THESIS

2
2004
63050002

This is to certify that the
dissertation entitled

FINANCIAL ECONOMETRIC MODELING OF RISK IN
COMMODITY MARKETS

presented by

Jeongseok Song

has been accepted towards fulfillment
of the requirements for the

Ph.D degree in Economics

Major Professor's Signature

05.20.2004

Date

MSU is an Affirmative Action/Equal Opportunity Institution

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
SEP 0 3 2006		
0 5 1 4 0 6		

6/01 c:/CIRC/DateDue.p65-p.15

FINANCIAL ECONOMETRIC MODELING OF RISK IN COMMODITY MARKETS

By

Jeongseok Song

A DISSERTATION

Submitted to
Michigan State University
In partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

2004

ABSTRACT

FINANCIAL ECONOMETRIC MODELING OF RISK IN COMMODITY MARKETS

By

Jeongseok Song

This dissertation is composed of three interrelated body chapters. Its goal is to identify underlying sources of return volatility movement and analyze important problems in the economics of commodity markets by applying various time series econometric models to commodity market price data.

Chapter 2 investigates stochastic properties of daily cash price changes for six commodities: corn, soybeans, live cattle, live hogs, unleaded gasoline, and gold. We use the FIGARCH conditional variance model and the semi-parametric local Whittle estimation method to explore the daily cash return volatility behavior. We apply the long memory models to the temporally aggregated daily cash returns and compare the volatility dynamics at various sample frequencies.

Chapter 3 is concerned with commodity futures return volatility at daily and higher sample frequencies. In particular, strong intra-day periodicity in the high frequency return volatility is observed. We examine the high frequency futures return volatility pattern after removing the intra-day seasonality using the Flexible Fourier Form (FFF) filter and compare the volatility movement with the daily futures return volatility process.

Chapter 4 introduces a newly suggested volatility measure, the realized volatility, and applies the volatility measure to commodity futures market price data. The realized volatility is calculated as the sum of high frequency squared returns and exhibits some ideal statistical properties. Taking advantage of the stochastic properties of the realized volatility measure allows us to study important economic determinants for commodity futures market risk behavior.

Dedicated to My Parents

Acknowledgements

This dissertation would not have been completed without support of many people. However, my first and greatest thanks must go to my advisor, Professor Richard T. Baillie for his superb guidance and encouragement. He provided insightful and invaluable advices as well as financial supports during my graduate studies.

Also, I wish to thank Professor Robert Myers in agricultural economics department for extending my knowledge to broader area and suggesting many helpful comments. In addition, I am grateful to Professors Christine Amsler and Anna Maria Herrera in economics department for their patience and generosity with respect to my dissertation drafts.

Due to too many to list, I mention a few of colleagues in alphabetical order of their last names to thank: Wooseok Choi, Young-Wook Han, Seung Won Kim, Jaewha Lee, Han Sang Yi, and Kyeongwon Yoo. They delighted and assisted me for a variety of reasons through a long journey of graduate studies.

My parents have waited for me to complete my graduate studies for a decade. It is impossible for me to compensate their care and support by any means.

My final thanks remain for Minkyung, my beloved daughter for yielding her time to play with me, and Minah Park, my wife, for doing many of my duties so as to let me concentrate on my works.

[TABLE OF CONTENTS]

[LIST OF TABLES]

[LIST OF FIGURES]

[CHAPTER 2]

[CHAPTER 3]

[CHAPTER 4]

CHAPTER 1

INTRODUCTION

This dissertation is concerned with the application of some modern financial econometric techniques to daily and high frequency commodity markets. The econometric methods are applied to the cash and futures markets. These cash and futures markets are an active and important financial institution in the modern economy, and the volatility associated with commodity futures markets is an important factor for study in risk management and in commodity trading. The market in recent years has observed remarkable growth in trading volume, the variety of contracts, and the range of underlying commodities. Market participants are also becoming increasingly sophisticated about recognizing and exercising operational contingencies embedded in delivery contracts. For all of these reasons, there is a widespread interest in models for pricing and hedging commodity-linked contingent claims. Despite these facts, relatively little attention has been paid to commodity markets, in comparison with the enormous recent empirical analyses of the currency and equity markets. While commodity markets are smaller and possibly lack the glamour of currency and equity markets, they are nevertheless important for the agricultural sector of the economy and for maintaining overall supply and demand conditions in the macro economy.

Chapter 2 is concerned with commodity cash market price risks. The cash markets are characterized by the unique physical properties of commodities since cash prices are determined by supply and demand for commodities that are subject to various unique factors such as weather and other environmental determinants. We introduce the long memory property with its characteristic self-similarity and study daily cash return volatility dynamics with reflection on various characteristics of commodities such as annual seasonality for agricultural products and distinguishing trading patterns for livestock. Accounting for those characteristics, our empirical investigation uses the FIGARCH and local Whittle semi-parametric estimation method to reveal that the long memory property is evident in the daily cash return volatility.

Chapter 3 is concerned with investigating the possible existence of the long memory feature in commodity markets. Apart from the study by Cai et al. (2002), this thesis appears to be the first systematic study of the phenomenon and its applicability to and implications for commodity markets. Hence, in chapter 3 we apply the FIGARCH and the local Whittle estimator to commodity futures market price data and also report the estimates of various long memory volatility models. We find overwhelming evidence for the phenomenon, which is consistent with the evidence found in the securities and currency markets. We also investigate and discuss the property of self-similarity in commodity markets, generally finding our empirical results to be consistent with self-similarity. It turns out to be very important to consider issues of time to maturity when modeling volatility in these markets. Further, chapter 3 deals with high frequency commodity futures market data. We first discuss meaningful ways of constructing high frequency returns, and then describe the empirical properties of these. Much emphasis is placed on the particularly unusual intra-day periodicity that occurs in these futures markets and its elimination through the application of Gallant's (1981) FFF filtering method. Our finding is supportive of self-similarity for high frequency commodity futures return volatility.

Chapter 4 is concerned with the relatively new measure of realized volatility, which has recently become a competitor with the dominant GARCH model of Engle (1982). We find some interesting features of very persistent autocorrelation, or long memory, in the realized volatility series. The realized volatility series are also partly determined by USDA announcement effects and the local market conditions of time to maturity. We introduce the new concept of information flow, which is measured using trading intensity built from a high frequency time dimension. We consider information flow and time-to-maturity effects to explain realized volatility. Even allowing for these effects, the long memory effects in the realized volatility series tend to remain. Chapter 4 also investigates the patterns of dependencies between the realized volatility series of several different commodities. We discuss these results in the context of fractional integration and the existence of a common structural long memory trend in the generation of the realized volatility series.

We summarize our studies and conclude this dissertation with possible future research in chapter 5.

CHAPTER 2

MODELING COMMODITY CASH RETURNS

2.1. Introduction

This chapter is concerned with the stochastic properties of daily commodity cash prices for corn, soybeans, live cattle, live hogs, unleaded gasoline, and gold. This type of analysis is an important precursor for many financial market applications, including calculation of optimal hedge ratios, computation of Value at Risk (VaR), etc. While previous studies have investigated the time series properties of commodity cash prices using stable GARCH models, we are unaware of any previous investigation of the long memory properties of daily cash series. For the successful application of financial market analysis and policy, an investigation of the detailed properties of these asset prices seems long overdue.

Commodities are physical products and possibly not involved in trading for possible swift arbitrage. Commodity trading involves some transaction costs attributable to storage and transportation that are not relevant to most financial assets. In particular, commodity cash prices are directly determined by the supply and demand for actual products, while commodity futures contracts are traded in order to reduce uncertainty for the underlying commodities. Therefore, commodity cash markets are quite different from other financial markets. Baillie and Myers (1991), and Cecchetti, Cumby, and

Figlewski (1988) used commodity cash and futures prices for the optimal hedge ratio calculation. Mackey (1989), Yang and Brorsen (1992), and Burton (1993) documented daily commodity cash prices by using nonlinear dynamic models. Yang and Brorsen (1992) and Burton (1993) compared GARCH models with chaos models to explain complicated commodity cash price volatility dynamics. Yang and Brorsen (1992) considered GARCH, mixed diffusion-jump, and deterministic chaos models of cash commodity prices and concluded that the GARCH volatility process provided the best fit. Mackey (1989) suggested a theoretical model to argue that supply and demand for commodities may cause nonlinear price dynamics.

The long memory property is well known to occur in squared returns, absolute returns, and various transformations of volatility such as conditional variances and stochastic volatility models. There are several plausible reasons for the occurrence of long memory in absolute returns, conditional variances and other measures of volatility. First, Granger (1980) showed how the contemporaneous aggregation of independent AR(1) processes can lead to a long memory process as the number of cross section units gets large. This result depends upon the autoregressive parameters having a beta distribution in the interval (0, 1). Usually, the aggregation of N independent AR(1) processes leads to an ARMA(N, N-1) model. However, Granger (1980) shows that this tends to follow fractional white noise as N gets large. Extension of this aggregation argument to volatility models is less than straightforward. Ding and Granger (1996) showed that if each asset's return is a martingale with stable GARCH(1,1) innovations, then the autocorrelations of the squared returns of the contemporaneously aggregated assets will exhibit hyperbolically decaying autocorrelations, and hence the long memory property. Also, Andersen and Bollerslev (1997a) claimed that long memory can result from aggregated heterogenous information components in line with the Mixture-of-Distribution hypothesis noted by Clark (1973) and Tauchen and Pitt (1983). A further suggestion of Parke (2000) is that long memory can arise from the aggregation of shocks, each with a different duration time. Indeed, financial markets are subject to numerous economic factors and considerably responsive to the vast amount of information available in the markets.

This chapter adds to the literature by investigating the long memory for commodity market price risk and examining self-similarity to verify the long-run temporal dependence as an original property of commodity cash price changes. We start with the daily cash price in this chapter and continue with daily and intra-day futures prices in chapter 3.

The remainder of this chapter proceeds as follow. Section 2 provides a brief theoretical background for long memory, self-similarity, and temporal aggregation. In section 3, we document empirical findings for the long memory and self-similarity by using the FIGARCH and the local Whittle semi-parametric model for the temporally aggregated daily cash returns motivated in section 2. Section 4 concludes the chapter.

2.2. Long Memory, Self-Similarity, and Temporal Aggregation

In this section, we discuss definitions of long memory and relate it to the concept of self-similarity. One possible definition of long memory is as follows: if the population autocorrelation of a time series process at lag j, denoted by ρ_j, has the following property,

$$\lim_{n\to\infty} \sum_{j=-n}^{n} \left|\rho_j\right| = \infty, \tag{2.1}$$

the process is said to exhibit long memory. For a sufficiently large number of lags j, a process with autocorrelation function $\rho_j \approx cj^{2d-1}$ and a positive constant c, and where $-0.5 < d < 0.5$, can be formally defined as a stationary long memory process. We call d the long memory parameter. Autocorrelation for such a type of process decays very slowly over long time lags.

Granger and Joyeux (1980), Granger (1980), and Hosking (1981) have developed the Autoregressive Fractionally Integrated Moving Average (henceforth, ARFIMA) model to represent a time series process with the long memory property. Baillie (1996) provides a comprehensive survey of the long memory theories and applications in macroeconomics and finance. As suggested by Granger and Joyeux (1980), Granger (1980), and Hosking (1981), the ARFIMA model takes the following form,

$$\phi(L)(1-L)^d (y_t - \mu) = \theta(L)\varepsilon_t, \tag{2.2}$$

where all the roots of the p'th order polynomial in the lag operator $\phi(L)$ and the q'th order polynomial in the lag operator $\theta(L)$ are assumed to lie outside the unit circle. The process ε_t is white noise. The operator $(1-L)^d$ is the fractional difference operator defined as follows:

$$(1-L)^d \equiv \left\{1-dL+\frac{d(d-1)}{2!}L^2-\frac{d(d-1)(d-2)}{3!}L^3+\ldots\right\}. \qquad \textbf{(2.3)}$$

The ARFIMA process combines the stationary and invertible ARMA model which generates I(0) behavior with the above fractional difference operator, which adds on the long memory behavior for the time series process. For a large lag j, there is hyperbolic decay in the autocorrelations of the ARFIMA process and $\rho_j \approx cj^{2d-1}$ where $c>0$. To describe another important property of the ARFIMA process, we consider the impulse response weights, following Campbell and Mankiw (1987). The impulse response weights are defined by first differencing the ARFIMA process, y_t, to obtain

$$(1-L)y_t = A(L)\varepsilon_t \qquad \textbf{(2.4)}$$

where $A(L)=(1-L)^{1-d}\phi(L)^{-1}\theta(L)$. We can express the lag polynomial $A(L)$ in terms of the hypergeometric functions as

$$A(L)=F(d-1,1,1;L)\phi(L)^{-1}\theta(L) \qquad \textbf{(2.5)}$$

where $F(a,b,c;z)\equiv\{\Gamma(c)/[\Gamma(a)\Gamma(b)]\}\left\{\sum_{i=1}^{\infty}[\Gamma(c+i)/\Gamma(i+1)]\right\}$ and $\Gamma(\bullet)$ is a Gamma function. Since $F(d-1,1,1;L)=0$, as Gradszteyn and Ryzhnik (1980) show, we have $A(1) = F(d-1,1,1;1)\phi(1)^{-1}\theta(1) = 0$ for $d < 1$. The impact of a unit innovation at time t on the process y_{t+k} is then given by

$$1 + \sum_{j=1,k} A_j . \quad \textbf{(2.6)}$$

Therefore, a fractionally integrated process with $d < 1$ is mean reverting. In particular, y_t for $0.5 < d < 1$ is still mean reverting, although the process is not covariance stationary. The long memory feature provides a flexible way of describing complicated volatility temporal patterns, while conventional ARMA class models capture only short-run dynamics in modeling time series and may be too strict in uncovering longer term persistence for the series.

Another important property of the long memory process is self-similarity. The general notion of self-similarity is that some random variables behave identically when they are viewed at different scales on a dimension. The dimension may be space or time, and, particularly, will be time when we analyze time series data. Consider a process y_t following the long memory property with autocorrelation $\rho_j \approx cj^{2d-1}$ for j lags. Given the autocorrelation function, the corresponding spectral density for the associated process can be expressed as follows,

$$f(\lambda) = \frac{\sigma^2}{2\pi}\sum_{-\infty}^{\infty} \rho_j e^{ik\lambda} . \quad \textbf{(2.7)}$$

Then, the spectral density is approximately of the form $\rho_j \approx cj^{2d-1}$ with a constant c as $\lambda \to 0$ where λ represents Fourier frequencies. More formally, y_t is called self-similar with a self-similar parameter *H*, if for any positive stretching factor *c* the rescaled process $c^{-H} y_{ct}$ has the same distribution as the original process y_t.

Following the formal definition and basic concept of self-similarity, we proceed with temporal aggregation. Let $R_t^{(k)} \equiv \sum_{l=0,(k-1)} R_{tk-l}$ denote temporally aggregated returns at a *k*-day sample frequency. For simplicity, assume that $R_t = \sigma_t z_t$ and z_t are independent and identically distributed and σ_t represents a positive and measurable time-varying function. According to many previous empirical findings for the long memory property for squared asset returns, we can assume that $\rho\left([R_t]^2, \left[R_{t-j}\right]^2\right) \approx j^{2d-1}$ for $0 < d < 0.5$ [1]. The temporally aggregated returns, $R_t^{(k)}$, can be expanded as follows:

$$\left[R_t^{(k)}\right]^2 = R_{tk}^2 + R_{tk-1}^2 + R_{tk-2}^2 + \cdots + R_{tk-k+1}^2 + 2\sum_{l \neq m} R_{tk-l} R_{tk-m} \,. \qquad \textbf{(2.8)}$$

Since we assume that z_t are independent and identically distributed, $R_{tk-l}R_{tk-m}$ terms for $l \neq m$ should not matter in considering the autocorrelation below. Then, the *j*-th order autocorrelation $\rho\left(\left[R_t^{(k)}\right]^2, \left[R_{t-j}^{(k)}\right]^2\right)$ is simply the sum of autocorrelations for all the

[1] The long memory process with $0 < d < 0.5$ shows all positive autocorrelations decaying at a hyperbolic rate; see Baillie (1996).

possible pairs of the squared terms underlying $\left[R_t^{(k)}\right]^2$ and $\left[R_{t-j}^{(k)}\right]^2$. In other words, the j-th order autocorrelation $\rho\left(\left[R_t^{(k)}\right]^2,\left[R_{t-j}^{(k)}\right]^2\right)$ can be obtained by summing the autocorrelations between $\left[R_t^{(k)}\right]^2$ and $\left[R_{t-jk-h}^{(k)}\right]^2$ for $h=-k+1,-k+2,\cdots,k-1$. After some straightforward algebra, we have

$$\begin{aligned}\rho\left(\left[R_t^{(k)}\right]^2,\left[R_{t-j}^{(k)}\right]^2\right) &= k^{-2}\sum_{h=-k+1}^{k-1}\left(k-|h|\right)\rho\left(\left[R_t^{(k)}\right]^2,\left[R_{t-j\cdot k-h}^{(k)}\right]^2\right)\\ &= k^{-2}\sum_{h=-k+1}^{k-1}\left(k-|h|\right)\left(jk+h\right)^{2d-1}\end{aligned} \tag{2.9}$$

where $h=-k+1,-k+2,\cdots,k-1$. Note that $k^{-2}\sum_{h=-k+1}^{k-1}\left(k-|h|\right)=1$. Further, if time lag j is sufficiently large, we have

$$k^{-2}\sum_{h=-k+1}^{k-1}\left(k-|h|\right)\left(jk+h\right)^{2d-1}\approx\left(jk\right)^{2d-1}\sim j^{2d-1} \tag{2.10}$$

Consequently, we have

$$\rho\left(\left[R_t^{(k)}\right]^2,\left[R_{t-j}^{(k)}\right]^2\right)\sim j^{2d-1} \tag{2.11}$$

According to this result, we assess that temporally aggregated squared returns theoretically exhibit identical decaying rates for their autocorrelations regardless of the sampling frequencies k. In other words, for sufficient lags j, the autocorrelation between $\left[R_t^{(k)}\right]^2$ and $\left[R_{t-j}^{(k)}\right]^2$ takes an identical form to the autocorrelation between R_t^2 and R_{t-j}^2 for different values of k. Concretely, if R_t^2 exhibits long memory, then $\left[R_t^{(k)}\right]^2$ also shows the same degree of long run temporal dependence. Consistent with the self-similarity notion discussed above, temporally aggregated squared returns show identical long memory behavior if their underlying squared returns follow the long memory process. This result can carry over to a temporally aggregated absolute return case as below.

Especially if we assume further that σ_t from $R_t = \sigma_t z_t$ follows log-normal distribution, then $\left|R_t^{(k)}\right|^{2\theta}$ for all $\theta > 0$ will exhibit identical decaying rates for autocorrelation behavior. This result has been noted by Granger and Newbold (1976) and recently confirmed by Andersen (1994). Since $\left[R_t^{(k)}\right]^2 = \left|R_t^{(k)}\right|^2$, temporally aggregated squared returns should be one particular case of power transformed absolute returns, $\left|R_t^{(k)}\right|^{2\theta}$ for $\theta = 1$. Further, identical decaying rates for autocorrelations of $\left|R_t^{(k)}\right|^{2\theta}$ for all $\theta > 0$ imply that temporally aggregated absolute returns also yield the same autocorrelation behavior as temporally aggregated squared returns, since temporally aggregated absolute return $\left|R_t^{(k)}\right|^{2\theta}$ is just another case for $\theta = 0.5$.

The theoretical self-similarity of temporally aggregated squared returns motivates the application of FIGARCH conditional variance model to temporally aggregated daily cash returns. Following Baillie, Bollerslev, and Mikkelsen (1996), the FIGARCH model is defined as follows:

$$\sigma_t^2 = \omega + \beta(L)\sigma_t^2 + \left[1 - \beta(L) - (1 - \phi(L))(1 - L)^d\right]\varepsilon_t^2 \cdot \tag{2.12}$$

where $\beta(L) = \beta_1 L + \beta_2 L^2 + \cdots + \beta_p L^p$ and $\phi(L) = \phi_1 L + \phi_2 L^2 \cdots + \phi_q L^q$. Baillie, Bollerslev, and Mikkelsen (1996) incorporated slow hyperbolic decay into the conditional variance modeling. The FIGARCH process considers a slowly decaying autocorrelation for lagged squared innovations and allows for persistent impulse response weights without involving the never-dying-out cumulative impulse response weights. The FIGARCH model can describe conditional variance in a more flexible way by allowing for $0 < d < 1$ while the IGARCH model yields unrealistically everlasting volatility persistence and the GARCH process considers only short run dynamics for conditional variances. To describe the features that distinguish the FIGARCH from the GARCH and the IGARCH process, we consider the impulse response functions. Another expression for the FIGARCH process is

$$\{1-\phi(L)\}(1-L)^d \varepsilon_t^2 = \omega + \{1-\beta(L)\}\upsilon_t \tag{2.13}$$

where $\upsilon_t = \varepsilon_t^2 - \sigma_t^2$. Analogously to the impulse response function for the ARFIMA process mentioned above, we express the first differenced ε_t^2 as,

$$(1-L)\varepsilon_t^2 = \omega + \gamma(L)\upsilon_t . \tag{2.14}$$

Then, we have the impulse response weights for the FIGARCH process such that

$$\gamma(L) = (1-L)^{1-d}\phi(L)^{-1}\{1-\beta(L)\}. \tag{2.15}$$

By the same token, the impulse response weights for the GARCH process and the IGARCH are $\gamma(L) = (1-L)\{1-\alpha(L)-\beta(L)\}^{-1}\{1-\beta(L)\}$ and $\gamma(L) = \{1-\beta(L)\}$, respectively. Since the limit of the cumulative impulse response weights is $\gamma(1)$, the impact of past shocks on the FIGARCH volatility process from equation (2.15) is zero, as for the GARCH process. Note that $\gamma(1) = 1-\beta > 0$ particularly for the IGARCH $(1,1)$. Further, we consider the FIGARCH cumulative impulse response weights for lag j as

$$\lambda_j = \sum\nolimits_{0,j} \gamma_j . \tag{2.16}$$

According to Stirling's approximation, the cumulative impulse response weight for lag j, λ_j for the FIGARCH process is

$$\lambda_j \approx \left[(1-\beta)/\Gamma(d)\right] j^{d-1}. \tag{2.17}$$

Therefore, the hyperbolic decay component is present in the cumulative impulse response weights so that a shock to the squared residuals will decay at a very slow rate although all the past shocks eventually will die out.

Another class of models to describe long memory volatility was suggested by Breidt, Crato, and de Lima (1993) and Harvey (1998). They model long memory for conditional volatility series as follows:

$$y_t = \sigma_t \xi_t \tag{2.18}$$

and

$$\sigma_t^2 = \sigma^2 \exp(h_t), \tag{2.19}$$

where ξ_t is normal and independently distributed. Estimation of the stochastic volatility model uses the state space representation and Quasi-Maximum Likelihood Estimation (QMLE) via the Kalman filter.

In this chapter, by using both parametric and semi-parametric models, we investigate whether daily commodity cash return volatility follows the long memory process. The FIGARCH model is used to identify the long memory behavior in daily cash return volatility parametrically while the local Whittle estimation method is chosen for a semi-parametric counterpart to estimate the long memory parameter for the absolute daily cash returns.

2.3. Application to Daily Cash Returns

In the previous section, we considered the theoretical relationship among the long memory process, the self-similarity property, and temporal aggregation. We plan to study the theoretical relationship empirically by using daily cash price data for various types of commodities: corn, soybeans, live cattle, live hogs, gasoline, and gold.

We apply the FIGARCH model to the temporally aggregated daily cash returns to analyze the return volatility temporal patterns across various daily sample frequencies. We choose one-day through five-day sample frequencies because 5 trading days usually form a week of business days. We temporally aggregate returns $R_t^{(k)} \equiv \sum_{l=0,(k-1)} R_{tk-l}$ at k-daily frequency by summing one-day cash returns over k-daily periods for $k = 1, 2, 3, 4, 5$. For the conditional mean, we choose MA(1) to capture the usually small but significant autocorrelations of return levels at the first few lags[2]. The generic MA(q)-FIGARCH(p,d,q) model to estimate for daily cash returns is as follows:

[2] In the high frequency context, this MA(1) term is related to the market microstructure noise issue. We will discuss this more in chapter 3 where high frequency commodity futures returns are considered.

$$y_t = 100\Delta\ln(P_t) = \mu + \varepsilon_t + \theta\varepsilon_{t-1} \tag{2.20}$$

$$\varepsilon_t = z_t\sigma_t \tag{2.21}$$

$$\sigma_t^2 = \omega + \beta(L)\sigma_t^2 + \left[1 - \beta(L) - (1 - \phi(L))(1 - L)^d\right]\varepsilon_t^2 \tag{2.22}$$

where P_t is the commodity cash price, z_t is an i.i.d.(0,1) random variable, $\beta(L) = \beta_1 L + \beta_2 L^2 + \cdots + \beta_p L^p$, $\phi(L) = \phi_1 L + \phi_2 L^2 \cdots + \phi_q L^q$, and L is the lag operator.

Before proceeding further, we will briefly describe some characteristics of the commodities considered here. Daily prices for cash commodities are cash prices for the delivery location and specifications included in the corresponding futures contracts. These were obtained from the Futures Industry Institute data center. The agricultural product cash markets for corn and soybeans especially seem to display different volatility patterns due to their inherent attributes. Figure 2-1 plots the sample autocorrelations for absolute returns of daily cash prices for all the commodities considered. The horizontal axis represents daily lags up to 1000 days in order to consider approximately four years of trading days. In particular, the unique patterns of corn and soybean cash return volatility in their sample autocorrelations are worth notice. In figure 2-1, the dotted line represents the sample autocorrelations of the absolute (raw) *cash* returns for corn and soybeans. As shown in the figure, there seems to be some pronounced yearly seasonality for the original daily cash return volatility for corn and soybeans. The peaks are observed almost every 250-day interval, which approximately corresponds to a year of trading days. Typical yearly planting and harvesting cycles for the agricultural products may be responsible for the seasonality. Such seasonality may impede proper analysis of inherent volatility patterns. To cope with the annual periodic patterns in daily cash return volatility, we apply the Fourier flexible functional filtering, as introduced by Gallant (1981). The Fourier flexible functional filtering is formally discussed in chapter 3 when we consider high frequency commodity futures price data, since we apply the FFF filtering to cope with strong intraday seasonality for all of the commodities. For the other types of commodities, seasonal patterns are not observed for the sample autocorrelations of absolute daily cash returns. Hence, the other commodities do not require filtering before we apply the FIGARCH model to those time series data. The solid lines for the correlograms of corn and soybeans represent the sample correlations for the filtered returns for those commodities. As shown in Figure 2-1, seasonal patterns seem to be markedly reduced by the FFF filtering. Another striking feature is the unusual autocorrelation patterns of live cattle cash returns. The sample autocorrelations of live cattle absolute cash returns appear to be very different from the others and repeatedly deviate very much from zero.

Tables 2-1 through 2-6 present the results of applying the FIGARCH model to cash returns for (filtered) corn, (filtered) soybeans, gasoline, cattle, hogs, and gold at various daily frequencies. Specification tests are performed by applying the Ljung-Box portmanteau statistic on the standardized residuals resulting from quasi-maximum likelihood estimation for the FIGARCH model on the grounds that the test statistics asymptotically follow χ^2_{m-k} distribution.

The estimated long memory parameter in Tables 2-1 through 2-6 is strongly statistically significant for the cash return series for corn, soybeans, gasoline, and gold. In contrast, the long memory estimates for live cattle and hogs seem to be less significant than those for the other commodities. In particular, daily cash prices for live cattle seem to be constant for Wednesdays, Thursdays, and Fridays mainly, while most of the daily cash price changes seem to occur on Mondays and sometimes on Tuesdays, according to our preliminary data analysis. This odd data feature may be responsible for the unusual sample autocorrelation patterns as shown in Figure 2-1. The cash prices for live hogs also seem to involve some unusual characteristics. Although the live hog cash price changes are found quite evenly over the week's days, the changes seem to have strong day-of-week effects. To capture possible day-of-week effects on daily cash price changes, we include dummy variables for Monday, Tuesday, Thursday, and Friday in the conditional variance specification. From our pre-estimation, we found that there are considerable day-of-week effects for live hog cash return volatility. The robust *t*-values for Monday, Tuesday, Thursday, and Friday[3] dummies are 3.867, 4.797, 1.694, and 2.507, respectively. Also, the mean level of live hog daily cash returns exhibits a significant level of serial correlations during the course of the MA-FIGARCH estimation. To capture such a strong serial correlation in the mean level of live hog cash returns, we impose MA(15)[4] for the conditional mean model.

Apart from the unusual features mentioned above for the livestock, the long memory estimates from the FIGARCH conditional variance specification from (2.20) to (2.22) seem to be significant, and the model performs fairly in fitting the daily cash return

[3] To avoid a dummy trap, we drop dummies for Wednesday.
[4] Our informal experiment revealed that beyond 15 time lags did not seem to be statistically significant.

volatility. Particularly, we practiced a robust Wald test of the stationary GARCH(1,1)[5] null hypothesis versus a FIGARCH(1,d,1) alternative hypothesis. Under the null, the robust Wald test statistic *W* will have an asymptotic χ_1^2 distribution. Especially for a one-day sample frequency, we reject the null hypothesis for $d = 0$, and thus the GARCH(1,1) model is rejected for most of the commodities, with the exception of the livestock. For the crops, gasoline, and gold, at many temporal aggregation levels the formal statistical test supports the conclusion obtained both here and in Jin and Frechette (1994) that the FIGARCH is superior to the GARCH for modeling commodity return volatilities[6]. On the other hand, the *W* statistics for live cattle and hogs seem to be extremely low and less likely to reject the null hypothesis of GARCH specification at a 5-day (i.e., weekly) sample frequency. Again, this feature can be attributed to inactive spot market trading and the possible day-of-week effects for the livestock discussed above.

In addition, the long memory estimate levels themselves appear to be very stable across different sample frequencies for most of the commodities, with few exceptions. Our results imply that conditional variances of daily cash returns for each commodity may demonstrate a similar degree of persistence at different sample frequencies. This finding seems to be supportive of the self-similarity property discussed in section 2.

The semi-parametric local Whittle estimation methods have been suggested by Kunch (1987) and Robinson (1995). As a robustness check for the FIGARCH estimation results, we apply the local Whittle estimation for the long-memory parameter by using the absolute daily cash returns. One of the motivations for the semi-parametric

[5] For some instances, we test the null hypothesis for different GARCH specification other than GARCH(1,1).
[6] In fact, Jin and Frechette (1994) have used commodity futures price data.

estimation method is that, while the long memory volatility parameter estimation results using parametric models such as ARFIMA or FIGARCH specification may be affected by any possible short run dynamics, the semi-parametric estimation method affords general treatment of short run temporal dependence[7]. We discuss the local Whittle estimation separately in more detail in chapter 3. Table 2-7 reports the estimates of the long memory parameter by using absolute daily cash returns.

For the absolute daily cash returns, the semi-parametric long memory estimates seem to be qualitatively similar to the FIGARCH estimation results. For example, the low long memory estimate levels for live cattle and hogs can be found for the local Whittle estimation results similarly as in the FIGARCH long memory estimates. Also, the local Whittle estimates for the long memory parameter seem to be stable, as we found from the FIGARCH estimation results, and supportive of self-similarity for temporally aggregated absolute returns, as the FIGARCH estimates are stable across different sample frequencies.

2.4. Conclusion

The long run volatility dynamics for prices of physical commodities have been considered in this chapter. By using both parametric and semi-parametric long memory models, we confirmed that long memory exists for daily cash return volatility and, further, that the long memory behaviors are consistently witnessed across various daily frequencies for most of the commodities. We observed this evidence for temporally aggregated absolute returns and squared returns in common. This feature is consistent

[7] In general, semiparametric estimation methods may be somewhat controversial due to their poor performance in terms of bias and mean squared error.

with the theoretical self-similarity property of long memory, which implies that the autocorrelation of the long memory process decays at the same rate regardless of the sample frequency. Despite distinct aspects of commodity cash markets, the cash return volatility seems to exhibit the long memory property with exceptions only for livestock, as found in previous studies for many financial markets.

More practically, a proper understanding of cash price risks is important information for the hedge ratio of commodity futures, since the optimal hedge ratio is the conditional covariance between cash and futures returns divided by the conditional variance of futures returns. Therefore, studies of conditional moments of cash price change are very related to futures hedge modeling. Analysis of commodity futures return volatility, using both daily and high frequency return data, follows this chapter.

Table 2-1: Estimated MA-FIGARCH Models for Daily Cash Returns for Corn

(The sample period: 1/02/80 – 3/30/01)

	1 day	**2 days**	**3 days**	**4 days**	**5 days**
T	5362	2681	1787	1340	1072
μ	0.0009 (0.0018)	0.0025 (0.0037)	0.0035 (0.0055)	0.0042 (0.0074)	0.0032 (0.0095)
θ	0.0215 (0.0162)	0.0006 (0.0220)	0.0162 (0.0259)	0.0237 (0.0313)	0.0269 (0.0321)
d	0.2720 (0.0438)	0.2992 (0.0675)	0.2641 (0.0618)	0.3215 (0.0808)	0.2702 (0.0827)
ω	0.0026 (0.0008)	0.0050 (0.0017)	0.0086 (0.0032)	0.0102 (0.0038)	0.0156 (0.0066)
β	0.1730 (0.0470)	0.1607 (0.0702)	0.1170 (0.0787)	0.1040 (0.0917)	0.0864 (0.0916)
m_3	-0.500	-0.471	-0.372	-0.394	-0.455
m_4	6.463	5.534	4.505	5.514	5.919
Q(20)	31.662	28.498	26.122	23.336	16.723
$Q^2(20)$	5.745	7.504	11.648	8.343	8.542
W	38.492	19.631	18.286	15.846	10.666

Key: ln(L) is the value of the maximized log likelihood; Q(20) and $Q^2(20)$ are the Ljung-Box statistics with 20 degree of freedom based on the autocorrelations of the standardized residuals and autocorrelations of the squared standardized residuals. The sample m_3 and m_4 are also based on the standardized residuals. T is the number of observations.

Table 2-2: Estimated MA-FIGARCH Models for Daily Cash Returns for Soybean (The sample period: 1/02/80 – 12/29/00)

	1 day	2 days	3 days	4 days	5 days
T	5300	2650	1766	1325	1060
μ	-0.0020	-0.0017	-0.0014	-0.0003	0.0024
	(0.0016)	(0.0031)	(0.0046)	(0.0061)	(0.0076)
θ	-0.0336	-0.0368	-0.0368	-0.0474	-0.0677
	(0.0155)	(0.0210)	(0.0254)	(0.0309)	(0.0303)
d	0.3291	0.3397	0.3904	0.2899	0.3498
	(0.0488)	(0.0649)	(0.1103)	(0.0671)	(0.0988)
ω	0.0016	0.0029	0.0036	0.0080	0.0078
	(0.0004)	(0.0009)	(0.0015)	(0.0033)	(0.0035)
β	0.2723	0.2753	0.3189	0.1102	0.1999
	(0.0620)	(0.0754)	(0.1311)	(0.0910)	(0.1201)
m_3	-0.265	-0.256	-0.005	0.080	0.059
m_4	5.152	4.538	3.772	3.509	3.761
Q(20)	22.361	26.364	15.869	19.815	21.595
$Q^2(20)$	34.198	25.785	21.693	21.277	24.398
W	45.485	27.369	12.532	18.657	12.528

Key: As for table 2-1

Table 2-3: Estimated MA-FIGARCH Models for Daily Cash Returns for Live Cattle

(The sample period: 1/02/80 – 12/29/00)

	1 day	**2 days**	**3 days**	**4 days**	**5 days**
T	4800	2400	1600	1200	960
μ	0.0037 (0.0233)	0.0019 (0.0195)	0.0122 (0.0275)	-0.0024 (0.1231)	-0.0018 (0.0619)
θ	0.0297 (0.0161)	-0.0091 (0.0169)	-0.0421 (0.0293)	-0.0493 (0.0357)	-0.0855 (0.0363)
d	0.1768 (0.0930)	0.1534 (0.0668)	0.1385 (0.0696)	0.0661 (0.0819)	0.0668 (0.0792)
ω	0.1546 (0.1050)	0.6173 (0.3500)	1.0479 (0.5645)	1.2694 (0.9834)	2.6452 (1.6237)
β	0.5832 (0.0679)	0.1557 (0.0656)	0.1354 (0.0638)	0.3828 (0.4618)	0.1077 (0.6031)
ϕ	0.4086 (0.0643)	______	______	0.4615 (0.4907)	0.1891 (0.6487)
m_3	-1.538	-0.873	-0.630	-0.553	-0.426
m_4	40.253	18.439	11.856	9.119	7.564
Q(20)	28.860	18.392	25.561	27.665	41.104
$Q^2(20)$	24.741	18.294	9.771	14.416	10.133
W	3.167	5.269	3.962	0.652	0.711

Key: As for table 2-1

Table 2-4: Estimated MA-FIGARCH Models for Daily Cash Returns for Live Hogs

(The sample period: 1/02/80 – 12/29/00)

	1 day[8]	2 days	3 days	4 days	5 days
T	4551	2275	1517	1137	910
μ	-0.0037	-0.0152	-0.0212	-0.0174	0.0177
	(0.0279)	(0.0578)	(0.0994)	(0.1071)	(0.0860)
θ_1	-0.1524	-0.2878	-0.0692	0.0120	0.1001
	(0.0170)	(0.0277)	(0.0284)	(0.0335)	(0.0369)
θ_2	-0.2078	0.1482	0.0830	0.1104	0.1384
	(0.0153)	(0.0232)	(0.0275)	(0.0319)	(0.0343)
θ_3	0.0463	0.0326	0.1200	0.0832	0.0086
	(0.0159)	(0.0228)	(0.0268)	(0.0312)	(0.0346)
θ_4	0.1229	0.0854	0.0862	0.0158	0.0006
	(0.0159)	(0.0229)	(0.0274)	(0.0306)	(0.0341)
θ_5	-0.0096	0.0734	0.0033	0.0147	-0.0947
	(0.0165)	(0.0232)	(0.0269)	(0.0329)	(0.0341)
θ_6	-0.0061	-0.0115	0.0426	-0.0253	-0.0444
	(0.0157)	(0.0234)	(0.0280)	(0.0350)	(0.0347)
θ_7	0.0562	0.0514	-0.0304	-0.0402	0.0230
	(0.0158)	(0.0225)	(0.0284)	(0.0350)	(0.0362)
θ_8	0.0394	-0.0287	0.0223	-0.0161	-0.0255
	(0.0171)	(0.0224)	(0.0277)	(0.0368)	(0.0363)
θ_9	0.0510	0.0562	-0.0412	0.0341	0.0337
	(0.0144)	(0.0225)	(0.0308)	(0.0333)	(0.0363)
θ_{10}	0.0196	-0.0223	0.0327	0.0195	-0.0246

[8] For live hogs at one-day sample frequency, we could cope with higher $Q^2(20)$ statistics by including day-of-week dummy variables. All the coefficient estimates are 0.9437 with standard error, 0.2440 for Monday; 1.3074 with standard error, 0.2725, for Tuesday; 0.3487 with standard error, 0.2058 for Thursday; and 0.6577 with standard error, 0.2623 for Friday. We do not consider such day-of-week effects since they seem to collapse by temporal aggregation beyond one-day sample frequency.

	(0.0174)	(0.0233)	(0.0321)	(0.0320)	(0.0342)
θ_{11}	-0.0151	-0.0014	-0.0301	0.0118	-0.1198
	(0.0159)	(0.0238)	(0.0301)	(0.0357)	(0.0422)
θ_{12}	-0.0155	-0.0094	-0.0214	-0.0250	-0.1273
	(0.0165)	(0.0222)	(0.0295)	(0.0340)	(0.0404)
θ_{13}	0.0496	-0.0305	0.0731	-0.0002	-0.0421
	(0.0158)	(0.0217)	(0.0289)	(0.0353)	(0.0407)
θ_{14}	0.0434	0.0049	0.0326	-0.0337	-0.1069
	(0.0153)	(0.0209)	(0.0279)	(0.0339)	(0.0377)
θ_{15}	-0.0498	-0.0127	-0.0055	-0.0730	-0.0202
	(0.0153)	(0.0212)	(0.0300)	(0.0342)	(0.0309)
d	0.2083	0.1789	0.1361	0.1570	0.0981
	(0.0548)	(0.0395)	(0.0411)	(0.0574)	(0.1203)
ω	0.1718	1.8094	3.0760	0.1614	1.1651
	(0.1838)	(0.4762)	(0.8904)	(0.0798)	(2.2573)
β_1	0.5239	0.0914	0.0197	0.9654	0.7838
	(0.2096)	(0.0442)	(0.0524)	(0.0179)	(0.3272)
β_2[9]	0.0230	______	______	______	______
	(0.0266)				
ϕ.	0.4150	______	______	0.9498	0.8272
	(0.1801)			(0.0263)	(0.2455)
m_3	-0.044	-0.001	-0.146	0.130	0.008
m_4	3.544	3.296	3.762	3.476	3.266
Q(20)	17.972	2.9234	9.510	11.983	6.753
$Q^2(20)$	29.283	21.082	24.578	21.542	22.423
W	27.616	20.516	10.967	7.475	0.665

Key: As for Table 2-1

[9] FIGARCH (2, d, 1) seems to fit the live hog cash daily return volatility at one-day sample frequency fairly relative to FIGARCH (1, d, 1) or FIGARCH (1, d, 0) conditional variances specifications.

Table 2-5: Estimated MA-FIGARCH Models for Daily Cash Returns for Gasoline
(The sample period: 1/02/91 – 12/29/00)

	1 day	2 days	3 days	4 days	5 days
T	2509	1254	836	627	501
μ	-0.0192	-0.0632	-0.0970	-0.1201	-0.1151
	(0.0451)	(0.0949)	(0.1442)	(0.1834)	(0.2230)
θ	0.0926	0.0574	-0.0107	-0.0781	-0.0829
	(0.0215)	(0.0329)	(0.0429)	(0.0458)	(0.0487)
d	0.2900	0.2968	0.3715	0.2309	0.2105
	(0.0694)	(0.0979)	(0.3187)	(0.1524)	(0.0845)
ω	0.8453	1.9729	1.0488	5.8006	7.9115
	(0.2650)	(0.7972)	(1.3298)	(4.9808)	(4.1986)
β	0.1726	0.1372	0.6309	0.1930	0.1898
	(0.0770)	(0.1211)	(0.1643)	(0.2527)	(0.1424)
ϕ	______	______	0.4135	______	______
			(0.1720)		
m_3	-0.103	-0.389	-0.214	-0.289	-0.160
m_4	4.739	4.320	3.734	3.931	3.371
Q(20)	28.076	24.839	32.326	23.792	26.302
$Q^2(20)$	17.416	16.398	9.190	15.708	17.395
W	17.463	23.160	1.359	2.295	6.204

Key: As for table 2-1

Table 2-6: Estimated MA-FIGARCH Models for Daily Cash Returns for Gold

(The sample period: 1/02/80 – 12/29/00)

	1 day	**2 days**	**3 days**	**4 days**	**5 days**
T	5283	2641	1761	1320	1056
μ	-0.0167	-0.0316	-0.0677	-0.0796	-0.0643
	(0.0096)	(0.0207)	(0.0298)	(0.0392)	(0.0559)
θ	-0.0583	-0.0124	-0.0043	-0.0132	-0.0309
	(0.0169)	(0.0301)	(0.0298)	(0.0356)	(0.0360)
d	0.2905	0.3438	0.2942	0.4093	0.3160
	(0.0351)	(0.0574)	(0.0434)	(0.0953)	(0.1374)
ω	0.0755	0.1374	0.2068	0.1316	0.4705
	(0.0250)	(0.0572)	(0.1155)	(0.1176)	(0.5367)
β	0.1512	0.1787	0.1071	0.2872	0.2477
	(0.0490)	(0.0694)	(0.1134)	(0.1398)	(0.1521)
m_3	0.086	0.959	0.419	0.637	1.411
m_4	9.563	15.589	8.085	10.058	18.805
Q(20)	39.257	19.918	18.735	12.869	22.962
$Q^2(20)$	11.015	2.982	10.029	6.215	8.135
W	68.634	35.854	45.890	68.634	5.291

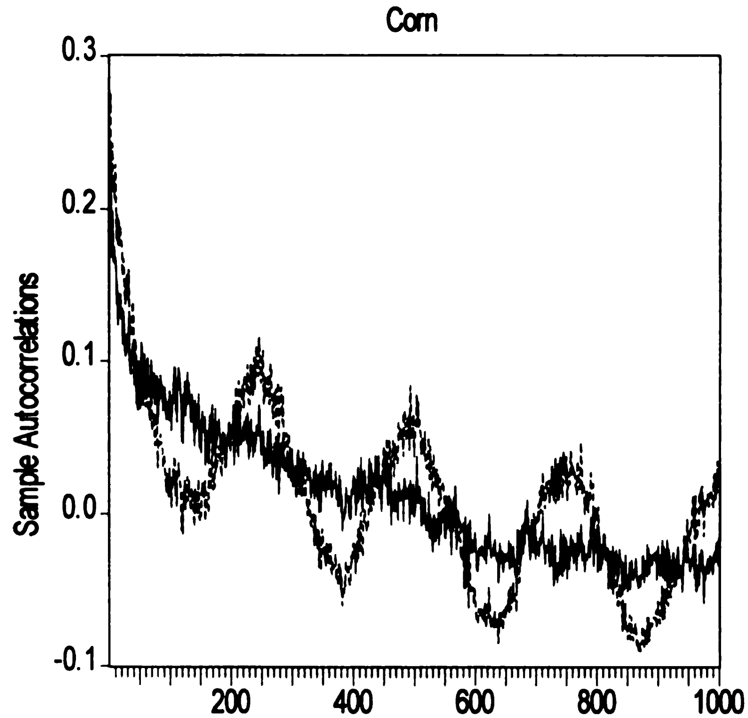
Key: As for table 2-1

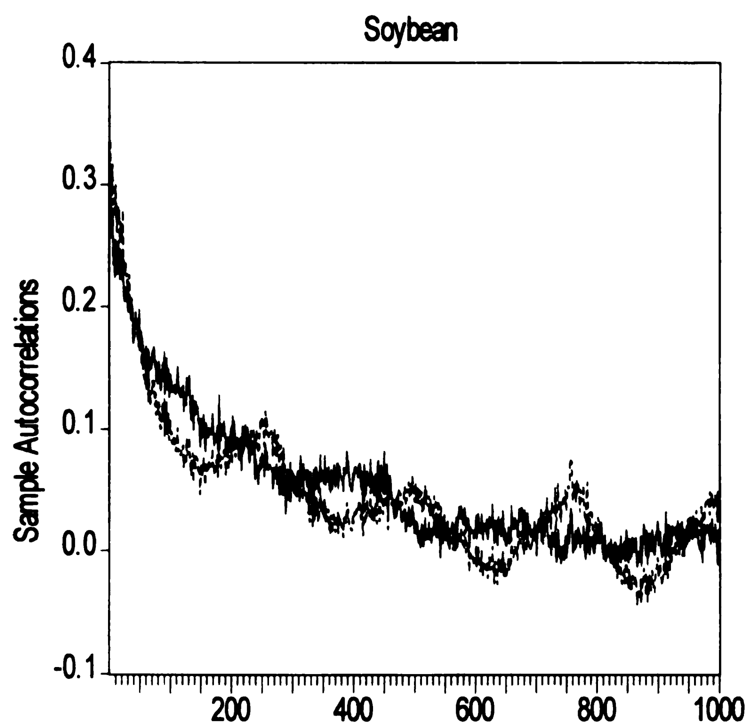
Table 2-7. Semi-Parametric Long Memory Parameter Estimation: Absolute Daily Cash Returns at Different Daily Sample Frequencies.

	1 day	2 days	3 days	4 days	5 days
Corn (Filtered)					
Local Whittle	0.3397	0.3633	0.3238	0.2823	0.2536
	(0.0376)	(0.0471)	(0.0539)	(0.0592)	(0.0635)
Soybean (Filtered)					
Local Whittle	0.3853	0.4693	0.4629	0.4568	0.4282
	(0.0378)	(0.0474)	(0.0541)	(0.0592)	(0.0638)
Live Cattle					
Local Whittle	0.1534	0.1457	0.1452	0.1995	0.1992
	(0.0390)	(0.0489)	(0.0599)	(0.0612)	(0.0660)
Live Hog					
Local Whittle	0.2418	0.2562	0.2224	0.1947	0.1492
	(0.0397)	(0.0497)	(0.0569)	(0.0625)	(0.0672)
Gasoline					
Local Whittle	0.2899	0.3027	0.3590	0.4061	0.4133
	(0.0481)	(0.0603)	(0.0689)	(0.0760)	(0.0818)
Gold					
Local Whittle	0.4436	0.4817	0.4626	0.4810	0.5073
	(0.0378)	(0.0474)	(0.0541)	(0.0595)	(0.0638)

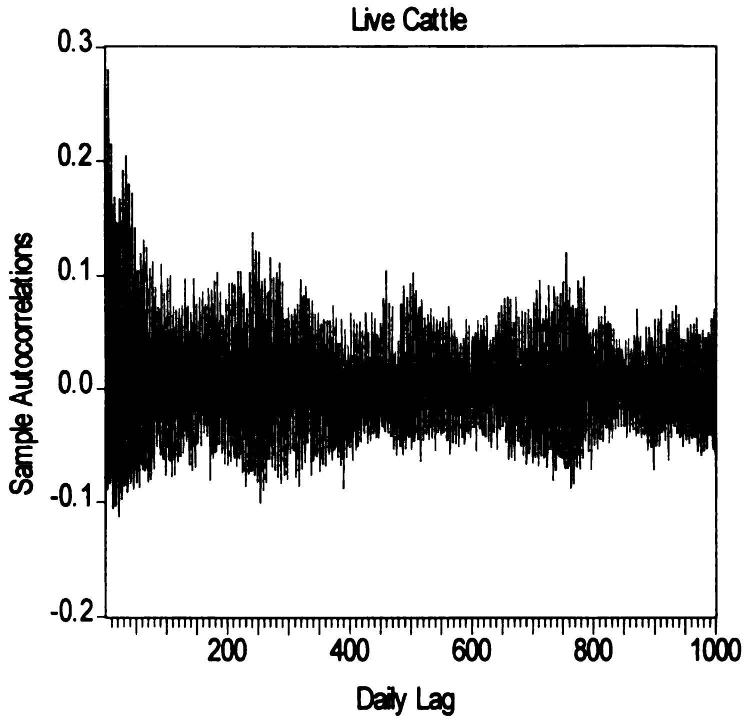
Key: Asymptotic standard errors are in parentheses below corresponding parameter estimates.

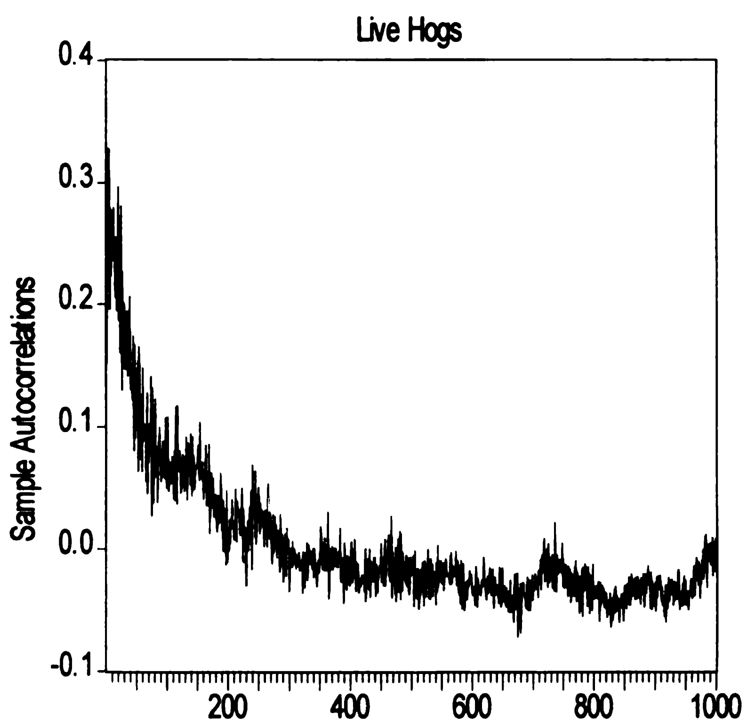
Fig. 2-1 Correlograms for Absolute Daily Cash Returns



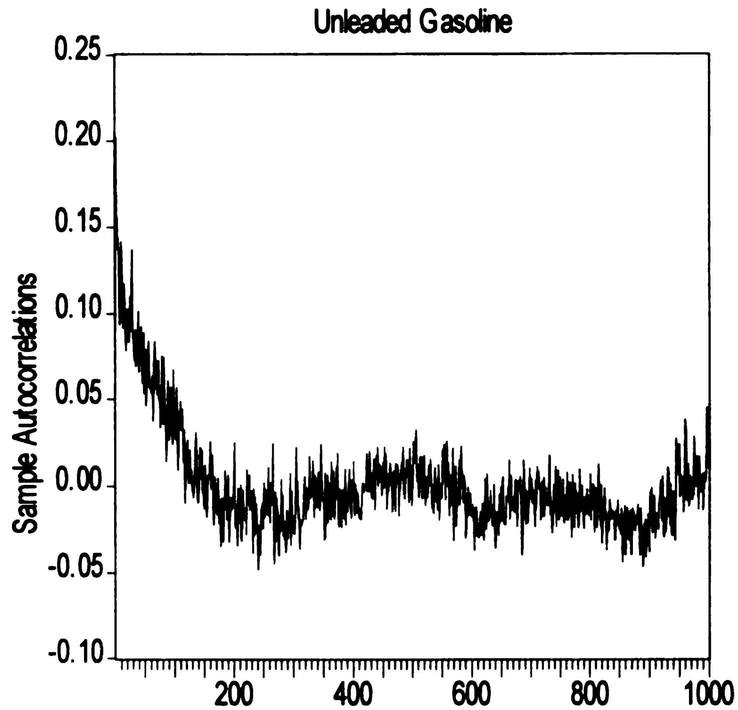


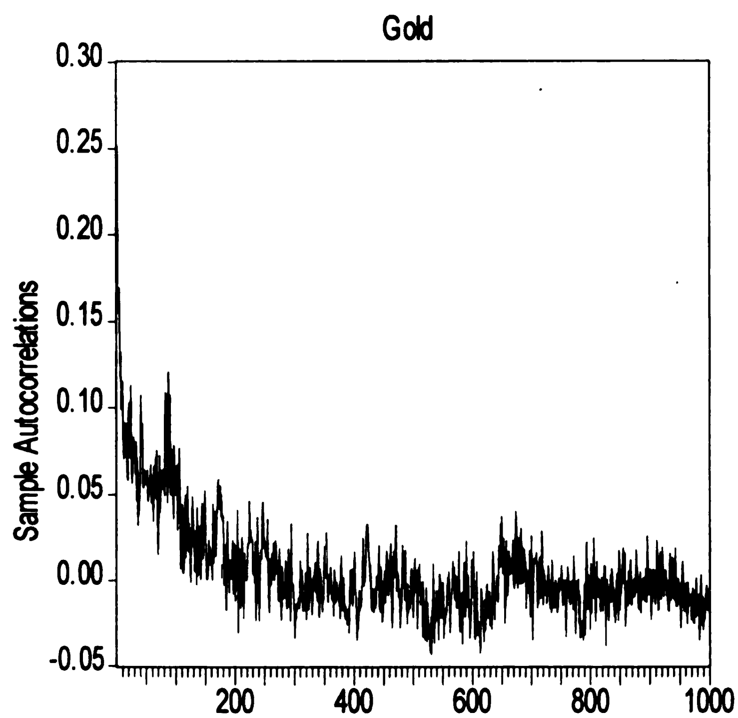
Key: Dotted line and solid line indicate the sample autocorrelations for absolute daily raw and filtered cash returns.





Key: Dotted line and solid line indicate the sample autocorrelations for absolute daily raw and filtered cash returns.





Key: Dotted line and solid line indicate the sample autocorrelations for absolute daily raw and filtered cash returns.

CHAPTER 3

MODELING DAILY AND HIGH FREQUENCY COMMODITY FUTURES RETURNS

3.1. Introduction

This chapter is concerned with the stochastic properties of commodity futures prices and applies some recent developments in volatility modeling, in particular the FIGARCH long memory volatility model, to commodity futures returns. The volatilities of daily futures returns are found to be well described by the FIGARCH model, with relatively similar estimates of the long memory parameter across commodities. The conditional means of the daily returns are close to being uncorrelated, with small departures from martingale behavior being represented by low order moving average models. We also estimate FIGARCH models for high frequency commodity futures returns based on intra-day tick data. These high frequency commodity returns are dominated by strong intra-day periodicity, hypothesized to be a result of repeated trading day cycles resulting from the institutional features of the futures exchanges where trades are taking place. The intra-day periodicity is removed using a deterministic Fourier Flexible Form (FFF) filter. The filtered high frequency futures returns are also well described by the FIGARCH process. The results of the chapter have important implications for our understanding of the stochastic properties of commodity prices, and hence for empirical applications such as optimal hedge ratio estimation, tests for futures market efficiency, tests for the announcement effect of market news, option valuation, farm risk portfolio management, etc.

The FIGARCH model has already been applied gainfully to exchange rates, stock returns, inflation rates, and a range of other economic data; for examples see Baillie, Bollerslev and Mikkelsen (1996), Bollerslev and Mikkelsen (1996), Baillie, Han and Kwon (2002), etc. However, there have been few applications of the model to commodities. Crato and Ray (2000) study long memory in the daily volatilities of several agricultural commodity futures returns, along with a stock index return, currencies, metals, and heating oil. They find strong evidence of long memory in daily commodity futures prices, though they do not explicitly estimate FIGARCH models. Jin and Frechette (2004) estimate FIGARCH volatility models for 14 agricultural futures series and find that FIGARCH fits the data significantly better than a traditional GARCH volatility model. While these studies have provided valuable information on the long memory properties of commodity futures price volatilities, much more work remains to be done.

This chapter adds to our understanding of long memory in commodity price volatilities in three main ways. First, while Jin and Frechette (2004) argue in favor of the FIGARCH model over the GARCH model for commodity futures volatilities, they did not undertake a formal statistical test comparing the two models. Here we undertake a robust Wald test, which formally compares the fit of the GARCH and FIGARCH models. Second, in addition to the standard quasi-maximum likelihood estimator (QMLE), we also apply the semi-parametric local Whittle estimator of the long memory parameter. This provides additional information on the robustness of long-memory inferences concerning daily commodity price volatilities. Third, in addition to daily returns we study high frequency returns on futures contracts using intra-day tick data. This study is the first to systematically examine volatility using high frequency commodity futures data.[10] We find that estimated models at different sampling frequencies are consistent with the theory that commodity futures returns are "self-similar" processes, and hence have long memory parameters that are invariant to the sampling frequency; see Beran (1994). The "self-similarity" of the estimates of the long memory volatility parameter across relatively short spans of high frequency data strongly suggests that the long memory property is an intrinsic feature of the system rather than being caused by exogenous shocks or regime shifts.

The plan of the rest of the chapter is as follows. Section 2 discusses the application of the long memory FIGARCH volatility model to daily futures returns. Similar to Jin and Frechette (2004), we find the FIGARCH models to be econometrically superior to regular stable GARCH models. Section 3 describes the results from the analysis of high frequency futures returns and compares them to the daily return results. Section 4 presents an analysis of semi-parametric local Whittle estimation of the long memory parameter as a robustness check, and also compares estimates of the long memory parameter across a range of different sampling frequencies. This shows that the commodity return series display self-similarity. Section 5 offers a brief conclusion.

[10] Cai, Cheung and Wong (2001) have analyzed high frequency gold futures. However, their approach is somewhat informal and does not include either FIGARCH or local Whittle estimation of the long memory parameter.

3.2. Analysis of Daily Commodity Returns

This section is concerned with the analysis of daily futures returns for different commodities. We examine six commodities: corn, soybeans, cattle, hogs, gasoline, and gold. Corn and soybeans are major annual crops that are of critical importance to U.S. agriculture. These crops are related in the sense that they can be substitutes in production and both are used heavily as animal feed. They are different, however, in that most corn is produced in the northern hemisphere, while soybeans have a significant southern hemisphere harvest in Brazil and Argentina. This southern hemisphere harvest may influence seasonal price and volatility patterns. Cattle and hogs are both important livestock commodities in U.S. agriculture, but their different life cycles mean different inherent price dynamics, even though we would expect a lot of similarity in the stochastic properties of prices for these two livestock commodities. Gasoline is included to see if results are markedly different for a natural resource-based commodity, and gold is included as a commodity that has a central role as a store of wealth.

Data were obtained from the Futures Industry Institute data center.[11] The daily data are daily closing futures prices on major U.S. futures markets for the relevant commodity, in particular, the Chicago Board of Trade for corn and soybeans, the Chicago Mercantile Exchange for live cattle and hogs, and the New York Mercantile Exchange for unleaded gasoline and gold. Returns are defined in the conventional manner as continuously compounded rates of return and calculated as the first difference of the natural logarithm of prices. To compute the futures returns, nearby contracts were used,

[11] The Futures Industry Institute is now called the Institute for Financial Markets. For more information and data availability, see http://www.theifm.org.

and then the data was switched to the next available contract nearby on the first day of the month in which the current nearby contract expires. For consistency, returns are always defined using the same futures contract.[12] The use of nearby futures contracts to define our futures return series has the advantage that we are using the most actively traded contracts to generate our return data. However, if volatility depends on time to maturity, as might be expected in at least some instances, then switching from an expiring futures contract to the next nearby maturity may introduce jumps into the volatility process because of jumps in time to maturity at the switch points. We will discuss how we allowed for the effects of these jumps in time to maturity when we outline the econometric model further below.

The details of the sample periods used for each commodity are provided in Table 3-1, along with some summary statistics for daily returns over these periods. All the daily data begin at the first trading day of January 1980, except for gasoline. For gasoline, we exclude data from January 1980 through December 1990 and begin the sample period the first trading day of January 1991. This is to avoid two periods of exceptional volatility in gasoline prices that we argue are a result of structural shifts in the volatility process for this commodity. The first period is 1986-87, a period in which Saudi Arabia expanded its oil production significantly in order to discipline other OPEC countries. The second period extends from August 1990 to December 1990 and is caused by the Iraqi invasion of Kuwait and the subsequent Gulf War. By starting the gasoline price series in January of 1991 we avoid having to model these structural breaks in the volatility process. All of

---

12 That is, at each point when the data switch to the next nearby maturing contract, the futures return is defined as the difference between the natural logarithm of today's futures price for a contract maturing at the next nearby and yesterday's futures price for a contract with exactly the same maturity date. In this way, daily returns are never defined using prices from two different contracts with different maturity dates.

the daily data end at the last trading day of December 2000, except for corn, which ends the last trading day in March of 2001. In all cases we used the most recent data that was provided in the data set obtained from the Futures Industry Institute.

Previous studies by Cecchetti, Cumby and Figlewski (1988), Baillie and Myers (1991), and Yang and Brorsen (1992) have argued that most daily cash and futures commodity returns are well described as martingales with GARCH effects. The possibility of mixed diffusion-jump processes has also been suggested as a way to characterize volatility in commodity prices. Yang and Brorsen (1992) compared GARCH, mixed diffusion-jump, and deterministic chaos models of cash commodity prices and concluded that the GARCH volatility process provided the best fit. It is only more recently that studies such as Crato and Ray (2000) and Jin and Frechette (2004) have begun to investigate the long memory properties of commodity volatilities.

Figures 3-1 and 3-2 plot the sample autocorrelations for the returns, squared returns and absolute returns in daily futures prices for two representative commodities, namely live cattle and corn. There is one noticeable difference between the crop commodity and the livestock commodity, namely that: both squared and absolute daily returns for corn exhibit strong yearly seasonality in their sample autocorrelations, while this does not occur for live cattle. To conserve space, the corresponding graphs for the other commodities are not shown. However, it was observed that soybeans also display seasonality in volatility (though not as pronounced as in the case of corn, perhaps because of the influence of a southern hemisphere harvest for soybeans) while live hogs, gasoline and gold display no seasonality in volatility (similar to live cattle). In order to analyze the intrinsic stochastic properties of the daily corn and soybean return volatilities we filter out the seasonality by using a FFF filter.[13] The sample autocorrelations for the returns, squared returns and absolute returns for the filtered daily corn futures price series is provided in Figure 3-3. Notice that the FFF filter has been quite effective in removing the seasonality in the squared and absolute corn futures returns. In all subsequent analysis of the corn and soybean return volatilities we use the filtered volatility models.

Plots of the live cattle sample autocorrelations (Figure 3-1), the FFF filtered corn sample autocorrelations (Figure 3-3), and other commodity return sample autocorrelations (not shown) reveal a familiar lack of autocorrelation in returns and the marked persistence in autocorrelations of squared and absolute returns that was first noticed by Ding, Granger and Engle (1993) for the case of stock market returns. In particular, the autocorrelation functions for the squared and absolute returns do not display the usual exponential decay associated with the stationary and invertible class of ARMA models, but rather appear to be generated by a long memory process with hyperbolic decay.

More formally, the autocorrelation at lag k, ρ_k, tends to satisfy $\rho_k \approx ck^{2d-1}$ as k gets large, where c is a constant and d is the long memory parameter. This type of persistence is consistent with the notion of hyperbolic decay and is sometimes called the "Hurst phenomenon." The Hurst coefficient is defined as $H = d + 0.5$. If $d = 1$, so that $H = 1.5$, then the autocorrelation function does not decay and the series has a unit root. If $d = 0$, so that $H = 0.5$, then the autocorrelation function decays exponentially and the series is stationary. But for $0 < d < 1$, i.e. $0 < H < 1.5$, the series is sufficiently flexible to allow for slower hyperbolic rates of decay in the autocorrelations. While many stochastic

---

[13] See the appendix for the details of the FFF filter.

processes could potentially exhibit the long memory property, the most widely used process is the ARFIMA(p, d, q) process of Granger and Joyeux (1980), Granger (1980), and Hosking (1981). In the ARFIMA process, a time series x_t is modeled as $a(\mathrm{L})(1-\mathrm{L})^d x_t = b(\mathrm{L})\varepsilon_t$ with $a(\mathrm{L})$ and $b(\mathrm{L})$ being p'th and q'th order polynomials in the lag operator L, with all their roots lying outside the unit circle, while ε_t is a white noise process. The ARFIMA process is stationary and invertible in the region of $-0.5 < d < 0.5$. At high lags, the ARFIMA(p, d, q) process is known to have an autocorrelation function that satisfies $\rho_k \approx ck^{2d-1}$, so that the autocorrelations may decay at a slow hyperbolic rate, as opposed to the required exponential rate associated with the stationary and invertible class of ARMA models. The sample autocorrelation function of the squared and absolute daily filtered futures corn returns appears to be very consistent with the above properties, and analogous plots for the other commodity returns were found to be extremely similar.

Virtually all studies of daily asset returns, including commodity assets, have found return y_t to be stationary with small autocorrelations at the first few lags, which can be attributed to a combination of a small time-varying risk premium, bid-ask bounce, and/or non-synchronous trading phenomena; see Goodhart and O'Hara (1997) for a description of this issue in high frequency currency markets. On the other hand, volatility has been found to be very persistently autocorrelated with long memory hyperbolic decay. A model that is consistent with these stylized facts is the MA(n)-FIGARCH(p, d, q) process,

$$y_t = 100\Delta\ln(P_t) = \mu + b(L)\varepsilon_t, \tag{3.1}$$

$$\varepsilon_t = z_t \sigma_t, \tag{3.2}$$

and

$$[1-\beta(L)]\sigma_t^2 = \omega + \left[1 - \{1-\beta(L)\} - \phi(L)(1-L)^d\right]\varepsilon_t^2, \tag{3.3}$$

where P_t is the asset price, z_t is an i.i.d.(0,1) random variable, and the polynomial in the lag operator associated with the moving average process is $b(L) = 1 + b_1 L + b_2 L^2 + ... + b_n L^n$. The FIGARCH model in equation (3.3) can be best motivated from noting that the standard GARCH(p, q) model of Bollerslev (1986) can be expressed as

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2,$$

where the polynomials are $\alpha(L) \equiv \alpha_1 L + \alpha_2 L^2 + ... + \alpha_q L^q$ and $\beta(L) \equiv \beta_1 L + \beta_2 L^2 + ... + \beta_p L^p$. The GARCH(p, q) process can also be expressed as the ARMA[max(p, q), p] process in squared innovations as

$$[1-\alpha(L)-\beta(L)]\varepsilon_t^2 = \omega + [1-\beta(L)]\upsilon_t$$

where $\upsilon_t \equiv \varepsilon_t^2 - \sigma_t^2$ and is a zero mean, serially uncorrelated process which has the interpretation of being the innovations in the conditional variance. The FIGARCH(p, d, q) process in equation (3.3) can also be written as

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + [1-\beta(L)]\upsilon_t, \tag{3.4}$$

where $\phi(L) = [1-\alpha(L)-\beta(L)](1-L)^{-d}$ is a polynomial in the lag operator. Equation (3.4) can be easily shown to transform to equation (3.3), which is the standard representation for the conditional variance in the FIGARCH(p, d, q) process. Further details concerning the FIGARCH process can be found in Baillie, Bollerslev and Mikkelsen (1996). The parameter d characterizes the long memory property of hyperbolic decay in volatility because it allows for autocorrelation decay at a slow hyperbolic rate. The attraction of the FIGARCH process is that for $0 < d < 1$, it is sufficiently flexible to allow for intermediate ranges of persistence, between complete integrated persistence of volatility shocks associated with $d = 1$ and the geometric decay associated with $d = 0$.

The volatility model in equation (3.3) has to be slightly adjusted to accommodate the potential jumps in volatility that can occur at contract switching points, when futures return data are computed from a sequence of nearby futures contracts. The long spans of daily futures returns are constructed from contracts with different maturities, and the resulting variations (and jumps) in time to maturity may have an influence on the volatility process. To account for possible time to maturity effects we introduce a time to maturity variable in the formulation of the FIGARCH(1, d, 1) model in (3.3), which then becomes

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \gamma TM_t + [1 - \beta L - (1-\phi L)(1-L)^d]\varepsilon_t^2, \tag{3.5}$$

where TM represents the time to maturity on the contract used to construct the futures return for period t, and γ is the associated parameter.

The above model (3.1), (3.2), and (3.5) is estimated for futures returns on our six commodities of interest by maximizing the Gaussian log likelihood function

$$\ln(L;\Theta) = -(0.5T)\ln(2\pi) - 0.5\sum_{t=1}^{T}\left[\ln(\sigma_t^2) + \varepsilon_t^2\sigma_t^{-2}\right], \tag{3.6}$$

where $\Theta' = (\mu, \theta_1, ...\theta_n, \omega, \beta_1, ..\beta_p, \phi_1, ...\phi_l)$ is the vector of unknown parameters. However, it has long been recognized that most asset returns are not well represented by assuming z_t in equation (3.2) is normally distributed; for examples, see McFarland, Pettit and Sung (1982) and Booth (1987). Consequently, inference is usually based on the quasi-maximum likelihood estimator (QMLE) of Bollerslev and Wooldridge (1992), which is valid when z_t is non-Gaussian. Denoting the vector of parameter estimates obtained from maximizing (3.6), using a sample of T observations on equations (3.1), (3.2) and (3.5), with z_t being non-normal by $\hat{\Theta}_T$, then the limiting distribution of $\hat{\Theta}_T$, is

$$T^{1/2}(\hat{\Theta}_T - \Theta_0) \rightarrow N[0, A(\Theta_0)^{-1}B(\Theta_0)A(\Theta_0)^{-1}], \tag{3.7}$$

where A(.) and B(.) represent the Hessian and outer product gradient, respectively, and Θ_0 denotes the vector of true parameter values. Equation (3.7) is used to calculate the robust standard errors that are reported in the subsequent results in this chapter, with the Hessian and outer product gradient matrices being evaluated at the point $\hat{\Theta}_T$ for practical implementation.

Table 3-2 presents the results of applying the above model (3.1), (3.2), and (3.5) to daily futures returns for the six commodities discussed earlier. The exact parametric specification of the model which best represents the degree of autocorrelation in the conditional mean and conditional variance of daily commodity returns, vari by commodity. The exact model specification for each commodity is indicated by the number of non-zero estimates provided for the polynomial in the lag operator terms in Table 3-2. For corn and soybean futures returns, we apply FIGARCH estimation to the FFF filtered returns (see the Appendix). Results from Box-Pierce portmanteau statistics on the standardized residuals are at the bottom of the table. The standard portmanteau test statistic, $Q(m) = T(T+2)\sum_{j=1}^{m} r_j^2 /(T-j)$, where r_j is the j'th order sample autocorrelation from the residuals, is known to have an asymptotic χ^2_{m-k} distribution, where k is the number of parameters estimated in the conditional mean. Similar degrees of freedom adjustments are used for the portmanteau test statistic based on the squared standardized residuals when testing for omitted conditional heteroscedasticity. This adjustment is in the spirit of the suggestions by Diebold (1988) and others. The sample skewness and kurtosis of the standardized residuals (m_3 and m_4), are also provided at the bottom of Table 3-2.

The Ljung-Box portmanteau statistics show that the models specified for each commodity do a good job of capturing the autocorrelations in the mean and volatility of the commodity return series. In each case there is no evidence of additional autocorrelation in the standardized residuals or squared standardized residuals, indicating that the chosen model specification provides an adequate fit. It is interesting that autocorrelation in the mean tends to persist more for the livestock commodities of live cattle and hogs than for the other commodities (i.e., more MA terms in the mean are required for an adequate fit). Furthermore, these commodities also seem to require more flexible models to capture their autocorrelation in volatility as well (i.e., more GARCH terms required for an adequate fit). The standardized residuals from all commodities, except perhaps live cattle and hogs, exhibit the usual features of excess kurtosis of daily asset returns. However, this is accommodated through use of the QMLE standard errors for inference.

The estimated MA-FIGARCH models reported in Table 3-2 seem to fit the data well. For each commodity there is weak evidence of small moving average effects in the mean returns. As stated earlier, this may be attributed to a combination of a small time-varying risk premium, bid-ask bounce, and/or non-synchronous trading phenomena. The volatility autocorrelation parameters in $\beta(L)$ and $\phi(L)$ indicate strong evidence of significant serial correlation in volatilities, which is consistent with previous findings of autocorrelated volatility in commodity returns; see Baillie and Myers (1991), Jin and Frechette (2004), and Yang and Brorsen (1992). Furthermore, the time to maturity parameter is statistically significant for all commodities except gold. Gold may not experience a time to maturity effect in volatility because its special role as a store of wealth means that cash and futures prices move very closely together, irrespective of the time to maturity on the futures contract. It is interesting that the time to maturity effect is negative for corn, soybeans and gasoline, but positive for cattle and hogs. This indicates that the upward jumps in time to maturity that occur at contract switching points reduce the volatility of returns for corn, soybeans, and gasoline, but increase volatility in live cattle and hogs. Apparently, live cattle and hogs are relatively more volatile further away from the maturity date, while corn, soybeans and gasoline are relatively more stable.

In this chapter we are primarily interested in the long memory parameter d. The estimated long memory parameters reported in Table 3-2 are strongly statistically significant for all six futures return series, and the hypotheses that $d = 0$ (stationary GARCH) and also $d = 1$ (integrated GARCH) are consistently rejected for all commodities using standard significance levels. Table 3-2 also reports robust Wald test statistics, denoted by W, for testing the null hypothesis of GARCH versus a FIGARCH data generating process. Under the null, W will have an asymptotic χ_1^2 distribution and, from Table 3-2, the GARCH model is rejected for every commodity at standard significance levels. This formal statistical test supports the conclusion obtained both here and in Jin and Frechette (2004) that FIGARCH is superior to GARCH for modeling the conditional variances of commodity returns. Evidently, long memory is a characteristic feature of daily commodity futures returns, and FIGARCH represents a significant improvement over GARCH.

3.3. Analysis of High Frequency Commodity Returns

Considerable previous work has examined the properties of high frequency returns in equity and currency markets, but to date very little analysis has been done on high frequency commodity returns. The only study we are aware of is Cai, Cheung and Wong (2001) who studied high frequency gold futures prices. Their study analyzed 5-minute gold futures returns between 1994 and 1997, and they discovered slow hyperbolic decay associated with the autocorrelation function of the returns. However, they used an informal method for approximating the long memory parameter and did not estimate formal FIGARCH models. This section of the chapter represents a first attempt at extensive analysis of the volatility properties of high frequency commodity futures returns using FIGARCH models.

The raw futures tick data for the analysis were obtained from the Futures Industry Institute data center along with the daily data (see footnote 2), and correspond to the same six commodities studied in the previous section. The prices are for real-time transaction records, which we initially convert to 5-minute price intervals by using the last price quoted before the end of every 5-minute interval over the trading day. For 5-minute intervals that have no price recorded we linearly interpolate between surrounding intervals to fill in the missing data. As with all high frequency asset price analyses, there are potential problems with data unreliability due to the sheer amount of data being used and the fact that there is considerable noise in the series because of little trade occurring at some of the recorded prices. However, we minimize these problems by running the data through a filter to identify and adjust anomalous observations. This was done by locating return observations greater than three standard deviations and evaluating these as possible data errors. A careful check and evaluation of these observations revealed a small number of what appeared to be data errors in the high frequency gold returns. These were then eliminated and replaced with a linearly interpolated value using the two contiguous observations. No errors were detected in high frequency commodity returns other than gold. Furthermore, instead of analyzing the 5-minute interval data (which will be the most susceptible to data errors and noise) we convert the data to lower frequencies (10-minute for corn and soybeans, and 15-minute for live cattle and hogs, gasoline, and gold) to undertake the analysis. Different intervals were chosen for different commodities because they are traded on markets that have different trading day lengths. Hence, in order to make sure interval returns could be computed that would exhaust the recorded daily price change but not use consecutive intervals that stretched over two different trading days, it was convenient to use 10-minute intervals for corn and soybeans but 15-minute intervals for live cattle, live hogs, gasoline, and gold.

An interval return during day t is defined as $y_{t,n} = 100\ [ln(P_{t,n})-ln(P_{t,n-1})]$, where $P_{t,n}$ is the futures price for the n-th intra-day interval during trading day t. As with many analyses of high frequency asset price returns, it was found that the high frequency commodity returns display considerable intra-day periodicity, which is usually attributed to institutional trading features. This periodicity was removed using the FFF filtering method, which is explained in detail in the Appendix.

Figure 3-6 plots the sample autocorrelations for lags of up to 5 trading days in 5-minute intervals displayed in the horizontal axis for the absolute returns of the unadjusted (raw) and the filtered 5-minute intervals for all the commodity futures returns series. The dotted line represents sample autocorrelations for the unfiltered absolute 5-minute returns, while the solid line indicates the autocorrelations for the filtered absolute 5-minute returns. The FFF filter seems to remove much of intra-day periodicity present in the raw absolute returns. As usual, there is a small negative but significant first-order autocorrelation in returns, which may be due to the non-synchronous trading phenomenon, while higher order autocorrelations are not significant at conventional levels. The autocorrelation functions of the absolute returns also exhibit a pronounced U shape, suggesting substantial intra-day periodicity. Similar U-shaped patterns are found in the equity markets (Harris, 1986; Wood et al., 1985; Chang et al., 1995; and Andersen and Bollerslev, 1997a). Figure 3-7 shows the average absolute filtered intra-day returns within a trading day. For all the commodities, the intra-day volatility patterns display an U-shaped pattern. Unless otherwise indicated, all remaining analyses were done on the filtered series.

The MA-FIGARCH model (3.1) through (3.3) was estimated based on the filtered high frequency filtered returns. As with the daily data, the orders of the MA and GARCH polynomials in the lag operator were chosen to be as parsimonious as possible but still provide an adequate representation of the autocorrelation structure of the high frequency data. For the high frequency data, MA(1)-FIGARCH(1,d,1) models proved adequate for all commodities. Long high frequency series were constructed by splicing several nearby futures contracts together, in the same way as described for the daily data. A time to maturity effect in volatility was tested, similar to that found in the daily return series. For the high frequency return data, however, the time to maturity effect was not statistically significant and so the time to maturity effect was restricted to zero. One possible reason for this result is that there are many fewer contract switches in the high frequency series, which combines a smaller number of futures contracts than the daily futures return series. The number of trading days and the number of intra-day periods are different across the different commodities. This information is provided in Table 3-3.

Details of the estimated MA(1)-FIGARCH(1,d,1) high frequency models for the six commodities are reported in Table 3-4. All the models have small but significant MA(1) parameter estimates, which are usually attributed to the non-synchronous trading phenomenon. Similar features for high frequency exchange rate returns have been noted by Andersen and Bollerslev (1997a), Goodhart and Figliuoli (1992), Goodhart and O'Hara (1997), and Zhou (1996). The estimated long memory volatility parameter d ranges from 0.2 to 0.3 for most of the commodities considered and is generally statistically significant.

Similar to the daily return results, we found significant long memory volatility in the high frequency returns data as well. In general, the long memory estimates for intra-day return volatilities are slightly lower than those for daily returns. Furthermore, as in the daily return models, the robust Wald statistics in Table 3-3 show strong evidence in favor of the FIGARCH specifications against the GARCH specifications in the high frequency model.

Details for the FIGARCH estimation results for various daily and intradaily sample frequencies are recorded in tables A-1 through A-12 in the Appendix to this chapter. Another remarkable observation from the detailed estimation results is that the Robust Wald statistics W for testing the null hypothesis of GARCH specification seem to be proportional to the sample frequency. This finding could imply that the long memory feature becomes more pronounced as we observe price changing more frequently within a particular sample period, while the long memory estimate levels themselves remain similar across different sample frequencies. Therefore, we can conjecture that, as higher sample frequencies are considered, the FIGARCH conditional variance specifications become superior to a simple GARCH model that does not implement long memory volatility.

3.4. Local Whittle Estimation and Self-Similarity

An alternative to the parametric long memory models used so far in this chapter is the application of the semi-parametric, local Whittle estimator for estimation of long memory parameters. The advantage of this estimator is that it allows for quite general forms of short run dynamics while ARFIMA and FIGARCH models are potentially sensitive to the specifications used to represent the short-run dynamics; see Kunch (1987) and Robinson (1995). Of course, semi-parametric estimation has its own problems, as it is very data intensive and often exhibits poor performance in terms of bias and mean square error. We apply local Whittle estimation as a robustness check on the FIGARCH parametric estimates of the long memory parameter *d*.

A characteristic of long memory that is independent of parametric model specification is that the spectrum of the series will be given by $f(\omega) \sim G\omega^{-2d}$, as $\omega \to 0+$ and G is a constant. This suggests a useful objective function for estimating d would be

$$Q = \ln\left[(1/m)\sum_{j=1}^{m}\omega_j^{2d} I(\omega_j)\right] - (2d/m)\sum_{j=1}^{m}\ln(\omega_j)$$

where $I(\omega_j)$ is the periodogram of the series at frequency ω_j (see Robinson, 1995). Solving this objective function numerically gives the local Whittle estimator of *d*. Note that it is not necessary to specify the short run dynamics of the process in order to estimate *d* in this framework. As shown by Robinson (1995) and others, the main decision variable is *m*, the choice of the number of ordinates of the periodogram. For consistency, it is necessary that $\left[(1/m)+(m/T)\right]\rightarrow 0$ as $T\rightarrow\infty$. For asymptotic normality, it is required that $(1/m)+m^{1+2\beta}\left[\ln(m)\right]^2 T^{-2\beta}\rightarrow 0$ as $T\rightarrow\infty$. In the empirical results reported in this chapter, *m* is chosen as $T^{0.80}$. Note that the asymptotic variance of the local Whittle estimator is given by $(1/4m)$.

Local Whittle estimation of the long memory volatility parameter *d* was applied to both the daily and high frequency returns for all six commodities studied earlier. Furthermore, both MA-FIGARCH and local Whittle estimation of *d* were undertaken for a range of alternative frequencies (1-day, 2-day, 3-day, 4-day and 5-day using the daily data, and various return frequencies between 10 minutes and 2 hours using the high frequency data). Estimation was undertaken over multiple frequencies to check for the self-similarity feature. Self-similarity occurs when the magnitude of the long memory parameter does not change across sampling frequencies; see Beran (1994). If the long memory parameter is invariant across frequencies, then it suggests that the long memory property is an intrinsic feature of the data and does not result from regime shifts or exogenous external shocks. The self-similarity property is technically extremely difficult to test empirically. However, one can subjectively evaluate changes in long memory parameter estimates across frequencies to see whether the self-similarity feature seems to hold in general.

Results of both FIGARCH and local Whittle estimation of the long memory parameter *d* are shown for a range of daily return frequencies in Table 3-5 and for a range of intra-day return frequencies in Table 3-6. Numbers in parentheses below the estimates are the estimated standard errors. The first thing to notice is that FIGARCH and local Whittle estimates of *d* appear quite consistent with one another, with *d* estimated in the range supporting long memory in commodity return volatilities. Hence, previous conclusions about the existence of the long memory property in commodity return volatilities using FIGARCH appear robust to specification of alternative representations of short-run dynamics. The second thing to notice in Tables 3-5 and 3-6 is that the long memory parameter estimates are generally quite consistent across different return frequencies, irrespective of whether we look at daily returns or intra-day returns. This result is consistent with the notion of self-similarity and suggests that long memory and hyperbolic decay are intrinsic features of commodity return data.

3.5. Conclusions

This chapter has examined the long memory volatility properties of both daily and high frequency intra-day futures returns for six important commodities. The absolute and squared returns all possess very significant long memory features and their volatility processes are found to be well described as FIGARCH fractionally integrated volatility processes. We also find small departures from the martingale in mean property. The long memory property in absolute returns was also undertaken by semi-parametric local

Whittle estimation of the long memory parameter. The estimation of MA-FIGARCH models and the application of the local Whittle estimators to absolute returns were also computed for a range of different sample frequencies using both the daily and intra-day high frequency returns. The long memory parameter estimates are found to be quite robust both across estimators and across sample frequencies. This is consistent with a finding of self-similarity, which implies that long memory in volatility is a pervasive and consistent feature of commodity returns, and is not just being caused by shocks or regime shifts to the underlying price processes.

Our findings suggest that any future empirical application using daily or intra-day commodity futures returns (for example, optimal hedge ratio estimation, tests for futures market efficiency, tests for the announcement effect of market news, option valuation, farm risk portfolio management, etc.) will need to account for the long memory property in commodity return volatilities.

Table 3-1: Summary Statistics of Returns

	Corn	Soybean	Cattle	Hogs	Gasoline	Gold
First Day	1/02/80	1/02/80	1/02/80	1/02/80	1/02/91	1/02/80
Last Day	3/30/01	12/29/00	12/29/00	12/29/00	12/29/00	12/29/00
Sample Size	5362	5300	5306	5306	2509	5283
Mean	-0.016	-0.005	0.037	0.042	0.0406	-0.0298
High	8.606	7.806	2.867	6.307	12.107	9.745
Low	-10.472	-11.665	-2.812	-7.632	-30.987	-9.909
Std. Dev.	1.279	1.341	0.898	1.403	1.9594	1.227

Key: The above statistics refer to $100\Delta \ln(P_t)$, where P_t is the price of the asset in time period t.

Table 3-2: Estimated MA-FIGARCH Models for Daily Futures Returns

	Corn	Soybeans	Cattle	Hog	Gasoline	Gold
μ	-0.0171	-0.0229	0.0456	0.0524	0.0097	-0.0367
	(0.0152)	(0.0152)	(0.0117)	(0.0217)	(0.0337)	(0.0102)
θ	0.0618	-0.0220	*	*	0.0695	-0.0247
	(0.0151)	(0.0144)			(0.0211)	(0.0163)
d	0.3154	0.3451	0.3718	0.3687	0.3179	0.2969
	(0.0362)	(0.0493)	(0.0422)	(0.0609)	(0.0577)	(0.0261)
ω	0.2036	0.2727	0.0185	0.0621	0.7625	0.0399
	(0.0473)	(0.0607)	(0.0141)	(0.0386)	(0.2151)	(0.0288)
β_1	0.2542	0.3313	0.3603	0.3420	0.2852	0.1923
	(0.0442)	(0.0597)	(0.0466)	(0.0639)	(0.0650)	(0.0438)
β_2	———	———	0.0819	0.1206	———	———
			(0.0212)	(0.0202)		
γ	-0.1820	-0.4218	0.0701	0.1933	-1.0264	0.0595
	(0.0890)	(0.1226)	(0.0383)	(0.0994)	(0.2978)	(0.0587)
m_3	-0.003	0.016	-0.170	-0.142	-0.166	-0.097
m_4	4.218	4.917	3.100	3.079	3.916	8.750
Q(20)	20.232	21.446	29.906	17.147	25.765	22.476
$Q^2(20)$	30.887	34.976	16.875	21.426	13.407	19.697
W	76.092	48.950	77.698	36.699	30.406	55.693

Key: Robust standard errors based on QMLE are in parentheses below the corresponding parameter estimates. The diagnostic statistics Q(20) and $Q^2(20)$ are the Ljung-Box statistics based on the first 20 autocorrelations of the standardized residuals and the autocorrelations of the squared standardized residuals respectively. The statistics m_3 and m_4 are the sample skewness and kurtosis respectively of the standardized residuals. The symbol * indicates that MA(5) and MA(10) models respectively were estimated for live cattle and live hogs respectively. The parameter estimates are not reported to conserve space.

Table 3-3: Summary Statistics for Five Minute Futures Returns

	Number of trading days	Number of intraday intervals	First time period	Last time period
Corn	471	44	9:40	13:15
Soybeans	409	44	9:40	13:15
Gasoline	401	63	10:00	15:00
Live Cattle	405	45	9:20	13:00
Live Hogs	400	45	9:20	13:00
Gold	401	72	8:30	14:25

	Corn	**Soybean**	**Cattle**	**Hogs**	**Gasoline**	**Gold**
First Day	5/03/99	5/03/99	5/03/99	5/03/99	5/03/99	5/03/99
Last Day	3/30/01	12/28/00	12/28/00	12/28/00	12/28/00	12/28/00
Sample Size	20724	17996	18225	18000	25263	25842
Mean	-0.003	-0.001	0.013	0.027	0.033	-0.009
High	14.706	14.721	7.231	22.422	33.416	25.168
Low	-15.783	-14.846	-7.160	-23.530	-32.308	-28.664
Standard Dev.	1.658	1.571	0.819	1.982	2.463	1.052

Table 3-4: Estimated MA-FIGARCH model for Filtered High Frequency Futures Returns

	Corn	Soybean	Cattle	Hog	Gasoline	Gold
Sample frequency	10 min.	10 min.	15 min.	15 min.	15 min.	15 min.
μ	-0.0030	-0.0023	0.0031	0.0091	0.0145	-0.0037
	(0.0017)	(0.0021)	(0.0015)	(0.0032)	(0.0039)	(0.0011)
θ	-0.1560	-0.0659	-0.0525	-0.0490	-0.0274	-0.0750
	(0.0112)	(0.0120)	(0.0144)	(0.0158)	(0.0127)	(0.0134)
d	0.2429	0.2213	0.2097	0.3503	0.1843	0.2047
	(0.0368)	(0.0329)	(0.0367)	(0.0620)	(0.0218)	(0.0421)
ω	0.0014	0.0018	0.0024	0.0030	0.0449	0.0026
	(0.0004)	(0.0005)	(0.0019)	(0.0012)	(0.0062)	(0.0006)
β	0.8866	0.8736	0.4234	0.7242	0.0572	0.2534
	(0.0339)	(0.0309)	(0.3885)	(0.0816)	(0.0291)	(0.1666)
ϕ	0.8314	0.8279	0.3450	0.5485	______	0.3573
	(0.0462)	(0.0417)	(0.3810)	(0.0964)		(0.1726)
m_3	0.366	0.106	-0.111	-0.199	-0.134	0.048
m_4	6.882	7.335	4.728	6.138	5.151	8.508
Q(20)	28.673	19.973	17.685	23.944	34.336	25.893
$Q^2(20)$	16.592	15.556	7.917	24.340	17.650	11.566
W	43.548	45.173	32.716	31.944	71.753	23.619

Key: As for Table 3-2

Table 3-5: Long Memory Parameter Estimation at Different Daily Sample Frequencies.

	1 days	2 days	3 days	4 days	5 days
Corn					
FIGARCH	0.3154	0.2734	0.3096	0.3312	0.2510
	(0.0362)	(0.0460)	(0.0670)	(0.0833)	(0.0671)
Local Whittle	0.4072	0.3446	0.3052	0.3122	0.2359
	(0.0376)	(0.0471)	(0.0539)	(0.0592)	(0.0635)
Soybeans					
FIGARCH	0.3451	0.3403	0.4052	0.3096	0.3294
	(0.0493)	(0.0780)	(0.1385)	(0.0783)	(0.0921)
Local Whittle	0.3902	0.3688	0.3918	0.3356	0.3260
	(0.0378)	(0.0474)	(0.0541)	(0.0592)	(0.0638)
Live Cattle					
FIGARCH	0.3718	0.4399	0.4208	0.4335	0.4747
	(0.0422)	(0.0863)	(0.0821)	(0.0984)	(0.1395)
Local Whittle	0.3866	0.3383	0.3361	0.3234	0.3226
	(0.0378)	(0.0472)	(0.0539)	(0.0592)	(0.0603)
Live Hogs					
FIGARCH	0.3687	0.3041	0.3085	0.2900	0.2411
	(0.0609)	(0.0578)	(0.0659)	(0.0935)	(0.0805)
Local Whittle	0.4061	0.3416	0.3609	0.2987	0.2455
	(0.0378)	(0.0472)	(0.0539)	(0.0592)	(0.0603)
Gasoline					
FIGARCH	0.3179	0.3140	0.2851	0.2999	0.2052
	(0.0577)	(0.0707)	(0.1041)	(0.1430)	(0.0874)
Local Whittle	0.2935	0.2548	0.2967	0.2722	0.2400
	(0.0481)	(0.0603)	(0.0689)	(0.0760)	(0.0818)
Gold					
FIGARCH	0.2969	0.3435	0.2754	0.3357	0.3108
	(0.0261)	(0.0520)	(0.0400)	(0.0470)	(0.0857)
Local Whittle	0.4323	0.3766	0.3565	0.3659	0.3432
	(0.0378)	(0.0474)	(0.0541)	(0.0595)	(0.0638)

Table 3-6: Long Memory Parameter Estimation at Different Intraday Sample Frequencies.

Corn	10 min.	20 min.	55 min.	1 hr. 50 min.
FIGARCH	0.2429	0.1919	0.2196	0.0814
	(0.0368)	(0.0430)	(0.0951)	(0.1556)
Local Whittle	0.1941	0.2037	0.1622	0.1652
	(0.0255)	(0.0324)	(0.0462)	(0.0595)

Soybeans	10 min.	20 min.	55 min.	1 hr. 50 min.
FIGARCH	0.2213	0.2689	0.2431	0.3111
	(0.0329)	(0.0560)	(0.0758)	(0.1378)
Local Whittle	0.2533	0.2365	0.1706	0.1448
	(0.0268)	(0.0340)	(0.0486)	(0.0625)

Live Cattle	15 min.	25 min.	45 min.	1 hr. 15 min.
FIGARCH	0.2097	0.2580	0.2519	0.2483
	(0.0367)	(0.0492)	(0.0806)	(0.0986)
Local Whittle	0.2128	0.1833	0.1796	0.1421
	(0.0307)	(0.0366)	(0.0451)	(0.0540)

Live Hogs	15 min.	25 min.	45 min.	1 hr. 15 min.
FIGARCH	0.3503	0.3987	0.3936	0.4045
	(0.0620)	(0.0835)	(0.1127)	(0.1405)
Local Whittle	0.2993	0.3414	0.2988	0.2802
	(0.0308)	(0.0368)	(0.0453)	(0.0543)

Gasoline	15 min.	35 min.	45 min.	1 hr. 45 min.
FIGARCH	0.1843	0.2672	0.2215	0.2191
	(0.0218)	(0.0556)	(0.0590)	(0.0828)
Local Whittle	0.2243	0.1876	0.1870	0.2436
	(0.0274)	(0.0376)	(0.0401)	(0.0543)

Gold	15 min.	45 min.	1 hr. 30 min.	2 hr. min.
FIGARCH	0.2047	0.2870	0.3167	0.4403
	(0.0421)	(0.3087)	(0.1180)	(0.2742)
Local Whittle	0.3832	0.3818	0.2704	0.2486
	(0.0261)	(0.0381)	(0.0486)	(0.0540)

Figure 3-1. Autocorrelation of Daily Live Cattle Futures

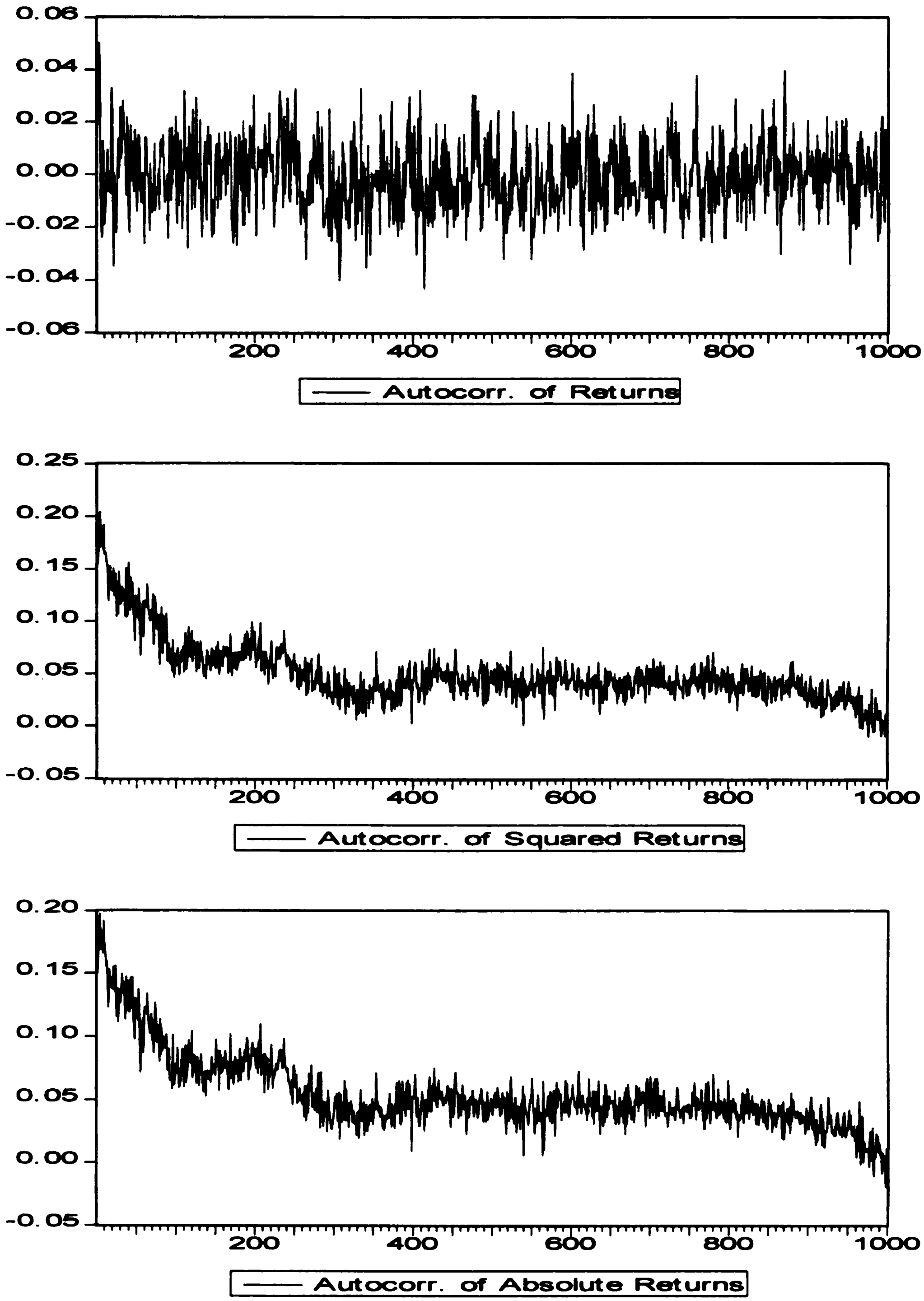


Figure 3-2. Autocorrelation of Daily Corn Futures

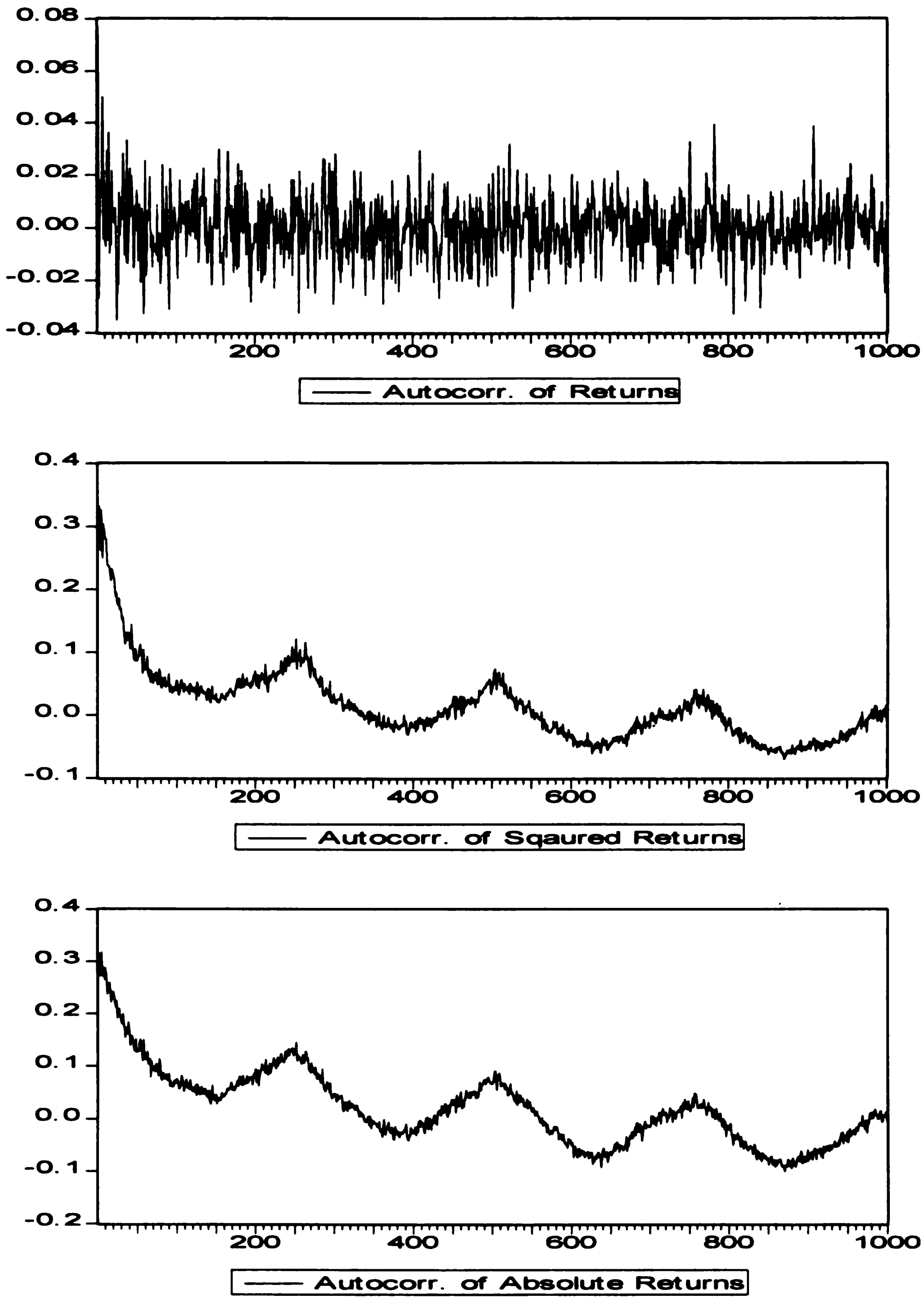


Figure 3-3. Autocorrelation of Filtered Daily Corn Futures

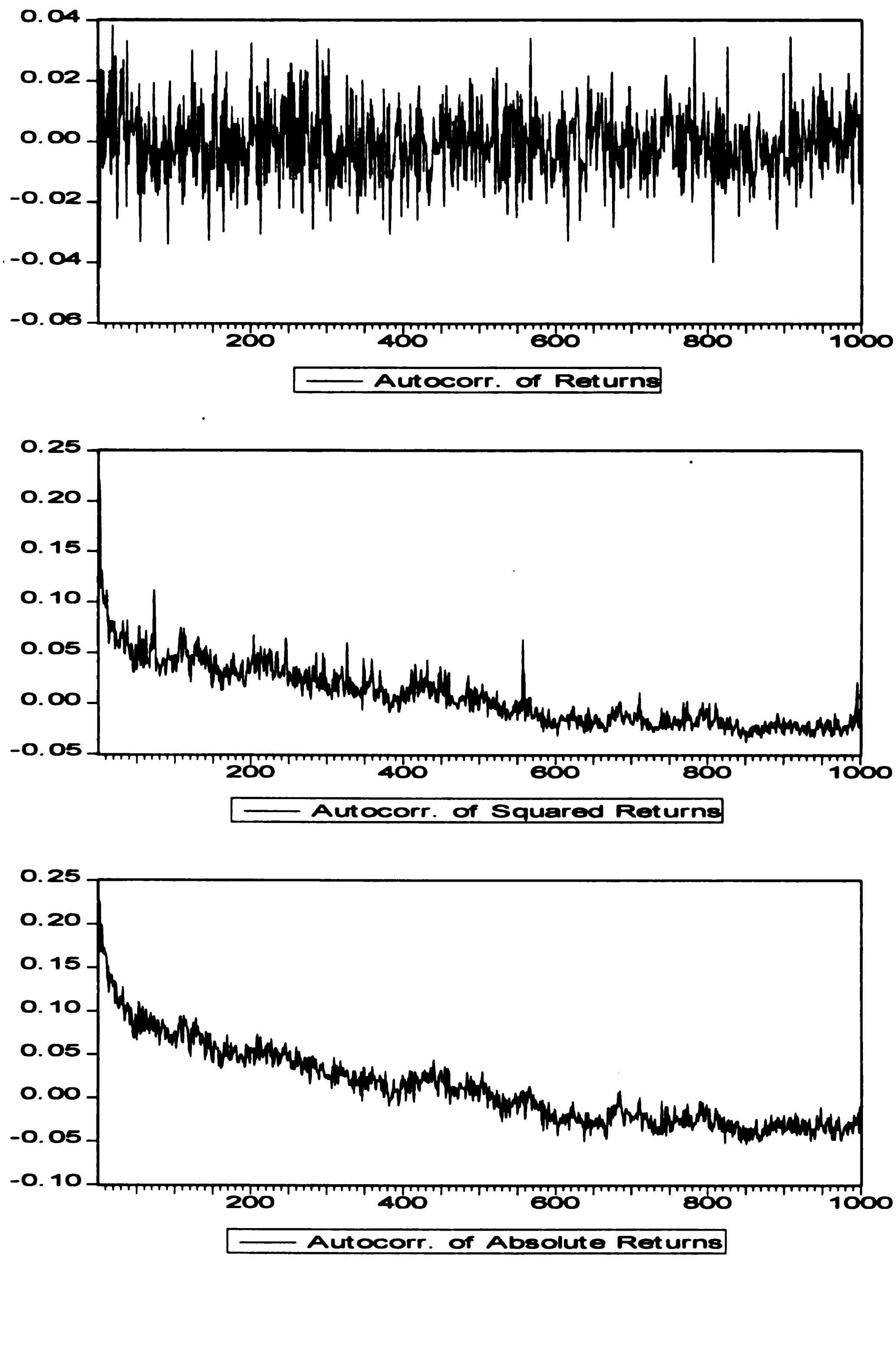


Figure 3-4. Autocorrelation of Daily Soybean Futures

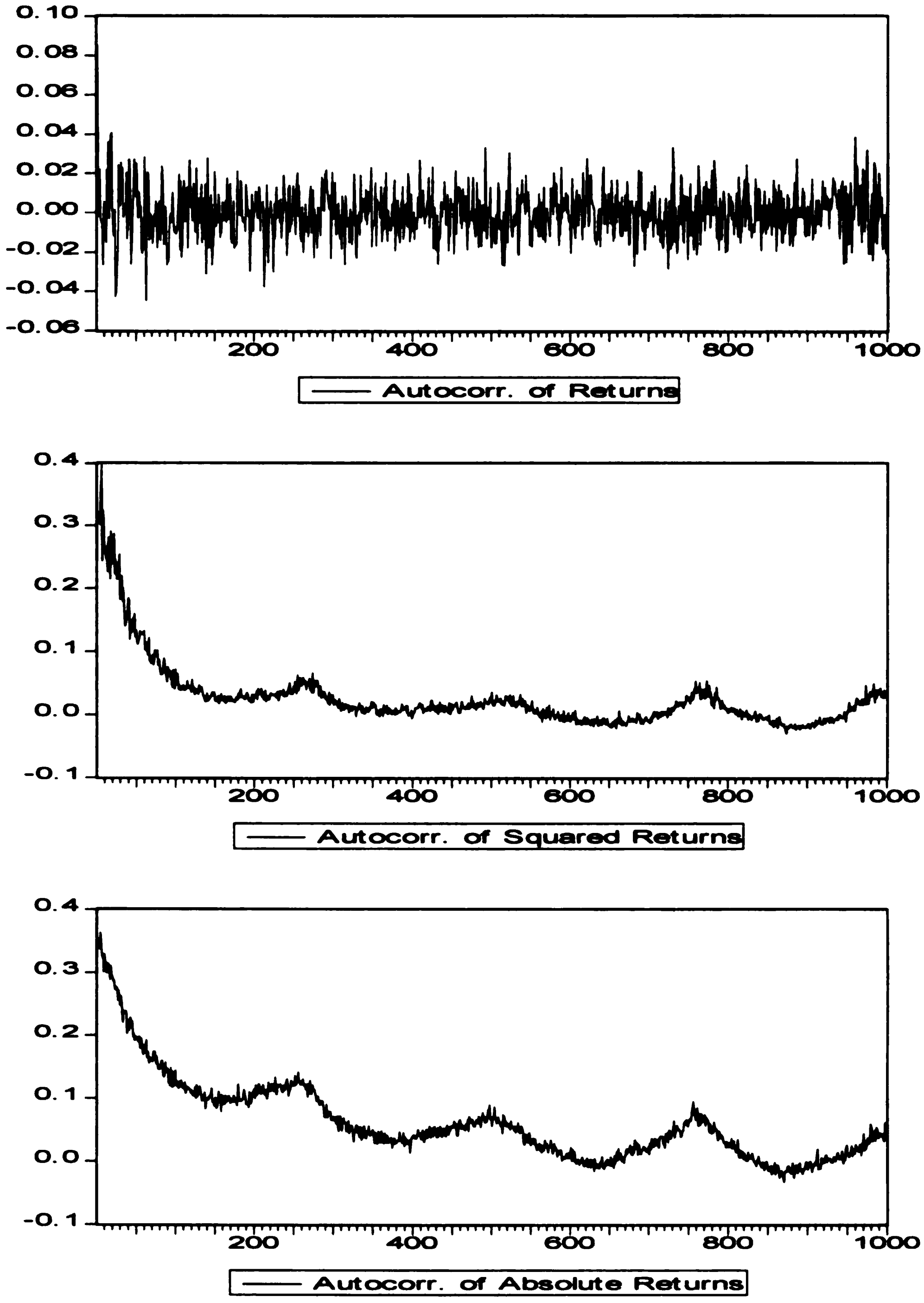


Figure 3-5. Autocorrelation of Filtered Daily Soybean Futures

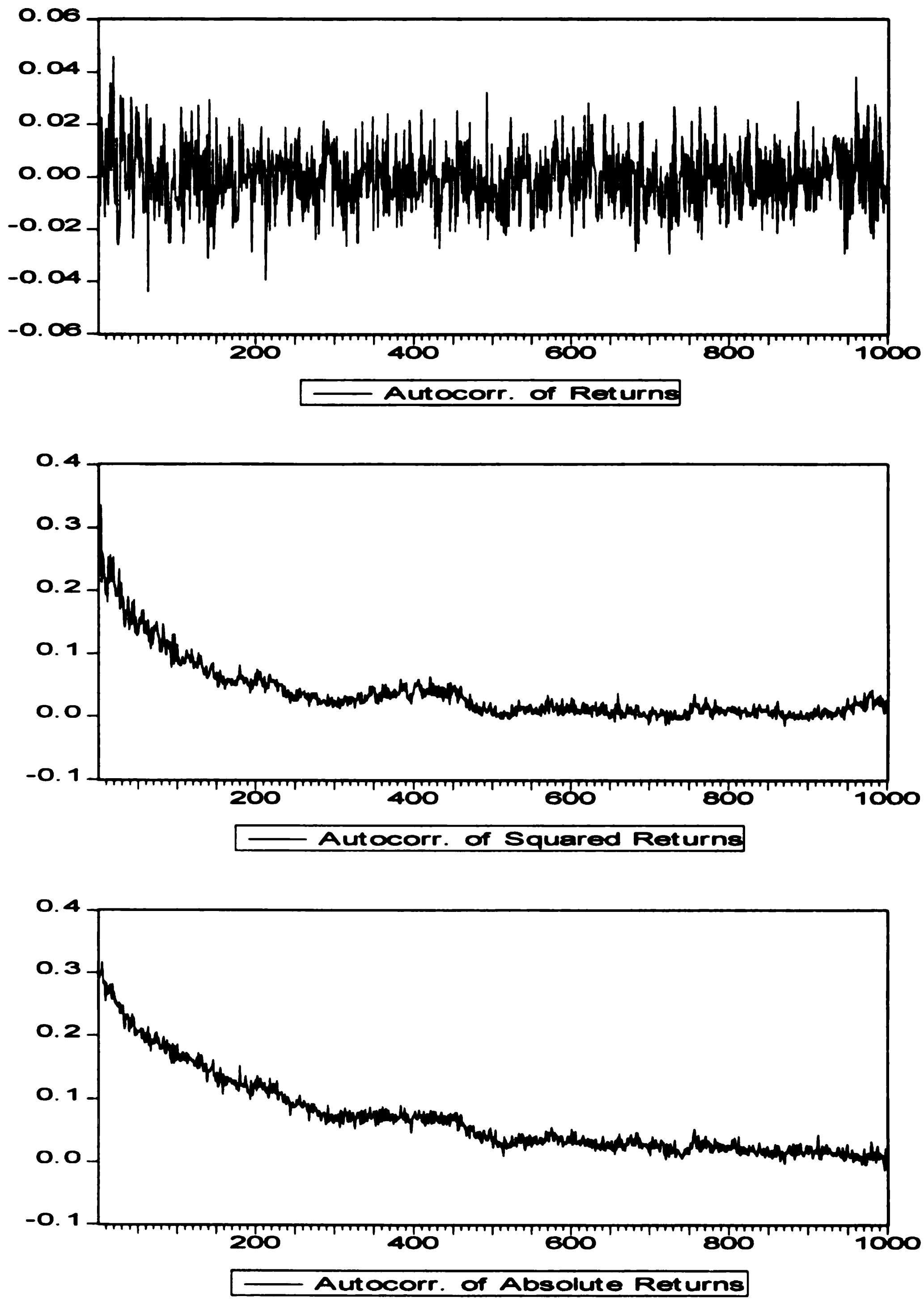
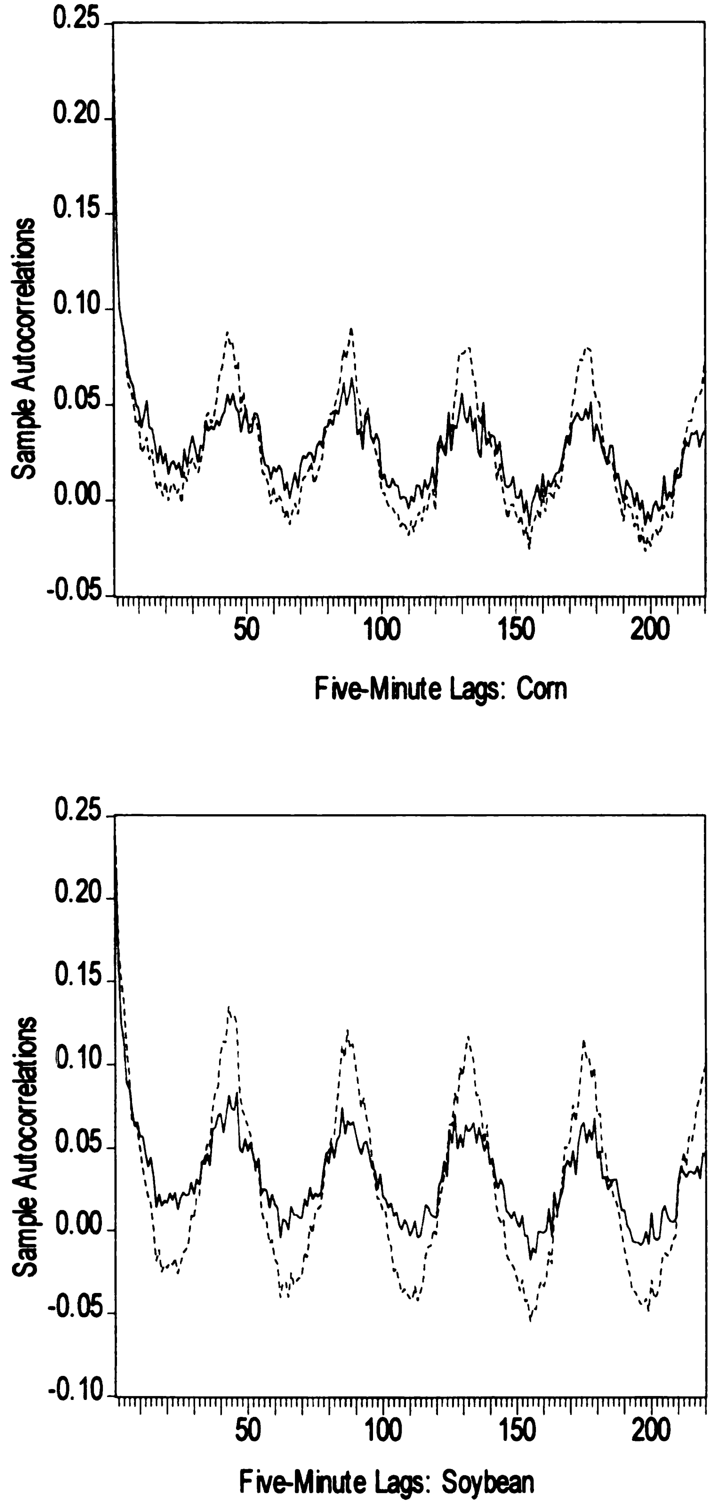
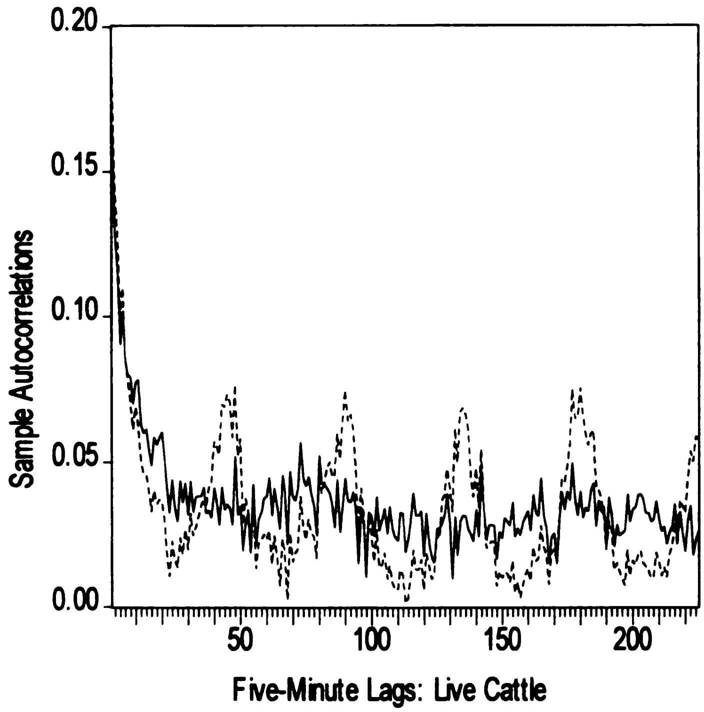
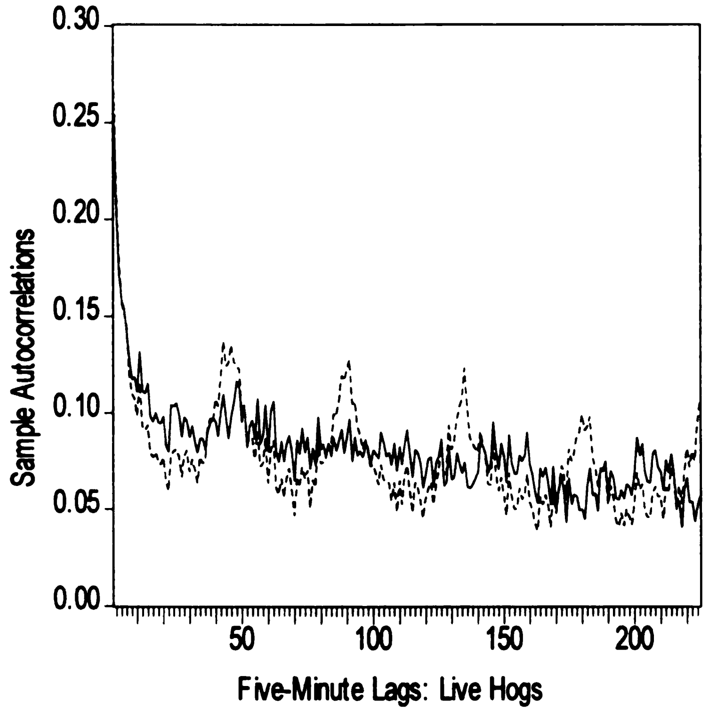


Figure 3-6 Correlograms for Absolute Raw and Filtered Five-minute Returns

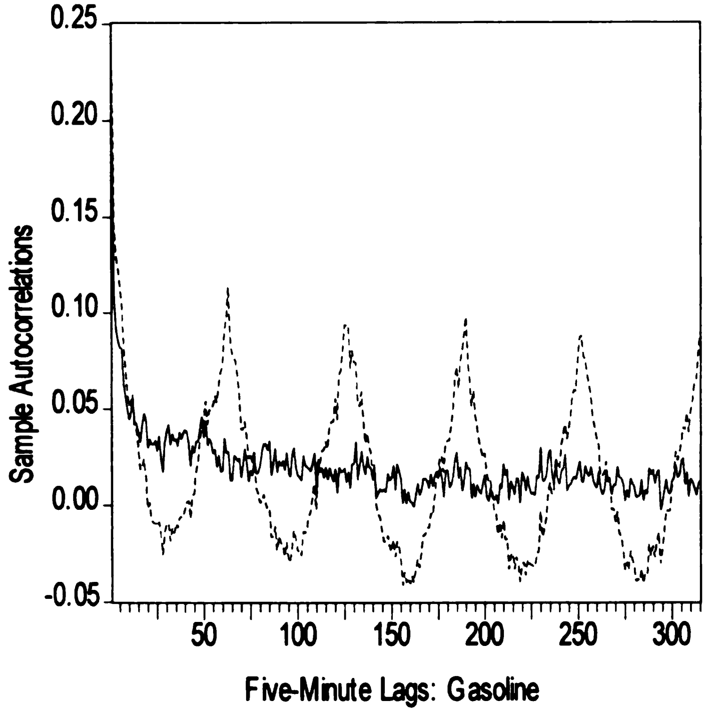


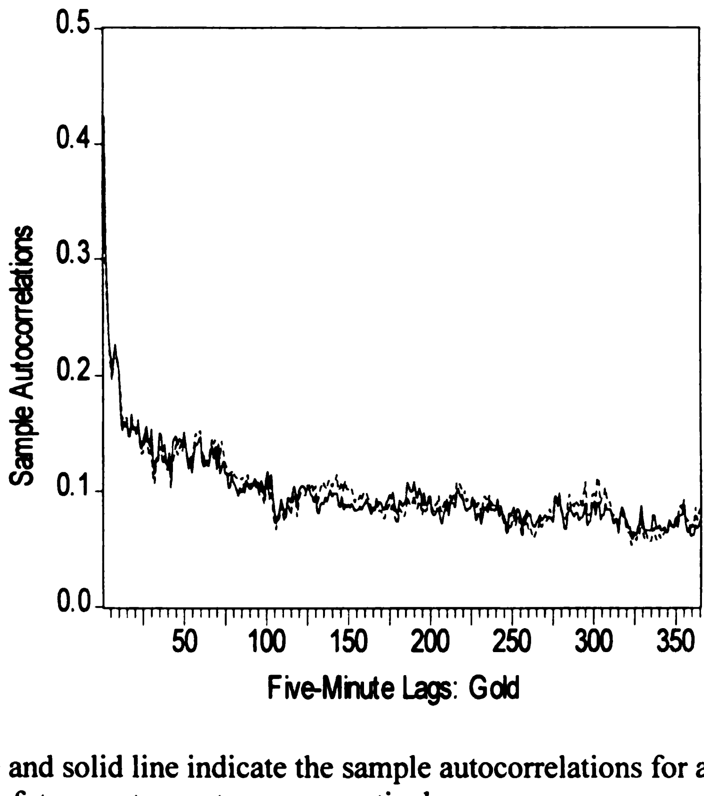
Key: Dotted line and solid line indicate the sample autocorrelations for absolute raw and filtered 5-minute futures return returns, respectively.





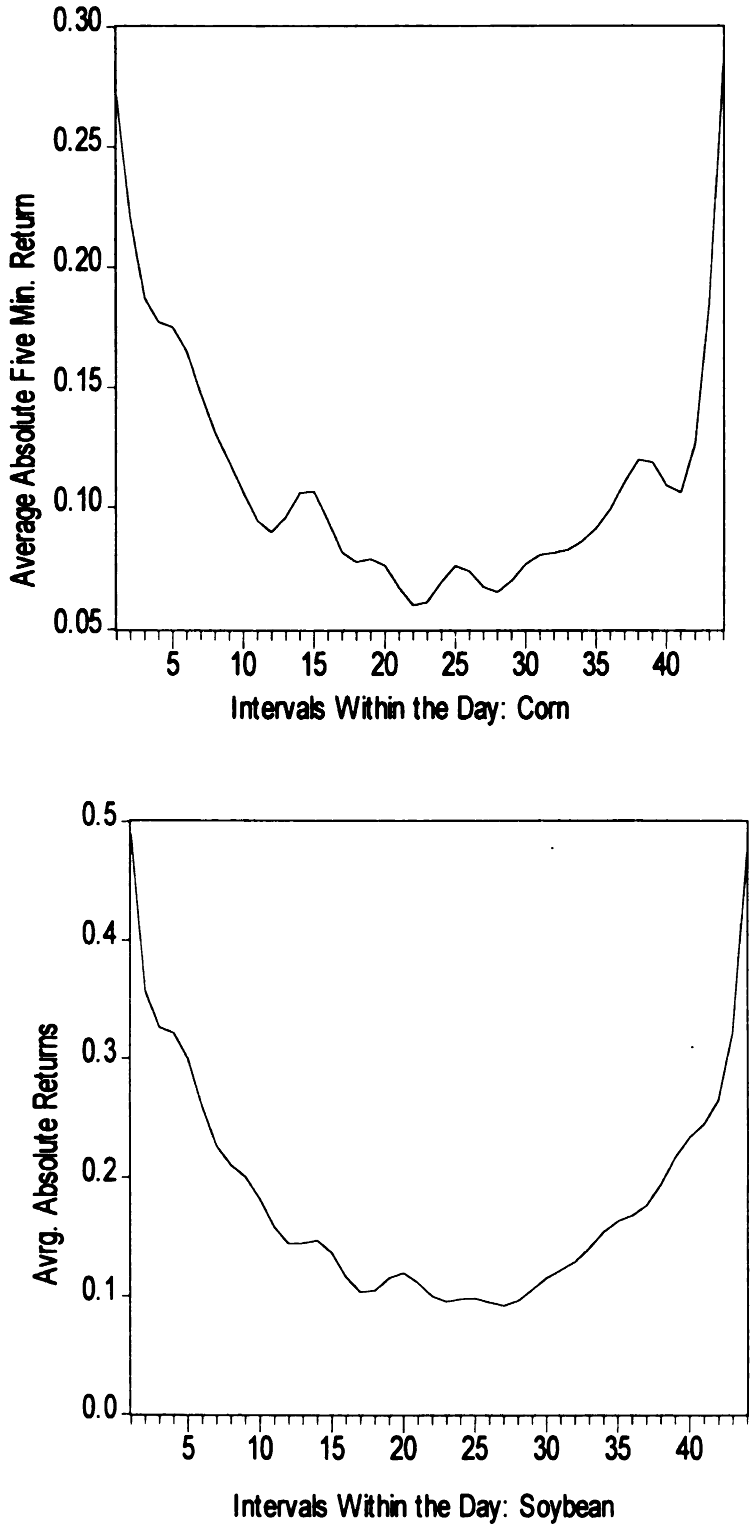
Key: Dotted line and solid line indicate the sample autocorrelations for absolute raw and filtered 5-minute futures return returns, respectively.

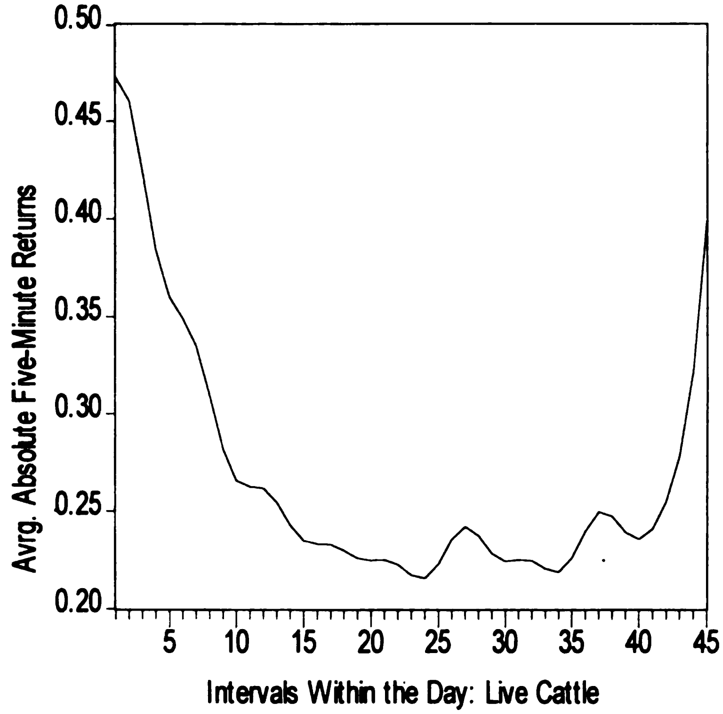




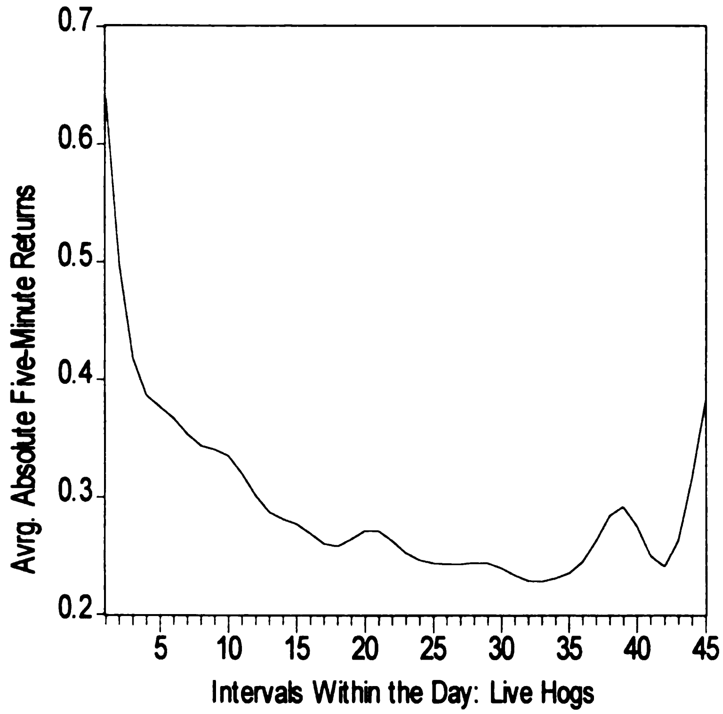
Key: Dotted line and solid line indicate the sample autocorrelations for absolute raw and filtered 5-minute futures return returns, respectively.

Figure 3-7 Fitted Intraday Volatility Pattern by the FFF filtering

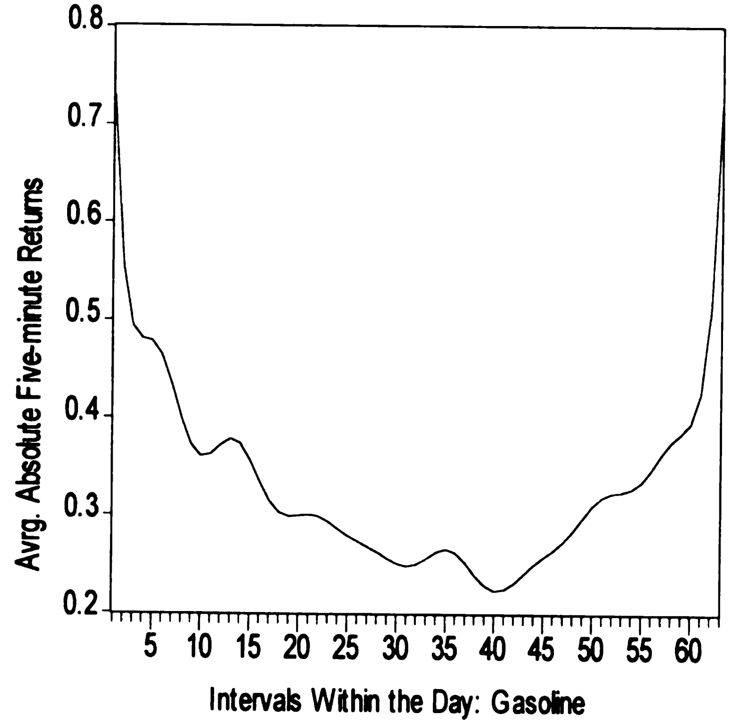




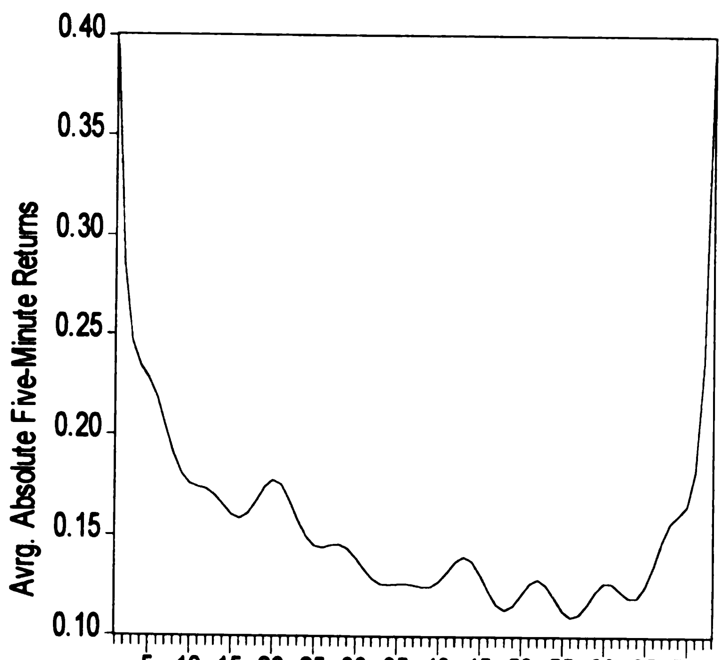
0.50
0.45
0.40
0.35
0.30
0.25
0.20
Avrg. Absolute Five-Minute Returns
5 10 15 20 25 30 35 40 45
Intervals Within the Day: Live Cattle



0.7
0.6
0.5
0.4
0.3
0.2
Avrg. Absolute Five-Minute Returns
5 10 15 20 25 30 35 40 45
Intervals Within the Day: Live Hogs



0.8
0.7
0.6
0.5
0.4
0.3
0.2
Avrg. Absolute Five-minute Returns
5 10 15 20 25 30 35 40 45 50 55 60
Intervals Within the Day: Gasoline



0.40
0.35
0.30
0.25
0.20
0.15
0.10
Avrg. Absolute Five-Minute Returns
5 10 15 20 25 30 35 40 45 50 55 60 65 70
Intervals Within the Day: Gold

Appendix

The regular opening and closing of commodity markets and the institutionalized features of lunch hours and so forth give rise to strong intra-day periodicity that is readily observable from the recurrent U-shaped patterns in the correlograms of the squared and absolute returns data. This is similar to but different from the currency markets where world-wide trading occurs. Following Andersen and Bollerslev (1998), we first remove these deterministic intra-day periodicities by applying Gallant's Flexible Fourier Form (FFF) filter; see Gallant (1981) and (1982). The estimated model becomes

$$y_{t,n} = E\left(y_{t,n}\right) + \left(\sigma_t \, s_{t,n} \, z_{t,n} N^{-1/2}\right) \qquad \textbf{(A1)}$$

where $E(y_{t,n})$ is the unconditional mean of returns, σ_t is the conditional variance of daily returns, $s_{t,n}$ is a deterministic function to represent intra-day seasonality, $z_{t,n}$ is an i.i.d(0,1) process, which is independent of the daily volatility process σ_t, and N is the number of return intervals per day. From equation (A1),

$$x_{t,n} = 2\ln|y_{t,n} - E(y_{t,n})| - \ln(\sigma_t^2) + \ln(N) = \ln(s_{t,n}^2) + \ln(z^2_{t,n}).$$

The observable variable $x_{t,n}$ is regressed on a nonlinear function of the time interval *n*, and daily volatility σ_t is pre-estimated from the MA-FIGARCH model using the daily futures return, equivalently,

$$x_{t,n} = f(\theta;t,n) + u_{t,n},$$

where $u_{t,n} = \ln(z^2{}_{t,n}) - E\,[\ln(z^2{}_{t,n})]$ is an i.i.d.(0,1) process and the functional form for f is

$$f(\theta;t,n) = \sum_{j=0}^{J} \sigma_t^j \{\mu_{0j} + \mu_{1j}\frac{n}{N_1} + \mu_{2j}\frac{n^2}{N_2}$$

$$+\sum\nolimits_{p=1,k}\left[\delta_{c,p}.\cos(p2\pi n/N) + \delta_{s,p}.\sin(p2\pi n/N)\right] \qquad \textbf{(A2)}$$

where $N_1 = (1/N)\sum_{i=1}^{N} i = (N+1)/2$, and $N_2 = (1/N)\sum_{i=1}^{N} i^2 = (N+1)(2N+1)/6$. On taking the variable $x_{t,n}$ as the dependent variable, the parameters in equation (A2) were estimated by OLS. The intra-day periodicity for interval n, on day t is then estimated as

$$\hat{s}_{t,n} = T.\left[\exp(f_{t,n}/2)\right]/\left[\sum\nolimits_{t=1,(T/N)}\sum\nolimits_{n=1,N}\exp(f_{t,n}/2)\right]. \qquad \textbf{(A3)}$$

The 10- or 15-minute high frequency returns are then filtered by the estimated intra-day periodicity series $\hat{s}_{t,n}$ to generate the filtered returns, which are defined as

$$\tilde{y}_{t,n} = y_{t,n}/\hat{s}_{t,n}. \qquad \textbf{(A4)}$$

The same filtering approach is also used to remove yearly seasonality existing in daily absolute returns. We use the sum of squared daily returns for each year as a substitute for the conditional volatility factor of the corresponding year since the number of sample years is less than 30 for all the commodities and is too short to model conditional variances properly. Alternatively, since we have a sufficient number of daily return observations within each year, the sum of squared daily returns yield desirable *ex-post* volatility measures for the associated year. The volatility measure is called "realized volatility" in the literature, including Andersen, Bollerslev, Diebold, and Labys (2001, 2003). They provided theoretical support for the use of the realized volatility measure and empirically showed the forecasting and modeling performance of the volatility measures in comparison to parametrically estimated conditional variances. The realized volatility series for the commodity futures market is analyzed in chapter 4. The details for the Fourier flexible functional regressions for the filtering are not reported here, but are available upon request.

Table A-1: Estimated MA-FIGARCH Models for Temporally Aggregated Daily Futures Returns for Corn

(The sample period: 1/02/80 – 3/30/01)

	1 days	**2 days**	**3 days**	**4 days**	**5 days**
T	5362	2681	1787	1340	1072
μ	-0.0171	-0.0271	-0.0305	-0.0226	-0.0684
	(0.0152)	(0.0321)	(0.0481)	(0.0674)	(0.0853)
θ	0.0618	0.0206	0.0122	0.0225	0.0172
	(0.0151)	(0.0212)	(0.0249)	(0.0291)	(0.0309)
d	0.3154	0.2734	0.3096	0.3312	0.2510
	(0.0362)	(0.0460)	(0.0670)	(0.0833)	(0.0671)
ω	0.2036	0.5326	0.7719	0.6386	0.9331
	(0.0473)	(0.1660)	(0.2935)	(0.3850)	(0.7374)
β	0.2542	0.1788	0.2451	0.2848	0.1426
	(0.0442)	(0.0583)	(0.0855)	(0.1083)	(0.0969)
γ	-0.1820	-0.3458	-0.7737	-0.2777	0.5884
	(0.0890)	(0.2751)	(0.4779)	(0.7623)	(1.2967)
m_3	-0.003	-0.020	0.060	0.116	0.139
m_4	4.218	3.955	3.932	3.836	4.002
Q(20)	20.232	23.658	19.151	17.220	13.021
$Q^2(20)$	30.887	21.600	21.215	15.462	15.388
W	76.092	35.281	21.362	15.793	13.399

Key: As for Table 3-2

Table A-2: Estimated MA-FIGARCH Models for Temporally Aggregated Daily Futures Returns for Soybean

(The sample period: 1/02/80 – 12/29/00)

	1 days	**2 days**	**3 days**	**4 days**	**5 days**
T	5300	2650	1766	1325	1060
μ	-0.0229	-0.0360	-0.0614	-0.0623	-0.0218
	(0.0152)	(0.0302)	(0.0464)	(0.0598)	(0.0750)
θ	-0.0220	-0.0234	-0.0090	-0.0250	-0.0457
	(0.0144)	(0.0206)	(0.0252)	(0.0297)	(0.0304)
d	0.3451	0.3403	0.4052	0.3096	0.3294
	(0.0493)	(0.0780)	(0.1385)	(0.0783)	(0.0921)
ω	0.2727	0.5874	0.9308	1.5733	1.2153
	(0.0607)	(0.1934)	(0.3575)	(0.5494)	(0.6612)
β	0.3313	0.2894	0.3571	0.1907	0.1695
	(0.0597)	(0.0961)	(0.1625)	(0.0981)	(0.1164)
γ	-0.4218	-0.8660	-1.9046	-2.3028	-1.0082
	(0.1226)	(0.3349)	(0.4880)	(0.8459)	(1.1923)
m_3	0.016	0.033	0.104	0.238	0.215
m_4	4.917	4.347	3.700	3.426	3.862
Q(20)	21.446	18.607	14.427	17.265	15.823
$Q^2(20)$	34.976	27.228	16.050	18.887	17.339
W	48.950	19.027	8.561	15.644	12.791

Key: As for Table 3-2

Table A-3: Estimated MA-FIGARCH Models for Temporally Aggregated Daily Futures Returns for Live Cattle (The sample period: 1/02/80 – 12/29/00)

	1 days	2 days	3 days	4 days	5 days
T	5306	2653	1768	1326	1061
μ	0.0456	0.0948	0.1482	0.1882	0.2244
	(0.0117)	(0.0213)	(0.0308)	(0.0446)	(0.0525)
θ			*		
d	0.3718	0.4399	0.4208	0.4335	0.4747
	(0.0422)	(0.0863)	(0.0821)	(0.0984)	(0.1395)
ω	0.0185	0.0485	0.1161	0.3017	0.3993
	(0.0141)	(0.0360)	(0.0801)	(0.1507)	(0.1989)
β_1	0.3603	0.3867	0.3401	0.4168	0.4051
	(0.0466)	(0.0901)	(0.0878)	(0.1037)	(0.1380)
β_2	0.0819	0.1174	0.0959	0.0617	0.0164
	(0.0212)	(0.0270)	(0.0310)	(0.0324)	(0.0513)
γ	0.0701	0.0304	-0.0674	-0.6927	-0.7675
	(0.0383)	(0.1186)	(0.2362)	(0.3897)	(0.5539)
m_3	-0.170	-0.186	-0.195	-0.225	-0.310
m_4	3.100	3.275	3.175	3.274	3.290
Q(20)	29.906	13.801	15.224	15.729	9.886
$Q^2(20)$	16.875	14.327	12.698	15.723	17.826
W	77.698	25.975	26.259	19.404	11.583

Key: As for table 3-2. (*) indicates that we omitted MA(5) coefficient estimates here since they are not important to our argument in the current chapter.

Table A-4: Estimated MA-FIGARCH Models for Temporally Aggregated Daily Futures Returns for Live Hogs (The sample period: 1/02/80 – 12/29/00)

	1 days	2 days	3 days	4 days	5 days
T	5306	2653	1768	1326	1061
μ	0.0524	0.0987	0.1526	0.2184	0.2408
	(0.0217)	(0.0467)	(0.0625)	(0.0867)	(0.1164)
θ			*		
d	0.3687	0.3041	0.3085	0.2900	0.2411
	(0.0609)	(0.0578)	(0.0659)	(0.0935)	(0.0805)
ω	0.0621	0.3920	0.5817	1.5508	2.9756
	(0.0386)	(0.1905)	(0.3116)	(0.5974)	(1.0538)
β_1	0.3420	0.2895	0.2526	0.2225	0.1020
	(0.0639)	(0.0634)	(0.0731)	(0.0964)	(0.0797)
β_2	0.1206	0.0335	0.0168	-0.0349	-0.0680
	(0.0202)	(0.0291)	(0.0500)	(0.0471)	(0.0645)
γ	0.1933	0.0117	0.0930	-1.3551	-2.5735
	(0.0994)	(0.4768)	(0.5817)	(0.9273)	(1.2072)
m_3	-0.142	-0.278	-0.248	-0.349	-0.226
m_4	3.079	3.636	3.622	3.734	3.648
Q(20)	17.747	16.775	11.670	8.097	11.780
$Q^2(20)$	21.426	19.681	16.352	21.553	6.912
W	36.699	27.644	21.951	9.615	8.974

Key: As for table 3-2. (*) indicates that we omitted MA(10) coefficient estimates here since they are not important to our argument in the current chapter.

Table A-5: Estimated MA-FIGARCH Models for Temporally Aggregated Daily Futures Returns for Gasoline (The sample period: 1/02/91 – 12/29/00)

	1 days	2 days	3 days	4 days	5 days
T	2508	1254	836	627	501
μ	0.0097	0.0148	0.0252	0.0806	0.1140
	(0.0337)	(0.0692)	(0.1063)	(0.1366)	(0.1820)
θ	0.0695	0.0303	-0.0176	0.0017	0.0050
	(0.0211)	(0.0286)	(0.0386)	(0.0373)	(0.0472)
d	0.3179	0.3140	0.2851	0.2999	0.2052
	(0.0577)	(0.0707)	(0.1041)	(0.1430)	(0.0874)
ω	0.7625	1.6346	3.4137	3.8882	8.2542
	(0.2151)	(0.5864)	(1.6571)	(2.6813)	(3.4729)
β	0.2852	0.2949	0.2026	0.2829	0.0715
	(0.0650)	(0.0835)	(0.1189)	(0.1727)	(0.1221)
γ	-1.0264	-2.2562	-4.1021	-5.1878	-7.9280
	(0.2978)	(0.8571)	(2.0028)	(3.2309)	(4.8035)
m_3	-0.166	-0.132	-0.270	0.037	0.086
m_4	3.916	3.444	3.716	3.467	3.375
Q(20)	25.765	20.823	17.493	22.675	24.154
$Q^2(20)$	13.407	20.628	23.396	12.190	14.627
W	30.406	19.745	7.495	4.400	5.509

Key: As for Table 3-2

Table A-6: Estimated MA-FIGARCH Models for Temporally Aggregated Daily Futures Returns for Gold
(The sample period: 1/02/80 – 12/29/00)

	1 days	2 days	3 days	4 days	5 days
T	5283	2641	1761	1320	1056
μ	-0.0367	-0.0702	-0.1050	-0.1295	-0.1351
	(0.0102)	(0.0200)	(0.0301)	(0.0393)	(0.0550)
θ	-0.0247	-0.0166	-0.0342	-0.0216	-0.0557
	(0.0163)	(0.0262)	(0.0384)	(0.0390)	(0.0380)
d	0.2969	0.3435	0.2754	0.3357	0.3108
	(0.0261)	(0.0520)	(0.0400)	(0.0470)	(0.0857)
ω	0.0399	-0.0368	-0.0526	-0.3467	-0.6776
	(0.0288)	(0.0465)	(0.1867)	(0.1012)	(0.3057)
β	0.1923	0.2221	0.0805	0.1890	0.2219
	(0.0438)	(0.0465)	(0.1426)	(0.0837)	(0.1269)
γ	0.0595	0.3543	0.5537	1.1286	2.9869
	(0.0587)	(0.1567)	(0.3100)	(0.3815)	(1.6484)
m_3	-0.097	0.118	0.031	0.248	0.726
m_4	8.750	7.674	6.833	6.734	12.135
Q(20)	22.476	23.539	26.282	16.538	23.584
$Q^2(20)$	19.697	26.551	14.605	10.474	8.722
W	55.693	61.185	22.961	42.882	5.570

Key: As for Table 3-2

Table A-7: Estimated MA(1)-FIGARCH(1,d,1) model for Temporally Aggregated Filtered High Frequency Futures Returns for Corn (The sample period: 5/03/99 – 3/30/01)

	10 minute	20 minute	55 minute	110 minute
T	10362	5181	1884	942
μ	-0.0030	-0.0072	-0.0126	-0.0220
	(0.0017)	(0.0038)	(0.0110)	(0.0222)
θ	-0.1560	-0.0512	-0.0082	-0.0142
	(0.0112)	(0.0158)	(0.0276)	(0.0370)
d	0.2429	0.1919	0.2196	0.0814
	(0.0368)	(0.0430)	(0.0951)	(0.1556)
ω	0.0014	0.0147	0.0135	0.0934
	(0.0004)	(0.0107)	(0.0134)	(0.1416)
β	0.8866	0.4406	0.7952	0.6742
	(0.0339)	(0.3437)	(0.1824)	(0.3633)
ϕ	0.8314	0.3647	0.7172	0.6372
	(0.0462)	(0.3341)	(0.2076)	(0.3489)
m_3	0.366	0.068	0.026	-0.265
m_4	6.882	9.105	6.102	6.463
Q(20)	28.673	16.375	12.699	17.424
$Q^2(20)$	16.592	13.106	6.181	19.453
W	43.548	19.845	5.340	0.391

Key: As for Table 3-2

Table A-8: Estimated MA(1)-FIGARCH(1,d,1) Model for Temporally Aggregated Filtered High Frequency futures returns for Soybean (The sample period: 5/03/99 – 12/28/00)

	10 minute	20 minute	55 minute	110 minute
T	8998	4499	1636	818
μ	-0.0023	-0.006	-0.0136	-0.0231
	(0.0021)	(0.0043)	(0.0122)	(0.0247)
θ	-0.0659	0.0027	-0.0042	-0.0314
	(0.0120)	(0.0158)	(0.0257)	(0.0385)
d	0.2213	0.2689	0.2431	0.3111
	(0.0329)	(0.0560)	(0.0758)	(0.1378)
ω	0.0018	0.0036	0.0184	0.0274
	(0.0005)	(0.0014)	(0.0123)	(0.0232)
β	0.8736	0.8377	0.6989	0.7143
	(0.0309)	(0.0414)	(0.1262)	(0.0783)
ϕ	0.8279	0.7182	0.5075	0.4165
	(0.0417)	(0.0537)	(0.1210)	(0.1446)
m_3	0.106	0.049	0.033	0.107
m_4	7.335	6.480	4.802	4.409
Q(20)	19.973	14.269	18.634	17.421
$Q^2(20)$	15.556	20.063	26.673	25.046
W	45.173	23.065	10.302	5.399

Key: As for Table 3-2

Table A-9: Estimated MA(1)-FIGARCH(p,δ,q) Model for Temporally Aggregated Filtered.5-minute returns for Live Cattle futures (The sample period: 5/03/99 – 12/28/00)

	15 minute	25 minute	45 minute	75 minute
T	6075	3645	2025	1215
μ	0.0031 (0.0015)	0.0056 (0.0026)	0.0080 (0.0049)	0.0151 (0.0086)
θ	-0.0525 (0.0144)	-0.0253 (0.0180)	-0.0431 (0.0230)	0.0696 (0.0351)
d	0.2097 (0.0367)	0.2580 (0.0492)	0.2519 (0.0806)	0.2483 (0.0986)
ω	0.0024 (0.0019)	0.0020 (0.0012)	0.0029 (0.0016)	0.0046 (0.0032)
β	0.4234 (0.3885)	0.6557 (0.1583)	0.7716 (0.0882)	0.7722 (0.0887)
ϕ	0.3450 (0.3810)	0.5147 (0.1613)	0.6269 (0.1211)	0.6064 (0.1105)
m_3	-0.111	-0.029	-0.042	0.133
m_4	4.728	4.862	4.894	3.927
Q(20)	17.685	18.159	25.909	23.868
$Q^2(20)$	7.917	13.639	13.347	10.007
W	32.716	27.537	9.775	6.108

Key: As for Table 3-2

Table A-10: Estimated MA(1)-FIGARCH(p,δ,q) Model for Temporally Aggregated Filtered.5-minute returns for Live Hog futures (The sample period: 5/03/99 – 12/28/00)

	15 minute	25 minute	45 minute	75 minute
T	6000	3600	2000	1200
μ	0.0091 (0.0032)	0.0155 (0.0058)	0.0270 (0.0105)	0.0402 (0.0171)
θ	-0.0490 (0.0158)	-0.0124 (0.0214)	-0.0389 (0.0252)	0.0063 (0.0349)
d	0.3503 (0.0620)	0.3987 (0.0835)	0.3936 (0.1127)	0.4045 (0.1405)
ω	0.0030 (0.0012)	0.0033 (0.0014)	0.0088 (0.0050)	0.0083 (0.0076)
β	0.7242 (0.0816)	0.7698 (0.0474)	0.6719 (0.1079)	0.7349 (0.0956)
ϕ	0.5485 (0.0964)	0.5081 (0.0658)	0.4046 (0.0921)	0.4788 (0.1297)
m_3	-0.199	-0.232	-0.283	-0.100
m_4	6.138	6.064	5.465	5.058
Q(20)	23.944	20.766	17.932	15.106
$Q^2(20)$	24.340	9.111	23.959	25.850
W	31.944	22.779	12.205	8.287

Key: As for Table 3-2

Table A-11: Estimated MA(1)-FIGARCH(p,δ,q) Model for Temporally Aggregated Filtered.5-minute returns for Gasoline futures (The sample period: 5/03/99 – 12/28/00)

	15 minute	35 minute	45 minute	105 minute
T	8421	3609	2807	1203
μ	0.0144	0.0295	0.0337	0.0683
	(0.0039)	(0.0092)	(0.0121)	(0.0294)
θ	-0.0270	-0.0127	-0.0320	0.0209
	(0.0127)	(0.0184)	(0.0191)	(0.0258)
d	0.1725	0.2672	0.2215	0.2191
	(0.0180)	(0.0556)	(0.0590)	(0.0828)
ω	-0.0745	0.0220	0.0522	0.1721
	(0.0121)	(0.0077)	(0.0242)	(0.1181)
β	-0.6661	0.6872	0.5713	0.4257
	(0.1391)	(0.0675)	(0.1114)	(0.1646)
ϕ	-0.7085	0.5473	0.3915	0.1805
	(0.1249)	(0.0832)	(0.1067)	(0.1596)
m_3	-0.129	-0.214	-0.152	-0.333
m_4	5.078	4.804	4.498	4.213
Q(20)	34.733	22.744	30.990	20.415
$Q^2(20)$	15.111	10.238	12.435	25.223
W	71.753	36.224	17.067	8.795

Key: As for Table 3-2

Table A-12: Estimated MA(1)-FIGARCH(p,δ,q) Model for Temporally Aggregated Filtered.5-minute returns for Gold futures (The sample period: 5/03/99 – 12/28/00)

	15 minute	45 minute	1 hr. 30 min.	2 hr.
T	9624	3208	1604	1203
μ	-0.0037	-0.0134	-0.0195	-0.0284
	(0.0011)	(0.0037)	(0.0107)	(0.0114)
θ	-0.0750	-0.0368	-0.0130	-0.0130
	(0.0134)	(0.0365)	(0.2437)	(0.0481)
d	0.2047	0.2870	0.3167	0.4403
	(0.0421)	(0.3087)	(0.1291)	(0.2742)
ω	0.0026	0.0032	0.0024	0.0223
	(0.0006)	(0.0471)	(0.0029)	(0.0152)
β	0.2534	0.6250	0.7714	0.1646
	(0.1666)	(5.3126)	(0.2648)	(0.1522)
ϕ	0.3573	0.6230	0.6747	______
	(0.1726)	(5.0333)	(0.3695)	
m_3	0.048	0.552	1.354	0.792
m_4	8.508	13.644	18.805	9.259
Q(20)	25.893	27.094	29.200	21.872
$Q^2(20)$	11.566	33.598	12.597	31.362
W	23.619	0.8643	6.0146	2.5788

Key: As for Table 3-2

CHAPTER 4

REALIZED VOLATILITY IN COMMODITY FUTURES MARKETS

4.1. Introduction

This chapter considers the new concept of realized volatility (RV), which is constructed from high frequency returns. We initially describe the new measurement and its properties and then apply the idea to commodity futures markets for the six important commodities considered in chapters 2 and 3. This study appears to be the first analysis using these concepts for commodity futures markets.

One interesting finding in this study is that the pure volatility measure known as realized volatility has almost ideal long memory features, which is consistent with previous work of Anderson, Bollerslev, Diebold and Labys (2002), who examined currency markets. We find that the commodity realized volatility is very well described as a fractionally integrated process and, furthermore, appears to follow a Gaussian distribution. At this level, our results are quite similar to the preceding literature, which has applied the concept of realized volatility to currency markets. However, unlike previous studies, we suggest particular factors that may possibly generate and interact with realized volatility series. In the context of commodity markets, these factors include the time to maturity of the futures contract and also the arrival of important economic news. Also, and particularly importantly, we consider a new concept: information flow, which depends on the total number of transactions at each high frequency interval of the market being active. This information flow variable turns out to be simple to compute and highly correlated with the measurement of realized volatility. In addition, this chapter examines the dependency structures between the realized volatilities for the different commodity futures data. This gives a clear indication of the mutual dependencies between the factors driving agricultural-type commodities such as corn and soybeans, while there is unsurprisingly little relationship between less related commodities. There is also some evidence of fractional cointegration between the realized volatility of corn and that of soybeans.

In this chapter, section 2 provides formal and theoretical background information for the concept of realized volatility. In section 3, we introduce and briefly discuss all of the possible issues relevant to realized volatility before the empirical investigation below. Section 4 investigates the stochastic properties of realized volatility to model and forecast the volatility measurement. Various important economic factors relevant to commodity futures markets are considered in section 5. Section 6 concludes the chapter.

4.2. Statistical Foundations of Realized Volatility:

Before defining the concept of realized volatility to be used in this chapter, it is important to recognize that, historically, there has been an awareness of the desirability of measuring the volatility associated with a continuous time diffusion process. In particular, Merton (1980) and Nelson (1992) argue that, under the theoretical assumption of a continuous diffusion process, the inherent volatility can be best measured by integrating high frequency returns data. Indeed, the finance literature has long focused on issues like instantaneous variance in the context of option pricing. The relevant previous studies include Hull and White (19887), Melino (1994), Scott (1987), and Wiggins (1987). The basic idea of realized volatility is that it can approximate the theoretical quadratic variation when the sample frequency within the time interval considered is sufficiently high, and in turn, it can provide a consistent estimator for the true latent volatility factor. These ideas are descendants of the approach of Porterba and Summers (1986), French, Schwert, and Stambaugh (1987), and Schwert (1989), who used daily returns to construct a measure of monthly volatility. Hence, the high frequency returns data can be used to construct a measure of integrated volatility by summing squared intra-day returns.

However, the measurement of realized volatility is difficult due to the fact that high frequency returns have a number of contaminating or complicating factors. In particular, as seen in chapter 3 of this dissertation, the high frequency commodity returns data are intimately involved with market microstructure factors, including a bid-ask spread and pronounced intra-day periodicity. Hence, the high frequency returns are first filtered by Gallant's (1981) Flexible Fourier Form (FFF) method before subsequent analysis, as was done in chapter 3. There is also the issue of spreads and jumps occurring at certain times. Before discussing the issues and practical problems with the implementation of the concept of realized volatility in commodity markets, we will first define the mathematical foundations of this concept.

The quadratic variation theory provides the theoretical foundation of realized volatility as a model-free unbiased estimator of conditional variance. Quadratic variation is a measure of the sample path oscillation for a special class of stochastic processes, known as semi-martingale processes, which have finite variations along their paths. For semi-martingale process $X(t)$ along the sample path $t \in [0, T]$ with a positive integer T, we define the quadratic variation process as follows:

$$[\![X(t), X(t)]\!] \equiv X(t)^2 - 2\int_{0,t} X(s_-)\,dX(s), \quad 0 \le t \le T, \tag{4.1}$$

where the notation X_- indicates the process whose value at s is $X_- = \lim_{u \to s, u<s}(X_u)$. We assume that these processes have a finite variation on [0, T]. We also assume that the stochastic integral $\int H dX = \left\{ \int_0^t H(s)\,dX(s) \right\}_{t \in [0,T]}$ is well defined for semi-martingale processes X. Further, we define $X(t,h) \equiv X(t) - X(t-h)$ for $0 \le h \le t \le T$. We proceed to the following important properties in interpreting the quadratic variation as a volatility measure.

Property (i):

If we define an increasing sequence of $\{0, \tau_{m,0}, \tau_{m,1}, \ldots\}$ so that $0 \le \tau_{m,0} \le \tau_{m,1} \le \ldots$, over the fixed time interval $[0,T]$ with $\sup_{j \ge 1}(\tau_{m,j+1} - \tau_{m,j}) \to 0$ and $\sup_{j \ge 1} \tau_{m,j} \to T$ for $m \to \infty$ with probability one, then we have

$$\lim_{m \to \infty} \left\{ X(0)Y(0) + \sum_{j \ge 1} \left[X(t \wedge \tau_{m,j}) - X(t \wedge \tau_{m,j-1}) \right] \left[X(t \wedge \tau_{m,j}) - X(t \wedge \tau_{m,j-1}) \right] \right\}$$

$$\to [\![X(t), X(t)]\!], \tag{4.2}$$

where the convergence is uniform on $[0,T]$ in probability and $(z \wedge x)$ denotes the minimum of the quantities z and x.. The left-hand side of equation (4.2) above can be approximated by the sum of squared returns at a sufficiently fine sample frequency, and the right-hand side in (4.2) represents the quadratic variation measure according to the definition in (4.1).

Property (ii):

If $X(t)$ is a locally square integrable local martingale,

$$E\left[X(t,h)^2 - \left(\llbracket X(t),X(t)\rrbracket - \llbracket X(t-h),X(t-h)\rrbracket\right) \mid F_{t-h}\right] = 0, \quad 0 < h \le t \le T, \quad \textbf{(4.3)}$$

where F_{t-h} is the information set available at time (*t-h*).

The formal description of quadratic theory can be utilized for a deep appreciation of the properties of a continuous time return process. A continuous arbitrage-free price process for general financial assets is well known to be a special type of semi-martingale, and thus, quadratic theory can be applied to this process under the assumptions mentioned above. The price process can be decomposed into a local martingale and a predictable finite variation process. The local martingale process is an "*unpredictable*" innovation. Then we can express the arbitrage-free logarithmic price process $p(t)$ over the interval $[0,T]$ as follows:

$$p(t) - p(0) = M(t) + A(t), \quad \textbf{(4.4)}$$

where $M(t)$ is a local martingale and $A(t)$ is a locally integrable and predictable process of finite variation which is deterministic drift for the price process. Since $A(t)$ is fully predictable and deterministic, $[\![A(t), A(t)]\!] = 0$ can be implied. This argument is intuitively equivalent to the fact that the variance of deterministic components should be zero and conditional mean is of no import in considering conditional variance. Then, we are allowed to focus only on martingale terms $M(t)$ in considering the quadratic variation of $p(t)$ as follows:

$$[\![p(t), p(t)]\!] = [\![M(t), M(t)]\!]. \tag{4.5}$$

According to the quadratic variation theory and the assumption of a semi-martingale price process, we can define the *h*-period quadratic variation for the continuous price process as follows:

$$\begin{aligned} Q\,\mathrm{var}_h(t) &\equiv [\![p(t), p(t)]\!] - [\![p(t-h), p(t-h)]\!] \\ &\equiv [\![M(t), M(t)]\!] - [\![M(t-h), M(t-h)]\!] \end{aligned}. \tag{4.6}$$

From equation (4.2), it is implied that the quadratic variation can be approximated by the sum of squared high frequency returns for a given interval $[t-h, t]$. Based on this notion, we define the *h*-period Realized Volatility measure at time *t*,

$$RV_{(h)}(t) \equiv \sum_{i=1,mh} r_{k,(m)}^2 \left(t-h+(i/m)\right) \qquad \text{for } i = 1,2,\ldots,m(h-1),mh, \tag{4.7}$$

where $r_{k,(m)}\left(t-h+(i/m)\right)$ is equal to $p\left(t-h+(i/m)\right)-p\left(t-h+(i-1)/m\right)$. In fact, the realized volatility in equation (4.7) is an empirical approximation for the left-hand side of equation (4.2) over the interval $[t-h,t]$ and, in turn, converges to the *h*-period quadratic variation defined in equation (4.6) by the property in (4.2). Consequently, the realized volatility is a *consistent* estimator of the theoretical volatility measure measured by the quadratic variation.

Another important notion of the realized volatility is that the volatility measure provides a model-free unbiased estimator of the conditional variance. This fact can be clarified from the property stated in equation (4.3). If we assume that $M(t)$ is a locally square integrable martingale and make use of the property shown in equation (4.3), then the conditional variance of $p(t)$ is reduced to the conditional expectation of the squared martingale term as follows:

$$\begin{aligned} \operatorname{var}\left(p(t,h) \mid F_{t-h}\right) &\equiv E\left(M(t)^2 \mid F_{t-h}\right) - E\left(M(t) \mid F_{t-h}\right)^2 \\ &= E\left(M(t)^2 \mid F_{t-h}\right) \\ &= E\left[\left(\left[\!\left[M(t),M(t)\right]\!\right] - \left[\!\left[M(t-h),M(t-h)\right]\!\right]\right) \mid F_{t-h}\right]. \end{aligned} \tag{4.8}$$

The first identity in equation (4.8) is simply a definition of the conditional variance for the arbitrage price process, and the second equality is due to the assumption of a martingale $M(t)$. The last equality in equation (4.8) results from equation (4.3). The term inside the expectation on the right hand side of the last equality in equation (4.8) is the same as the volatility measure defined in (4.6). Consequently, the conditional variance of the compounded return process over [*t-h*, *t*] interval is equivalent to the conditional expectation of the quadratic variation over the interval. As discussed above, the quadratic variation can be approximated by the realized volatility, as the number of sub-sample periods within a given interval is sufficiently large. Thus, we state that the realized volatility is an unbiased estimator for the conditional variance of the compounded returns. We have not specified any functional form in our claim, (4.8) and have instead utilized only the properties of the quadratic variation measure and the assumption of a martingale price process based on the arbitrage-free price. On the other hand, conventional GARCH models assume a parametric form to model conditional variances. Thus, the realized volatility can be said to be a model-free unbiased estimator for conditional variances.

In the theoretical asset and derivative pricing studies, it is frequently assumed that logarithmic prices follow an univariate diffusion.

$$dp(t) = \mu(t)dt + \sigma(t)W(t) \tag{4.9}$$

where $W(t)$ is a standard Brownian motion. Equivalently, we can rewrite the equation as follows:

$$p(t)-p(t-h)=\int_{t-h,t}\mu(s)ds+\int_{t-h,t}\sigma(s)dW(s). \quad \textbf{(4.10)}$$

By using the standard stochastic differential equation algebra and Ito's Lemma, it follows that

$$\left[dp(t)\right]^2=\int_{t-h,t}\sigma(s)^2\,ds. \quad \textbf{(4.11)}$$

$\sigma(t)^2$ can be termed an instantaneous volatility under the diffusion set-up, and, by using the volatility defined in (4.6), $Q\,\mathrm{var}_h(t)/h$ is close to $\sigma(t)^2$. Therefore, the integral of $\sigma(t)^2$ over the interval [*t-h*, *t*] is approximately equal to $Q\,\mathrm{var}_h(t)$ as follows:

$$Q\,\mathrm{var}_h(t)\equiv[\![M(t),M(t)]\!]-[\![M(t-h),M(t-h)]\!]=\int_{t-h,t}\sigma^2(s)ds. \quad \textbf{(4.12)}$$

Taking a conditional expectation for equation (4.12), then we have

$$E\left(Q\mathrm{var}_h(t)\mid F_{t-h}\right)=E\left(\int_{t-h,t}\sigma^2(s)ds\mid F_{t-h}\right). \quad \textbf{(4.13)}$$

The expected value of the integral metric on the right-hand side of equation (4.13), called "integrated volatility" in the literature, is of especially central interest in option pricing studies, as in Hull and White (1987), Melino (1994), Scott, and Wiggins (1987). As shown above, the realized volatility over the interval [*t*-*h*, *t*] is an unbiased estimator of the conditional expectation of $Q\,\mathrm{var}_h(t)$. Accordingly, the realized volatility provides an unbiased estimator of the integrated volatility measure for pricing derivatives securities and options.

In particular, for our applications to commodity markets, we consider the case of a one-day horizon being indicated by *h,* since the daily time horizon is the sample frequency that is of central interest for risk management, asset pricing, and portfolio allocation. In particular, the realized volatility in our study is defined as follows:

$$RV_t = 0.5\ln\left(\sum_{i=1,mh} r^2_{k,(m)}\left(t-h+\left(i/m\right)\right)\right). \qquad \textbf{(4.14)}$$

By using the properties of the quadratic variation under the assumption of a continuous arbitrage-free price process, we found that the realized volatility is a consistent estimator of true latent volatility and is a simple, unbiased estimator of conditional variances.

4.3. Practical Issues in the Calculation of Realized Volatility

High frequency commodity return data have some unique features and also some features which may be shared with other asset markets, such as the possibility of jumps and discontinuities in the volatility process arising from major economic announcements. One of the most important issues in realized volatility calculation is determining how to model the persistence of the realized volatility. We pursue this issue by applying the long memory model to the volatility measures. We also investigate the distributional properties of a realized volatility that is constructed from five-minute commodity futures returns. The realized volatility measure is also used to derive the distribution of daily commodity futures returns standardized by the new volatility measure. In addition, we analyze various important economic factors that may affect realized volatility dynamics by considering commodity-specific announcements, the time-to-maturity for the commodity futures contracts, and an information flow variable. Our approach is influenced by several previous studies. In particular, Andersen and Bollerslev (1998a) and Andersen, Bollerslev, Diebold, and Vega (2002) have documented news announcement effects on the five-minute return volatility for the US Dollar-Deutsch Mark exchange rate. Also, Bauwens, Omrane and Giot (2003) directly analyzed the news announcement effects on the realized volatility of Euro-Dollar foreign exchange returns. In this chapter, we consider commodity-specific announcements as we analyze the news effects on realized volatility using various time series methods. It appears from this study that some commodities seem to depend on specific announcements, while major foreign currencies rely on common macro news.

Unlike conventional financial assets, the volatility of commodity futures contracts with different delivery dates appears to have its own discernible characteristics due to possible seasonal patterns of the underlying physical products. Samuelson (1965) argues that futures price volatility is likely to increase as the contract approaches maturity, which has become known as the "Samuelson effect." In that sense, we consider the time-to-maturity effect on commodity futures return volatility. This time-to-maturity effect has been documented in a variety of commodity futures market studies, such as Anderson (1985), Milonas (1986), and Serletis (1992). Further, we consider information flow together with time to maturity in order to study the relationship between the realized volatility and the time to maturity.

Another economic issue for the realized volatility analysis in this chapter is mutual interdependence between the volatility measures for different commodities. Intuitively, it seems reasonable that various physical aspects of the commodities underlying the futures market may be related to their futures return volatilities. For example, it is possible that the commodity futures return volatilities belonging to similar commodity products have some dependence on one another. Later, we consider a time series model to describe the contemporaneous interdependence between different commodities' futures return realized volatilities.

As noted before, realized volatility from commodity futures is one representation of volatility that does not require a parametric model and can be easily forecasted by a simple time series econometric model.

4.4. Stochastic Properties for Realized Volatility for Modeling and Forecasting

4.4.1. The Distributional Facts of the Realized Volatility

The distributions of asset returns have been an important issue since unconditional distributions of most asset returns are usually fat-tailed, and such a feature has motivated conditional distributions relevant to various GARCH conditional variance modelings. However, the conditional distributions still remain leptokurtic, although they are less leptokurtic relative to the unconditional distributions. Turning to the distribution issue, we describe distributional characteristics of the realized volatility process for the commodities in Tables 4-2 and 4-3, while Table 4-1 shows that the unconditional distribution of daily commodity futures returns is leptokurtic. In contrast to the raw daily commodity returns, (i) realized volatilities appear to follow a normal distribution, and (ii) daily commodity returns standardized by the realized volatility are also close to normal random variables. To be precise, we assume that return process can be modeled as follows:

$$r_t = \sigma_t \varepsilon_t, \tag{4.15}$$

where r_t represents return at time t, σ_t denotes the time-varying conditional standard deviation, and ε_t is independently and identically distributed with a zero mean and unit variances for simplicity. The traditional GARCH model estimates conditional variance, σ_t^2, by assuming parametric form. As found in many previous studies, the distributions of returns standardized by the GARCH estimated conditional variances seem to still have higher kurtosis than a normal distribution, although the kurtosis seems to be lower than an unconditional return distribution. As shown in Table 4-3, the kurtosis of daily commodity returns noticeably decreases after standardization by the realized volatility. In particular, the excess kurtosis for gold futures returns is remarkably reduced from 28.9 to 3.39 by standardization using the realized volatility. For live hogs, the associated kurtosis is decreased from 7.04 to 2.92. The fact that standardized daily commodity futures returns are normally distributed is supportive of the theoretical assumption of an underlying continuous-time diffusion, which is found in many mathematical finance studies. As shown from Figure 4-1 (a), the kernel density graphs for corn, soybean, cattle, and gasoline realized volatilities are supportive of a normal distribution, while the density graphs for live hog and gold realized volatilities are still quite leptokurtic. The kernel densities for the distributions of daily returns standardized by the realized volatilities are presented in Figure 4-1 (b). For the standardized futures returns, the kernel density functions look similar to those for the normal distributions for all the commodities considered. From the realized volatility levels in Figures 4-2 (a) through (f), we can observe some peaks in the realized live hog volatility and pronounced jumps for the realized gold volatilities. This feature is consistent with the exceptional leptokurtic distribution of the realized volatility for live hogs and gold. Other than this abnormal data feature, the realized volatility seems to follow a normal distribution.

4.4.2. The Long Memory of Realized Volatility

One of the well-known facts of asset return volatility is that it is very persistent, while returns underlying the volatility are serially uncorrelated. As discussed in Baillie (1996), the ARFIMA model is a conventional parametric form used to describe slowly decaying time series processes. The theoretical background for long memory has been discussed in more detail in chapters 2 and 3. In the current chapter, we estimate a simple ARFIMA(0,d,0) model to estimate the long memory parameter for the realized volatility. For completeness of the long memory estimation, we also use the local Whittle semi-parameter approach that is explained in chapter 3. The estimation results for the ARFIMA and the local Whittle method are shown in Tables 4-4 and 4-5, respectively. The long memory parameter estimates for the commodities are in the range of 0.2 to 0.3 in most cases. Some exceptions can be found for estimates for hogs and gold greater than 0.3. The long memory estimate values for all the commodities seem to be very similar for both parametric and semi-parametric methods. As shown in Figure 4-2, the realized volatilities for hogs and gasoline seem to include some significant changes. Their higher long memory estimates may be due to some possible structural breaks. This issue would benefit from independent research but is not pursued further here. For the other commodities without unusual data features, the long memory estimates are within a stable range.

4.4.3. Forecast for Realized Volatility

Based on the theoretical background discussed in section 2, the realized volatility generated from a sufficient number of high frequency sample returns is a consistent estimator of true latent volatility factor under the assumption of an arbitrage-free price process. In that sense, we are allowed to treat the realized volatility as an observable proxy for the true underlying volatility and assess that the future realized volatility measure is the "volatility" to be forecasted. We can evaluate various volatility forecasts by considering which model provides the closest forecast to the realized volatility measure. Following Andersen, Bollerslev, Diebold, and Labys (2003) and Andersen and Bollerslev (1998b), we evaluate the forecasting performance by using a simple least square regression. The regression approach was originally employed in the literature to evaluate forecasting of the conditional mean in Mincer and Zarnowitz (1969). The generic evaluation regression can be set up as follows:

$$V_t = a_0 + a_1 CV_{t-1} + \varepsilon_t, \tag{4.16}$$

where V_t is the future volatility at time t, CV_{t-1} is the one-period ahead forecast generated under alternative conditional variance models, and ε_t is an error term. In principle, it can be implied that a_0 and a_1 should be equal to zero and unity, respectively, if we correctly specify a forecasting model for the future volatility factor, for which $E(V_t \mid \Omega_{t-1}) = a_0 + a_1 CV_{t-1}$. As we assess how the future realized volatility is to be forecasted, a natural candidate for a good volatility forecast is the one generated from the past realized volatility time series, since they are very persistent, as illustrated by the long memory estimation results for the ARFIMA and the semi-parametric estimation shown in Tables 4-4 and 4-5. The other alternative volatility forecast is generated from the GARCH-estimated conditional variances that have been elaborated by many previous studies since Engle (1982) and Bollerslev (1986).

We empirically compare the forecasting performance of the ARFIMA forecast with the GARCH forecast by running the following three OLS regression set-ups as shown in Table 4-6.

$$RV_t = b_0 + b_1 RV_{ARFIMA,t-1} + \varepsilon_t \qquad \textbf{(4.17)}$$

and

$$RV_t = b_0 + b_2 \sigma_{GARCH,t-1} + \varepsilon_t, \qquad \textbf{(4.18)}$$

where $RV_{ARFIMA,t-1}$ denotes one period ahead forecasts from the ARFIMA(0,d,0) model using the past realized volatility series and $\sigma_{GARCH,t-1}$ forecasts from the GARCH(1,1) model using the compounded daily futures returns. Based on the robust standard errors in Table 4-6, all the b_1 estimates for the regression set-up (4.17) are not significantly different from one another, although the b_0 estimates are significantly different from zero only for live cattle and gasoline. Our finding implies that the ARFIMA forecasting model is correctly specified in most cases. The forecast ability of the GARCH estimated conditional variance model[14] can be evaluated by using the set-up (4.18). According to Table 4-6, all the b_2 estimates for the GARCH forecasts are more different from one another than the corresponding estimates for the ARFIMA forecasts. In particular, the b_2 estimates for the regression (4.18) seem to be significantly different from one another for live cattle, live hogs, gasoline, and gold. The GARCH conditional variance model for those commodities seems to be mis-specified. Thus, it can be implied that using historical realized volatility series with the ARFIMA model seems to provide more correctly specified conditional variances than the GARCH model. We have some mixed evidence for R squares, since the R squares for (4.17) are higher than for (4.18) for corn, soybeans, and live cattle, while we found the opposite to be true of the other commodities.

For a more fair comparison of the forecasting performance of the ARFIMA model and the GARCH model, we regress the realized volatility of the ARFIMA forecast and the GARCH estimated conditional variances jointly as follows:

$$RV_t = b_0 + b_1 RV_{ARFIMA,t-1} + b_2 \sigma_{GARCH,t-1} + \varepsilon_t . \qquad \textbf{(4.19)}$$

[14] We also performed comparisons of the ARFIMA forecasts with the FIGARCH estimated conditional variances. The results made no meaningful difference and, thus, are not reported separately in the current paper.

Including both types of forecasts seem to improve forecasting performance quite significantly relative to the individual regressions of (4.17) and (4.18), since the adjusted R squares are higher than those for (4.17) and (4.18)[15]. All of the b_1 estimates, except for those for corn, seem to be significantly different from one. The b_2 coefficient estimates for the GARCH forecast are also significantly different from one. However, the sum of the b_1 and b_2 estimates for the regression equation (4.19) seems to be close to one for the commodities, with the exception of gold futures. This result implies that a linear combination of the ARFIMA forecast and the GARCH forecast may jointly serve to specify the correct forecasting model and may yield improved forecasting ability with higher R squares. Our findings suggest that the ARFIMA forecasts can provide a correct forecasting model when we forecast the future realized volatility for commodity futures, and their forecasting performance is not inferior to the GARCH model.

4.5. Economic Factors for the Commodity Futures Realized Volatility

An important possibility is that economic variables are relevant factors with which to describe commodity futures return volatility. To consider various types of economic factors under an integrated framework, we estimate a simple ARFIMA(0,d,0) model for each realized volatility with announcement dummies, time-to-maturity variables, and another commodity's realized volatility. This is a joint estimation for all of the considered coefficient estimates, including the long memory parameter. The estimation model takes the following form:

---

[15] We report the adjusted R squares for (4-8) for appropriate interpretation, since R squares generally tend to increase with the number of regressors.

$$(1-L)^d\left(y_t-\mu-\sum_{i=-1,2}\delta_i I_i-\gamma\cdot TM_t-\beta x_t\right)=\varepsilon_t, \qquad \textbf{(4.20)}$$

where TM_t is the time-to-maturity variable, I_i indicates i-days after the relevant announcements, δ_i denotes the coefficient for I_i, and x_t is the realized volatility of a counterpart commodity that is considered. The variable TM_t is calculated as the ratio of the number of remaining trading days (as of day t) before the futures contract's expiration to the total number of trading days within the "nearby" contract, so that the time-to-maturity variable is scaled between zero and one.

4.5.1. Announcement Effects

Intuitively, it is reasonable that commodity-specific announcements may have a meaningful relation with the relevant commodity volatility, while more general economic announcements can indirectly affect commodity markets. Recent empirical studies, including Andersen, Bollerslev, Diebold, and Vega (2003) and Andersen and Bollerslev (1998a), document the effect of macroeconomic announcements and news on the five-minute DM-US dollar return volatility. Cai, Cheung, and Wong (2001) studied how various relevant economic announcements influence five-minute gold futures return volatility. In this section, we focus on more specific announcements for the commodities in order to investigate the announcement effect together with other economic factors. An important extension from the previous studies is that we use the realized volatility measure.

We use the monthly announcement for our analysis since, for example, quarterly announcements are too sparse over the current sample period to extract sufficient information for daily volatility, and also, weekly announcements are not available for the commodities considered here. Since including more general announcement dummies may reduce parameter estimation efficiency and the degrees of freedom without making a difference for the results of the estimation, we therefore only include the most relevant announcements for each commodity. We select the following announcements for each commodity:

(1) the monthly crop production report for corn and soybeans;

(2) the monthly cattle report for live cattle;

(3) the utility capacity report for unleaded gasoline; and

(4) the production price index announcement for gold.

The hog announcements are quarterly rather than monthly-based for the sample periods of our data set, so we do not consider the announcement effects for live hogs in the current section. Since we are only considering other economic effects jointly with the announcements, we do not consider live hogs any further in the rest of this chapter. To analyze dynamic patterns of the announcement timing effect on the volatility, we classify announcement timing effects further as (1) pre-announcement effects, (2) contemporaneous effects, and (3) post-announcement effects. The pre-announcement argument is to capture any possible news-leakage effect prior to the announcements. To consider these three types of announcement timing effects, we assign dummy variables for one day before, the day of, one day after, and two days after the relevant announcements. The estimation results are presented in Tables 4-7 through 4-11, in panels (a) and (b). We mainly discuss empirical findings recorded in panel (a) in Tables 4-7 through 4-11, since those announcement results are qualitatively similar to those in panel (b) in Tables 4-7 through 4-11. We practiced the likelihood ratio test to evaluate if inclusion of the announcement dummies could yield meaningful different estimation results. The hypothesis to be tested here is that all the coefficients for the announcement dummies are equal to zero. For the corn futures realized volatility, the coefficient estimates seem to be significant and negative for one day before the announcement. According to these results, the realized volatility of corn seems to decrease one day before the monthly crop reports. Particularly for the realized soybean volatility, the likelihood ratio test statistics for the hypothesis $\delta_{-1} = \delta_0 = \delta_1 = \delta_2 = 0$ are significant, and therefore the hypothesis is rejected. Therefore, including the announcement dummies in the ARFIMA estimation of the realized soybean volatility seems to make a significant difference in terms of the maximized log likelihood values. For soybeans, the contemporaneous announcement and one day after announcement dummy coefficient estimates are significant and positive, while the estimates for the pre-announcement dummy variable are significant but negative. Such a negative pre-announcement effect is a commonality with the corn realized volatility mentioned above. For the live cattle realized volatility, the hypothesis $\delta_{-1} = \delta_0 = \delta_1 = \delta_2 = 0$ is for some instances rejected at a 10 percent significance level. Thus, including the announcement dummies can contribute to some degree to an explanation of the realized volatility. However, our findings show that none of the individual announcement coefficient estimates for live cattle are significant. For the gasoline realized volatility, the individual coefficient estimates for one day before and two days after announcements appear to be significant and negative, while the hypothesis $\delta_{-1} = \delta_0 = \delta_1 = \delta_2 = 0$ cannot be rejected by the likelihood ratio test. We found marginally significant one day after announcement estimates for the gold futures realized volatility but insignificant likelihood ratio test statistics for the hypothesis $\delta_{-1} = \delta_0 = \delta_1 = \delta_2 = 0$.

We have mixed evidence for announcement effects for the realized volatilities in the presence of the time-to-maturity effect and possible relationships between different commodities' realized volatilities. Especially for soybeans, we found that the monthly crop production reports seem to significantly affect this crop's realized volatility.

4.5.2. Time to Maturity and Information Flow

We spliced multiple futures contracts to construct a long series of futures return series. Switching from an expiring futures contract to the next nearby maturity may introduce jumps into the volatility process because of jumps in time to maturity at the switch points. In the current section, we consider the effect on the realized volatility series of varying times to maturity from different contracts. One well-known hypothesis is that futures return volatility rises as contracts approach their maturity, as Samuelson (1965) assessed. The intuition behind the Samuelson hypothesis is that there is little information flow that resolves uncertainty about futures prices in the far distant future. On the other hand, as we come closer to the maturity date, we become more sensitive to information that influences the final level of the futures price. Some previous studies support the Samuelson hypothesis, while other research failed to find evidence for the hypothesis. Anderson and Danthine (1983) and Anderson (1985) argue that the Samuelson hypothesis is generally not true unless we have information flows incorporated into the model. Anderson and Danthine (1983) introduce information flows into their theoretical model and demonstrate that the resolution of uncertainty is the source of increased volatility in futures prices. Milonas (1986) and Galloway and Kolb (1996) found some mixed evidence for the Samuelson hypothesis by using futures price series of financial assets and commodities. Chen, Duan, and Hung (1999) tested and considered the Samuelson effect to model optimal hedging under the GARCH framework by using daily spot and futures stock index data. They found the Samuelson hypothesis unsupported by their empirical findings.

According to the results in panel (a) in Tables 4-7 through 4-11, the time-to-maturity coefficient estimates for corn, soybeans, and cattle realized volatilities are significant and positive. It can be implied that, for example, the realized soybean volatility is reduced as the contracts approach their expiration dates. Seemingly, our findings may be inconsistent with the Samuelson hypothesis, since the numerical sign of the estimates should be negative according to that hypothesis. However, we use only nearby parts of the commodity futures contracts and thus do not observe the early time periods of the contracts. Therefore, we should be careful in interpreting the resulting significantly positive coefficient estimates for time-to-maturity variables. More relevant discussion of this topic will follow when we consider information flow. Particularly, lower futures return volatility immediately prior to expiration may be due to the liquidity effect. On the other hand, the time-to-maturity coefficient estimates are not significant and are negative-signed for the realized volatilities associated with gasoline and gold futures returns.

To consider an underlying factor in the relation between time to maturity and realized volatility, we introduce information flow following Anderson & Danthine (1983) and Anderson (1985), who argued the importance of information flow in explaining the Samuelson hypothesis. In a different line of previous studies including Andersen (1998), Tauchen and Pitt (1983), and Clark (1973), daily trading volumes were used as an information flow measure on the grounds of the mixture-of-distribution hypothesis. Andersen and Bollerslev (1997a) supported the long memory of high frequency US-DM return volatility based on the mixture-of-distribution hypothesis. In our study, we use the trading intensity within each day, which can be informative about how often trading occurs due to new information arriving at the relevant market. To measure trading intensity, we simply calculate the percentage of five-minute intervals associated with actual transactions out of the total sub-period intervals for each day. Not all of the intra-day intervals involve real-time trading since transactions occur unevenly. This is so-called "non-synchronous trading" in market microstructure literature.

To diagnose any possible relations among realized volatility, trading intensity, and time to maturity, we display correlations between those factors in Table 4-12. We focus on corn, soybeans, and cattle to reflect on their features which are seemingly inconsistent with the Samuelson hypothesis, as we mentioned above. We chose contracts with relatively long lifetimes for those commodities. From Table 4-12, we observe a negative correlation between time to maturity and realized volatility when we consider the whole contract periods. All the signs for the correlation data are negative, excepting only the November 2000 soybean contract. This feature is consistent with the Samuelson hypothesis. In contrast, the correlations are positive for most of the commodities if we use nearby contract,s as shown from the ARFIMA model estimation (4.20). An apparently positive correlation between time to maturity and realized volatility may be caused by the use of nearby contracts rather than a real inconsistency with the Samuelson hypothesis.

Another noteworthy thing is that the correlation between trading intensity (information flow) and realized volatility is very strong and positive, if we consider the whole contract periods. Motivated by this fact, we add the information flow variable into the ARFIMA model (4.20) and specify another time series model as follows:

$$(1-L)^d\left(y_t-\mu-\sum\nolimits_{i=-1,2}\delta_i I_i-\gamma\cdot timat_t-\beta x_t-\theta\cdot flow_t\right)=\varepsilon_t, \quad \textbf{(4.21)}$$

where $flow_t$ is the number of five-minute intervals with actual transactions divided by the number of total subperiods within a trading day. The coefficient estimates for the trading intensity variable from the ARFIMA estimation for (4.21) are presented in panel (b) in Tables 4-7 through 4-11, and they are statistically significant and positive for all of the commodities. It is worth notice that the time-to-maturity coefficient β estimates for corn, soybeans, and cattle are no longer significant using the set-up that includes the information flow variable, although those commodities showed significant time-to-maturity estimates for the original model (4.20). On the other hand, the time-to-maturity coefficient estimates for the gasoline and gold realized volatilities are insignificant for the estimation model (4.20) without considering the information flow variable. However, those estimates become significant and negative after including the information flow variable in the ARFIMA model as in (4.21).

According to our findings, trading intensity as information flow proxy seems to explain a sizeable portion of realized volatility, while time to maturity has become less relevant to the volatility for those commodities. This result is consistent with the theoretical claim by Anderson and Danthine (1983) that time to maturity could matter to futures return volatility since information flow is linked to the volatility. According to the theory, information flow is a driving factor channeling between time to maturity and realized volatility, and therefore, once information flow is taken into account in explaining the volatility measure, we may be able to observe some genuine time-to-maturity effects on the realized volatility, if any.

Another theory relevant to our findings is the mixture-of-distribution hypothesis. In particular, Clark (1973) theoretically asserted that daily returns are generated from many intra-day returns within a day, and variance in the daily price change is proportional to the number of daily transactions, although he used the daily trading volume to embody the number of daily transactions. Our finding of a significantly positive relation between the realized volatility and the trading intensity within a day can be theoretically justified by the mixture-of-distribution hypothesis since trading intensity reflects effective intra-day price changes, and those intra-day price changes underlie the realized volatility.

In addition, the comparison between the realized volatility and the squared daily return, one of the usual daily volatility measures, can be made in terms of time to maturity and information flow issues. The last two columns in Table 4-12 show (i) a correlation between trading intensity and squared daily returns, and (ii) a correlation between the time-to-maturity variable and squared daily returns. Table 4-12 shows that correlations between trade intensity and squared daily returns are very low and even negative in some instances. This is markedly in contrast with the strong and high correlations between trading intensity and the realized volatility. The average value of the correlations between trading intensity and the realized volatility is 0.5902. From the correlation results, we can imply that squared daily returns may not fully reflect the relationship between information flow and daily futures volatility. On the other hand, time to maturity and squared daily returns are negatively correlated to a less significant degree than time to maturity and realized volatilities if we consider the whole contract periods. The data for the correlations between time to maturity and squared returns are even positive for five out of 13 instances, while only one correlation between time to maturity and the realized volatility is positive. From the correlation check, we can imply that the realized volatility measure is more consistent with the Samuelson hypothesis, as well as the theoretical linkage between volatility and information flow, than the squared daily return.

4.5.3. Interdependence Between the Realized Volatilities for Different Commodities

Another economic dimension in analyzing the realized volatility is possible interdependence between different commodity volatilities. The volatility linkage between different markets is one of the active issues in empirical finance, since it is informative for portfolio management, derivative pricing, and risk management. Brunetti and Gilbert (2000) have found two similar gasoline price volatility processes correlated by using a bivariate FIGARCH model in a fractional cointegration context[16]. Fleming, Kirby, and Ostdiek (1998), Kodres and Pritsker (2002), and Fleischer (1998) examine volatility linkages by considering the relation between volatility and information flow. Fleming, Kirby, and Ostdiek (1998) argue that common information can cause a strong volatility linkage for stock, bond, and foreign currency markets. Andersen, Bollerslev, Diebold, and Labys (2003) applied the VAR model to fractionally differenced realized volatilities constructed from DM-US Dollar, Yen-Dollar, and DM-Yen exchange rate returns for the purpose of forecast modeling, and they found that many of the VAR coefficient estimates are significant. However, we use the fractional VAR estimation approach to diagnose any lead and lag relations across different commodities. We fractionally difference the realized volatilities by using the long memory parameter estimates from the ARFIMA (0,d,0) model shown in Table 4-4 and apply a VAR model to those fractionally differenced realized volatility series. The estimation results are presented in Table 4-13. According to our findings, most of the estimated VAR coefficients do not seem to be significant, therefore implying that, after we control the long memory feature for the realized volatility process, there may not be significant lead and lag interdependence between the commodities considered here. Based on this finding, we allow ourselves to concentrate on contemporaneous relations and include a counterpart commodity in the ARFIMA model for each realized commodity volatility. Table 4-14 shows a correlation matrix for the realized volatilities for corn, soybeans, live cattle, live hogs, gasoline, and gold. Higher correlations between corn and soybeans and between live cattle and hogs seem to be sensible since they belong to the same category

[16] Since we can have an observable volatility measurement, it is possible to test fractional cointegration relations directly by using several semi-parametric approaches suggested by Robinson and Marinucci (1999). This is an area of possible future research, but is not pursued further here.

of products. As shown in Tables 4-7 through 4-11, we simultaneously estimate the long memory parameter and the coefficient for linear relations between two realized volatility series in the presence of the announcement dummy variables, time-to-maturity variables, and information flow variables discussed above.

In this chapter, we directly employ realized volatility measures to consider any possible volatility linkage across different or similar types of commodities, since the realized volatility is an observable volatility proxy, unlike the stochastic volatility model Fleming, Kirby, and Ostdiek (1998) used to estimate latent volatility factors. We find mixed evidence for contemporaneous relations across different commodity futures markets. Our results show that the realized volatilities of corn and soybeans exhibit mutually significant interdependence. The long memory estimates for the realized volatility of corn are lower when they are evaluated together with soybeans than they are when evaluated with other commodities. Hence, it can be implied that the realized volatilities of corn and soybeans share long memory time trends and that there is a possible fractional cointegration relation between corn and soybean realized volatilities[17]. This finding does not seem to change with respect to different specification choices, either in (4.20) or in (4.21). On the other hand, the long memory estimates for the cattle realized volatility are slightly lower when considered in conjunction with the hog realized volatility than when other commodities are considered as its counterparts.

---

[17] Recent literature on the fractional cointegration tests includes Marinucci and Robinson (2001) under a semi-parametric framework and Dueker and Starz (1998) using a vector ARFIMA model.

4.6. Conclusion

The cumulative sum of squared intra-day returns can be a model-free and consistent estimator of the true volatility factor on the grounds of quadratic variation theory and the assumption of a continuous arbitrage-free price process under conditions of regularity, as we discussed in this chapter. More importantly, the realized volatility measure provides an observable volatility factor. In this chapter, we identified stochastic properties of the realized volatility and used the volatility measure to consider economic factors in analyzing daily futures return volatilities in major commodity futures markets. The main statistical finding is that the commodity realized volatility process is normally distributed and exhibits slowly decaying temporal dependences. This is consistent with the long memory volatility findings from many previous studies of exchange and stock markets, for example, those which produce FIGARCH estimation results using exchange rates and stocks. Based on these stochastic properties of realized volatility, we used an ARFIMA type model to analyze the presence of the announcement effect, the time-to-maturity effect (accounting for information flow), and contemporaneous interdependence between different commodity markets. Our findings are that there is significant contemporaneous interdependence between the realized volatilities of corn and soybeans, and that information flow is a key factor in commodity realized volatility, consistent with the futures return volatility model suggested by Anderson and Danthine (1983) and Clark's (1973) mixture-of-distribution hypothesis. After all the factors have been elaborated, the long memory dynamic patterns remain, and thus, slowly decaying volatility seems to be intrinsic for the commodity futures markets considered in this chapter.

With access to longer sample periods of high frequency commodity price data, there would be more opportunities to study various aspects of realized volatility dynamics. First, the out-of-sample forecasting performance of a simple ARFIMA model could be evaluated using realized volatilities. Second, a data set encompassing a longer time span may contain further nonlinearities, such as structural breaks; it would be possible to test the realized volatility series being considered for these nonlinearities. If structural change in realized volatility series were to be detected, then it would be possible to adjust to account for the breaks and to reconsider the temporal dynamic patterns of realized volatility. One last but important research possibility would be to relate realized volatility to market micro-structure issues. For example, transaction occurrences at tick time intervals and the corresponding durations may provide more insights into the relationship between information flow and volatility. We leave these issues for future studies.

Table 4-1: Basic Descriptive Statistics: Unconditional Distribution of Daily Commodity Futures Returns

	Corn	Soybean	Live Cattle	Live Hogs	Gasoline	Gold
No. of Obs.	471	409	405	400	401	401
Mean	-0.0414	-0.0179	0.0538	0.1205	0.1899	-0.0615
Median	-0.0856	-0.0907	0.0649	0.1494	0.2843	-0.0878
Maximum	3.2232	3.9491	1.4069	6.1608	6.3237	8.6236
Minimum	-4.1582	-3.6473	-1.3549	-6.3229	-5.0272	-5.6716
Std. Dev.	0.9658	1.0958	0.5138	1.265	1.7899	0.9209
Skewness	0.1579	0.2837	0.1132	-0.1692	-0.2364	1.7325
Kurtosis	4.0318	3.9841	3.0026	7.0436	3.2606	28.9425

Table 4-2: Basic Descriptive Statistics: Distribution of Realized Volatility

	Corn	Soybean	Live Cattle	Live Hog	Gasoline	Gold
No. of Obs.	471	409	405	400	401	401
Mean	0.0818	0.0337	-0.6834	0.1138	0.5745	-0.4305
Median	0.0882	0.0155	-0.681	0.0829	0.5536	-0.5078
Maximum	1.1441	0.898	0.3411	1.5969	1.4124	1.7832
Minimum	-1.0067	-0.8194	-1.5099	-2.526	-0.1352	-1.1943
Std. Dev.	0.2925	0.295	0.2895	0.4084	0.2574	0.4136
Skewness	-0.0536	0.2203	0.0051	-0.3881	0.2253	1.5114
Kurtosis	3.7882	2.9894	3.316	7.0984	3.1453	6.6992

Table 4-3: Basic Descriptive Statistics: Daily Returns Standardized by Realized Volatility

	Corn	Soybean	Live Cattle	Live Hog	Gasoline	Gold
No. of Obs.	471	409	405	400	401	401
Mean	-0.0786	-0.0451	0.1033	0.1128	0.1362	-0.1159
Median	-0.0905	-0.1202	0.1311	0.1496	0.1789	-0.1550
Maximum	1.9920	2.5650	2.4143	2.9544	2.5809	3.2496
Minimum	-2.5425	-2.3841	-2.1058	-2.3745	-2.3058	-2.0908
Std. Dev.	0.7988	0.9280	0.9299	0.9128	0.9285	0.8119
Skewness	0.0353	0.1795	0.0708	-0.0467	-0.0754	0.3923
Kurtosis	2.8075	2.6010	2.4013	2.9226	2.5930	3.3921

Key: For the basic description of each realized volatility series, we use the whole sample period of high frequency data. In contrast, we should use the realized volatility series with common trading days for the joint estimation for the ARFIMA(0,d,0) below. Thus, the numbers of sample observations for the ARFIMA estimation in Tables 4-4 through 4-11 are different from the sample numbers in Tables 4-1 through 4-3 above.

Table 4-4: ARFIMA(0,d,0) Estimation for Realized Volatility Series

$$(1-L)^d(y_t - \mu) = \varepsilon_t$$

	Corn	Soybean	Cattle	Hog	Gasoline	Gold
Sample Period	(5/03/99 -12/28/00)	(5/03/99 -12/28/00)	(5/03/99 -12/28/00)	(5/03/99 -12/28/00)	(5/03/99 -12/28/00)	(5/03/99 -12/28/00)
μ	0.0699	0.0049	-0.6548	0.1204	0.5609	-0.5395
	(0.0545)	(0.0554)	(0.0540)	(0.1046)	(0.0402)	(0.1456)
d	0.2460	0.2591	0.2702	0.3409	0.2325	0.3961
	(0.0403)	(0.0385)	(0.0356)	(0.0607)	(0.0390)	(0.0550)
σ^2	0.0703	0.0725	0.0660	0.1093	0.0565	0.1148
	(0.0053)	(0.0051)	(0.0050)	(0.0211)	(0.0043)	(0.0133)
ln(L)	-35.582	-41.472	-23.255	-121.157	6.744	-130.604
m_3	0.075	0.094	0.025	-1.205	0.289	1.131
m_4	3.221	2.916	3.206	15.494	3.272	6.203
Q(20)	29.729	20.658	17.784	15.867	9.112	21.984
$Q^2(20)$	31.441	29.764	19.998	27.349	12.924	28.223

Key: Robust standard errors based on QMLE are in parentheses below the corresponding parameter estimates. The diagnostic statistics Q(20) and $Q^2(20)$ are the Ljung-Box statistics based on the first 20 autocorrelations of the standardized residuals and the autocorrelations of the squared standardized residuals respectively. The statistics m_3 and m_4 are the sample skewness and kurtosis respectively of the standardized residuals.

Table 4-5: The Local Whittle Estimation for the Long Memory Parameter

	Local Whittle Estimates for Realized Volatility
Corn	0.2644
Soybean	0.2419
Cattle	0.2805
Hog	0.3404
Gasoline	0.2374
Gold	0.3756

Key: We use the sample size powered to 0.8 for the bandwidth.

Table 4-6: Mincer-Zarnowitz Regressions for Realized Volatilities (One-day-ahead Forecast)

	b_0	b_1	b_2	R square[18]
Corn	0.0171	1.0013		0.1572
	(0.0153)	(0.1220)		
	-0.7726		0.8885	0.0745
	(0.1566)		(0.1655)	
	-0.2847	0.8758	0.3221	0.1599
	(0.1494)	(0.1452)	(0.1640)	
Soybeans	0.00146	1.03492		0.1625
	(0.0125)	(0.1314)		
	-0.8189		0.7875	0.1600
	(0.1194)		(0.1104)	
	-0.4804	0.6219	0.4569	0.1833
	(0.1260)	(0.1427)	(0.1196)	
Live Cattle	-0.20762	0.88849		0.2501
	(0.0392)	(0.0742)		
	-2.5065		3.6166	0.2141
	(0.2120)		(0.4191)	
	-0.9971	0.6473	1.31	0.2536
	(0.3587)	(0.1285)	(0.5994)	
Live Hogs	0.0021	1.0413		0.3494
	(0.0175)	(0.0987)		
	-0.8197		0.7660	0.3532
	(0.0847)		(0.0740)	
	-0.4676	0.5511	0.4286	0.3678
	(0.1313)	(0.1612)	(0.1244)	
Gasoline	0.2239	0.8345		0.1409
	(0.0481)	(0.1107)		
	-0.4683		0.5857	0.1485
	(0.1253)		(0.0686)	
	-0.415	0.5758	0.4199	0.1898
	(0.1067)	(0.1086)	(0.0631)	
Gold	-0.0561	0.9758		0.3255
	(0.0498)	(0.1019)		
	-0.7893		0.3845	0.3373
	(0.0428)		(0.0495)	
	-0.4508	0.5005	0.2284	0.3665
	(0.0998)	(0.1488)	(0.0445)	

Key: The table reports OLS parameter estimates for Mincer-Zarnowitz regressions of realized volatility on a constant and forecasts from different models. The OLS regression

[18] We used the adjusted *R* squares for the regression including both the ARFIMA forecast and the GARCH conditional variance for more accurate interpretation since *R* squares generally tends to increase with the number of regressors.

is $RV_t = b_0 + b_1 RV_{ARFIMA,t} + b_2 CV_{GARCH,t} + u_t$. The robust standard errors are reported in the parenthesis. RV_t is $0.5*\ln(\Sigma_{i=1,A} r_{t,i})$ where A is the number of intraday returns within each trading day. $RV_{ARFIMA,t}$ is the forecasted value of RV_t using ARFIMA(0,d,0) model and $CV_{GARCH,t}$ is the GARCH estimated conditional variances. To evaluate the ARFIMA forecast alone, we restrict $b_2 = 0$. To evaluate the GARCH forecast alone, we restrict $b_1 = 0$.

Table 4-7 (a): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, and Contemporaneous Dependence between Commodity Markets: Corn

Dependent variable: Corn realized volatility					
	Soybeans	Cattle	Hogs	Gasoline	Gold
μ	-0.0286	-0.0333	-0.0592	-0.0516	-0.0131
	(0.0450)	(0.0690)	(0.0620)	(0.0680)	(0.0607)
d	0.1658	0.2170	0.2227	0.2163	0.2220
	(0.0369)	(0.0387)	(0.0394)	(0.0390)	(0.0390)
σ^2	0.0568	0.0662	0.0656	0.0662	0.0658
	(0.0043)	(0.0049)	(0.0048)	(0.0049)	(0.0049)
δ_{-1}	-0.0699	-0.1177	-0.1166	-0.1182	-0.1094
	(0.0431)	(0.0497)	(0.0489)	(0.0497)	(0.0499)
δ_0	-0.0380	0.0197	0.0300	0.0208	0.0191
	(0.0503)	(0.0523)	(0.0518)	(0.0524)	(0.0521)
δ_1	-0.0291	0.0142	0.0186	0.0142	0.0204
	(0.0649)	(0.0631)	(0.0599)	(0.0639)	(0.0632)
δ_2	0.1046	0.0850	0.0892	0.0863	0.0889
	(0.0561)	(0.0627)	(0.0633)	(0.0625)	(0.0622)
γ	0.2104	0.2305	0.2421	0.2310	0.2231
	(0.0549)	(0.0664)	(0.0668)	(0.0664)	(0.0656)
β	0.3720	0.0138	0.0757	0.0161	0.0598
	(0.0465)	(0.0477)	(0.0431)	(0.0575)	(0.0343)
ln(L)	5.731	-23.866	-22.037	-23.860	-22.651
m3	-0.135	0.075	0.082	0.078	0.084
m4	3.170	3.137	3.112	3.153	3.139
Q(20)	30.454	29.537	26.063	29.358	29.183
Q2(20)	26.895	27.232	30.119	26.618	26.398
LR Test	6.57	7.104	7.602	7.242	6.762

Key: As for Table 4-4. The estimation is based on $(1-L)^d\left(y_t-\mu-\sum_{i=-1,2}\delta_i I_i-\gamma\cdot TM_t-\beta x_t\right)=\varepsilon_t$ where TM_t is time-to-maturity variable, and I_i indicates for i-days after the relevant announcements, and δ_i denotes for the coefficient for I_i. x_t is the realized volatility for a counterpart commodity considered. The variable TM_t is calculated as the ratio of the number of remaining trading days as of day t before the futures contract expiration to the total number of trading days within the "nearby" contract so that the time-to-maturity variable is scaled between zero and one.

Table 4-7 (b): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, Information Flow, and Contemporaneous Dependence between Commodity Markets: Corn

Dependent variable: Corn realized volatility					
	Soybeans	Cattle	Hogs	Gasoline	Gold
μ	-0.6857	-0.7836	-0.8174	-0.8380	-0.7799
	(0.1502)	(0.1541)	(0.1572)	(0.1639)	(0.1578)
d	0.1796	0.2296	0.2331	0.2292	0.2385
	(0.0444)	(0.0418)	(0.0420)	(0.0423)	(0.0434)
σ^2	0.0510	0.0586	0.0582	0.0586	0.0585
	(0.0036)	(0.0042)	(0.0042)	(0.0042)	(0.0042)
δ_{-1}	-0.0377	-0.0797	-0.0793	-0.0812	-0.0745
	(0.0437)	(0.0497)	(0.0493)	(0.0494)	(0.0498)
δ_0	-0.0049	0.0452	0.0588	0.0496	0.0490
	(0.0502)	(0.0518)	(0.0516)	(0.0521)	(0.0522)
δ_1	0.0085	0.0514	0.0549	0.0514	0.0551
	(0.0514)	(0.0509)	(0.0495)	(0.0519)	(0.0519
δ_2	0.1073	0.0861	0.0918	0.0906	0.0911
	(0.0540)	(0.0596)	(0.0599)	(0.0593)	(0.0589)
γ	-0.0545	-0.0937	-0.0784	-0.0908	-0.0956
	(0.0672)	(0.0773)	(0.0768)	(0.0772)	(0.0777)
β	0.3416	0.0496	0.0709	0.0449	0.0425
	(0.0444)	(0.0453)	(0.0374)	(0.0525)	(0.0339)
θ	0.8646	1.0322	1.0143	1.0266	1.0135
	(0.1733)	(0.1791)	(0.1770)	(0.1799)	(0.1808)
ln(L)	26.884	-0.134	1.187	-0.287	0.034
m3	0.240	0.387	0.397	0.407	0.400
m4	2.985	2.988	3.036	3.010	3.005
Q(20)	26.735	28.544	23.647	28.310	27.842
Q2(20)	15.995	28.977	31.075	28.753	29.019

Key: As for Table 4-4. The estimation is based on $(1-L)^d\left(y_t-\mu-\sum_{i=-1,2}\delta_i I_i-\gamma\cdot TM_t-\beta x_t-\theta\cdot flow_t\right)=\varepsilon_t$ where TM_t is time-to-maturity variable, and I_i indicates for i-days after the relevant announcements, and δ_i denotes for the coefficient for I_i. x_t is the realized volatility for a counterpart commodity considered. The variable TM_t is calculated as the ratio of the number of remaining trading days as of day t before the futures contract expiration to the total number of trading days within the "nearby" contract so that the time-to-maturity variable is scaled between zero and one. $flow_t$ is the number of five-minute intervals with actual transactions divided by the number of total subperiods within a trading day.

Table 4-8 (a): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, and Contemporaneous Dependence between Commodity Markets: Soybean

Dependent variable: Soybean realized volatility					
	Corn	Cattle	Hogs	Gasoline	Gold
μ	-0.0808	-0.0281	-0.0772	-0.1411	-0.0422
	(0.0526)	(0.0767)	(0.0645)	(0.0747)	(0.0684)
d	0.2188	0.2525	0.2528	0.2562	0.2547
	(0.0374)	(0.0377)	(0.0385)	(0.0364)	(0.0371)
σ^2	0.0590	0.0677	0.0679	0.0672	0.0676
	(0.0044)	(0.0051)	(0.0051)	(0.0051)	(0.0051)
δ_{-1}	-0.0711	-0.1193	-0.1189	-0.1247	-0.1111
	(0.0585)	(0.0655)	(0.0644)	(0.0653)	(0.0656)
δ_0	0.1656	0.1610	0.1730	0.1640	0.1672
	(0.0471)	(0.0491)	(0.0499)	(0.0479)	(0.0486)
δ_1	0.1248	0.1231	0.1247	0.1221	0.1293
	(0.0536)	(0.0495)	(0.0496)	(0.0489)	(0.0506)
δ_2	-0.0832	-0.0486	-0.0433	-0.0401	-0.0428
	(0.0458)	(0.0527)	(0.0526)	(0.0508)	(0.0510)
γ	0.1127	0.1436	0.1413	0.1380	0.1317
	(0.0535)	(0.0599)	(0.0599)	(0.0588)	(0.0588)
β	0.3684	0.0701	0.0312	0.1242	0.0582
	(0.0504)	(0.0511)	(0.0429)	(0.0561)	(0.0369)
ln(L)	-1.327	-28.238	-28.841	-26.651	-28.015
m3	0.246	0.186	0.201	0.245	0.193
m4	3.127	3.160	3.161	3.205	3.224
Q(20)	24.099	28.412	28.774	31.350	28.741
Q2(20)	19.150	25.281	27.717	25.461	26.377
LR Test	20.124***	19.230***	20.212***	19.856***	19.242***

Key: As for table 4-7 (a). (***) represents significant the Likelihood Ratio test statistic at one percent level.

Table 4-8 (b): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, Information Flow, and Contemporaneous Dependence between Commodity Markets: Soybean

Dependent variable: Soybean realized volatility					
	Corn	Cattle	Hogs	Gasoline	Gold
μ	-0.4625	-0.4930	-0.5338	-0.5603	-0.4870
	(0.1756)	(0.1831)	(0.1817)	(0.1800)	(0.1841)
d	0.2611	0.2902	0.2901	0.2922	0.2927
	(0.0488)	(0.0435)	(0.0442)	(0.0426)	(0.0431)
σ^2	0.0574	0.0653	0.0658	0.0653	0.0657
	(0.0044)	(0.0050)	(0.0051)	(0.0051)	(0.0051)
δ_{-1}	-0.0494	-0.0907	-0.0915	-0.0978	-0.0874
	(0.0560)	(0.0622)	(0.0614)	(0.0628)	(0.0628)
δ_0	0.1795	0.1749	0.1894	0.1799	0.1833
	(0.0468)	(0.0489)	(0.0498)	(0.0485)	(0.0492)
δ_1	0.1391	0.1403	0.1414	0.1378	0.1431
	(0.0568)	(0.0542)	(0.0544)	(0.0536)	(0.0549)
δ_2	-0.0662	-0.0309	-0.0250	-0.0238	-0.0263
	(0.0436)	(0.0488)	(0.0487)	(0.0473)	(0.0475)
γ	-0.0522	-0.0644	-0.0585	-0.0505	-0.0576
	(0.0919)	(0.0943)	(0.0941)	(0.0929)	(0.0940)
β	0.3554	0.0932	0.0363	0.1060	0.0390
	(0.0492)	(0.0511)	(0.0414)	(0.0550)	(0.0363)
θ	0.5073	0.6369	0.6065	0.5716	0.5795
	(0.2224)	(0.2202)	(0.2206)	(0.2191)	(0.2218)
ln(L)	3.837	-21.309	-22.523	-21.076	-22.438
m3	0.266	0.220	0.238	0.263	0.227
m4	3.303	3.280	3.312	3.359	3.370
Q(20)	24.637	28.745	28.692	30.750	28.825
Q2(20)	21.446	26.225	29.634	26.361	28.436

Key: As for table 4-7 (b)

Table 4-9 (a): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, and Contemporaneous Dependence between Commodity Markets: Cattle

	Dependent variable: Cattle realized volatility				
	Corn	Soybeans	Hogs	Gasoline	Gold
μ	-0.8165	-0.8182	-0.8207	-0.8399	-0.7899
	(0.0595)	(0.0588)	(0.0573)	(0.0644)	(0.0634)
d	0.2537	0.2521	0.2452	0.2555	0.2561
	(0.0324)	(0.0327)	(0.0343)	(0.0326)	(0.0328)
σ^2	0.0597	0.0595	0.0593	0.0596	0.0592
	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0041)
δ_{-1}	0.1139	0.1169	0.1074	0.1104	0.1197
	(0.0755)	(0.0748)	(0.0752)	(0.0752)	(0.0750)
δ_0	-0.0394	-0.0310	-0.0408	-0.0374	-0.0379
	(0.0597)	(0.0591)	(0.0599)	(0.0607)	(0.0611)
δ_1	-0.0008	0.0068	-0.0058	0.0026	0.0154
	(0.0530)	(0.0550)	(0.0544)	(0.0540)	(0.0545)
δ_2	-0.1170	-0.1075	-0.1165	-0.1109	-0.1126
	(0.0607)	(0.0623)	(0.0602)	(0.0610)	(0.0600)
γ	0.3241	0.3233	0.3156	0.3198	0.3315
	(0.0626)	(0.0627)	(0.0615)	(0.0633)	(0.0625)
β	-0.0070	0.0550	0.0644	0.0448	0.0683
	(0.0440)	(0.0446)	(0.0369)	(0.0572)	(0.0390)
ln(L)	-3.856	-3.173	-2.406	-3.510	-2.099
m3	0.018	0.035	-0.003	0.015	-0.013
m4	2.940	2.926	2.924	2.918	2.880
Q(20)	27.495	26.498	24.246	28.064	27.914
Q2(20)	17.899	16.619	18.749	18.605	15.898
LR Test	8.064*	7.620	7.644	8.438*	8.538*

Key: As for table 4-7 (a)

Table 4-9 (b): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, Information Flow, and Contemporaneous Dependence between Commodity Markets: Cattle

	Dependent variable: Cattle realized volatility				
	Corn	Soybeans	Hogs	Gasoline	Gold
μ	-2.2860	-2.2778	-2.3152	-2.2885	-2.2474
	(0.1473)	(0.1464)	(0.1474)	(0.1512)	(0.1445)
d	0.2248	0.2221	0.1924	0.2265	0.2299
	(0.0331)	(0.0332)	(0.0389)	(0.0337)	(0.0339)
σ^2	0.0480	0.0479	0.0472	0.0480	0.0476
	(0.0035)	(0.0035)	(0.0034)	(0.0035)	(0.0034)
δ_{-1}	0.1039	0.1099	0.0979	0.1042	0.1122
	(0.0612)	(0.0611)	(0.0621)	(0.0614)	(0.0608)
δ_0	-0.0614	-0.0553	-0.0662	-0.0624	-0.0623
	(0.0511)	(0.0500)	(0.0488)	(0.0511)	(0.0519)
δ_1	0.0701	0.0781	0.0651	0.0722	0.0852
	(0.0553)	(0.0556)	(0.0564)	(0.0556)	(0.0558)
δ_2	-0.0975	-0.0893	-0.0973	-0.0950	-0.0949
	(0.0564)	(0.0580)	(0.0556)	(0.0568)	(0.0557)
γ	0.0293	0.0317	0.0128	0.0301	0.0404
	(0.0591)	(0.0592)	(0.0578)	(0.0597)	(0.0590)
β	0.0283	0.0566	0.0954	0.0295	0.0634
	(0.0393)	(0.0405)	(0.0412)	(0.0517)	(0.0367)
θ	1.7114	1.7015	1.7387	1.6983	1.6970
	(0.1629)	(0.1625)	(0.1648)	(0.1628)	(0.1598)
ln(L)	38.403	39.097	41.906	38.370	40.093
m3	0.088	0.095	0.040	0.078	0.028
m4	3.044	3.025	3.033	3.019	3.006
Q(20)	33.359	32.712	29.739	34.787	32.848
Q2(20)	25.928	25.355	24.515	26.001	27.595

Key: As for table 4-7 (b)

Table 4-10 (a): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, and Contemporaneous Dependence between Commodity Markets: Gasoline

Dependent variable: Gasoline realized volatility					
	Corn	Soybeans	Cattle	Hogs	Gold
μ	0.5866	0.5838	0.6478	0.5826	0.6108
	(0.0507)	(0.0504)	(0.0607)	(0.0519)	(0.0536)
d	0.2373	0.2384	0.2468	0.2435	0.2425
	(0.0418)	(0.0412)	(0.0399)	(0.0412)	(0.0415)
σ^2	0.0555	0.0549	0.0550	0.0553	0.0552
	(0.0042)	(0.0041)	(0.0041)	(0.0042)	(0.0041)
δ_{-1}	-0.1097	-0.0961	-0.1170	-0.1137	-0.1056
	(0.0489)	(0.0477)	(0.0470)	(0.0500)	(0.0493)
δ_0	-0.0252	-0.0175	-0.0186	-0.0324	-0.0221
	(0.0468)	(0.0471)	(0.0477)	(0.0472)	(0.0458)
δ_1	-0.0015	0.0169	0.0024	-0.0050	-0.001
	(0.0539)	(0.0554)	(0.0559)	(0.0532)	(0.0527)
δ_2	-0.1133	-0.1048	-0.1105	-0.1141	-0.1093
	(0.0618)	(0.0594)	(0.0632)	(0.0609)	(0.0616)
γ	-0.0308	-0.0309	-0.0428	-0.0315	-0.0318
	(0.0514)	(0.0517)	(0.0506)	(0.0513)	(0.0511)
β	0.0007	0.0920	0.0862	0.0385	0.0532
	(0.0471)	(0.0452)	(0.0510)	(0.0356)	(0.0355)
ln(L)	10.348	12.438	12.000	10.907	11.531
m3	0.287	0.303	0.297	0.309	0.285
m4	3.213	3.189	3.126	3.240	3.173
Q(20)	9.099	9.671	9.034	8.865	9.719
Q2(20)	10.941	12.602	13.329	11.193	11.080
LR test statistic	6.866	6.112	7.362	7.108	7.246

Key: As for table 4-7 (a)

Table 4-10 (b): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, Information Flow, and Contemporaneous Dependence between Commodity Markets: Gasoline

Dependent variable: Gasoline realized volatility					
	Corn	Soybeans	Cattle	Hogs	Gold
μ	-1.0029	-0.9843	-0.9394	-0.9969	-0.9728
	(0.2086)	(0.2082)	(0.2106)	(0.2105)	(0.2098)
d	0.2024	0.2037	0.2132	0.2075	0.2071
	(0.0401)	(0.0397)	(0.0399)	(0.0425)	(0.0399)
σ^2	0.0471	0.0467	0.0467	0.0470	0.0469
	(0.0038)	(0.0037)	(0.0037)	(0.0038)	(0.0038)
δ_{-1}	-0.1489	-0.1389	-0.1557	-0.1513	-0.1461
	(0.0454)	(0.0444)	(0.0438)	(0.0457)	(0.0457)
δ_0	-0.0439	-0.0383	-0.0379	-0.0480	-0.0417
	(0.0470)	(0.0477)	(0.0475)	(0.0473)	(0.0462)
δ_1	-0.0422	-0.0283	-0.0387	-0.0443	-0.0418
	(0.0454)	(0.0471)	(0.0473)	(0.0451)	(0.0446)
δ_2	-0.1250	-0.1187	-0.1224	-0.1255	-0.1219
	(0.0502)	(0.0493)	(0.0517)	(0.0500)	(0.0507)
γ	-0.1864	-0.1841	-0.1969	-0.1861	-0.1861
	(0.0469)	(0.0475)	(0.0464)	(0.0469)	(0.0469)
β	-0.0079	0.0697	0.0784	0.0212	0.0393
	(0.0423)	(0.0424)	(0.0484)	(0.0398)	(0.0326)
θ	1.8066	1.7820	1.7981	1.7969	1.7920
	(0.2280)	(0.2281)	(0.2257)	(0.2308)	(0.2286)
ln(L)	42.291	43.687	43.878	42.471	43.049
m3	0.406	0.425	0.402	0.427	0.392
m4	3.510	3.468	3.430	3.523	3.493
Q(20)	7.158	7.873	7.378	7.391	7.658
Q2(20)	13.014	13.033	15.750	12.680	14.289

Key: As for table 4-7 (b)

Table 4-11 (a): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, and Contemporaneous Dependence between Commodity Markets: Gold

	Dependent variable: Gold realized volatility				
	Corn	Soybeans	Cattle	Hogs	Gasoline
μ	-0.5267	-0.5075	-0.4562	-0.5239	-0.5954
	(0.1471)	(0.1461)	(0.1499)	(0.1511)	(0.1548)
d	0.3997	0.3967	0.4006	0.4029	0.4006
	(0.0583)	(0.0583)	(0.0577)	(0.0638)	(0.0581)
σ^2	0.1126	0.1129	0.1131	0.1137	0.1128
	(0.0130)	(0.0131)	(0.0132)	(0.0129)	(0.0133)
δ_{-1}	-0.0167	-0.0329	-0.0331	-0.0325	-0.0427
	(0.0713)	(0.0724)	(0.0715)	(0.0724)	(0.0728)
δ_0	-0.0079	0.0000	-0.0018	0.0018	-0.0044
	(0.0700)	(0.0672)	(0.0660)	(0.0674)	(0.0686)
δ_1	-0.1239	-0.1274	-0.1256	-0.1248	-0.1229
	(0.0651)	(0.0646)	(0.0651)	(0.0659)	(0.0649)
δ_2	-0.0363	-0.0284	-0.0400	-0.0376	-0.0309
	(0.0622)	(0.0615)	(0.0631)	(0.0625)	(0.0613)
γ	-0.0173	-0.0365	-0.0267	-0.0292	-0.0182
	(0.1163)	(0.1159)	(0.1162	(0.1160)	(0.1130)
β	0.1340	0.1130	0.1012	0.0382	0.1322
	(0.0571)	(0.0592)	(0.0661)	(0.1030)	(0.0686)
ln(L)	-126.814	-127.382	-127.814	-128.734	-127.269
m3	1.123	1.138	1.168	1.109	1.131
m4	6.177	6.213	6.269	6.006	6.367
Q(20)	20.654	21.082	21.109	19.187	21.165
Q2(20)	28.617	28.985	28.738	28.937	26.186
LR test statistic	2.774	3.172	3.036	3.052	2.980

Key: As for table 4-7 (a)

Table 4-11 (b): ARFIMA(0,d,0) Estimation for the Announcement effect, Time-to-maturity effect, Information Flow, and Contemporaneous Dependence between Commodity Markets: Gold

Dependent variable: Gold realized volatility					
	Corn	Soybeans	Cattle	Hogs	Gasoline
μ	-2.2811	-2.2563	-2.2270	-2.3143	-2.3558
	(0.4323)	(0.4396)	(0.4409)	(0.4484)	(0.4259)
d	0.4030	0.4000	0.4035	0.4092	0.4021
	(0.0641)	(0.0637)	(0.0631)	(0.0714)	(0.0633)
σ^2	0.1004	0.1010	0.1009	0.1011	0.1005
	(0.0119)	(0.0121)	(0.0121)	(0.0117)	(0.0122)
δ_{-1}	-0.0486	-0.0631	-0.0632	-0.0622	-0.0723
	(0.0709)	(0.0714)	(0.0707)	(0.0718)	(0.0719)
δ_0	-0.0216	-0.0147	-0.0162	-0.0116	-0.0188
	(0.0695)	(0.0692)	(0.0668)	(0.0711)	(0.0689)
δ_1	-0.1194	-0.1225	-0.1208	-0.1187	-0.1183
	(0.0603)	(0.0608)	(0.0608)	(0.0622)	(0.0605)
δ_2	-0.0350	-0.0300	-0.0383	-0.0357	-0.0299
	(0.0603)	(0.0596)	(0.0605)	(0.0601)	(0.0595)
γ	-0.2022	-0.2156	-0.2114	-0.2174	-0.2021
	(0.1255)	(0.1249)	(0.1253)	(0.1279)	(0.1242)
β	0.1185	0.0776	0.0893	0.0557	0.1237
	(0.0552)	(0.0563)	(0.0640)	(0.1002)	(0.0658)
θ	1.9566	1.9455	1.9659	1.9921	1.9681
	(0.4324)	(0.4400)	(0.4437)	(0.4473)	(0.4299)
ln(L)	-104.679	-105.746	-105.559	-105.964	-104.892
m3	1.299	1.320	1.344	1.276	1.308
m4	6.473	6.557	6.596	6.178	6.685
Q(20)	26.239	26.326	26.951	24.996	27.286
Q2(20)	28.877	28.831	28.529	32.131	26.776

Key: As for table 4-7 (b)

Table 4-12: Correlation among the realized volatility, squared daily return, trading intensity, and time-to-maturity

Item	Contract	Corr. Between Realized Volatility and Time-to-maturity (Whole)	Corr. Between Realized Volatility and Time-to-maturity (Nearby)	Corr. Between Realized Volatility and Trading Intensity (Whole)	Corr. Between Time-to-maturity and Trading Intensity (Whole)	Corr. Between Squared Returns and Trading Intensity (Whole)	Corr. Between Squared Returns and Time-to-maturity (Whole)
Corn	2000.05	-0.313	0.220	0.501	-0.833	0.053	0.048
Corn	2000.07	-0.521	0.348	0.608	-0.684	0.112	-0.105
Corn	2000.09	-0.650	0.403	0.857	-0.794	0.240	-0.132
Soybean	2000.05	-0.067	0.202	0.221	-0.861	-0.030	0.109
Soybean	2000.07	-0.015	0.062	0.178	-0.630	0.008	0.035
Soybean	2000.08	-0.399	0.092	0.552	-0.856	-0.011	0.027
Soybean	2000.09	-0.477	-0.380	0.668	-0.852	0.079	-0.043
Soybean	2000.11	0.049	0.339	0.561	-0.128	0.056	0.038
Cattle	2000.04	-0.522	0.292	0.737	-0.856	0.179	-0.092
Cattle	2000.06	-0.400	0.635	0.665	-0.814	0.056	0.072
Cattle	2000.08	-0.429	0.426	0.662	-0.732	0.014	-0.038
Cattle	2000.10	-0.510	0.289	0.637	-0.836	0.169	-0.160
Cattle	2000.12	-0.755	0.056	0.820	-0.902	0.324	-0.321

Table 4-13: VAR Parameter Estimates (regression form)

X	Corn	Soybean	Cattle	Hogs	Gasoline	Gold
Const.	0.0099	-0.0129	-0.1596	0.0252	0.1597	0.0121
	(0.0337)	(0.0346)	(0.0326)	(0.0421)	(0.0306)	(0.0417)
Lag 1						
Corn	-0.0710	0.0233	0.0181	-0.1596	-0.0531	-0.0058
	(0.0580)	(0.0595)	(0.0562)	(0.0725)	(0.0526)	(0.0718)
Soybean	0.0487	-0.0544	0.027	-0.019	0.0671	0.1000
	(0.0567)	(0.0582)	(0.0550)	(0.0709)	(0.0515)	(0.0702)
Cattle	-0.021	-0.0471	-0.0658	-0.0487	0.0576	0.0784
	(0.0550)	(0.0565)	(0.0534)	(0.0689)	(0.0500)	(0.0681)
Hogs	-0.0224	-0.0509	-0.022	-0.0397	0.0351	-0.1881
	(0.0427)	(0.0439)	(0.0415)	(0.0535)	(0.0388)	(0.0529)
Gasoline	0.0879	-0.0458	0.0247	0.0360	-0.0281	-0.057
	(0.0598)	(0.0614)	(0.0580)	(0.0748)	(0.0543)	(0.0740)
Gold	-0.059	0.0338	-0.0056	-0.0352	0.0131	-0.0752
	(0.0430)	(0.0442)	(0.0417)	(0.0538)	(0.0391)	(0.0533)
Lag 2						
Corn	0.1144	0.0548	0.0532	-0.0201	0.0504	0.0964
	(0.0580)	(0.0596)	(0.0563)	(0.0726)	(0.0527)	(0.0718)
Soybean	-0.0119	-0.1047	-0.0046	0.0256	0.022	-0.0917
	(0.0568)	(0.0583)	(0.0551)	(0.0711)	(0.0516)	(0.0703)
Cattle	0.003	-0.0955	0.0036	-0.0366	-0.0336	-0.0783
	(0.0548)	(0.0563)	(0.0531)	(0.0686)	(0.0497)	(0.0678)
Hogs	-0.0097	0.0749	0.015	0.0208	0.0314	-0.0591
	(0.0429)	(0.0440)	(0.0416)	(0.0536)	(0.0389)	(0.0531)
Gasoline	-0.0969	0.0664	-0.0231	0.0561	0.0049	-0.1081

	(0.0594)	(0.0610)	(0.0576)	(0.0743)	(0.0539)	(0.0735)
Gold	0.0089	0.0144	0.062	0.0598	-0.0428	-0.0031
	(0.0429)	(0.0441)	(0.0416)	(0.0537)	(0.0390)	(0.0531)
Lag 3						
Corn	-0.0337	0.0402	-0.0132	0.0978	-0.0335	-0.0194
	(0.0582)	(0.0597)	(0.0564)	(0.0728)	(0.0528)	(0.0720)
Soybean	0.0884	0.0573	-0.0434	0.02	0.0291	-0.0752
	(0.0568)	(0.0583)	(0.0551)	(0.0710)	(0.0515)	(0.0703)
Cattle	-0.0968	-0.0124	-0.0347	0.0669	-0.0076	-0.0346
	(0.0551)	(0.0566)	(0.0534)	(0.0689)	(0.0500)	(0.0682)
Hogs	0.0289	0.0554	0.0412	0.0725	-0.0199	-0.0729
	(0.0431)	(0.0442)	(0.0418)	(0.0539)	(0.0391)	(0.0533)
Gasoline	0.0285	-0.0098	0.0193	-0.0977	-0.0485	-0.0571
	(0.0598)	(0.0614)	(0.0580)	(0.0748)	(0.0543)	(0.0740)
Gold	-0.011	0.0505	-0.0027	0.0036	-0.0287	0.1055
	(0.0426)	(0.0437)	(0.0413)	(0.0533)	(0.0386)	(0.0527)
Lag 4						
Corn	0.0326	0.0601	-0.0082	-0.036	0.1221	0.0235
	(0.0573)	(0.0589)	(0.0556)	(0.0718)	(0.0521)	(0.0710)
Soybean	-0.0113	0.0174	0.106	0.0171	-0.045	0.1459
	(0.0563)	(0.0578)	(0.0546)	(0.0705)	(0.0511)	(0.0697)
Cattle	-0.0283	0.0001	0.0256	0.1546	-0.0285	-0.0275
	(0.0555)	(0.0570)	(0.0539)	(0.0695)	(0.0504)	(0.0688)
Hogs	0.0216	0.0911	0.0142	0.0529	-0.011	-0.0583
	(0.0433)	(0.0444)	(0.0420)	(0.0541)	(0.0393)	(0.0536)
Gasoline	0.0906	0.0104	0.0552	-0.029	0.0497	-0.0481
	(0.0597)	(0.0613)	(0.0579)	(0.0747)	(0.0542)	(0.0739)

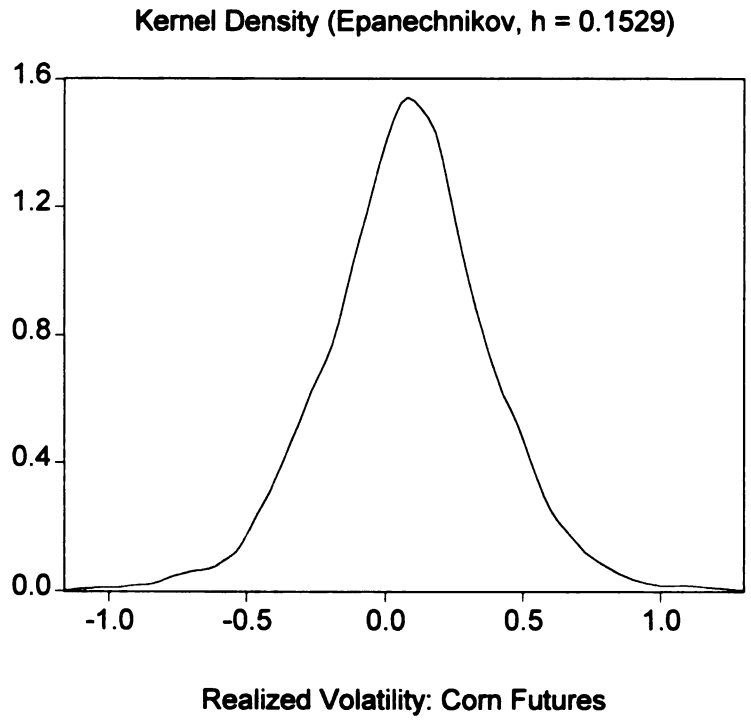
Gold	0.0215	0.0145	-0.0651	0.0427	0.018	0.005
	(0.0422)	(0.0434)	(0.0410)	(0.0529)	(0.0384)	(0.0523)
Lag 5						
Corn	-0.017	-0.0045	0.017	-0.1346	-0.0034	-0.0511
	(0.0571)	(0.0586)	(0.0554)	(0.0714)	(0.0518)	(0.0707)
Soybean	-0.0195	0.0327	-0.0682	0.0587	0.0264	-0.0355
	(0.0568)	(0.0583)	(0.0551)	(0.0711)	(0.0516)	(0.0703)
Cattle	0.1075	0.0706	-0.0084	-0.0303	0.0077	-0.0035
	(0.0554)	(0.0569)	(0.0537)	(0.0693)	(0.0503)	(0.0686)
Hogs	-0.0252	0.0535	0.0552	0.0421	-0.0083	-0.0251
	(0.0431)	(0.0442)	(0.0418)	(0.0539)	(0.0391)	(0.0533)
Gasoline	-0.0539	-0.0167	-0.0976	0.0662	-0.0305	-0.0963
	(0.0594)	(0.0610)	(0.0577)	(0.0744)	(0.0540)	(0.0736)
Gold	0.0279	-0.0458	0.0172	0.028	0.0301	0.1167
	(0.0411)	(0.0422)	(0.0399)	(0.0515)	(0.0373)	(0.0509)

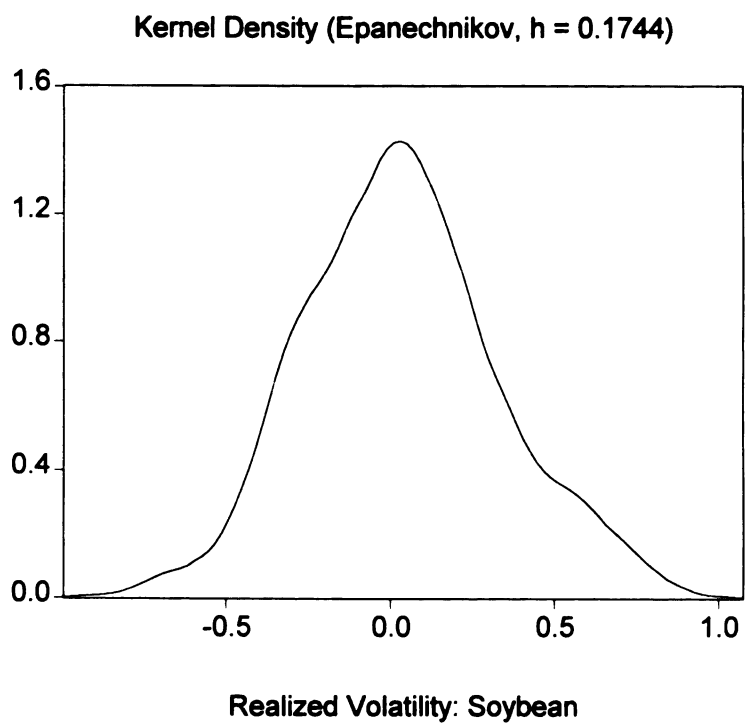
Key: The standard errors are in parentheses.

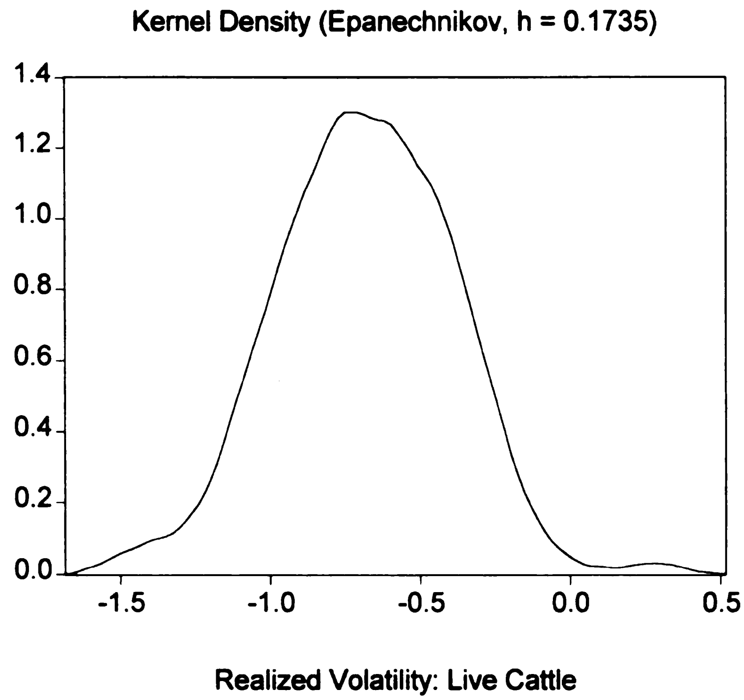
Table 4-14: Correlation matrix for six realized volatility series

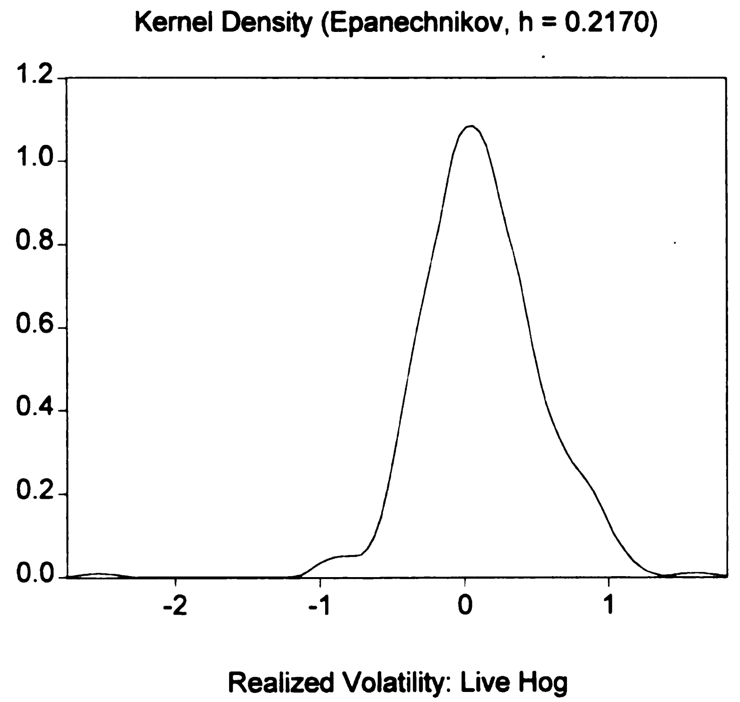
	Corn	Soybean	Cattle	Hog	Gasoline	Gold
Corn	1.00000	0.43224	-0.00395	0.02822	0.04000	0.07547
Soybean	0.43224	1.00000	0.12513	0.12368	0.08430	0.10974
Cattle	-0.00395	0.12513	1.00000	0.28375	-0.04924	0.03887
Hogs	0.02822	0.12368	0.28375	1.00000	-0.10063	-0.02247
Gasoline	0.04000	0.08430	-0.04924	-0.10063	1.00000	0.04274
Gold	0.07547	0.10974	0.03887	-0.02247	0.04274	1.00000

Figure 4-1(a). Kernel Density for Realized Volatility

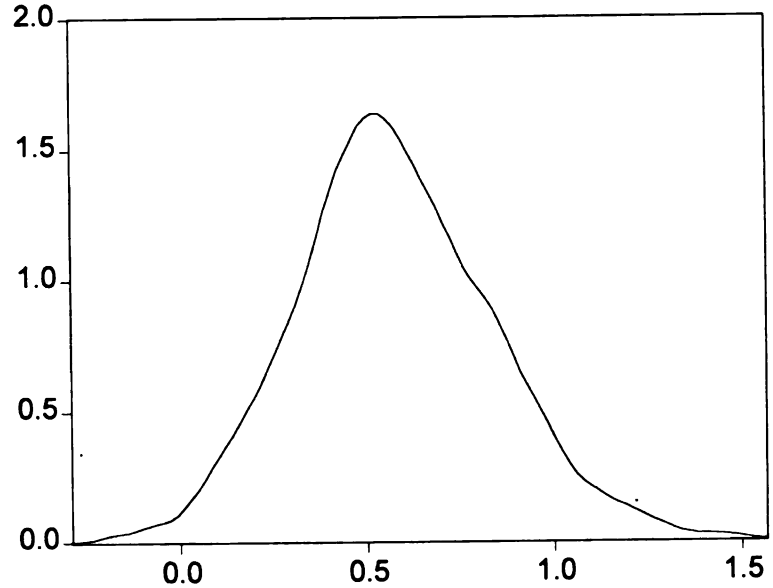





Kernel Density (Epanechnikov, h = 0.1735)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-1.5
-1.0
-0.5
0.0
0.5
Realized Volatility: Live Cattle

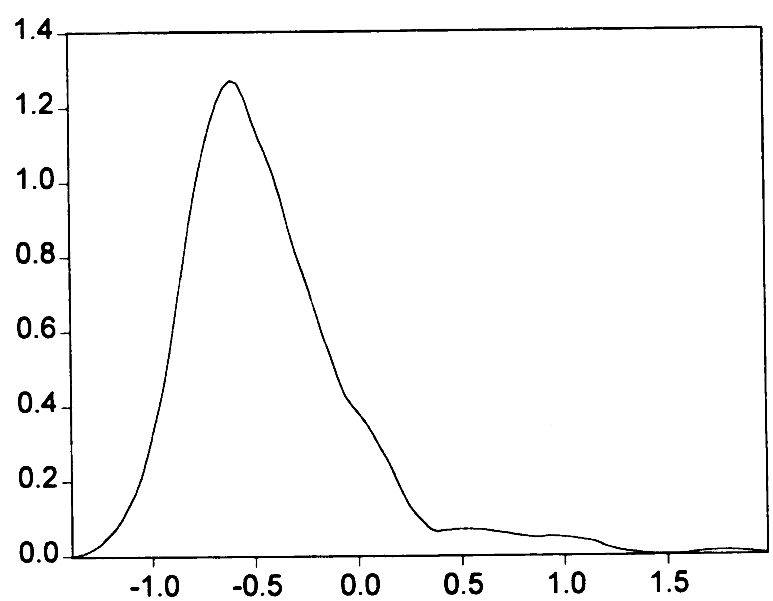

Kernel Density (Epanechnikov, h = 0.2170)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-2
-1
0
1
Realized Volatility: Live Hog

Kernel Density (Epanechnikov, h = 0.1533)



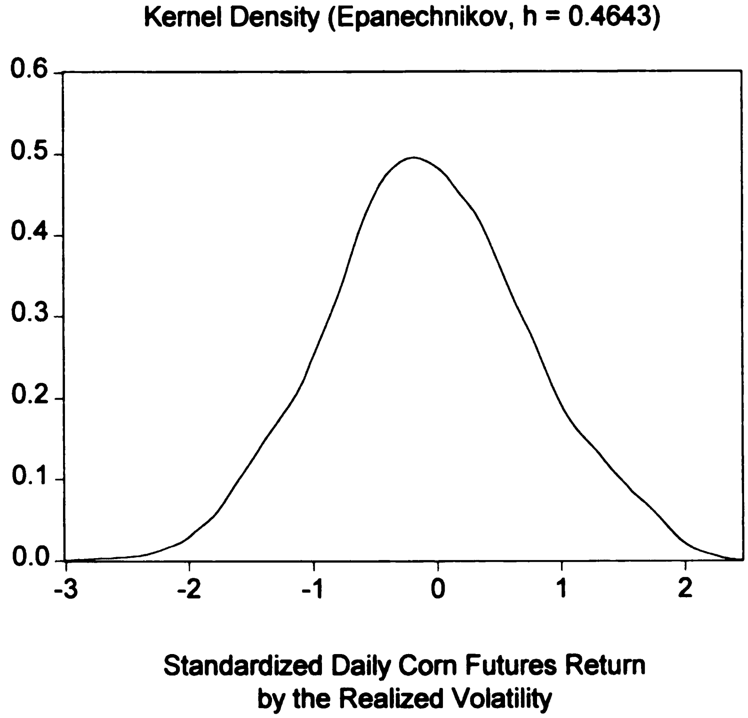
Realized Volatility: Unleaded Gasoline

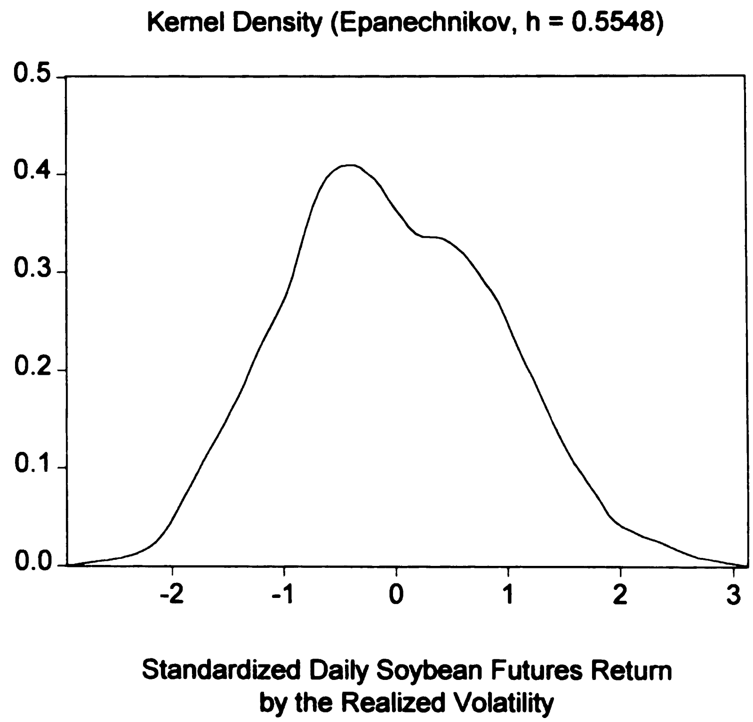
Kernel Density (Epanechnikov, h = 0.1976)



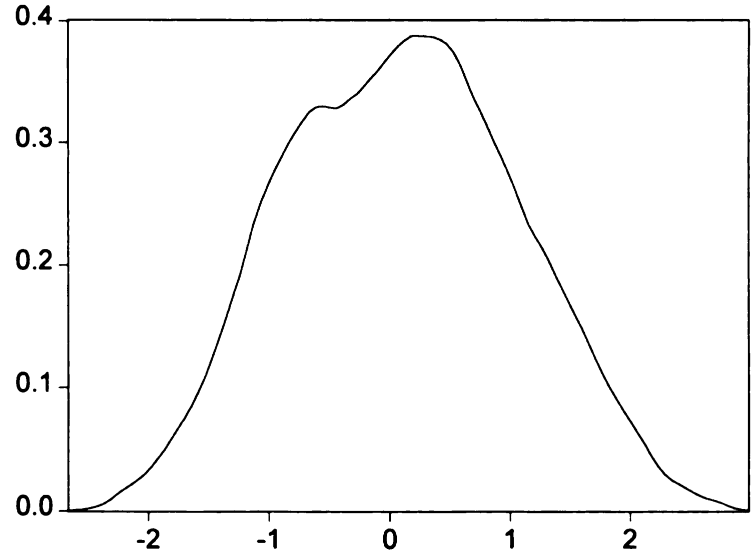
Realized Volatility: Gold

Figure 4-1(b). Kernel Density for Daily Returns Standardized by Realized Volatility



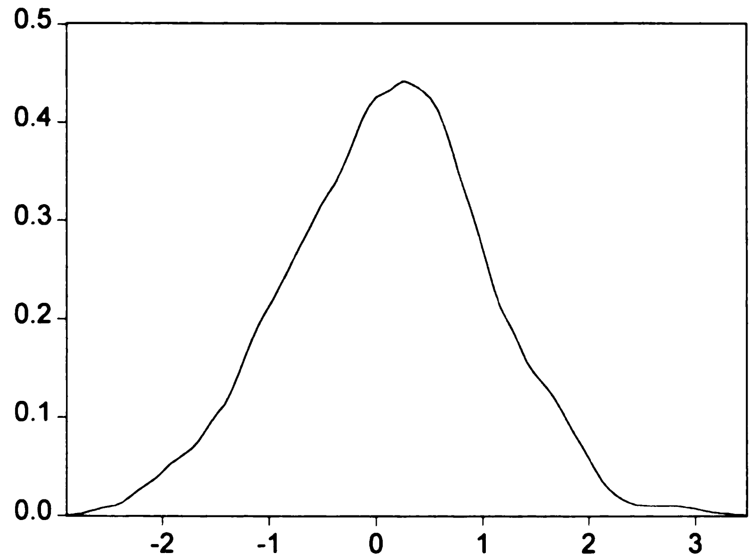


Kernel Density (Epanechnikov, h = 0.5570)



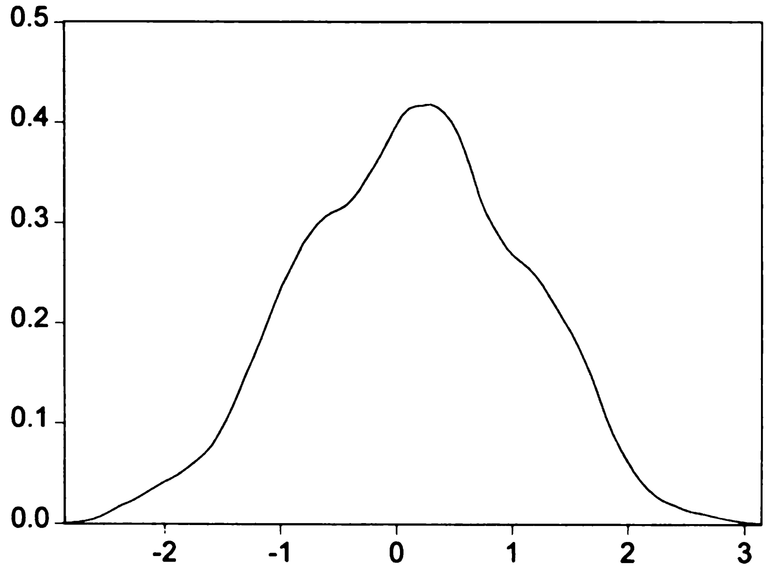
Standardized Daily Cattle Futures Return
by the Realized Volatility

Kernel Density (Epanechnikov, h = 0.5212)



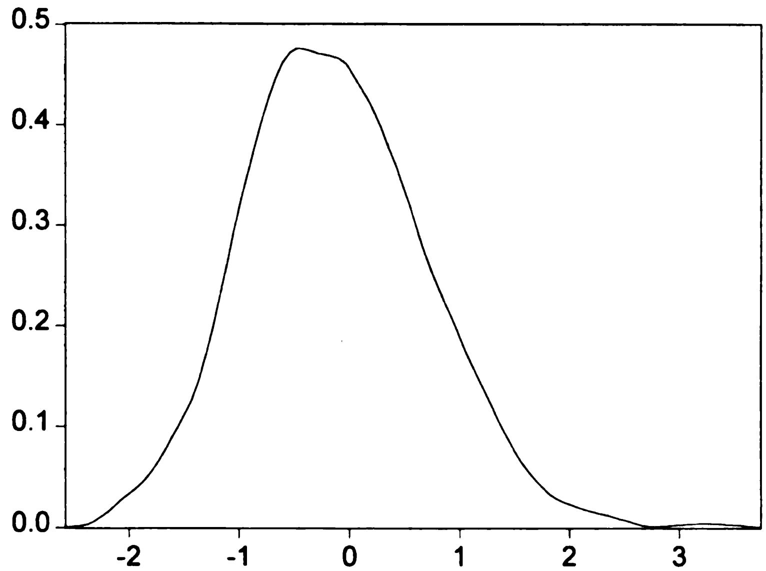
Standardized Daily Hog Futures Return
by the Realized Volatility

Kernel Density (Epanechnikov, h = 0.5573)



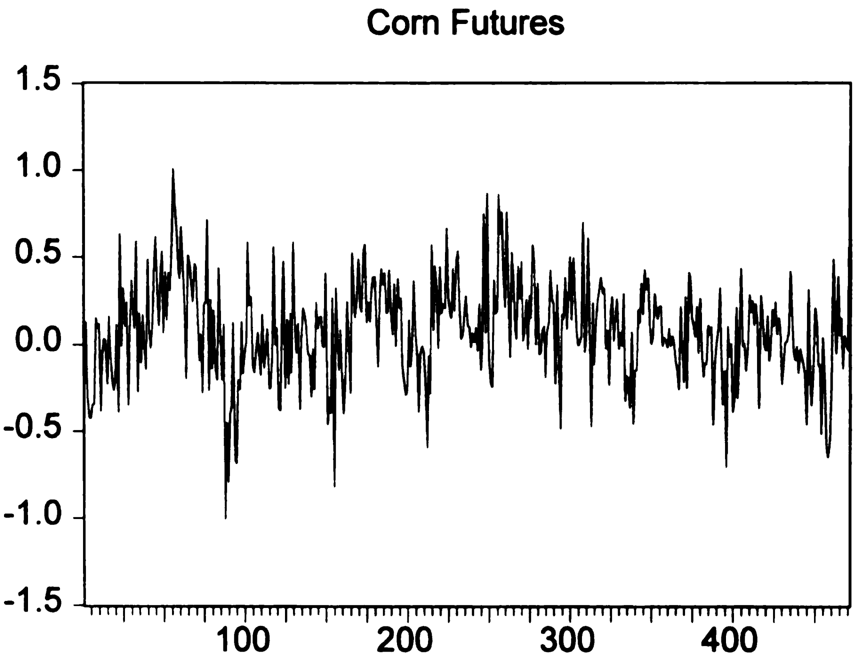
Standardized Gasoline Daily Futures Return
by the Realized Volatility

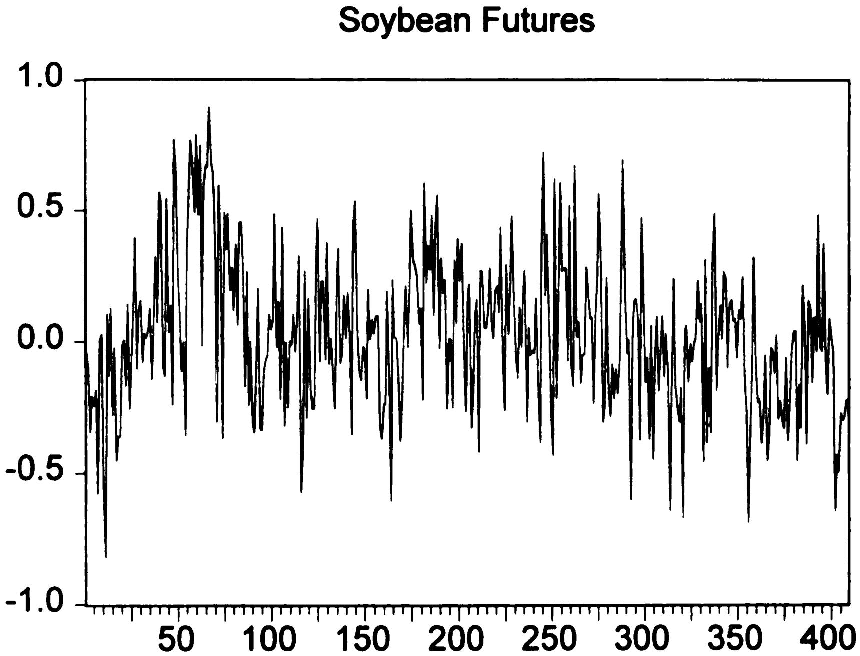
Kernel Density (Epanechnikov, h = 0.4844)

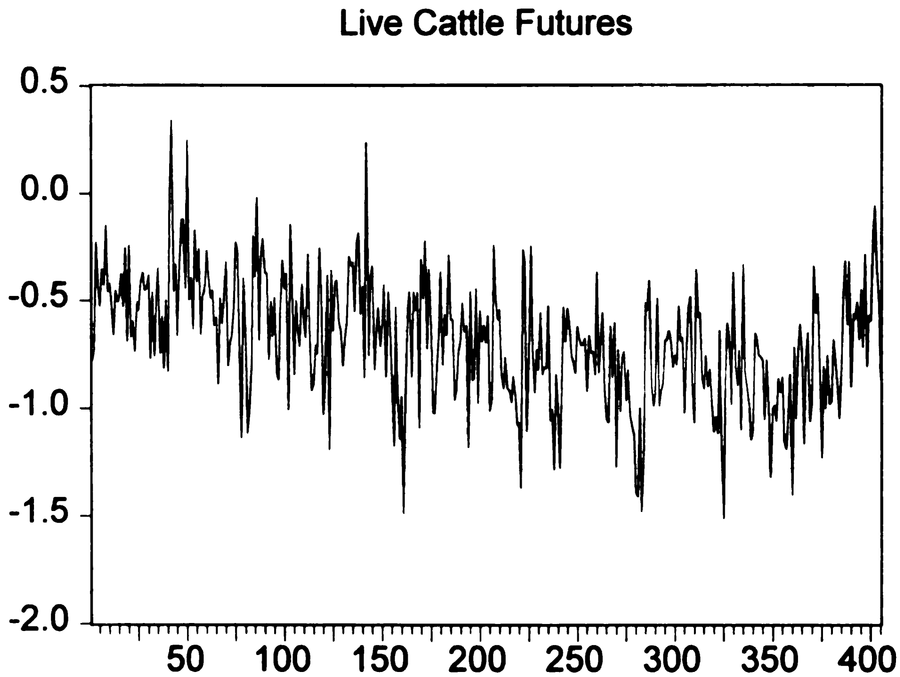


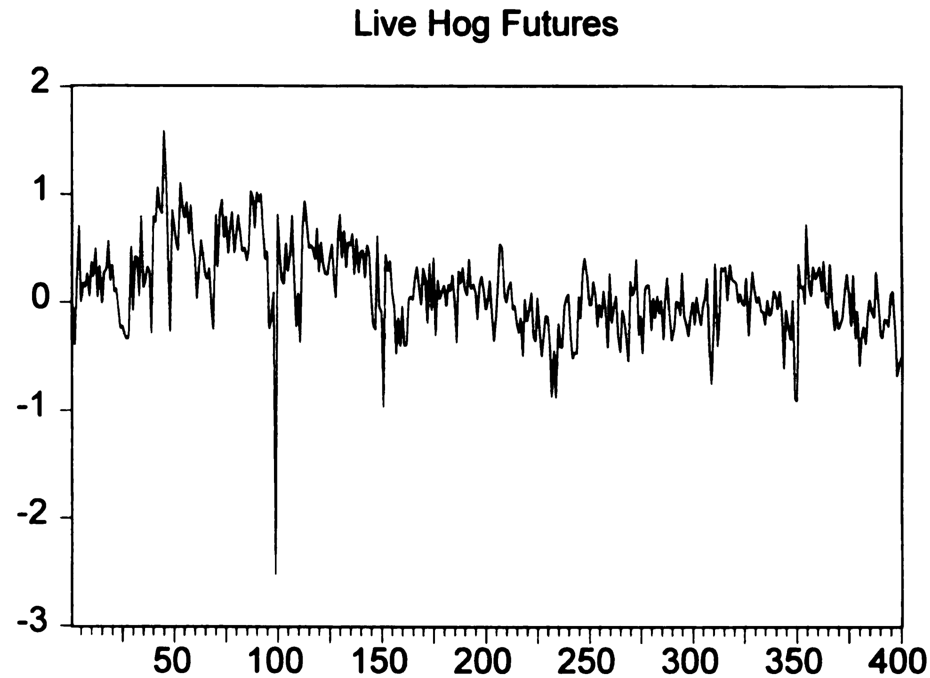
Standardized Gold Daily Futures Return
by the Realized Volatility

Figure 4-2 Realized Commodity Volatility Level

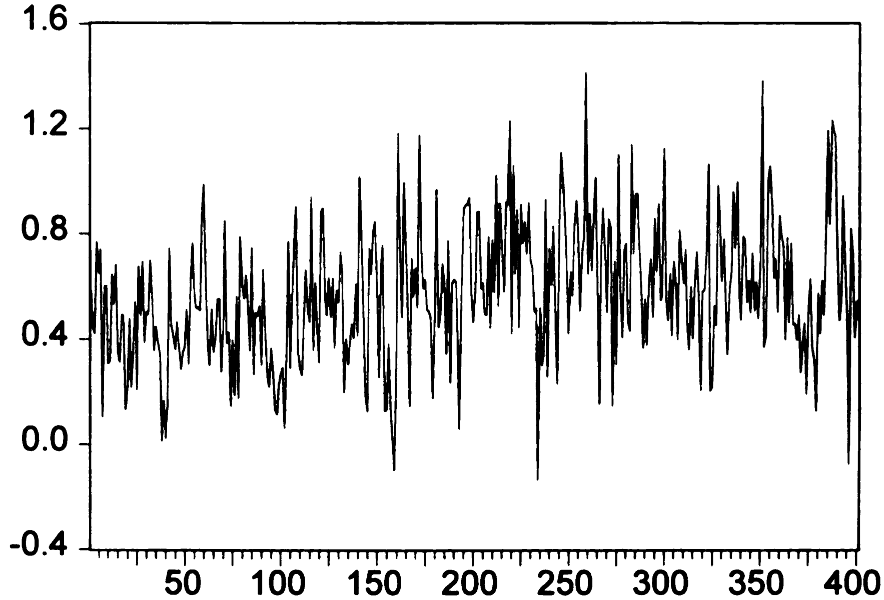




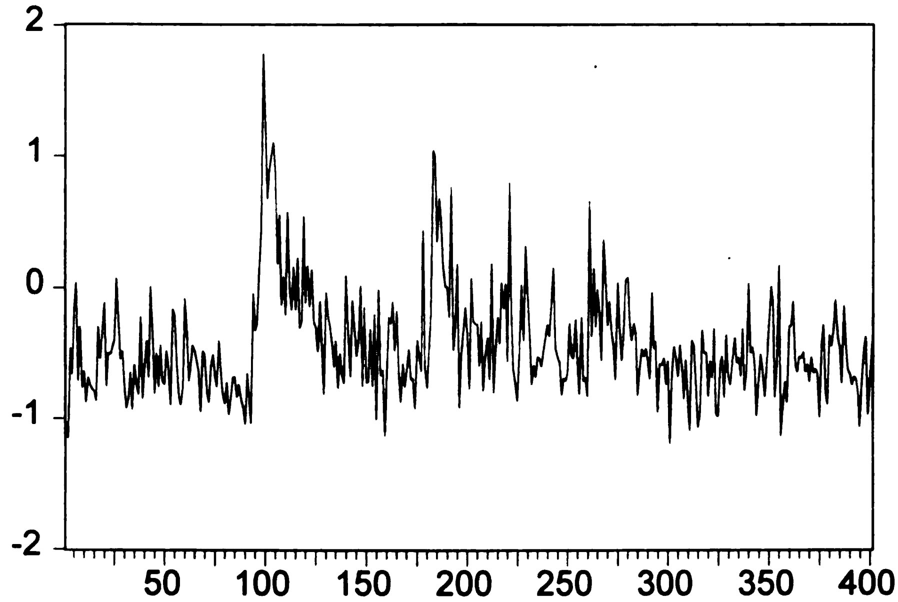

Live Cattle Futures
0.5
0.0
-0.5
-1.0
-1.5
-2.0
50 100 150 200 250 300 350 400


Live Hog Futures
2
1
0
-1
-2
-3
50 100 150 200 250 300 350 400

Gasoline Futures



Gold Futures



CHAPTER 5

CONCLUSION

This dissertation has studied commodity market price risks by using various time series econometric models. The empirical investigation in this dissertation is focused on commodity markets but is extensive in that we analyze the volatility dynamics for 1) daily cash and futures price changes, and 2) higher frequency futures returns. Further, we employed parametric, semi-parametric, and new volatility modeling in our investigations of the commodity return volatility movements. In this final chapter, we list and discuss important factors considered in this dissertation. Possible future research is discussed at the end of this chapter.

First, we found evidence for slowly decaying autocorrelations for daily commodity cash and futures as well as for intra-daily futures return volatility. This observation is consistent with much previous evidence from conventional financial asset return volatility studies. Our findings imply that commodity price risk patterns seem to be similar to financial asset risk behavior, despite some unique characteristics of the commodity markets based on the physical properties of the various types of products. In particular, we observed the long memory phenomena for cash and futures returns at various sample frequencies within the same sample period. This result is consistent with one of the theoretical properties of the long memory process, "self-similarity." We utilized both the parametric FIGARCH model and the semi-parametric local Whittle estimation to identify the long memory return volatility feature. Our findings in chapter 2 and 3 are supportive of long-run temporal dependence in commodity price risks at daily and high frequency sample frequencies.

We used high frequency price data in particular in this dissertation. Tick sample frequency data have become more available recently due to developments in computer technology. Motivated by the mixture-of-distribution hypothesis and empirically well-known volatility persistence, recent and active studies have highlighted high frequency return data to pursue deeper understandings of return volatility patterns. However, the use of high frequency data necessarily involves market microstructure issues to be resolved in order to analyze the intrinsic volatility dynamics. In chapter 3, we applied the Flexible Fourier Form filtering to remove strong intra-day volatility periodicity, one of the market microstructure biases.

Secondly, we applied a newly suggested volatility measure to increasingly available high frequency return data. The so-called "realized volatility" is easy to calculate since the measure is the sum of the squared high frequency returns. This volatility measure can provide important implications of data which are consistent with financial theory for option pricing and derivative modeling. According to formal quadratic variation theory, the realized volatility measure is a consistent estimator for true latent volatility factors. Taking advantage of this observable volatility measure, we can enrich volatility dynamics without relying on the complicated parametric form of traditional GARCH models. In chapter 4, we confirmed that the realized volatility measure can provide some enhanced frameworks for commodity market risk management. We considered important economic determinants such as commodity-specific announcements, time to maturity, information flow, and volatility linkage across markets.

After applying various classes of econometric models at different sample frequencies, we consistently witnessed long memory in the return volatility. The slowly decaying autocorrelations seem to be an intrinsic property of the commodity return volatility.

Since we have uncovered the commodity return volatility for both cash and futures markets, a stage for further risk management modeling is ready. Baillie and Myers (1991) noted that the optimal futures hedge ratio (OHR) is time-varying, and they calculated the OHR using a bivariate GARCH model. Since the optimal hedge ratio is defined to be a ratio of the conditional covariance between cash and futures to the conditional variance of futures, proper modeling of conditional moments is important in calculating the optimal hedge ratio. One possibility for further study is to build a bivariate futures hedge model implementing the long memory property. This idea faces some modeling issues since conditional variance matrices in multivariate contexts may involve additional time-varying components. In previous studies on regular GARCH models, constant conditional correlation matrices were assumed. We could allow for more flexible forms of conditional correlation structures that may yield more implications for the optimal hedge ratio modeling, as Tse and Tsui (2002) assume time-varying correlations for a multivariate GARCH model.

Although it is true that cash and futures prices are not necessarily cointegrated, there may be a possible cointegration relation between their squares. If cointegration exists between cash and futures price volatilities, it will be necessary to include a lagged error correction term in the bivariate FIGARCH model. In particular, Brunetti and Gilbert (2000) considered fractional cointegration in a bivariate FIGARCH set up to study the relationship among volatilities in closely-related oil markets. Cointegration analysis of commodity cash and futures return volatilities seems to create room for improvement of the optimal hedge ratio.

Another possible extension is to use the realized volatility (RV) for daily futures variance and any of the previous methods for computing the covariance between cash and futures price changes. As studied in chapter 4, the realized volatility constructed from high frequency commodity futures price data could provide relatively accurate conditional variances without relying on the parametric form of the GARCH model. Therefore, incorporating realized volatility into the hedge model may afford a simple framework for the optimal hedge ratio calculation. The authors are in the process of producing further work on the issues above.

References:

Admati, A.R. and P. Pfleiderer (1988), "A Theory of Intraday Patterns: Volume and Price Volatility", *Review of Financial Studies*, 1, 3-40.

Admati, A.R. and P. Pfleiderer (1989), "Divide and Conquer: A Theory of Intraday and Day-of-Week Mean Effects", *Review of Financial Studies*, 1, 189-224.

Akgiray, V. and G.G. Booth (1987), "Compound Distribution Models of Stock Returns: An Empirical Comparison", *Journal of Financial Research*, 10, 269-80

Akin, R.M. (2003), "Maturity Effects in Futures Markets: Evidence from Eleven Financial Futures Markets", UC Santa Cruz Working Paper.

Andersen, T.G. (1994), "Stochastic Autoregressive Volatility: A Framework for Volatility Modeling", *Mathematical Finance*, 4, 75-102.

Andersen, T.G. and T. Bollerslev (1997a), "Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns", *Journal of Finance* 52, 975-1005.

Andersen, T.G. and T. Bollerslev, (1997b), "Intraday Periodicity and Volatility Persistence in Financial Markets", *Journal of Empirical Finance*, 4, 115-158.

Andersen, T.G. and T. Bollerslev (1998a), " Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements and Longer Run Dependencies", *Journal of Finance*, 53, 219-265.

Andersen, T.G. and T. Bollerslev (1998b), "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts", *International Economic Review*, 39, 885-905.

Andersen, T.G., T. Bollerslev, and F.X. Diebold (2004), "Parametric and Nonparametric Volatility Measurement", In L.P. Hansen and Y. Ait-Sahalia (eds.), *Handbook of Financial Econometrics*, Amsterdam: Northholand, forthcoming.

Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens (2001a), "The Distribution of Realized Stock Return Volatility", *Journal of Financial Economics*, 61, 43-76.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2001b), "The Distribution of Realized Exchange Rate Volatility", *Journal of American Statistical Association* , 96, 42-55.

Andersen, T G, T Bollerslev, F.X. Diebold and P. Labys (2001c), "Exchange Rate Returns Standardized by Realized Volatility are (Nearly) Gaussian", *Multinational Finance Journal*, 4, 159-179.

Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2003), "Modeling and Forecasting Realized Volatility", *Econometrica*, 71, 529-626.

Andersen, T.G., T. Bollerslev, F.X. Diebold and C. Vega (2002), "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange", *American Economic Review*, 93, 38-62.

Anderson, R. (1985), "Some Determinants of the Volatility of Futures Prices", *Journal of Futures Markets*, 5, 331-348.

Anderson, R.W., and J. Danthine (1983), "The Time Pattern of Hedging and the Volatility of Futures Prices", *Review of Economic Studies* 50, 249-266.

Baillie, R.T. and T. Bollerslev (1989), "The Message in Daily Exchange Rates: A Conditional Variance Tale", *Journal of Business and Economic Statistics*, 7, 297-306.

Baillie, R.T. and T. Bollerslev (1991), "Intra-Day and Inter Market Volatility in Foreign Exchange Rates", *Review of Economic Studies*, 58, 565-585.

Baillie, R.T., Y. Han, and T. Kwon (2002), "Further Long Memory Properties of Inflationary Shocks", *Southern Economic Journal*, 68, 496-510.

Baillie, R.T. and R.J. Myers (1991), "Bivariate GARCH estimation of the optimal commodity futures hedge", *Journal of Applied Econometrics*, 6, 109-124.

Baillie, R T, Bollerslev, T and H-O Mikkelsen (1996), "Fractionally integrated generalized autoregressive conditional heteroskedasticity", *Journal of Econometrics*, 74, 3-30.

Bauwens, L., W.B. Omrane, and P. Giot (2003), "News announcements, market activity and volatility in the Euro-Dollar foreign exchange market ", *Journal of International Money and Finance*, forthcoming

Bhattacharya, R. N., Gupta, V. K., Waymire, E. (1983), "The Hurst effect under trends", *Journal of Applied Probability* 20, 649-662.

Beran, J. (1994), *Statistics for Long Memory Processes*, Chapman and Hall, London.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31, 307-327.

Bollerslev, T and J.M. Wooldridge (1992), "Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances", *Econometric Reviews*, 11, 143-172.

Booth, G.G. and V. Akgiray (1987), "Compound Distribution Models of Stock Returns: An Empirical Comparison.", *Journal of Financial Research*, (Fall 1987), 269-80.

Breidt, F.J., N. Crato, and P.J.F. de Lima (1993), "Modeling Long Memory Stochastic Volatility", *Journal of Econometrics*, 73, 325-334.

Brunetti, C. and Gilbert, C. (2000), "Bivariate FIGARCH and Fractional Cointegration", *Journal of Empirical Finance*, 7, 509-530.

Burton, M. (1993), "Some Illustrations of Chaos in Commodity Markets", *Journal of Agricultural Economics*, 44, 38-49.

Cai, J, Cheung Y-L and M.C.S. Wong (2001), "What moves the gold market?", *Journal of Futures Markets*, 21, 257-278.

Campbell, J.Y. and N.G. Mankiw (1987), "Are Output Fluctuation Transitory?", *Quarterly Journal of Economics*, 102, 857-880.

Cecchetti, S.G, Cumby, R.E. and S. Figlewski (1988), "Estimation of optimal futures hedge", *Review of Economics Statistics*, 70, 623-630.

Change, E.C., Jain, P.C., and P.R. Locke (1995), "S&P 500 Index Futures Volatility and Price Around the NYSE Close", *Journal of Business*, 68, 61-84.

Christensen, B.J. and N.R. Prabhala (1998), "The Relation Between Implied Volatility and Realized Volatility", *Journal of Financial Economics*, 50, 125-150.

Chen, Y.J., J.C. Duan, and M.W. Hung, 1999, "Volatility and Maturity Effects in the Nikkei Index Futures", *Journal of Futures Markets*, 19(8), 895-909.

Clark, P.K. (1973), "A Subordinate Stochastic Process Model with Finite Variance for Speculative Prices", *Econometrica*, 41, 135-155.

Dacorogna, M.M., U.A. Muller, R.J. Nagler, R.B. Olsen and O.V. Pictet (1993), "A geographical model for daily and weekly seasonal volatility in the foreign exchange markets", *Journal of International Money and Finance,* 12, 413-438.

Diebold, F.X. (1988), *Empirical Modeling of Exchange Rate Dynamics*, New York, Springer Verlag.

Ding, Z. and C.W.J. Granger. (1996), "Modeling volatility persistence of speculative returns", *Journal of Econometrics*, 73, 182-215.

Ding, Z., C.W.J. Granger, and R.F. Engle. (1993), "A Long memory property of stock market returns and a new model", *Journal of Empirical* Finance, 83-106.

Drost, F.C. and T.E.Nijman (1993), "Temporal aggregation of GARCH processes", *Econometrica*, 61, 909-927.

Dueker, M. and R. Startz (1998), "Maximum-Likelihood Estimation of Fractional Cointegration With An Application to U.S. and Canadian Bond Rates", *Review of Economics and Statistics*, 83, 420-426.

Ederington, L., Lee, J.H. (1993), "How markets process information: news releases and volatility", *Journal of Finance*, 1160 – 1191.

Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica*, 50, 987-1008.

Engle, R.F., and T. Bollerslev (1986), "Modeling the Persistence of Conditional Variances", *Econometric Reviews*, 5, 1-50.

Fama, E.F. (1965), "The Behavior of Stock Market Prices", *Journal of Business*, 14, 319-338.

Fleming, J., C. Kirby, and B. Ostdiek (1998) "Volatility and Common Information in Stock, Bond, and Money Markets", *Journal of Financial Economics*, 49, 111-137.

Fleischer, P. (1998),"Volatility and Information Linkages Across Markets and Countries", *Australian Journal of Management*, 28, 251-272.

Flemming, J., C. Kirby, and B. Ostdiek (1998), "Information and volatility linkages in the stock, bond, and money markets", *Journal of Financial Economics*, 49, 111-137.

French, K.R., G.W. Schwert, and R.F. Stambaugh (1987), "Expected Stock Returns and Volatility", *Journal of Financial Economics*, 19, 3-29.

Gallant, A.R. (1981), "On the Bias in Flexible Function Forms and an Essentially Unbiased Form: the Fourier Flexible Form", *Journal of Econometrics*, 15, 211-245.

Gallant, A.R. (1982), " Unbiased Determination of Production Technologies", *Journal of Econometrics*, 20, 285-323.

Galloway, T. M., and R.W. Kolb, 1996, "Futures Prices and the Maturity Effect", *Journal of Futures Markets,* 16(7), 809-28.

Goodhart, C. and L. Figliuoli (1992), "The geographical location of the foreign exchange market: a test of the island hypothesis", *Journal of International and Comparative Economics,* 1, 13-28.

Goodhart, C.A.E. and M. O'Hara (1997), "High frequency data in financial markets: issues and applications", *Journal of Empirical Finance*, 4, 73-114.

Gradszteyn, I.S. and I.M. Ryzhnik (1980), "Tables of Integrals, Series, and Products", 4th ed. Academic Press, New York, NY.

Granger, C.W.J. (1980), "Long Memory Relationships and the Aggregation of Dynamic Models", *Journal of Econometrics*, 14, 227-238.

Granger, C.W.J. (1981), "Some Properties of Time Series Data and Their Use in Econometric Model Specification", *Journal of Econometrics*, 16, 121-130.

Granger, C.W.J. and Z. Ding (1996), "Variety of Long Memory Models", *Journal of Econometrics*, 73, 61-77.

Granger, C.W.J. and R.F. Engle (1993), "A Long Memory Property of Stock Market Returns and A New Model", *Journal of Econometrics*, 73, 182-215.

Granger, C.W.J. and P. Newbold (1976), "Forecasting Transformed Series", Journal of the Royal Statistical Society, Series B 38, 189-203.

Han, L., J. Kling, and C. Sell (1999), "Foreign Exchange Futures Volatility: Day-of-the-Week, Intraday, and Maturity Patterns in the Presence of Macroeconomic Announcements", *Journal of Futures Markets*, 19(6), 665-693.

Harris, L. (1986), "A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns", *Journal of Financial Economics*, 16, 99-117.

Havey, A.C. (1998), "Long Memory in Stochastic Volatility", in J. Knight and J. Satchell (eds.), *Forecasting Volatility in Financial Markets*, London, UK, Butterworth-Heinemann.

Hong, H. (2000), "A Model of Returns and Trading in Futures Markets", *Journal of Finance*, 55(2), 959-988.

Hosking, J.R.M. (1981), "Fractional Differencing", *Biometrika*, 68, 165-176.

Hsieh, D.A. (1989a), "Testing for Nonlinear Dependence in Daily Foreign Exchange Rates", *Journal of Business*, 62, 339-368.

Hsieh, D.A. (1989b), " Modeling Heteroscedasticity in the Daily Foreign Exchange Rates", *Journal of Business and Economic Statistics*, 3, 307-317.

Hurvich, C.M, and W.W. Chen (2000), "An Efficient Taper for Potentially Overdifferenced Long Memory Time Series", *Journal of Time Series Analysis*, 21, 155-180.

Hull, J. and A. White (1987), "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance*, 42, 381-400.

Jin, H.J. and D.L. Frechette (20002), "Forecasting Agricultural Cash Price Movements in the Presence of Long-Term Memory", *American Journal of Agricultural Economics*, 86, 432-443

Kodres, L.E. and M. Pritsker (2002), "A rational expectations model of financial contagion", *Journal of Finance*, 57, 769–799.

Kunsch, H.R. (1987), "Statistical Aspects of Self-similar Process", In *Proceedings of the First World Congress of the Bernoulli Society*, vol.1, 67-74.

Lamourex, C.G. and W. D. Lastrapes (1990), "Persistence in Variance, Structural Change, and the GARCH model", *Journal of Business and Economic Statistics*, 8, 225-234.

LeBaron, B. (1999), "Technical Trading Rule Profitability and Foreign Exchange Intervention", *Journal of International Economics*, 49, 125-143.

Lo, A.W. (1991), "Long Terms Memory in Stock Market Prices", *Econometica*, 59, 1279-1313.

Mackey, M.C. (1989), "Commodity Price Fluctuations: Price Dependent Delays and Nonlinearities as Explanatory Factors", *Journal of Economic Theory*, 48, 497-509.

Marinucci, D. and P.M. Robinson (2001), "Semiparametric Fractional Conitegraion Analysis", *Journal of Econometrics*, 73, 303-324.

Martens, M., and J. Luu (2002), "Testing the mixture of distributions hypothesis using "realized" volatility", Erasmus University Working Paper.

Mandelbrot, B.B. (1963), "*The Variation of Certain Speculative Prices*", Journal of Business, 36, 394-419.

McFarland, J W, Pettit, R. Richardson, Sam K Sung (1987), "The Distribution of Foreign Exchange Price Changes: Trading Day Effects and Risk Measurement", *Journal of Finance*, 42, 189-194.

Maddala, G.S. and H. Li (1996), "Bootstrap Based Tests in Financial Models", in G. S. Maddala and C R Rao (eds), Handbook of Statistics, 14, 463-488, Elsevier.

Melino, A. (1994), "The Estimation of Continuous-Time Models in Finance", in C.A. Sims, ed., *Advanced in Econometrics, Sixth World Congress Volume II*, Cambridge University Press.

Merton, R.C. (1980), "On Estimating the Expected Returns on the Market: An Exploratory Investigation", *Journal of Financial Economics*, 8, 323-361.

Milhoj, A. (1987), "A Conditional Variance Model for Daily Observations of an Exchange Rate", *Journal of Business and Economic Statistics*, 5, 99-103.

Milonas, N.T. (1986), "Price Variability and The Maturity Effect in Futures Markets", *Journal of Futures Markets*, 3, 443-460.

Mincer, J. and V. Zarnowiz (1969), "The Evaluation of Economic Forecasts", in J Mincer, ed., *Economic Forecasts and Expectations* (New York: NBER).

Muller, U.A., M.M. Dacorogna, R.B. Olsen, O.V. Pictet, M. Schworz and C. Morgenegg (1990), "Statistical study of foreign exchange rates, empirical evidence of price change law and intraday analysis", *Journal of Banking and Finance*, 14, 1189-1208.

Nelson, D.B. (1992), "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model", *Journal of Econometrics*, 52, 61-90.

Parke, W.R. (1999) "What Is Fractional Integration?", *Review of Economics and Statistics*, 81, 628-632

Porterba, J. and L. Summers (1986), "The Persistence of Volatility and Stock Market Fluctuations", *American Economic Review*, 76, 1124-1141.

Robinson, P.M. (1995), "Gaussian Semiparametric Estimation of Long Range Dependence", *Annals of Statistics*, 23, 1630-1631.

Scott, L. (1987), "Option Pricing when the Variance Changes Randomly: Theory, Estimation, and An Application", *Journal of Financial and Quantative Analysis*, 22, 419-436.

Samuelson, P.A. (1965), "Proof that Properly Anticipated Price Fluctuate Randomly", *Industrial Management Review*, 6, 41-49.

Serletis, A. (1992), "Maturity Effects in Energy Futures", *Energy Economics*, April, 150-157.

Schwert, G.W. (1989), "Why Does Stock Market Volatility Change Over Time?", *Journal of Finance*, 44, 1115-1153.

Tauchen, G.E. and M. Pitts (1983), "The Price Variability-Volume Relationship on Speculative Markets", *Econometrica*, 51, 485-505.

Taqqu, M.S. and V. Teverovsky (1997), "Robustness of Whittle Estimators for Time Series with Long Range Dependence", *Stochastic Models*, 13, 723-757.

Taqqu, M.S. and V. Teverovsky (1998), "On Estimating the Intensity of Long Range Dependence in Finite and Infinite Variance Time Series", 177-217, A Practical Guide to Heavy Tails", (edited by R Adler, R. Friedman and M.S. Taqqu), Birkhauser, Boston, MA.

Tse,Y.K. and Tsui, A.K.C. (2002) "A multivariate GARCH model with time-varying correlations", *Journal of Business and Economic Statistics*, 20, 351 - 362.

Wiggins, M. (1987), "Option Values under Stochastic Volatility: Theory and Empirical Estimates", *Journal of Financial Economics*, 35, 23-45..

Wood, R.A, McInish, T.H, and J.K. Ord (1985), "An Investigation of Transaction Data for NYSE stocks", *Journal of Finance*, 40, 723-741.

Yang, S., and B.W. Brorsen (1992), "Nonlinear Dynamics of Daily Cash Prices", *American Journal of Agricultural Economics*, 74, 706-715.

Zhou, B (1996), "High Frequency Data and Volatility in Foreign Exchange Rates", *Journal of Business and Economic Statistics*, 14, 45-52.

MICHIGAN STATE UNIVERSITY LIBRARIES
3 1293 02504 3476