$ from the martingale; the expected value of tomorrow's price equals today's price, if the expected price change (or the expected inflation) is zero]. Therefore, $E[S(T) \mid \Psi(t)] > F(t)$ or $S(t) > F(t)$. In this case, risk premium is positive [since $RP = E[S(T) \mid \Psi(t)]-F(t) > 0$]. The positive risk premium is thought of as the expected payment to a speculator buying at time t one contract from a hedger who is insuring a commitment to sell later on the spot to reduce his unwanted risk. The risk premium plays the role of an insurance premium. Thus, the speculators will buy futures contracts and hedgers will sell short. Therefore, the positive risk premium also implies that the futures price rises over time, i.e., normal backwardation.

On the other hand, if the commodity is traded in the unstable (or seasonal) market, the transactions of futures contracts will be done through storage of commodity. Thus, we have to consider carrying costs. The futures price will be the sum of spot price and carrying costs, that is, $F(t) = S(t) + C(t,T)$. This suggests that the futures price is greater than the spot price, that is, $F(t) > S(t)$ or $F(t) > E[S(T)|\ \Psi(t)]$. [from the martingale, $E[S(T)|\ \Psi(t)] = S(t)$]. Therefore, risk premium is negative [since $RP = E[S(T)|\ \Psi(t)] - F(t) < 0$]. The negative risk premia also implies that the current futures price is downward over time, i.e., contango. The speculators will take a short position and hedgers will hold long positions. According to the definition of risk premium, the above two risk premia can be summarized as

$RP = E[S(T)|\ \Psi(t)] - F(t)$: normal backwardation

$-RP = E[S(T)|\ \Psi(t)] - F(t)$: contango

In the financial market, if both variables { $F(t)$ and $E[S(T)|\ \Psi(t)]$ } do not change by the same degree, the risk premium will change over time. Therefore, the above equations do not hold if $F(t)$ and $E[S(T)|\ \Psi(t)]$ vary over time. To overcome this problem we must consider that risk premium varies over time as futures and spot prices vary over time.

Chapter III

TIME VARYING RISK PREMIA AND THE HEDGING MODEL

This chapter shows how to introduce the time varying risk premia into a hedging model. The procedures are as follows: (1) apply the first order condition of the intertemporal utility maximization model; (2) utilize index arbitrage; (3) derive the covariance of the expected return of an asset and an intertemporal marginal rate of substitution of consumption; (4) introduce the time varying risk premia into the hedging model; and (5) apply the GARCH-M model.

3.1 *Intertemporal Utility Maximization Model*

To explain why price volatility occurs, intertemporal asset pricing model which was developed by Lucas(1978) is employed. Let us briefly discuss Lucas's (1978) main theory.* which examined the stochastic behavior of equilibrium asset prices in a pure exchange economy. Consider an economy with a representative consumer for a large number of identical consumers. The consumer maximizes his expected utility

(3-1) $$E\{\sum_{t=0}^{\infty} \beta^t U[c(t)]\} \qquad 0 < \beta < 1,$$

* For this part, see Thomas J. Sargent "Dynamic Macroeconomic Theory" (1987) PP.92-115.

where c(t) is a stochastic process representing consumption of a single good, β is a discount factor, $U(\cdot)$ is a current period utility function and $E\{\cdot\}$ is an expectations operator. The utility function $U(\cdot)$ is assumed to be strictly concave, increasing and twice continuously differentiable.

In this economy, the only durable good is a set of "trees"(or stocks), which are equal to the number of people in the economy. Each period, each tree yields fruit or dividends in the amount D(t) to its owner at the beginning of period t. The fruit is nonstorable, but the tree is perfectly durable. Let p(t) be the price of a tree (or stock price) in period t, measured in units of consumption goods per tree. We also define s(t) as the number of trees (or stocks) owned at the beginning of period t. Each investor starts life at time t with one tree [p(t)*s(t)] and with its initial dividend of the consumption good [D(t)*s(t)]. Therefore, the initial wealth is [p(t)+D(t)]s(t). The consumer's consumption at time t is constrained by

$$A(t+1) \leq R_{t,t+1}\ [A(t)-c(t)]$$

where c(t) is consumption of an agent at time t, A(t) is the amount of a single earning asset valued in units of the consumption good, held at the beginning of period t [A(t) = {p(t) + D(t)}s(t)], $R_{t,t+1}$ is real gross rate of return on the asset between t and t+1 (gross yield on stock) [$R_{t,t+1}$ = p(t+1) + D(t+1) / p(t)]. We shall imagine that $R_{t,t+1}$ is a random process and that $R_{t,t+1}$ becomes known to the investor only at time t+1. Substituting these two terms into the above equation gives

$$[p(t+1) + D(t+1)]s(t+1) \leq [\ p(t+1) + D(t+1) / p(t)\ ]\ [\{p(t) + D(t)\}s(t) - c(t)].$$

Rearranging this equation gives

(3-2) $c(t) + p(t)s(t+1) \leq [p(t) + D(t)]s(t)$.

Each consumer maximizes his expected lifetime utility subject to (3-2). Let the state of the economy be x. The state of the economy in Lucas's model includes the current dividend D(t) and all information available at time t that helps predict future dividends. The dividend D(t) is assumed to be governed by a Markov process with a time-invariant transition probability distribution function given by $prob[D(t+1) \leq x(t+1) | D(t)=x(t)] = F[x(t+1),x(t)]$,where x(t+1) = future state(good or bad state of dividends), x(t)= today's state(good or bad state of dividends). The stock price p(t) is

(3-3) $p(t) = h[x(t)]$

where h is a continuous, bounded function defined on the domain of the current state x(t), where in Lucas's model the state x(t) equals current dividends D(t). Stock price p(t) depends upon unknown x(t). Today's state can be good or bad state of dividends. Therefore, today's stock price depends upon new information of the current dividend D(t). Similarly, tomorrow's stock price also depends on tomorrow's dividend D(t+1).

The maximization of the optimal control can be solved by the Bellman equation for dynamic programming. The control variables are the rate of consumption [c(t)] and the quantity of investment at time t+1 [s(t+1)]. The state variables are $\{[p(t)+D(t)]s(t)\}$ which are this period's wealth of the individual agent. Let $V\{s(t)[p(t)+D(t)]\}$ be the optimal value function for (3-2). Sargent(1987) derived the real ex-dividend stock price from the common intertemporal utility maximization model. Bellman's equation is

(3-4) $V_t = \max_{s(t+1)} \{U[c(t)] + \beta E_t V_{t+1}\}$.

Substituting the value functions and consumption function into $V_t, V_{t+1}, c(t)$ respectively gives

(3-5)

$$V\{s(t)[D(t)+p(t)]\} = \max_{s(t+1)} \{U[p(t)+D(t)]s(t) - p(t)s(t+1)\}$$
$$+\beta E_t V\{s(t+1)[D(t+1)+p(t+1)]\}$$

where $V_t = V\{s(t)[D(t)+p(t)]\}$

$V_{t+1} = V\{s(t+1)[D(t+1)+p(t+1)]\}$

c(t)=$[p(t)+D(t)]s(t) - p(t)s(t+1)$.

(3-5) may be rewritten as

(3-6) $$V[s(p+x)] = \max_{s'} \{U[(p+x)s - ps'] + \beta E_t V[s'(p'+x')]$$

where primes denote next-period values, the absence of primes denotes this-period values, the variable s', p', x' are next period's holding of stocks, next period's stock price, and dividends, and the current dividend equals x. Now, substituting p = h(x) into (3-6) gives

(3-7) $$V[s(h(x)+x)] = \max_{s'} \{U[(h(x)+x)s - h(x)s'] + \beta \int_{-\infty}^{\infty} V[(h(x')+x')s']dF(x',x')\}.$$

The first order necessary condition arising from the intertemporal utility maximization model is

(3-8) $$U'[(h(x)+x)s - h(x)s'][-h(x)] + \beta \int_{-\infty}^{\infty} V'[(h(x')+x')s'][h(x')+x']dF(x',x) = 0$$

Let us write (3-8) as

$$(3\text{-}9)\quad U'[c(x)][h(x)] = \beta \int_{-\infty}^{\infty} U'[c(x')[h(x') + x']dF(x',x)$$

where c(x) = $[h(x)+x]s - h(x)s'$ (from the constraint condition)

$c(x') = [h(x')+x']s' - h(x')s''$

$V'[(h(x')+x')s'] = U'[(h(x')+x')s' - h(x')s''] = U'[c(x')]$

$V'[(h(x)+x)s] = U'[(h(x)+x)s - h(x)s'] = U'[c(x)]$ (from Benveniste and Scheinkman's formula).

(3-9) may be written as

$$(3\text{-}10)\quad h(x) = \beta \int_{-\infty}^{\infty} \frac{U'[c(x')]}{U'[c(x)]}[h(x') + x']dF(x',x).$$

In the competitive equilibrium consumption allocation of this economy, there can be no gains from trade because preferences and endowment patterns are the same across all individuals and there is only one source of goods, this period's dividends. So, $c_t = d_t$. In this equilibrium, $s = s' = 1$ indicates that there is one tree per person, and c(x)= $[h(x)+x]s - h(x)s'$ = x. Therefore, in equilibrium, c = x = d ($c_t = x_t = d_t$). Substituting these into (3-10) gives

$$(3\text{-}11)\quad h(x) = \beta \int_{-\infty}^{\infty} \frac{U'[x']}{U'[x]}[h(x') + x']dF(x',x)$$

$$= \beta \int_{-\infty}^{\infty} \frac{U'[c']}{U'[c]}[h(x') + x']dF(x',x)$$

Another way to represent (3-11) that is more convenient for purposes of empirical testing is

$$(3\text{-}12)\quad P(t) = \beta E_t \frac{U'[c(t+1)]}{U'[c(t)]}[P(t+1) + D(t+1)]$$

where $\beta \frac{U'[c(t+1)]}{U'[c(t)]}$ is the marginal rate of substitution of consumption between period t+1 and period t

E_t denotes the mathematical expectation conditional on information available at time t.

The above equation shows that changes in expectations regarding future dividends, in the discount factor, and the marginal rate of substitution of consumption will make stock prices volatile. Unexpected changes in the riskless interest rate will also affect the discount factor and stock prices. Since expectations are conditional upon available information, new information can change expectations about these variables and affect stock prices. Therefore, asset price volatility occurs because of new information shocks. (3-12) can be written as

$$(3\text{-}13)\quad P(t) = E_t\{MRS_{t,T}[P(t+1) + D(t+1)]\}$$

where $MRS_{t,T} = \beta \frac{U'(c_{t+1})}{U'(c_t)}$ (marginal rate of substitution between known period t and expected period t+1 consumption). This is the first order condition arising from the intertemporal utility maximization model.

Rearranging (3-13), we have

$$(3\text{-}14)\quad 1 = E_t[MRS_{t,t+1} \frac{P(t+1)+D(t+1)}{P(t)}]$$

or

$$(3\text{-}15)\quad 1 = E_t[MRS_{t,t+1} R_{t,t+1}]$$

where $\frac{P(t+1)+D(t+1)}{P(t)} = R_{t,t+1}$ (real gross rate of return on the asset between time t and t+1). The equation (3-15) is the present value of the cash flow which is rolled over by one unit of an asset. If there is only one period remaining before the maturity date, then $R_{t,t+1} = \prod_{k=t}^{T-1} R_k$ (that is, the known dollar return from investing a dollar at time t in a t+1 single period bond is equal to the product of the k one-period returns). Thus, we define

$$(3\text{-}16)\quad R_{t,t+1} = \prod_{k=t}^{T-1} R_k .$$ (3-16) gives

$$(3\text{-}17)\quad 1 = E_t[MRS_{t,t+1} \prod_{k=t}^{T-1} R_k].$$

Suppose that current time is t and the future time is t+1. (3-17) can be rewritten as

$$(3\text{-}18)\quad 1 = E_t[MRS_{t,t+1} R_t]$$

where $R_k = R_{T-1} = R_{t+1-1} = R_t$ and R_t represents one plus the risk-free rate between time t and time t+1. We can rewrite this as

$$(3\text{-}19)\quad 1/R_t = E_t[MRS_{t,t+1}].$$

As Richard and Sundaresan [1981, equation (30)] note, the asset pricing paradigm in (3-15) provides a present value operator that converts dollar payoffs at time t+1 into dollar values at time t. Cox, Ingersoll and Ross (1981, proposition 2) and French (1983) demonstrate that the futures price is the present value of the product of the expected spot price and the gross return from rolling over one period bond. Thus, the futures price, F(t) is the present value at time t of a payoff of $S(t+1)R_t$ at time t+1. Equation (3-18) may be used to price futures contracts.

(3-20) $F(t) = E_t[MRS_{t,t+1} S(t+1) R_t]$

where F(t) is the futures price at time t for a contract that matures at time t+1, S(t+1) is the spot price at time t+1 [the notation of spot price is changed from P(t) to S(t)] and R_t (= 1+R, where R is risk free interest rate) is accumulated interest rate in one period bond from t to from t+1. Note that $S(t+1)R_t$ is the payoff in the futures market at time t+1.

3.2 *Index Arbitrage in the Stock Index Futures*

Index arbitrage is utilized for the equilibrium value of the intertemporal asset pricing model, which is developed by Lucas (1978). We can derive index arbitrage from the cost of carry model, which is based on the principle of arbitrage. I shall explain the cost of carry model to derive index arbitrage. Like most financial futures, stock index futures essentially trade in a full carry market. The full carry market is a market with a condition such that the futures price must equal the spot price plus the cost of carrying the spot to the delivery date of the futures contracts. To fulfill a full carry market, some conditions such as an ease of short selling, large supply of an asset, and high storability are needed. Financial assets fulfill these conditions and they tend to be full carry assets. Therefore, the cost of carry model provides a complete understanding of stock index futures pricing. The cost of carry model for a perfect market* with unrestricted short selling is:

(3-21) $F(t) = S(t) + C$

where $F(t)$ = futures price at time t for a futures contract that expires at time T.

$S(t)$ = spot price at time t.

C = carrying cost including storage, insurance, etc.

Hanson and Kopprasch(1984) explained that the forward price should equal the spot price today plus net carry adjustment (finance cost minus the dividend paid by the underlying security) by the use of arbitrage pricing.

(3-22) $F(t) = S(t) + S(t)R - S(t)D$

or

(3-23) $F(t) = S(t) + S(t)(R - D)$

Where $R = r(T/360)$ and $D = d(T/360)$

T = time remaining, measured in days, in the forward contract

r = cost of financing (risk-free rate of interest per annum at time t)

d = dividend yield rate

We develop a simple index arbitrage argument under the following assumptions:

(a) There are no transaction costs, no taxes, and no restrictions on short selling;

(b) The yield on the security is known, and paid at the end of the futures contract;

(c) Interest rates are nonstochastic.

It is well known that forward and futures prices will not be exactly equal if interest rates

* A perfect market is a market with no transaction costs, no restrictions on short selling, and no limitations to storage. It is also assumed that borrowing and lending rates are equal.

are stochastic because futures contracts are settled daily and forward contracts are not settled until the contract matures. A major difference between a futures contract and a forward contract results from variation margin, the daily marking to the market of an open futures position. If an investor is long a futures contracts and its settlement price on a given day is below that of the previous day, he is required to pay the price difference in cash. If the price rises, he will receive the price difference in cash. However, with forward contracts, no cash changes hands until expiration of the contracts. Under the above assumptions, forward and futures prices are equal. When the conditions of the cost of carry model are violated, arbitrage opportunities arise. Suppose the futures price exceeds the cost of purchasing the stock and carrying it until the futures contract matures [$F(t) > S(t) + S(t)(R-D)$]. An arbitrager could undertake the following trades: at time t,

(i) this trader borrows a certain amount of money [$S(t)$] at an interest rate r for which he/she is obliged to repay the value of $S(t)(1+r)$ (principal plus interest) at time T. At the same time, this trader buys (or invests) the stock for $S(t)$ which offers a yield of d at time T.

(ii) The trader sells one futures contract which will be delivered to the futures buyer at $F(t)$ at time T since the stock price is not certain at time T,

This is called cash and carry strategy. Since the initial value of a futures contracts is zero, no cash is required to create this portfolio.* The trader's cash flow is as follows:

* Although margin deposits are required to begin futures contracts, these deposits may be posted in the form of U.S. Treasury bills. Therefore, the opportunity cost of the margin requirement is zero.

Time	Cash Market		Future Market	
t	a. Borrow S(t)	+S(t)	Sell one forward contract	+F(t)
	b. Buy the stock	-S(t)		
T	a. Sell the stock	+S(T)	Deliver the forward contract	-F(T)
	b. Collect the dividend	+S(t)D		
	c. Repay debt	-S(t)(1+R)		
Profit (or loss)	S(T)+S(t)D-S(t)(1+R)		F(t) - F(T)	

Net profit: S(T)+S(t)D-S(t)(1+R)+F(t)-F(T) = S(t)D - S(t)(1+R) + F(t)

Since F(T) = S(T) (that is, the futures price equals the spot price at the maturity date.)

Because the stock index futures contracts are settled in cash, the spot and futures values at the close of trading for the contract converge. This ensures that the basis goes to zero. At the equilibrium the net profit must be equal to zero, i.e., S(t)D - S(t)(1+R) + F(t) = 0. Otherwise, there exists the possibility of arbitrage profit without investing the trader's own money. Then this is the same results as equation (3-22).

If the futures price is less than the cost of purchasing the stock and delivering it at time T, F(t) < S(t) + S(t) (R - D), arbitrageurs will reverse his position by making the following strategy:

(i) Short the stock and invest the proceeds at the risk-free rate.

(ii) Enter into one long futures contract.

This is called reverse cash-and carry strategy. In the transaction, the arbitrager sells the stock short. In the stock market, this transaction is called short selling because one sells a security that he or she does not actually own. Suppose that an investor contracts a broker to short 100 Microsoft shares. The broker immediately borrows the stocks from another client and sell them in the open market in usual way, depositing the sale proceeds to the investor's account. In other words, the investor with a short position invests all of the proceeds from the short sale at the risk-free rate. Once the stocks are borrowed, the investor can continue to maintain the short position for as long as desired. At some stage, however, the investor will be asked to close out the position by the broker. Sometimes, the investor is short-squeezed and must close out the position immediately even though he or she may not be ready to do so, when the broker runs out of shares to borrow. The broker then uses funds in the investor's account to buy the stock and replace them in the account of the client from which the stocks were borrowed. The investor makes a profit if the stock price goes down and a loss if it goes up. A broker requires significant initial margin from the investors with short positions. As with futures contracts, additional margin may be required if there are increases in the stock price. The investor with a short position must pay the dividends that would normally be received on the stocks that have been shorted. The broker will transfer this to the account of the client from whom the stocks have been borrowed. The trader's cash flow is as follows:

time	Cash Market		Future Market	
t	a. Sell S(t)	S(t)	Buy one futures contract	-F(t)
	b. Invest the proceeds from the short sale	-S(t)		
T	a. Buy the stock	-S(T)	Short the futures contract	F(T)
	b. Pay the dividend	-S(t)D		
	c. From the bonds	S(t)(1+R)		
Profit (or loss)	-S(T)-S(t)D+S(t)(1+R)		-F(t) + F(T)	

Net profit: -S(T)-S(t)D+S(t)(1+R)-F(t)+F(T) = -S(t)D + S(t)(1+R) - F(t)

Then this is the same result as equation (3-22). These two above strategies in stock index futures are called index arbitrage.

We can also rearrange Equation (3-22) to

$$\frac{F(t).-S(t)}{S(t)} = R - D \qquad (3\text{-}24)$$

We can explain index arbitrage in different way, using Equation(3-24). The result is the same. If R is greater than D, then the basis, [F(t)-S(t)]/S(t), is positive. The basis indicates the difference between spot price and the underlying futures price. Hence, profits can be made by buying the stocks underlying the index and shorting futures contracts because F(t) > S(t) [by using the idea of "buy low, sell high" principle]. As a result, the increasing demand for spot will drive the spot price up and the increasing supply of futures will drive futures price down. This will continue until all arbitrage profits disappear. This is called a cash and carry strategy. The cash-and-carry strategy is attractive when spots are priced too low relative to low relative to the futures. On the other hand, if D>R, then the basis is negative and profits can be made by shorting the stocks and taking a long position in the futures contracts. The reverse cash-and-carry strategy is attractive when spots are priced too high relative to the futures. These strategies are well known as index arbitrage.

A numerical example may help to illustrate this simple cash-and -carry arbitrage strategy. Suppose that the time span is only one period, for example, current time t to the future time t+1. Let S(t)= $110.00, F(t)=$115.00, and S(t+1)=F(t+1)=$111.00 denote the cash price for underlying the spot price at time t, the stock index futures price at time t, and the spot and futures price at time t+1 (maturity date), respectively.

Numerical Example of Simple Index Arbitrage

Assumptions: a. There is no tax and no transaction cost.

b. Dividend rate is 4% (annual rate) and paid at time t+1 (maturity date).

c. Interest rate is 10% (annual rate).

d. The price of Stock A rises by $1.00, to $111.00.

Time	Cash Markets	Futures Markets
t	- Borrow $110 at 8%	Sell 1 futures contract for $115.00.

- Buy 1 share of Stock A
  for $110

t+1 - Sell stock A for $111.00. | Buy 1 futures contract for $111.00.

- Receive dividends from
  Stock A [$110*0.04*(90/360)=$1.10]
- Repay the debt.[$110*{1+0.10*(90/360)}=$112.75]

Loss: S(t+1)+S(t)D-S(t)(1+R)
= $111.00+$1.10-$112.75 =-$.65

Profit: F(t)-F(t+1)
=$4.00

Net Profit (or payoff) at time t+1: S(t+1)+S(t)D-S(t)(1+R)+F(t)-F(t+1)=$4.00-$.65=$3.35. So this trader' loss from the spot transaction is $.65. However, the profit from the futures transaction is $ 4.00. On the whole, this trader made a profit because of hedging a futures contract.

By the use of the index arbitrage condition in the stock index futures market, we can substitute $S(t+1)R_t = S(t+1) + S(t)D - S(t)(1+R) + F(t) - F(t+1)$ into Equation (3-20). As I mentioned earlier, the initial value of futures contracts is zero [F(t) = 0]. Thus, the net present value of this arbitrage transaction is worth zero at time t since the traders do not invest their own money in the futures market at time t. For convenience D(t+1) = S(t)D, R(t) = S(t)(1+R), and we do not cancel out F(t+1) and S(t+1). Then F(t) = 0 and (3-20) becomes

(3-25) $F(t) = E_t[MRS_{t,t+1}S(t+1) - F(t+1) + F(t) + D(t+1) - R(t)] = 0$

where $S(t+1)R_t = S(t+1) - F(t+1) + F(t) + D(t+1) - R(t)$

= payoff at time t+1

This equation may be rewritten as

(3-26) $E_t[MRS_{t,t+1}S(t+1) - F(t+1) + F(t) + D(t+1) = E_t[MRS_{t,t+1}R(t)]$

$= E_t[MRS_{t,t+1}R_tS(t)] = E_t[MRS_{t,t+1}R_tS(t)] = E_t[MRS_{t,t+1}R_t][S(t)]$

[Notice that R(t) = S(t) (1+R) and $R_t = (1+R)$ from (3-17)].

(3-27) $E_t[MRS_{t,t+1}S(t+1) - F(t+1) + F(t) + D(t+1) = S(t)$

by (3-17) [since $E_t[MRS_{t,t+1}R_t] = 1$]. Rearrange (3-27) as

(3-28) $E_t[MRS_{t,t+1}S(t+1) - F(t+1)] = S(t) - E_t[MRS_{t,t+1}F(t) + D(t+1)]$.

This equation can be simplified to (3-29)

(3-29) $E_t[MRS_{t,t+1}F(t+1) - S(t+1)] = E_t[MRS_{t,t+1}][F(t) + D(t+1)] - S(t)$

$= [F(t) + D(t+1)] / R_t - S(t)$

by (3-18) [since $E_t[MRS_{t,t+1}] = 1/R_t$]. Multiplying R_t on the both sides gives

(3-30) $E_t\{MRS_{t,t+1}R_t[F(t+1) - S(t+1)]\} = F(t) + D(t+1) - S(t)R_t$.

By the definition of covariance, the left-hand side of (3-30) can be written as

(3-31)

$$E_t[MRS_{t,t+1}R_t]E_t[F(t+1) - S(t+1)] + Cov[MRS_{t,t+1}R_t, F(t+1) - S(t+1)]$$
$$= F(t) + D(t+1) - S(t)R_t$$

Then, (3-32)

$$E_t[F(t+1) - S(t+1)] + Cov[MRS_{t,t+1}R_t, F(t+1) - S(t+1)] = F(t) + D(t+1) - S(t)R_t$$

[since $E_t[MRS_{t,t-1}R_t] = 1$]

(3-32) shows that if the index arbitrage condition holds, the right hand side of (3-32) will be zero. Then, (3-32) becomes

$$E_t[F(t+1) - S(t+1)] = -Cov[MRS_{t,t+1}R_t, F(t+1) - S(t+1)]$$

This implies that the basis risk (difference between expected spot prices and expected futures prices) is equal to the risk premium.

3.3 *Derivation of Time Varying Risk Premia*

As I mentioned earlier (from Chapter 2.3), the risk premium can be defined as by the difference between expected futures price and today's futures price [RP = F(t) - E[F(t+1)]] and the difference between the expected spot price and today's futures price [RP = F(t) - E[S(t+1)]]. Note that E[F(t+1)]=E[S(t+1)]. From the definition of the risk premium, Hodrick and Srivastava [1987, proposition 1] derived the time varying risk premia as

(3-33) $F(t) - E_t[F(t+1)] = Cov_t[MRS_{t,t+1}R_t, F(t+1)]$

where $Cov_t[MRS_{t,t+1}R_t, F(t+1)] =$ the covariance of the marginal rate of substitution between time t and time t+1 with known return of futures contracts R_t and expected futures prices at time t+1. The derivation is as follows: Applying a covariance decomposition to (3-20), yields:

$$F(t) = E_t[MRS_{t,t+1} \prod_{k=t}^{T-1} R_k\ ]\ E_t[F(t+1)] + Cov[MRS_{t,t+1} \prod_{k=t}^{T-1} R_k ; F(t+1)]$$

$$= E_t[F(t+1)] + Cov[MRS_{t,t+1} \prod_{k=t}^{T-1} R_k ; F(t+1)]$$

[since $E_t[MRS_{t,t+1} \prod_{k=t}^{T-1} R_k\ ] = 1$ from (3-16)]. Then, the above equation becomes

$$F(t) - E_t[F(t+1)] = Cov_t[MRS_{t,t+1} R_t, F(t+1)].\quad \text{Q.E.D.}$$

(3-32) may be rewritten as

(3-34)

$$F(t) + D(t+1) - S(t)R_t - E_t[F(t+1) - S(t+1)] = Cov[MRS_{t,t+1} R_t, F(t+1) - S(t+1)].$$

Dviding S(t) on both sides gives

(3-35)

$$\frac{F(t) + D(t+1)}{S(t)} - R_t - E_t[\frac{F(t+1) - S(t+1)}{S(t)}] = Cov_t[MRS_{t,t+1} R_t, \frac{F(t+1) - S(t+1)}{S(t)}].$$

This equation is similar to the equation (3-33). $[F(t) + D(t+1)]/S(t) - R_t$ for F(t) and

$E_t[\frac{F(t+1) - S(t+1)}{S(t)}]$ for $E_t[F(t+1)]$. So, (3-35) may be written as

(3-36) $$\frac{F(t)}{S(t)} + \frac{D(t+1)}{S(t)} - R_t = E_t[\frac{F(t+1)}{S(t)}] - E_t[\frac{S(t+1)}{S(t)}] + RP_{t,t+1}$$

where in this study, the covariance of equation (3-35) is defined as TVRP. Hence,

$$RP_{t,t+1} = Cov_t[MRS_{t,t+1} R_t, \frac{F(t+1) - S(t+1)}{S(t)}].$$

This equation shows that if an investor expects large profits from the short sale in futures (expects the futures prices to go down relative to the spot prices), then the investor will substitute today's consumption for investing more risky assets. The risk averse investor will ask for more compensations, if the deviation in the marginal substitution and the expected return increases and the basis risk increases.

3.4 *Hedging Model*

Duffie (1989) defines futures hedging as taking a position in futures contracts that offsets a pre-existing spot price risk. One calculates the optimal allocation of the number of futures contracts held and spot contracts held. This allocation is called hedge ratio, which indicates h = futures position / cash market position.

Figlewski (1984) employs a hedging model to stock index futures. The derivation is as follows: Let $\tilde{R}_p$, $\tilde{R}_I$, and $\tilde{R}_F$ denote the random variable returns on the portfolio, the spot index, and the index futures contract, respectively, assuming a holding period of length T. The return on the portfolio is

$$(3\text{-}37)\quad \tilde{R}_p = \frac{\tilde{V}_T - V_0 + \tilde{D}_p}{V_0}$$

where

V_0 and $\tilde{V}_T$ = the beginning and ending market values for the portfolio,

$\tilde{D}_p$ = the cumulative values as of T of the dividends paid out on the portfolio during the period, assuming reinvestment at the riskless rate of interest from the data of payout until T. The dividend payout is a random variable because the amount, its timing, and the reinvestment rate are all uncertain as of time t. The return on the index portfolio is

(3-38) $\widetilde{R}_I = \dfrac{\widetilde{I}_T - I_0 + \widetilde{D}_I}{I_0}$

where variables are defined analogously to (3-37). The return on the index futures contract is

(3-39) $\widetilde{R}_F = \dfrac{\widetilde{F}_T - F_0}{I_0}$

The rate of return on a futures contract is not a well-defined concept, since taking a futures position does not require an initial outlay of capital. The initial margin deposit to open a futures position does not represent an investment of capital since it can be posted in the form of interest bearing Treasury bills (risk-free interest rate). For convenience we define the rate of return on the futures as the change in the futures price divided by the initial level of the spot index.

Suppose the return on a hedged portfolio in which futures contracts on N_F index "shares"* have been sold short against the long portfolio of stocks. Let N_F be the number of index units sold short in the futures markets. Let $N_S = \dfrac{V_0}{I_0}$ be the number of index units of value in the spot portfolio. So, the return on a hedged portfolio is

(3-40) $\widetilde{R}_H = \dfrac{\widetilde{V}_T - V_0 + \widetilde{D}_P}{V_0} - \dfrac{N_F * I_0}{V_0} * \dfrac{\widetilde{F}_T - F_0}{I_0} = \widetilde{R}_p - \dfrac{N_F}{N_S}\widetilde{R}_F = \widetilde{R}_P - h\widetilde{R}_F$,

* An index share is defined to be an amount of the index portfolio whose market value is equal to $1 times the spot index. Most currently traded stock index futures have contract sizes of 500 index shares.

where h (= - N_F / N_S) is hedge ratio which is a short (or long) position in h units of futures per units of the spot position. Because a hedging ratio usually consists of opposite positions, i.e., short in futures and long in spot, the signs will be $N_F \langle 0$ and $N_S \rangle 0$, so that h>0. Since both futures and spot prices tend to move in the same direction, holding the same amounts of the contracts with the opposite positions will offset the risk of the unknown future price movements.

The expected rate of return on a hedged portfolio and variance of return for the hedged position are given by

(3-41) $\overline{R}_H = \overline{R}_P - h\overline{R}_F$

(3-42) $Var[\widetilde{R}_H] = Var[\widetilde{R}_P] + h^2 Var[\widetilde{R}_F] - 2hCov[\widetilde{R}_P, \widetilde{R}_F]$

where bars represent expectations. Notice that $\overline{R}_H = \overline{R}_P$ and $Var[\widetilde{R}_H] = Var[\widetilde{R}_P]$ if h=0 which means unhedged position.

To find the optimal hedge ratio which minimizes risk, we set the derivative of (3-42) with respect to h equal to zero and obtain

(3-43) $h^* = \dfrac{Cov[\widetilde{R}_p, \widetilde{R}_F]}{Var[\widetilde{R}_F]} = \rho_{pF} \dfrac{\sigma_P}{\sigma_F}$.

This is easily computed in practice by running a regression of $R_p on R_F$ using historical data. The slope coefficient in the equation is h*. Thus the easiest way to find the risk-minimizing hedge ratio is to estimate the following regression (3-44).

Figlewski (1986) suggested that optimal hedge ratio comes from the regression

(3-44) $\Delta S = \alpha + h\Delta F_t + \varepsilon_t$

where ΔS =S(t+1)-S(t), ΔF =F(t+1)-F(t), h = the risk-minimizing hedge ratio = the ratio of the size of the position taken in the futures contracts to the size of the spot position, α = the constant term, and ε_t = forecast error. The hedge ratio equals the sample covariance between the independent variable and dependent variable divided by the sample variance of the independent variable. In other words, the optimal hedge ratio is the product of the coefficient of correlation between ΔS and ΔF and the ratio of the standard deviation of ΔS to the standard deviation of ΔF. For example, if ρ = 1 and $\sigma_F = \sigma_p$, the optimal hedge ratio, h, is 1.0. This implies that the futures price mirrors the spot price perfectly. If ρ = 1 and $\sigma_F = 2\sigma_p$, the optimal hedge ratio is 0.5. This implies that the futures price changes by twice as much as the spot price. The time varying hedge ratio is similar to the conventional OLS hedge ratio except that time varying conditional moments replace the time invariant unconditional moments. The hedge ratio will change over time because conditional moments change as new information arrives to the market. This time varying hedge ratio is the same as the conventional OLS hedge ratio if the joint distribution of spot and futures is constant over time. However, I did not derive the time varying hedge ratio because it is not my major concern. The constant term of the regression does not appear in this hedging model although Figlewski includes the constant term in the regression. Another way to write the equation (3-36) is

$$(3\text{-}45)\quad E_t[\frac{S(t+1)}{S(t)}]+\frac{D(t+1)}{S(t)}-R_t = E_t[\frac{F(t+1)}{S(t)}]-\frac{F(t)}{S(t)}+RP_{t,t+1}.$$

This equation (3-45) is analogous to the optimal hedging model (3-44). The equation (3-45) will be utilized to test for the presence of TVRP for the GARCH-M model.

3.5 Application of GARCH-M model

The GARCH-M model is defined as

$$y_{t+1} = b'x_{t+1} + \lambda h_{t+1} + \varepsilon_{t+1}$$

$$\varepsilon_{t+1}|\Psi_t \sim N(0,h_{t+1})$$

$$h_{t+1} = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon^2_{t+1-i} + \sum_{j=1}^{p}\beta_j h_{t+1-j}$$

where in finance models, dependent variable y_{t+1} is denoted by the mean equation, which can be considered as the excess return, x_{t+1} is a kx1 vector of exogenous variables, λh_{t+1} is the risk premium, which means the increase in the expected rate of return due to an increase in the variance of the return, ε_{t+1}, conditional on the realized values of the set of variables $\Psi_t = [y_t, x_t, y_{t-1}, x_{t-1}, \cdots]$, is normally distributed with conditional mean zero $E[\varepsilon_{t+1}|\varepsilon_t] = 0$ and conditional variance h_{t+1}, α, β , and b are vectors of unknown parameters to be estimated, h_{t+1} is clearly function of the elements of Ψ_t which is generated by the lagged errors $\varepsilon^2{}_{t+1-i}$ and the past conditional variances h_{t+1-j}(i=1,....,q and j=1,...,p). If changing risks rise (the conditional variances increase), the expected returns (or mean equation) will increase because risk averse investors need more compensations when they hold riskier assets. GARCH-M model can be viewed as a statistical implementation of the mean-variance analysis in finance. This model also provide a good measure to estimate the TVRP. It is possible to use the conditional variance to test the existence of TVRP by letting y_{t+1} be the expected returns and letting λ be the effect of the risk premia. In other words, it is possible to test the validity of the rational expectations hypothesis to allow risk averse investors.

After applying this GARCH-M model (see chapter 4 for more details) to (3-45), an empirical estimated equation with TVRP is

$$\text{(3-46)}\ \frac{S(t+1)}{S(t)} + \frac{D(t+1)}{S(t)} - R_t = \mu + b'[\frac{F(t+1) - F(t)}{S(t)}] + \lambda' f(h_{t+1}) + \varepsilon_{t+1}$$

or equivalently,

$$y_{t+1} = \mu + b'x_{t+1} + \lambda' f(h_{t+1}) + \varepsilon_{t+1}$$

where μ , a constant term, is added to pick up some information such as trend movement. The intercept term μ is also the mean of the excess returns on the stock index and has the dimension of an average daily risk premium in (3-46). h_{t+1} is a conditional variance, is assumed to be an GARCH stochastic process, and is a good proxy variable for estimating the risk because it includes only unpredictable components and consists of the information sets Ψ_t generated by ε_{t+1}. $f(h_{t+1})$ is an unknown functional form of TVRP. λ measures the effect of the risk premia. If λ =0, then the TVRP do not change the behavior of dependent variable of (3-46). b indicates the hedging ratio of the spot position against the futures contracts. $[\frac{F(t+1)-F(t)}{S(t)}] = x_{t+1}$ = futures returns.

$\frac{S(t+1)}{S(t)} + \frac{D(t+1)}{S(t)} - R_t = y_{t+1}$ = excess returns on a stock index with dividend from the transaction of spot markets [since $\frac{S(t+1)}{S(t)} + \frac{D(t+1)}{S(t)}$ = real spot returns between time t and t+1 and R_t is one plus risk-free rate, see equation (3-15)]. F(t+1) is a random variable and it is assumed to be independent of ε_{t+1} and does not involve the parameters of b and h_{t+1} to obtain the unbiased estimation. ε_{t+1} is the difference between the ex-ante and ex-post rate of return which in efficient markets would be unpredictable. Equation (3-46) is not a traditional optimal hedging model. However, we can call this equation an modified hedging model because (3-46) still indicates that the spot positions are hedged by futures contracts.

First, to measure risk aversion from excess returns on a stock index, we estimate the time varying conditional variance without x_{t+1} term.

$$\text{(3-47)} \quad \frac{S(t+1)}{S(t)} + \frac{D(t+1)}{S(t)} - R_t = \mu + \lambda f(h_{t+1}) + \varepsilon_{t+1}$$

Second, we add x_{t+1} term (which is the futures returns) to model (3-47). Because the futures price is an important determinant of the excess returns on a stock index. Market prices reveal information among the traders and the prices of risky assets (futures prices) will be a function of the information that the economic agents have. If futures prices could contain all available information to yield the unbiased estimators of the underlying expected spot prices, then the expected spot prices would be the function of futures prices. That's why we apply the futures prices to (3-47).

Chapter IV

ECONOMETRIC MODELS

4.1 *Time Series Behavior of Spot and Futures Returns*

Distributions of Spot and Futures Returns

Almost all studies in the distribution of futures prices agree that changes in futures prices are not normally distributed, but that the distribution of percentage changes in futures price is leptokurtic. Leptokurtosis is the tendency for a distribution to have many extreme observations relative to a normal distribution. The greater frequency of extreme observations makes the leptokurtic distribution have "fat tail" and high peak.

We also test for the normality of the disturbances and use the following the Bera-Jarque (1982) test statistic:

$$\text{Bera-Jarque Test} = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right]$$

where n is sample size, and S and K are the sample skewness and kurtosis. Since for a normal distribution the value of skewness is zero and the value of kurtosis is 3, (K-3) represents excess kurtosis. Under the null hypothesis that the residuals are normally distributed, this statistic also asymptotically follows χ^2 distribution with two degrees of freedom. If the B-J statistic is greater than the critical values for χ^2 statistics [9.21(1%) and 5.99(5%), respectively], one can reject the hypothesis that the residuals are normally distributed. Otherwise, we can not reject the normality assumption.

Autocorrelations of Spot and Futures Returns

A time series is autocorrelated if the value of one observation in the series is statistically related to another. In first-order autocorrelation, for example, one observation is related to the immediately preceding observation. This question has considerable practical importance. For example, if futures returns show positive first-order autocorrelation, then positive returns in one period tend to be followed by positive returns in the next period. Similarly, negative returns tend to be followed by subsequent negative returns. If the correlation were strong enough, it would be possible to devise trading strategies to profit from this follow-on tendency.

Any time series data can be thought of as being generated by a stochastic process. A stochastic process is said to be covariance stationary if its mean and variance are constant over time and the value of covariance between two time periods depends on lag between the two time periods.* Let Y(t) be a stochastic time series with these properties:

Mean: $E[Y(t)] = \mu$

Variance: $Var[Y(t)] = E[Y(t) - \mu]^2 = \sigma^2 = \gamma_0$

Covariance: $Cov[Y(t), Y(t+k)] = E[\{Y(t) - \mu\}\{Y(t+k) - \mu\}] = \gamma_k$

where γ_k, the autocovariance at lag k, is the covariance between the values of Y(t) and Y(t+k). Note that if k=0, we obtain γ_0, which is the variance of Y; if k=1, γ_1 is the covariance between two adjacent values of Y. If Y(t) is to be stationary, the mean,

* In the time series literature such a stochastic process is known as a weakly stationary stochastic process.

variance, and autocovariances of Y(t+k) must be same as those of Y(t). If a time series is not stationary, it is called a nonstationay time series. One simple test of stationarity is based on the autocorrelation function (ACF). The ACF at lag k is defined as

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\text{covariance}}{\text{variance}}.$$

The statistical significance of any ρ_k can be judged by its standard error. If a time series is purely random, the sample autocorrelation coefficients are approximately normally distributed with zero mean and variance 1/T, where T is the sample size. For our data T = 1300, implying a standard error of $1/\sqrt{1300}$ =0.0277. Then, the 95% confidence interval for any ρ_k is $\pm 1.96(0.0277)$ = ± 0.0543 on either side of zero. Thus, if an estimated ρ_k falls outside the interval ± 0.0543, we can reject the null hypothesis that the true ρ_k autocorrelation coefficients are equal to zero.

To more investigate the hypothesis that all the autocorrelation coefficients are simultaneously equal to zero, we can use the Ljung-Box (1978) statistic, which is defined as

$$\text{LB} = n(n+2)\sum_{k=1}^{m}[\rho_k^2 / (n-k)] \sim \chi_m^2$$

where n = sample size and m = lag length (in my study, m = 24). The LB statistic is approximately distributed as the chi-square distribution with m df. If the computed LB exceeds the critical value, we can reject the null hypothesis of no autocorrelation.

Unit Root Test of Stationarity

An alternative test of stationarity is known as the unit root test. Consider the following simple AR(1) model with zero mean, white noise innovations:

(4-1) $Y_t = Y_{t-1} + \varepsilon_t$

where ε_t is the stochastic error term that has zero mean, constant variance σ^2, and is non-autocorrelated. Such an error term is also known as a white noise error term. The equation (4-1) is a AR(1), regression in that we regress the value of Y at time t on its value at time t-1. Now using the lag operator L so that $LY_t = Y_{t-1}, L^2Y_t = Y_{t-2}$ and so on, we can write the equation (4-1) as $(1 - L)Y_t = \varepsilon_t$. The term unit root refers to the root of the polynomial in the lag operator. With ρ, the equation (4-1) can be rewritten as

(4-2) $Y_t(1 - \rho L) = \varepsilon_t$

If ρ=1, the we can say that the stochastic variable Y_t has a unit root problem, i.e., a nonstationary situation. In this case, the stochastic variable Y_t that has a unit root is known as a random walk. A random walk is an example of a nonstationary time series. Equation (4-2) is often expressed in an alternative form as

(4-3) $\Delta Y_t = (\rho - 1)Y_{t-1} + \varepsilon_t = bY_{t-1} + \varepsilon_t$

where $b = (\rho - 1)$ and $\Delta Y_t = Y_t - Y_{t-1}$. If $b = 0$, then we can say that the first difference of a random walk time series ($= \varepsilon_t$) are a stationary time series because by assumption ε_t is purely random. If a time series becomes stationary after the first difference, we say that the original random walk series is integrated of order 1, denoted by I(1).

To find out if a time series Y_t is non-stationary, run the regression (4-3) and find out if ρ is statistically equal to 1 or, equivalently, estimate (4-3) and find out if the null hypothesis that $b = 0$ on the basis of the t statistic. Unfortunately, the t value obtained does not follow Student's t distribution even in large samples. The appropriate test statistic is the pseudo t-statistic whose critical value can be obtained from MacKinnon (1991). If the computed absolute Pseudo t values is less than Mackinnon ADF absolute critical value, then we reject the hypothesis that the given time series is stationary. If , on the other hand, it is more than the critical value, the time series is stationary.

A convenient reformation of the I(1) with the inclusion of the constant is

(4-4) $\Delta y_t = a + by_{t-1} + \varepsilon_t$

With this formation, Dickey and Fuller (1981) test for a unit root is carried out by testing the hypothesis that b equals to zero. If the error term ε_t is autocorrelated, one modifies (4-4) as follows:

(4-5) $\Delta y_t = a + by_{t-1} + \sum_{i=1}^{p} c_i \Delta y_{t-i} + \varepsilon_t$

where $\Delta y_{t-i} = y_{t-i} - y_{t-i-1}$ and p is the number of lagged values of first differences Δy_t. The number of lagged difference terms to include is often determined empirically, the idea being to include enough terms so that the error term in (4-5) is serially independent. (4-5) is called the Augmented Dickey - Fuller (ADF) test (to examine whether estimated residuals are stationary.

White's Test

In this study, White (1980) test provides a test for the presence of unconditional heteroskedasticity. The procedure of White (1980) test is as follows:

(4-6) $Y_t = bX_t + \varepsilon_i$ (i = 1,2,3,.......,T)

(Step 1) Estimate the (4-6) and obtain the residuals, $\hat{\varepsilon}_i$.

(Step 2) Run the auxiliary regression:

$$\hat{\varepsilon}_t^2 = \alpha_1 X_t + \alpha_2 X_t^2$$

That is, the squared residuals from the original regression are regressed on the original X variable and their squared value.

(Step 3) Under the null hypothesis that there is no heteroscedasticity, it can be shown that sample size (T) times the R^2 obtained from the auxiliary regression asymptotically follows the chi-square distribution with degree of freedom equal to the number of regressors (excluding the constant term) in the auxiliary regression. That is, T* $R^2 \sim \chi^2_{df}$ which has, asymptotically, a chi-square distribution with 2 df (in this study, there are 2 df since there are 2 regressors in the auxiliary regression). If T* R^2 is greater than critical Chi-square value for 2 df (9.21 for 1%, 5.99 for 5%, and 4.61 for 10%), we conclude that there is heteroscedasticity, which is to say that in the auxiliary regression, $\alpha_1 \neq 0$ and $\alpha_2 \neq 0$.

4.2 *ARCH , GARCH, and ARCH-M Models*

Ordinary regression analysis assumes that the error variance is the same for all observations. When the error variance is not constant, the data are said to be heteroskedastic. Since OLS regression assumes constant error variance, heteroscedasticity causes the OLS estimates to be inefficient. The Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized Autoregressive conditional Heteroscedasticity (GARCH) models provide a means of estimating and correcting for the changing variability of the data.

Heteroscedasticity is usually associated with cross sectional data, while homoscedasticity has been dealt with time series data. However, the volatility of time-series data from financial markets indicate that variances may not be constant over time. Engle (1982) first introduced the ARCH model to explain heteroscedasticity(volatility) in time series data. Engle's results suggest that heteroscedastic error term (the variances are not constant but may change over time) is explained by the conditional variance, which depends upon past information and may be a random variable.

Engle (1982) derived the Lagrange Multiplier(LM) test to detect the ARCH effects [or to test whether the residuals ε_t from a regression model show time-varying heteroscedasticity] before the estimation is performed. ARCH test proceeds as follows. First, for example, the regression of equation (4-6) is estimated by OLS for all available observations to obtain b and ε_t and the OLS sample residuals $\hat{\varepsilon}_t$ are saved. Next, $\hat{\varepsilon}_t^2$ is regressed on a constant and q of its own lagged values.

(4-7) $h_t = Var(\varepsilon_t) = \hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + ... + \alpha_q \varepsilon_{t-q}^2$

where the α 's are the coefficients of the ARCH process and t = 1,2,...,T. The null hypothesis is $H_0 : \alpha_1 = \alpha_2 .... \alpha_p = 0$, in which case $h_t = \alpha_0$, that is the case of homoscedastic error variance. The sample size T times the R_ε^2 from the regression of equation (4-7) converges in distribution to a χ^2 with q degrees of freedom under the null hypothesis.

The simple first-order ARCH model is

(4-8)
$$y_t = b' x_t + \varepsilon_t$$
$$h_t = Var(\varepsilon_t) = \hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$
$$\varepsilon_t | \Psi_{t-1} \sim N(0, h_t)$$

where α and b' are vectors of unknown parameters and Ψ_{t-1} is all available information at time t-1. The first equation is called the mean equation. The second, which is called the variance equation, parameterizes the conditional variance as a function of previous forecast errors. This variance equation is an ARCH stochastic process. The conditional forecasting error, given past information set Ψ_{t-1}, is normally distributed with conditional mean zero $E[\varepsilon_t | \varepsilon_{t-1}] = 0$ and conditional variance h_t. If $\alpha_1 = 0$, ε_t will be Gaussian white noise and if it is a positive number, successive observations will be dependent through higher-order moments. If α_1 is too large, the variance of the process will be infinite. The first order ARCH process also generates data with fatter tails than the normal density distribution.

First of all, we examine the difference between unconditional and conditional variance. The unconditional variance of the dependent variable can be written as $Var(y_t) = Var(b'x_t) + Var(\varepsilon_t)$ so that unconditional variances include predictable components $[Var(b'x_t)]$ as well as unpredictable components $[Var(\varepsilon_t)]$, which may cause an inaccurate measurement of risk. Therefore, this unconditional variance is not good measure of uncertainty. However, the conditional variance measures only unpredictable components, $Var(y_t|y_{t-1}) = Var(\varepsilon_t|\varepsilon_{t-1}) = h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2$

Therefore, the conditional variance is more useful variable measuring risk because current information is considered in predicting over the future variances which may change over time. If the process generating the disturbances is stationary variance, the unconditional variance is not changing over time. Then, $Var(\varepsilon_t) = Var(\varepsilon_{t-1})$.

$$Var(\varepsilon_t) = \alpha_0 + \alpha_1 Var(\varepsilon_{t-1}) \rightarrow Var(\varepsilon_t) - \alpha_1 Var(\varepsilon_{t-1}) = \alpha_0 \rightarrow Var(\varepsilon_t)[1-\alpha_1] = \alpha_0$$

[Since $Var(\varepsilon_t) = Var(\varepsilon_{t-1})$]

$$\therefore Var(\varepsilon_t) = \frac{\alpha_0}{1-\alpha_1}$$

Therefore, the conditions for a first-order linear ARCH process to be covariance stationary can be generalized for qth-order process as follows:

$$Var(\varepsilon_t) = \frac{\alpha_0}{1-\sum_{i=1}^{q}\alpha_t}$$

Therefore, the model obeys the classical assumptions [$E(\varepsilon_t) = 0$ (unconditionally)

$Var(\varepsilon_t) = \dfrac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i}$ and $Cov[\varepsilon_t, \varepsilon_s] = 0$ (for all $t \neq s$)], and OLS is the most efficient linear estimator of b. But there is a more efficient nonlinear estimator. The log-likelihood function for this model is given by Engle(1982).[see 4.3]. Engle also shows the conditions for the process to be stationary as follows:

For integer r, the first-order linear ARCH process with $\alpha_o > 0, \alpha_1 \geq 0$ is stationary only if

$\alpha_1^r \prod_{i=1}^{r} (2i - 1) < 1$ [see the theorem 1&2 of Engle(1982)].

The qth-order linear ARCH process, ARCH(q), is a short memory process in that only the most recent q squared residuals are used to estimate the changing variance. The ARCH(q) is $h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$.

Bollerslev(1986) introduces the conditional variances by adding the past conditional variances in the conditional variance equation. This model is called Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model allowing for a much more flexible lag structure. The GARCH model allows long memory process which use all past conditional variances to estimate the current variance. The GARCH model is defined as follows: Conditioned on an information set at time t-1, the distribution of the disturbance is assumed to be $\varepsilon_t | \Psi_{t-1} \sim N(0, h_t)$,

and the conditional variance is

(4-9) $h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$

where $p \geq 0, q > 0, \alpha_0 > 0, \alpha_i \geq 0 (i = 1,\cdots,q), \beta_j \geq 0 (j = 1,\cdots,p)$

Note that for p=0 GARCH(q,0) is equivalent to ARCH(q), and for p=q=0 ε_t is white noise. In the ARCH(q) process the conditional variance is specified as a linear function of past sample variances only, whereas the GARCH(q,p) process allows lagged conditional variances. The extension of the ARCH model to the GARCH model is similar to the extension of the AR model to ARMA model in traditional time series using the square of error terms instead of error terms. Autocorrelation and partial autocorrelation are useful tools for identifying the behavior of the GARCH process. There is no general test for the GARCH model, either. Bollerslev (1986) also shows the necessary and sufficient conditions for the stationarity of the GARCH(q,p) process. The GARCH(q,p) process is covariance stationary with $E(\varepsilon_t) = 0$ (unconditionally) $Var(\varepsilon_t) = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j}$ and $Cov[\varepsilon_t, \varepsilon_s] = 0$ (for all $t \neq s$) if $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$ [Theorem 1 of Bollerslev (1986)], where q = lag of past sample variance and p = lag of the conditional variances. (4-9) shows that this implies an ARMA process for ε_t^2 where the autoregressive coefficient is given by $\alpha_i + \beta_j$. This ARMA process for ε_t^2 would have a unit root if $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j = 1$. Engle and Bollerslev (1986) referred to a model satisfying this condition as an integrated GARCH process (IGARCH). If ε_t follows an

IGARCH process, then the unconditional variance of ε_t is infinite, so neither ε_t nor ε_t^2 satisfies the condition of a covariance-stationary process.

Engle, Lilien, and Robins(1987) introduced the ARCH in Mean (ARCH-M) regression model allowing the conditional variance to be a regressor of the mean equation. The ARCH-M model is defined as

(4-10) $y_t = b'x_t + \lambda h_t + \varepsilon_t$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

$$\varepsilon_t | \Psi_t \sim N(0, h_t)$$

The ARCH -M (and GARCH-M) model is very useful in a financial model because the conditional variance can measure the degree of risk. Changing risk(conditional variance) can affect the expected returns (mean equation)on holding assets. When risk rises(the conditional variance increase), the expected returns will be increased since risk averse agents require more compensation if they hold riskier assets(trade-off of risk and return). It is possible to use the conditional variances to test the existence of TVRP by letting the mean equation be the excess returns and testing whether or not $\lambda = 0$.

4.3 Estimating and Testing the GARCH and ARCH-M model

Estimating and Testing the GARCH Model

The following estimation procedure were developed by Engle(1982) and Bollerslev(1986). The GARCH model is given by

$$y_t = b'x_t + \varepsilon_t$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} = z_t'\omega$$

$$\varepsilon_t | \Psi_t \sim N(0, h_t)$$

where $z_t' = (1, \varepsilon_{t-1}^2, \cdots, \varepsilon_{t-q}^2, h_{t-1}, \cdots, h_{t-p}), \omega' = (\alpha_0, \alpha_1, \cdots, \alpha_q, \beta_1, \cdots, \beta_p), \theta = (b', \alpha', \beta')$.

For normally distributed disturbances, the log likelihood function for a sample of T observations is

$$\text{(4-11)} \quad \ln L = \sum_t -\frac{1}{2}[\ln(2\pi) + \ln \sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2}] = -\sum_t \frac{1}{2}[\ln(2\pi) + \ln h_t + \varepsilon_t^2 h_t^{-1}] = \sum_t l_t(\theta)$$

where $\varepsilon_t = y_t - b'x_t$. Differentiating with respect to the variance parameter gives

$$\frac{\partial l_t}{\partial \omega} = \frac{1}{2} h_t^{-1} \frac{\partial h_t}{\partial \omega} [\frac{\varepsilon_t^2}{h_t} - 1],$$

$$\frac{\partial l_t^2}{\partial \omega \partial \omega'} = [\frac{\varepsilon_t^2}{h_t} - 1] \frac{\partial}{\partial \omega'} [\frac{1}{2} h_t^{-1} \frac{\partial h_t}{\partial \omega}] - \frac{1}{2} h_t^{-2} \frac{\partial h_t}{\partial \omega} \frac{\partial h_t}{\partial \omega'} \frac{\varepsilon_t^2}{h_t}$$

where $\frac{\partial h_t}{\partial \omega} = z_t + \sum_{j=1}^{p} \beta_j \frac{\partial h_{t-j}}{\partial \omega}$ [(21) of Bollerslev (1986)].

The difference between the ARCH and the GARCH model is the inclusion of the recursive part in (21) of Bollerslev (1986). Note that $\beta_1 < 1$ guarantees that (21) of Bollerslev (1986) is stable. To obtain maximum likelihood estimates, an iterative procedure is needed. The recursive terms complicate the use of Newton's method of scoring for the GARCH model. Instead the iterative method developed by the Berndt, Hall, Hall, and Hausman(BHHH) (1974) algorithm is employed to obtain the consistent estimates of the variance parameters. BHHH is an iterative method in which updated terms are calculated by a regression of one's on the scores of the likelihood function. BHHH algorithm is a type of Gauss-Newton method, where the Hessian matrix is approximated by the outer product of the scores and it estimates the maximum value of the likelihood function by verifying the numerical derivatives of the first-order conditions. The method of BHHH is

$$\theta^1 = \theta^0 + H^{-1}g^0$$

where $H = [\sum_t \frac{\partial l_t}{\partial \theta}\frac{\partial l_t}{\partial \theta'}]^{-1} = [g'g]^{-1}, g'i = \frac{\partial L_t}{\partial \theta}$, i is a T*1 unit vector. H indicates the Hessian of θ and g indicates the first derivative vector of lnL with respect to θ. The starting values are taken from the classical linear regression. Therefore, BHHH method for estimating the variance parameters would be

$$\theta^{(i+1)} = \theta^{(i)} + \lambda_i [\sum_{t=1}^{T} \frac{\partial l_t}{\partial \theta}\frac{\partial l_t}{\partial \theta'}]^{-1} \sum_{t=1}^{T} \frac{\partial l_t}{\partial \theta},$$

where $\frac{\partial l_t}{\partial \theta}$ is evaluated at $\theta^{(i)}, \lambda_i$ is a variable step length chosen to maximize the likelihood function in the given direction. The maximum likelihood estimate $\hat{\theta}^T$ is strongly consistent for θ^0 and asymptotically normal with mean θ^0 and covariance matrix

$$H^{-1} = -E[\frac{\partial^2 l_t}{\partial \theta \partial \theta'}]^{-1}.$$

To test the presence of GARCH(or ARCH) process, the LM test is employed. The LM tests are convenient for testing restrictions in either the mean or the variance specification because it is simple if the model has already been estimated under the null hypothesis and easily constructed from the matrix of scores. The LM test statistic for $H_0: \beta_j = 0$ is given by

$$LM = \frac{1}{2} f_0' Z_0 (Z_0' Z_0)^{-1} Z_0' f_0 = TR^2$$

where $f_0 = [\varepsilon_1^2 h_1^{-1} - 1, \ldots, \varepsilon_T^2 h_T^{-1} - 1]$

$$Z_0 = [h_1^{-1} \frac{\partial h_1}{\partial \omega}, \ldots, h_T^{-1} \frac{\partial h_T}{\partial \omega}]'$$

This statistic will asymtotically be chi square with the number of degrees of freedom when the null is true. R^2 is the squared multiple correlation coefficient between $f_0 and Z_0$. This corresponds to TR^2 from the OLS regression in the first BHHH iteration for the general model starting at the maximum likelihood estimates under the null. For example, an LM test of ARCH(q) against the hypothesis of no ARCH effects[ARCH(0), the classical model] can be tested by computing $\chi^2 = TR^2$ in the regression of ε_t^2 on a constant and q lagged values. The statistic is asymtotically chi squared with q degrees of freedom. Values larger than the critical value give evidence of the presence of ARCH effects.

Estimating and Testing the ARCH-M model

The GARCH-M model is an extension of ARCH-M model by adding the past conditional variances in the conditional variance equation. To estimate and test the GARCH-M model, the ARCH-M model is explained for the GARCH-M model. The estimation procedure for ARCH -M model has been the maximum likelihood approach. The following estimation procedure were developed by Engle et al. (1987). For normally distributed disturbances, the log likelihood function for the first-order ARCH-M model for a sample of T observations is

$$(4\text{-}12)\quad \ln L = \sum_t -\frac{1}{2}[\ln(2\pi) + \ln\sigma_t^{\,2} + \frac{\varepsilon_t^2}{\sigma_t^2}] = -\sum_t \frac{1}{2}[\ln(2\pi) + \ln h^2{}_t + \varepsilon_t^2 h_t^{-2}]$$

$$= -\frac{1}{2}\sum_t \ln[\alpha_0 + \alpha_1(y_t - b'x_t - \lambda h_t)^2] - \frac{1}{2}\sum_t \frac{(y_t - b'x_t - \lambda h_t)^2}{[\alpha_0 + \alpha_1(y_t - b'x_t - \lambda h_t)^2]}$$

where $\varepsilon_t = y_t - b'x_t - \lambda h_t$, $\theta = (b', \lambda', \alpha')$

Maximum likelihood estimates (MLEs) are obtained by maximizing the log-likelihood function, the equation (4-11), with respect to the unknown parameters $b', \lambda, \alpha_0, \alpha_1$. If $\lambda = 0$, the OLS estimator of b is consistent, and if the x_t can be treated as fixed constants, the least squares standard errors are correct. Neither of these statements is true if $\lambda \neq 0$. The first order conditions for a maximum likelihood estimation (differentiating with respect to b) gives

$$(4\text{-}13)\quad \frac{\partial L_t}{\partial b} = \frac{1}{2}[\frac{x_t'\varepsilon_t}{h_t^2} + \frac{1}{h_t}\frac{\partial h_t}{\partial b}(\frac{\varepsilon_t^2}{h_t^2} - 1) + \frac{\partial h_t}{\partial b}\frac{\lambda\varepsilon_t}{h_t^2}]$$

The first term in (4-13) correspond to the first order condition for an exogenous heteroskedastic correction. The second term comes from the ARCH model of Engle (1982), because the conditional variance is a function of the b's. The last term appears because the h_t is a part of the mean regression function. The expression in (4.13) gives the standard ARCH model when $\lambda = 0$. [see (20) of Engle (1982)]. The derivatives of h_t with respect to all the parameters of the model have a recursive structure. The derivatives of the likelihood function with respect to the other parameters have a similar form.

$$\frac{\partial L}{\partial \lambda} = \frac{1}{2}\left[\frac{\varepsilon_t}{h_t} + \frac{1}{h_t}\frac{\partial h_t}{\partial \lambda}\gamma_t\right],$$

$$\frac{\partial L}{\partial \alpha} = \frac{1}{2}\left[\frac{1}{h_t}\frac{\partial h_t}{\partial \alpha}\gamma_t\right],$$

where $\gamma_t = \frac{\varepsilon_t^2}{h_t^2} - 1 + \frac{\lambda\varepsilon_t}{h_t}$. To obtain maximum likelihood estimates, the iterative method developed by the Berndt, Hall, Hall, and Hausman(BHHH) (1974) algorithm is employed. This procedure is the same as that of estimating the GARCH model.

To test the presence of GARCH(or ARCH) process, the LM test is employed. The LM test is attractive if estimation under the alternative is complicated since it requires only estimation under the null. The LM test statistic for $H_0 : \alpha_i = \beta_j = 0$ is given by

$$LM = i'g(g'g)^{-1}g'i = TR^2$$

This statistic will asymtotically be chi square with the number of degrees of freedom when the null is true. R^2 is the squared multiple correlation coefficient between g and i. This corresponds to TR^2 from the OLS regression in the first BHHH iteration for the general model starting at the maximum likelihood estimates under the null. This convenient test statistic can be also computed to test high-order ARCH or GARCH specification.

4.4 *Specification Tests*

The LM test is applied to test the specification of the mean and conditional variance equations. The procedure is as follows:

H_0: Restricted Model;

(4-14) $Y_{t+1} = b_1 + b_2 X_{t+1} + \lambda f(h_{t+1}) + \varepsilon_{t+1}$

where $Y_{t+1} = \frac{S(t+1)}{S(t)} + \frac{D(t+1)}{S(t)} - R_t$, $X_{t+1} = \frac{F(t+1) - F(t)}{S(t)}$

It is assumed that $b_3 = \mu = 0$.

H_1: Unrestricted Model;

(4-15) $Y_{t+1} = b_1 + b_2 X_{t+1} + b_3 Y_t + \lambda f(h_{t+1}) + \mu\varepsilon_t + \varepsilon_{t+1}$

where $Y_t = \frac{S(t)}{S(t-1)} + \frac{D(t)}{S(t-1)} - R_{t-1}$=lagged dependent variable. The moving average term [MA (1)] is added as an additional regressor to consider additional forecasting error causing from the difference of time reported between the spot and futures prices. As noted by Bollerslev, Engle, and Wooldridge (1988), the next test considers the lagged excess holding yield as explanatory variable for the risk premia because the lagged dependent variable helps forecast returns. Agents may use information in addition to past innovations in forming their expectations. Tests may be performed to determine whether or not lagged dependent variable and the moving average term with lag one belong in (4-14) as additional regressors. If they belong in the equation (4-14), then the lagged dependent variable and MA(1) bear important information and the equation (4-14) should be respecified. The equation (4-15) with the lagged dependent variable, the MA(1), and the conditional variance with GARCH (or ARCH) specification is called ARMA with GARCH(or ARCH). The LM tests are applied to test whether or not these additional variables belong to (4-14). If we compare (4-14) with (4-15), the former is a restricted version of the latter. The restricted regression assumes that coefficients of lagged dependent variable and MA term are equal to zero. To test this, the LM test proceeds as follows:

(Step 1) Estimate the restricted regression (4-14) by OLS and obtain the residuals, $\hat{\varepsilon}_{t+1}$.

(Step 2) If the unrestricted regression (4-15) is the true regression, the residuals obtained in (4-14) should be related to constant, lagged dependent variable, and MA term.

(Step 3) This suggests that we regress the $\hat{\varepsilon}_{t+1}$ obtained in step 1 on all the regressors, which means (4-16)

(4-16) $\hat{\varepsilon}_{t+1} = \beta_1 + \beta_2 X_{t+1} + \beta_3 Y_t + \lambda f(h_{t+1}) + \mu\varepsilon_t + u_{t+1}$

where u_{t+1} is an error term with the usual properties. (Step 4) For large- sample size, Engle has shown that T (the sample size) times the R^2 estimated from the regression (4-16) follows the chi-square distribution with df equal to the number of restrictions imposed by the restricted regression. Symbolically, we write

(4-17) $TR^2 \sim \chi^2_{(restrictions\#)}$

(Step 5) If the chi - square value obtained from (4-17) exceeds the critical chi - square value at the chosen level of significance, we reject the restricted regression. Otherwise, we do not reject it.

Next is to test the specification of the conditional variance equation. Bollerslev (1986) suggests a LM statistic which is surprisingly simple to compute. The LM test for

ARCH(q) against GARCH(q,p) can be carried out by referring T times the R^2 to the chi-squared critical value with q degrees of freedom. The statistic is computed as T R^2 of the first step OLS iteration of the BHHH algorithm for the general model with initial parameter values given by estimates under the null. This test statistic is used to test GARCH (1,2) and GARCH(2,1) specifications as alternatives against GARCH(1,1) specification. Note that the test for ARCH(q) against GARCH(q,p) is exactly the same as that for ARCH(q) against ARCH(q+p). For example, let the null hypothesis be ARCH(1) against the alternative hypothesis GARCH(1,1), which has one additional parameter. To test the above hypothesis, ARCH(1) is estimated first. Then, setting the parameter of the past conditional variance at zero, we estimate GARCH(1,1). The next step is to calculate TR^2, where the value of R^2 is simply obtained from the first iteration of BHHH of the unrestricted model (H_1) setting the additional parameters equal to zero. Then, TR^2 is asymptotically distributed as a chi-square with one degree of freedom (the degrees of freedom are the difference between the number of parameters for the null and the alternative hypotheses). If the value of TR^2 is larger (smaller) than the critical value of the distribution, then that null hypothesis is rejected (accepted).

4.5 *Hypotheses Tests*

Let the equation (3-46) with GARCH (p,q) stochastic process be specified as

$$\frac{S(t+1)}{S(t)}+\frac{D(t+1)}{S(t)}-R_t=\mu+b[\frac{F(t+1)-F(t)}{S(t)}]+\lambda f(h_{t+1})+\varepsilon_{t+1}\text{(mean equation)}$$

$$h_{t+1}=\alpha_0+\sum_{i=1}^{q}\alpha_i\varepsilon_{t+1-i}^2+\sum_{j=1}^{p}\beta_j h_{t+1-j} \qquad \text{(variance equation)}$$

$$\varepsilon_{t+1}|\Psi_t \sim N(0,h_{t+1})$$

where the first mean equation is an GARCH in mean (or ARCH in mean) model in which the conditional variance is included in the mean equation as an exogenous variable. The ARCH in mean (ARCH-M) model is originally presented by Engle et al. (1987). The variance equation parameterizes the conditional variance as a function of previous forecast errors and previous conditional variances and is defined as the ARCH and GARCH process, respectively. The third equation means that the conditional one step forecasting error, given available information set Ψ_t, is normally distributed with conditional mean zero and conditional variance h_{t+1}.

The coefficients to be estimated are b (the hedge ratio), λ (effects of TVRP). The coefficients of the GARCH process are $\alpha_0, \alpha_i, and \beta_j$. The null hypotheses are

a) If H_0: b=0, then futures contracts are not effective hedge instruments to avoid the underlying spot price risk.

If H_0: b=1, then the amounts of the underlying futures contracts held will be the same as the amounts of the spot position.

b) If $H_0: \alpha_i = \beta_j = 0$, then no ARCH (and no GARCH);

if there is no ARCH(and GARCH) effect (means nonstationary variance processes), then α_i $and \beta_j$ estimated in h_{t+1} will be statistically insignificant. $\alpha_i = \beta_j = 0$ for all I and j corresponds to the OLS model. Engle's ARCH(q) process is also a special case with $\beta_j = 0$, for all j.

c) If $H_0 : \lambda = 0$, then no TVRP;

if the estimated λ is statistically insignificant under the presence of GARCH disturbances, then the TVRP may not affect the normalized spot prices(the underlying spot market). If $\lambda > 0$, the expected risk premium is proportional to the variance of stock index returns. Positive λ s also means that the risk premia rise when the spot market's volatility rises since risk averse traders require more compensation to get into the increasingly volatile markets.

Chapter V

EMPIRICAL RESULTS

5.1 Data Description

The daily data of spot prices, dividend rates, and three-month Gensaki rates for the period from Jan. 4th, 1991 to Dec. 29th, 1995 are obtained from the Samsung Economic Research Institute supplied by the Data Stream. The futures contract on the Nikkei Stock Index are from various issues of the London Financial Times and the Knight Ridder. The data consist of daily closing prices for the underlying stock index and daily settlement prices for all futures contract.

The first Japanese stock index contract appeared in an "offshore" market. Futures contracts based on the Nikkei 225 Stock Index began trading on the Singapore Monetary Exchange (SIMEX) on September 3rd, 1986. The Nikkei 225 stock index is a simple, price weighted index similar to the Dow Jones Industrial Average: The sum of the prices of the component shares is scaled and adjusted for stock splits with a divisor factor. The Nikkei index includes 225 of the largest Japanese firms on the First Section of the Tokyo Stock Exchange. The futures contract trades on weekdays and some Saturdays. The contract has four maturities (March, June, September, and December). The last trading day is the third Wednesday of the contract month. The settlement of the contract is in cash, based on the closing value of the Nikkei Stock Index Average.

The dividend rate* of the 225 stocks in the index are obtained by dividing the aggregate cash dividends (based on the latest known annual rate) by the aggregate market value of the index. The actual dividend payments are announced after the ex-dividend dates. The actual dividends paid over the life of the futures contract are used as a proxy for the expected dividend to be paid on the Nikkei 225 stocks on a per contract per day basis. That is, it is assumed that dividends and ex-dividend dates over the life of the contract are known at time t.

The interest rate we use is the three-month Gensaki rate (a repurchase rate) which is similar to the short-term Treasury bill rate in the U.S. The Gensaki rate is very close to the risk-free rate and is suitable for use here. These short-term bills which are traded in an over-the-counter market are very actively traded, highly liquid, and largely unrestricted by the Central Bank. The securities are issued by the municipalities and backed by the government.

The daily data is used due to the effect of marking to market, which is a daily settlement imposed by the SIMEX. The specification of TVRP is based on one day trading. It is assumed that the available information over the weekend is the same as that of a weekday, weekend effect will not change the information set Ψ_t because the same information set Ψ_t is required for all t for the estimation of ARCH. The trading date of futures and the interest rates are adjusted to active trading days of spot rates and the values of the day before are substituted for the non-traded day.

* This model assumes the dividend yield is a constant flow. Unlike American corporations, Japanese companies generally pay dividends twice a year (occasionally once). These payments are concentrated at the end of March and Sept. However, the impact on futures prices is not significant because Japanese dividend yields are very small.

The dependent and the independent variables of an GARCH-mean model is

$$\frac{S(t+1)}{S(t)}+\frac{D(t+1)}{S(t)}-R_t=\frac{S(t+1)}{S(t)}+\frac{S(t)*D}{S(t)}-(1+R)$$

$$=\frac{S(t+1)}{S(t)}+[d*(90/360)]-[1+r*(90/360)] \text{ and } \frac{F(t+1)-F(t)}{S(t)}, \text{ respectively.}$$

where S(t) = spot price at time t, S(t+1) = spot price at time t+1 [=F(t+1)], F(t) = futures price at time t for a futures contract that expires at time t+1, r = 3 month Gensaki interest rate, and d = dividend rate.

The trading day starts on January 4th, 1991 at the time subscript t=1, and ends on December 29th, 1995 at t = 1301. The actual sample size of independent and dependent variables is 1300 (T=1300) since one of the observations is the difference between the two. To avoid problems of thin markets and convergence near the maturity of the contract dates, the daily futures rates with intervals of three months are collected and the maturity months of the contracts are ignored. For example, there are a March and June futures contract on March 1st. In this case, a March futures contract is ignored to avoid problems of thin market and convergence. The data from March, April, and May are used for a June futures contract and June, July, and August for a September futures contract, etc.

Time Series Properties of Spot and Futures Returns

(Table 1) is a glossary and definition of variables. The spot and futures returns are based on the endogenous and exogenous variables [see chapter 3.5 and 5.1]. All unit root tests are based on the spot and futures prices, while the other statistics are based on the endogenous and exogenous variables for the differenced spot and futures prices. The daily excess holding period returns to the Nikkei 225 stock index, which is expressed as the dividend inclusive spot price, [S(t+1) + D(t+1)] / S(t), minus one plus the risk-free interest rate (3-month Gensaki rate) , is used for the spot returns.

Several tests on the time series of the Nikkei 225 spot and futures returns were performed and results are reported in Tables 2 - 4. Figure (2) and (4) give plots of the Nikkei 225 stock index level and the Nikkei 225 stock index futures level. Both levels fell dramatically and stood at about 15,000 by the end of the second quarter of 1995. In return series [Figure (3) and Figure (5)] the clustering of fluctuations is a little apparent: Large changes tend to be followed by large changes, and small by small. This is a typical phenomenon for many financial data sets. Chan and Karolyi (1991) show that intraday returns on the S&P 500 stock index and the S&P 500 stock index futures indicate negative skewness and high kurtosis. Both stock and futures returns series also showed empirical distributions with heavy tails and sharp peaks at the center compared to the normal distribution. In other words, both series were leptokurtic. The descriptive statistics (Table 2) include the sample size, mean, standard deviation, skewness, kurtosis, t-value, and Bera-Jarque test for the stock index and futures returns. T-statistics relating to whether the mean returns are significantly different from zero are also included. (Table 2) also shows that the unconditional distributions of the Nikkei stock index futures returns series are non-normal, as evidenced by high skewness, high kurtosis, and significant Bera-Jarque statistics. Interestingly, the futures returns show high kurtosis compared to the spot returns. Kurtosis values suggest that the distribution of the Nikkei stock index futures contracts are very leptokurtic compared to the distribution of the Nikkei stock index. Zero excess kurtosis is rejected for both series. We also test for the normality of the disturbances and use the Bera-Jarque (1982) test statistic. Based on the B-J statistic we reject normality because the B-J statistics are greater than the critical values for χ^2 statistics [9.21(1%) and 5.99(5%), respectively].

The autocorrelation coefficients up to the 4th order [from the lag 5 through 325, the averages are computed] are computed for the spot and futures returns. Mackinlay and Ramaswamy (1988) found evidence of positive autocorrelation in the S&P 500 spot index price change series ranging from 0.038-0.41 but not the S&P 500 futures price change series. On the other hand, Lim (1991) did not find positive autocorrelation in the Nikkei 225 spot index and futures price change series. It was observed that the autocorrelation for both the Nikkei 225 spot index and futures price changes were not significantly different from zero. In this study, there is no autocorrlations in the stock index and futures returns based on the log-diffrenced prices. However, (Table 3) shows evidence of positive autocorrelations from my definition of stock index returns and futures returns because of big scale. All autocorrelation coefficients for the spot and futures returns are statistically significant, that is, significantly different from zero. All autocorrelation coefficients up to lag 164 are outside the 95% confidence bounds. The autocorrelation in the stock index and futures returns do diminish over time. To more investigate the hypothesis that all the autocorrelation coefficients are simultaneously equal to zero, the Ljung-Box (1978) statistic was used. Table (3) also shows that there is strong evidence of autocorrelation in the spot and futures returns. This suggests that returns on the Nikkei stock index and futures are predictable based on past information about the returns.

To estimate conditional variance and a hedge ratio, we need to test whether or not the endogenous and exogenous variables (spot and futures returns) are stationary. The unit root tests are based on the spot and futures prices. Endogenous and exogenous variables are used for the differenced spot and futures prices. Pseudo t values which determine whether each series is a nonstationary I(1) process are presented in (Table 4). Dickey and Fuller (1981) and Augmented Dickey-Fuller unit root tests are employed to test the levels and differences of the series. It is assumed that there is no time trend present in the data generation process. All pseudo t values of the level and difference series are negative. The pseudo t values of the spot and futures returns in the level series are well above (in absolute value) the DF and ADF critical values computed by Mackinnon of -3.5073 (1%), -2.8951 (5%), and -2.5844 (10%), respectively, so that we can not reject the null hypothesis that the given time series is stationary ($b = 0$). This implies that the spot and futures prices time series are stationary. When ADF tests are applied to the difference series, the pseudo t values can not reject the null hypothesis that the given time series is stationary ($\dot{b} = 0$). This shows that the difference series are also stationary. To examine the residuals Ramsey (1969) RESET (regression specification test) is used for model specification errors. For example, the augmented regressions of spot price in the level produce [see Table 4, B,(a)]

$$\Delta S_t = 5.14e-002 - 3.28e-006 S_{t-1} + 0.469 \Delta S_{t-1}$$

(12.12) (-14.11) (19.14) Rsq=0.7630

and

$$\Delta S_t = -0.0077 + 0.6947 \Delta S_{t-1}$$
(-10.42) (34.88) Rsq=0.7266

The observations are 1299. Based on Ramsey's RESET test, F statistic = (0.7630-0.7266)/1/(1-0.7266)/(1299-3) = 182.17. The F statistic is larger than the standard critical value of 3.84(5%) and the Dickey-Fuller value of 4.59(5%). Therefore, once again, we can not reject the null hypothesis and may conclude that the spot and futures prices are stationary.

5.2 ARCH Tests

Engle's LM test is used to detect whether the residuals from a regression model show heteroskedasticity over time. The results of ARCH tests for daily excess holding period returns to the Nikkei 225 stock index are shown in (Table 5) and (Table 6). To show the possibility of the presence of higher-order ARCH, we performed the following test for q=1, 2, 3, and 4. For the whole period of (Table 5) and (Table 6), ARCH (q) is significant up to q=4. Note that the values of χ^2 increase as we go from q=1 to q=2, 3, and 4. For the sub periods of (Table 5), in case of 1992 and 1993, χ^2 statistics are very significant. After 1994, χ^2 statistics is very small. This implies that clearly the periods of 1992 through 1993 shows substantially more volatility than periods of 1994 and 1995. For the sub periods of (Table 6), except in case of 1994, χ^2 statistics are very significant. Overall, we may conclude that if ARCH effect is present, it is of order 1. To get higher values of ARCH process, we estimated ARCH(1), ARCH(2), ARCH(3), ARCH(4), and ARCH(5) models. But only ARCH(1) model for the whole period have the highest value and proved to be significant. For example, if ARCH(5) selected, weighted sum of past squared errors, which is $\sum_{i=1,5} w_i = \sum_{i=1,5} (6-i)/15$, will be used. One can discount older innovations in this weighting scheme. When the length of the lags is selected as 1 (q=1), the procedure is as follows: (Step 1) Regression of Y_{t+1} on a X_{t+1} for 1991-1995 from (Table 6)

(Step 2) Regression of the squared residual on a constant and the lagged squared residual: $\varepsilon_{t+1}^2 = 0.00017 + 0.4884\varepsilon_t^2, R^2 = 0.2387, T = 1299$

We see that $TR^2 = (1299)(0.2387) = 310.07$, which is approximately χ_1^2 with 1 df. The critical value of the chi-square distribution with one degree of freedom (χ_1^2) given type I error at 0.01(1%) are 6.63. Therefore, the null hypothesis: $\alpha_i = 0$ [for all i = 1...q] is rejected. This indicates the presence of an ARCH process. Thus, we reached the conclusion of strong evidence of heteroscedasticity in the error variance on the basis of the ARCH test.

5.3 Empirical Results for Hedging model

The excess holding spot returns, Y_{t+1} was calculated as:

$$Y_{t+1} = \frac{S(t+1)}{S(t)} + \frac{D(t+1)}{S(t)} - R_t = \frac{S(t+1)}{S(t)} + d - (1+r)$$

where S(t) and S(t+1) are the spot price at time t and t+1, d is dividend rate, and r is 3 month Gensaki interest rate.

(5-1) $Y_{t+1} = -0.0254 + \varepsilon_{t+1}$
(-33.96)

(Table 7) presents the mean and standard deviation of stock index returns for the whole period and for the sub-periods. The mean of the excess holding spot returns over the sample is -0.0254 percent at annual rates while the standard deviation is 0.0269 at annual rates. Bollerslev, Engle, and Wooldridge (1988) showed that the large negative intercepts for stocks (using the returns on the value-weighted NYSE index including dividend) and bonds are not surprising because reduced capital gains taxes on long term assets provide incentives to hold these assets even at otherwise unfavorable rates of return. (Table 7) shows negative excess spot returns and they are significantly different from zero. Spot returns are real return on the stock between time t and time t+1 from the transaction of spot market. So, they can be negative because real return is adjusted for inflation. During 1993 the Nikkei stock index dropped from 18,863.16 to 16,798.94, giving 11 percent decline. If measured in real terms, the decline is even more serious. In the return series, the clustering of fluctuations is apparent. This suggests that not only do the conditional mean excess holding spot returns vary over time, but also the conditional variances seem to be changing over time.

(Table 8) shows ARCH model for daily excess holding period returns to the Nikkei 225 stock index. From the ARCH tests we select the first-order of ARCH model. ARCH (1) gives:

(5-2) $Y_{t+1} = -0.0176 + \varepsilon_{t+1}$
(-25.65)

$$h_{t+1} = Var(\varepsilon_{t+1}) = \hat{\varepsilon}_{t+1}^2 = \underset{(16.60)}{0.00028} + \underset{(14.96)}{0.6571}\varepsilon_t^2 \qquad L = 3,000.11$$

The ARCH effect is very strong, showing a t statistic of 14.96. Thus, there is a strong relation between recent squared errors and the estimate of volatility. The ARCH process shows stationarity since the sum of the estimated parameters of conditional variance is less than one; the ARCH processes will be stationary if $\sum_{i=1}^{q} \alpha_i < 1$ [see Engle (1982)]. In this study, the ARCH processes in (5-2) satisfy the above stationarity conditions since $\alpha_1 = 0.6571 < 1$.

Second, to represent a series with changing volatility, we use a generalized ARCH model. Empirical studies frequently suggest that a GARCH(1,1) is adequate in modeling conditional variance, e.g. Bollerslev (1986, 1987), Engle and Bollerslev (1986). GARCH (1,1) for the whole period gives:

(5-3) $$Y_{t+1} = \underset{(-23.85)}{-0.0153} + \varepsilon_{t+1}$$

$$h_{t+1} = Var(\varepsilon_{t+1}) = \hat{\varepsilon}_{t+1}^2 = \underset{(2.71)}{5.65e-6} + \underset{(3.62)}{0.0957}\varepsilon_t^2 + \underset{(31.94)}{0.8945}h_t$$

$$L = 3,137.19$$

Engle and Bollerslev (1986) shows that GARCH (p,q) process is weakly stationary since the mean, variance, and autocovariance are finite and constant over time if $\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$. The estimates of the GARCH model in (Table 9) indicates that there is a strong evidence for the presence of integrated GARCH effects where $\alpha_1 + \beta_1$ is close to unity. If unity is the true value of this parameter, then shocks to volatility persist forever. In this study, $\alpha_1 + \beta_1$ for the whole period equals 0.9902 which is close to unity. This indicates persistence of a forecast of the conditional variance over all future horizons, and also implies an infinite variance for the unconditional distribution of ε_t. In other words, shocks to the stock index returns volatility persist forever.

Antoniou and Holmes (1995) explained that α_1 reflects the impact of recent news. α_1 is the coefficient relating to the lagged squared error term. The lagged error term relates to changes in the spot price (or futures price) on the previous day which are attributable to market-specific factors. Assuming that markets are efficient, then these price changes are due to the arrival in the market of items of information. Hence, α_1 relates to the impact of yesterday's market-specific price changes on price changes today. Given that this relates to the arrival of information yesterday, α_1 can thus be viewed as a "news" coefficient, with a higher value implying that recent news has a greater impact on price changes. β_1 is the coefficient on the lagged variance term and as such is picking up the impact of price changes relating to days prior to the previous day and thus to news which arrived before yesterday. Thus, β_1 can be thought of as reflecting the impact of "old news". This old news will have less impact on today's price changes. That' why we discount older innovations. In sum, (Table 8) and (Table 9) indicates that there is the time varying risk premium in the variance equation and represents misspecification. The hypothesized true model is GARCH (1,1)-M model:

(5-4) $Y_{t=1} = \mu + \lambda h_{t+1}^{1/2} + \varepsilon_{t+1}$

$$h_{t+1} = \sigma_{t+1}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \sigma_t^2$$

$$\varepsilon_{t+1} | \Psi_t \sim N(0, h_{t+1})$$

The functional form of the risk premium is defined as $f(h_{t+1}) = h_{t+1}^{1/2}$. From the (Table 10), estimating the model (5-3) with h_{t+1}, $\log(h_{t+1})$, $\exp(h_{t+1})$, and $h^2{}_{t+1}$ produced t statistic and values of log-likelihoods. The one-second power of the conditional variance is selected since this form gives the highest values of log likelihoods from (Table 10) and it yields the estimation results. The maximum likelihood estimates and their t statistics for the whole period are shown in (Table 11):

(5-5) $Y_{t=1} = -0.0070 - 0.4973 h_{t+1}^{1/2} + \varepsilon_{t+1}$
(-7.09) (-9.63)

$$h_{t+1} = 4.81e-6 + 0.1015\varepsilon_t^2 + 0.8918\sigma_t^2$$
(2.69) (3.425) (30.44)

L = 3,186.79

The above estimation results indicate strong evidence (in terms of t- statistics) of both ARCH (and GARCH) effect ($\alpha_1 = \beta_1 \neq 0$) and time varying risk premium ($\lambda \neq 0$). The sign of the estimated coefficients of the conditional variances, λ's, are negative mean-variance ratio for the whole period and for the sub period of 1995 in the GARCH- M model. This suggests that when the conditional volatility rises, the spot returns (return from investment in the spot market) decrease. For the period up to 1990, the Nikkei index had a spectacular run-up and increased to over 38,000. However, for the period post 1990, the Nikkei fell dramatically and stood at about 15,000 by the end of the second quarter of 1995. The Japanese stock market is going back to the efficient market from the bubble market. In the mean time, risk-averse investors are reluctant to enter the market due to downward stock market, resulting in a reduction in the search for information. This could lead to pricing inefficiency because information affect the price. Unexploited speculative activities in turn led to abnormal returns. While the question of efficiency is not major concerns in this research, this findings does suggest that the efficiency of the market for the Nikkei 225 stock index needs to addressed. The one spike in (Figure 6) corresponds to period of Feb., 24 1993 Nikkei 225 stock index collapse (from 18863.16 to 16,798.94).

Now, we turn to a more careful analysis of this hedging model. The futures prices are an important determinant of the excess returns on a stock index. That's why we apply this futures price to (5-4). First, we use the hedge model to estimate the optimal hedge ratio. The results in (Table 12) & (Table13) are based on a homoscedastic (or OLS) model (no presence of the ARCH process: $h_{t+1} = \alpha_0$) and a heteroscedastic model (presence of either ARCH or GARCH processes), respectively. The results of the OLS regression for the whole period were shown in (Table 12):

(5-6) $$Y_{t+1} = -0.0169 + 1.2656 X_{t+1} + \varepsilon_{t+1}$$

$$t = (-30.13) \quad (38.62) \qquad\qquad R^2 = 0.5347, \quad \text{D.W} = 0.4044$$

The results reveal that the estimated slope coefficient is significant at the 5% level on the basis of t test and the spot returns is linearly related to the futures returns. The corresponding coefficient and t statistic in Figlewski (1986) for the OLS technique [from the equation (3-44)] to estimate the optimal hedge ratio are 0.7957 and 12.63. The above results (5-6) also show that there is evidence of positive autocorrelation for the whole period because Durbin-Watson d statistic is close to 0. The estimated hedge ratio b is 1.2656. This indicates that each dollar of the spot market position should be hedged with $1.2656 dollars in the futures position by selling $1.2656 in the futures market. Thus, a bigger futures position is required to hedge the spot price risk. The OLS results from (Table 12) show that the White test indicates the presence of strong unconditional heteroskedasticity for the whole period and for the sub periods. The normality test statistics for the whole period and for the sub period are significant. In the presence of nonnormality and heteroskedasticity, the estimates of the OLS hedge ratios (between 0.72 and 1.27) are inefficient and the t-statistics are unreliable.

Adding the futures returns to GARCH(1,1)-M model (5-4) gives:

(5-7) $$Y_{t+1} = \mu + bX_{t+1} + \lambda h_{t+1}^{1/2} + \varepsilon_{t+1}$$

$$h_{t+1} = \sigma_{t+1}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \sigma_t^2$$

$$\varepsilon_{t+1} | \Psi_t \sim N(0, h_{t+1})$$

where $Y_{t+1} = \dfrac{S(t+1)}{S(t)} + \dfrac{D(t+1)}{S(t)} - R_t$, $X_{t+1} = [\dfrac{F(t+1) - F(t)}{S(t)}]$. The maximum likelihood estimates and their t statistics for the whole period are shown in (Table 13):

(5-8) $$Y_{t+1} = -0.0053 + 1.0959X_{t+1} - 0.4359h_{t+1}^{1/2} + \varepsilon_{t+1}$$
(-11.18) (45.82) (-11.12)

$$h_{t+1} = 2.05e-6 + 0.2513\varepsilon_t^2 + 0.7719\sigma_t^2$$
$$(2.71)\ (6.25)\quad (23.04)$$

L = 3,790.99

(Table 13) is the final preferred model including the futures prices. In this model all coefficients are significant and the log likelihood is substantially above that of (5-4). The estimation results (Table 13) indicate strong evidence (in terms of t- statistics) of both ARCH (and GARCH) effect ($\alpha_1 \neq 0\ \beta_1 \neq 0$) and time varying risk premium ($\lambda \neq 0$) where the hedging model based on (5-7) is applied. This implies that an investor's risk premium changes over time. $\alpha_1 + \beta_1$ for the whole period and for the sub-periods of 1991 and 1993 shows that there is strong evidence for the presence of IGARCH effects. This implies the persistence of shocks to the Nikkei 225 stock index returns volatility. The estimated coefficients of the b's, which indicate the hedge ratios, are statistically more significant (between 0.72 and 1.10) for heteroscedastic model. The null hypothesis, $H_0{:}b = 0$ (futures contracts are not effective hedge instruments to avoid the underlying spot price risk) is rejected. The estimated GARCH-M hedge ratio b for the whole period is 1.10. This indicates that each dollar of the spot market position should be hedged with \$1.10 dollars in the futures position by selling \$1.10 in the futures market. Thus, a smaller futures position is required to hedge the spot price risk, compared to OLS hedge ratios (1.27). The hypothesis, $H_0{:}b = 1$ (futures contracts are perfect hedges against spots), is accepted for the sub-periods of 1992 and 1995 under the OLS and GARCH-M specifications.

If $\lambda > 0$, the expected risk premium is proportional to the variance of stock returns. In this study, however, the sign of the estimated coefficients of the conditional variances, λ's, are negative mean-variance ratio for the whole period and for most sub period in the heteroscedastic model, except the subperiod of 1991. This results is somewhat puzzling because it is expected the risk premia (or expected returns) will go up when the market's volatility rises since risk averse investors ask for more compensation to get into volatile market. This result may be explained by the following. The Nikkei index fell dramatically and stood at about 15,000 in 1995 from 38,000 in 1990. The Japanese stock market is going back to the efficient market from the bubble market. In the mean time, risk averse investors are reluctant to enter the market due to downward stock market, resulting in a reduction in search for information. This could lead to pricing inefficiency because the information affect the price. Unexploited speculative activities in turn led to abnormal returns. While the question of efficiency is not major concerns in this research, this findings does suggest that the efficiency of the market for the Nikkei 225 stock index needs to addressed. In addition to that, speculative activities dominate the market outcomes because the futures prices are considered and the futures markets are very speculative, even though investors are risk-averse investors. This unexploited arbitrage opportunities in turn led to abnormal returns. French, Schwert, and Stambaugh (1987) and Chou (1987) found a statistically significant positive relationship between conditional variance and excess returns on the NYSE stock index and the S&P composite portfolio using GARCH-M model. My results contradict their results and show statistically significant negative relationship between conditional volatility and excess returns on the Nikkei 225 stock index using GARCH-M model. Some studies question the existence of a positive mean-variance ratio. Abel (1988) found that the mean-variance ratio is not necessarily positive when the investor's preference is not logarithmic. Glosten, Jagannathan, and Runkle (1989) obtain a negative ARCH-M parameter when they include the nominal risk-free rate in the conditioning information set. The findings of this paper also support that an increase in riskiness can either increase or decrease stock index returns. (Figure 7) shows the estimated risk premia (conditional variance) from the GARCH-M model in (Table 13). The one spike in (Figure 7) corresponds to period of Feb.,1993 Nikkei index collapse.

5.4 Specification Tests

The LM test is applied to test the specification of the mean and conditional variance equations. (Table 14) and (Table 15) shows the results of TR^2 for the specification tests of the mean equation and the variance equation (GARCH processes) for the hedging model (5-7). The level of type I error is chosen at .01(1%). The critical values of chi square distribution with the degree of freedom one is 6.63 at the level of type I error =.01. The first one is to test for the specification of the mean equation. The lagged dependent variable is added as an additional regressor to find the volatility of spot returns. The moving average term [MA (1)] is added as an additional regressor to consider additional forecasting error caused from the difference of time reported between the spot and futures prices. From the LM test, TR^2 of the moving average with one lag and the one-lagged dependent variable (normalized spot returns) in the mean equations are also significant. The values of the t statistics (under the null = 0) with MA(1) and lagged dependent variable are presented in (Table 14). The null hypotheses was also rejected from the t tests. These results indicate that the basic model (5-7) requires modification to empirically represent the actual data.

Next is to test the specification of the conditional variance equation. The results of the LM tests for the variance equation are given in (Table 15). The results show that the value of TR^2 for GARCH(q=1,p=2) is significant at the 1% level. Then the null hypothesis of GARCH(1,1) is rejected. Therefore, this result indicates that GARCH(1,1) has the problem of misspecification. However, the value of TR^2 for GARCH(q=2,p=1) is not significant. In this case, the null hypothesis of GARCH(1,1) cannot be rejected.

Chapter VI

SUMMARY AND CONCLUSIONS

Two important questions are asked in this research. The first is to find whether or not an investor's risk attitude (or risk premium) changes over time. The second is to investigate how the market's volatility affects a investor's expected returns and to investigate the existence of a negative GARCH-M parameter (or mean -variance ratio).

To investigate the above questions, we derive TVRP which is based on the rational expectation hypothesis. The intertemporal asset pricing model was modified to allow for risk aversion. A simple index arbitrage in the stock index futures market is employed for the equilibrium value of intertemporal asset pricing model which was developed by Lucas (1978). We then apply the hedging and GARCH-M model to the Nikkei 225 stock index futures market. The conditional variances are utilized to directly test for the presence of time varying risk premia. TVRP can measure the degree of riskiness in the volatile stock markets. The changing conditional variances over time can be one of the factors which contributes to the volatility of financial markets. Generalized Autoregressive Conditional Heteroscedasticity in mean (GARCH-M) models of excess returns on the Nikkei 225 stock index are used to estimate the time varying conditional variances and a mean-variance ratio that represents the risk-return trade off. In the ARCH model, the nature of heteroscedastic error terms over time is captured by the conditional variance, which depend upon past available information. Thus, the conditional variance is useful variable for measuring uncertainty and provides a model of TVRP.

The empirical results in Chapter 5 show that there was strong evidence for the time varying risk premia, as reflected in heteroscedastic error terms over time in the hedging model using data on spot and futures prices of the Nikkei 225. This suggests that an investor's risk premium changes over time. Another finding is the persistence of shocks to the Nikkei 225 stock index returns volatility. Naka (1989) showed that the returns on the spot currency exchange markets using five major currencies can be positive or negative when the TVRP increases. Abel (1988) claims that in a general equilibrium the mean-variance relationship is not necessarily positive when the investor's preference is not logarithmic. Glosten, Jagannathan, and Runkle (1989) also describes the asymmetry in the conditional volatility of Japanese stock returns and obtains a negative ARCH-M parameter when they include the nominal risk-free rate in the conditioning information set. My results also show a statistically significant negative relationship between conditional volatility and excess returns on the Nikkei 225 stock index using GARCH-M model. The findings of this paper support that an increase in riskiness can either increase or decrease spot returns.

The specification tests in section 5.4 demonstrate that the hedging model of equation 5.7 needs to be improved to reflect the actual data used. The expected return on the stock index is influenced by its own lagged observations. Therefore, the lagged dependent variables contain important information which the theoretically developed models in chapter 3 cannot capture, and the models can be improved by adding the lagged dependent variables and MA(1). The GARCH (1,1) process is rejected in favor of the GARCH(1,2) in the hedging model.

Although the empirical results support the main hypotheses based on theoretically developed models, the specification tests suggest that better models can be constructed for the hedging model. An extension of this research can be summarized as follows: 1) estimate the hedging model including all exogenous variables (macroeconomic variables such as industrial production and changes in interest rates) which belong to the theoretically developed models; 2) introduce 2SLS model; 3) create a CAPM model based on different assumptions and apply other ARCH models; 4) collect additional data to expand the available information set.

(Table 1)

Glossary and Definitions of Variables

Symbol	Variable	Definition
	Basic Data Series	
S(t+1),S(t)	Spot Price	daily closing prices for the Nikkei 225 stock index
F(t+1), F(t)	Futures Price	daily settlement prices for the Nikkei stock index futures
D	Dividend yield	dividend rate obtained by dividing aggregate cash dividends by the aggregate market value of the index
R	Risk-free interest rate	3 month Gensaki rate (a repurchase rate) which is similar to the short-term Treasury bill in the U.S.
y_{t+1}	Stock index returns	the daily excess holding period returns to the Nikkei 225 stock index or the Nikkei stock index returns minus one plus the risk-free rate
x_{t+1}	Futures returns	difference of futures prices between time t+1 and t divided by spot price at time t
$\sum_i w_i$	Weighted sum of past squared errors	
	Figures	
X		Nikkei 225 stock index futures returns
Y		Nikkei 225 stock index returns
H1		Predicted risk premium to the Nikkei 225 stock index from the daily GARCH-M model in [Table 11] for the conditional variance
H2		Predicted risk premium to the Nikkei 225 stock index from the daily GARCH-M model in [Table 13] for the conditional variance

(Table 2)

Descriptive Statistics*

	Spot Returns (or Endogenous Var.)	Futures Returns (or Exogenous Var.)
Sample Period	1991-1995	1991-1995
Sample size	1300	1300
Mean	-0.0254	-0.0067
(t-statistic)	(-33.96)	(-15.53)
Std. Dev.	0.0269	0.0156
Skewness	-0.37	0.62
Kurtosis	0.11	4.85
Normality Test (Bera-Jarque Test)	482.58**	268.80**

* Returns are based on the endogenous and exogenous variables [see 5.1]. They are adjusted to a annual returns.

** Significant at the 1% and 5% level. B-J is distributed χ^2 under the null of normality. The critical values for χ^2 statistics are 9.21 and 5.99, respectively.

(Table 3)

Autocorrelations of Spot and Futures Returns

	Spot Returns (or Endogenous Var.)	Futures Returns (or Exogenous Var.)
Autocorrelations (ρ_k):		
Lag k		
1	0.6941*	0.2368*
2	0.6694*	0.2292*
3	0.6899*	0.2039*
4	0.6933*	0.1917*
5-36	0.6520*	0.1368*
37-68	0.5921*	0.1452*
69-100	0.5291*	0.1113*
101-132	0.4513*	0.0958*
133-164	0.3687*	0.0747*
165-196	0.3116*	0.0414
197-228	0.2546*	0.0358
229-260	0.2242*	0.0121
261-292	0.1813*	0.0096
293-325	0.1460*	-0.0075
Ljung-Box Q(325)	78,356**	3,942**

* Significant at the 5% level. Asymtotic standard errors for the autocorrelation coefficients can be calculated as the square root of the reciprocal of the number of observations (e.g., ± 0.0277 for 1300 observations) under the null hypothesis of zero autocorrelation. Then, the 95% confidence interval for any ρ_k is ± 1.96(0.0277) = ± 0.0543 on either side of zero. From the lag 5 through lag 325, the averages was computed.

** Significant at the 1% and 5% level. The critical values for χ^2 statistics are 135.807 and 124.342, respectively.

(Table 4)

Stationary Tests: Unit Root Test* on Spot and Futures Prices

H_0: b=0

A. Dickey - Fuller Specification

(a) Single(levels) unit root test

Spot: $\Delta S_t = a + bS_{t-1} + \varepsilon_t$
Futures: $\Delta F_t = a' + b' F_{t-1} + \varepsilon'_t$

b	Pseudo t - statistics	Critical Value(5%)
Nikkei Spot price	-31.0465	-2.8951
Nikkei Futures price	-14.1556	

(b) Two (Differences) unit root tests

Spot: $\Delta^2 S_t = a + b\Delta S_{t-1} + \varepsilon_t$
Futures: $\Delta^2 F_t = a' + b'\Delta F_{t-1} + \varepsilon'_t$

b	Pseudo t - statistics	Critical Value(5%)
Nikkei Spot price	-15.3223	-2.8951
Nikkei Futures price	-28.3413	

* All unit root tests are based on the spot and futures prices. Endogenous and exogenous variables are used for the differenced spot and futures prices.

B. Augmented Dickey - Fuller Specification

H_0: b=0

Sample Period: 1991:01:04 through 1995:12:29

(a) Single (levels) unit root test

Spot: $\Delta S_t = a + bS_{t-1} + c\Delta S_{t-1} + \varepsilon_t$
Futures: $\Delta F_t = a' + b' F_{t-1} + c'\Delta F_{t-1} + \varepsilon_t'$

b	Pseudo t - statistics	Critical Value(5%)	F Statistic*
Nikkei Spot Price	-14.1153	-2.8951	182.17
Nikkei Futures Price	-11.6438	-2.8951	125.74

(b) Two (Differences) unit root tests

Spot: $\Delta^2 S_t = a + b\Delta S_{t-1} + c\Delta^2 S_{t-1} + \varepsilon_t$
Futures: $\Delta^2 F_t = a' + b' \Delta F_{t-1} + c'\Delta^2 F_{t-1} + \varepsilon_t'$

b	Pseudo t - statistics	Critical Value(5%)	F Statistic
Nikkei Spot Price	-9.7515	-2.8951	89.35
Nikkei Futures Price	-18.5710	-2.8951	264.27

* The critical values for F ratio (standard) and F ratio (Dickey-Fuller) are 3.84 (5%) and 4.59(5%), respectively. Degrees of freedom for the numerator and denominator are 1 and T-3. Ramsey's RESET(regression specification test) test is used for a diagnostic statistics on the residuals.

$$F = \frac{(R^2_{new} - R^2_{old})/\#\,of\ new\ regressor}{(1 - R^2_{new})\,/\,(T - \#\,of\ \text{parameters in the new model})}$$

(Table 5)

ARCH Tests for daily excess holding period returns to the Nikkei 225 stock index

$$Y_{t+1} = \mu + \varepsilon_{t+1}$$

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_i \sum_{i=1}^{q} \varepsilon_{t+1-i}^2$$

	q=1	q=2	q=3	q=4
1991.1-1995.12				
TR^2	139.89*	236.24*	255.37*	270.64*
1991.1-1991.12				
TR^2	0.15	1.20	1.23	1.71
1992.1-1992.12				
TR^2	13.47*	14.08*	14.81*	19.17*
1993.1-1993.12				
TR^2	0.06	50.88*	50.89*	65.82*
1994.1-1994.12				
TR^2	0.16	0.29	4.00	4.26
1995.1-1995.12				
TR^2	1.115	1.119	1.473	1.632
Critical Value(1%)	6.63	9.21	11.34	13.28
Critical Value(5%)	3.84	5.99	7.81	9.49

* Significant at the 1% level. It follows chi-square distribution with degrees of freedom equal to number of the length of the lags (q). ARCH test is Engle's (1982) LM test. For example, ARCH(1) is LM test for 1st order ARCH effects. The following equation is estimated for the ARCH(1): $\hat{\varepsilon}_{t+1}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2$, where $\hat{\varepsilon}_{t+1}^{\,2}$ is the square of the residuals are obtained from the regression estimated based on equation (5-1).

(Table 6)

ARCH Tests in the hedging model for daily excess holding period returns to the Nikkei 225 stock index

$$Y_{t+1} = \mu + bX_{t+1} + \varepsilon_{t+1}$$

$$\sigma^2_{t+1} = \alpha_0 + \alpha_i \sum_{i=1}^{q} \varepsilon^2_{t+1-i}$$

	q=1	q=2	q=3	q=4
1991.1-1995.12				
TR^2	310.07*	314.74*	330.18*	332.48*
1991.1-1991.12				
TR^2	13.92*	22.42*	23.18*	23.87*
1992.1-1992.12				
TR^2	15.92*	19.39*	19.44*	19.57*
1993.1-1993.12				
TR^2	49.09*	56.74*	57.63*	57.86*
1994.1-1994.12				
TR^2	0.41	0.44	2.98	3.19
1995.1-1995.12				
TR^2	40.84*	40.96*	46.08*	48.12*
Critical Value(1%)	6.63	9.21	11.34	13.28
Critical Value(5%)	3.84	5.99	7.81	9.49

* Significant at the 1% level. It follows chi-square distribution with degrees of freedom equal to number of the length of the lags (q). ARCH test is Engle's (1982) LM test. For example, ARCH(1) is LM test for 1st order ARCH effects. The following equation is estimated for the ARCH(1): $\hat{\varepsilon}^2_{t+1} = \alpha_0 + \alpha_1 \varepsilon^2_t$, where $\hat{\varepsilon}_{t+1}{}^2$ is the square of the residuals are obtained from the regression estimated based on equation (5-6).

(Table 7)

Means and Standard Deviations of Stock Index Returns

$$Y_{t+1} = \mu + \varepsilon_{t+1}$$

Sample size	Mean	Std. Dev.	t-Value
Jan.,1991- Dec.,1995:1300	-0.0254	0.0269	-33.96*
Jan.,1991-Dec.,1991: 257	-0.0635	0.0304	-68.73*
Jan.,1992-Dec.,1992: 261	-0.0333	0.0355	-26.41*
Jan.,1993-Dec.,1993: 260	-0.0180	0.0316	-18.01*
Jan.,1994-Dec.,1994: 260	-0.0107	0.0270	-14.75*
Jan.,1995-Dec.,1995: 262	-0.0019	0.0319	-1.869*

* Significant at the 5% level. The critical values for t-statistics are 1.66. t-statistics relate to the hypothesis that mean returns equal zero.

Notes: Returns are adjusted to an annual returns. Mean returns on futures contracts are -0.0067 for the whole period.

(Table 8)
ARCH model for daily excess holding period returns to the Nikkei 225 stock index

$$Y_{t+1} = \mu + \varepsilon_{t+1}$$

$$\sigma^2_{t+1} = \alpha_0 + \alpha_i \sum_{i=1}^{q} w_i \varepsilon^2_{t+1-i}$$

	Coefficients			
	μ	α_0	α_i	L*
Jan.,1991- Dec.,1995:				
	-0.0176	0.00028	0.6571	3,000.11
	(-25.65)	(16.60)	(14.96)	
Jan.,1991-Dec.,1991:				
	-0.0637	0.00021	0.0638	718.75
	(-66.63)	(8.25)	(0.65)	
Jan.,1992-Dec.,1992:				
	-0.0333	0.00031	0.2284	652.93
	(-28.1)	(8.79)	(2.62)	
Jan.,1993-Dec.,1993:				
	-0.0177	0.00019	0.1541**	725.09
	(-19.37)	(10.25)	(2.41)	
Jan.,1994-Dec.,1994:				
	-0.0113	0.00004	0.0969***	822.31
	(-19.46)	(3.69)	(1.32)	
Jan.,1995-Dec.,1995:				
	-0.0016	0.00022	0.2073	706.89
	(-1.586)	(8.36)	(2.15)	

* L stands for log-likelihood. Numbers in parentheses are t-statistics.

** ARCH(2) is applied for the higher values. $\sum_{i=1,2} w_i = \sum_{i=1,2} (3-i)/3$.

*** ARCH(5) is applied for the higher values. $\sum_{i=1,5} w_i = \sum_{i=1,5} (6-i)/15$.

(Table 9)

GARCH (1,1) model for daily excess holding period returns to the Nikkei 225 stock index

$$Y_{t+1} = \mu + \varepsilon_{t+1}$$
$$\sigma^2_{t+1} = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \sigma_t^2$$

	Coefficients					
	μ	α_0	α_1	β_1	$\alpha_1 + \beta_1$	L*
Jan.,1991- Dec.,1995:						
	-0.0153	5.65e-6	0.0957	0.8945	0.9902	3,137.19
	(-23.85)	(2.710)	(3.619)	(31.94)		
Jan.,1991-Dec.,1991:						
	-0.0642	0.00011	0.0976	0.3411	0.4387	708.69
	(-67.75)	(9.707)	(8.958)	(39.52)		
Jan.,1992-Dec.,1992:						
	-0.0333	0.00006	0.1513	0.6759	0.8272	674.49
	(-27.33)	(2.514)	(2.811)	(7.166)		
Jan.,1993-Dec.,1993:						
	-0.0179	0.00025	6.64e-23	9.81e-7	-	705.28
	(-18.05)	(2.452)	(.)	(0.000)		
Jan.,1994-Dec.,1994:						
	-0.0114	4.79e-6	0.1441	0.8202	0.9643	844.61
	(-19.54)	(1.790)	(3.413)	(18.76)		
Jan.,1995-Dec.,1995:						
	-0.0009	0.00018	0.2329	0.1344	0.3673	687.05
	(-0.823)	(3.460)	(2.288)	(0.774)		

* L stands for log-likelihood. Numbers in parentheses are t-statistics.

(Table 10)

Estimating the GARCH (1,1)-M models with other types of conditional variances

$$Y_{t+1} = \mu + \lambda f(h_{t+1}) + \varepsilon_{t+1}\ ^{*}$$

	Coefficients(λ)	Std. Error	t-ratio	Log-likelihood
$h_{t+1}^{1/2}$:	-0.4972	0.0516	-9.630	3,186.79
h_{t+1}:	-0.717	30.442	-1.6235	3,159.19
$h_{t+1}{}^{2}$:	-0.1917	25.87	-0.0074	3,151.50
$\log h_{t+1}$:	-4.82e-003	6.955e-004	-6.9322	3,181.01
$\exp h_{t+1}$:	-0.476	0.421	-1.1283	3,157.75

* $f(h_{t+1})$ is the unknown functional form of the risk premium.

(Table 11)

GARCH(1,1)-M models for daily excess holding period returns to the Nikkei 225 stock index

$$Y_{t+1} = \mu + \lambda h_{t+1}^{1/2} + \varepsilon_{t+1}$$

$$\varepsilon_{t+1}|\Psi_t \sim N(0, h_{t+1})$$

$$h_{t+1} = \sigma^2_{t+1} = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \sigma_t^2$$

		Coefficients			
μ	λ	α_0	α_1	β_1	L*
Jan.,1991- Dec.,1995:					
-0.0070	-0.4973	4.81e-6	0.1015	0.8918	3,186.79
(-7.086)	(-9.63)	(2.687)	(3.425)	(30.44)	
Jan.,1991-Dec.,1991:					
-0.0800	1.0287	1.63e-6	0.0724	0.9294	734.16
(-18.63)	(3.34)	(0.744)	(2.981)	(48.95)	
Jan.,1992-Dec.,1992:					
-0.0378	0.2344	0.00006	0.1417	0.6895	656.71
(-2.224)	(0.286)	(2.302)	(2.339)	(6.508)	
Jan.,1993-Dec.,1993:					
-0.0179	1.14e-6	0.00025	-2.06e-25	9.37e-7	705.28
(.)	(.)	(2.449)	(.)	(.)	
Jan.,1994-Dec.,1994:					
-0.0135	0.2430	5.35e-6	0.1359	0.8236	817.90
(-5.688)	(0.995)	(2.086)	(3.347)	(19.739)	
Jan.,1995-Dec.,1995:					
0.0019	-0.1934	0.00017	0.2272	0.1621	707.17
(0.144)	(-0.254)	(2.344)	(2.278)	(0.664)	

* L stands for log-likelihood. Numbers in parentheses are t-statistics.

(Table 12)

OLS in the hedging models for daily excess holding period returns to the Nikkei 225 stock index

$$Y_{t+1} = \mu + bX_{t+1} + \varepsilon_{t+1}$$

	Coefficients μ	b	DW	Rsq	Normality	White Test
Jan.,1991-Dec.,1995:	-0.0169 (-30.13)	1.2656 (38.62)	0.4044	0.5347	215.46*	195.77**
Jan.,1991-Dec.,1991:	-0.0507 (-42.45)	0.7153 (13.26)	0.9410	0.4082	84.13*	24.23**
Jan.,1992-Dec.,1992:	-0.0262 (-41.78)	0.9799 (30.86)	1.2466	0.7875	71.43*	0.506
Jan.,1993-Dec.,1993:	-0.0153 (-21.03)	0.7452 (15.99)	1.2328	0.4959	16,765.73*	36.18**
Jan.,1994-Dec.,1994:	-0.0084 (-16.13)	0.8853 (16.93)	1.7408	0.5264	148.57*	37.35**
Jan.,1995-Dec.,1995:	0.00053 (1.184)	1.0356 (33.007)	0.5522	0.8073	121.47*	4.60

* Significant at the 1% level (9.21). Bera-Jarque test follows χ^2 with 2 df.

** Significant at the 1% level (9.21). The above white test follows χ^2 with 2 df equal to the number of regressors in the auxiliary regression.

(Table 13)

GARCH(1,1)-M models in the hedging model for daily excess holding period returns to the Nikkei 225 stock index

$$Y_{t+1} = \mu + bX_{t+1} + \lambda h_{t+1}^{1/2} + \varepsilon_{t+1}$$

$$\varepsilon_{t+1} | \Psi_t \sim N(0, h_{t+1})$$

$$h_{t+1} = \sigma^2{}_{t+1} = \alpha_0 + \alpha_1 \varepsilon_t^2 + \beta_1 \sigma_t^2$$

Coefficients μ	b	λ	α_0	α_1	β_1	Normality Test	Log-likelihood
Jan.,1991-Dec.,1995:							
-0.0053	1.0959	-0.4359	2.05e-6	0.2513	0.7719	3,657*	3,790.99
(-11.18)	(45.82)	(-11.12)	(2.71)	(6.25)	(23.04)		
Jan.,1991-Dec.,1991:							
-0.0646	0.7365	1.14257	1.90e-6	0.1296	0.8665	10.17*	826.79
(-37.15)	(16.82)	(7.890)	(1.31)	(4.69)	(27.48)		
Jan.,1992-Dec.,1992:							
-0.0251	0.9817	-0.0761	0.00002	0.2261	0.5367	15.86*	859.29
(-6.937)	(33.68)	(-0.211)	(2.04)	(2.61)	(3.271)		
Jan.,1993-Dec.,1993:							
-0.0152	0.7958	-0.0223	4.25e-6	0.4930	0.5750	93.50*	906.30
(-17.06)	(37.15)	(-0.389)	(2.41)	(4.84)	(5.056)		
Jan.,1994-Dec.,1994:							
-0.0086	0.7240	-0.1041	4.66e-6	0.2464	0.7115	142.67*	910.73
(-5.070)	(16.02)	(-0.430)	(2.67)	(2.49)	(8.620)		
Jan.,1995-Dec.,1995:							
-0.0079	0.9714	-0.7607	3.63e-6	0.2272	0.6970	182.25*	952.34
(-11.18)	(45.82)	(-11.12)	(2.71)	(6.25)	(23.04)		

* Significant at the 1% level (9.21). B-J statistic follows χ^2 with 2 df.

(Table 14)

Specification for GARCH-M (1,1) in the hedging model

Variables Omitted from the Mean Equation

Variables	TR^2	Probability Distribution		t-statistic
MA(1)	1,090.19*	~	χ_1^2	47.03**
One Lagged Dependent Var.	982.69*	~	χ_1^2	21.92**

(Table 15)

Specification of GARCH Stochastic Process in the hedging model

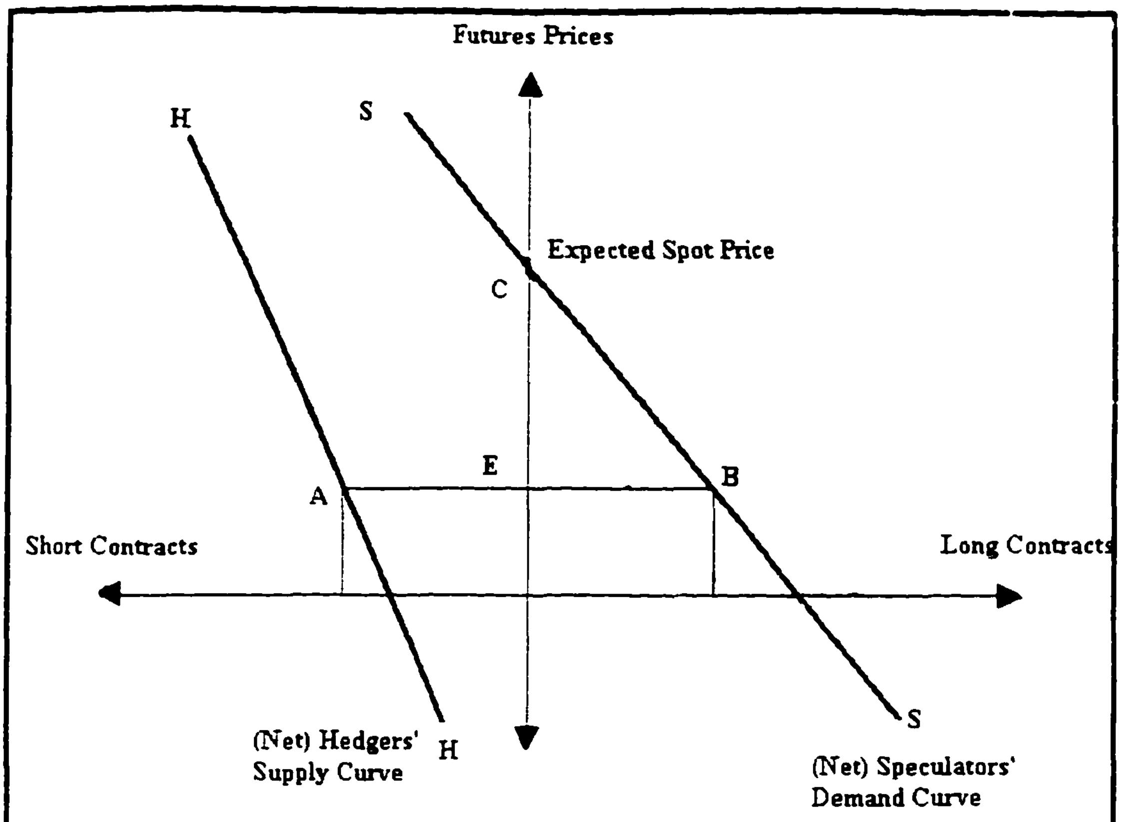
Variables Omitted from the Variance Equation

Alternative Hypothesis	parameter	TR^2	Probability Distribution	
GARCH(q=1,p=2)	β_2	550.03*	~	χ_1^2
GARCH(q=2,p=1)	α_2	0.42	~	χ_1^2

* Significant at the 1% level. The critical value of chi square distribution with the degree of freedom one is 6.63 at the level of type I error =.01. The * indicates that the null hypothesis is rejected at type I error = .01.

** The ** indicates that the null hypothesis is rejected at type I error = .01, where the critical value at $t_{.01}$ =2.576..

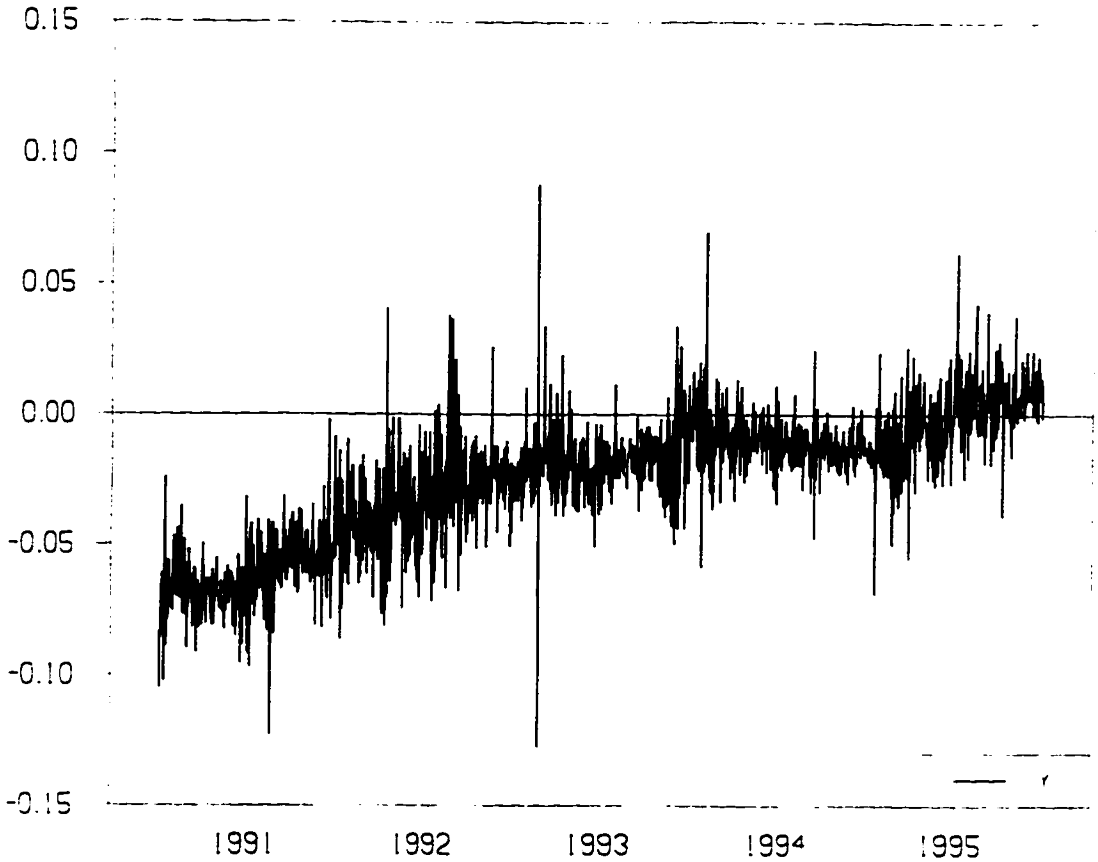
APPENDIX



(Figure 1) Hedgers' and Speculators' Curve

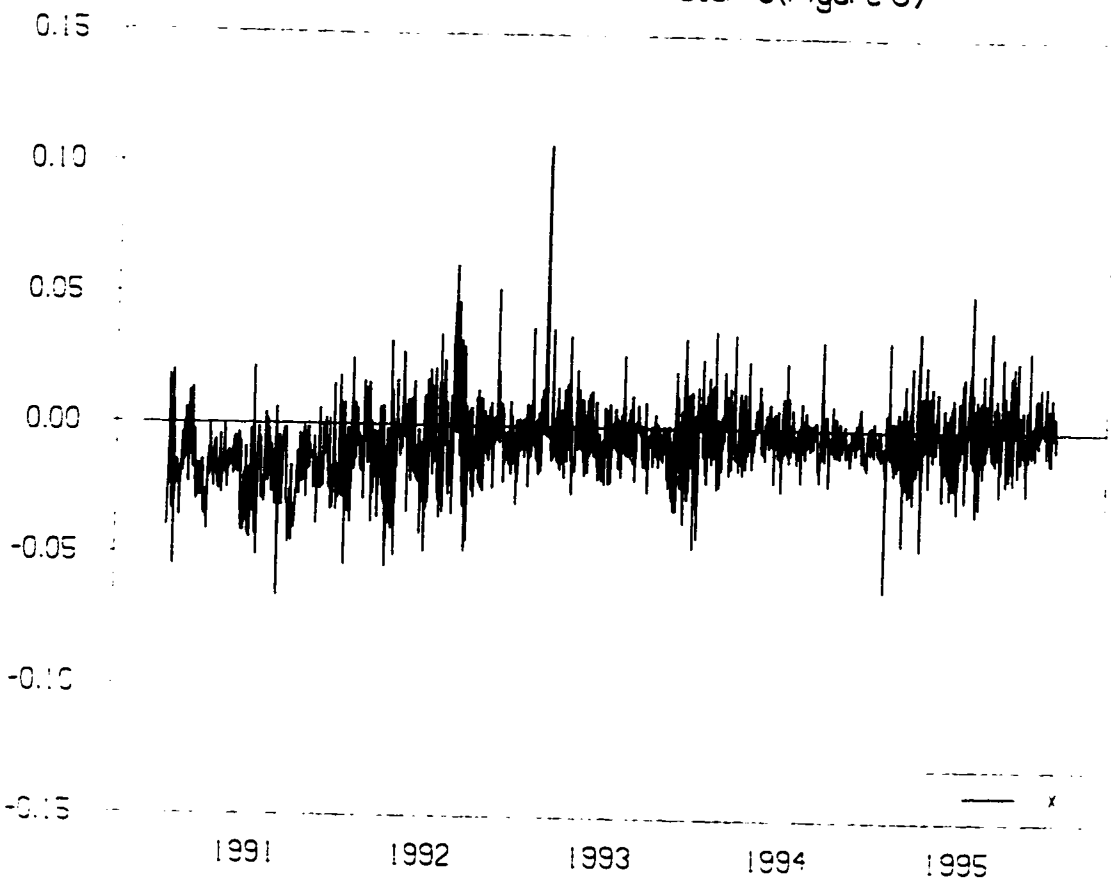
Nikkei 225 Stock Index (Figure 2)

Nikkei 225 Stock Index Returns(Figure 3)

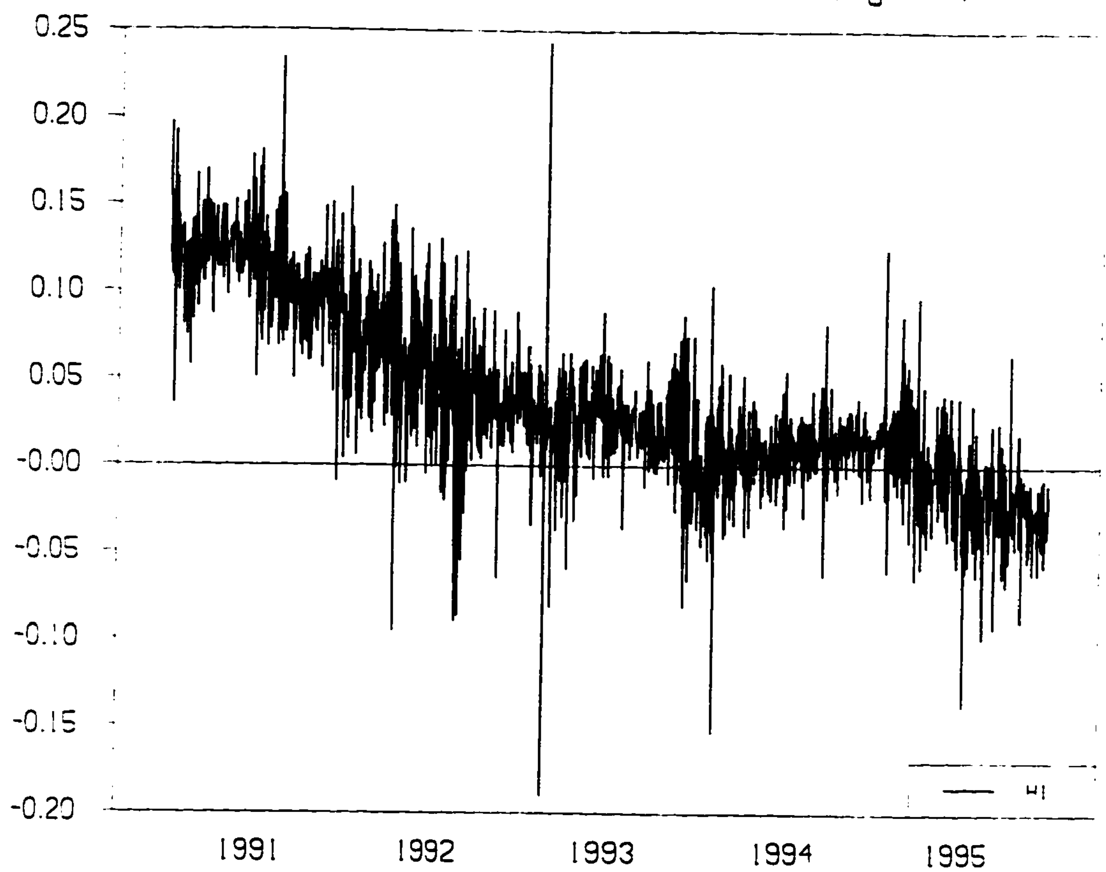



Nikkei 225 Stock Index Futures Contracts (Figure 4)
28000
26000
24000
22000
20000
18000
16000
14000
1991
1992
1993
1994
1995
FUTURES

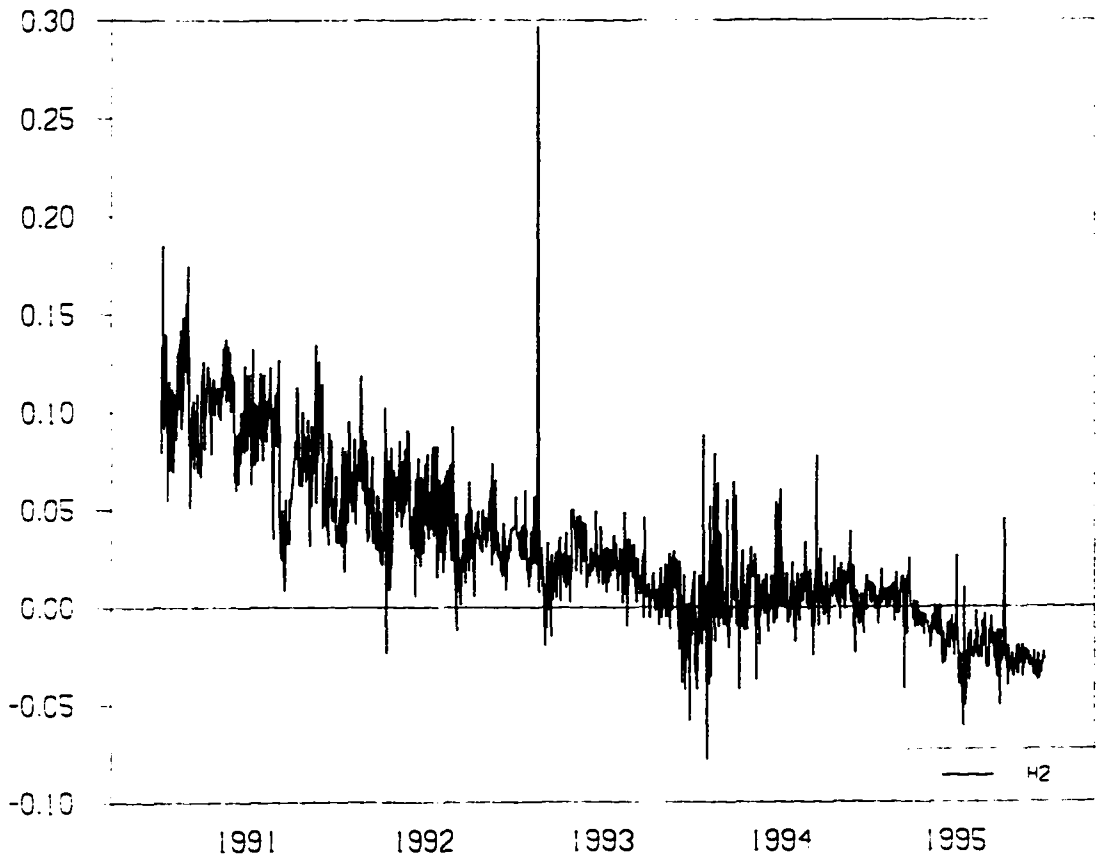
Nikkei 225 Stock Index Futures Returns(Figure 5)



Conditional Variance for the Nikkei Stock Index (Figure 6)



Conditional Variance for the Nikkei Stock Index in the Hedging Model (Figure 7)



BIBLIOGRAPHY

Abel, Andrew B., 1988, "Stock Prices Under Time Varying Dividend Risk : An Exact Solution in an Infinite-Horizon General Equilibrium Model," Journal of Monetary Economics, 22, 375-393.

Antonious, Antonios, and Holmes, Phil, 1995, "Futures Trading, Information and Spot Price Volatility : evidence for the FTSE-100 Stock Index Futures Contract Using GARCH," Journal of Banking and Finance, 19, 117-129.

Bailey, Richard T., Myers, Robert J., 1991, "Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge," Journal of Applied Econometrics, 6, 109-124.

Bailey, Warren, 1989, "The Market for Japanese Stock Index Futures : Some Preliminary Evidence, " The Journal of Futures Markets, 9, 283-295.

Baillie, Richard T. and DeGennaro, Ramon P., 1990, " Stock Returns and Volatility," Journal of Financial and Quantitative Analysis, 2, 203-214.

Bera, Anil, Bubnys, Edward L., and Park, Hun Y., 1993, "ARCH Effects and Efficient Estimation of Hedge Ratios For Stock Index Futures," Advances in Futures and Options Research, 6, 313-328.

Bera, Anil K., and C. Jarque, 1982, "Model Specification Tests: A Simultaneous Approach," Journal of Econometrics, 20.

Bera, Anil K., and Higgins, Matthew L., 1993, "ARCH Models : Properties, Estimation and Testing, " 7, 305-362,

Berndt, E. K., Hall, B.H., Robert E., and Hausman, Jerry A., 1974, "Estimation and Inference in Nonlinear Structural Models," Ann. Econ. And Soc. Measurement 3, 653-665.

Bollerslev, Tim, 1986, " Generalized Autoregressive Conditional Heteroskedasticity," Journal of Econometrics, 31, 307-328.

Bollerslev, Tim, 1987, " A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," Review of Economics and Statistics, 69, 542-547.

Bollerslev, Chou, and Kroner, 1992, "ARCH modelling in finance," Journal of Econometrics, 52, 5-59.

Bollerslev,T., R.F. Engle, and J.M. Wooldridge, 1988. " A Capital Asset Pricing Model With Time Varying Covariance", Journal of Political Economy, 116-131.

Brenner, Subrahmanyam, and Uno, 1989, "The behavior of Prices in the Nikkei Spot and Futures Market, " Journal of financial Economics, 23, 363-383.

Cecchetti, Cumby, and Figlewski, 1988, "Estimation of the Optimal Futures Hedge," The Review of Economics and Statistics, 623-630.

Chan, Kalok, and Chan, K.C., and Karoyi, G. Andrew, 1991, "Intraday Volatility in the Stock Index and Stock Index Futures Markets," The Review of Financial Studies, 4, 657-684.

Chou, Ray Yeutien, 1988, "Volatility Persistence and Stock Valuation: Some Empirical Evidence Using GARCH," Journal of Applied Econometrics, 3, 279-294.

Chou, Engle, and Kane, 1992, "Measuring risk aversion from excess returns on a stock index," Journal of Econometrics, 52, 201-224.

Cooper,Rick, 1991. "Risk Premia in the Futures and Futures Markets", Journal of Futures Market, 357-370.

Cornell, Bradford, and French, Kenneth R., 1983, "The Pricing of Stock Index Futures," The Journal of Futures Markets, 3, 1-14.

Cox,J.,J.Ingersoll, and S.Ross, 1981. "The Relation Between Forward Prices and Futures Prices", Journal of Financial Economics,321-346.

Dickey, D., and Fuller, W., 1981, "The Liklihood Ratio Statistics for Autoregressive Time Series With a Unit Root," Econometrica, 49, 1057-1072.

Dixit, Avinash K., and Pindyck, Robert S., 1994, Investment Under Uncertainty, Princeton University Press : Princeton.

Domowitz, I. and C.S. Hakkio,1985, "Conditional Variance and the Risk Premium in the Foreign Exchange Market", Journal of International Economics,47-66.

Duffie, Darrell, 1989, Futures Markets, Prentice-Hall, Inc. : New Jersey.

Dusak, K ,1973," Futures Trading and Investor Returns: An Investigation of Commodity Market Risk Premium", Journal of Political Economy, 81:1387-1406.

Ederington, Louis H, 1979, "The Hedging Performance of the New Futures Markets," Journal of Finance, 34, 157-70.

Edwards,Franklin R. 1988, " Futures Trading and Cash Market Volatility: Stock Index and Interest Rate Futures", Journal of Futures Market, 421-439.

Engle, Lilien, and Robins, 1987," Estimating Time Varying Risk Premia in the Term Structure", Econometrica,391-407.

Engle, R.F.,1982." Autoregressive Conditional Heteroskedacity with Estimates of the Variance of United Kingdom Inflation", Econometrica, 987-1007.

Engle, R.F., 1991, "Statistical Models for Financial Volatility" UCSD Discussion Paper, 91-32, 1-7.

Engle, R.F., and Bollerslev, T., 1986, "Modelling the persistence of conditional variances, " Econometric Reviews, 5, 1-50.

Figlewski, S., 1984, "Hedging Peformance and Basis Risk in Stock Index Futures, " Journal of Finance, 657-669.

Figlewski, S.,1986. Hedging with Financial Futures for Institutional Investors: From Theory to Practice, Ballinger Publishing Company: Cambridge,MA.

French, Kenneth R., 1983, " A Comparison of Futures and forward Prices," Journal of Financial Economics, 12, 311-342.

Ghosh, Asim, 1993, "Cointegration and Error Correction Models : Intertemporal Causality Between Index and Futures Prices," The Journal of Futures Markets, 13, 193-198.

Ghosh, Asim, 1993, "Hedging with Stock Index Futures : Estimation and Forecasting with Error Correction Model," The Journal of Futures Markets, 7, 743-752.

Glosten, L., R. Jagannaghan, and D. Runkle, 1989, "Relationship between the expected value and the volatility of the nominal excess return on stocks, Manuscript (Northwestern University, Evanston, IL).

Green, William H, 1993, Econometric Analysis, Macmillan: London.

Gujarati, Damodar N., 1995, Basic Econometrics, Mcgraw-Hill, Inc.: New York.

Hamilton, James D., 1994, Time Series Anaylysis, Princeton University Press: Princeton.

Hansen, H.N. and Kopprasch, Robert W. 1984, Pricing of Stock Index Futures, Dow

Jones: Irwin, IL.

Hodrick, Robert J. and Sanjay Srivastava, 1987, "Foreign Currency Futures", Journal of International Economics,1-24.

Hull, John C.,1993, Options, Futures, and Other Derivative Securities, second edition, New Jersey, Prentice-Hall, Inc.

Keynes, J.M.,1930, Treatise on Money, Macmillan: London.

Kolb, Robert W., 1994, Futures, Options, and Swaps, Miami, Florida, Kolb Publishing Co.

Lim, Kian-Guan, 1990, "Arbitrage and Price Behavior of the Nikkei Stock Index Futures," The Journal of Futures Market, 1-11.

Ljung, G. M. and Box, G.P.E., 1978, " On a measure of Lack of Fit in Times Series Models, " Biometrika, vol. 66, 66-72.

Lucas, Robert E. Jr., 1978, "Asset Prices in a Exchange Economy", Econometrica 46,1429-1445.

Mackinnon, J. 1991, "Critical Values for Cointegration Test," In Long-Run Economic Relationships Readings in Cointegration, Engle, R., and Granger, C.(eds.), New York : Oxford University Press, 267-276.

Mackkinlay, A. Craig, and Ramaswamy, Krishna, 1988, "Indes-Futures Arbitrage and the behavior of Stock Index Futures Prices," The Review of Financial Studies, 1, 137-158.

Mendelbrot, B., 1963, " The Variation of Certain Speculative Prices," Journal of Business, 36, 394-419.

Modest, David M., and Sundaresan, Mahadevan, 1983, "The Relationship Between spot and Futures Prices in Stock Index Futures Markets : Some Preliminary Evidence, "The Journal of Futures Markets, 3, 15-41.

Naka, Atsuyuki,1989, "The Volatility of Financial Markets: A Time Series Analysis of Foreign Exchange Futures", The University of Arizona, Dissertation, Unpublished.

Nefci, Salih N., 1996, An Introduction to the Mathematics of Financial Derivatives, Academic Press : San Diego.

Nelson. Daniel B., 1991, "Conditional Heteroskedasticity in Asset Returns : A New Approach. " Econometrica, 59, 347-370.

Park, Tae H., and Switzer, Lorne N., 1995, "Time-varying distributions and the optimal hedge ratios for stock index futures," Applied Financial Economics, 5, 131-137.

Ramsey, J., 1969, "Tests for Specification Errors in Classical Linear Least Squares Regression Analysis," Journal of the Royal Statistical Statistical Society, Series B, 31, 350-371.

Richard, Scott F., and Sundaresan, M., 1983 "A Continuous Time Equilibrium Model of Forward Press and Futures Prices in a Multigood Economy," Journal of Financial Economics, 9, 347-371.

Rossi, Peter E., 1996, Modelling Stock Market Volatility, Academic Press : San Diego.

Sargent, Thomas J. 1987, Dynamic Macroeconomic Theory, Harvard University Press: London.

Samuelson, P. 1965, " Proof That Properly Anticipated Prices Fluctuate Randomly", Industrial Management Review, 41-49.

Schwert, G. William, 1989, "Why Does Stock Market Volatility Change Over Time.," The Journal of Finance, 5, 1115-1153.

White, H., 1980, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." Econometrica 48.

Williams, Jeffrey, 1986, The Economic Function of Futures Markets, Cambridge University Press : Cambridge.

Yoo, Jisoo, and Maddala G.S., 1991, "Risk Premia and Price Volatility in Futures Markets, " The Journal of Futures Markets, 11, 165-177.