ALTERNATIVE MEASURES OF VOLATILITY IN AGRICULTURAL FUTURES MARKETS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Yuanfang Wang, M.A.

* * * * *

The Ohio State University
2005

Dissertation Committee:  Approved by
Dr. Matthew Roberts, Adviser
Dr. Brian Roe
Dr. Stanley Thompson

Graduate Program in Agricultural, Environmental, and Development Economics
ABSTRACT

The three essays of this thesis focus on modeling and forecasting agricultural futures market volatility utilizing two nonparametric volatility measures, realized volatility and range-based volatility. Special attention is given to comparing the performance of time series models used previously in the literature with these two measures.

The first essay investigates the properties of realized volatility in the soybean futures market. The results indicate that the distributional properties of realized volatility based on 5-minute returns largely correspond with existing literature. The findings of three volatility measures confirm that the Mixture-of-Distributions-Hypothesis (MDH) is valid. In contrast, the standardized daily returns display some different properties compared with stock and exchange rate data. Moreover, the parametric ARFIMA and GARCH models reflect same patterns as described in nonparametric analysis.

The second essay compares the performance of GARCH models, range-based GARCH models, and log-range based simple regression models in terms of their forecasting abilities. Realized volatility is used as the forecasting evaluation criterion. In-sample fitting results reveal that range-based GARCH models outperform standard GARCH models. Out-of-sample tests indicate that GARCH models extended with daily ranges work better than the log-range based ARMA models.
The third essay examines the economic value of realized volatility in the context of hedging. Specifically, it provides an analysis of incorporating information contained in intraday prices into multivariate GARCH models and evaluates the statistical performance of the augmented GARCH models. Then, effectiveness of different optimal hedge ratios depending on different estimation procedures is compared by examining the in-sample and out-of-sample performance of the ratios. Results reveal that the use of the augmented GARCH models, while statistically appropriate, provides only marginal gain to the hedger.
ACKNOWLEDGMENTS

First and foremost, I am deeply grateful to my adviser, Dr. Matthew Roberts, for his encouragement and patience through my dissertation writing stage. Dr. Roberts’ advice based on his broad knowledge on econometrics and programming always helps me avoid detours in research. Without his support, this dissertation would not even have started.

My greatest thanks also go to my committee members, Dr. Brian Roe and Dr. Stanley Thompson. Their suggestions are helpful for improving this thesis.

I also wish to thank Dr. Carl Zulauf for his excellent lectures on futures and options, which have significantly broadened my research horizon.

Finally, I would like to extend my special thanks to Dr. Larry Libby. His warm-hearted help has firmly supported me during the past several years. I have learned a lot from his attitude toward life.
VITA

1996………………  B.A. Economics, Capital University of Economics and Business
1999………………  M.A. Industrial Economics, Capital University of Economics and Business
2000………………  M.A. Economics, The Ohio State University
1999 – Present…… Graduate Teaching and Research Associate, The Ohio State University

FIELDS OF STUDY

Major Field: Agricultural, Environmental, and Development Economics
## TABLE OF CONTENTS

Abstract ........................................................................................................................................... ii
Acknowledgments ......................................................................................................................... iv
Vita ................................................................................................................................................... v
List of Tables ................................................................................................................................... viii
List of Figures .................................................................................................................................. ix
Chapters:

1. Introduction ........................................................................................................................... 1
2. Realized Volatility in the Agricultural Futures Market ......................................................... 3
   2.1 Introduction ....................................................................................................................... 3
   2.2 Realized Volatility Measurement ..................................................................................... 7
      2.2.1 Theory ....................................................................................................................... 7
      2.2.2 Data Source and Construction ................................................................................. 9
      2.2.3 Selecting Time Interval ............................................................................................ 10
      2.2.4 Price Limits ............................................................................................................. 13
   2.3 Distributions of Realized Volatility and Returns ............................................................... 16
      2.3.1 Results of Existing Studies ..................................................................................... 16
      2.3.2 Distributional Properties of Volatilities .................................................................. 20
         2.3.2.1 Unconditional Distributions .............................................................................. 20
         2.3.2.2 Temporal Dependence ..................................................................................... 22
      2.3.3 Distributions of Standardized Returns ...................................................................... 24
      2.3.4 Other Properties ....................................................................................................... 26
         2.3.4.1 ARCH-M Effect ............................................................................................... 26
         2.3.4.2 Asymmetric Volatility ..................................................................................... 26
   2.4 Modeling Realized Volatility and Daily Returns ............................................................... 28
      2.4.1 ARMA and ARFIMA Models .................................................................................... 29
      2.4.2 ARMA and ARFIMA Model Estimation Results ..................................................... 30
      2.4.3 GARCH Model ....................................................................................................... 31
      2.4.4 Model Estimation Results ....................................................................................... 33
   2.5 Conclusion ......................................................................................................................... 35
3. Forecasting Daily Volatility Using Range-based Data ............................................................ 50
   3.1 Introduction ....................................................................................................................... 50
   3.2 Range-based Volatility Estimators .................................................................................... 52
      3.2.1 Review of Range-based Volatility Estimators ......................................................... 53

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.2 Properties of the Log-Range Volatility Proxy</td>
<td>55</td>
</tr>
<tr>
<td>3.3 Time Series Models</td>
<td>56</td>
</tr>
<tr>
<td>3.3.1 GARCH Models</td>
<td>56</td>
</tr>
<tr>
<td>3.3.2 GARCH Models Extended with Additional Information</td>
<td>57</td>
</tr>
<tr>
<td>3.3.3 Simple Regression Model</td>
<td>58</td>
</tr>
<tr>
<td>3.4 Data and In-Sample Fit</td>
<td>59</td>
</tr>
<tr>
<td>3.4.1 Data Description</td>
<td>59</td>
</tr>
<tr>
<td>3.4.2 In-Sample Fit</td>
<td>61</td>
</tr>
<tr>
<td>3.4.2.1 Estimation of GARCH Models</td>
<td>61</td>
</tr>
<tr>
<td>3.4.2.2 GARCH Models Extended with Daily Range</td>
<td>63</td>
</tr>
<tr>
<td>3.4.2.3 Range-based ARMA Models</td>
<td>65</td>
</tr>
<tr>
<td>3.5 Out-of-Sample Daily Volatility Forecasts</td>
<td>65</td>
</tr>
<tr>
<td>3.5.1 Forecast Evaluation Criteria</td>
<td>65</td>
</tr>
<tr>
<td>3.5.2 Results</td>
<td>67</td>
</tr>
<tr>
<td>3.6 Conclusion</td>
<td>68</td>
</tr>
<tr>
<td>4. Estimation of the Optimal Commodity Hedge Ratio</td>
<td>76</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>76</td>
</tr>
<tr>
<td>4.2 The Optimal Hedge Ratio</td>
<td>79</td>
</tr>
<tr>
<td>4.2.1 The Minimum Variance Hedge Ratio</td>
<td>79</td>
</tr>
<tr>
<td>4.2.2 Alternative Methods for Estimating the Optimal Hedge Ratio</td>
<td>82</td>
</tr>
<tr>
<td>4.3 Estimation of Bivariate GARCH models</td>
<td>84</td>
</tr>
<tr>
<td>4.3.1 Data and Basic Statistics</td>
<td>84</td>
</tr>
<tr>
<td>4.3.2 Bivariate GARCH Models</td>
<td>86</td>
</tr>
<tr>
<td>4.3.3 Bivariate GARCH Models for Soybean Data</td>
<td>88</td>
</tr>
<tr>
<td>4.3.3.1 Basic BEKK GARCH Model</td>
<td>88</td>
</tr>
<tr>
<td>4.3.3.2 Augmented BEKK GARCH Model</td>
<td>88</td>
</tr>
<tr>
<td>4.3.3.3 Model Estimation and Results</td>
<td>89</td>
</tr>
<tr>
<td>4.4 The Optimal Hedge</td>
<td>92</td>
</tr>
<tr>
<td>4.4.1 Derivation of the Optimal Hedge Ratio</td>
<td>92</td>
</tr>
<tr>
<td>4.4.2 Optimal Hedge Results</td>
<td>94</td>
</tr>
<tr>
<td>4.4.2.1 In-Sample Results</td>
<td>94</td>
</tr>
<tr>
<td>4.4.2.2 Out-Of-Sample Results</td>
<td>97</td>
</tr>
<tr>
<td>4.4.2.3 Hedging Effectiveness: In-Sample and Out-Of-Sample</td>
<td>99</td>
</tr>
<tr>
<td>4.5 Conclusion</td>
<td>103</td>
</tr>
<tr>
<td>5. Summary and Conclusion</td>
<td>112</td>
</tr>
</tbody>
</table>

Bibliography                                                              | 114  |
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Summary Statistics of Daily Realized Variance, Realized Standard Deviation and Logarithmic Standard Deviation</td>
<td>37</td>
</tr>
<tr>
<td>2.2</td>
<td>Ljung-Box and Augmented Dickey-Fuller Test Statistics</td>
<td>37</td>
</tr>
<tr>
<td>2.3</td>
<td>Descriptive Statistics for Daily Returns</td>
<td>37</td>
</tr>
<tr>
<td>2.4</td>
<td>News Impact Function Estimates</td>
<td>38</td>
</tr>
<tr>
<td>2.5</td>
<td>ARMA and ARFIMA Model Estimates</td>
<td>38</td>
</tr>
<tr>
<td>2.6</td>
<td>FIGARCH, FIGARCH-M and FIEGARCH Estimation Results</td>
<td>39</td>
</tr>
<tr>
<td>3.1</td>
<td>Summary Statistics for Soybean Futures Returns (100 × ln(P_t / P_{t-1}))</td>
<td>69</td>
</tr>
<tr>
<td>3.2</td>
<td>Summary Statistics for Soybean Futures Log Absolute Returns and Log Range</td>
<td>70</td>
</tr>
<tr>
<td>3.3</td>
<td>GARCH Estimation Results</td>
<td>71</td>
</tr>
<tr>
<td>3.4</td>
<td>Estimation Results of GARCH Extended with Daily Range</td>
<td>72</td>
</tr>
<tr>
<td>3.5</td>
<td>Daily Volatility Forecast Performance</td>
<td>73</td>
</tr>
<tr>
<td>4.1</td>
<td>Augmented Dickey-Fuller Test for Cash and Futures Prices &amp; Cash and Futures Returns</td>
<td>105</td>
</tr>
<tr>
<td>4.2</td>
<td>Summary Statistics of Cash and Futures Returns: Sub-samples and Full Sample</td>
<td>105</td>
</tr>
<tr>
<td>4.3</td>
<td>Estimation Results of the Bivariate GARCH Models</td>
<td>106</td>
</tr>
<tr>
<td>4.4</td>
<td>Descriptive Statistics for Hedge Ratios</td>
<td>107</td>
</tr>
<tr>
<td>4.5</td>
<td>Hedging Effectiveness Comparison: In-Sample and Out-of-Sample Results</td>
<td>108</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Volatility Signature Plot for Different Time Intervals</td>
<td>40</td>
</tr>
<tr>
<td>2.2</td>
<td>5-Minute Return Series</td>
<td>40</td>
</tr>
<tr>
<td>2.3</td>
<td>Autocorrelations for 5-minute Return Series</td>
<td>41</td>
</tr>
<tr>
<td>2.4</td>
<td>Realized Volatility and Squared Daily Returns</td>
<td>41</td>
</tr>
<tr>
<td>2.5</td>
<td>Kernel Density Estimates</td>
<td>42</td>
</tr>
<tr>
<td>2.6</td>
<td>Q-Q Plots of Daily Realized Volatilities</td>
<td>43</td>
</tr>
<tr>
<td>2.7</td>
<td>Time Series Plots of Daily Realized Volatilities</td>
<td>44</td>
</tr>
<tr>
<td>2.8</td>
<td>Autocorrelation Functions of Daily Realized Volatilities</td>
<td>45</td>
</tr>
<tr>
<td>2.9</td>
<td>Time Series Plot of Daily Returns</td>
<td>46</td>
</tr>
<tr>
<td>2.10</td>
<td>Q-Q Plots of Daily Returns</td>
<td>47</td>
</tr>
<tr>
<td>2.11</td>
<td>Realized Volatilities Vs. Daily Returns</td>
<td>48</td>
</tr>
<tr>
<td>2.12</td>
<td>News Impact Curve</td>
<td>49</td>
</tr>
<tr>
<td>3.1</td>
<td>Soybean Futures Prices (1985/01-1989/12)</td>
<td>74</td>
</tr>
<tr>
<td>3.2</td>
<td>Soybean Futures Returns (1985/01-1989/12)</td>
<td>74</td>
</tr>
<tr>
<td>3.3</td>
<td>Q-Q Plots of Log Range and Log Absolute Returns</td>
<td>75</td>
</tr>
<tr>
<td>4.1</td>
<td>Time-Varying Hedge Ratios vs. Constant Hedge Ratio for Year 1996</td>
<td>109</td>
</tr>
<tr>
<td>4.2</td>
<td>Time-Varying Hedge Ratios vs. Constant Hedge Ratio for Year 1999</td>
<td>110</td>
</tr>
<tr>
<td>4.3</td>
<td>Out-of-Sample Forecasts of Optimal Hedge Ratios</td>
<td>111</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Users of financial markets always require accurate estimates of volatility. Volatility is a key factor in risk management, security valuation and option pricing. The application of misspecified volatility has the potential to induce inappropriate or even wrong assessment of asset risk and portfolio selection. Thus, not surprisingly, seeking good volatility estimators has drawn increased attention from financial academics and practitioners. Recently, two volatility proxies, realized volatility and range-based volatility, have been increasingly used to address a range of issues in financial markets. However, most studies just focus on examining their distributional properties. Although these studies have demonstrated that these two volatility proxies have many appealing properties, these findings are essentially statistical. If the realized volatility and the range-based volatility approaches can yield substantial economic benefits, it will hold promise for practical values of these two proxies in asset pricing and financial risk management applications.

For the above reason, this thesis investigates these two volatility proxies and examines their implications for volatility forecasting and risk management in agricultural futures markets. Chapter two investigates the properties of realized volatility in the soybean futures market. The results indicate that the distributional properties of realize
volatility based on 5-minute returns largely correspond with existing literature. The findings of three volatility measures confirm that the Mixture-of-Distributions-Hypothesis (MDH) is valid. In contrast, the standardized daily returns display some different properties compared to stock and exchange rate data. Moreover, the time series ARFIMA and GARCH models reflect same patterns as described in nonparametric analysis.

Chapter three compares the performance of GARCH models, range-based GARCH models, and log-range based simple regression models in terms of their forecasting abilities. Realized volatility is used as the forecasting evaluation criterion. In-sample fitting results reveal that range-based GARCH models perform better compared with standard GARCH models. Out-of-sample results reveal that the log-range based ARMA models may have disadvantages in comparison to GARCH models extended with daily ranges.

Chapter four examines the economic value of realized volatility in the context of hedging. Specifically, it provides an analysis of incorporating information contained in intraday prices into multivariate GARCH models and evaluates the statistical performance of the augmented GARCH models. Then, effectiveness of different optimal hedge ratios depending on different estimation procedures is compared by examining the in-sample and out-of-sample performance of the ratios. The results indicate that the augmented GARCH models only provide marginal gains to hedgers. Contrary to previous research, this chapter finds less support for the realized volatility approach.

Chapter five offers concluding remarks and summary of the thesis.
CHAPTER 2

REALIZED VOLATILITY IN
THE AGRICULTURAL FUTURES MARKET

2.1 Introduction

Efficient estimation of market volatility is very important to financial research. However, because asset price volatility is not directly observable, much effort has been devoted to extracting volatility from other observable market activities. One common approach, which a voluminous literature has employed, is to estimate the latent volatility using time-varying volatility models. These time series models fall into one of two categories, the ARCH family, first introduced by Engel (1982), and the stochastic volatility (SV) family, which traces its roots to Clark (1973). Both ARCH and SV models treat volatility as an unobservable, time-varying process. The features of SV models are similar to the ARCH specification but involve a contemporaneously stochastic component. Another widely-used approach is to extract implied volatility from options or other derivatives prices. For example, information of volatility can be extracted with the well-known Black-Scholes option pricing formula.

Estimating ARCH models is relatively easy since ARCH models have closed-form likelihood functions in spite of variance being unobserved. However, it is often difficult
to draw sharp distinctions between competing ARCH models. Andersen, Bollerslev, Diebold, and Labys (2001, ABDL, hereafter) claim that the existence of multiple competing models suggests misspecification and the robustness of the volatility measures based on these models is uncertain. Compared with ARCH models, stochastic volatility models are relatively difficult to estimate since closed-form likelihood functions don’t exist for stochastic volatility models. Moreover, the estimates based on option implied volatility models and statistical models are only valid under specific assumptions. Thus, all of these approaches have distinct weaknesses.

An alternative approach to measure market volatility is a model-free estimator, which uses simpler techniques to provide estimates of the ex post realized volatility. The traditional model-free estimator of volatility is based on the calculation of squared returns over the relevant return horizon. A natural proxy of daily volatility, $\sigma_t^2$, is $r_t^2$, where $r_t$ is the daily return $r_t = \ln(p_t) - \ln(p_{t-1})$ and $p_t$ is the closing price on day $t$. The daily squared return constitutes an unbiased estimator for the latent volatility factor. However, Andersen and Bollerslev (1998b) show that it is also a very noisy volatility estimator and does not provide reliable inferences regarding the underlying latent volatility in daily samples.

The limitations of traditional procedures motivate many different approaches for measuring and analyzing financial market volatility. Parkinson (1980) proposes the extreme value method for estimating the variance of the rate of return. This approach utilizes the information of daily high and low prices and the efficiency of this estimator is much higher than the average of daily squared returns. Garman and Klass (1980) improve on this procedure by constructing a new volatility estimation measure based on the
historical opening, closing, high, and low prices. They claim that the new method is more efficient than Parkinson’s volatility estimator. Kunitomo (1992) improves Parkinson’s estimation method by allowing the presence of drift terms in stochastic processes. Although these alternative procedures lead to estimates of varying statistical efficiency, the derivation of all of these estimators depends on the assumption of continuous price paths. When the sample path of prices can only be observed at discrete transactions, all the statistics will be biased (Garman and Klass, 1980). Furthermore, Ball and Torous (1984) point out that estimators based on daily high and low prices rely only on the first few moments of the joint probability distribution of daily high, low and closing prices and additional information must be provided for more efficient volatility estimators.

Most recently, the availability of intraday financial databases has had an important impact on research in financial market volatility. These intraday data, also called tick-by-tick data, have been actively recorded in several exchanges, such as Chicago Board of Trade (CBOT), New York Stock Exchange (NYSE), American Stock Exchange (AMEX) or the National Association of Security Dealers Automated Quotation system (NASDAQ). Traditional financial databases usually provide daily or weekly data, but intraday data give much more information about the market and its associated characteristics. The availability of these new datasets has shed new light on the modeling of volatility. Taylor and Xu (1997) and Andersen and Bollerslev (1997, 1998a) provide thorough descriptions of intraday data and intraday volatility. They show that high frequency intraday returns contain valuable information for the measurement of volatility at the daily level.

Based on these earlier studies, Andersen and Bollerslev (1998b) introduce a new and complementary volatility measure, termed realized volatility. Realized volatility
estimates volatility by summing squared intraday returns. Andersen, Bollerslev, Diebold and Ebens (2001, ABDE, hereafter) show that volatility estimates so constructed are close to the underlying integrated volatility. Thus, the volatility of a price process can be treated as an observable process.

Most of the recent research focuses on describing the distributional properties of realized volatility and modeling realized volatility. For example, ABDL (2000) study daily volatility of DM/$ and Yen/$ exchange rates; Ebens (1999) analyzes the Dow Jones Industrial Average (DJIA) index; ABDE (2001) examine 30 individual stocks; Areal and Taylor (2002) study the distributional properties of FTSE-100 futures price; Thomakos and Wang (2003) consider four futures contracts: the Deutsche Mark, the S&P 500 index, US Bonds and the Eurodollar. Although these findings provide support for the realized volatility approach, there is an obvious void in the existing literature. Exchange and stock markets have drawn most researchers’ attention. Although there are several studies on futures market, no papers have explored the properties of realized volatility in non-financial asset prices.

The purpose of this paper is to address the above issue. The other aim is to compare distributional properties of realized volatilities of non-financial futures with the ones obtained on currency markets and financial futures markets. This study has five sections. The next section explains some theoretical background of realized volatility, describes the data and discusses the choice of sampling interval. Section three introduces existing studies on distributional properties of realized volatilities and presents the results for soybean futures data. In section four, time series models are fitted to realized volatility measures and soybean futures daily returns. The final section concludes.
2.2 Realized Volatility Measurement

2.2.1 Theory

The idea of using higher frequency data to generate measures of lower frequency volatility traces its origin to French, Schwert, and Stambaugh (1987), Schwert (1989, 1990a) and Schwert and Seguin (1991). They construct monthly realized equity price volatilities by using squared daily returns. Schwert (1990b) uses the standard deviations of intraday returns to study volatility. Schwert (1998) constructs daily stock market volatilities relying on 15-minute returns. These studies lack formal justification and theoretical underpinnings for such measures. It was not until Andersen and Bollerslev (1998b) that realized volatility was formalized.\textsuperscript{1}

Let \( p_{n,t} \) denote the time \( n \) logarithmic price at day \( t \), where \( n = 1, \ldots, N \), and \( t = 1, \ldots, T \). Assume it follows a continuous-time stochastic volatility diffusion,

\[
dp_{n,t} = \sigma_{n,t} dW_{n,t}
\]

(2.1)

where \( W_{n,t} \) denotes a standard Brownian motion. The discretely observed time series of returns with \( n \) observations per day, or a return horizon of \( 1/n \), is defined by

\[
r_{n,t} \equiv p_{n,t} - p_{n-1,t}
\]

(2.2)

Given the sample path of variance, \( \{\sigma_{n,t}\}_{n=1,\ldots,N,t=1,\ldots,T} \), then the variance of daily returns is

\[
\sigma^2_t \equiv \int_1^N \sigma^2_{n,t} dn
\]

(2.3)

And the sum of intraday squared returns, the realized volatility, is defined as

\textsuperscript{1} The realized volatility measure is also formally developed in an independent work by Barndorff-Nielsen and Shephard (2002). ABDL (2001) provide rigorous theoretical underpinnings for realized volatility in the context of special semi-martingales.
Using the theory of quadratic variation, Andersen and Bollerslev (1998b) show that the quadratic variation of the returns in (2.4) converges to the integrated volatility of (2.3) almost surely for all \( t \) as the sampling frequency of the returns increases, or \( n \to \infty \),

\[
\operatorname{plim}_{n \to \infty} \sum_{n=1}^{N} r_{n,t}^2 = \int_{0}^{\infty} \sigma_{n,t}^2 \, dn = \sigma_t^2
\]  

(2.5)

It follows, therefore, that by using intradaily returns, nonparametric, model-free estimates of volatility can be constructed. The usefulness of the realized volatility estimator has been proved in several studies. For example, Andersen and Bollerslev (1998b) show that increasing the sampling frequency provides a much better volatility forecast criteria than the squared daily returns.

An alternative approach, based on a discrete-time framework, is established by Ebens (1999) with the same consistency results. Hence, volatility may be treated as observed by sampling intraday returns sufficiently frequently. Generally, \( n = 1 \) is normalized to represent one trading day. The intraday hourly, 10-minute, 5-minute, and continuous returns are thus denoted by \( n = 24, 144, 288, \infty \). As \( n \) increases, volatility estimates become more precise. However, in practice, because of data limitations and market microstructure features, such as price discreteness and nonsynchronous trading, increasing \( n \) too far also induces noise into the measurement. Nonetheless, five minute intervals have been widely used in extant studies. ABDL (2001) propose that this sampling frequency balances two competing factors: the benefits derived from large

\[ \text{These frequencies are based upon currency trading, i.e., a 24 hour market. Hourly data for futures does not produce 24 observations per day.} \]
sample and the contamination of market microstructure effects. Of course, this does not mean that the choice of the size of $n$ is predetermined for all markets since most of the literature focuses on currency markets. Taken together, the calculation of realized volatility is an empirical issue that depends on market characteristics.

2.2.2 Data Source and Construction

Intraday Chicago Board of Trade soybean futures prices were obtained from the Futures Industry Institute. The data are time and sales transaction prices, recorded by the exchange. The full sample consists of 4,949,175 high frequency prices from January 2, 1990 through July 31, 2001 for all futures maturities. From 1990 to 1999, the data specify the transaction time to an accuracy of one second, and for the final two years, only to an accuracy of one minute. The price record covers the full CBOT floor trading from 9:30 a.m. to 1:15 p.m. (Central Standard Time).

Dates on which there was a span of at least 25 minutes without trades were omitted. Since most trading activity is usually concentrated in the contract nearest to delivery, the calculation of the returns is based on the nearby futures contract over consecutive intervals. However, returns are calculated from the second nearby contract when the nearby contract is in the delivery month. In order to calculate a continuous sequence of futures returns, an interpolation method is employed. Specifically, returns are calculated using the last recorded logarithmic price before and the consecutive price after each five-minute mark. This interpolated average is weighted linearly by the inverse relative distance to each time mark. The first return for the trading day is deleted since it is an
overnight return.\(^3\) For example, the returns from 9:30 a.m. to 9:35 a.m. are deleted when calculating 5-min returns.

All in all, these corrections result in a sample of 2909 days. One normal trading day consists of 44 intraday 5-minute returns or 23 intraday 10-minute returns or 15 intraday 15-minute returns and so on. The realized volatility is calculated according to equation (2.4).

### 2.2.3 Selecting Time Interval

The usefulness of realized volatility computed from high-frequency data depends on sampling frequency. The theory suggests that realized volatility is effectively an error-free volatility measure provided that returns are sampled sufficiently frequently. However, market microstructure effects prevent sampling too frequently. The price discreteness, the nonsynchronous trading and quotation effect may induce negative autocorrelation in the interpolated returns (ABDL, 2001). Consequently, a problem arises: a reasonable choice of sampling frequency may not be simply the highest available. It may be some value that can balance the microstructure frictions and the measurement errors. As briefly reviewed in section 2.2.1, most of the existing studies focus on using 5-minute returns to obtain the daily realized volatility. The authors believe that this frequency balances between the cost of market microstructure biases and the need of sampling at very high frequencies. However, there are no standard criteria that provide formal justifications about how to choose sampling frequency. Some authors use the median length of duration between

---

\(^3\) Martens (2002) shows that the realized volatility excluding the overnight return works better than that including the overnight return for S&P 500 index futures contract.
price changes to construct returns series. Some believe that autocorrelations between high
frequency returns provide useful information in selecting time intervals. Others insist
checking the market liquidity and trading volumes. The absence of a uniform selection
method leads to different conclusions.

ABDL (1999) develop a simple graphical diagnostic, the volatility signature plot.
They propose that microstructure bias tends to manifest itself as sampling frequency
increases by distorting the average realized volatility. Thus, plotting average realized
volatility against sampling frequency may be useful in selecting the optimal frequency.
However, this method tends to choose lower frequencies for both liquid and illiquid
assets. Figure 2.1 shows the volatility signature plots. From this graph, average realized
volatility remains stable as sampling frequency increases up to approximately 30-minute
returns. This result is consistent with ABDL’s finding for the Deutschemark-U.S. Dollar
and Japanese Yen-U.S. Dollar exchange rates. Theoretically, the selection of 30-minute
interval represents a reasonable tradeoff between the need of sampling at high
frequencies and the cost of market microstructure biases. However, in practice, the open
outcry session of soybean futures only spans from 9:30 am to 1:15 pm. The total pit time
is 225 minutes. If 30-minute returns are used to estimate the realized volatility, there are
only 7 observations for each trading day. Thus, the realized volatility estimates may have
measurement errors due to few observations.

ABDL (1999) also suggest that high frequency return autocorrelations provide
complementary information for constructing realized volatility intervals. The reason is
straightforward. Tick data are not regularly time-spaced. However, in order to calculate
continuous returns series, interpolation methods are necessary for constructing regularly
time-spaced data. ABDL (2001) conclude that irregular spacing of the data induces negative autocorrelation in the fixed-interval return series. Bid-ask bounce effects may exaggerate the spurious negative dependence. Thus, negative serial correlations signify the existence of measurement errors.

Since one of the purposes of this study is to compare the properties of realized volatility of grain futures market with those in extant literature, where 5-minute intervals are frequently used, the remainder of this study will focus on 5-minute return series. To check whether this is the proper frequency for computing realized volatility for the current dataset, autocorrelations of 5-minute returns are investigated. Figure 2.2 plots the 5-minute return series. The basic features of this series are quite similar to those observed at the daily level. For example, tranquil (volatile) periods tend to be followed by tranquil (volatile) periods. Thus, the clustering effect is obvious. Figure 2.3 shows the ACF of the returns series. The first two sample autocorrelations are well outside the 95% confidence interval.

Ebens (1999) provides an approach to determine the negative effect that spurious correlations may induce. Applying the standard MA(q) model to the return series,

$$t_{n,t} = \xi_{n,t} + \theta_{1,t} \xi_{n-1,t} + \theta_{2,t} \xi_{n-2,t} + \ldots + \theta_{q,t} \xi_{n-q,t}$$  \hspace{1cm} (2.6)

where $\xi_{n,t}$ is a white noise process, Ebens (1999) shows that the relationship between realized volatility and daily volatility can be expressed as,

$$E(\sigma_n^2) = (1 + \sum_{i=1}^{q} \theta_{i,t}^2) \sigma_t^2$$  \hspace{1cm} (2.7)

where $\sigma_t^2$ denotes actual daily volatility.
From equation (2.7), it is obvious that spurious dependence between returns will result in larger estimates of the actual volatility. If unevenly spaced data and bid-ask bounce are present and significant, the realized volatility estimate is erroneous and the consistency results may not hold. Given the fact that the first two sample autocorrelations are significant judged by the 95% confidence interval, applying an MA(2) model to soybean data obtains that $\hat{\theta}_1 = -0.0240$ and $\hat{\theta}_2 = -0.0177$. Equation (2.7) reveals that the measurement error induced by serial correlation is only 0.0009. In other words, the realized volatility estimate scales up the actual volatility by 1.0009. The order of magnitude is small and thus can be ignored.\(^4\) Hence, 5-minute return series are used in place of 30-minute returns for the remaining of this study. Figure 2.4 plots squared daily returns and 5-minute-based realized volatility. The two series reflect similar patterns. However, squared daily returns are much more volatile than realized volatility.

### 2.2.4 Price Limits

Futures markets use price limits as one of the regulation tools to guarantee market integrity. Price limits restrict transaction prices to lie between a symmetric range around the previous day’s settlement price. The origin of such limits can be traced to the desire of authorities to reduce the default risk and lower the margin requirement. The Chicago Board of Trade formally applied daily limits in 1925. No trade can take place outside of the limit bounds. For some futures contracts, however, the price limits may be expanded or removed after the contract is locked limit. Also, limits are lifted 2 business days before

---

\(^4\) Ebens (1999) applies equation (2.6) and (2.7) to DJIA data and finds a factor of 1.0029. He asserts that this factor is small enough that no adjustments need to be made.
the delivery month. Advocates of price limits believe that price limits decrease price volatility, reduce default risks and margin requirement, and do not interfere with trading activity. On the other hand, critics claim that price limits create higher volatility levels on subsequent days, prevent prices from reaching the equilibrium level and interfere with trading activities. There has been a large amount of empirical research related to price limits. However, empirical research does not provide conclusive support for either position.

Hall and Kofman (2001) claim that price limits affect market participants’ expectation and decisions. Traders adjust their trading behavior by revising their order flow. Correspondingly, the volatility of prices is affected. As revealed in previous sections, the effectiveness of realized volatility depends on its ability to capture the real price discovery process. If price limits affect the underlying generating process for the equilibrium prices or delay price discovery, realized volatility may be a biased estimator of the market volatility. Furthermore, while some of the markets previously studied in the realized volatility literature do have trading limits in place, such as the S&P 500, they are much less frequently invoked than in physical commodity markets. When the equilibrium price moves beyond the trading limits, trading ceases. Since no trades are recorded during these moves, trading limits naturally bias realized volatility downwards. Consequently, price limits should be taken into account in modeling realized volatility.

To investigate whether price limits may lower the effectiveness of realized volatility, price limit days must be identified at first. Unfortunately, the dataset used in this study does not provide information about price limit days.
A procedure is used in this article to mitigate the inadequacy of data. Three types of dates are defined as follows: (1) Touch days. A touch day occurs when the price limit is touched, but prices do not close at the limit. These are characterized by three conditions:

\[ |H_t - C_{t-1}| = \Delta_t \text{ or } |L_t - C_{t-1}| = \Delta_t; \Delta_t \geq 30 \text{ cents; } \text{Mod}(\Delta_t, 5) = 0 \]

where \( H_t \) and \( L_t \) are high and low prices at day \( t \); \( C_{t-1} \) is the closing price at day \( t-1 \); \( \Delta_t \) is the price change which must be greater than 30 cents; Mod denotes the modulus. (2) Closing days. On closing days, prices close at the limit price. They are identified using the three conditions for touch days plus the following condition:

\[ C_t = H_t \text{ or } C_t = L_t \]

(3) Limit move days. Limit move days occur when all trades occur at the limit move price. It satisfies all of the characteristics of closing days, plus

\[ O_t = C_t \]

where \( O_t \) denotes the opening price at day \( t \).

Based on the above definitions, 34 touch days, 11 closing days, and 0 limit move days are found in the current sample. The ratios of three different types of days to the total number of observations are 1.17%, 0.38%, 0% respectively.

To further check the effect of price limits on realized volatility, the relationship between realized volatility and conditional variances of the GARCH model is checked using the following equation,

\[
\sigma_{rv}^2 = a_0 + a_1 \sigma_{GARCH}^2 + a_2 D_{\text{touch}} + a_3 D_{\text{touch}} \sigma_{GARCH}^2 + \varepsilon
\]

where \( \sigma_{GARCH}^2 \) denotes conditional variances estimated from the GARCH (1,1) model; \( D_{\text{touch}} \) denotes dummy variables, \( D_{\text{touch}} = 1 \) if touch day occurs and \( D_{\text{touch}} = 0 \) otherwise.
The estimates of $a_1$, $a_2$ and $a_3$ are 0.402, 1.247 and -0.1 respectively. The first two estimates are highly significant and the estimate for $a_3$ is not significant. The regression results for closing days are similar to touch days except the fact that the estimate for the parameter $a_3$, -0.432, is significant. The positive sign of $a_1$ indicates that realized volatility and the conditional variances of the GARCH model have positive relationship. However, the negative sign of $a_3$ reveals that the GARCH model may have some information that realized volatility does not have.

Given the fact that detailed information on price limits are insufficient and the proportion of price limit days in the whole sample is small, price limits are not considered in the rest of this essay. However, the simple analysis conducted in this section demonstrates the worthiness to study price limits. From the regression results in this section, the effect of price limits on realized volatility in commodity markets is interesting for future research.

2.3 Distributions of Realized Volatility and Returns

2.3.1 Results of Existing Studies

Realized volatility computed from high-frequency intraday returns creates new opportunities to understand, analyze, and forecast market volatility. However, as briefly reviewed in the introduction, this approach is relatively new and much remains to be learned. Thus, it is natural to provide characterizations of both the unconditional and conditional distributions of realized volatility as the first step. Such backgrounds are necessary for a deep understanding of realized volatility measures. Meanwhile, the
characterization of the distribution of asset returns is important for risk management. For example, the distributional properties of volatility are relevant for pricing derivative instruments since different specifications will give different theoretical prices. As a second example, direct descriptions of volatilities and distributional properties facilitate constructing Value-at-Risk statistics.

Distributions of market volatilities have been studied extensively by researchers, traders and regulators. Clark (1973) uses a stochastic process to describe cotton futures volatility. He argues that the variance of the daily volatility is a random variable with a mean proportional to the mean number of daily transactions. He then proposes that the distribution for daily variance may be described as lognormal and that daily returns standardized by daily variance are conditionally normal. Given these two properties, cotton futures returns have been modeled by a leptokurtic mixture distribution with fat tails. The distributional assumptions advocated by Clark are named as the Mixture-of-Distributions-Hypothesis (MDH). Several later studies concentrate on testing this hypothesis. For example, Tauchen and Pitts (1983) study the price variability for 90-day T-bill futures. Again, the lognormal family of probability distributions is better than alternatives for empirical modeling of the price change. And the trading volume and the price change have a bivariate normal mixture distribution. However, as pointed out by Areal and Taylor (2002), “empirical investigation of Clark’s conjectures using daily returns has limited potential to provide decisive conclusions because daily volatility is then an unobservable latent variable.” Thus, the emergence of intraday realized volatility provides a promising way for modeling volatility distributions since daily volatility can be treated as observed by summing intraday returns sufficiently frequently.
ABDL (2000) study daily volatility of DM/$ and Yen/$ exchange rates by using 30-minute returns. They find that returns of these two major dollar exchange rates standardized by the realized volatilities are very nearly Gaussian. ABDL (2001) further explore the time series of 5-minute DM/$ and Yen/$ returns and find that the distributions of realized daily variances and standard deviations are skewed to the right and leptokurtic. In contrast, the distributions of log standard deviations are approximately Gaussian. They also confirm that strong volatility clustering exist in daily returns.

Ebens (1999) analyzes the Dow Jones Industrial Average (DJIA) index by focusing on 5-minute returns. The results regarding the distribution and persistency are remarkably similar to those obtained in ABDL (2001). In addition, he finds the asymmetric volatility effect, which is generally observed for equities. ABDE (2001) examine 30 individual stocks. For most of the stocks, the unconditional distributions of the realized daily variances are highly non-normal and skewed to the right and the logarithms of the realized volatility are approximately normal. Furthermore, the daily returns normalized by the realized standard deviations are also close to normal. Strong dynamic dependence exists for all stocks.

Areal and Taylor (2002) apply realized volatility to a futures market for the first time. They study 5-minute returns from FTSE-100 futures. They show that a long memory process describes the realized volatility very well. And the divergence of the distribution of the logarithm of variance and that of returns standardized by realized volatility from normal is minor. Thomakos and Wang (2003) consider four futures contracts: the Deutsche Mark, the S&P 500 index, US Bonds and the Eurodollar. They find that the distributions of the raw returns for all four futures contracts are highly non-Gaussian and
leptokurtic. After normalizing the raw returns by the realized standard deviations, the standardized return distributions are nearly Gaussian. Meanwhile, the distributions of realized volatilities are all highly non-Gaussian and skewed to the right. However, the distributions of the logarithmic standard deviations for all the four futures contracts become statistically indistinguishable from normal distributions. In addition, realized volatility exhibits long-memory dynamics consistent with a fractionally-integrated process.

All of the above results are generally consistent across assets and are quite compelling. Basically, they can be characterized by Clark’s conjectures. Although these findings provide support for the realized volatility approach, there is an apparent void in the existing literature, that is, no papers have studied futures contracts of physical commodities. Agricultural futures markets, especially grain futures, are an important component of the CBOT. The trading of agricultural futures contracts has some distinct properties. For example, according to the regulations of the Chicago Board of Trade (CBOT) the normal pit hours for financial and equity futures are from 7:20 a.m. to 2:00 p.m. and 7:20 a.m. to 3:15 p.m. respectively. By contrast, the market for CBOT grain futures contracts opens at 9:30 a.m. and closes at 1:15 p.m. For financial and equity futures, macroeconomic news play an important role in trading activities, and most news is reported at 8:30 a.m. Thus, U.S. macroeconomic news is announced during the trading hours of the financial and equity futures. The important reports for the grain markets are released when trading is closed. Andersen and Bollerslev (1998a) demonstrate that macroeconomic announcements have an important effect on daily volatility. Different timing may result in different effects on the distributional properties of the realized volatility.
volatility. Thus, measuring the realized volatility of agricultural futures contract will be a
good extension of and compliment to the existing studies.

2.3.2 Distributional Properties of Volatilities

2.3.2.1 Unconditional Distributions

In time series applications, volatility can be described by variance, standard deviation
or logarithmic variance/standard deviation. In previous sections, $\sigma_{rv}^2$ is used to present
realized volatility. To avoid any confusion caused by this notation, $\sigma_{rv}^2$ stands for realized
variance hereafter. $\sigma_{rv}$ and $\log(\sigma_{rv})$ denote the realized standard deviation and logarithmic
realized standard deviation, respectively. Table 2.1 provides summary statistics of the
unconditional distributions of the realized variances, standard deviations and logarithmic
standard deviations.

The first column of the first panel in Table 2.1 provides mean, variance, skewness
and kurtosis of the daily realized variance. The standard deviation indicates that the
realized daily volatilities fluctuate significantly through time. Moreover, it is obvious that
the distribution is right skewed with the coefficient 2.9108. Another obvious result is that
the distribution is extremely leptokurtic with the kurtosis coefficient 16.6940. These
observations are reinforced by the kernel density plot of Figure 2.5 and Q-Q plot of
Figure 2.6. The concave shape of the Q-Q plot indicates that the distribution of the series
is positively skewed with a long right tail. ABDL (2001) believe that the significant right
skewness and fatter tails are due to high serial dependence between intraday returns. Thus,
even sampled frequently, the realized daily variance is still far from Gaussian.
The summary statistics of the standard deviations also indicate that the distribution has fatter tails than the normal distribution and are skewed to the right. But both values are reduced, with the skewness coefficient 1.2613 and kurtosis coefficient 5.6652. The second panels of Figure 2.5 and Figure 2.6 provide visual proof of these reductions. Taken together, although it is a more easily explainable volatility proxy than the variance, the standard deviation still retains non-Gaussian properties.

Also shown in Table 2.1 are the distributional characteristics for the realized logarithmic standard deviation. The sample skewness and kurtosis coefficient are 0.0410 and 3.1106 respectively. Both numbers suggest the Normal distribution is a close approximation to the logarithmic standard deviation. This point is illustrated by the kernel density plot in the third panel of Figure 2.5. The graph appears symmetric. Furthermore, the Q-Q plot falls nearly on a straight line. All in all, the results obtained for the logarithmic standard deviation contrast sharply with those of realized variance and standard deviation.

The second panel of Table 2.1 displays the Jarque-Bera test statistic. The small p-values resulted from the realized variance and standard deviation series and the larger value from the logarithmic series reveal same information as discussed above.

In summary, the unconditional distributional characteristics of the realized variance, realized standard deviation and the logarithmic standard deviation for the soybean futures data are consistent with the findings from ABDL (2001) for the exchange rates, ABDE (2001) for the Dow Jones stocks and Thomakos and Wang (2003) for equity futures.
2.3.2.2 Temporal Dependence

The existence of volatility clustering at different frequencies has been extensively documented in the finance literature. This high degree of volatility persistence suggests that financial market volatility is highly predictable. Previous studies usually rely on ARCH models to estimate variance, which is a latent process. As realized volatility is a direct measure of actual market volatility, it provides straightforward descriptions of the conditional dependence in volatility.

Figure 2.7 displays the time series plots of the realized variance, realized standard deviation and the logarithmic standard deviation. The basic properties are in line with those implied by ARCH effects. The three volatility measures seem positively serially correlated and the strong persistence is evident for all three series. The visual impression of the strong clustering effect is confirmed by the highly significant Ljung-Box test statistics and small p-values reported in the upper panel of Table 2.2. The null hypothesis of no serial correlation is overwhelmingly rejected for all three series.

Figure 2.8 provides autocorrelation functions for the realized variance, standard deviation and the logarithmic standard deviation. Except reinforcing the persistent correlations of the three series, Figure 2.8 also indicates that autocorrelations of realized volatilities tend to exhibit slow, hyperbolic decay. For the variance series, the autocorrelation starts around 0.5 and decay very slowly to about 0.04 at lag 90. After that, the autocorrelations tend to fall in the 95% confidence interval. At the 200 day offset, the autocorrelation is 0.0311. Similarly, the realized standard deviation begins around 0.53 and decreases to about 0.04 at the 100-day displacement. However, at some lags after 100, autocorrelations fall out of the 95% confidence band. At lag 200, the autocorrelation is
0.0724, out of the confidence interval. Moreover, the decaying rate of the standard deviation is a little bit slower than that of the variance process.

The correlogram of the logarithmic standard deviation tells a quite different story compared with the first two series. First, the autocorrelations are systematically greater than the 95% confidence level. There is no point where the series tends to fall in the 95% interval. Second, it is obvious that the realized logarithmic standard deviation decays more slowly than the realized variance and standard deviation. Finally, even at a lag of 200, where the correlation value is 0.115, the series is still well above the 95% critical value.

The results outlined above are different from the findings of Ebens (1999), in which both the variance and standard deviation have autocorrelations above the 95% bands. The pattern of the standard deviation is similar to the logarithmic variance instead of the variance.

The low first-order autocorrelations of the three series may indicate that they do not exhibit unit-roots. However, the slow decay may suggest the presence of a unit root. The Augmented Dickey-Fuller test, allowing for a constant with 10 lagged difference terms, routinely and soundly rejects the unit-root hypothesis. As shown in the lower panel of Table 2.2, the test statistics are -7.0267, -7.1378 and -7.6654 for three volatility measures and the 1% and 5% critical values are -3.4324 and -2.8623. Based on these results, the autocorrelations of the three series are best characterized by long memory processes. This assertion is consistent with the existing findings in exchange rates, stock index, and equity futures.
Based on the results in 2.3.2.1 and 2.3.2.2, it is clear that the properties of the realized variance and standard deviations are similar in essence. For the distribution aspect, they are asymmetric and leptokurtic. For the temporal dependence, their correlograms demonstrate quite similar patterns. In contrast, the logarithmic variance is distributed approximately Gaussian, and its autocorrelations decay much more slowly.

### 2.3.3 Distributions of Standardized Returns

The daily returns series is constructed by taking the first difference of the last recorded prices for the whole sample. Then, this series is standardized to a mean zero process. The following analysis is based on the transformed returns, which are denoted as $r_t$. Figure 2.9 displays the time series plot of the daily returns. The first column of Table 2.3 presents a summary of the unstandardized returns. Consistent with previous findings, the unconditional distributions of the futures returns are approximately symmetric but highly leptokurtic, with the sample skewness -0.0341 and the sample kurtosis 5.5787. The upper panel of Figure 2.10 confirms this point.

In time series applications, the daily returns are always described as $r_t = \sigma_t \eta_t$, where $\eta_t \sim iid(0,1)$ and $\sigma_t$ is the conditional standard deviation of returns. If a time series model, ARCH or Stochastic Volatility model, is correctly specified, the standardized returns, $r_t / \sigma_t$, should account for the tail thickness. However, it is well known that the standardized returns from ARCH models still display large kurtosis. The second column of Table 2.3 provides the sample moments of the standardized returns. Specifically, this series is calculated by dividing the unstardardized returns by estimates of standard
deviations of Normal-GARCH (1,1) model. It is evident that the distribution appears fat-tailed, although the value of the sample kurtosis has decreased to 4.6241. Moreover, the skewness coefficient is positive, which indicates a right-skewed curve. Finally, the middle panel of Figure 2.10 provides visual impression of the above findings.

In contrast, the diagnostic statistics in the third column of Table 2.3 verifies that the distribution of the realized volatility standardized daily returns is closer to a standard normal. The coefficient of kurtosis has decreased to 3.7122, although it is still slightly leptokurtic. Furthermore, the Q-Q plot in Figure 2.10 is nearly a straight line except at the right end.

In summary, the values of mean, standard deviation, and skewness for the three volatility measures are close. The differences in the sample kurtosis are significant. The result from the realized volatility standardized returns stands in sharp contrast to the nonstandardized returns and GARCH standardized returns. The findings about the standardized returns largely correspond with extant literature. However, an interesting phenomenon is that the kurtosis coefficient of the standardized returns differs in value for different markets. For example, Ebens (1999) finds that the kurtosis estimate for DJIA is 2.75; ABDL (2000) determine the coefficients are 2.406 and 2.414 for DM/$ and Yen/$ respectively. Those values indicate that the distributions appear platykurtic. Thomakos and Wang (2003) find that the sample kurtosis coefficients for DM, E-Dollar and S&P futures are 3.0147, 3.1579, and 3.0521 and only T-bonds futures is less than 3.

All in all, except differences in magnitude with previous findings, the results in 2.3.2 and 2.3.3 confirm that the Mixture-of-Distributions-Hypothesis (MDH) advocated by Clark is valid in the soybean futures market.
2.3.4 Other Properties

2.3.4.1 ARCH-M Effect

In some financial applications, the expected return on an asset is related to the expected asset risk (volatility). This relationship has been widely explored in portfolio theory, for example, the Markowitz Portfolio analysis and the Capital Asset Pricing Model. Many argue that an increase in volatility results in an increase in the expected rate of return. Engel, Lilien and Robbins (1987) propose the ARCH in mean (ARCH-M) model, where the conditional variance is included into the mean equation. Two variants of the ARCH-M specification use the conditional standard deviation and the log variance (standard deviation) in place of the conditional variance.

Figure 2.11 displays the relationships between daily returns, $r_i$, and three different volatility measures. The straight line in each graph is obtained by regressing volatility on daily returns using least squares. Also shown in Figure 2.11 are the scatter plots of returns and three volatility measures. It reveals the non-linear relationship between returns and volatility. The three graphs have a similar pattern. The linear relationship between current daily returns and volatilities is not obvious. The coefficients of multiple determination, $R^2$, are 0.0014, 0.0009 and 0.0005 for three regression lines respectively. These diagnostic statistics confirm that impression. Thus, the ARCH-M effect can be ignored for current data.

2.3.4.2 Asymmetric Volatility

In 2.3.4.1, the relationship between volatility and current returns was investigated. In time series applications, the asymmetric response of volatility to past returns is also of
interest. French, Schwert and Stambaugh (1987) and Schwert (1990b) find that stock volatility is negatively related to stock returns. Nelson (1991) argues that for equities, negative returns are followed by higher volatilities than positive returns of the same magnitude. This phenomenon is known as “leverage effect”. In most applications, leverage effects have been probed for stock returns. For futures data, since the leverage hypothesis cannot be applied, the term “asymmetric volatility” is used to describe an asymmetry in the relation between volatility and returns. The Threshold ARCH (TARCH) model is introduced independently by Zakoian (1990) and Glosten, Jagannathan and Runkle (1991) to describe asymmetric volatility. The following regression models are based upon the specification for the conditional variance of TARCH models. They are used to investigate the asymmetric effects on the realized variance, the standard deviation and the log standard deviation.

\[
\sigma^2_{rv,t} = \sigma^2_{rv,t-1} + \alpha \sigma^2_{rv,t-1} + \beta r_{t-1}^2 + \gamma r_{t-1}^2 d_{t-1} + \epsilon_t \\
\sigma_{rv,t} = \sigma_{rv,t-1} + \beta r_{t-1} + \gamma r_{t-1} d_{t-1} + \epsilon_t \\
\log(\sigma_{rv,t}) = \log(\sigma_{rv,t-1}) + \beta r_{t-1} + \gamma r_{t-1} d_{t-1} + \epsilon_t
\]

(2.9a), (2.9b), and (2.9c)

where \( d_t = 1 \) if \( r_{t-1} < 0 \) and 0 otherwise.

In models (2.9a), (2.9b) and (2.9c), good news \( (r_{t-1} > 0) \), and bad news \( (r_{t-1} < 0) \), have different effects on the volatility. Good news has an impact of \( \beta \), while bad news has an impact of \( (\beta + \gamma) \). If \( \gamma \neq 0 \), asymmetric volatility exists.

Table 2.4 reports the regression estimates with their standard errors. All estimates are significant at 1% level. The \( \beta \) coefficients are positive and the \( \gamma \) coefficients are

---

5 See Schwert (1990b) for details about two hypotheses of the leverage effects.
negative for all three volatility measures. Thus, these results support the existence of asymmetric volatility. This parallels the findings of Thomakos and Wang (2003) for Deutsche Mark, Eurodollar and T-bonds futures contracts. And Ebens (1999) and ABDE (2001) all point toward the presence of asymmetries in stock returns.

However, although statistically significant for all three volatility measures, the $\gamma$ coefficients are smaller in absolute magnitude. To further check the asymmetric effect, the concept of the “News Impact Curve” is employed. Engle and Ng (1991) introduce a News Impact Curve with asymmetric response to good and bad news. One method to plot this curve is to use estimates of equations 2.9(a), 2.9(b) and 2.9(c) and fix last period’s volatilities $\sigma^2_{r_{j-1}}$, $\sigma_{r_{j-1}}$ and $\log(\sigma_{r_{j-1}})$ to the corresponding medians of $\sigma^2_r$, $\sigma_r$, and $\log(\sigma_r)$. Figure 2.12 displays the scatter plots for the volatilities against the lagged returns. The solid lines represent News Impact Curves, with different lines representing positive and negative returns. In each case, the lines are more steeply sloped for positive returns than negative returns.

### 2.4 Modeling Realized Volatility and Daily Returns

Based upon the fact that realized volatility can be treated as an observable process of the true market volatility, the findings in Section 2.3 can be used as inferences to select time series models which focus on depicting volatility. That is, the candidate time series models should account for the facts outlined in Section 2.3. The usefulness of existing time series models can be verified if the results from time series models correspond with the findings based on realized volatility.
2.4.1 ARMA and ARFIMA Models

As repeated in previous sections, realized volatility has several appealing properties. First, since realized volatility can be treated as observable, modeling realized volatility is equivalent to modeling return volatility process. Thus, it is not necessary to treat volatility as a latent process as in ARCH and stochastic volatility models. This idea greatly simplifies the estimation process. Second, according to the distributional properties described above, the log standard deviation is nearly Gaussian. Therefore, it is natural to use simpler time series models like the usual Gaussian ARMA family, to directly model the logarithmic standard deviations.

Also, from the stylized facts depicted above, the long memory characteristics of realized volatilities, realized standard deviations and logarithmic standard deviations are obvious. This is a well-known fact in existing literature. Related work includes ABDL (1999), Ebens (1999), Areal and Taylor (2002) and Thomakos and Wang (2003), among others. Given this fact, it is possible that standard parsimonious ARMA models may not account for the high order autocorrelations. Consequently, a fractionally integrated ARMA (ARFIMA) model may be applied to model the realized volatility series. In contrast to ARMA models, which explain the typical exponentially decaying autocorrelations, ARFIMA models capture the stylized fact of long memory processes, which have slower decay of the autocorrelations. Granger and Joyeux (1980) introduced the ARFIMA process. The general representation of the ARFIMA (p, d, q) model is,

\[ \phi(L_p)(1-L)^d y_t = \varphi(L_q) \xi_t \]  

(2.10)
where \( L \) is the lag operator; \( d < 1 \) is the integration parameter; \( p \) and \( q \) are order parameters; \( \phi(L_p) = \sum_{i=1}^{p} \phi_i L^i \) and \( \varphi(L_q) = \sum_{i=1}^{q} \varphi_i L^i \). This representation includes AR and MA processes as two special cases. When \( d = 1 \), equation (2.10) reduces to the ARIMA model.

### 2.4.2 ARMA and ARFIMA Model Estimation Results

Table 2.5 reports the estimation results of ARMA (5,4) and ARFIMA (5,d,4) models. These two models are selected as representatives of two types of competing families. Various models have been tested based on log likelihood values and Q-tests. For models, where \( p > 5 \) and \( q > 4 \), convergence becomes more difficult to obtain. Besides, the improvement of Q-test result is not obvious and sometimes even worse. For models, where \( p < 5 \) and \( q < 4 \), the Q-test results are not satisfactory. For example, for the ARFIMA (1,d,1) model the value of Q(55) is 92.478, with p-value 0.007 and the AIC (Akaike information criteria) is 0.2757; for the ARFIMA (2,d,2) model, the Q(55) is 90.549, with p-value 0.0066 and the AIC is 0.2765; for the ARFIMA (2,d,3) model, the Q(55) is 84.071 with p-value 0.0177 and the AIC is 0.2737; the corresponding values for ARFIMA (4,d,4) are 74.077, 0.0532 and 0.2718. The ARFIMA (5,d,4) model has p-value of 0.110 and AIC of 0.2709. Thus, ARFIMA (5,d,4) provides the most promising specification for characterizing long memory process. ARMA models have similar fitting results. Taken together, ARMA (5,4) and ARFIMA (5,d,4) balance the requirements of model fitting and parsimony.

All of the models (including ARFIMA and GARCH families) were estimated in OX. The corresponding packages include Arfima 1.0 and GARCH 3.0 (Doornik and Ooms (1998), Laurent and Peters (2002), and Doornik (2002)).
It is evident that the estimate for the fractional integration parameter is highly significant. The value of this parameter is 0.4593, which is consistent with the findings of ABDL (1999). ABDL (1999) suggest that the estimate of the fractional integration coefficient tends to be in the neighborhood of 0.4 for many realized volatilities.

Compared with the ARMA (5,4) model, the ARFIMA (5,4) model has several advantages. First, most estimates are statistically significant, which indicates good in-sample fitting. Second, the Q-test results are better. The p-values at Lag 35, 45, 55 and 65 are 0.006, 0.020, 0.060, 0.050 for the ARMA (5,4) model. And the corresponding values for the ARFIMA (5,4) model are 0.023, 0.065, 0.163 and 0.110. Finally, less obvious, the log likelihood represents a small improvement.

The higher order models described above are not coincidences. For example, Thomakos and Wang (2003) fit an ARFIMA (5,d,5) model to D/M, E-dollar, S&P 500 and T-bonds futures data. However, Ebens (1999) finds that ARFIMA (0,d,0) model is good enough to capture a long memory process for DJIA data. Since the history of realized volatility is relatively short, it is still uncertain to conclude whether the higher-order model reveals any special properties of futures data.

2.4.3 GARCH Models

Numerous findings have demonstrated the persistence of volatility in financial markets. Engel (1982) introduces ARCH model to capture this stylized fact. Bollerslev (1986) generalizes it to the GARCH case. \(^7\) The GARCH \((p,q)\) model is in the form of,

\(^7\) For an extensive review of ARCH and related models, see Bera and Higgins (1993).
\[ y_i = c + \varepsilon_i \quad (2.11) \]

\[ \varepsilon_i = \sigma_i \eta_i \]

\[ \sigma_i^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{i-1}^2 + \sum_{i=1}^{q} \beta_i \sigma_{i-1}^2 \]

where \( \eta_i \sim iid(0,1) \).

GARCH family has many extensions. Before choosing any model specifications, it is important to guarantee that candidate models can account for the facts outlined in Section 2.3. As far as modeling returns and volatilities, the main findings of previous section include: (1) returns are not correlated with current volatilities and thus the ARCH-M effect is not obvious; (2) the asymmetric effect exists for current data and thus EGARCH or TGARCH models may be considered; (3) volatility may be described as a long memory process. Based on these results, FIGARCH and FIEGARCH models are possible model specifications for the soybean data. FIGARCH-M model is also considered here to check whether this parametric model reflect the same pattern as the realized volatility approach.

Engel and Bollerslev (1986) consider a special class of GARCH models, in which

\[ \sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i = 1 \]. They name this type of models as integrated GARCH (IGARCH) models. As integrated in mean processes, a shock persists in the future variance for IGARCH processes. Baillie, Bollerslev, Mikkelsen (1996) extend IGARCH models to accommodate fractional integration. The new model is titled as FIGARCH. The variance equation of a FIGARCH \((p,d,q)\) model is,

\[ \sigma_i^2 = \omega[1 - \beta(L)]^{-1} + \{1 - [(1 - \beta(L))]^{-1} \phi(L)(1 - L)^d \} \varepsilon_i^2 \quad (2.12) \]
where \( L \) is the lag operator; \( d < 1 \) is the integration parameter; \( \alpha(L) = \sum_{i=1}^{p} \alpha_i L^i \),

\[ \beta(L) = \sum_{i=1}^{q} \beta_i L^i, \quad \text{and} \quad \phi(L) = (1 - \alpha(L) - \beta(L))(1 - L)^{-1}. \]

The FIGARCH-M model has the same variance equation as the FIGARCH model specified above. The only difference is that the mean equation becomes,

\[ y_t = c + \sigma_t^2 + \epsilon_t \quad (2.13) \]

Nelson (1991) proposes the EGARCH model to accommodate the leverage effect and the asymmetry in the conditional variance. Bollerslev and Mikkelsen (1996) refine the EGARCH model as follows,

\[ \ln \sigma_t^2 = \omega + [1 - \beta(L)]^{-1}[1 + \alpha(L)]g(\eta_{t-1}) \quad (2.14) \]

where \( g(\eta_t) = \theta \eta_t + \gamma \cdot E[\eta_t | -E | \eta_t] \).

Combining the idea of fractional integration with EGARCH type of model, Bollerslev and Mikkelsen (1996) introduce the FIEGARCH model. The variance equation of FIEGARCH \((p,d,q)\) is specified as follows,

\[ \ln \sigma_t^2 = \omega + \phi(L)^{-1}(1 - L)^{-d}[1 + \alpha(L)]g(\eta_{t-1}) \quad (2.15) \]

### 2.4.4 Model Estimation Results

For the soybean futures returns, \( y_t = r_t \) and the constant \( c \) is excluded from the mean equation since the daily returns has been demeaned prior to estimation. Table 2.6 reports the estimation results of FIGARCH \((1,d,1)\), FIGARCH-M \((1,d,1)\), and FIEGARCH \((0,d,1)\) models. For FIGARCH and FIGARCH-M models, the ARCH innovations \( \eta_t \) are
the Student-t density; for FIEGARCH models, the innovations are the generalized error distribution (GED).

Several regularities emerge from the estimates presented in Table 2.6. First, the parameter estimates of $d$ are highly significant for all models. For FIGARCH (1,d,1) and FIGARCH-M (1,d,1) models, the estimates are 0.4571 and 0.4585 respectively. Again, this magnitude is consistent with the assertion of ABDL (1999). However, Jin and Frechette (2004) show that the value of $d$ is 0.517 using the FIGARCH (1,d,1) model for soybean data, which is higher than the value obtained in this study. For the FIEGARCH model, the value of $d$ is much higher, 0.5650. This is in line with the one reported by Ebens (1999), who finds $d = 0.585$ for the FIEGARCH model for the DJIA data.

Second, consistent with the findings of previous sections, the ARCH-M effect is not obvious and thus can be ignored. This point is confirmed through comparison between the FIGARCH (1,d,1) and the FIGARCH-M (1,d,1) models. There are no significant differences in goodness-of-fit for both models as implied by several test statistics. For example, the likelihood values of the FIGARCH model and the FIGARCH-M model are -4402.162 and -4401.930 respectively. And the AIC (Akaike information criteria) and SC (Schwartz criteria) imply that the FIGARCH model is a little better than the FIGARCH-M. Specifically, the AIC are 3.0312 and 3.0316; the SC are 3.0413 and 3.0439. Taken together, FIGARCH-M model is not a good choice for modeling the daily soybean returns.

Finally, as expected, the estimates of the two additional parameters in the FIEGARCH model are highly significant. And the sign of $\theta$ is positive, which indicates that a positive correlation between the past return and subsequent volatility. This point is
consistent with the result in previous sections. Compared with the FIGARCH model, the FIEGARCH is a more promising GARCH specification for characterizing daily returns and volatilities. The addition of two parameters is not only highly significant from the log likelihood ratios but is also preferred from AIC and SC perspective.

All in all, the results from parametric models largely correspond with the findings in previous nonparametric analysis based on the realized volatilities. And the values of estimates are highly consistent with extant literature.

2.5 Conclusion

The use of intraday returns as a direct measure of market volatility is a relatively new field. Previous studies of realized volatility generally focus on equities or exchange rates. The point of this study is not to offer a new theoretical approach for modeling volatility, instead, it is to investigate the properties of realized volatility in the grain futures market.

The results indicate that realized volatility based on 5-minute returns largely correspond with existing literature. Specifically, the properties of the realized variance, the standard deviation and the log standard deviation have quite similar patterns as those observed in stock market or exchange rate market, although there are some discrepancies in magnitude. The findings of three volatility measures confirm that the Mixture-of-Distributions-Hypothesis (MDH) advocated by Clark (1973) is valid in the soybean futures market.

In contrast, the standardized daily returns display some different properties compared with stock and exchange rate data. The asymmetric effect exists and the news impact
curves are more steeply sloped to the right of the origin, which indicates that a positive correlation between the past return and subsequent volatility.

The long memory characteristics of realized volatilities, especially for the logarithmic standard deviations are obvious. Thus an ARFIMA model is used to describe the volatility process. And the results indicate that the ARFIMA model is better than the corresponding ARMA model.

Moreover, the parametric GARCH models, FIGARCH, FIGARCH-M and FIEGARCH, reflect the patterns described by nonparametric analysis. The implication of this conclusion is that the existing time series models can provide good in sample fits and may result in good forecasts since the validity and usefulness of realized volatility has been thoroughly explored.
<table>
<thead>
<tr>
<th></th>
<th>Realized Variance</th>
<th>Realized Std. Deviation</th>
<th>Logarithmic Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.8628</td>
<td>0.8686</td>
<td>-0.2071</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>0.7340</td>
<td>0.3294</td>
<td>0.3628</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.9108</td>
<td>1.2613</td>
<td>0.0410</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>16.6940</td>
<td>5.6652</td>
<td>3.1106</td>
</tr>
<tr>
<td>Jarque-Bera Test</td>
<td>26837.5600</td>
<td>1632.2530</td>
<td>2.2991</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.3168</td>
</tr>
</tbody>
</table>

Table 2.1: Summary Statistics of Daily Realized Variance, Realized Standard Deviation and Logarithmic Standard Deviation

<table>
<thead>
<tr>
<th></th>
<th>Realized Variance</th>
<th>Realized Std. Deviation</th>
<th>Logarithmic Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q-Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp; P-Value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 1</td>
<td>709.31</td>
<td>0.0000</td>
<td>834.18</td>
</tr>
<tr>
<td>Lag 5</td>
<td>2540.10</td>
<td>0.0000</td>
<td>3247.40</td>
</tr>
<tr>
<td>Lag 15</td>
<td>5197.80</td>
<td>0.0000</td>
<td>69996.7</td>
</tr>
<tr>
<td>Lag 20</td>
<td>5910.80</td>
<td>0.0000</td>
<td>8097.20</td>
</tr>
<tr>
<td>Lag 25</td>
<td>6419.80</td>
<td>0.0000</td>
<td>8995.50</td>
</tr>
<tr>
<td>Lag 30</td>
<td>6740.50</td>
<td>0.0000</td>
<td>9625.20</td>
</tr>
<tr>
<td>Lag 35</td>
<td>6945.50</td>
<td>0.0000</td>
<td>10094.00</td>
</tr>
<tr>
<td>ADF Test*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp; P-Value</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 1</td>
<td>-7.6654</td>
<td>0.0000</td>
<td>-7.1378</td>
</tr>
<tr>
<td>Lag 5</td>
<td></td>
<td></td>
<td>-7.0267</td>
</tr>
<tr>
<td>Lag 15</td>
<td></td>
<td></td>
<td>-7.0267</td>
</tr>
<tr>
<td>Lag 20</td>
<td></td>
<td></td>
<td>-7.0267</td>
</tr>
<tr>
<td>Lag 25</td>
<td></td>
<td></td>
<td>-7.0267</td>
</tr>
<tr>
<td>Lag 30</td>
<td></td>
<td></td>
<td>-7.0267</td>
</tr>
<tr>
<td>Lag 35</td>
<td></td>
<td></td>
<td>-7.0267</td>
</tr>
</tbody>
</table>

* The 1% and 5% critical values for the ADF test are -3.4324 and -2.8623.

Table 2.2: Ljung-Box and Augmented Dickey-Fuller Test Statistics

<table>
<thead>
<tr>
<th></th>
<th>( r_t ) *</th>
<th>( r_t / \sigma_{GARCH} )</th>
<th>( r_t / \sigma_{rv} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0000</td>
<td>0.0130</td>
<td>-0.0085</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>1.1942</td>
<td>1.0000</td>
<td>1.2634</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.0341</td>
<td>0.0015</td>
<td>-0.0497</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>5.5787</td>
<td>4.6241</td>
<td>3.7122</td>
</tr>
</tbody>
</table>

* The return series is multiplied by 100.

Table 2.3: Descriptive Statistics for Daily Returns
The table reports the ordinary least squares regression for the model (9). Standard errors are based on Newey-West heteroscedasticity and autocorrelation consistent estimators.

* All estimates are significant at 1% level.

Table 2.4: News Impact Function Estimates

<table>
<thead>
<tr>
<th></th>
<th>ARMA</th>
<th>ARFIMA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>d-parameter</strong></td>
<td></td>
<td>***0.4593</td>
</tr>
<tr>
<td><strong>AR-1</strong></td>
<td>***-1.0170</td>
<td>0.2335</td>
</tr>
<tr>
<td><strong>AR-2</strong></td>
<td>***0.6735</td>
<td>0.2023</td>
</tr>
<tr>
<td><strong>AR-3</strong></td>
<td>***1.1362</td>
<td>0.1941</td>
</tr>
<tr>
<td><strong>AR-4</strong></td>
<td>0.1532</td>
<td>0.2261</td>
</tr>
<tr>
<td><strong>AR-5</strong></td>
<td>***-0.0626</td>
<td>0.0306</td>
</tr>
<tr>
<td><strong>MA-1</strong></td>
<td>1.2801</td>
<td>0.2322</td>
</tr>
<tr>
<td><strong>MA-2</strong></td>
<td>-0.1898</td>
<td>0.2586</td>
</tr>
<tr>
<td><strong>MA-3</strong></td>
<td>*-0.9004</td>
<td>0.1013</td>
</tr>
<tr>
<td><strong>MA-4</strong></td>
<td>-0.2391</td>
<td>0.1853</td>
</tr>
<tr>
<td><strong>Q-test &amp; P-value</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 35</td>
<td>47.400</td>
<td>0.006</td>
</tr>
<tr>
<td>Lag 45</td>
<td>55.504</td>
<td>0.020</td>
</tr>
<tr>
<td>Lag 55</td>
<td>61.754</td>
<td>0.060</td>
</tr>
<tr>
<td>Lag 65</td>
<td>74.516</td>
<td>0.050</td>
</tr>
</tbody>
</table>

Log likelihood

-386.822

**Table 2.5: ARMA and ARFIMA Model Estimate**

* *, **, *** significant at 10%, 5% and 1%.
<table>
<thead>
<tr>
<th></th>
<th>FIGARCH (1,d,1)##</th>
<th>FIGARCH-M (1,d,1)</th>
<th>FIEGARCH (0,d,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega$</td>
<td><strong>0.1101</strong></td>
<td><strong>0.1094</strong></td>
<td>0.1304</td>
</tr>
<tr>
<td>$\alpha$</td>
<td><em>0.1313</em>*</td>
<td><em>0.1305</em>*</td>
<td>0.0668</td>
</tr>
<tr>
<td>$\beta$</td>
<td>*<strong>0.5688</strong></td>
<td>*<strong>0.5694</strong></td>
<td>0.7728</td>
</tr>
<tr>
<td>$d - \text{Figarch}$</td>
<td>*<strong>0.4571</strong></td>
<td>*<strong>0.4585</strong></td>
<td>0.5650</td>
</tr>
<tr>
<td>ARCH-in-mean</td>
<td>-0.0096</td>
<td>0.0139</td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td></td>
<td>0.0406</td>
<td>0.131</td>
</tr>
<tr>
<td>$\gamma$</td>
<td></td>
<td>*<strong>0.1089</strong></td>
<td>0.0221</td>
</tr>
<tr>
<td>AIC</td>
<td>3.0311</td>
<td>3.0316</td>
<td>3.0276</td>
</tr>
<tr>
<td>SC</td>
<td>3.0413</td>
<td>3.0439</td>
<td>3.0399</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-4402.162</td>
<td>-4401.930</td>
<td>-4396.144</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.0078</td>
<td>-0.0090</td>
<td>-0.0492</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.7133</td>
<td>4.7148</td>
<td>4.5396</td>
</tr>
<tr>
<td>Q-test &amp; P-Value#</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 10</td>
<td>11.975</td>
<td>11.993</td>
<td>11.921</td>
</tr>
<tr>
<td>Lag 15</td>
<td>16.132</td>
<td>16.093</td>
<td>15.854</td>
</tr>
</tbody>
</table>

*, **, *** significant at 15%, 5% and 1%.

# Q-test for the standardized residuals.

# #The ARCH innovations $\eta_t$ are conditional on the student-t distribution for the FIGARCH and FIGARCH-M models and the innovations are conditional on GED for the FIEGARCH model.

Table 2.6: FIGARCH, FIGARCH-M and FIEGARCH Estimation Results
Figure 2.1: Volatility Signature Plot for Different Time Intervals

Figure 2.2: 5-Minute Return Series
Figure 2.3: Autocorrelations for 5-minute Return Series

Figure 2.4: Realized Volatility and Squared Daily Returns

(Solid line denotes squared daily returns and dashed line denotes realized volatilities.)
Figure 2.5: Kernel Density Estimates
Figure 2.6: Q-Q Plots of Daily Realized Volatilities

(x-label denotes volatility quantile and y-label denotes Normal quantile for each plot)
Figure 2.7: Time Series Plots of Daily Realized Volatilities
Figure 2.8: Autocorrelation Functions of Daily Realized Volatilities
Figure 2.9: Time Series Plot of Daily Returns
Figure 2.10: Q-Q Plots of Daily Returns

(x-label denotes return quantile and y-label denotes Normal quantile for each plot)
Figure 2.11: Realized Volatilities Vs. Daily Returns
Figure 2.12: News Impact Curve
CHAPTER 3

FORECASTING DAILY VOLATILITY
USING RANGE-BASED DATA

3.1 Introduction

Users of agricultural markets always need to establish accurate representations of expected future volatility. For example, future volatility is the main ingredient in calculating optimal hedge ratios. The application of misspecified future volatility has the potential to induce inappropriate assessment of asset risk and portfolio selection. Thus, not surprisingly, seeking improved forecasts of agricultural market price volatility has drawn increased attention from financial academics and practitioners.

On the one hand, the existence of volatility clustering at different frequencies has been extensively documented in the finance literature. This high degree of volatility persistence suggests that financial market volatility is predictable. On the other hand, forecasting the future level of volatility is challenging for several reasons. For example, volatility is not directly observable; therefore the choice of evaluation metric for forecasting performance is uncertain. Establishing an appropriate framework for volatility forecasting is an important theme for financial academics and is of great relevance to practitioners.
The fact that range-based volatility estimators are highly efficient has been acknowledged by many authors (For example, Parkinson (1980), Garman and Klass (1980), Beckers (1983), Ball and Torous (1984), Kunitomo (1992), and Yang and Zhang (2000)). Their findings raise the question as to whether the forecasting ability of extant ARCH models can be improved with range-based data. However, these works only focus on constructing efficient volatility estimators and little attention is paid to the application of these estimators.

Schwert (1990a), Gallant, Hsu and Tauchen (1999), and Chou (2001) adopt range-based estimators in the extant time series models and find that the forecasting ability of these models is significantly improved. However, these earlier works do not exploit the essential properties of range-based estimators. Moreover, the formats of range-based volatility proxies in these earlier studies are different. Thus it is difficult to conclude which format has appealing statistical properties.

It was not until Alizadeh, Brandt and Diebold (2002) that the usefulness of a simple volatility proxy, the log range, was formally established and applied to time series models. They clarify that the log range, defined as the log of the difference between the high and low log prices during the day, is nearly Gaussian, robust to microstructure noise and much less noisy than alternative volatility measures such as log absolute or squared returns. Compared with earlier studies, their work fully exploits the distributional properties of the log range estimators and thus provides a theoretical underpinning for using the Gaussian ARMA class of models.

As applications of range-based time series models are not extensively explored, it remains unclear whether range-based volatility proxies have practical uses in asset
markets. Thus it is important to investigate the forecasting performance of range-based models. Motivated both by the appeal of range-based models and by the practical need to check their forecasting ability in agricultural commodity futures market, this study will investigate whether ARCH models extended with the range data provide better out of sample forecasts of daily volatility and whether a simple ARMA model of the log range has competitive forecasting ability. More importantly, this study will check the forecasting performance under different criteria using realized volatility.

The remainder of this paper is organized as follows. Section 3.2 reviews range-based estimators and describes the properties of the log range proxy. Section 3.3 presents the time series models used in this study. Section 3.4 summarizes the data and the in sample fit of the models. Section 3.5 presents the forecasting results and compare results under different criteria. Section 3.6 concludes.

3.2 Range-based Volatility Estimators

In recent times, two volatility proxies, realized volatility and range-based volatility, have been increasingly used to address a range of issues in financial markets. Since Andersen and Bollerslev (1998) proposed the concept of realized volatility, several papers have shown that vast improvements can be made to volatility forecasts by using intraday data (For example, Ebens (1999), Andersen, Bollerslev, Diebold and Ebens (2001), Andersen, Bollerslev, Diebold, and Labys (2001)). Although high-frequency data contain more information about financial markets, microstructure effects may plague volatility estimates and forecasts when high-frequency data is used. Problems may be more serious for intraday data, since bid-ask bounce and non-synchronous trading may
create spurious autocorrelation and thus distort the realized volatility estimator.

Christoffersen (2002) shows that the realized volatility is a very noisy estimator if trading is infrequent or the bid-ask spread is large. Alizadeh, Brandt and Diebold (2002) demonstrate that due to price discreteness and bid-ask bounce biases, the performance of realized volatility deteriorates sharply as the return interval shrinks. Building on theoretical and Monte Carlo analysis, they claim that a volatility proxy based on daily price range is much more robust to microstructure noise than realized volatility measures.

Furthermore, intraday data have only recently become readily available and data on most securities are only available at daily or lower frequencies (Forsberg and Bollerslev, 2001). In contrast, high and low prices have long been recorded along with the daily closing and opening prices for many assets. The access to intraday high and low prices readily results in the application of the simple range-based volatility proxy.

### 3.2.1 Review of Range-based Volatility Estimators

In recent years, there has been renewed interest in using range-based volatility estimators. This is not surprising. Since the price range utilizes information from the entire sample path in contrast to estimators based only on closing prices, it is expected to be superior. The application of the price range as a volatility proxy dates to Parkinson (1980). Under the assumption that price is governed by a lognormal diffusion and there are no opening jumps or drifts, Parkinson (1980) proposes a volatility estimator using the high and low prices only. Define $H_t$ as the highest price and $L_t$ as the lowest price on day $t$. Parkinson’s estimator is
Parkinson (1980) shows that the range-based estimator is approximately 5 times more efficient than the return-based estimator, \( r^2 \), where \( r \) is the first order difference of logarithmic close prices. Garman and Klass (1980) extend Parkinson’s estimators and conclude that the efficiency gain from including opening and closing prices along with high and low prices is significant. Define \( p_t^o \) as the opening price and \( p_t^c \) as the closing price on day \( t \), the simple version of Garman and Klass’ estimator (Bollen and Inder, 2002) is

\[
\hat{\sigma}_{gk} = 0.5\left(\ln(H_t/L_t)\right)^2 - 0.39\left(\ln(p_t^r/p_t^o)\right)^2
\]

Beckers (1983) applies Parkinson’s estimator on actual market data and finds that it has a downward bias when step size is finite. Contrary to Parkinson (1980), Kunitomo (1992) assumes that the geometric Brownian motion may have nonzero drift terms and proposes a new method for estimating the volatility parameter. Ball and Tourus (1984) derive a maximum likelihood equivalent to Garman and Klass’ estimator. Yang and Zhang (2000) present a new unbiased estimator based on multiple periods of open/close/high/low prices. Their estimator is independent of both the drift and opening jumps assumption.

The literature described above focuses exclusively on constructing the volatility estimator and improving its efficiency. Little attention has been paid to the application of these estimators. And since the Parkinson and Garman-Klass estimators are quite complicated in comparison with the standard volatility proxies, such as log absolute or squared returns, the application of the range-based measures is very limited even though they are highly efficient. Moreover, the literature surveyed above assumes constant
volatility of financial asset returns. However, it is widely known today that volatility is
time varying and predictable (Andersen and Bollerslev, 1997). Finally, the performance
of range-based estimators in the context of forecasting the asset return volatility is not
investigated in the above studies.

3.2.2 Properties of the Log-Range Volatility Proxy

Schwert (1990a) applies the spread (range), defined as the difference between the
highest and lowest log prices over a fixed sampling interval, to two simple forecasting
frameworks. He finds that the range data do not help predict stock returns but do add
significant information in predicting volatility. Gallant, Hsu and Tauchen (1999) prove
the usefulness of range data in a stochastic volatility framework. Chou (2001) constructs
the conditional autoregressive range (CARR) model and finds that the CARR model is
superior in volatility forecasting. Although they provide promising results, none of these
studies addresses the essential properties of the range-based estimators.

In Alizadeh, Brandt and Diebold (2002), the properties of the log range volatility
proxy are extensively explored. They claim that the log range has three important
properties. First, the log range is much more efficient than the standard volatility proxies,
the log absolute or squared returns--the measurement error associated with the log range
is about one quarter of the error of the standard volatility estimators. Second, the log
range is nearly Gaussian. Finally, under certain conditions, the realized volatility is more
efficient than the range. Bollen and Inder (2002) claim that these conditions are based on
restrictive assumptions about the data generating process for intraday returns. For
example, the discretely sampled returns are serially uncorrelated, the sample path for
volatility is continuous and the returns can be sampled sufficiently frequently. If these conditions do not hold, the realized volatility estimator is erratic. In contrast, the range estimator is less likely to be seriously contaminated by microstructure noises such as bid-ask spreads and price discreteness.  

The above three properties raise the question of whether this simple range-based volatility proxy is useful in practice. Brandt and Jones (2002) examine the forecasting performance of EGARCH models in combination with the log range volatility proxy. Their study suggests substantial gains from the use of range data.

3.3 Time Series Models

3.3.1 GARCH Models

There are two major categories of time-varying volatility models, the ARCH family and the stochastic volatility (SV) family. The autoregressive conditional heteroscedasticity (ARCH) model was introduced by Engle (1982). Compared with previous econometric models, ARCH processes are specifically designed to model and forecast conditional variances. This property of the ARCH model makes it appealing for modeling the volatility of economic time series. Bollerslev (1986) proposed an extension of the conditional variance function and introduced the generalized ARCH (GARCH) model. For many applications, the GARCH (1,1) model has been proved to be a parsimonious representation that fits data well. The representation of the GARCH (1,1) is

---

8 Lildholdt (2002) derives the joint density of normalized close, high, and low prices and applies this result to GARCH models. However, the efficiency gain from combining the range with the open and close prices is minor and unnecessary (Alizadeh, Brandt and Diebold, 2002).
\[ y_i = c + \varepsilon_i \]  
\[ \varepsilon_i = \sigma_i \eta_i \]
\[ \sigma_i^2 = \omega + \alpha \varepsilon_{i-1}^2 + \beta \sigma_{i-1}^2 \]

where \( \eta_i \) is a mean-zero, unit-variance, i.i.d. random variable.

Engle and Kraft (1983) were the first to consider the effect of ARCH (GARCH) on forecasting. Akgiray (1989) was the first to apply the GARCH model to forecast volatility. After that, numerous papers have employed ARCH (GARCH) models for forecasting. However, the forecasting performance of the ARCH family is disappointing in many studies. Different conclusions have been drawn for different sample periods and different speculative markets. Andersen and Bollerslev (1998b) provide a few insights into these results. They reveal that the coefficient of multiple determination, \( m^2 \), is low when the daily squared returns are used as a measure of ex-post volatility. Realized volatility, the sum of intraday squared returns, is a much more efficient volatility proxy. Andersen and Bollerslev (1998b) find that ARCH family performs better if the ex-post volatility is estimated by the sum of intraday squared returns.

**3.3.2 GARCH Models Extended with Additional Information**

Andersen and Bollerslev (1998b) demonstrate that the daily squared return is a very noisy estimator. If the previous trading day is quite volatile, but the closing price happens to be the same as the opening price, the lagged daily squared return would be zero. Thus by extending the daily GARCH model with information related to the real volatility dynamics, the new model will provide a reasonable explanation that the previous day was
volatile. The conditional variance equation may be extended to allow for the inclusion of additional regressors. For example, Bessembinder and Seguin (1993) include daily volume. Laux and Ng (1993) include the number of price changes. Taylor and Xu (1997) and Martens (2002) use intraday returns. Martens (2001) use the daily range. The specification of the variance equation for the extended GARCH models is,

\[
\sigma_i^2 = \omega + \alpha \epsilon_{i-1}^2 + \beta \sigma_{i-1}^2 + \zeta I_{i-1}
\]

(3.4)

where \( I_i \) presents any trade related variables such as the traded volume, the sum of squared intraday returns or the daily range.

### 3.3.3 Simple Regression Model

It would be interesting to explore whether alternative volatility proxies, such as the log range and squared intraday returns, fit the class of ARMA models. The logic is that if an estimator is highly efficient, it is possible to extract valuable information about the future value of volatility by just using simple technique. Given the findings described by Alizadeh, Brandt and Diebold (2002), it is natural to assume that the log-range process falls within the Gaussian ARMA models. If true, this will greatly reduce the computational costs. Standard forecasting techniques may be applied to generate predictions of future log range. Through simple transformations, the forecasts of volatility can be obtained. Specifically, the SR model for the range data is,

\[
R_i = \alpha_1 R_{i-1} + \alpha_2 R_{i-2} + ... + v_i
\]

(3.5)

where \( v_i \sim iid(0,1) \); \( R_i \) denotes the log range.
3.4 Data and In-Sample Fit

3.4.1 Data Description

The data set consists of Chicago Board of Trade (CBOT) soybean futures intraday transaction prices and daily prices. The daily data were obtained from the Bridge/CRB Futures Database. The sample consists of daily soybean futures high/low/closing prices from January 2, 1985 to July 31, 2001. The intraday data are time and sales transaction prices, which were obtained from the Futures Industry Institute. The full sample covers the period January 2, 1990 to July 31, 2001. The first 1,264 trading days (January 2, 1985—December 29, 1989) are used to estimate the parameters of the various models. The next 2,909 trading days (for which intraday data are available) are used to test the out-of-sample forecasting performance.

In calculating the returns series, the nearby contracts are used to construct the continuous returns series. However, returns are calculated from the second nearby contract when the nearby contract is in the delivery month. This switch guarantees that returns are nearly always calculated from the prices of the contract that has the highest trading volume.

Figure 3.1 plots the prices of the futures data. Figure 3.2 plots the returns series actually used in this study, which is $100 \times \ln(P_t / P_{t-1})$ of the futures data. Table 3.1 reports summary statistics for daily returns. Soybean futures returns conform to several stylized facts which have been extensively documented for financial variables. The distribution of the returns is almost symmetric and has fat tails and a substantial peak at zero. Excess kurtosis of the series indicates that the distribution of daily returns is far
from Gaussian. The autocorrelations of returns are close to zero. The Q-statistics are smaller than the critical values at 5% level. In contrast, the squared returns are significantly autocorrelated. Figure 3.2 reflects another stylized fact, the clustering effect. Variances of returns change over time and large (small) changes tend to be followed by large (small) changes.

Since the price path is not observable when the market is closed. Garman and Klass (1980) suggest a method to mitigate the effect of discontinuous observations. Given the idea suggested by Garman and Klass (1980), the daily range is defined as,

$$\text{Range}_t = \text{Max}(h_t, c_{t-1}) - \text{Min}(l_t, c_{t-1})$$

(3.6)

where $h_t$ and $l_t$ denote highest and lowest prices on day $t$ respectively and $c_{t-1}$ represents the closing price on day $t - 1$. Since the current soybean daily data only cover the full floor trading from 9:30 a.m. to 1:15 p.m., equation (3.6) captures information about the overnight market activity. While some of the markets previously studied in the literature do have trading limits in place, such as the S&P 500, they are much less frequently invoked than in physical commodity markets. When the equilibrium price moves beyond the trading limits, trading ceases. Since no trades are recorded during these moves, equation (3.6) provides a reasonable range proxy for limit move days. Following Alizadeh, Brandt and Diebold (2002), the daily log range is defined as,

$$R_t = \log[\log(\text{Max}(h_t, c_{t-1})) - \log(\text{Min}(l_t, c_{t-1}))]$$

(3.7)

The volatility literature primarily uses absolute or squared returns as volatility proxies. To justify the superior efficiency of the log range, Table 3.2 presents descriptive statistics for log absolute returns and the log range. Firstly, the log range is preferable in terms of its smaller standard deviation. Secondly, the skewness and kurtosis of the log range are
0.2405 and 2.9312, respectively. These values are closer to the corresponding values of 0 and 3 for a normal random variable compared with those for log absolute returns. This conclusion is confirmed by checking the Jarque-Bera statistic. It is more obvious by looking at Figure 3.3, which shows the quantile-quantile (Q-Q) plot. The Q-Q plot for the log range falls nearly on a straight line and indicates that the log range has a distribution close to Normal. In contrast, the Q-Q plot of the log absolute returns curves downward at the left end and upward at the right. Finally, the log range proxy is superior in terms of its time series dynamics. The large and slowly decaying autocorrelations of the log range clearly manifest strong volatility persistence. The erratic fluctuation of log absolute returns masks the volatility persistence.

3.4.2 In-Sample Fit

3.4.2.1 Estimation of GARCH Models

Maximum likelihood estimation of the GARCH model is easy to implement once the density function of $\varepsilon_t$ is specified. If the residuals are not conditionally normally distributed, a quasi-maximum likelihood (QML) estimator will still be consistent provided that the mean and variance functions are correctly specified.

The seasonal effects of price volatility are widely documented in many surveys. In time series modeling, one can take care of seasonality first and fit a model with the deseasonalized data. Or a model can be estimated for seasonally unadjusted data by adding a seasonal component in the model. This study follows the second approach. Roberts (2001) models the seasonal effects in volatility by including a Fourier expansion.
for the intercept of the GARCH volatility equation. The specification of the GARCH model is thus of the form, \(^9\)

\[
100 \times \ln(P_t / P_{t-1}) = \mu + \epsilon_i
\]

\[
\epsilon_i = \sigma_i \eta_i
\]

\[
\sigma_i^2 = \omega_i + \alpha \epsilon_{i-1}^2 + \beta \sigma_{i-1}^2
\]

\[
\omega_i = \kappa + \sum_{m=1}^{M} [\phi_m \sin(2m\pi \tau) + \psi_m \cos(2m\pi \tau)] \quad 0 \leq \tau \leq 1
\]

where \(\tau\) denotes the time of year of the observation.

Estimation results (based on 1985-1989 daily data) for GARCH (1,1) models are given in Table 3.3. All of the specifications capture well the autocorrelation in the volatility of returns. For the GARCH (1,1) model without seasonality, the estimates of parameters \(\alpha\) and \(\beta\) are highly significant. The persistence in volatility is quite large, with \(\alpha + \beta\) larger than 0.98. The Ljung-Box portmanteau test statistic for up to tenth order serial correlation in the standardized residuals \(\eta_i\) takes the value \(Q(10) = 8.0294\), which is not significant for the \(X^2_{10}\) distribution. However, the Q-tests suggest that there exists serial dependence in the squared residuals at lag 5 and lag 10. Furthermore, the unconditional sample kurtosis for the residuals is 4.0142, which exceeds the normal value of three. And the residuals continue to display asymmetry.

The second set of results in Table 3.3 includes a first order seasonal expansion. Only one of the two seasonal parameters is significant at 5% level. However, the LR test

\(^9\) The GARCH (1,1) model is considered to be a parsimonious representation, since results not reported here show that higher orders have nothing extra to offer.
statistic equals 9.04, which is significant at 2.5% level in the corresponding asymptotic 
\( \chi^2 \) distribution. The addition of two parameters is also preferred from an AIC (Akaike 
information criteria) perspective. Moreover, the inclusion of a Fourier series reduces the 
sample kurtosis in residuals.

For the second order seasonality, only two seasonal parameters are significant at 5% 
level although the LR test and AIC prefer the inclusion of two additional parameters. 
Including a third order seasonality is rejected not only from a LR test perspective but also 
from a t-test perspective.

Note, the estimated value for \( \alpha \) decreases as more parameters are added into the 
variance equation. This indicates that the reliance of the conditional volatility on the 
previous date is reduced. From the above results, a first order seasonality model 
represents a reasonable tradeoff between the need of model fit and the need of parsimony. 
Also the GARCH (1,1) model with no seasonality is used as the reference basis for 
forecast analysis.

3.4.2.2 GARCH Models Extended with Daily Range

In this section, the daily soybean data are fitted to the GARCH models extended with 
daily range. The intercept of the variance equation is defined as,

\[
\omega_t = \kappa + \zeta I_{t-1} + \sum_{m=1}^{M} \phi_m \sin(2m\pi \tau) + \psi_m \cos(2m\pi \tau) \] 

(3.9)

where 
\( I_t = Max(h_t, c_{t-1}) - Min(l_t, c_{t-1}) \).

Table 3.4 presents the estimation results. First, the log-likelihood values of four 
range-based GARCH models are greater than those of the corresponding GARCH models
in section 3.4.2.1. This result suggests that range data improve in-sample model fitting and reflects the greater precision of the range as a volatility proxy. Second, range-based GARCH models are also desirable from the AIC and SC (Schwartz criteria) perspectives since both values of AIC and SC fall as the daily range is added in. Third, the estimates of $\alpha$ become insignificant at the 1% level or even at the 5% level for second order and third order seasonality. In contrast, the estimates of the range parameter $\zeta$ are highly significant. This result is foreseeable since $\varepsilon^2_{t-1}$ and $I_{t-1}$ are competing factors to represent last period’s variance and the inclusion of range data reduces the importance of $\varepsilon^2_{t-1}$ in accounting for last period’s volatility. This result also confirms the fact that daily range is a relatively less noisy volatility proxy than daily squared returns.

Omitting range, the estimation results for the range-based GARCH models are quite similar to the results for the GARCH models in terms of seasonality. AIC, SC and LR tests all suggest that range-based GARCH (1,1) with third order seasonality is readily rejected. The addition of one order of seasonality performs better than the second order seasonality model in terms of LR test. This conclusion is confirmed by the estimates of second order seasonality model. $\phi_2$ is significant at 5% level and $\psi_2$ is not significant judged by the standard errors. Additionally, the values of AIC are close for both models. Finally, the skewness and kurtosis coefficients of the standardized residuals for the three models do not provide much information about model selection. The p-values for the Q-test are also close and tell a similar story for the four models. All in all, the first order seasonality model works best for range-based GARCH models. However, in order to
avoid the possible over-fitting problem in forecasting, the range-based GARCH model without seasonality is also included as one of the forecasting frameworks.

3.4.2.3 Range-based ARMA Models

Based on equation (3.5), different orders of ARMA models have been fitted to the log range data. Standard linear Gaussian modeling techniques suggest that a low-order ARMA (1,1) model provides good in-sample fit. Since the model selection procedures are analogous to those in the previous sections, the results are not reported in detail.

From the results of Alizadeh, Brandt and Diebold (2002),

\[ R_t = \ln(\sigma_t) + E[\ln R_t^*] \]  \hspace{1cm} (3.10)

\[ R_t = \ln(\sigma_t) + (0.43 + \ln(\sqrt{1/252})) \]  \hspace{1cm} (3.11)

where \( R_t \) denotes the daily log range, defined as in equation (3.7). \( R_t^* \) denotes the standardized process with volatility=1. From equation (3.11), volatility forecasts can be readily derived based on the forecasts of daily log range.

3.5 Out-of-Sample Daily Volatility Forecasts

3.5.1 Forecast Evaluation Criteria

It is difficult to compare the forecasting performance of competing models since there is a variety of evaluation criteria used in the literature. Statistical analysis is one of the evaluation measures frequently used. Poon and Granger (2001) suggest that utility-based economic criteria are costly to apply and statistical analysis provides a practical way for forecast evaluation. West and Cho (1995) consider alternative statistical measures.
Basically, statistical measures evaluate the difference between forecasts at time $t$ and realized values at time $t + k$. However, asset price volatility is not directly observable and measuring the realized values of volatility is challenging. Much effort has been devoted to extracting volatility from other observable market activities. The daily squared return has been widely used in the literature as ex-post volatility. However, Andersen and Bollerslev (1998b) show that it is a very noisy volatility estimator and does not provide reliable inferences regarding the underlying latent volatility in daily samples. They introduce a new volatility measure, termed realized volatility. Realized volatility estimates volatility by summing squared intraday returns. Volatility estimates so constructed are close to the underlying integrated volatility. Thus, the volatility of a price process can be treated as an observable process. In this study, realized volatility is calculated based on 5-minute return series.

For the performance evaluation, two forecast evaluation criteria, Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), are defined by,

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (\sigma_{t}^{^2} - \sigma_{rv,t}^{^2})^2}$$

(3.12)

where $T$ denotes the forecast horizon. $\sigma_{t}^{^2}$ denotes one step ahead daily forecast and $\sigma_{rv,t}^{^2}$ denotes realized volatility.

$$MAE = \frac{1}{T} \sum_{t=1}^{T} | \sigma_{t}^{^2} - \sigma_{rv,t}^{^2} |$$

(3.13)
The second metric used to evaluate daily volatility forecasts is the regression-based method. The coefficient of determination ($m^2$) of the regression of realized volatility on forecasted volatility results from,

$$\sigma_{\text{rv},t}^2 = \varphi_0 + \varphi_1 \hat{\sigma}_t^2 + e_t$$  \hspace{1cm} (3.14)

### 3.5.2 Results

Table 3.5 reports the out of sample forecasts based on evaluation criteria RMSE, MAE, and $m^2$. The forecasts are based on parameter estimates from rolling samples with fixed sample size of 1264 days.

A number of conclusions may be drawn. Noticeably, range-based ARMA (1,1) models provide poor out of sample forecasts compared with GARCH models. Results are qualitatively consistent across three different statistical measures. Evidence on the GARCH models has several implications. First, the GARCH (1,1) model is inferior to the other three GARCH models: the first order seasonality GARCH (1,1) model, the GARCH (1,1) model extended with range data, and the first order seasonality GARCH model extended with daily range. The daily GARCH (1,1) model has the smallest regression $m^2$ and highest values for RMSE, MAE among the four GARCH models. Second, the regression based method and summary statistics both suggest that GARCH (1,1) models extended with the difference between daily high and low are better than the use of GARCH models ignoring daily range. Third, including seasonality improves the out of sample forecasts of the daily GARCH (1,1) model. The coefficient of determination $m^2$, RMSE and MAE all point out that. Fourth, interestingly, the first order seasonality
GARCH model extended with daily range is not the best model based on $m^2$ and MAE. The first order seasonality GARCH model extended with daily range has an $m^2$ of 0.2370, whereas the GARCH (1,1) model extended with daily range has an $m^2$ of 0.2756. The MAE is 0.6515 for the GARCH (1,1) model extended with daily range, whereas it is 0.6569 for the first order seasonality GARCH model extended with daily range. The use of Fourier series does not lead to a superior forecasting performance for the extended GARCH (1,1) models.

### 3.6 Conclusion

Previous studies reveal that the range-based volatility estimator is highly efficient. However, little attention is paid to the application of these estimators. This paper compares the performance of GARCH models, range-based GARCH models, and log-range based ARMA models in terms of their forecasting abilities. Realized volatility is used as the forecasting evaluation criterion. In-sample fitting results reveal that range-based GARCH models perform better than standard GARCH models. Out-of-sample results reveal that the log-range based ARMA models may have disadvantages in comparison to GARCH models extended with daily ranges. The empirical analysis here also makes the following points. For forecasting soybean futures market volatility it is important to include the daily range in the GARCH models. However, for the extended GARCH models, the adding of seasonality becomes less important. Since the effectiveness of forecast estimates will differ depending on the forecast horizon and forecast error measures, whether similar findings arise from other applications of range data is a question for further research.
### Summary Statistics for Soybean Futures Returns

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.0097</td>
<td>1.3233</td>
<td>-0.4045</td>
<td>6.6273</td>
<td>-8.5892</td>
<td>5.7397</td>
</tr>
</tbody>
</table>

### Q-Test Results

<table>
<thead>
<tr>
<th>Lags</th>
<th>Returns</th>
<th>Squared Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q-Statistics</td>
<td>P-Value</td>
</tr>
<tr>
<td>Lag 1</td>
<td>0.332</td>
<td>(0.565)</td>
</tr>
<tr>
<td>Lag 2</td>
<td>0.340</td>
<td>(0.844)</td>
</tr>
<tr>
<td>Lag 5</td>
<td>5.936</td>
<td>(0.313)</td>
</tr>
<tr>
<td>Lag 10</td>
<td>12.306</td>
<td>(0.265)</td>
</tr>
<tr>
<td>Lag 15</td>
<td>19.861</td>
<td>(0.177)</td>
</tr>
</tbody>
</table>

Table 3.1: Summary Statistics for Soybean Futures Returns \(100 \times \ln(P_t / P_{t-1})\)
<table>
<thead>
<tr>
<th></th>
<th>Log Absolute Returns</th>
<th>Log Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.4717</td>
<td>-4.2941</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1.0154</td>
<td>0.5600</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.5119</td>
<td>0.2405</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.0743</td>
<td>2.9312</td>
</tr>
<tr>
<td>Jarque-Bera Statistics &amp; P-Value</td>
<td>54.22   (0.000)</td>
<td>12.43 (0.002)</td>
</tr>
<tr>
<td>Autocorrelations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 1</td>
<td>0.170</td>
<td>0.504</td>
</tr>
<tr>
<td>Lag 2</td>
<td>0.122</td>
<td>0.488</td>
</tr>
<tr>
<td>Lag 5</td>
<td>0.207</td>
<td>0.486</td>
</tr>
<tr>
<td>Lag 10</td>
<td>0.155</td>
<td>0.427</td>
</tr>
<tr>
<td>Lag 20</td>
<td>0.148</td>
<td>0.392</td>
</tr>
</tbody>
</table>

Table 3.2: Summary Statistics for Soybean Futures
Log Absolute Returns and Log Range
### Table 3.3: GARCH Estimation Results

<table>
<thead>
<tr>
<th></th>
<th>No Seasonality</th>
<th>First Order Seasonality</th>
<th>Second Order Seasonality</th>
<th>Third Order Seasonality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
<td>Std. Error</td>
</tr>
<tr>
<td>( \mu )</td>
<td>-0.0167</td>
<td>0.0271</td>
<td>-0.0183</td>
<td>0.0268</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>***0.0885</td>
<td>0.0150</td>
<td>***0.0786</td>
<td>0.0139</td>
</tr>
<tr>
<td>( \beta )</td>
<td>***0.9020</td>
<td>0.0149</td>
<td>***0.9085</td>
<td>0.0151</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>**0.0176</td>
<td>0.0079</td>
<td>**0.0220</td>
<td>0.0088</td>
</tr>
<tr>
<td>( \phi_1 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \psi_1 )</td>
<td>**-0.0170</td>
<td>0.0071</td>
<td>**-0.0137</td>
<td>0.0069</td>
</tr>
<tr>
<td>( \phi_2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \psi_2 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi_3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \psi_3 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>1.5456</td>
<td>1.5436</td>
<td>1.5422</td>
<td>1.5420</td>
</tr>
<tr>
<td>SC</td>
<td>1.5537</td>
<td>1.5558</td>
<td>1.5584</td>
<td>1.5624</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.17813</td>
<td>-0.22178</td>
<td>-0.22772</td>
<td>-0.23814</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.0142</td>
<td>3.9193</td>
<td>3.8482</td>
<td>3.7216</td>
</tr>
<tr>
<td>Q-test P-values#</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 1</td>
<td>0.2238</td>
<td>0.5243</td>
<td>0.1350</td>
<td>0.6166</td>
</tr>
<tr>
<td>Lag 2</td>
<td>0.4769</td>
<td>0.6369</td>
<td>0.3266</td>
<td>0.6930</td>
</tr>
<tr>
<td>Lag 3</td>
<td>0.2674</td>
<td>0.7768</td>
<td>0.1958</td>
<td>0.7926</td>
</tr>
<tr>
<td>Lag 5</td>
<td>0.4171</td>
<td>0.1028</td>
<td>0.3461</td>
<td>0.0691</td>
</tr>
<tr>
<td>Lag 10</td>
<td>0.6260</td>
<td>0.0946</td>
<td>0.5894</td>
<td>0.0787</td>
</tr>
</tbody>
</table>

*, **, *** Significant at the 10%, 5%, and 1% level, respectively.  # P-values for \( \eta \) and \( \eta^2 \).
<table>
<thead>
<tr>
<th></th>
<th>No Seasonality</th>
<th>First Order Seasonality</th>
<th>Second Order Seasonality</th>
<th>Third Order Seasonality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
<td>Estimate</td>
<td>Std. Error</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.0005</td>
<td>0.0270</td>
<td>-0.0003</td>
<td>0.0270</td>
</tr>
<tr>
<td>$\alpha$</td>
<td><strong>0.0547</strong></td>
<td>0.0242</td>
<td><strong>0.0391</strong></td>
<td>0.0175</td>
</tr>
<tr>
<td>$\beta$</td>
<td>***0.8604</td>
<td>0.0376</td>
<td>***0.8641</td>
<td>0.0265</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.0000</td>
<td>0.0215</td>
<td>0.0000</td>
<td>0.0153</td>
</tr>
<tr>
<td>Range</td>
<td>***0.0125</td>
<td>0.0040</td>
<td>***0.0144</td>
<td>0.0042</td>
</tr>
<tr>
<td>$\phi_1$</td>
<td>-0.0048</td>
<td>0.0087</td>
<td>0.0046</td>
<td>0.0085</td>
</tr>
<tr>
<td>$\psi_1$</td>
<td><strong>-0.0236</strong></td>
<td>0.0112</td>
<td><strong>-0.0249</strong></td>
<td>0.0124</td>
</tr>
<tr>
<td>$\phi_2$</td>
<td>-0.0136</td>
<td>0.0082</td>
<td><strong>-0.0215</strong></td>
<td>0.0108</td>
</tr>
<tr>
<td>$\psi_2$</td>
<td>0.0094</td>
<td>0.0091</td>
<td>0.0116</td>
<td>0.0093</td>
</tr>
<tr>
<td>$\phi_3$</td>
<td>0.0114</td>
<td>0.0093</td>
<td>0.0076</td>
<td>0.0086</td>
</tr>
<tr>
<td>$\psi_3$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>1.5374</td>
<td>1.5350</td>
<td>1.5346</td>
<td>1.5348</td>
</tr>
<tr>
<td>SC</td>
<td>1.5476</td>
<td>1.5492</td>
<td>1.5529</td>
<td>1.5572</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-1938.3298</td>
<td>-1933.1845</td>
<td>-1930.7353</td>
<td>-1928.9753</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.14806</td>
<td>-0.18658</td>
<td>-0.19939</td>
<td>-0.20334</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>3.7991</td>
<td>3.6683</td>
<td>3.6506</td>
<td>3.5381</td>
</tr>
<tr>
<td>Q-test P-values#</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag 1</td>
<td>0.1963</td>
<td>0.5978</td>
<td>0.1138</td>
<td>0.6012</td>
</tr>
<tr>
<td>Lag 2</td>
<td>0.4320</td>
<td>0.5765</td>
<td>0.2861</td>
<td>0.5941</td>
</tr>
<tr>
<td>Lag 3</td>
<td>0.3017</td>
<td>0.7389</td>
<td>0.2104</td>
<td>0.7500</td>
</tr>
<tr>
<td>Lag 5</td>
<td>0.4655</td>
<td>0.0141</td>
<td>0.3697</td>
<td>0.0069</td>
</tr>
<tr>
<td>Lag 10</td>
<td>0.6568</td>
<td>0.0180</td>
<td>0.6150</td>
<td>0.0118</td>
</tr>
</tbody>
</table>

*, **, *** Significant at the 10%, 5%, and 1% level, respectively. # P-values for $\eta$ and $\eta^2$.

Table 3.4: Estimation Results of GARCH Extended with Daily Range
<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
<th>MAE</th>
<th>$m^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH (1,1)</td>
<td>1.0183</td>
<td>0.7122</td>
<td>0.2450</td>
</tr>
<tr>
<td>First Order Seasonality GARCH (1,1)</td>
<td>0.9906</td>
<td>0.7081</td>
<td>0.2466</td>
</tr>
<tr>
<td>GARCH (1,1) Extended with Daily Range</td>
<td>0.8731</td>
<td>0.6515</td>
<td>0.2756</td>
</tr>
<tr>
<td>First Order Seasonality GARCH (1,1) Extended with Daily Range</td>
<td>0.8718</td>
<td>0.6569</td>
<td>0.2370</td>
</tr>
<tr>
<td>Range-based ARMA (1,1)</td>
<td>2.2416</td>
<td>0.9417</td>
<td>0.0120</td>
</tr>
</tbody>
</table>

Note: $m^2$ denotes the coefficient of multiple determination.

Table 3.5: Daily Volatility Forecast Performance
Figure 3.1: Soybean Futures Prices (1985/01-1989/12)

Figure 3.2: Soybean Futures Returns (1985/01-1989/12)
Figure 3.3: Q-Q Plots of Log Range and Log Absolute Return
CHAPTER 4

ESTIMATION OF THE OPTIMAL COMMODITY HEDGE RATIO

4.1 Introduction

The availability of intraday data has provided a convenient and efficient way to model market volatility. Realized volatility can be constructed by summing squares of intraday high-frequency returns. Volatility estimates so constructed are close to the underlying true volatility. A number of papers have implemented this approach. However, most studies focus on examining the properties of realized volatility. For example, Andersen, Bollerslev, Diebold and Labys (2001) examine exchange rates, Andersen, Bollerslev, Diebold and Ebens (2001) examine stocks, and Areal and Taylor (2002) examine stock index futures. These studies have demonstrated that realized volatility has many appealing properties and can be treated as a proxy of market volatility. However, these findings are essentially statistical. To quote Fleming, Kirby and Ostdiek (2003) (FKO, hereafter),\(^\text{10}\)

“A separate question is whether the gains in precision are sufficient to have a meaningful impact on decisions that depend on conditional volatility estimates. Presumably, applications such as risk management should benefit because performance in this context depends largely on the statistical properties of the estimates. It is not clear, however, whether using realized volatility

\(^{10}\) Page 474.
leads to more accurate option prices or better investment management decisions. Perhaps standard volatility models provide a sufficient representation of volatility dynamics for these purposes so that switching to realized volatility yields only small benefits."

If the realized volatility approach can yield substantial economic benefits, it will hold promise for practical values of realized volatility in asset pricing and financial risk management applications. In Chapter 2, realized volatility estimators based on intraday soybean data are examined. Statistical tests apparently confirm that they are useful in modeling the conditional distribution of soybean futures returns. As noted above, however, whether realized volatility can be treated as a successful risk management tool demands more empirical evidence.

Furthermore, end-users of agricultural markets may not see additional benefits from intraday models when compared to extant daily or weekly models. Information and transaction costs may make intraday portfolio adjustments impractical. Thus, an appropriate approach is to extend the existing daily or weekly models with intraday information.

The purpose of this study is to investigate the usefulness of realized volatility in the context of hedging. Specifically, this study presents an analysis of incorporating information contained in intraday prices into multivariate GARCH models and evaluates the statistical performance of the augmented GARCH models. Then, effectiveness of different optimal hedge ratios depending on different estimation procedures is compared by examining the in-sample and out-of-sample performance of the ratios.

For users of agricultural futures markets who have a long position in a physical commodity, the conventional recommendation is to take a short position of equal size in the futures market. However, many studies suggest that a hedge ratio of one is not always
optimal. A hedge ratio that is less than one indicates that participants in futures markets may not be fully hedged. Johnson (1960) depicts a risk-minimizing (or minimum variance) optimal hedge ratio as the ratio of the covariance between cash and futures returns to the variance of the futures returns. Benninga, Eldor, and Zilcha (1984) demonstrate that this risk-minimizing hedging rule is also a utility-maximizing optimum under the assumption that futures markets are unbiased.

Many empirical studies use the simple regression approach to estimate optimal hedge ratios, in which the slope coefficient of an OLS regression of cash returns on futures returns is equivalent to the optimal ratio. The advantage of this approach is its ease of implementation. However, this simple technique has several problems. For example, Myers and Thompson (1989) point out that the OLS model does not take account of the conditioning information required by optimal hedging. Cecchetti, Cumby and Figlewski (1988) (CCF, hereafter) claim that simple OLS ignores the time-varying nature of optimal hedge ratios. Thus, the estimates of optimal futures hedges based on simple regressions are of limited value.

As an alternative to simple regression, ARCH models are used by CCF (1988) to estimate time-varying optimal futures hedges. Whereas they address the time-varying dependencies, their approach assumes a constant correlation coefficient between cash and futures returns. Baillie and Myers (1991) employ multivariate GARCH models to accommodate a time-varying conditional covariance matrix of portfolio returns while relaxing the constant correlation assumption. The multivariate GARCH model successfully overcomes the shortcomings of the simple regression approach and thus is treated as an appropriate framework for this study.
This paper has five sections. The next section reviews the concept of the optimal hedge ratio. Section three introduces the realized volatility based bivariate GARCH model and compares it to the existing bivariate GARCH model. Both models are fitted to soybean cash and futures data. The fourth section calculates different optimal hedge ratios and compares their performance. The economic value of the augmented GARCH model is investigated. The final section offers a summary and conclusion.

4.2 The Optimal Hedge Ratio

4.2.1 The Minimum Variance Hedge Ratio

When an agent takes both spot and futures positions, the agent has to decide the role that hedging plays in his overall market activities and choose a hedging strategy that reflects his objectives. A major problem for the agent is to choose an optimal hedge ratio (OHR). The hedge ratio is defined as the ratio of the size of the position taken in futures contracts to the size of the cash position.

The theory on the optimal futures hedge traces its origin to Telser (1955) and Johnson (1960). Telser (1955) asserts that if traders participate in both cash and futures markets, they can always find an optimal hedging strategy if they follow a safety-first rule of decisions. He believes that knowing the form of probability distributions of spot and futures prices is not a necessary requirement to derive these optimal rules. Instead, it is the second moments of the probability distributions of spot and futures prices that determine the optimal rules.

Johnson (1960) reformulates the theory of hedging. He proposes that hedging activities are motivated primarily by the desire to reduce risk, but hedgers also attempt to
maximize the expected profits simultaneously. Thus, a hedger always has a desire to find an optimal hedge ratio in a mean-variance framework. Although the hedging and speculation model used in Johnson (1960) is in a mean-variance framework, he actually only derives a risk-minimizing hedge. As a result, the optimal size of a futures position in a ratio with the given holdings in a cash market is measured by the covariance between futures returns and cash returns relative to the variance of futures returns. Johnson also suggests a method to evaluate the effectiveness of this optimal futures hedge, which is interpreted as the percent reduction in the variance of his portfolio.

Telser and Johnson argue that the traditional view of hedging as taking a position in a futures market that is equal in magnitude but of opposite sign to a position in the spot market is not always correct. The traditional one-to-one (or naïve) hedge will not minimize risk if futures and cash prices do not move together precisely. Telser and Johnson emphasize that it may be optimal under some circumstances for participants in futures markets to hedge only partially. Both of them attempt to set up analytical models of optimal hedging. However, neither of them tests their models in a futures market.

Ederington (1979) applies the methods advocated by Johnson to four futures markets: the 8% GNMA (Government National Mortgage Association) certificate, T-bill, wheat and corn. More specifically, she calculates the risk-minimizing optimal hedge ratios and evaluates the corresponding hedging effectiveness for two-week and four-week hedges. A striking result is that the proportion hedged is significantly different from one even for pure risk-minimizers. This is certainly consistent with Telser and Johnson’s conclusion. The empirical results also suggest that the reduction in variances associated with the
optimal future hedges is significant in magnitude for the GNMA, wheat and corn futures markets.

There is a distinction between the studies of Rolfo (1980), Anderson and Danthine (1981) and those of Johnson (1960) and Ederington (1979). Rolfo (1980) and Anderson and Danthine (1981) assume that a hedger’s utility is determined by the expected return and variance of his portfolio. Thus, they calculate optimal futures hedges that maximize traders’ expected utility. In contrast, Johnson (1960) and Ederington (1979) only focus on minimizing risk. Rolfo (1980) applies his models to the world cocoa market. The results also suggest that the ratio of the optimal hedge is not unity, which is consistent with the conclusion based upon risk-minimizing futures hedges.

Advocates of utility-maximizing hedge ratios argue that a hedge ratio which does not account for expected returns cannot be optimal. Since the relation between risk and return is a tradeoff, hedgers always take into consideration this tradeoff when they make hedging decisions. Thus, an optimal hedging rule must be in a mean-variance framework and maximize the expected utility which is defined by return and variance. The traditional theory by Johnson (1960) and related empirical studies, for example, Ederington (1979), are rigorously challenged.

Benninga, Eldor, and Zilcha (1984) (BEZ, hereafter) show that the risk-minimizing hedge ratio is optimal if futures markets are assumed unbiased, or the expected returns to taking a futures position are zero. Their conclusion is general in the sense that this optimum does not depend on any assumptions about the hedger’s utility function or the degree of risk aversion. If the assumption of unbiasedness of futures markets is valid, the minimum variance hedge position is always optimal. The BEZ theoretical result has been
reinforced by many empirical studies. For instance, Hudson, Leuthold, and Sarassoro (1987) examine futures price changes for wheat, soybean and live cattle futures markets. Their results demonstrate that futures prices change effectively in response to the inflow of new information and thus these three futures markets can be treated as unbiased.

Based on the above discussion, the remainder of this article will focus on estimation of the risk-minimizing hedge ratio and treat it as the optimal hedge ratio.

4.2.2 Alternative Methods for Estimating the Optimal Hedge Ratio

To implement optimal hedge ratios, the relevant covariance and variance must be estimated. The conventional approach in previous literature is to regress historical series of spot price changes or returns on those for the futures contract and use the slope coefficient as a measure of the optimal hedge ratio (For example, Ederington (1979), Carter and Loyns (1985), Brown (1985)). Several researchers have noted that this simple regression approach is not appropriate for optimal hedge ratio estimation.

The simple regression approach has two shortcomings. One is that this standard technique is based on the information of unconditional moments. Myers and Thompson (1989) show that the covariance and variance in the optimal hedging rule should be conditional moments because traders utilize all market information available at the time they make their hedging decisions. However, the simple regression approach provides only a ratio of the unconditional covariance between the spot and futures returns to the unconditional variance of the futures returns. Myers and Thompson test their assertion in corn, soybean and wheat futures markets. Their results suggest that simple regression approaches using price levels or returns are misleading.
The other problem of the standard regression approach, as suggested by Myers (1991), is that it implies a time-invariant hedge ratio. CCF (1988) reveal that the joint distribution of cash and futures prices changes over time and the optimal hedging strategy should reflect this fact. The conventional OLS method, however, does not consider the problem of time-varying distributions and thus a time-invariant hedge ratio is unsatisfactory. CCF (1988) use ARCH (3) models to estimate optimal hedge ratios, which take proper account of relevant conditioning information and time variation in the joint distribution of cash and futures returns. A shortcoming of their study is that they assume that the conditional correlations between cash and futures returns are constant.

Realizing the deficiencies of previous studies, Baillie and Myers (1991) use bivariate GARCH models to characterize the conditional joint distribution of cash and futures prices and apply the corresponding results to estimate optimal hedge ratios. Their approach has two desirable properties. First, bivariate GARCH models provide convenient assumptions about the conditional density of cash returns or futures returns. The second advantage of bivariate GARCH models is that they allow for the direct estimation of a time-varying covariance matrix. As an illustration, Baillie and Myers describe how the optimal hedge ratios could be estimated for six different commodities. Their results indicate that the improved optimal hedge ratios based on bivariate GARCH models perform best in in-sample and out-of-sample periods.

Following Baillie and Myers, several researchers improve the basic GARCH framework for optimal hedging. For example, Haigh and Holt (2002) develop an approach which combines hedging strategies in a dynamic programming (DP) framework with GARCH specifications. Their model allows for optimal updating variance and
covariance over a multi-period hedging horizon. In contrast, a hedge is not updated in the basic GARCH framework for each horizon. In this study, however, the DP-GARCH approach is not employed in spite of its theoretical advantage. Haigh and Holt (2002) focus on 4-week hedging horizon. The DP-GARCH optimally updates the hedge each week whereas the basic GARCH uses a constant hedge on each week for a given hedging horizon. Haigh and Holt believe that the DP-GARCH is superior in that it adjusts to short-run market volatility. Since the current study is based on daily and intraday data instead of weekly data, it is supposed to strengthen the standard GARCH in characterizing short-run volatility.

Because of its practicality and ease of use, the research on optimal hedging has relied heavily on the multivariate GARCH model. For example, Park and Switzer (1995) and Tong (1996) study stock index futures; Kroner and Sultan (1993) and Brooks and Chong (2001) study foreign currency futures; Gagnon and Lypny (1995) study interest rate futures; Baillie and Myers (1991), Myers (1991), Garcia, Roh, and Leuthold (1995), Haigh and Holt (2000), and Moschini and Myers (2002) study commodity futures. As such, this study will use the methodology developed by Baillie and Myers (1991) to estimate time-varying optimal hedge ratios and extends their work by adding intraday information.

4.3 Estimation of Bivariate GARCH models

4.3.1 Data and Basic Statistics

In this study, the agent is assumed to have a perpetual stock of soybeans. Without loss of generality, his stock is normalized to 5,000 bushels (1 contract). Suppose the agent
participates in both cash and futures markets. The object of the agent is to hedge his long position in the cash market by selling soybean futures. At any one time, soybean futures contracts with different maturities are traded at the Chicago Board of Trade. The agent is assumed to use the nearby contract and rolls over to the next nearest contract in the delivery month. Similar assumptions can be found in Brown (1985) for wheat, corn, and soybeans, García, Roh and Leuthold (1995) for the soybean complex and Moschini and Myers (2002) for corn. Moreover, the agent is assumed to adjust his portfolio on a daily basis to incorporate changes of information in both markets. Therefore daily data are used in this study.

Futures data consist of Chicago Board of Trade (CBOT) soybean futures intraday transaction prices and daily prices.\textsuperscript{11} The futures daily return is computed from the change in the logarithmic closing prices. The cash price is the official Illinois river daily cash price. The daily cash return is the change in the logarithmic price. The data span from January 1990 to July 2001, with a length of time more than 10 years, and a total of 2909 daily observations. This sample period is split into two parts. The first part is for estimation and in-sample performance evaluation, which runs from January 1990 to July 2000. The post-sample period from August 2000 to July 2001 is used to examine model performance. The agent is assumed to hold 5,000 bushels of soybeans continuously over the whole sample period.

Statistics of augmented Dickey-Fuller (ADF) tests are shown in Table 4.1. Both series, the log of cash price and the log of futures price, are first examined for existence of a unit root. The ADF tests fail to reject the null hypothesis of a unit root at any significance.

\textsuperscript{11} For a complete description of intraday and daily futures data, see Chapter 2, 2.2.2 and Chapter 3, 3.4.1.
levels. Then the two return series, changes in the logarithm of either cash or futures, are examined for the same test. The result suggests that both return series are stationary.

The sample autocorrelations of cash and futures returns up to 15 lags are calculated. Almost all autocorrelation coefficients are within the 95% bounds. This indicates that the coefficients are not significantly different from zero at 5% significance level. Thus, no evidence for the presence of autocorrelation is found.

Table 4.2 gives basic descriptive statistics of returns for each year and the full sample. Time variation in the standard deviation of returns and the correlation between the returns series is obvious. This observation suggests that the assumption of constant joint distribution of cash and futures returns is not appropriate. For the full sample and for all sub-samples, the mean returns for both cash and futures are close to zero. The zero mean return of futures suggests that the soybean futures market in this study can be treated as unbiased.

### 4.3.2 Bivariate GARCH Models

Since constructing optimal hedge ratios relies on the conditional distributions of cash and futures returns, knowledge of these distributions is very important. According to Bera and Higgins (1993), multivariate GARCH models are useful in estimating several coefficients related to risk management, for example, the beta coefficient and the hedge ratio. The multivariate GARCH modeling technique provides a way to characterize the joint distribution of cash and futures prices and then a time-varying variance and covariance matrix of portfolio returns is readily available. Therefore period-by-period
estimates of the time-dependent optimal hedge ratios are attainable by retrieving elements from the variance and covariance matrix.

Because a fully specified bivariate GARCH model has 21 parameters, restrictions have to be imposed on parameters to make estimation easier. There are numbers of specifications of the bivariate GARCH model. A common specification is the diagonal representation GARCH model, introduced by Bollerslev, Engle and Wooldridge (1988). The diagonal representation assumes that each element in the conditional variance matrix depends only on its own past values and squared residuals. As addressed by Bera and Higgins (1993), one disadvantage of diagonal models is that the covariance matrix is not always positive definite, and therefore the numerical optimization of the likelihood function will fail. Another popular model specification is proposed by Baba, Engle, Kraft and Kroner (1989). The BEKK model is expressed as,

\[ y_t = \mu + \epsilon_t, \]
\[ \epsilon_t | \Omega_{t-1} \sim N(0, H_t) \]
\[ H_t = A'A + C'C + D'\epsilon_{t-1}\epsilon_{t-1}'D \]

where \( y_t \) is a (2×1) vector; \( H_t \) is a (2×2) conditional covariance matrix; \( \Omega_{t-1} \) is the information available at time \( t - 1 \); \( A, C, D \) are (2×2) matrices, with \( A \) being symmetric and \( C \) and \( D \) are unrestricted matrices. Due to the quadratic nature of the terms in the variance equation, positive definiteness is guaranteed.

To further alleviate the burden of model estimation, additional restrictions may be imposed on the parameters in (4.1). A more parsimonious representation can be obtained by imposing the diagonal assumption on the parameter matrices. In other words, by assuming that matrices \( C \) and \( D \) are symmetric, the BEKK-diagonal model specification
can be established. This reduces the number of conditional covariance matrix parameters to 9, and therefore $A$, $C$, and $D$ have 3 parameters each.

### 4.3.3 Bivariate GARCH Models for Soybean Data

#### 4.3.3.1 Basic BEKK GARCH Model

The following equations present the BEKK-diagonal bivariate GARCH(1,1) model used in this study,

$$100 \times r_t = \varepsilon_t$$

$$\varepsilon_t | \Omega_{t-1} \sim N(0, H_t)$$

$$H_t = A' A + C' H_{t-1} C + D' \varepsilon_{t-1} \varepsilon_{t-1}' D$$

where $r_t = (s_t, f_t)'$ is a $(2 \times 1)$ vector containing cash and futures returns. $A, C, D$ are $(2 \times 2)$ symmetric matrices. Other variables are as defined previously. Specifically, the variance equation is expressed as,

$$H_t = \begin{bmatrix} a_1 & a_2 \\ a_2 & a_3 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ a_2 & a_3 \end{bmatrix} + \begin{bmatrix} c_1 & c_2 \\ c_2 & c_3 \end{bmatrix} \begin{bmatrix} h_{11,t-1} & h_{12,t-1} \\ h_{12,t-1} & h_{22,t-1} \end{bmatrix} \begin{bmatrix} c_1 & c_2 \\ c_2 & c_3 \end{bmatrix} + \begin{bmatrix} d_1 & d_2 \\ d_2 & d_3 \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t-1}^2 & \varepsilon_{1,t-1} \varepsilon_{2,t-1} \\ \varepsilon_{1,t-1} \varepsilon_{2,t-1} & \varepsilon_{2,t-1}^2 \end{bmatrix} \begin{bmatrix} d_1 & d_2 \\ d_2 & d_3 \end{bmatrix}$$

#### 4.3.3.2 Augmented BEKK GARCH Model

By accounting for intraday information and seasonality, an augmented GARCH model with realized volatility and seasonal factors is introduced. The resulting variance equation is thus complicated compared with the basic GARCH model. To add the second moment covariates, $H_t$ is extended as,
In this implementation, $A$, $B$, $C$ and $D$ are constrained to be symmetric matrices, and therefore have 3 parameters each. The covariates must be added in such a way as to preserve the positive definite of $H_i$. The matrix $B$ only appears if covariates are to be added in the second moment, and it must also be positive definite in order to guarantee that $H_i$ remains positive definite.

Given

$$B_i = \begin{bmatrix} \beta_{1,i} & \beta_{2,i} \\ \beta_{2,i} & \beta_{3,i} \end{bmatrix}$$

(4.5)

each element of $B_i$ is a linear function,

$$\beta_{i,j} = I_{i,j} b_i$$

(4.6)

where $I_i$ is a $(T \times K)$ matrix, $b_i$ is a $(K \times 1)$ vector. Then second moment covariates are available and $H_i$ remains positive definite. Specifically, by adding seasonality and intraday information, equation (4.6) can be expressed as,

$$\beta_{i,j} = b_{i,1} \sin(2\pi \tau) + b_{i,2} \cos(2\pi \tau) + b_{i,3} \sigma_{rv,i}$$

(4.7)

where $\tau$ denotes the time of year of the observation; $\sigma_{rv,i}^2$ denotes 5-minute based realized volatility as defined in Chapter 2; $I_i$ is a $(T \times 3)$ matrix.

4.3.3.3 Model Estimation and Results

By construction, the elements of $B_i$ have different effects on $H_i$. To see this, the second term on the right hand side of equation (4.4) is,

$$H_i = A' A + B_{i-1}' B_{i-1} + C' H_{i-1} C + D' \varepsilon_{i-1}' \varepsilon_{i-1} D$$

(4.4)
It is obvious that $\beta_1$ affects the cash variance and the covariance between cash and futures returns. Similarly, $\beta_3$ affects the futures variance and the covariance between cash and futures. In contrast, $\beta_2$ affects both components of $H_t$. Thus, different specifications of $B_t$ result in different $H_t$ behavior. For example, if $I_2$ is the only matrix used to add the second moment covariates, $\beta_2$ will become the only element in $B_t$ and appear in each element of $H_t$. How to choose the elements of $B_t$ is an empirical issue. Therefore different specifications of $B_t$ are tested. Given equation (4.7), it is appropriate to assume that seasonality factors and realized volatility in the futures market affect both cash and futures variances. Thus, $\beta_2$ is always included in $B_t$. As such, four possible specifications are left. For the sake of simplicity, four augmented bivariate GARCH models are named as GARCH-I$_2$, GARCH-I$_1$I$_2$, GARCH-I$_2$I$_3$, and GARCH-I$_1$I$_2$I$_3$ according to the corresponding $I_i$ matrix or matrices included.

Table 4.3 presents the estimation results of the basic BEKK-diagonal GARCH model and four different augmented GARCH models. Comparing the basic GARCH model with the augmented GARCH models is useful to decide whether there is any gain in using a more complex modeling technique. If there is no statistical difference detected between these models, the rewards to hedging using an augmented GARCH model with intraday information are likely to be low.

Likelihood ratio (LR) tests are first applied to compare the models. As reported in Table 4.3, the basic GARCH model is overwhelmingly rejected in favor of the
augmented GARCH models. Results indicate that the augmented GARCH models provide more explanatory power. The LR test statistic is significant at any reasonable level of probability. For instance, in comparing the basic GARCH model with the GARCH-I_2 model, the $\chi^2(3)$ test statistics is 832.82 and the critical value for the test at the 0.005 significance level is 12.84.

Results show that the coefficients, $c_1, c_2, c_3, d_1, d_2,$ and $d_3$ are all highly significant across the four augmented GARCH models. For the GARCH-I_2 model, the coefficients associated with seasonality and realized volatility, $b_{21}, b_{22},$ and $b_{23},$ are statistically significant at 95% confidence level, indicating the importance of adding second moment covariates. This observation also reinforces the previous assertion that $\beta_2$ plays an important role in influencing all the variance covariance elements. Whereas the GARCH-I_1I_2, GARCH-I_2I_3, and GARCH-I_1I_2I_3 models are superior to the GARCH-I_2 model in terms of the LR test, the adding of parameters leads to more insignificant estimates. For example, $a_1, a_2, b_{11}, b_{12}, b_{13}, b_{22},$ and $b_{31}$ are not significant in the GARCH-I_1I_2I_3 model. This implies that the adding of the second moment covariates results in collinearity. For this reason, the GARCH-I_1I_2I_3 is excluded. Given the fact that only futures intraday data are available, it is natural to assume that including $\beta_3$ instead of $\beta_1$ is a better specification for matrix $B_r$. The log-likelihood values for the GARCH-I_1I_2 and GARCH-I_2I_3 models also suggest this assumption. Therefore the GARCH-I_2I_3 model is used for estimation of optimal hedge ratios. Furthermore, as it provides significant estimates of parameters and thus a good parsimonious representation of the augmented GARCH models, the GARCH-I_2 model is also included for further analysis. Finally, the basic
GARCH model will be employed as a benchmark to measure the effectiveness of the two augmented models in the sense that it is identical to the augmented GARCH when the second moment covariates are not allowed. Overall, the statistical analysis in this section suggests that it may be advantageous for the hedger to utilize the intraday information. Therefore, the augmented GARCH models are potentially useful tools for obtaining optimal hedging strategies.

4.4 The Optimal Hedge

4.4.1 Derivation of the Optimal Hedge Ratio

The derivation of the optimal hedge ratio here largely depends on Baillie and Myers (1991). Suppose a hedger takes a long position in a cash commodity market and a short position in a futures market. He has to choose a hedge ratio in hedging the given cash position. If the agent hedges a fraction of the cash position in the futures market, the expected return on the agent’s portfolio is represented by:

$$E(r_i | \Omega_{t-1}) = E(s_i | \Omega_{t-1}) - \gamma_{t-1} E(f_i | \Omega_{t-1})$$  \hspace{1cm} (4.9)

where $r_i$ is the return of a hedged portfolio between $t$ and $t-1$; $s_i$ is the return in the spot market between $t$ and $t-1$; $f_i$ is the return in the futures market between $t$ and $t-1$; $\gamma_{t-1}$ is the hedge ratio.

The optimal hedge ratio is the value of $\gamma_{t-1}$ that maximizes the hedger’s mean-variance utility function. BEZ (1984) prove that the utility-maximizing hedging ratio can be represented by a simple minimum variance hedging rule if the futures market is unbiased. Specifically, if the expected return to take out the futures position is zero, the
hedge ratio will only affect the variance of the hedged portfolio but not the expected return of the portfolio. Thus, the minimum variance hedge ratio is equivalent to the utility-maximizing hedge ratio.

In section 4.3.1, it is shown that the sample mean of the futures return is close to zero and the futures return series follows a martingale. Therefore the soybean futures market can be treated as unbiased. Given $E(f_t) = 0$, the optimal hedge ratio is simply the value of $\gamma_{t-1}$ than minimizes the variance of the return to the hedged portfolio. The conditional variance is denoted by,

$$Var(r_t | \Omega_{t-1}) = Var(s_t | \Omega_{t-1}) + \gamma_{t-1}^2 Var(f_t | \Omega_{t-1}) - 2\gamma_{t-1} Cov(s_t, f_t | \Omega_{t-1}) \quad (4.10)$$

Because of the existence of basis risk, no hedge ratio can completely eliminate risk. An agent chooses $\gamma_{t-1}$ to minimize risk (the conditional variance of the hedged portfolio returns) by setting the derivative of (4.10) with respective to $\gamma_{t-1}$ equal to zero, which is given by,

$$\frac{\partial Var(r_t | \Omega_{t-1})}{\partial \gamma_{t-1}} = 2\gamma_{t-1} Var(f_t | \Omega_{t-1}) - 2 Cov(s_t, f_t | \Omega_{t-1}) = 0 \quad (4.11)$$

An optimal hedge ratio (OHR) is thus defined as the conditional covariance between cash and futures prices divided by the conditional variance of futures price, which is denoted by,

$$\gamma_{t-1} = \frac{Cov(s_t, f_t | \Omega_{t-1})}{Var(f_t | \Omega_{t-1})} \quad (4.12)$$

This ratio is widely used in the literature, for example, Ederington (1979), Carter and Loyns (1985), Brown (1985), Myers and Thompson (1989), Cecchetti, Cumby and

The conventional approach to estimate the hedge ratio is based on OLS technique. Specifically, the regression equation is,

\[ s_t = \alpha_0 + \alpha_1 f_t + e_t \]  

(4.13)

where \( s_t \) and \( f_t \) are defined as before and \( e_t \) is the error term. The static or constant hedge ratio is estimated as the slope coefficient \( \alpha_1 \).

Given the bivariate GARCH models presented in Section 4.3.3.1 and 4.3.3.2, the time-varying hedge ratio can be expressed as,

\[ \gamma_{t-1} = \frac{h_{12,t}}{h_{22,t}} \]  

(4.14)

4.4.2 Optimal Hedge Results

4.4.2.1 In-Sample Results

The descriptive statistics of hedge ratios generated from the four econometric specifications, the basic GARCH, GARCH-I\(_2\), GARCH-I\(_2\)I\(_3\), and OLS, are reported in the second column of Table 4.4. The average optimal hedge ratios for the basic GARCH, GARCH-I\(_2\), and GARCH-I\(_2\)I\(_3\) are 0.9526, 0.9621 and 0.9548, respectively. For the OLS model, the constant hedge ratio is 0.9639. There are modest differences in the mean hedge ratios between the bivariate GARCH models and the OLS model. The average hedge ratios for each of the econometric models imply that the hedger takes the short position (positive sign) in the futures market and about 96% of the exposed cash position
should be hedged to reduce the soybean price uncertainty. This result addresses the importance of the CBOT soybean contracts.

The standard deviation of time-varying hedge ratios estimated from the three GARCH models are 0.1440, 0.0880, and 0.0965, respectively. In terms of variability, the hedge ratios based on the GARCH-I$_2$ model and the GARCH-I$_2$I$_3$ is less variable than the basic GARCH model. Myers (1991) argues that traders have to adjust their positions by greater amounts when they are using less stable hedging rules. Since the basic GARCH hedge ratios are less stable, experiencing more pronounced fluctuations, the hedger has to adjust his futures position more often if he uses the basic GARCH hedge ratios.

In-sample results also reveal that for basic GARCH, GARCH-I$_2$ and GARCH-I$_2$I$_3$ portfolios, the hedger will at most hedge 151%, 141%, and 120% of the cash position, respectively. Noticeably, the minimum hedge ratios are negative. They reflect the fact that cash and futures prices may move in opposite directions in the short run, requiring the hedger to go long in the futures market to hedge the long cash position. Tong (1996) argues that the temporary spread between cash and futures markets causes a negative conditional covariance and thus a negative hedge ratio. However, the opposite co-movement is only temporary and the long-run relationship between cash and futures is still positive as implied by the positive average hedge ratios.

Although the average hedge ratios generated from the three bivariate GARCH models are close to the constant hedge ratio, the variance and covariance matrices of the commodity prices highlight the time-varying nature of the optimal hedge. Figure 4.1 and Figure 4.2 provide the graphical support for this point. Two sub-samples are analyzed, with 1996 representing a bull market and 1999 representing a bear market. Optimal hedge
ratio paths based on the three GARCH models are shown in Figure 4.1 (Year 1996) and Figure 4.2 (Year 1999), respectively. A significant difference between Figure 4.1 and Figure 4.2 is that the trader employs different hedge ratios in these two years. Specifically, Figure 4.1 suggests much more hedging on average than Figure 4.2. The large dip of the first quarter in Figure 4.2 emphasizes this judgment.

In each figure, the upper graph illustrates the hedge ratio path estimated with the basic GARCH model. The middle graph illustrates the hedge ratio path estimated with the GARCH-I\_2 model. The bottom graph illustrates the hedge ratio path estimated with the GARCH-I\_2I\_3 model. In each graph, the constant hedge ratio based on the OLS model is also shown for comparison. Both figures demonstrate that time-varying ratios are clustered around the constant hedge ratio. However, it is clear from the two figures that the bivariate GARCH models suggest that optimal hedge ratios vary considerably over time. In many periods there is an obvious difference between hedge ratios estimated from bivariate GARCH models and constant hedge ratios.

In each figure, a close inspection of the three time-varying hedge ratio paths indicates that they are similar but not identical. For 1996 (1999), the hedge ratio based on the basic GARCH fluctuates between 0.6133 (0.3018) and 1.4156 (1.2605). The hedge ratio based on the GARCH-I\_2 model has smaller fluctuations, ranging from 0.7037 (0.3422) to 1.3733 (1.1139) in 1996 (1999). The hedge ratio based on the GARCH-I\_2I\_3 model has smallest fluctuations, ranging from 0.6573 (0.3669) to 1.1800 (1.0541) in 1996 (1999).

Overall, the in-sample results reveal the time-varying nature of optimal hedge ratios. Yet, the long-term average hedge ratios are all close to one regardless of the models used.
In terms of variability, the augmented GARCH models are preferred to the basic GARCH model.

### 4.4.2.2 Out-Of-Sample Results

Brooks and Chong (2001) argue that complicated models tend to provide better in-sample fit to the data and thus lead to better hedging rules. They suggest that the out-of-sample or ex-ante evaluation is in a more realistic setting because traders are more concerned with future performance. Therefore, 249 observations from August 1, 2000 to July 31, 2001 are used to make an out-of-sample comparison. Parameters are re-estimated as each new observation becomes available. Then the new hedge ratio is derived from the forecast of the variance covariance matrix. The constant hedge ratio based on the OLS model also displays variability in the out-of-sample period since new observations are added in each day.

The out-of-sample average hedge ratios, along with standard deviations and minimum and maximum values, for each of the econometric specifications are reported in the third column of Table 4.4. Again, results show that the average hedge ratios generated from the four estimation procedures have modest differences. The recommended average hedge ratios are 0.9409, 0.9639, 0.9635 and 0.9642 for the basic GARCH, GARCH-I₂, GARCH-I₂I₃ and OLS models, respectively. The average hedge ratios generated by the augmented GARCH models and the OLS model are slightly more than those generated by the basic GARCH method. This implies a hedger would on average purchase fewer soybean contracts under the basic GARCH methodology. Overall, while some differenced emerged in the estimated models, the average hedge ratios for the four
models are relatively similar. Results reveal that hedge ratios are close to one for the soybean data. Haigh and Holt (2002) argue that a hedge ratio close to one indicates the corresponding basis is generally more predictable.

As with the in-sample analysis, the standard deviation of hedge ratios generated by the augmented GARCH models are less than that generated by the basic GARCH. In terms of variability, the complex time-varying hedge ratios are more stable than the basic GARCH hedge ratios. For instance, the volatility of the GARCH-I\textsubscript{2}I\textsubscript{3} hedge ratios is 0.0808 whereas the volatility of the basic GARCH hedge ratios is 0.1195. A less obvious finding is that the out-of-sample hedge ratios have smaller variance compared with their in-sample counterparts. To illustrate, the standard deviation drops from 0.1440 to 0.1195 for the basic GARCH hedge ratios.

Results also reveal that for the basic GARCH portfolio, the proportions hedged will at most be 120.38%. The corresponding numbers for the GARCH-I\textsubscript{2} and GARCH-I\textsubscript{2}I\textsubscript{3} portfolios will be 110.35% and 108.46% respectively.

Figure 4.3 illustrates the out-of-sample hedge ratios using OLS and bivariate GARCH estimation methods. Although the OLS model experiences some variability, with standard deviation of 0.0011, the out-of-sample OLS hedge ratios are nearly constant. In contrast, the hedge ratio paths based on the bivariate GARCH models show considerable variation. Again, three different GARCH hedge ratios display similar pattern.

Basically, the in-sample results are consistent with out-of-sample results. Over both the in-sample and out-of-sample periods, hedge ratios based on the augmented GARCH models exhibit less volatility.
4.4.2.3 Hedging Effectiveness: In-Sample and Out-Of-Sample

Whereas the above analysis suggests that from a statistical standpoint it is helpful to include realized volatility (or intraday information) in daily GARCH models, a hedger is more likely to be concerned with economic differences between the competing models. For this reason, the analysis of the in-sample and out-of-sample hedging effectiveness is imperative.

For purposes of comparison, four hedging strategies are formed: no hedge, the naïve (or full) hedge, the constant hedge and the time-varying hedge. The naïve hedge uses a hedge ratio of one. The constant hedge is constructed from the OLS model. The time-varying hedge ratios are constructed from the three different GARCH specifications. As indicated by Baillie and Myers (1991), comparison between the effectiveness of different hedging rules is made by comparing the variance of the hedged portfolios. If the variances derived from the augmented GARCH models do not differ much from their counterparts, the intraday information is assumed to have a limited effect. In this study, the portfolios are constructed as equation (4.9) with the estimated optimal hedge ratios.

In order to compare the effectiveness of different hedging strategies, a correct probability model for soybean cash and futures prices is needed. Baillie and Myers (1991), Myers (1991) and Haigh and Holt (2002) use the basic bivariate GARCH specification as the correct model to measure different hedging strategies. Their choice is based on the evaluation of alternative models in-sample fit. According to the analysis in 4.3.3.2 and 4.3.3.3, the augmented GARCH specification is treated as the correct probability model for soybean cash and futures prices in this application. Therefore, the
hedging effectiveness evaluation proceeds with two augmented GARCH specifications, GARCH-I_2 and GARCH-I_2I_3.

The variances derived from each model are calculated and then averaged over the relevant period to provide a summary measure of hedging effectiveness. The performance of the different hedge methods is then evaluated by computing the percentage reduction in the variance of the portfolio return compared with the unhedged case. The variance of the hedged portfolio is calculated by equation (4.10) with the estimated hedge ratios. The smaller the variance of the hedged portfolio is, the more effective the hedging strategy.

Table 4.5 shows the variance and the percentage change in variance of the hedged portfolio using different types of optimal hedge ratios. Results are provided for the in-sample and out-of-sample periods separately. It is not surprising that the no-hedge strategy has the highest variance for both periods. Hedging activities, in contrast, reduce portfolio variance. To illustrate, given that GARCH-I_2 is the correct probability model, the constant hedge reduces risk by roughly 86.47% and 89.31% for in-sample and out-of-sample periods respectively.

During the in-sample period, modest differences occur for the alternative hedging strategies. In the upper panel, the unitary hedge and the constant hedge reduce risk by 86.30% and 86.47% respectively. The basic GARCH hedge reduces the variance of returns by 86.54%, an improvement of only 0.07% from the OLS hedge ratio. The largest reduction in variance is associated with the augmented GARCH models which incorporate intraday information. The use of the GARCH-I_2 and GARCH-I_2I_3 models reduce the variance of returns by about 87.72% and 87.63% respectively. Although the time-varying strategy based on the augmented GARCH specification outperforms other
strategies, the improvement is only marginally better in terms of variance reduction. The results based on the GARCH-I_{2I_3} model are quite similar to those based on the GARCH-I_2. A striking feature of the in-sample findings is that the full hedge strategy is nearly as effective as other hedging strategies in reducing risk.

The out-of-sample results from both upper and lower panels reinforce the main findings of the in-sample analysis. Hedging effectiveness is almost identical for different strategies. The use of the augmented GARCH models results in the largest reduction in variance of returns only with negligible differences. For example, in the upper panel, the average variance with the GARCH-I_2 is 0.1469, whereas the corresponding values for the naïve hedge and the constant hedge are 0.1589 and 0.1569, representing 0.82% and 0.68% gain, respectively. In the lower panel, the average variance with the GARCH-I_2 is 0.1615 whereas the corresponding values for the naïve hedge and the constant hedge are 0.1717 and 0.1700, implying 0.69% and 0.58% gain, respectively. Again, the full hedge does a good job in reducing risk.

Taking the in-sample and out-of-sample hedging effectiveness together, Table 4.5 reveals that the augmented GARCH models provide marginal gains to hedging in terms of reduction in variance over the basic GARCH and OLS approaches. Furthermore, the reduction in variance is almost identical for the time-varying hedge, the full hedge and the constant hedge. The conclusion here is consistent with Myers (1991), Garcia, Roh, and Leuthold (1995), and Tong (1996).

According to Myers (1991), complicated time-varying models (such as the augmented GARCH) require the continual futures adjustments and thus result in extra commission charges. Therefore, a naïve hedge or a constant hedge may be a reasonable
choice in this application. Besides, when one considers that the estimation and forecast of the augmented GARCH can be complicated and costly, a simple technique may be appropriate.

Overall, hedging effectiveness testing suggests that the additional complexity of the GARCH approach is not justified. Referring back to the analysis in 4.4.2.1 and 4.4.2.2, the time-varying hedge ratios based on GARCH models are always clustered around the OLS constant hedge ratio. According to Tong (1996), “…asset fluctuations are important in the short run, but they tend to net out each other in the long run.” This suggests that a hedge ratio ignoring time-varying variance and a hedge ratio incorporating time-varying variance will yield similar results on average if the variance covariance matrix in (4.2) or (4.4) is reasonably stable over sample period. Since the OLS constant hedge ratio is always close to the value of one (the full hedge), this indirectly indicates that time-varying hedge ratios are also close to the full hedge ratio. Haigh and Holt (2002) argue that hedging ratios close to one imply that the market basis is more predictable. In this case, the gains to using more sophisticated models are not obvious compared with standard approaches. Based on their assertion, it may be more advantageous to use the augmented GARCH if the hedge ratios are smaller or the markets are more unpredictable. Yet, the current data imply a relatively stable basis. Therefore, it is reasonable to get similar hedging results from OLS, basic GARCH, and augmented GARCH models.

Myers (1991) suggests that over more extended periods of time, a constant hedge may provide an attractive alternative to the more general time-varying framework. The sample period used for estimation and forecast is approximately 11 years with 2659 daily observations, which is relatively a long horizon. This may partly explain the results.
4.5 Conclusion

This paper introduces a method to incorporate the intraday information into the extant GARCH specification and calculates the optimal hedge ratios based on this new approach. The purpose is to examine whether the additional effort required to estimate the augmented GARCH model provides significantly improved hedging performance, compared to the OLS or basic GARCH approach.

The paper further compares the hedging effectiveness of different methods for in-sample and out-of-sample periods. The markets under study are soybean cash and futures markets during the period January 1990-August 2001.

Results show that the augmented GARCH specification has obvious advantages in characterizing essential features of the data in contrast with the standard GARCH specification. The log-likelihood ratio tests and the t-statistics both suggest that the augmented GARCH model is potentially useful for generating a time-varying optimal hedge ratio.

The in-sample and out-of-sample estimation of optimal hedge ratios indicates long-term average hedge ratios are all close to one regardless of the models used. In terms of variability, the augmented GARCH models are preferred to the basic GARCH model.

The hedging effectiveness is investigated by comparing four strategies: no hedge, the naïve hedge, the constant hedge and the time-varying hedge. A comparison between the augmented GARCH, the basic GARCH, the full hedge and the constant hedge ratio portfolios indicates that they have similar variances. Results reveal that the use of the augmented GARCH models, while statistically appropriate, provides only marginal gains to the hedger.
Since the effectiveness of optimal hedge ratios will differ depending on the forecast horizon and data frequency, whether similar findings arise from other applications is a question for further research. The results listed above apply only to the soybean data studied here and different results may be found in other applications.
<table>
<thead>
<tr>
<th></th>
<th>ADF t-statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash Price*</td>
<td>-1.8898</td>
<td></td>
<td></td>
<td>-3.458</td>
</tr>
<tr>
<td>Futures Price</td>
<td>-1.9775</td>
<td></td>
<td></td>
<td>-2.871</td>
</tr>
<tr>
<td>Cash Return**</td>
<td>-16.2149</td>
<td>10% Critical Value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Futures Return</td>
<td>-16.7683</td>
<td></td>
<td></td>
<td>-2.594</td>
</tr>
</tbody>
</table>

* Price refers to the log price.
** Return refers to the log return.

Table 4.1: Augmented Dickey-Fuller Test for Cash and Futures Prices & Cash and Futures Returns

<table>
<thead>
<tr>
<th>Year</th>
<th>Futures Returns*</th>
<th>Cash Returns</th>
<th>Correlation of Cash and Futures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev</td>
<td>Mean</td>
</tr>
<tr>
<td>1990</td>
<td>-0.00062</td>
<td>0.01081</td>
<td>-0.0003</td>
</tr>
<tr>
<td>1991</td>
<td>-0.00032</td>
<td>0.01355</td>
<td>0.00031</td>
</tr>
<tr>
<td>1992</td>
<td>-0.00006</td>
<td>0.00908</td>
<td>-0.00007</td>
</tr>
<tr>
<td>1993</td>
<td>0.00030</td>
<td>0.00991</td>
<td>0.00054</td>
</tr>
<tr>
<td>1994</td>
<td>-0.00094</td>
<td>0.01173</td>
<td>-0.00094</td>
</tr>
<tr>
<td>1995</td>
<td>0.00073</td>
<td>0.01030</td>
<td>0.00113</td>
</tr>
<tr>
<td>1996</td>
<td>-0.00018</td>
<td>0.01227</td>
<td>-0.00021</td>
</tr>
<tr>
<td>1997</td>
<td>0.00082</td>
<td>0.01476</td>
<td>0.00011</td>
</tr>
<tr>
<td>1998</td>
<td>-0.00079</td>
<td>0.01067</td>
<td>-0.00094</td>
</tr>
<tr>
<td>1999</td>
<td>-0.00066</td>
<td>0.01433</td>
<td>-0.00030</td>
</tr>
<tr>
<td>2000</td>
<td>-0.00007</td>
<td>0.01193</td>
<td>0.00042</td>
</tr>
<tr>
<td>2001</td>
<td>0.00005</td>
<td>0.01277</td>
<td>0.00022</td>
</tr>
<tr>
<td>Full Sample</td>
<td>-0.00015</td>
<td>0.01192</td>
<td>0.00001</td>
</tr>
</tbody>
</table>

* Returns series refers to the log return.

Table 4.2: Summary Statistics of Cash and Futures Returns: Sub-samples and Full Sample
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Basic GARCH</th>
<th>Augmented GARCH-I</th>
<th>Augmented GARCH-I$_1$</th>
<th>Augmented GARCH-I$_2$</th>
<th>Augmented GARCH-I$_1$I$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. errors</td>
<td>Estimate</td>
<td>Std. errors</td>
<td>Estimate</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.1921</td>
<td>0.3388</td>
<td>0.0351</td>
<td>0.0520</td>
<td>-0.0829</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.2755</td>
<td>0.3304</td>
<td>0.2940</td>
<td>0.0474</td>
<td>0.2270</td>
</tr>
<tr>
<td>$a_3$</td>
<td>0.3955</td>
<td>0.2534</td>
<td>0.1807</td>
<td>0.0803</td>
<td>0.1651</td>
</tr>
<tr>
<td>$b_{11}$</td>
<td></td>
<td>0.1062</td>
<td>0.0489</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_{12}$</td>
<td></td>
<td></td>
<td>-0.0107</td>
<td>0.0294</td>
<td>-0.0913</td>
</tr>
<tr>
<td>$b_{13}$</td>
<td></td>
<td></td>
<td>0.1332</td>
<td>0.0646</td>
<td></td>
</tr>
<tr>
<td>$b_{21}$</td>
<td>-0.0974</td>
<td>0.0275</td>
<td>-0.0375</td>
<td>0.0205</td>
<td>-0.0913</td>
</tr>
<tr>
<td>$b_{22}$</td>
<td>-0.0511</td>
<td>0.0233</td>
<td>-0.0555</td>
<td>0.0154</td>
<td>-0.0464</td>
</tr>
<tr>
<td>$b_{23}$</td>
<td>0.0517</td>
<td>0.0296</td>
<td>0.1188</td>
<td>0.0468</td>
<td>0.0824</td>
</tr>
<tr>
<td>$b_{31}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.1532</td>
</tr>
<tr>
<td>$b_{32}$</td>
<td></td>
<td></td>
<td>-0.0539</td>
<td>0.0456</td>
<td>-0.0599</td>
</tr>
<tr>
<td>$b_{33}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.0774</td>
</tr>
<tr>
<td>$c_1$</td>
<td>0.8568</td>
<td>0.0557</td>
<td>0.5107</td>
<td>0.0216</td>
<td>0.521</td>
</tr>
<tr>
<td>$c_2$</td>
<td>0.0495</td>
<td>0.1451</td>
<td>0.4312</td>
<td>0.0202</td>
<td>0.4397</td>
</tr>
<tr>
<td>$c_3$</td>
<td>0.7916</td>
<td>0.1057</td>
<td>0.4717</td>
<td>0.0217</td>
<td>0.4883</td>
</tr>
<tr>
<td>$d_1$</td>
<td>0.4626</td>
<td>0.0868</td>
<td>0.4299</td>
<td>0.0275</td>
<td>0.3734</td>
</tr>
<tr>
<td>$d_2$</td>
<td>-0.1266</td>
<td>0.1392</td>
<td>-0.1563</td>
<td>0.0377</td>
<td>-0.2019</td>
</tr>
<tr>
<td>$d_3$</td>
<td>0.4960</td>
<td>0.0878</td>
<td>0.4455</td>
<td>0.0345</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Log-likelihood: -889.6345, -473.2249, -391.0606, -369.6593, -339.0155

Table 4.3: Estimation Results of the Bivariate GARCH Models
<table>
<thead>
<tr>
<th></th>
<th>In-sample</th>
<th>Out-of-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic GARCH</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.9526</td>
<td>0.9409</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.1440</td>
<td>0.1195</td>
</tr>
<tr>
<td>Min</td>
<td>-0.7089</td>
<td>0.2920</td>
</tr>
<tr>
<td>Max</td>
<td>1.5135</td>
<td>1.2038</td>
</tr>
<tr>
<td><strong>GARCH-I_{I_{2}}</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.9621</td>
<td>0.9639</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0880</td>
<td>0.0807</td>
</tr>
<tr>
<td>Min</td>
<td>-0.4852</td>
<td>0.4472</td>
</tr>
<tr>
<td>Max</td>
<td>1.4064</td>
<td>1.1035</td>
</tr>
<tr>
<td><strong>GARCH-I_{I_{2}I_{3}}</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.9548</td>
<td>0.9635</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0965</td>
<td>0.0808</td>
</tr>
<tr>
<td>Min</td>
<td>-0.6530</td>
<td>0.3576</td>
</tr>
<tr>
<td>Max</td>
<td>1.2019</td>
<td>1.0846</td>
</tr>
<tr>
<td><strong>OLS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.9639</td>
<td>0.9642</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.0011</td>
<td></td>
</tr>
<tr>
<td>Min</td>
<td>0.9613</td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>0.9662</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: Descriptive Statistics for Hedge Ratios
### Table 4.5: Hedging Effectiveness Comparison: In-Sample and Out-of-Sample Results

<table>
<thead>
<tr>
<th></th>
<th>In-Sample</th>
<th></th>
<th>Out-of-Sample</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>If GARCH-I\textsubscript{2} is the correct probability model:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Hedge*</td>
<td>1.5646</td>
<td></td>
<td>1.4680</td>
<td></td>
</tr>
<tr>
<td>Full Hedge**</td>
<td>0.2144</td>
<td>86.30%</td>
<td>0.1589</td>
<td>89.17%</td>
</tr>
<tr>
<td>OLS</td>
<td>0.2117</td>
<td>86.47%</td>
<td>0.1569</td>
<td>89.31%</td>
</tr>
<tr>
<td>Basic GARCH</td>
<td>0.2106</td>
<td>86.54%</td>
<td>0.1560</td>
<td>89.37%</td>
</tr>
<tr>
<td>GARCH-I\textsubscript{2}</td>
<td>0.1922</td>
<td>87.72%</td>
<td>0.1469</td>
<td>89.99%</td>
</tr>
<tr>
<td>GARCH-I\textsubscript{2}I\textsubscript{3}</td>
<td>0.1936</td>
<td>87.63%</td>
<td>0.1474</td>
<td>89.96%</td>
</tr>
<tr>
<td><strong>If GARCH-I\textsubscript{2}I\textsubscript{3} is the correct probability model:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Hedge</td>
<td>1.5105</td>
<td></td>
<td>1.4580</td>
<td></td>
</tr>
<tr>
<td>Full Hedge</td>
<td>0.2193</td>
<td>85.48%</td>
<td>0.1717</td>
<td>88.23%</td>
</tr>
<tr>
<td>OLS</td>
<td>0.2162</td>
<td>85.69%</td>
<td>0.1700</td>
<td>88.34%</td>
</tr>
<tr>
<td>Basic GARCH</td>
<td>0.2121</td>
<td>85.96%</td>
<td>0.1701</td>
<td>88.33%</td>
</tr>
<tr>
<td>GARCH-I\textsubscript{2}</td>
<td>0.1955</td>
<td>87.06%</td>
<td>0.1615</td>
<td>88.92%</td>
</tr>
<tr>
<td>GARCH-I\textsubscript{2}I\textsubscript{3}</td>
<td>0.1942</td>
<td>87.14%</td>
<td>0.1606</td>
<td>88.98%</td>
</tr>
</tbody>
</table>

* No hedge indicates that the size of futures position is zero.

** Full hedge indicates that the hedge ratio is one.
Figure 4.1: Time-Varying Hedge Ratios vs. Constant Hedge Ratio for Year 1996
Upper Graph: Basic GARCH vs. OLS
Middle Graph: GARCH-I\textsuperscript{2} vs. OLS
Bottom Graph: GARCH-I\textsuperscript{2}I\textsuperscript{3} vs. OLS
Figure 4.2: Time-Varying Hedge Ratios vs. Constant Hedge Ratio for Year 1999

Upper Graph: Basic GARCH vs. OLS
Middle Graph: GARCH-I\textsubscript{2} vs. OLS
Bottom Graph: GARCH-I\textsubscript{3}I\textsubscript{3} vs. OLS
Figure 4.3: Out-of-Sample Forecasts of Optimal Hedge Ratios
Upper Graph: Basic GARCH vs. OLS
Middle Graph: GARCH-I\(^2\) vs. OLS
Bottom Graph: GARCH-I\(^2\)I\(^3\) vs. OLS
CHAPTER 5

SUMMARY AND CONCLUSION

In this dissertation, two volatility measures are examined in terms of their
distributional properties and their effectiveness in volatility forecasting and optimal
hedging, respectively. Special attention is given to analyzing the performance of existing
time series models with the addition of these two proxies.

For distributional properties, the results indicate that the characteristics of realized
volatility in the soybean futures market largely correspond with existing literature.
Specifically, the properties of the realized variance, the standard deviation and the log
standard deviation have quite similar patterns as those observed in stock and foreign
exchange markets, although there are some discrepancies in size. In contrast, the
standardized daily returns display some different properties compared with stock and
exchange rate data. The long memory characteristics of realized volatilities, especially for
the logarithmic standard deviations, are obvious. As for the log range, the skewness and
kurtosis are closer to the corresponding values of a normal random variable compared
with those for log absolute returns.

For out-of-sample forecast estimates, the performance of GARCH models, range-
based GARCH models, and log-range based ARMA models are compared. The results
reveal that the log-range based ARMA models may have disadvantages in comparison to
GARCH models extended with daily ranges in terms of forecasting abilities. The
empirical analysis suggests that it is important to include the daily range in the GARCH
models for forecasting soybean futures market volatility. However, for the extended
GARCH models, adding seasonality becomes less important.

The economic value of realized volatility is examined in the context of hedging.
Results show that the augmented GARCH specification has advantages in characterizing
essential features of the data in contrast with the standard GARCH specification. The in-
sample and out-of-sample estimation of optimal hedge ratios indicates long-term average
hedge ratios are all close to one regardless of the models used. In terms of variability, the
augmented GARCH models are preferred to the basic GARCH model. The hedging
effectiveness is investigated by comparing four strategies: no hedge, the naïve hedge, the
constant hedge and the time-varying hedge. A comparison between the augmented
GARCH, the basic GARCH, the full hedge and the constant hedge ratio portfolios
indicates that they have similar variances. Results reveal that the use of the augmented
GARCH models, while statistically appropriate, provides only marginal gains to the
hedger.

The results listed above apply only to the soybean data studied here and different
results may be found in other applications. The relatively narrow applications contained
in this thesis preclude making any general conclusions. Each of the foregoing chapters
leaves questions to be answered. The results of chapters two, three and four may have
broad implications for future research in commodity price analysis, volatility forecasting
and empirical finance, respectively.


