A WAVELET-BASED ANALYSIS OF COMMODITY FUTURES MARKETS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Gabriel Power

August 2007

A WAVELET-BASED ANALYSIS OF COMMODITY FUTURES MARKETS

Gabriel Power, Ph. D.

Cornell University 2007

The time horizon of decision-making is an essential dimension of economic problems but is difficult to explicitly define. In this thesis, we use time series analysis augmented by wavelet transform methods to precisely identify distinct time horizons in economic data and measure their explanatory power. This enables us to address three timely and persistent questions in the literature on commodity derivatives markets are addressed. First, are findings of long memory (fractional integration) in commodity futures price volatility spurious, following Granger's conjecture? Yes, only two out of eleven commodities are characterized by true long memory and certain stochastic break models (e.g. Markov-switching) are found to be more plausible. Second, do large Index Traders such as commodity pools and pension funds increase futures price volatility through a large volume of trading activity? This appears to be true only for non-storable commodity contracts. Third, can we improve the accuracy of term structure models of futures prices by (i) including more state variables to better capture maturity and inventory effects, and (ii) filtering out what appears to be noise at the shortest time horizons? The results suggest that (i) three state variables is an optimal choice and (ii) estimates using filtered data are not improved and the noise may be economically meaningful.

BIOGRAPHICAL SKETCH

Gabriel John Power was born in Québec City, province of Quebec, Canada. He is the firstborn son of Claire Lapointe and Thomas Michael Power. He and his brothers Lawrence and Patrick grew up in the French Polynesian Islands, in New Zealand, in Gabon and in Canada. He attended Université de Moncton from 1995 to 1997 and Université Laval in Quebec City from 1997 to 2000, earning both Bachelors and Masters Degrees in Economics. He volunteered in Haifa, Israel, from 2001 to 2002 and worked in Switzerland from 2002 to 2003 as a Junior Research Fellow on a project funded by the United States Institute of Peace, Washington D.C. He was admitted to Cornell University on a University Fellowship to earn a Ph.D in Applied Economics, which he completed in July 2007. He married Cheryl Lynn van den Hoonaard in August 2004 and together they have a daughter, Carmel, born November 2005. Beginning in August 2007 he is Assistant Professor at Texas A&M University, College Station, where he teaches and conducts research in the area of futures and options.

To Cheryl

ACKNOWLEDGEMENTS

I thank first and foremost my advisor Calum Turvey, without whose encouragement, support and questions I could not have completed my thesis. He has been the ideal mentor and has taken care to give me many opportunities to mature as a researcher. Working with him has been both inspiring and exciting.

I thank my Special Committee members, Tim Mount and Hazem Daouk. They have provided invaluable help and raised important issues that helped guide my research to fruition.

I thank my family and my wife Cheryl's family for support and encouragement.

I thank the professors in the Graduate Faculty, who provided much advice during my time at Cornell, in particular Jon Conrad, Harry Kaiser, Daniel Simon, David Just, Bill Lesser, Greg Poe, Jeff Prince, Todd Schmit and Bill Tomek.

I thank the diligent and enthusiastic administrative staff at Warren Hall. I am especially grateful to Linda Morehouse, Steve Hatfield, Faye Butts, Anita Vogel, Judy Neno, Marge Arcangeli and Carol Thomson, as well as Janine Brace at the Graduate School.

I thank my Cornell classmates and Ithaca friends, especially Amgad, Shira, the Meads, the Brinn-Beers, Rob & Rachel, Fanja, Kerry & Chris, Kiran & Steve, Moody, Noma, Jordan, Dan, Marc, Travis, Koji, Satoru, Govind, J.V., Koralai, Pin, Jaclyn, Hope, Sudha, Annemie and many more.

I thank the donors whose generous contributions to AEM made it possible for me to study at Cornell and become a university professor.

TABLE OF CONTENTS

Biographical Sketch……………………………………………………..……… iii

Dedication…………………………………………………………..……... iv

Acknowledgements…………………………………………………..…..... v

Table of Contents…………………………………………………...…… vi

List of Tables…………………………………………………………..…… ix

List of Figures………………………………………………………………… xi

1. Introduction………………………………………………………………… 1

2. Wavelets and Time Series……………………………………………..……. 9

2.1 Introduction…………………………………………………………..…… 9

2.2 The Lifting Scheme Approach to Wavelets……………………………… 10

2.3 Desirable Wavelet Properties……………………………………………. 16

2.4 Standard and Translation-Invariant Discrete Wavelet Transforms……… 19

2.5 Wavelets and Long Memory…………………………………………….. 20

2.6 A Simulation Study of Wavelet Transform Reconstruction …………….. 22

2.7 Accuracy of Wavelet Time Series Reconstruction………………………. 23

2.8 Time Series Properties of Wavelet-Decomposed Data………...………… 26

2.9 Application: Variance Ratio Test across Timescales…………………….. 30

2.10 Conclusion………………………………………………………………. 31

3. Is Long Memory in Commodity Futures Data Spurious? ……………………. 32

3.1 Introduction…………………………………………………………….…. 32

3.2 Long Memory in Commodity Futures Prices and Volatility……………… 36

3.3 Commodity Futures Price Data…………………………………………… 38

3.4 The Option Pricing Bias from Long Memory……………………………. 39

3.5 A Log-Range Measure of Volatility……………………………………… 40

3.6 Wavelets Distinguish Short from Long Memory………………………… 43

3.7 Identifying Spurious Long Memory……………………………………… 50

3.8 Semi-parametric Wavelet Estimation of Long Memory………………….. 53

3.9 Parametric Wavelet Estimation of Long Memory………………………… 55

3.10 Exact Wavelet ML Estimates of Long and Short Memory……………… 63

3.11 Testing for Spurious Long Memory…………………………………….. 65

3.12 An Alternative Model of Futures Price Volatility………………………. 71

3.13 Conclusion………………………………………………………………. 73

4. Revealing the Impact of Index Traders on Commodity Futures Markets……… 75

4.1 Introduction………………………………………………………………… 75

4.2 Index Funds and the Commitment of Traders Report……………………... 77

4.3 Futures Market Volatility and the Long-Run Volume of Trade…………… 80

4.4 Data, Estimation and Identification Strategies……………………………... 82

4.5 Endogenously Biased Model and Hausman-Wu Test……………………… 87

4.6 Full Sample Unbiased Structural Model Estimates………………………… 90

4.7 Wavelet-Filtered Sample Unbiased Structural Model Estimates………….. 93

4.8 Estimation Results for Other Commodities……………………………….. 95

4.9 Do Index Traders Increase Futures Price Volatility? ……………………… 96

4.10 The Distribution of Trader Time Horizons………………………………. 99

4.11 Testing for Changes over Time in Trader Heterogeneity………………… 102

4.12 The Distribution of Trade Volume Volatility……………………………. 109

4.13 Conclusion……………………….…………………………………………. 114

5. Estimating the Term Structure of Commodity Futures Prices using Wavelet Thresholding…………………………………………………………………. 123

5.1 Introduction……………………………………………………………….... 123

5.2 The Term Structure of Commodity Futures Prices…………………………124

5.3 Recovering the Net Convenience Yield………….………………………. 130

5.4 Wavelet Thresholding…………………….……………………………. 132

5.5 State-Space Estimation with Wavelet Thresholding……………………… 134

5.6 Estimation Results for One-Factor to Four-Factor Models……………….. 138

5.7 Interpretation and Results of In-Sample Tracking………………………… 141

5.8 Conclusion…………………………………………………..…….. 145

6. Conclusion………………………………………………………………. 147

Appendices……………………………………………………………………… 151

A.1 Data Cleaning……………………..……………………………………… 151

A.2 Additional Estimation and Test Details………………………………….. 152

References………………………………………………………………………. 155

LIST OF FIGURES

2.1 Chicago Board of Trade March 2005 corn futures settlement price………… 28

2.2 Wavelet transformation of corn futures price data…………………………… 28

3.1 Log-range volatility, Chicago Mercantile Exchange lean hogs futures……… 44

3.2 Log-range volatility, Chicago Mercantile Exchange live cattle futures……… 44

3.3 Log-range volatility, Chicago Mercantile Exchange soybeans futures……..… 45

3.4 Log-range volatility, Chicago Mercantile Exchange corn futures…………..… 45

3.5 Log-range volatility, Kansas City Board of Trade wheat futures……….….... 46

3.6 Log-range volatility, Winnipeg Commodity Exchange canola futures……..... 46

3.7 Log-range volatility, New York Board of Trade cocoa futures……………..... 47

3.8 Log-range volatility, New York Board of Trade coffee futures……………..... 47

3.9 Log-range volatility, New York Board of Trade cotton futures……………..... 48

3.10 Log-range volatility, New York Board of Trade sugar no.11 futures……..... 48

3.11 Log-range volatility, New York Board of Trade orange juice futures……….. 49

3.12 Stationarity test results for Hurst parameter, KCBOT wheat futures……..… 57

3.13 Stationarity test results for Hurst parameter, NYBOT sugar futures……..… 58

3.14 Stationarity test results for Hurst parameter, CME hogs futures……...…..… 59

4.1 Daily trade volume, Chicago Board of Trade soybeans futures………………. 85

4.2 Actual and wavelet-filtered daily trade volume, CBOT soybean futures…….. 86

4.3 Andrews-Ploberger-Hansen sup-LM test result plot………….………………… 110

4.4 Distribution of trade volume variance, CME lean hogs futures ……………… 116

4.5 Distribution of trade volume variance, CME live cattle futures ……………… 117

4.6 Distribution of trade volume variance, KCBOT wheat futures ……………… 118

4.7 Distribution of trade volume variance, WCE canola futures …………………. 119

4.8 Distribution of trade volume variance, CBOT corn futures ………………….. 120

4.9 Distribution of trade volume variance, CBOT soybean futures ……………… 121

5.1 Example of term structure of futures prices…………………………………... 124

5.2 One factor model fit with and without wavelet thresholding………………….. 142

5.3 Two factor model fit with and without wavelet thresholding…………………..143

5.4 Three factor model fit with and without wavelet thresholding……………….. 143

5.5 Four factor model fit with and without wavelet thresholding……………..….. 144

LIST OF TABLES

2.1 Numerical accuracy of wavelet reconstruction, stationary process……….….. 25

2.2 Numerical accuracy of wavelet reconstruction, non-stationary process…..…. 26

3.1 Descriptive statistics, log-range price volatility……………………………… 41

3.2 Semi-parametric wavelet based estimation of Hurst parameter……………… 56

3.3 ARFIMA model estimates and hypothesis test results, non-NYBOT futures….66

3.3 ARFIMA model estimates and hypothesis test results, NYBOT futures…..…. 68

4.1 Volume-Price Volatility Model for CBOT corn futures……………………… 94

4.2 Volume-Price Volatility Model for CBOT soybean futures………………….. 97

4.3 Volume-Price Volatility Model for WCE canola futures…………………...… 98

4.4 Volume-Price Volatility Model for CME lean hogs futures………………...… 99

4.5 Volume-Price Volatility Model for CME live cattle futures………………...… 100

4.6 Regression of trade volume over wavelet time horizons, non-NYBOT……… 103

4.7 Regression of trade volume over wavelet time horizons, NYBOT…………… 104

4.8 Monte Carlo wavelet-based test results………………………………………. 108

4.9 Variance of futures trade volume, proportion explained by time horizon……. 112

5.1 Estimation results for one- to four-factor models……………………………. 139

CHAPTER 1

INTRODUCTION

This thesis addresses three timely problems in the literature on commodity derivatives markets. Novel insights with practical implications are provided on the causes and consequences of long memory, the impact of large Index Traders on market volatility, and the shape of the futures price term structure (forward curve). The empirical strategy consists in a combination of established time series analysis with statistical tools derived from wavelet transforms, a recently developed concept that has found widespread use in the physical sciences and in engineering. The following example illustrates what is a wavelet.

Consider the problem of a commodity producer who participates in the futures market because she wishes to hedge her position against price risk. She examines historical data available on futures prices for different maturities. Each time series may be considered individually as a univariate signal contaminated by measurement noise. Theory, however, does not suggest a unique model to explain what the true data generating process might be.

In the absence of a well-motivated structural model, one approach to better understand the data is to find an approximation of the time series using elaborate but deterministic functions. Two well-established methods are sinusoids (functions of sines and cosines) and splines (polynomial knots). In both cases, the idea is that any signal can be approximated by an arbitrarily large sum of deterministic terms. The difficulty is that this sum tends to be prohibitively large, especially in the case of asset prices where singularities are the norm and not the exception. Instead of sinusoids and splines, a better building block would be a deterministic function that is itself shaped somewhat like the data: short, asymmetrical waves containing spikes and cusps. In other words, the ideal building block may be the wavelet.

A second approach to learn from the observed data is to consider the problem of filtering out noise to see the true signal more clearly. This is a special case of the vast class of signal extraction problems that spans several fields of research. In the case of futures prices, the principal difficulty is how to distinguish measurement noise from short-lived but economically meaningful variability. There is clearly a trade-off to be considered, but in general the application of filters has led to results where either the signal is "over-smoothed" or the noise is insufficiently reduced. Here too, research in the natural and experimental sciences suggests that using wavelets to define a filtering criterion may be useful. In those fields, striking advances have been made, but the nature of the data (deterministic, experimental, with controlled measurement error) is different enough from economic and financial data (stochastic, observational, with significant measurement error).

Risk management increasingly relies on the use of sophisticated instruments that provide diverse types of insurance. Market participants face various forms of risk, including price, yield, credit, weather, and income/revenue. Accurate modeling of volatility is vital to the success of commodity markets and by implication, to a light and more efficient regulatory presence. Correct pricing of options, optimal storage and inventory decisions and hedging risk in general all depend on the ability to track and forecast volatility well enough. Although the volume of research on futures markets is large, too much emphasis appears to have been placed on narrow, technical questions, and too little on fundamental, unsolved economic problems. In his Editor's note, Powers (1994) writes: “Deeper insights are needed into the structure, conduct and performance of the industry; the purpose, relevance, costs and benefits of the regulatory structures; the implications of legal decisions and tax and accounting rules on market efficiency; market usage and risk management.”

This dissertation presents three essays on some persistent and timely questions on commodity derivatives with practical implications for market participants. New insights and results are obtained from the empirical analysis of commodity futures and options time series supported by statistical methods based on wavelet transforms. The emerging field of wavelet analysis is well suited to help with the empirical identification of effects and causes specific to particular time horizons of decision-making.

Wavelets are a class of mathematical functions that satisfy specific regularity conditions that make them ideally suited for three broad types of problems: (1) approximating complicated functions by a finite sum of simpler functions (i.e. wavelets), (2) decomposing an empirical time series dataset into asymptotically independent components, and (3) creating an ideal filter that is specific to the data under scrutiny and where the wavelets are tailored to match the data’s characteristics. Wavelet analysis is much more flexible than Fourier analysis and is more economically intuitive. For instance, a timescale decomposition can be interpreted as isolating the different time horizons across which is distributed the variable of interest (e.g. futures prices).

Following the introductory chapter is a short chapter written to cover the essential results from wavelet theory as it provides insights into the statistical methods used.

Chapters 3, 4 and 5 consist of three independent essays on persistent problems on commodity derivatives markets. Few applications of wavelet methods have been made to economics so far. Yet many problems are better understood when the time horizon of decision-makers is explicitly considered.

The essays are connected by their emphasis on the identification of time horizon-specific influences. It is well understood in economics and finance that, for instance, individuals make decisions not only on the basis of immediate costs and benefits but also based on long-run consequences. It is, however, difficult to precisely characterize the different time horizons, ranging from short term to long term.

A helpful, qualitative interpretation of the importance of economic time horizons is provided by Peters (1994) and suggests many testable hypotheses if a precise definition of different time horizons can be given:

- “Markets are stable when they contain investors with large numbers of different time horizons, thus ensuring ample liquidity.”
- “If the validity of fundamental information changes, long-term investors either stop trading or trade on technical factors. However, the market becomes less stable without the long-term horizon investors.”
- “Prices reflect a combination of short-term and long-term valuations, where short-term valuations are more volatile.”
- “If [an asset] has no tie to the economic cycle (e.g. currency), there is no long-term trend, so trading, liquidity, and short-term information dominate.”

Wavelets provide an intuitive, theoretically sound, and computationally tractable framework in which to define and empirically identify different economic time horizons. Time horizons in economics are generally empirically identified as sub-samples or moving windows of a dataset. For example, data may be aggregated as yearly, quarterly, or monthly. In contrast, wavelets enable a simultaneous analysis of how much variation in the data occurs in large number of time horizons or timescales. A simple image is that of a novel made up of several chapters. Suppose someone reads only the introduction and the conclusion. The reader knows how the story begins and ends, but not how events unfold during the intermediate chapters, their occurrence over time and the speed at which events unfold.

Our explicit identification of time horizons is to our knowledge the first such use of the methodology in the literature. It is made possible by decomposing the original data using wavelet transforms (Mallat 1992; Meyer 1992; Daubechies 1993). This thesis provides new, empirically-supported answers to three timely and persistent problems in the literature on commodity futures markets. The methodological contribution of the thesis is the application of wavelet transform-based time series analysis adapted for economics from their original engineering and applied mathematics purposes.

The principal motivation for adapting wavelet methods for economic time series analysis is to enable the identification of the impact of distinct time horizons as explanatory factors driving the unknown stochastic process that underlies observed economic time series such as daily commodity futures prices. In Chapter 2, numerical examples and an intuitive step-by-step construction approach are used to explain the concept of wavelet functions and resulting wavelet transforms. This approach has the advantage of avoiding a discussion of Fourier analysis with no loss of accuracy. After defining the key concepts, it is argued on the basis of clear criteria that a specific class of wavelet functions is best suited for empirical time series analysis. Much of the analysis in the thesis depends on the accuracy of the computational wavelet transform when applied to real data. Therefore, a numerical simulation is presented, where two time series are generated from pre-specified processes (a stationary ARMA process and a non-stationary, long memory process) and decomposed through wavelet transform analysis. It is then shown that the loss of statistical information from the transformation is limited by the software machine precision (double precision in Matlab or R). Lastly, this chapter considers the time series properties of the wavelet-obtained components of the includes an analysis of the properties of a typical commodity futures price time series

Chapters 3, 4 and 5 consist of three essays on timely problems in the literature on commodity derivatives markets. The main research questions asked and answers obtained in this thesis are the following:

- In Chapter 3, we ask whether the literature's findings of long memory (persistence) in futures price volatility are spurious. True long memory may allow arbitrage, undermine the efficiency of futures markets, and induce a substantial bias in the price of options on futures. If they are spurious, is the illusion of persistence caused by short memory, fragile estimators, or the presence of random breaks in the data process? Using a robust estimator in a joint model of both short and long memory effects, we find that long memory estimates are significant and are explained neither by short memory bias nor by the choice of estimator. However, an application of recently-developed tests based on the properties of true long memory shows that for nine out of eleven commodities studied, long memory is spurious. A more plausible model that is fitted to the data is a Markov-switching model.

- In Chapter 4, we test the hypothesis that Index Traders, a class of large investment funds (e.g. pension funds) that has increasingly invested in commodities, have increased price volatility. This widely-held claim has motivated the Commodity Futures Trading Commission (CFTC) in 2007 to begin reporting separately the positions of Index Traders from the positions of large Commercial and Non-Commercial traders in its weekly Commitment of Traders report separately. In the absence of confidential data on trader-level positions, this chapter adopts a "revealed" methodology to evaluate the impact of Index Traders on market volatility. The CFTC's research shows that Index Traders do not engage in short-run trading. We therefore filter out from a dataset on daily futures trading volume all variation occurring at time horizons shorter than one month and use this filtered data in a joint model of trade volume and price volatility. Filtering is enabled by wavelet transform analysis (see Chapter 2). A Hausman-Wu test confirms that volume and volatility are endogenous, so we estimate the joint model by 2SLS using both the original data and the wavelet-filtered data. Comparing the two sets of estimates, the evidence suggests that Index Traders have increased price volatility for non-storable commodities (meats), but not for storable commodities (grains). The chapter's second contribution is to estimate, for all major agricultural commodities and over the time period 1981-2006, the explanatory power of all distinct time horizons on futures trade volume. We find that non-storable commodities generally trade at shorter time horizons than do storable commodities, and also that, perhaps as a result of Index Traders, intermediate and long run time horizons have gained importance in the last five to ten years. Two tests of structural breaks and change-points are used: one wavelet-based Monte Carlo and the other in the Andrews-Ploberger-Hansen sup-Wald class.

- Chapter 5 looks at the problem of forecasting the constellation of futures prices and volatility. To make this problem tractable, we estimate a state space dynamic term structure model using the Kalman filter. This model is explained by a small number of latent factors or state variables and provides computed parameter values for drift, diffusion, mean-reverting speed, risk premia, convenience yield, cost of carry, and seasonality. This chapter considers the ability of two alternative approaches to improve efficiency. The first is to increase the number of state variables (and parameters). The second is to apply, before estimating a parsimonious state space model, the statistical method of wavelet thresholding to pre-filter the data and remove mean zero noise below a threshold that is not arbitrary but rather endogenously determined. If this noise is indeed of no economic significance, the resulting estimates must be both more accurate and more efficient. However, the evidence suggests that what appears to be short-run noise in fact contains information that helps obtain good parameter estimates. The results also suggest that including more than three state variables model does not improve estimation accuracy enough to warrant the greater computational burden.

CHAPTER 2

WAVELETS AND TIME SERIES

2.1 Introduction

This chapter provides a selective review of wavelet theory as it applies to time series analysis. A thorough treatment of wavelet methods in statistics is contained in Ogden (1996), Percival and Walden (2001) and Vidakovic (1998). Seminal contributions include Daubechies (1988, 1992, 1993), Mallat (1998), Meyer (1985, 1993), Strang and Nguyen (1996) and Stromberg (1985).

Two detailed surveys of wavelet methods for economic time series analysis are Crowley (2007) and Gencay, Selcuk and Whitcher (2001). Yet these sources as well as all economics papers introduce wavelets through Fourier analysis and vector spaces (e.g. Luenberger 1969). While these concepts are familiar to economists, they are not commonly used and therefore do not provide a suitably clear introduction to wavelets, particularly since wavelets have been designed in part as an alternative to Fourier analysis. Therefore, wavelets are instead introduced in this chapter based on the lifting scheme method developed by Sweldens (1994). Essential results from the theory of wavelets applied to time series analysis are presented to provide a unifying framework for the three essays in this dissertation.

A simple example illustrates the construction of basic wavelets, following which the main technical conditions are defined and described in the context of empirical time series research. A first empirical application using a variant of the Variance Ratio test is made in this chapter to determine differences across timescales (or time horizons) in the persistence of daily innovations to futures prices. For the interested reader, an outline of wavelet theory results for time series analysis presented using Fourier analysis concepts is included in the Appendix.

Also included is a section of results of simulation-based wavelet analysis done using pre-determined Data Generating Processes (DGP) fully known to the researcher. These simulations consider the analysis of a few stylized, canonical time series models frequently used in economics and finance. The aim of this section is to provide a baseline or benchmark against which to evaluate the results obtained from the analysis of actual data.

Applications of wavelets to economics and finance have been limited so far. In his survey of wavelet methods for economics, Crowley (2007) cites eleven journal articles and ten working papers. Pioneering contributions include Ramsey and Lampart (1998a,b) who investigate the macroeconomic causal relationship between money and income as well as Davidson, Labys and Lesourd (1998), who apply a nonparametric wavelet regression to study volatility at different time horizons in international aggregate monthly commodity prices. A recent example of an economic application of wavelets is Lien and Shrestha (2006), who use wavelet-based methods to compute the optimal hedge ratio by time horizon for several commodity futures markets.

2.2 The Lifting Scheme Approach to Wavelets}

Wavelets are functions that satisfy specific regularity conditions and form a basis (to be precise, a frame) in a vector space (see e.g. Luenberger 1969). Any function in a general class "can be written as a linear combination of the wavelets" (Sweldens 1994). Wavelets have been widely and successfully used in mathematics, engineering and in the natural and physical sciences. The first generation of wavelets (Daubechies

1988, 1992, Mallat 1992, Meyer 1992, Strömberg 1981) relies on a Fourier analysis framework. The mathematical motivation for using the Fourier framework is that wavelet operations become simple algebra in the Fourier domain. Since Fourier analysis is used less frequently by economists than by physicists and engineers, our presentation draws from Sweldens's (1996, 1997) "second generation" wavelet framework which makes no reference to Fourier analysis and is more general and flexible than the earlier approach. To our knowledge, all economics and finance papers have introduced wavelets in the Fourier language.

In addition to making the concepts and their construction more intuitive, the lifting scheme framework provides a more general method of working with wavelets. This means it can be applied to situations where the traditional wavelet approach cannot. Some relevant examples include the construction of wavelet transforms ideally suited to bounded domains, such as intervals (e.g. finite-length time series data) or for application to irregularly sampled data such as ultra-high-frequency tick data. Jensen and la Cour-Harbo (2001) provide a textbook introduction to wavelets based on the lifting scheme.

One particularly useful application of wavelets is to allow us to decompose a signal or time series dataset into explanatory shares attributed to each time horizon. The time horizons are arbitrarily determined but can be interpreted as approximate economic time horizons.

The following example is inspired by Jensen and la Cour-Harbo (2001). Consider a sequence of daily futures settlement prices F_t in U.S. dollars per unit contract:

$$\{60, 66, 72, 64, 68, 70, 74, 70\}$$

Suppose to get closer to the true data generating process, we would like to represent the data in a more efficient form. This is not unlike the engineering problem of optimal data compression. Consider representing the data as a correctly time-localized sequence of means and deviations from means. If done correctly, there will be no statistical loss of information, and the original data sequence can be reconstructed as perfectly as the software level of precision permits.

We believe the time series data are correlated, and correlation should be higher among nearby observations than among distant ones. The goal is to compute a new vector of the same length (that is, eight observations) consisting of four pairwise means and four pairwise deviations from means. We group the observations into four pairs:

$$\{60, 66\}, \{72, 64\}, \{68, 70\}, \{74, 70\}$$

Then we compute the four pairwise means:

$$\{63, 68, 69, 72\}$$

Lastly we compute the pairwise differences (for each pair, this is the odd observation minus the pairwise average):

$$\{-3, 4, -1, 2\}$$

The data are now represented as both a long run mean and time-localized deviations from this mean. If we use a large dataset, we can obtain a large number of levels of deviations-from-means. Each level is associated with a different timescale or time horizon, for example deviations at the daily timescale or at the annual timescale.

This simple example is a trivial wavelet transform, and we would like to find an optimal wavelet transform. Optimality in this case means the wavelet class possesses a number of desirable properties that are determined by whether the wavelet function satisfies specific regularity conditions. A large mathematics literature on wavelets shows how different regularity conditions are derived to ensure a number of properties that are ideal for applications ranging from statistics to physics and engineering. Optimal wavelet properties are described in a later section of this chapter, and it is concluded that the Daubechies (1992, 1993) family of wavelets is best suited overall for typical economic time series data. An important exception is irregularly sampled data such as ultra high frequency tick-by-tick financial data, for which is well suited Sweldens's lifting scheme method for custom-designed wavelets.

Sweldens's lifting scheme begins with a "trivial" wavelet such as the mean and deviations operations, and then "lifting" is applied to produce a better wavelet transform. Stages of lifting allow the transform to be tailor-made for the application and data used. The lifting scheme also nests all traditional wavelet transforms.

It is also possible to set a threshold below which deviations are considered minor and therefore safely deleted. Such a thresholding rule allows us to reconstruct the data using only a subset of the computed differences, and it may be easier to approximate the underlying Data Generating Process (DGP). This procedure is discussed further and applied in Chapter 5.

Lifting involves (a) splitting, (b) predicting, and (c) updating. Consider some data $\lambda_{0,k}$. The first step is to split the data into smaller subsets $\lambda_{(-1,k)}$ and $\gamma_{(-1,k)}$. The convention is that index order reflects the size of the dataset. No restriction is imposed except that some method must exist to reconstruct the original data from the two subsets. The second step, prediction, involves finding a prediction operator *P* that is independent of the data such that we can predict the subset $\gamma_{(-1,k)}$ using the other subset $\lambda_{(-1,k)}$:

$$\gamma_{(-1,k)} = P[\lambda_{(-1,k)}]$$

In the third step, we consider repeating the procedure, and end up with a sequence:

$$\{\lambda_{(-n,k)}, \gamma_{(-1,k)}, \gamma_{(-2,k)}, \ldots, \gamma_{(-n,k)}\}$$

where the first vector $\lambda_{(-n,k)}$ represents the long-run trend of the data, and where the vectors $\{\gamma_{(-1,k)}, \gamma_{(-2,k)}, \ldots, \gamma_{(-n,k)}\}$ each represent variation occurring at a distinct timescale, which in economics is interpreted as a time horizon of decision-making.

Suppose a researcher is working with a time series dataset of a single random variable. The random variable is continuous but recorded at discrete intervals (let's assume for now that intervals are equally spaced). This vector of data could be for example the end-of-the-day settlement price, in dollars per unit contract, of a traded commodity.

The researcher wishes to model the underlying (unknowable) Data Generating Process (DGP) in order to analyze, interpret and forecast. Economic and financial theory suggests candidate structural models for the DGP which usually require obtaining other data as proxies for the explanatory variables. Alternatively, assuming the data are well-behaved (e.g. covariance-stationary), statistical inference is valid and a reduced-form Box-Cox framework can be used instead. This ARIMA model provides estimates of parameters and explains or forecasts the random variable using only information about itself.

Trying to model the unknown DGP is a closely related problem to the challenge of data compression in the engineering literature. If our data are completely random, no data compression is possible because there does not exist a correlation structure to exploit. In economic time series, we would say there is no meaningful DGP, and the data are at least white noise, perhaps IID.

Consider a vector of data *f(t).* Let's denote by *k* a specific sample point, e.g. *k*={1, 2, 3, 4, ...}. We can define our original vector of data as $\lambda_{(0,k)}$ where 0 means it is the original. A very simple, naive approximation is to sub-sample only the even observations, so let's define $\lambda_{(-1,k)} = \lambda_{(0,2k)}$. What have we lost? This vector of errors from the naive approximation is defined as $\gamma_{(-1,k)}$ and these are precisely the wavelet coefficients. The simplest possible wavelet is indeed to let the wavelet coefficients be precisely the odd observations from the original data: $\gamma_{(-1,k)} = \lambda_{(0,2k+1)}$. This means to reach the most efficient representation we want the highest correlation between the initial subsets $\lambda_{(-1,k)}$ and $\gamma_{(-1,k)}$.

Can we predict the odd observations using only the even observations? We can use the fact that in a typical economic or financial time series, correlation is stronger among nearby observations than between distant observations. Consider taking the average of neighboring observations to create a predictor:

$$\lambda_{(-1,2k+1)} = 0.5(\lambda_{(-1,k)+}\ \lambda_{(-1,k+1)})$$

As a result, our wavelet coefficients become:

$$\gamma_{(-1,k)} = \lambda_{(0,2k+1)} - 0.5(\lambda_{(-1,k)} + \lambda_{(-1,k+1)})$$

An iterative procedure is obtained by applying the method first to $\lambda_{(-1,k)}$ which yields $\lambda_{(-2,k)}$, then to the newly obtained $\lambda_{(-2,k)}$ and so on. This approach however leads to a problem called aliasing. Intuitively, this means some variation in the data may be "double-counted." We would like the λ terms to capture low frequencies and the $\gamma_{(-1,k)}$ terms to capture frequency. To avoid aliasing, we impose the condition that the average of the coefficients $\lambda_{(j,k)}$ must be the same equal for each level *j*. It is beyond the scope of this section to provide the mathematical results behind the optimality of specific wavelet functions. Much of the mathematical literature on wavelets concerns this problem, and a seminal collection of papers is found in Daubechies (1993).

2.3 Desirable Wavelet Properties

In this section, we describe the properties that make particular wavelets optimal for a given application as well as trade-offs involved in the selection of an ideal wavelet. In time series analysis, desirable wavelet properties include symmetry, moment preservation, orthogonality between levels of decomposition, perfect reconstruction, correct time alignment (linear/zero phase), minimization of spurious artifacts and boundary effects, and compact support.

To illustrate the usefulness of these properties, we focus on the Daubechies (1988) wavelet class, which the literature has found to be the best for empirical time series work using economic and financial data. We also discuss properties of the original wavelet, discovered by Haar (1910), which is the simplest to construct and also a nested special case of the Daubechies wavelet. A large number of wavelets have been defined but only those of Daubechies and Haar appear to be consistently useful to economists. A thorough treatment of wavelet properties is found in Daubechies (1992, 1993), Ogden (1996) and Vidakovic (1998).

The four key properties for wavelets in time series analysis are:

1. A nonzero number of vanishing moments
2. Compact support
3. Orthogonality and orthonormality
4. Linear phase

To explain the importance of a nonzero number of vanishing moments, we introduce the two principal conditions of a wavelet. First, a wavelet is a function $\psi(\cdot)$ defined on the extended Real line such that the admissibility condition is satisfied:

$$\int_{\mathbb{R}} \psi(t)dt = 0 \tag{2.1}$$

Second, a wavelet is generally required to satisfy the unit energy (variance) condition:

$$\int_{\mathbb{R}} \psi^2(t)dt = 1 \tag{2.2}$$

Then, a greater requirement is for the wavelet to have a number N of vanishing moments such that, for $k = \{0, \ldots, N\text{-}1\}$ the wavelet satisfies:

$$\int_{\mathbb{R}} t^k \psi_0(t)dt \equiv 0 \tag{2.3}$$

A greater number of vanishing moments is particularly important for the wavelet-based analysis of long-range dependence (see Chapter 3), because it provides the long-range parameter estimator with robustness against contamination by nonlinear and potentially non-stationary trends (Teyssiere and Abry 2006). The literature also refers to filters associated with wavelet transforms and the length of a filter is precisely twice its number of vanishing moments. A large number of vanishing moments increases however the size of the wavelet and may generate spurious artifacts in the transformed data. The Daubechies regular and least asymmetrical wavelets among others have an arbitrary number of vanishing moments such that the researcher can select the most appropriate number. In contrast, the simple Haar wavelet has zero vanishing moments as it is piecewise linear.

Compact or finite support captures local variation more accurately. The wavelet oscillates locally and quickly fades away on the left and on the right. In contrast, sines and cosines oscillate indefinitely. The Haar and Daubechies (regular and least asymmetrical) are three of the only four wavelets that are both compactly supported and orthogonal wavelets.

Orthogonality means that for a wavelet timescale representation of the data, the different levels are uncorrelated which implies the perfect reconstruction property holds. Suppose we want to know how much of a time series variance is explained by variation at the short run, medium run, and long run. Orthogonality implies that the perfect reconstruction property holds and therefore enables an accurate deconstruction of a time series into different levels or time horizons. Orthonormality further ensures unit energy (variance), which means the decomposed data remains accurate to scale. Both the Daubechies and Haar wavelets are orthonormal.

Linear phase ensures correct time localization. For example, we may wish to determine the precise date of a mean or variance change-point in a time series. Linear phase is also a necessary and sufficient condition for perfect symmetry, a property that only the Haar wavelet possesses. Since excessive asymmetry is undesirable, Daubechies developed a Least Asymmetrical wavelet that has essentially correct time localization and is therefore often used in economic applications.

As with nonparametric regression and frequency domain analysis, wavelet analysis involves dealing with the problem of boundary effects. The theory behind wavelets has been developed under the assumption of an infinite number of observations, but sampled data in economics and other non-experimental sciences are necessarily finite. If no correction is made, the computed wavelet coefficients will be overstated at the beginning and end of the sample. Two general solution methods are, first, to discard those biased observations by truncating the sample a few observations after the beginning and before the end and, second, to artificially extend the time series for purposes of wavelet analysis but only include the true observations in the economic analysis and interpretation of results. The time series can be extend by padding with zeros, reflecting (symmetrically) the observations at the sample's endpoints, or assuming the sample repeats periodically. Cohen et al. (1993) have found that zero-padding creates large artifacts in the data and reflecting the data causes the orthonormality property to be lost. Periodization is therefore the least harmful method unless the researcher can afford to discard some observations at both endpoints.

2.4 Standard and Translation-Invariant Discrete Wavelet Transforms

To obtain a frequency domain representation of time series data suitable for spectral analysis, the Fourier transform is applied to the data (see e.g. Hamilton 1994). The workhorse of wavelet-based time series analysis is the Discrete Wavelet Transform (DWT). Unlike the Fourier transform, which is unique, wavelet transforms are numerous because each one is constructed from a specific wavelet function and filter length. For all wavelets, the resulting Discrete Wavelet Transform is the inner product (convolution) of the data with translations and dilations of the wavelet function. The outcome is a wavelet coefficient vector of the same length as the original data. The wavelet coefficients contain information in both the time and scale domain, where the scale corresponds to different length time periods. For example, if the original data are daily observations, then the scales would include daily, weekly, monthly and so forth. Assuming the property of orthonormality holds,

In this thesis, data are sampled daily over a period of two decades. This means the wavelet transform requirement of a sample of dyadic length (base two) is not overly restrictive. Many economic datasets, however, consist of much shorter time series where each observation matters. This is the case, for example, with many macroeconomic time series. This transform, also called the maximum overlap discrete wavelet transform, may be applied to data of any length. The downside is that it loses the orthonormality property, which implies a loss of efficiency and a more conservative interpretation of the results.

The second reason to use the translation-invariant wavelet transform is that, as implied by its name, its localization in time remains accurate, whereas the basic discrete wavelet transform has a small bias. For instance, after it is found that there exist in the data one or more change-points or structural breaks, the translation-invariant transform should be used to actually date the change-point or break.

2.5 Wavelets and Long Memory

In this section, wavelets are discussed in the context of the most frequently used time series models. The conventional framework for time series analysis in economics is the autoregressive moving average (ARMA) representation of the data. Using this model, the time series data under scrutiny is described as a function of its own weighted lags as well as weighted lags of the innovation (error) term, which is assumed to be at least mean zero white noise (uncorrelated) and possibly identically and independently distributed (IID). The autoregressive and moving average terms are considered "short memory" because their effect on innovations is short-lived and the autocorrelation function and impulse response function decay geometrically (exponentially). Likewise, plain and generalized autoregressive conditional heteroskedasticity (ARCH and GARCH) models are designed to capture simple nonlinear dynamics in the volatility of the time series data and describe well the volatility clustering stylized fact observed in a large number of economic and financial time series data.

In contrast, “long memory” (usually called long-range dependence in the statistics literature) implies a slow, hyperbolic decay in the autocorrelation function and in the impulse response function, which means the effect of shocks or innovations on the data is long-lived. This concept originates with Hurst’s (1951) seminal Rescaled Range analysis (R/S) and the mathematics literature on fractals applied to time series data by Mandelbrot (1963) and Mandelbrot and van Ness (1968). A well known and extensively studied special case of long memory in economics is permanent memory, equivalently the unit root (Phillips 1987; Perron and Phillips 1988). In the ARMA framework, a unit root in the autoregressive lag polynomial implies that innovations have a permanent effect on the data process and results in non-stationarity. Generally, by non-stationarity is meant covariance-non-stationarity, such that the variance/covariance is time-dependent. A stronger definition of non-stationarity that is however not testable considers all existing moments of the data generating process to be time-homogeneous. A non-stationary time series process is said to be integrated of order one, or I(1), and can be modeled as Autoregressive Integrated Moving Average (ARIMA), while the stationary case is defined as I(0). Greater orders of integration are possible but rarely found in economics.

Fractional orders of integration, defined as $d \in (-1, 1)$, have been suggested by Granger (1980) and Granger and Joyeux (1981) to provide a link between the Hurst coefficient of long memory and the conventional time series ARMA and GARCH models. For $d \in (-1, 1)$, $H=0.5+d/2$. The general extensions are called Autoregressive Fractionally Integrated Moving Average (ARFIMA) and Fractionally Integrated General Autoregressive Conditional Heteroskedasticity (FIGARCH, e.g. Bollerslev and Mikkelsen 1996). Hosking (1981, 1984) provides formal results on the fractional difference operator d and conditions for stationarity and invertibility. Tanaka (1999) contributes important refinements on the fractional unit root and several of his results are used in this thesis. Baillie (1996) provides an early survey of results on long memory models in economics, but this is an active area of both theoretical and applied research.

Wavelets can be used to represent the original data in the timescale domain based on some objective criterion. The wavelet property of orthogonality between timescales implies that a self-similar pattern such as a fractal signature (Mandelbrot 1963) should be evident across timescales if the data are characterized by true long-range dependence (long memory). In addition to enabling a graphical or visual test of long-range dependence, wavelets are ideally suited to construct a variety of estimators and tests. Examples include parametric estimators (Jensen 2000), semi-parametric estimators (Teyssiere and Abry 2006), tests for intractable serial correlation (Hong and Lee 2005) and tests for multivariate higher order moment dependence (Duchesne 2006).

2.6 A Simulation Study of Wavelet Transform Reconstruction

Wavelets make it possible to decompose a data signal, stochastic process or function into additively orthogonal levels (or timescales in the wavelet time series literature). When applied to economic time series data, an intuitive interpretation can be made. Each level is a time horizon to which is associated a proportion of the variation in the data. In a rural economic setting, time horizons may have a more immediate geographic interpretation: long-run horizons imply national, macroeconomic causal forces, medium-run horizons regional forces and short-run horizons local forces. This method makes it possible to explicitly identify distinct time horizons and investigate economic hypotheses that concern the incidence of effects across time horizons of decision-making or across different depths of underlying economic forces.

To verify the accuracy of the numerical wavelet transforms used to decompose the data, we simulate some time series data consistent with two plausible futures prices data generating processes, namely a stationary ARMA(2,2) and a non-stationary, long memory fractional Brownian Motion with Hurst parameter of 0.75. Application of a discrete wavelet transform produces wavelet coefficients, which is a representation of the data in the wavelet time-scale domain. Applying an inverse wavelet transform to subsets of the wavelet coefficients results in a perfect decomposition of the original data into several orthogonal time series, each of which has the same length as the original time series and which can be simply added to yield the original time series. These artificial time series vectors cannot be used as regressors to explain the original time series data because the perfect reconstruction property implies by definition that all explanatory variable coefficients must equal one.

The original data is compared to the reconstructed data and we compute the approximation error caused by transforming the data back and forth. The loss function used are is the root mean squared (approximation) error and we also consider as criterion the first four sample moments of the distribution of approximation errors.

2.7 Accuracy of Wavelet Time Series Reconstruction

This section presents the results of a simulation study on the accuracy of the wavelet transform to decompose and reconstruct time series data with no loss of information. Two samples of data are generated from a pre-determined process, decomposed into timescale wavelet coefficients using a discrete wavelet transform, and finally the original data is recovered using the inverse of the discrete wavelet transform previously used. To guarantee the existence of an inverse, an orthonormal wavelet must be selected to construct the transform, therefore we use the Daubechies wavelet. As explained earlier, we may choose an arbitrary number of vanishing moments for this wavelet, which results in a specific filter length. We experiment with filter lengths ranging from 2 to 20 and find that the length 8 or 10 appears best.

The first simulated data generating process is a linear Autoregressive Moving Average model with two lags of each type, i.e. ARMA (2,2), with an intercept of 100 and no deterministic or stochastic trend (no unit root). In this model, the dependent variable "today" is explained by its own two most recent lags as well as an innovation term and the innovation's two most recent lags. The serial correlation has a "short memory" and the persistence of shocks is short-lived. The number of observations used is T=512 observations, with 712 observations generated and the first 200 dropped, which is called the "burn in" stage. The Auto-Regressive and Moving Average parameters are $\phi = (0.6, -0.3)$ and $\theta = (0.4, 0.2)$.

Using simulated data with IID Normal innovations and a Daubechies wavelet, which has the orthonormality property, we expect to find that the first four moments of the distribution of approximation errors are Gaussian Normal. The loss function selected is the root mean squared error. It is the square root of the average, over all T observations, of all squared approximation errors, defined as the reconstructed data point minus the true data point, for all T observations.

The results, summarized in Table 2.1, suggest that all wavelet reconstructions are unbiased and the approximation errors are close to Gaussian Normal as desired (skewness=0, kurtosis=3).

Table 2.1: Numerical accuracy of wavelet reconstruction for ARMA(2,2) process using Daubechies wavelet with filter length 2 to 16

Wavelet	**Root mean squared error**	**Approximation error mean**	**Error SD**	**Error skewness**	**Error kurtosis**
dau2	2.048e-07	-3.095e-13	3.153e-14	0.0219	2.4881
dau4	1.9233e-07	-2.054e-13	3.937e-14	0.3850	2.8761
dau6	2.2280e-07	-3.234e-13	5.25e-14	0.1242	2.7510
dau8	2.0728e-07	7.234e-13	6.921e-14	0.6317	2.509
dau10	3.551e-13	3.799e-13	6.383e-14	0.4599	2.8303
dau12	2.0567e-07	4.428e-13	7.225e-14	0.0841	2.5783
dau16	1.9750e-07	3.404-13	7.098e-14	-0.2651	2.8467

The second simulated process consists of fractional Brownian motion with a Hurst long memory coefficient of 0.75. It is a non-stationary, persistent (long memory) process with innovations that are distributed not IID Normal or as white noise but rather as fractional white noise. Fractional white noise increments over time are stationary but not independent of each other.

A total of 712 time series observations are generated from a fractional Brownian motion process with a starting value of 100. The first 200 observations are discarded as a “burn-in” stage. Observations 201 to 712 inclusive are saved for a total of 512 data points. Again, the Daubechies wavelet is used with different filter lengths.

The results shown in Table 2.2 suggest that the Daubechies-based wavelet transform for any filter length will provide outstanding reconstruction with only a trivial loss of statistical information, even for a challenging process such as non-stationary fractional Brownian motion.

Table 2.2: Numerical accuracy of wavelet reconstruction for fractional Brownian motion process using Daubechies wavelet with filter length 2 to 16

Wavelet	**Root mean squared error**	**Approximation error mean**	**Error SD**	**Error skewness**	**Error kurtosis**
dau2	1.4149e-14	-2.002e-14	4.063e-15	-0.2694	3.1030
dau4	9.1089e-15	-8.269e-15	4.608e-15	-0.3150	3.0254
dau6	1.2539e-14	-3.122e-14	7.112e-15	-0.2656	3.1984
dau8	4.0599e-14	7.596e-14	1.377e-14	-0.4040	2.8241
dau10	1.8447e-14	-4.182e-14	7.940e-15	-0.3954	2.7796
dau12	2.5605e-14	-3.915e-14	9.823e-15	-0.4953	2.4945
dau 16	2.0923e-14	-4.185e-14	1.655e-14	-0.0489	2.3809

2.8 Time Series Properties of Wavelet-Decomposed Data

In the previous section it was found that applying a wavelet transform to time series data does not cause a loss of statistical information beyond machine precision. However, to conduct meaningful hypothesis testing of economic models using wavelet-transformed data, we need to verify whether the stationarity of data is preserved. For example, suppose we extract from a stationary time series dataset several timescale levels. Will any of these levels be non-stationary and therefore at risk of leading to spurious regressions in the Granger-Newbold (1974) sense? Also, if the original data are non-stationary, do the wavelet-computed levels inherit this property? To answer these questions, we analyze in this section a typical futures contract price time series dataset before and after wavelet decomposition.

Consider the price of the CBOT corn futures contract expiring in March 2005. This contract begins trading on 26 June 2003 and stops trading on 14 March 2005, for a total of 440 business daily observations. Figure 2.1 shows the daily price of this contract over the entire time period. It is customary in the research literature to exclude observations from the contract's own expiry month (here the last ten observations). To focus on the period of most active trading, 256 observations are used, dated from 23 February 2004 to 28 February 2005. An Augmented Dickey-Fuller test (computed using one to eight lags) suggests the null hypothesis of a unit root cannot be rejected, whether or not a deterministic time trend is included. The test procedure and optimal lag length selection follow Ng and Perron (2001) and Elliott, Rothemberg and Stock (1996).

Applying a discrete wavelet transform to the data produces wavelet coefficients that allow us to construct several orthogonal, nearly independent time series, each of which corresponds to a distinct time horizon, from daily variation occurring in the data to long-term (here semestrial). Figure 2.2 illustrates each of the artificial time series. Adding together the artificial time series results in the original time series data.

Augmented Dickey-Fuller test results suggest that the price components associated with the daily time horizon and with time horizons of one month and longer are stationary, but that the price components of time horizons greater than a day and less than a month are non-stationary. Therefore, non-stationarity in the original data translates into non-stationarity in some but not all wavelet-computed artificial time series. Stationarity in the original data implies stationarity in the wavelet-computed series.

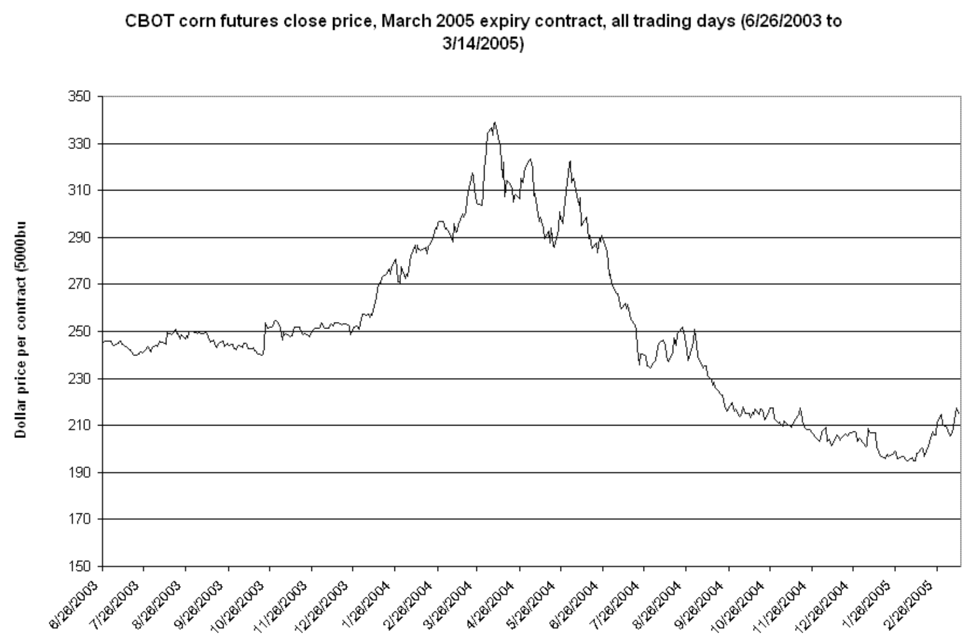


Figure 2.1: Chicago Board of Trade March 2005 corn futures settlement price, 6/26/03 to 03/14/05

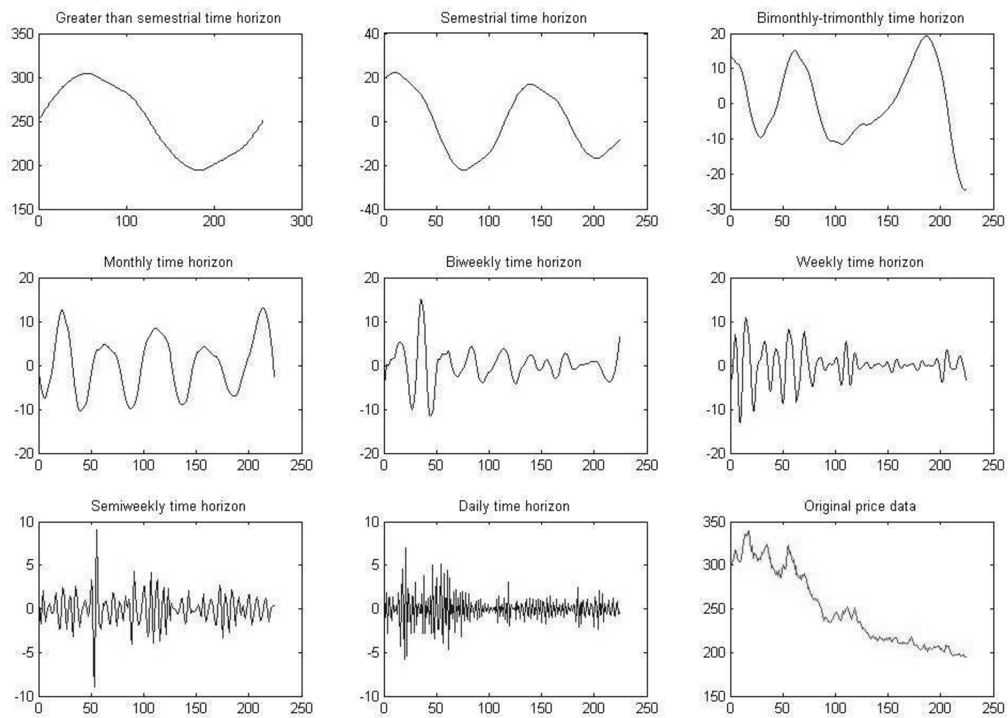


Figure 2.2: Wavelet transformation of corn futures price data into orthogonal, additive, time horizon-specific time series, $price/contract

While testing for serial correlation is relatively simple, evaluating serial dependence in higher order moments is difficult and an area of active research. A number of nonparametric tests exist, but these tend to have low power (Hong 2004).

Consider a economic time series process and suppose there exist opposing economic influences at different time horizons that result in the appearance of a constant variance ratio. This result suggests a random walk. For example, Turvey (2007) finds that for medium- to long-run samples, the null of a random walk in prices cannot be rejected for all but two agricultural commodities.

To further illustrate the meaning of wavelet-estimated timescales (time horizons in an economic setting), a test of the random walk hypothesis is performed on each timescale data series to answer the question: is the random walk result explained by opposing persistent/antipersistent forces at different horizons?

The data used consist of the daily settlement price for the Chicago Mercantile Exchange live cattle futures contract over the time period 2/1989 to 12/2004 inclusive. A total of 4096 observations are used. The Variance Ratio test used is Kim's (2006) wild bootstrap test which has been shown to possess generally superior size and power properties, and the holding periods used are {2, 4, 6, 8, 10, 12, 16, 32, 64} days. The holding period is the subsample used to compute a variance estimate and which is compared to the variance as computed normally. The results suggest the following interpretation. Daily and semiweekly time horizon variation are strongly mean-reverting (antipersistent). Weekly and biweekly variation are persistent for holding periods of up to two weeks, but mean-reverting for longer holding periods. Longer time horizons are persistent for any holding period.

2.10 Conclusion

In this chapter we presented a introduction to wavelets in a time series context using the lifting scheme framework developed by Sweldens (1994), which, unlike other approaches to wavelets, does not require using concepts from Fourier analysis. A number of important wavelet properties were defined and illustrated using the two most commonly used wavelet functions in time series analysis, the Haar and Daubechies. We also provided simulation-based empirical evidence that wavelet-based data transformations of typical economic and financial time series do not cause loss of information and do not induce non-stationarity.

CHAPTER 3

IS LONG MEMORY IN COMMODITY FUTURES DATA SPURIOUS?

3.1 Introduction

In this chapter, we consider the large research literature that claims to have identified significant estimates of long memory in commodity futures prices and price volatility. The implication is that the modeling assumption of geometric Brownian motion should be abandoned in favor of substantially more complicated fractional Brownian motion models. It also implies that options on commodity futures are likely to be severely mispriced. This chapter asks whether findings of long memory are spurious and can be explained by inconsistent and inefficient estimation procedures and by the presence of structural breaks or level shifts. Several steps are taken to make the results more robust. A less noisy measure of volatility is computed from the log-range of prices instead of the traditional price log-returns. The wavelet-based likelihood estimator is preferable to previously used GPH and FIGARCH methods on the basis of consistency, efficiency and coefficient interpretation. The wavelet MLE is also capable of distinguishing short memory effects from long memory, which otherwise would bias the results. It is argued based on this new evidence that in the case of agricultural commodities, long memory is most likely an artifact of the data. Implications for option pricing are that the Black-Scholes solution, adjusted for seasonality and major structural breaks, remains applicable. Semiparametric wavelet estimators of long memory are also presented and applied, but it is argued that these are of limited usefulness to economists because neither analytical nor bootstrap standard errors/confidence intervals are reliable.

The price of a financial option is frequently quoted in terms of its implied volatility. This is an unobserved parameter that solves, for a going price and a set of observable characteristics, the famous and widely-used Black-Scholes-Merton formula (Black and Scholes 1973; Merton 1973). The measurement of volatility remains an active and diverse area of research in both academia and industry. A central concern is whether volatility rapidly or slowly recovers from shocks that affect its magnitude. The main contribution of this chapter is to provide, to the best of the author's knowledge, the first systematic and informative test of spurious long memory in commodity futures price volatility data. The results presented in this chapter contribute to an active and growing literature in agricultural economics on the relationship between commodity futures and options through improved models of price volatility and measures of serial dependence.

In commodity markets, options are written on futures contracts. A number of papers found futures prices to be persistent, a finding that appeared to challenge the efficiency of commodity futures markets (e.g. Corazza, Malliaris, and Nardelli 1997). More recent work suggests however that persistence (long memory) in commodity prices is better explained by a combination of level shifts (a one-time increase or decrease in the mean of the process) in the data and long memory in the volatility of futures (Tomek 1994; Wei and Leuthold 2000; Smith 2005). As a result, the question of long memory in prices has been settled and the literature now focuses on whether price volatility is characterized by long memory. How does long memory in volatility affect the underlying asset price? Modern asset pricing models in the tradition of Black and Scholes (1973) consider that price is a function of a deterministic drift term (trend), a stochastic or random diffusion term (volatility) and possibly a stochastic jump process that may help explain level shifts and structural breaks. Long memory in volatility implies dependence between increments of the diffusion term and therefore has an impact on the price path over time.

A large estimate of long memory in futures price volatility also implies a potentially large bias in the classic Black-Scholes option pricing method. Option pricing based on the Black-Scholes model assumes that the underlying asset (here, the commodity futures contract) is reasonably well described as Geometric Brownian Motion (GBM), which means the natural logarithm of the asset price behaves in the continuous-time limit as an IID Normal random walk with drift. Long memory in volatility implies that the correct option pricing solution is based on fractional rather than geometric Brownian motion (Rogers 1997; Sottinen 2001). Such an option pricing model is substantially more difficult to use, which may further discourage the adoption and use of options in the agribusiness sector.

This chapter therefore addresses one set of causes and consequences of option pricing bias in commodity markets, namely long memory in futures price volatility. The principal aim of this work is to determine whether empirical findings of long memory in commodity futures prices and volatility are spurious. Alternative explanations are considered including the effect of correlated short memory dynamics (generally measured as ARMA parameters) and the presence of structural breaks or level shifts in the data (Smith, 2005; Banerjee and Urga, 2005; Perron, 2006).

The main finding of this chapter is that apparent long memory in commodity futures price volatility is only true for two out of eleven commodities, but is not caused by the effect of short memory dynamics. Rather, the data would be better described by a

Markov-switching or stochastic break model, either of which could generate spurious long memory.

The chapter takes the following steps to answer the question. A measure of volatility is constructed using the daily price range following Alizadeh, Brandt and Diebold (2002). While less accurate than the realized volatility computed from intra-day high-frequency tick data, this measure has been found to be asymptotically superior to the traditionally used volatility measures, absolute or squared logreturns. This volatility proxy is justified by the use of more than 4000 observations for each commodity and the difficulty and cost of obtaining reliable tick data for most agricultural commodity futures. To estimate the long memory parameter d in the canonical fractionally integrated time series model (ARFIMA), a wavelet-based estimator is used (McCoy and Walden 1996; Jensen 2000).

Wavelets are ideally suited to distinguish short from long memory and also to detect the fractal signature of long memory because, as explained in Chapter 2, they are self-similar across time-scales or time horizons and their orthonormality property ensures zero correlation between time-scales. As a result, the wavelet-based estimator is consistent, efficient in its class, and unbiased by the presence of short memory dynamics, unlike for example the frequently-used Geweke-Porter-Hudak (GPH, 1983) estimator. The GPH estimator conveniently requires only an OLS linear regression in the frequency domain, but has been found to be inconsistent, inefficient and biased (Agiakloglou, Newbold and Wohar 1992; Robinson 1995; Smith 2005). The wavelet-based estimate of long memory can be directly interpreted and tested in the standard ARFIMA framework. For $d<0.5$, the process is stationary and the most natural null hypothesis, tested using e.g. Tanaka's (1999) Wald statistic, is then $d=0$ or equivalently white noise innovations (increments) against fractional white noise innovations and $d>0$. Standard errors are computed from Tanaka's (1999) analytical covariance formula that incorporates both the short memory and long memory Information Matrices as well as cross-dependencies. Previous results in the literature appear to generally not account for these cross-dependencies and as a result the standard errors are understated.

Model robustness checks include a separate estimation using only Wednesday observations (i.e. weekly sampling) to account for "day of the week effects" as well as estimates from different wavelet-based long memory estimators. Simple Likelihood Ratio tests are computed to evaluate whether the long memory parameter is significant and the results are contrasted with the evidence from Wald and modified KPSS and Phillips-Perron tests that are designed to consider the presence of spurious long memory. Semi-parametric wavelet-based long memory estimators in the tradition of the Hurst-Mandelbrot R/S analysis are considered, but recent work suggests that for the Hurst long memory parameter neither bootstrap nor Monte Carlo standard errors and confidence intervals are reliable. Weak evidence of long memory is found but it is not possible to confidently test the null hypothesis in this case.

3.2 Long Memory in Commodity Futures Prices and Volatility

Understanding the behavior of futures prices is central to commodity risk management (Tomek 1997; Tomek and Peterson 2001). Futures prices influence hedging and inventory decisions, spot price discovery, and the use of commodity options written on futures. An important question, which motivated the unit root literature in econometrics and particularly in empirical macroeconomics is whether the influence of economic shocks or innovations is short-lived or permanent (Nelson and Plosser

1982; Phillips 1987; Phillips and Perron 1988). It is now well-established that agricultural commodity price time series are unlikely to contain a unit root (Wang and Tomek 2007). This conclusion is supported both by theoretical work (Deaton and Larocque 1992; Tomek, 1994) and by the econometric literature on the low power of unit root tests in the presence of either structural breaks or long memory (e.g. Cochrane 1987).

The concept of long memory, originally given an economic definition by Granger (1980) and Granger and Joyeux (1981), considers that shocks may be so persistent that they are in short time series observationally equivalent with shocks from a unit root process. Moreover, the spurious regression result of Newbold and Granger (1974) is likely to hold for stationary processes with long memory (Tsay and Chung 2000). This means it is not sufficient to verify only stationarity of two time series for which a dynamic economic relationship is being considered. Long memory in time series is characterized by a hyperbolic (slow) rate of decay in the autocorrelation and impulse response functions, instead of the usual geometric (faster) rate of decay. In the standard ARFIMA time series framework, a long memory process is defined as I(d), or fractionally integrated of order $d \in (-1,1)$. The case $d=1$ is the well-known case of a unit root and permanent memory.

A large and active literature suggests that long memory or persistence in commodity futures price volatility is significant and of practical consequence (Baillie et al. 2007; Corazza, Malliaris and Nardelli 1998; Crato and Ray 2000; Cromwell, Labys and Kouassi 2000; Elder and Jin 2007; Helms, Kaen and Rosenman 1984; Jin and Frechette 2004; Peterson, Ma and Ritchey 1992; Wei and Leuthold, 2000). In contrast, although she does not test for spurious long memory, Lordkipandize (2004,

p. 82) finds that soybean and corn futures price volatility is primarily caused by seasonality and maturity effects rather than by long memory.

The main contribution of this chapter is to determine whether findings of long memory in agricultural commodity futures price volatility are spurious and to suggest an alternative explanation based on evaluating different causes of spurious long memory. This chapter provides robust estimates of the long memory parameter for eleven commodity futures contract time series in a joint model with short memory and seasonal model parameters. The long memory estimator is unbiased by the presence of short memory effects. Correct standard errors are computed using the complete Information Matrix accounting for cross-dependencies with short memory. To evaluate whether findings of long memory are significant, asymptotic tests (Wald, Likelihood Ratio) are applied, but since these tests have incorrect size, we also use recently developed tests for spurious long memory.

3.3 Commodity Futures Price Data

The data consist of business daily observations of agricultural commodity futures prices for contracts on coffee, cotton, cocoa, sugar no.11, frozen concentrated orange juice, hard red winter wheat, soybeans, corn, canola, live cattle, and lean hogs (formerly live hogs). Commodity futures contracts are traded until the 15th of the contract month (or the last business day before the 15th). To avoid near-maturity effects and delivery risk bias, observations for contracts in their own expiry month are discarded. Contracts are therefore rolled-over (spliced) approximately 15 days before they expire.

The observations cover the years 1988-2007, varying slightly across commodities. Data for the years 2005, 2006 and 2007 are reserved for out-of-sample forecasting, which implies at least 500 observations for each commodity, and leaves more than 4000 observations for each commodity for the estimation of long and short memory parameters. Precisely 4096 observations are used for estimation purposes.

The contracts include both storable and non-storable commodities. Storable commodities have inventory stocks while by definition non-storable commodities do not. This suggests a testable hypothesis that price and volatility dynamics will differ between storable and non-storable commodities (Williams and Wright 1984, 1989).

3.4 The Option Pricing Bias from Long Memory

One typical violation of the Black-Scholes model in futures price sample data is volatility clustering (Myers and Hanson 1993), generally addressed by using ARCH and GARCH models (Engle 1982; Bollerslev 1986). This short-range dependence however does not appear to substantially affect option pricing solutions (Roberts 2002).

Long-range dependence, or long memory, implies the Black-Scholes option pricing solution is fundamentally biased (Rogers 1997; Sottinen 1998), as the underlying asset is better described by fractional Brownian motion, a more general stochastic process that nests geometric Brownian motion as a special case (Cox and Miller 1965). How important is the bias caused by long memory on option pricing? Ohanissian, Russell and Tsay (2004) find that it can cause options to be mispriced by as much as 67%.

3.5 A Log-Range Measure of Volatility

The two traditional measures of daily or weekly volatility in the commodity spot and futures prices literature are absolute and squared logreturns, computed as deviations from the long-run mean. If the underlying asset price at time t is F_t then the logreturn is defined as: $r_t = \ln(F_t) - \ln(F_{t-1})$ and volatility is defined as either $|r_t|$ or as $(r_t)^2$. Improved efficiency and no significant bias follow from assuming the long-run mean is zero.

Though both measures are frequently used, Granger (2000) argues on the basis of Nyquist's (1983) L_p norm argument that squared logreturns should only be used if the data are approximately Gaussian Normal, which is seldom true in economic and financial logreturn data. Since these data display excess kurtosis, absolute logreturns are more appropriate.

In this chapter, the log-range of daily futures prices is used as a measure of volatility instead of absolute price logreturns. There are several reasons why this is warranted. Alizadeh, Brandt and Diebold (2002) and Yang and Zhang (2002) provide theoretical and empirical evidence for the asymptotic optimality of the log-range as an estimator of volatility in economic and financial time series data. Regarding the asymptotic validity of the result, all of our commodity time series consist of more than 4000 observations. Absolute logreturns are a particularly noisy proxy for price variation and are more heavily contaminated by measurement error (Parkinson, 1980; Garman and Klass 1980; Rogers and Satchell 1991). As a result, the log-range based volatility measure is more efficient than are absolute logreturns. Anderson and Bollerslev (1998) show that the range-based volatility measure is nearly as accurate as computing realized volatility from ultra high-frequency tick data, the latter which is the ideal measure of daily integrated volatility (Andersen, Bollerslev, Diebold et al. 2001a, 2001b, 2003; Barndorff-Nielsen and Sheppard 2002). The lower volume of trade in commodity markets indeed makes the realized volatility approach difficult to implement.

The log-range is very well approximated by the Gaussian Normal distribution, which improves both efficiency and accuracy in maximum likelihood estimation (Alizadeh, Brandt and Diebold 2002; Brandt and Jones 2006). In particular, quasi-MLE estimation using a logreturn-based volatility is highly inefficient (Andersen and Sorensen 1997; Kim, Shephard and Chib 1998). Lastly, absolute or squared logreturns are not well supported by choice theory as proxies for risk (Machina, 1987; Levy, 1992).

The log-range, for a time increment *t* that can be a day or an intra-daily time period, is defined as:

$$\mathrm{h_t} = \ln(\sup F_t - \inf F_t) \tag{3.1}$$

Parkinson (1977, 1980) shows that the log-range is closely related to the diffusion term σ in the geometric Brownian motion (Black-Scholes) asset price model and option price solution. This result is based on Feller's (1951) definition of the Moment Generating Function of a random variable that behaves as a daily range of prices. Open and close prices are not incorporated as they do not improve accuracy of results and they introduce undesirable market microstructure effects (Brown, 1990; Alizadeh, 1998).

Descriptive statistics for the log range volatility measure are presented in Table 3.1. International commodities traded at the New York Board of Trade, such as cocoa, coffee and cotton, are more volatile, skewed and leptokurtic (heavy-tailed) than are principally domestic commodities such as Chicago Board of Trade grains and Chicago Mercantile Exchange meats. Kim and White's (2001) measures of skewness and kurtosis are used, which are more robust to the presence of outliers and therefore provide a better description of the data's first four sample moments.

Table 3.1: Descriptive statistics of log-range price volatility in commodity futures contract time series data, T=4266, daily observations from 2/1988 to 1/2005

Futures contract	Mean	Std dev.	Skewness (Normal=0)	Kurtosis (Normal=3)
CBOT corn	0.015	0.008	2.070	8.417
CBOT soybeans	0.015	0.008	1.932	6.719
CME lean hogs	0.017	0.009	2.315	14.013
CME live cattle	0.011	0.005	1.494	3.077
KCBOT wheat	0.015	0.009	1.690	4.776
WCE canola	0.012	0.007	1.544	4.368
NYBOT cocoa	0.022	0.013	1.723	5.149
NYBOT coffee	0.028	0.018	2.175	9.204
NYBOT FCOJ	0.020	0.014	3.071	19.649
NYBOT cotton	0.018	0.011	2.295	12.069
NYBOT sugar#11	0.026	0.016	2.562	16.308

A number of robustness checks are performed. To control for calendar effects such as the "weekend" anomaly (French 1980; Thaler 1987; Gibbons and Hess 1981; Kamara 1997), we repeat estimation for a small number of commodities using only the Wednesday observation (i.e. weekly sampling). Two reasons suggest however that calendar effects need not be a problem. Empirical work has found that these anomalies have essentially disappeared since 1975 (Connolly 1989) or since 1987 (Fortune 1998), and the earliest data used in this chapter begins in 1988. Also, once unintentional data snooping is accounted for, calendar effects have been found to be in general not statistically significant (Sullivan, Timmermann and White 2001).

Standard time series diagnostic tests are performed on the data (Augmented Dickey-Fuller, Phillips-Perron, KPSS, Variance Ratio) to evaluate its sample properties and ensure that our data are comparable with data used in previous research. Test results suggest that in levels we cannot reject the null of a unit root (ADF test) but in differences we cannot reject the null of no unit root (KPSS test). Such findings are standard in the literature, but Wang and Tomek (2007) warn that commodity prices in levels should not in theory be characterized by a unit root. Rather, such test results are the consequence of low test power caused by mis-specification of the test, omission of level shifts in the data or both. The data in log-return or log-range form are stationary but ARCH effects (volatility clustering) are present. Test details are provided in the Appendix. The data are not deflated by the Prices Paid Farmers Index (Tomek 1997) because this Index has an annual frequency while the data are daily, therefore spurious effects risk being introduced. Figures 3.1 to 3.11 present time series plots of the nearby futures contract volatility data for the eleven commodities studied in this chapter.

3.6 Wavelets Distinguish Short from Long Memory

A substantial difficulty associated with estimating the long memory parameter (H or d) is that it is, even asymptotically, correlated with short memory dynamics such as AR and MA parameters (Tanaka 1999). As a result, both the point estimate of the long memory parameter and its standard errors are biased.

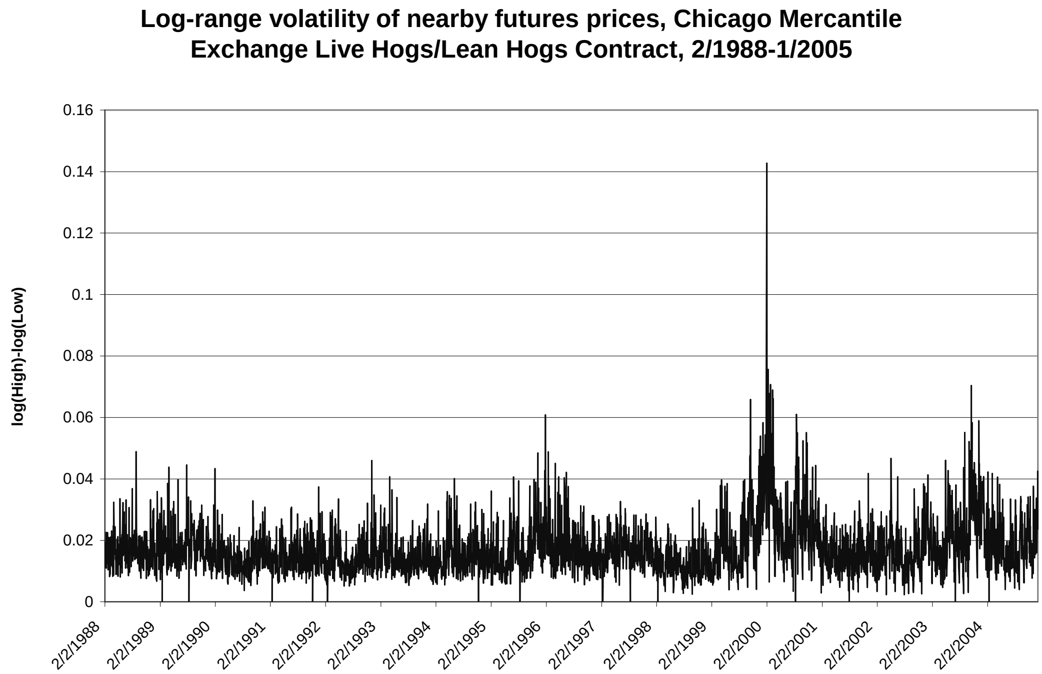


Figure 3.1: Time series plot of daily log-range price volatility, CME lean hogs futures

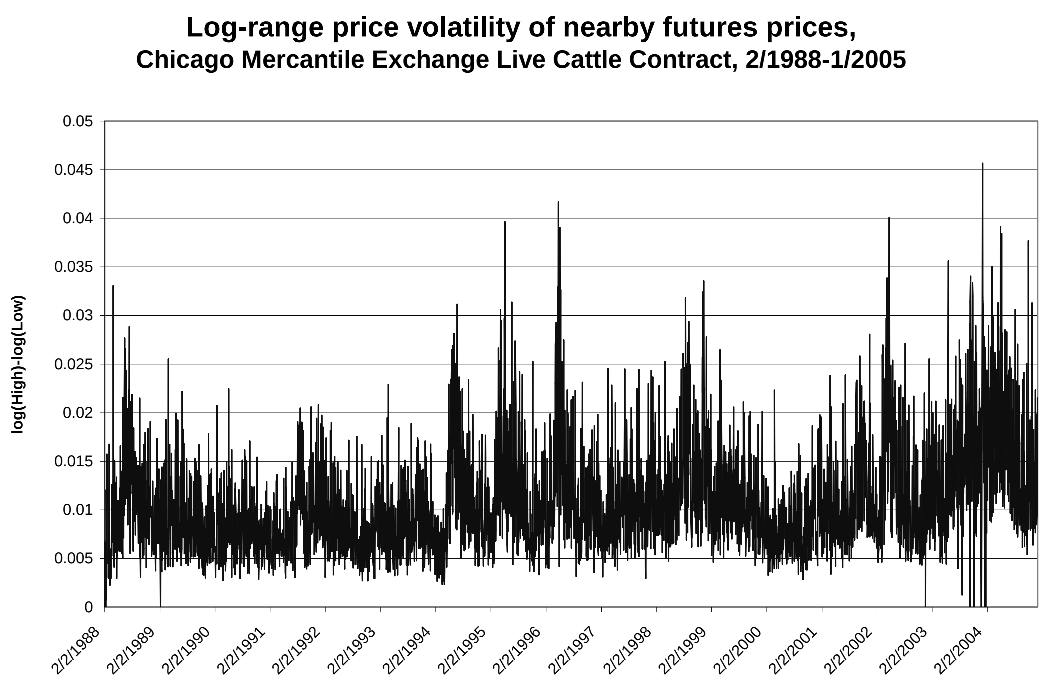


Figure 3.2: Time series plot of daily log-range price volatility, CME live cattle futures

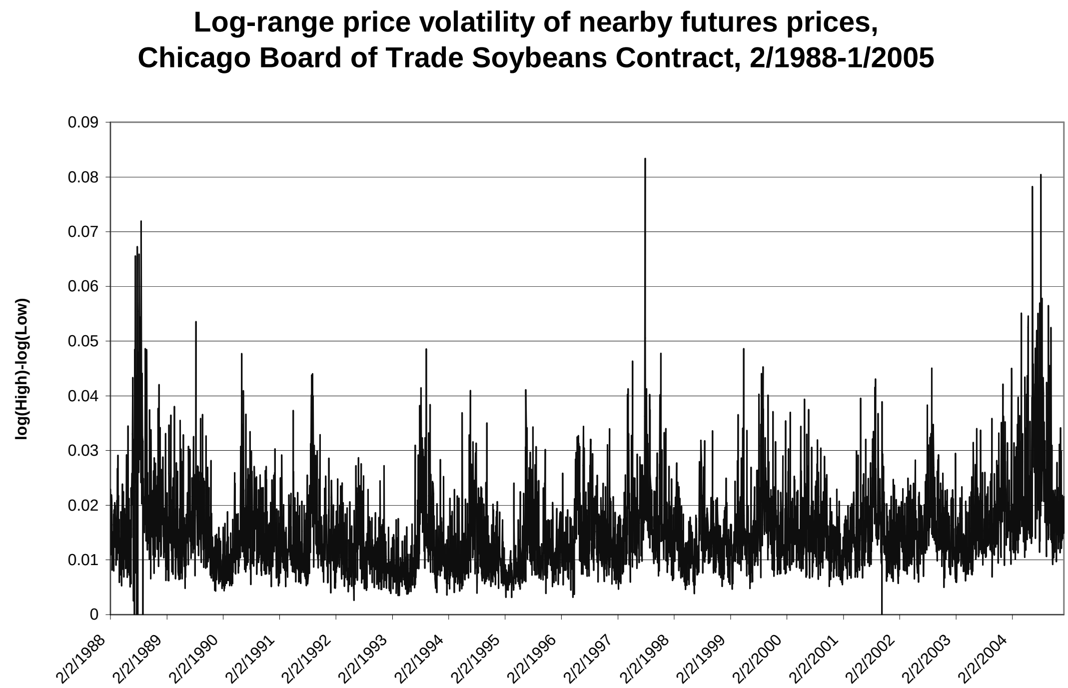


Figure 3.3: Time series plot of daily log-range price volatility, CBOT soybeans futures

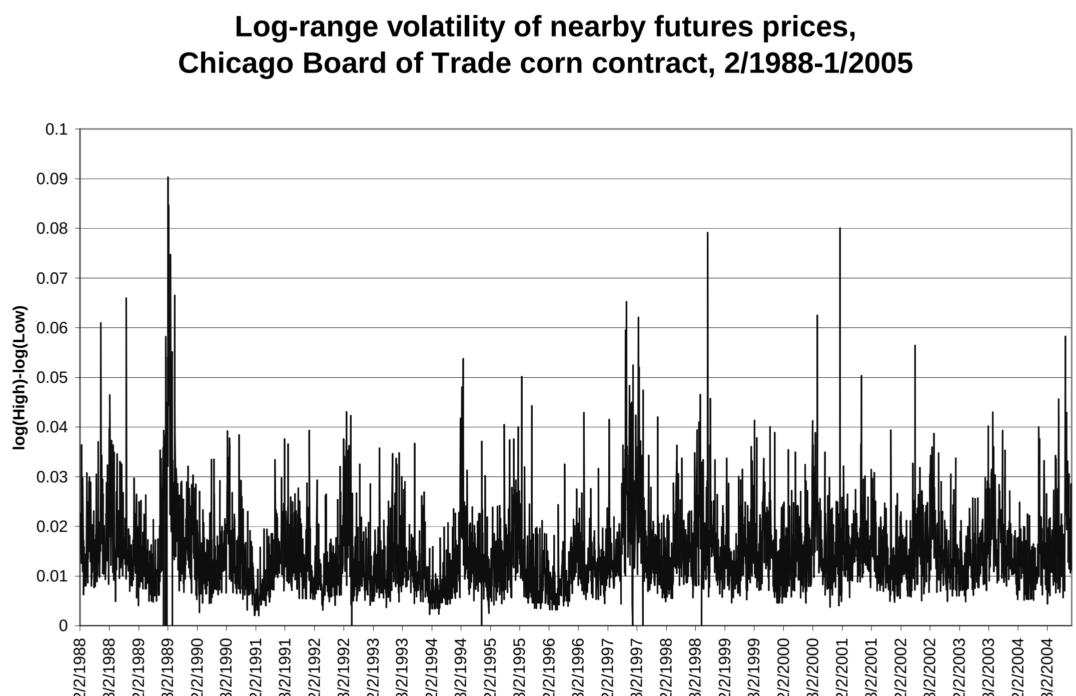


Figure 3.4: Time series plot of daily log-range price volatility, CBOT corn futures

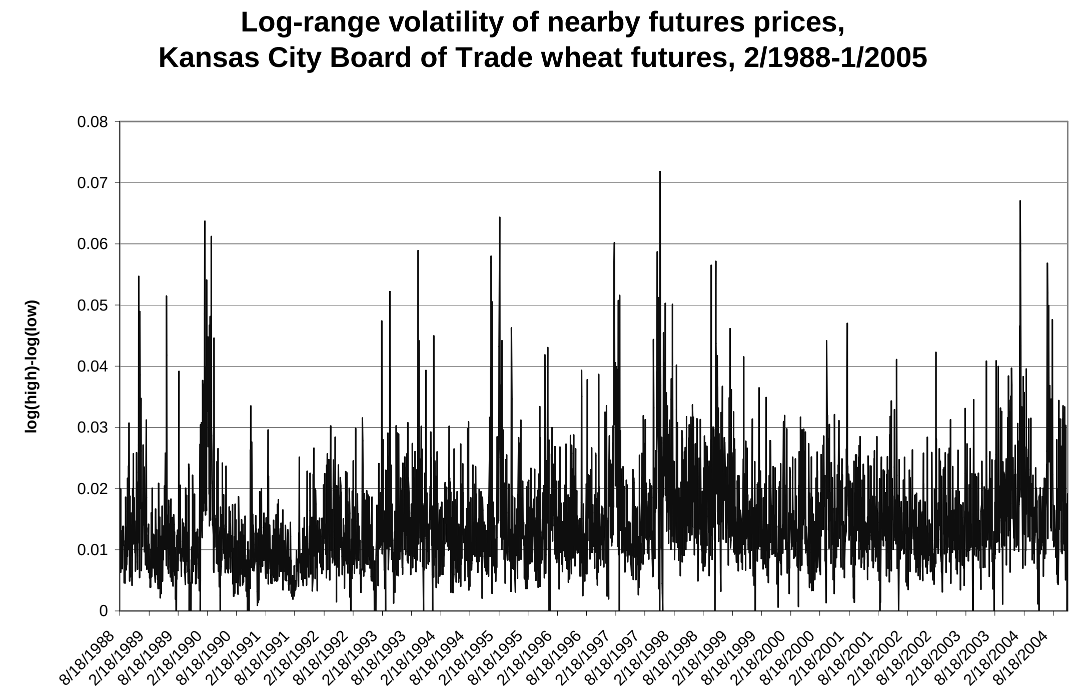


Figure 3.5: Time series plot of daily log-range price volatility, KCBOT wheat futures

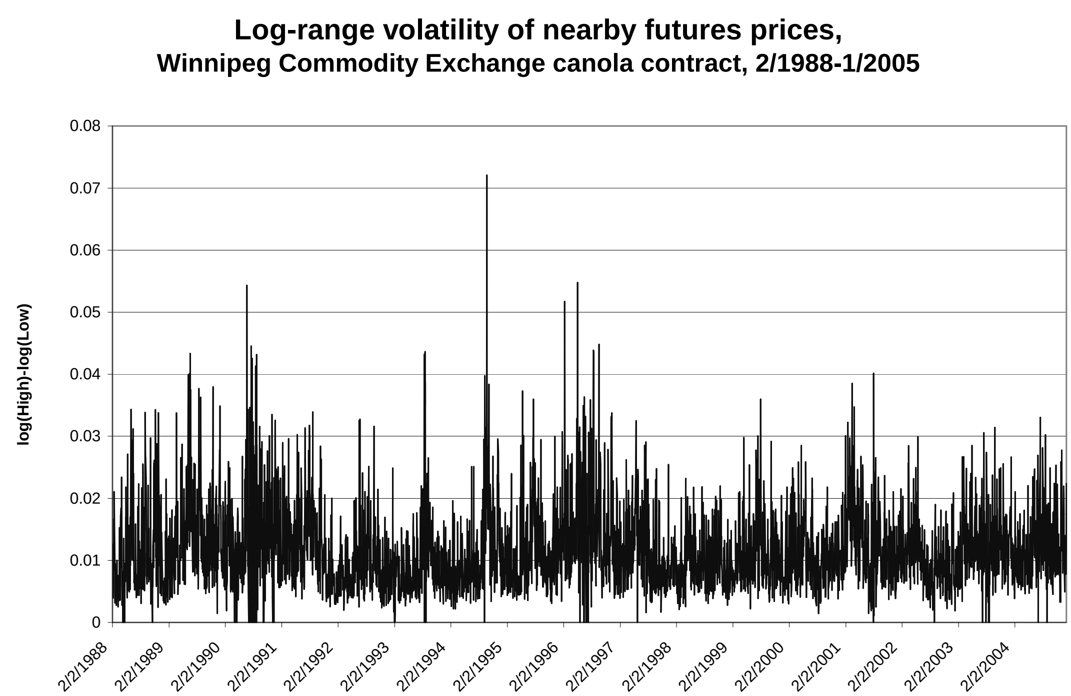


Figure 3.6: Time series plot of daily log-range price volatility, WCE canola futures

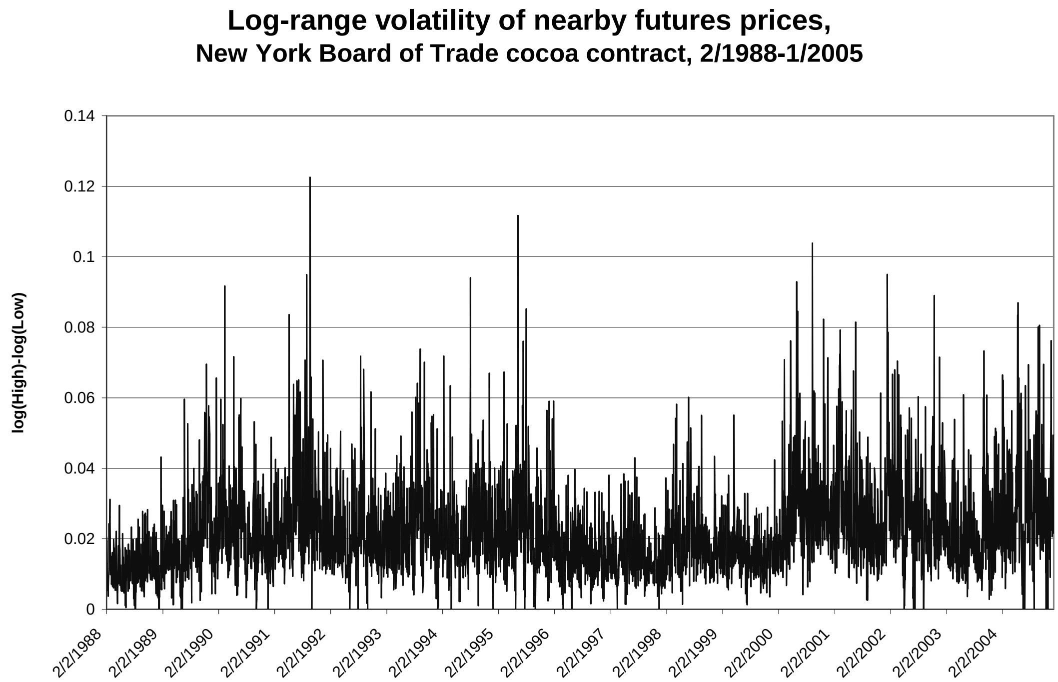


Figure 3.7: Time series plot of daily log-range price volatility, NYBoT cocoa futures

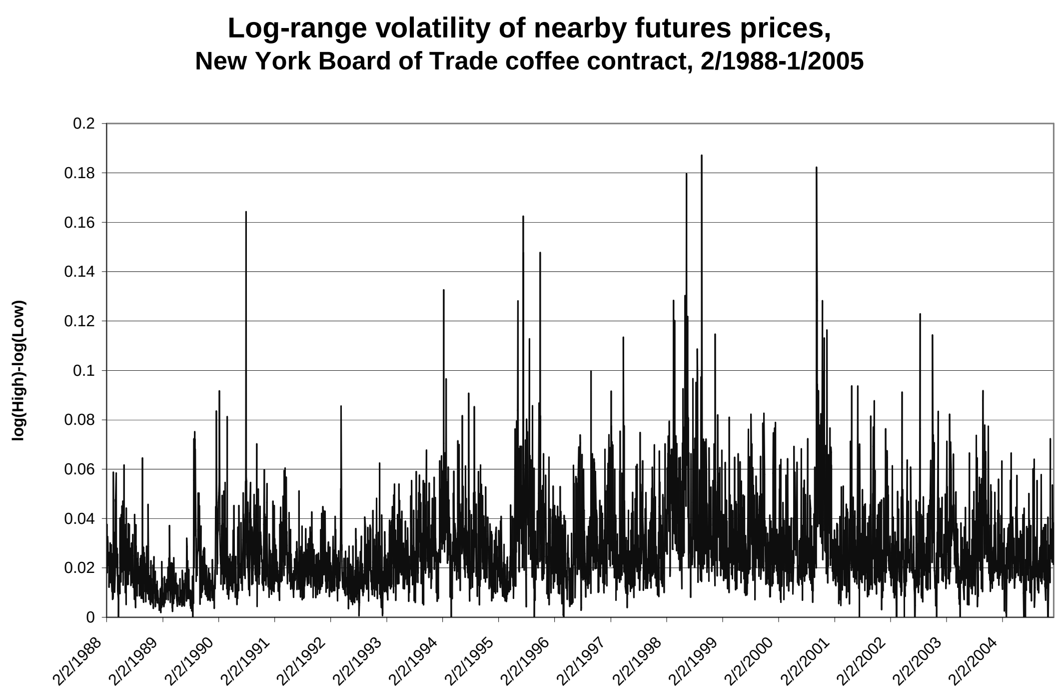


Figure 3.8: Time series plot of daily log-range price volatility, NYBoT coffee futures

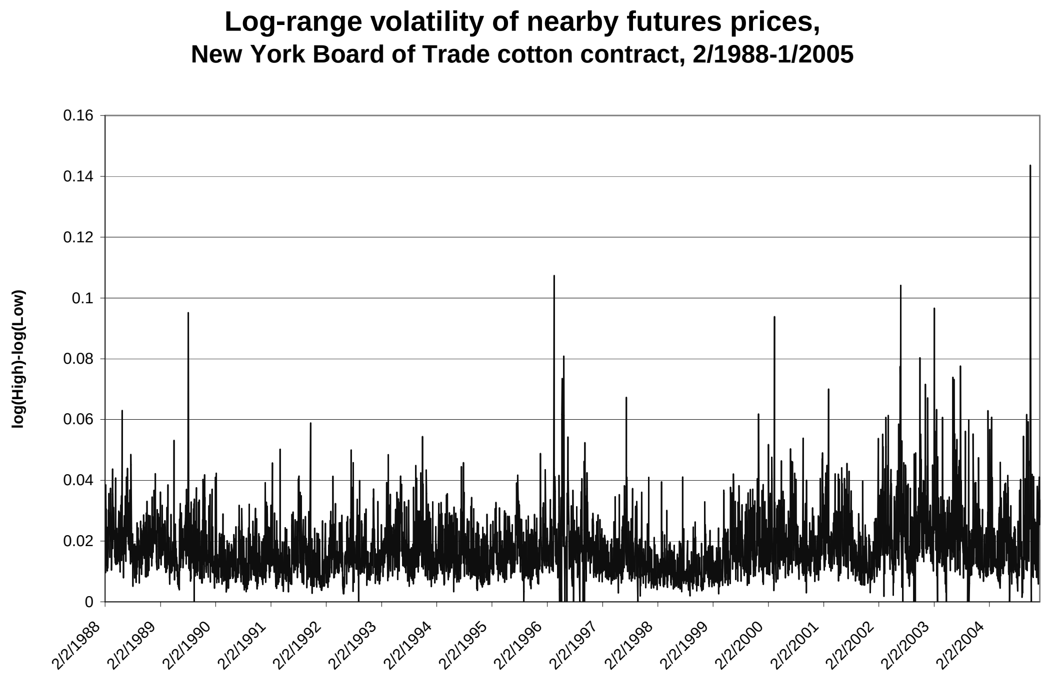


Figure 3.9: Time series plot of daily log-range price volatility, NYBoT cotton futures

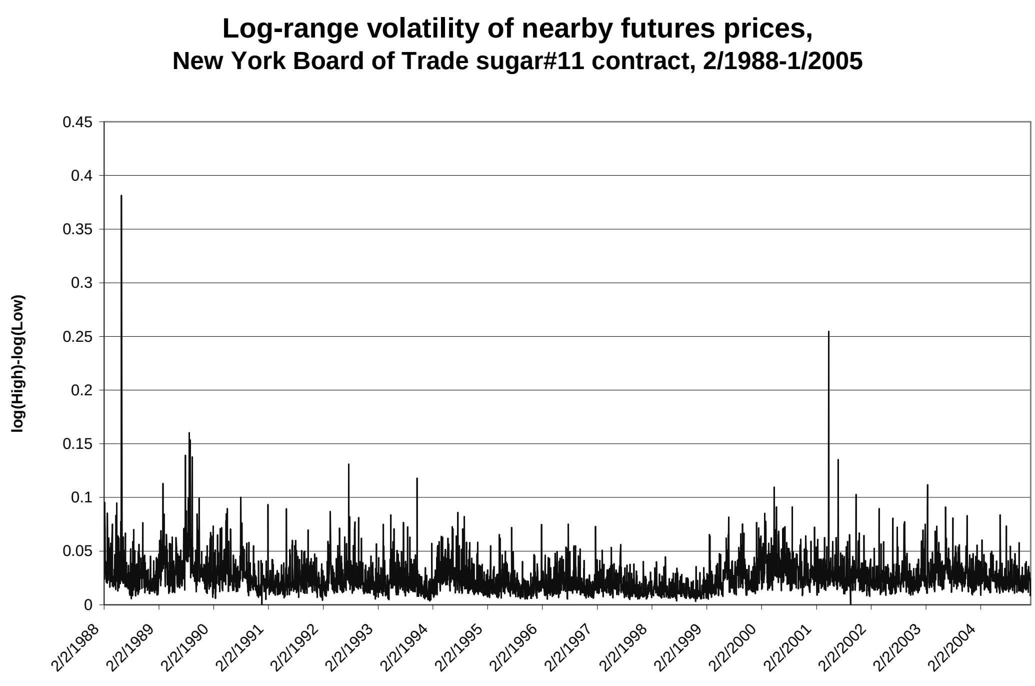


Figure 3.10: Time series plot of daily log-range price volatility, NYBoT sugar#11 futures

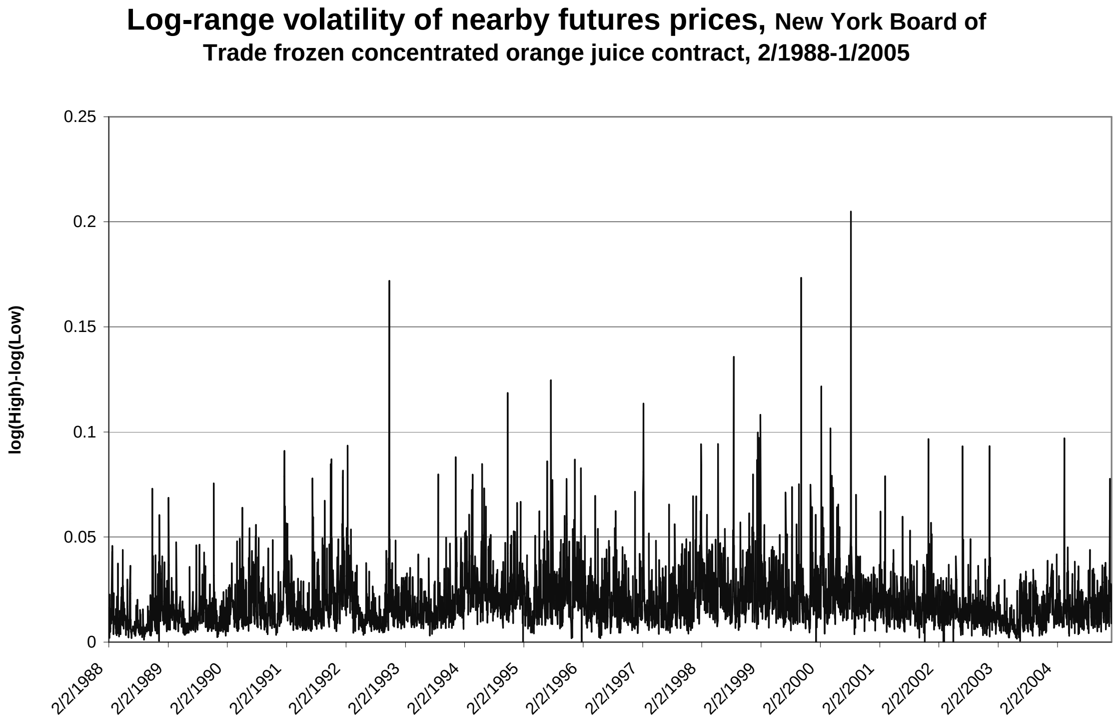


Figure 3.11: Time series plot of daily log-range price volatility, NYBoT frozen concentrated orange juice futures

Many papers in the literature do not appear to account for the impact of short memory dynamics on their estimates of long memory. One solution to this problem is to use an estimation method based on wavelet functions (Gencay, Selcuk, and Whitcher 2001). This is because wavelets are by design able to separate long memory from short memory dependence, or more generally, variation in a signal or time series that occurs at different timescales (Percival and Walden 2001).

This implies a wavelet-based estimate of long memory will be unbiased when the time series data short memory parameters are either ignored or inaccurately estimated. Moreover, long memory and short memory parameters can be independently and accurately estimated. To the best of our knowledge, the only work that has considered wavelet-based estimators to examine long memory in agricultural commodity price data is Elder and Jin (2007). But, where they focus on comparing results of long memory estimation using wavelet and non-wavelet based methods, this chapter tests for spurious long memory and determines both the cause of spurious long memory and what alternative model better describes the data. To establish the significance of the long memory parameter *d*, other papers in the literature such as Jin and Frechette (2004) only use, for example, non-robust Likelihood Ratio tests. Lastly, Elder and Jin (2007) use a logreturn-based volatility variable while the daily log-range is used in this chapter.

3.7 Identifying Spurious Long Memory

A persistent problem in the literature is the accurate identification of long memory in futures data. Early evidence of long memory in cash and futures prices has been reconsidered and recent advances have focused on long memory in the volatility of futures prices (e.g. Baillie et al. 2007; Jin and Frechette 2004).

In time series econometrics, long memory is generally defined as fractional integration, or I(d), which is only one type of long memory process (Granger 2000). It is well understood that the aggregation of short memory (e.g. ARMA) time series data may result in the appearance of long memory (Granger 1980, 1990). Indeed, Chambers (1998) proves that true long memory processes have a long memory parameter that is invariant under time aggregation, a useful fact for hypothesis testing. The illusion of long memory can also be the consequence of structural breaks and level shifts, two phenomena that are better supported than is long memory by economic theory (Diebold and Inoue 2001; Granger 2005).

Disentangling stochastic, singular shocks such as structural breaks from long memory in time series data is a difficult task and in many cases the two classes of models are observationally equivalent (see e.g. Banerjee and Urga, ed., 2005, *Journal of Econometrics* symposium; Perron 2006). Spurious findings of long memory may be caused by a biased or inconsistent estimation procedure, by level shifts, structural breaks and regime switches, or by inefficient standard errors and confidence intervals (Chambers 1998; Diebold and Inoue 2001; Shimotsu 2006; Zivot and Andrews 1992). Level shifts can occur for example when the first moment (mean) of the data generating process suddenly changes while the rest of the distribution is unaffected (Smith 2005). Structural breaks occur when the values of some or all coefficients in the model change at some point in the time series. Regime switching is generally described by a time series process whose distribution is stationary for a given state of nature, and for which the state in each time period is determined by a probabilistic, e.g. Markov, transition matrix.

Daily volatility of stock logreturns is characterized by autocorrelograms that are significant beyond 3000 (day) lags, even after removing outlier observations (Granger 1999). Estimates of the fractional difference parameter d using large data samples generally fall below but near 0.5, but for sub-samples of shorter length the estimates vary between 0.3 and 0.7, which suggests it is not true long memory. Another reason to doubt that economic or financial time series are generated by a true fractional integration process is that it is difficult to reconcile estimates of d with the data's sample moments. For example, for daily absolute logreturns of financial data, it would be necessary to assume the innovations (errors) are distributed as fractional Chi-Squared.

The estimators generally used in the literature are not necessarily robust. The popular Geweke-Porter-Hudak semi-parametric estimator is both inconsistent and inefficient (Robinson 1995a,b). Smith (2005) moreover shows that the GPH estimator is heavily biased in the presence of level shifts in the data and suggests a new, nearly unbiased GPH-type estimator. This bias explains for example an apparently large (d=0.79) estimate of long memory in relative soybean prices. As a result, once level shifts have been accounted for, estimates of long memory are not statistically different from zero.

A second widely used long memory estimator designed for volatility data is the Fractionally Integrated GARCH model (Bollerslev 1986);Bollerslev and Mikkelsen 1996). The FIGARCH estimator is however both fragile in the presence of mis-specified short memory parameters and also unreliable as a measure of long memory (Davidson 2004). A third case of a problematic long memory estimator is the Quasi-MLE estimator for stochastic volatility with long memory (e.g. Breidt, Crato and de Lima 1998), which is generally non-robust (Alizadeh, Brandt and Diebold 2002; Andersen and Sorensen 1997).

As for the large class of semi-parametric Hurst long memory parameter H estimators (e.g. Lo 1991), Riedi (2003) shows that confidence intervals around H are only reliable under overly restrictive conditions, and Franco and Reisen (2007) use simulation to show that bootstrapped standard errors of the long memory parameter are not accurate. Turvey (2007) shows that for all but two agricultural commodities, the data generating process is consistent with white noise innovations rather than fractional Gaussian noise (as would be the case under long memory).

These results suggest the need for a more robust investigation of long memory.

3.8 Semi-parametric Wavelet Estimation of Long Memory

Research has found that semi-parametric estimators, frequently used in the natural sciences, are superior to parametric (maximum likelihood) estimators when the model is likely to be mis-specified (Boes et al. 1989), but that MLE is preferable when the model is correctly specified (Cheung 1993). The classic Rescaled-Range analysis (R/S) of Mandelbrot and Van Ness (1968) obtains an estimate of Hurst's *H* long memory parameter. Lo (1991) improved upon Hurst's and Mandelbrot's R/S estimator by making it robust to heteroskedasticity in the data, but interpretation and hypothesis testing appear unreliable (Teverovsky, Taqqu and Willinger 1999).

The properties of wavelets, in particular scale-invariance, make them ideally suited to detect the self-similar fractal signature of several types of long memory, including fractional Brownian motion. A wavelet-based semi-parametric estimator of the Hurst parameter can be implemented and provides results that are superior to traditional R/S analysis (Teyssiere and Abry 2006). This semi-parametric estimator can be applied to all timescales without adjustment and is has been found to be unbiased and efficient in its class. The wavelet orthogonality property makes this estimator robust to the presence of a trend and to non-stationary singularities.

The Hurst coefficient can be easily obtained from an application of a wavelet transform to time series data (see e.g. Taqqu 2003). The method consists of first applying a Discrete Wavelet Transform to the time series data, which produces a vector of wavelet coefficients. Then the wavelet coefficients, each of which is associated with a timescale, are squared and regressed over the base-2 (dyadic) logarithm of the timescales. The slope coefficient is directly proportional to *H*. A similar method by Jensen (1999) can be used to obtain an OLS estimator of the fractional difference parameter *d*, which is directly related to the Hurst coefficient *H.*

Improved, unbiased semi-parametric estimators of long range dependence *H* have been developed by Abry, Veitch and Flandrin (1998), Veitch and Abry (1999), and Teyssiere and Abry (2006). These jointly estimate the long-range dependence parameters α and *C* and also compute a tailored goodness-of-fit statistic. Their approach has the advantage of using a pre-filtering algorithm to correct the bias caused by the discrete sampling of the data (Veitch, Taqqu, and Abry 2000). In addition, Veitch and Abry (1999) propose a test for true long-range dependence that relies on the self-similar properties of wavelets. This is a test of the stationarity of the long memory parameter computed over a number of sub-samples. In this chapter, we consider 16 sub-samples of 256 observations each. This corresponds to estimating *H* approximately once per year for every year in the sample and finding out if this parameter changed over time. Results are presented for the three commodities for which the stationarity of *H* is rejected graphically in Figures 3.12 to 3.14.

The results, presented in Table 3.2, show that for all commodities, the null hypothesis of H=0.5 cannot be rejected at the standard 5% level of significance. This evidence supports the recent findings of Turvey (2007), that increments of the data are consistent with a white noise process (not necessarily Gaussian) rather than long range dependence such as fractional Brownian motion.. For all but three commodities, we cannot reject the null hypothesis that the long memory parameter *H* has been constant over the entire sample (1988-2004). The stationarity of *H* is however clearly rejected (at the 1% level) for CME lean hogs, KCBOT wheat and NYBOT sugar #11.

Although straightforward to compute and frequently used in the natural sciences, the semi-parametric wavelet approach is of limited usefulness in economics because it has been shown that both analytical and bootstrap standard errors and confidence intervals are unreliable for this estimator (Riedi 2003; Franco and Reisen 2007).

3.9 Parametric Wavelet Estimation of Long Memory

Following Granger's (1980) and Hosking's (1981) formal definitions of long memory in the ARMA time series framework, Sowell (1992) obtained an exact maximum likelihood estimator for fractionally integrated processes. Its computation requires, however, inverting a dense covariance matrix at every step of the procedure, which is unrealistic for large datasets. For this reason, approximate frequency domain estimators such as those by Geweke and Porter-Hudak (GPH, 1983) or Fox and Taqqu (1986) are frequently used. However, the GPH estimator is both inconsistent and inefficient (Agiagoglou, Newbold and Wohar 1992; Robinson 1995) while the Fox-Taqqu estimator is systematically biased. Feasible exact ML estimators suffer from a large bias as the sample size grows because they are not robust to a mis-specified mean or trend (Cheung and Diebold, 1994). Indeed, the sample mean is an inaccurate estimator of the population mean in the presence of long memory (Beran 1994). Robinson (1995) suggests instead a semi-parametric local Whittle estimator based on Kunsch (1987).

Hosking (1984) derives a Cumulative Sum of Squares (CuSum) estimator that is asymptotically equivalent to Sowell's (1992) exact MLE, but the CuSum estimator is severely biased in small to moderate-sized samples (Chung and Baillie 1993). Chung (1996a,b) derives asymptotic results for the CuSum estimator of a generalized

ARFIMA(p,d,q) process including an analytical formula for standard errors which is used in this chapter.

Table 3.2: Semi-parametric wavelet-based estimation of Hurst long memory parameter *H*

Commodity	H estimate (wavelet)	Std. error	Reject Ho: H=0.5?	p-value for Ho: stationary H	Reject Ho: stationary H?	H estimate (Turvey)	Reject Ho: H=0.5?
Coffee (NYBOT)	0.483	0.015	No	0.088	No	0.402	
Cocoa (NYBOT)	0.492	0.015	No	0.496	No	0.465	
Corn (CBOT)	0.519	0.015	No	0.747	No	0.348	**
Cotton (NYBOT)	0.523	0.016	No	0.186	No	N/A	
Lean hogs (CME)	0.483	0.015	No	0.0066	***	0.438	
Live cattle (CME)	0.516	0.015	No	0.115	No	0.272	***
FCOJ (NYBOT)	0.507	0.015	No	0.054	No	0.458	
Canola (WCE)	0.496	0.015	No	0.053	No	0.396	
Soybeans (CBOT)	0.482	0.024	No	0.296	No	0.332	**
Sugar#11 (NYBOT)	0.506	0.015	No	0.0015	***	0.543	
Wheat (KCBOT)	0.493	0.015	No	0.0017	***	N/A	

** reject 5%, *** reject 1%

Notes: The estimator is based on Abry and Veitch (1998, 1999, 2002) with a pre-filtering correction for discretely sampled data and using the Daubechies(10) wavelet function. Test is for Ho: H=0.5 (independent increments) and test for stationarity of H over time. Comparison of estimates with results from Turvey (2007) Table 5 (Sample=940 days).

The ability of wavelet functions to decorrelate time series data across timescales helps distinguish long memory from short memory (ARMA) components as well as from change-points or structural breaks (Percival and Walden 2001). Wavelet-based estimators of long memory are not affected by the presence of an unknown or mis-specified mean unlike exact ML estimators (Jensen 2000).

Recall that exact ML estimation of long memory involves inverting a dense covariance matrix at every step of the convergence procedure. Wavelets provide a sparse representation of the covariance matrix, which greatly simplifies this computational burden and introduces only a trivial bias. A large number of wavelet-based estimators of long memory have been developed. McCoy and Walden's (1996) presented an early wavelet-based exact MLE, which was improved upon by Percival and Bruce (1998) to include robustness to polynomial trends, by Jensen (2000) for robustness to contaminated (e.g., non-experimental) data and by Craigmile, Guttorp and Percival (2005) for robustness to trend contamination.

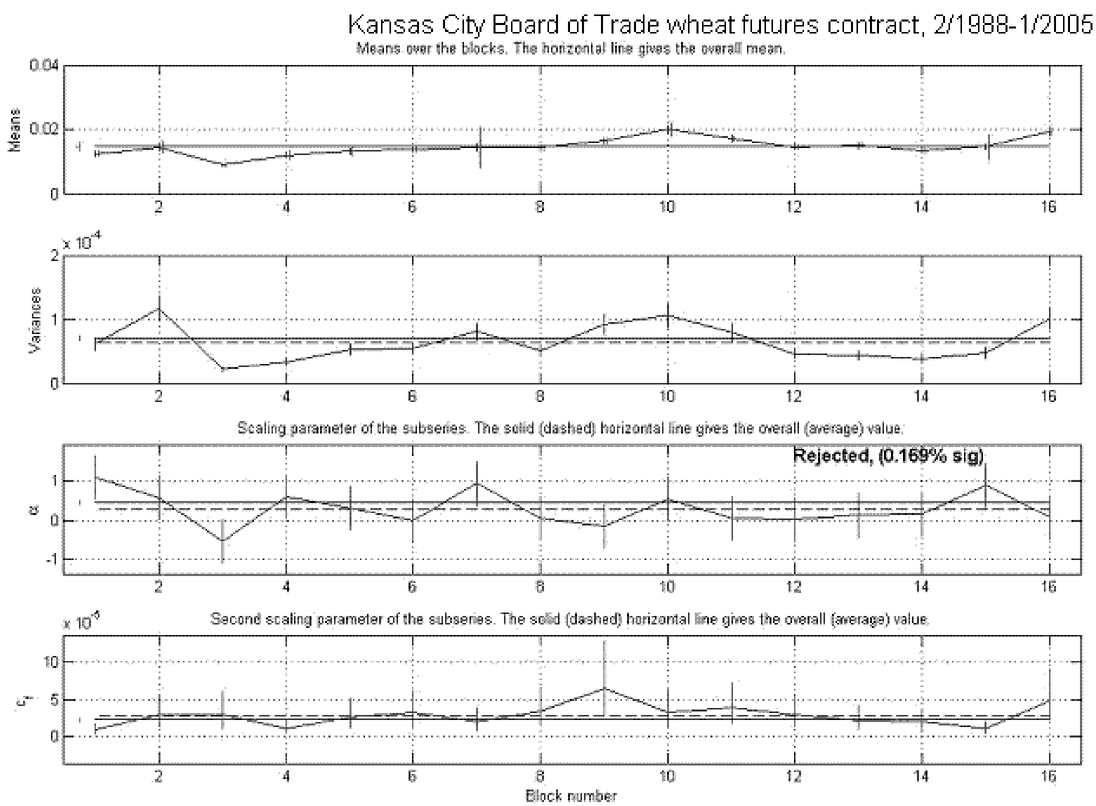


Figure 3.12: Stationarity test results for wavelet-based estimate of Hurst *H* parameter, KCBOT wheat futures contract

Whitcher (2004) introduces a seasonal component, and Jensen (1998, 1999) suggests a method to jointly estimate long memory and short memory parameters.

The general long memory process to be estimated is:

$$\phi(\mathrm{L})\ (1\text{-}2\eta\text{+}\mathrm{L}^2)^d\,(Y_t-\mu)=\theta(L)\varepsilon_t \qquad (3.2)$$

which includes both autoregressive ϕ(L) and moving average θ(L) polynomials, a fractional order of integration *d* (Hosking 1981) as well as a seasonal persistence process η (Gray, Zhang and Woodward 1989) which is equivalently a power series known as Gegenbauer polynomials (Rainville 1960).

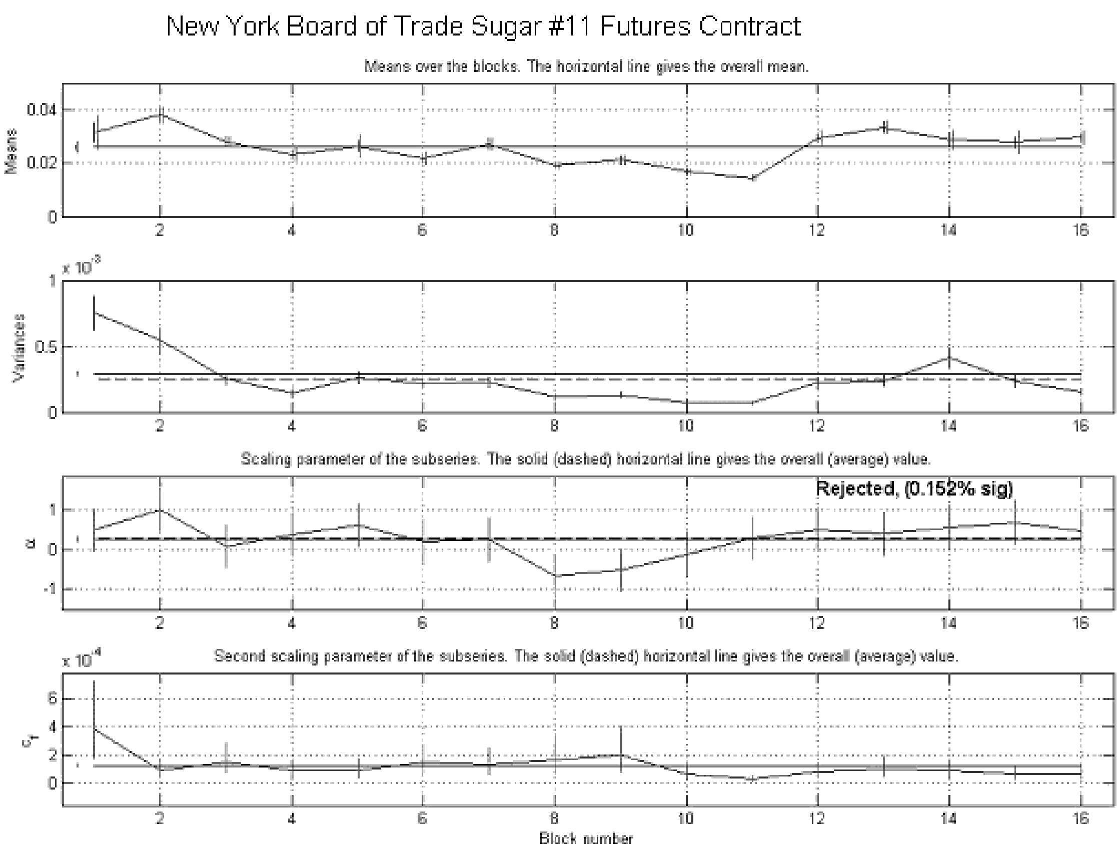


Figure 3.13: Stationarity test results for wavelet-based estimate of Hurst *H* parameter, NYBOT sugar no.11 futures contract

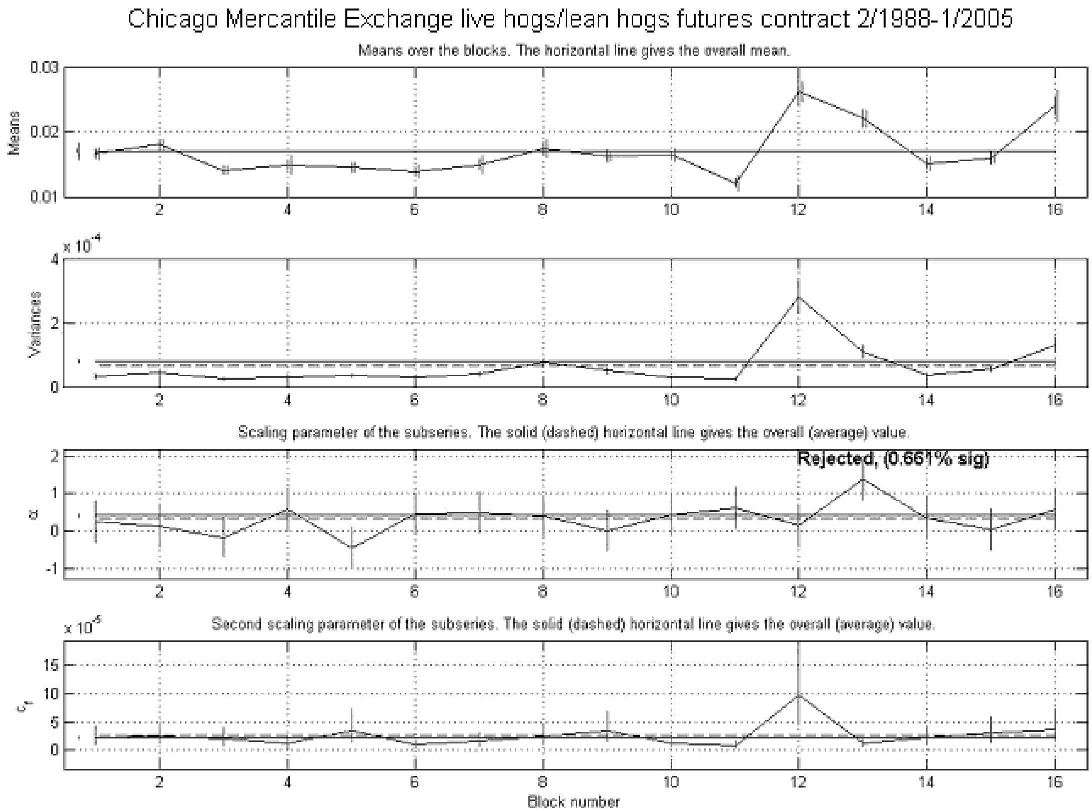


Figure 3.14: Stationarity test for wavelet-based estimate of Hurst *H* parameter, CME live hogs/lean hogs futures contract

Gegenbauer polynomials enable the long memory parameter to be associated with seasonality, an advantageous option to study certain economic time series. In this chapter, however, time series are sampled daily and we impose the restriction that long memory is not seasonally-dependent.

The exact wavelet ML estimator used in this chapter is based on the Haar(4) wavelet transform (Daubechies 1992), as described in Chapter 2. It has the smallest (Root) Mean Squared Error in its class, is computationally efficient, only slightly affected by the wavelet boundary effects caused by the finiteness of the data sample, and is robust to misspecification of trend and short memory (ARMA) parameters (Jensen 2000).

The estimator jointly provides values for the long memory and short memory parameters including seasonality coefficients that are useful for grain futures contracts.

For a general ARFIMA(p,d,q) process with white noise innovations, the concentrated log-likelihood (Jensen 2000) is:

$$\log \ell(\theta) = -\frac{1}{2}\sum_{m\in M}\sum_{N(m)}\left(\frac{\langle Y,\psi_{m,n}\rangle^2}{\sigma_{m,n}} + \ln 2\pi\sigma_{m,n}\right) \tag{3.3}$$

where $\theta=(d, \sigma^2, \sigma_\eta^2)$ is the vector of parameters (long memory, ARFIMA variance, and white noise variance) and $\langle Y, \psi^2\rangle$ is the vector of wavelet coefficients resulting from the convolution of wavelet functions with the original data. Standard errors for the long memory parameter are computed following analytical solutions from Chung (1996a,b) and Tanaka (1999). It is necessary to first estimate *d*, then estimate the ARMA parameters ($\phi(L)$, $\theta(L)$) and finally obtain the information matrix for the ARMA parameters (see e.g. Hamilton 1994, pp. 142-144). Only then can accurate standard errors for *d* be computed. If the short memory parameters are all zero, such that the process is ARFIMA(0,d,0) then the standard errors for *d* are computed as follows:

$$se(\hat{d}) = T^{-1/2}\sqrt{6/\pi^2} \tag{3.4}$$

For the case of seasonal persistence, standard errors are computed as follows:

$$se(\hat{d}) = T^{-1/2}\left[2\left(\frac{\pi^2}{3} - \pi\arccos(0.5) + \arccos^2\right)\right]^{-1/2} \tag{3.5}$$

A Wald test can be applied to evaluate the null that the fractional difference parameter *d* is zero, equivalently that there is no long memory (Tanaka 1999). Even though the wavelet-based estimator is robust to the presence of mis-specified short memory parameters, the test is only accurate if these ARMA (or GARCH) terms are included in the Information Matrix:

$$\frac{\hat{d} - d_0}{\sqrt{T^{-1}\omega^{-2}}} \tag{3.6}$$

where the complete Information matrix for both long and short memory parameter estimators is:

$$\omega^2 = \left(\frac{\pi^2}{6} - \kappa' \Im^{-1}(\phi, \theta)\kappa \right) \tag{3.7}$$

$\Im(\phi,\theta)$ is the Information matrix for only the ARMA terms and where κ has length (p+q+1) and is computed from the expansion of the ARMA(p,q) lag polynomials, assuming invertibility holds (see Tanaka 1999 for details). We compute the expansions using a simple tailor-made program in Matlab.

We compute the ARMA information matrix using the BHHH Hessian estimator (Berndt, Hall, Hall and Hausman 1974):

$$\hat{\Im}(\hat{\phi}, \hat{\theta}) = \hat{G}'\hat{G} \tag{3.8}$$

where **G** is the true asymptotic matrix of scores and the BHHH estimator uses numerically estimated scores.

The Wald test, however, has a size problem and tends to over-reject the null of *d*=0. Similarly, Likelihood Ratio tests are not effective against spurious long memory. They may be computed, however, to compare the restricted ARFIMA(0,d,0) to the unrestricted ARFIMA (p,d,q) model to evaluate the significance of the short memory parameters.

The wavelet-based estimates of the long memory fractional difference parameter are unaffected by short memory dynamics. This has the advantage of enabling a two-step estimation procedure, which improves the convergence of the likelihood by reducing computational burden. As a consequence, rather than estimate simultaneously all model parameters as does Jensen (1998), we first estimate the long memory parameter *d* and then estimate the short memory ARMA parameters (ϕ,θ) using the correctly fractionally differenced data. Once the short memory parameters are estimated, their Information matrix can be used to obtain the correct standard errors for the long memory parameter *d* and these are presented in Table 3.3 and 3.4.

Fractional differencing is similar to taking differences of a dataset that is originally in levels, as is frequently done with non-stationary time series to enable hypothesis testing. The main difference is that fractional differencing must be computed numerically. To fractionally difference the time series data, the following binomial formula due to Hosking (1981, 1984) is used:

$$\Delta^d Y_t = \sum_{j=1}^{k} \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} Y_{t-j} \tag{3.9}$$

Since working with Gamma functions is unwieldy, Stirling's approximation is used to simplify computations (Abramowitz and Stegun 1972, p. 257):

$$\frac{\Gamma(k+\alpha)}{\Gamma(k+\beta)} = \lim_{k\to\infty} k^{\alpha-\beta}(1+O(k^{-1})) \tag{3.10}$$

A fast numerical solution to this approximation is to use the Gauss hypergeometric function, which can be implemented in the statistical analysis language **R** (Reisen 1999; Fraley et al. 2006):

$$2F1(d,1;1;L)=\left(\frac{1}{1-L}\right)^{d} \tag{3.11}$$

Once the properly fractionally differenced data are obtained, a standard ARMA model is fitted by exact maximum likelihood (e.g. Hamilton 1994, pp. 132-133) in a state-space framework using the Kalman filter and assuming Gaussian innovations. Though it is not explored in this chapter, it would be straightforward to fit instead a GARCH model to the fractionally differenced data which may be more appropriate to describe a volatility variable. For data that are integrated of order $d<0.5$, the underlying process is stationary while it is non-stationary when the data are integrated of order $d\geq0.5$. The differenced data is found to be stationary based on an appropriate ADF-GLS test at the 1% level of significance (Elliott, Rothemberg and Stock 1996).

Model selection of ARMA parameters is based on pairwise Likelihood Ratio tests between a larger unrestricted model and a smaller restricted model, always using the 1% level of significance. The idea is to begin with a very large number of AR and MA lags and using LR tests reduce the number of lags until the tests suggest we have reached a parsimonious representation of the data. For most commodities, the resulting model contains three or four lags for both the AR and MA terms. Akaike and Schwartz Information Criteria are computed but these are generally less reliable because they have been found to over-parameterize the model.

3.10 Exact Wavelet Maximum Likelihood Estimates of Long and Short Memory

Results of the wavelet ML long memory estimation are presented in Tables 3.3 and 3.4. For each commodity futures time series are included the long memory parameter estimate with both naïve and correct standard errors, the AR and MA short memory parameter coefficient estimates with their White covariance robust standard errors, as well as five test results and the interpretation whether long memory is true or spurious.

The estimated long memory parameter *d* is 0.309 for live cattle, 0.320 for lean hogs, 0.321 for soybeans, 0.304 for corn, 0.431 for wheat, 0.436 for canola, 0.258 for coffee, 0.271 for cocoa, 0.290 for cotton, 0.194 for orange juice, and 0.279 for sugar #11. Since H=0.5 + d/2 this implies the Hurst coefficient for these commodities varies between 0.597 and 0.718, all of which suggest significant persistence. To control for the "day of the week effect", the analysis is repeated for corn futures using only Wednesday observations. Estimation results do not differ substantially.

Two sets of standard errors are presented in Tables 3.3 and 3.4. The naïve standard errors assume are computed under the (usually mistaken) assumption that short memory parameters are either zero or have no effect on long memory. They are identical for all commodities because they depend only on the number of observations, which is 4096 in all cases. The correct standard errors are computed from the complete Information matrix which accounts for the bias caused by short memory parameters (Tanaka 1999). The correct standard errors for the fractional difference parameter *d* are hardly affected by the presence short memory (ARMA) terms. For example, the correct standard error for the long memory parameter in soybean futures price volatility is 0.0187, while the naïve standard error is 0.0155.

Five test results are presented for each commodity futures time series. The theory and intuition behind each test is presented in the following section. As expected, the Wald test rejects the null of d=0 for all commodities but cannot distinguish between true and spurious long memory. The second test is also Wald but accounts for the bias caused by short memory. Again, the test rejects the null of *d*=0 for all commodities which suggests the appearance of long memory is not caused by the bias due to short memory. The third and fourth tests are standard KPSS and Phillips-Perron tests applied to the fractionally differenced data and as suggested by Shimotsu (2006) are very useful taken together. As explained in the previous section, for long memory to be true, we must fail to reject the KPSS null hypothesis (d=0) and reject the Phillips-Perron null hypothesis (d=1). The fifth test is a Hausman specification-type test suggested by Ohanission, Russell and Tsay (2005) with a null of true long memory.

3.11 Testing for Spurious Long Memory

The literature on testing between unit roots (or long memory) and structural breaks or level shifts is vast (Banerjee and Urga 2005; Perron 2006). We consider two simple but effective tests by Shimotsu (2006) and by Ohanissian, Russell and Tsay (2004).

Shimotsu (2006) suggests three useful tests based on the time aggregation invariance property of true long memory processes shown by Chambers (1998). The test selected for this chapter consists of fractionally differencing the data (using a robust estimate of *d*) and then subjecting these tests to the well-known KPSS test for a null of stationarity and Phillips-Perron test for a null of non-stationarity (unit root).

Three alternative data generating processes that are known to generate spurious long memory are considered: (i) a mean plus noise process, (ii) Engle and Smith's (1999) stochastic permanent break model, and (iii) a Markov-switching model.

Table 3.3: Log-range volatility ARFIMA (p,d,q) model estimates, standard errors and hypothesis test results, for CME, CBOT, KCBOT and WCE commodities

Commodity futures contract	CBOT corn	CBOT soybeans	CME lean hogs	CME live cattle	KCBOT wheat	WCE canola
Long memory *d*	0.304	0.321	0.320	0.309	0.431	0.436
Correct standard error for *d*	0.0187	0.0187	0.0187	0.0187	0.0188	0.0186
Naïve standard error for *d*	0.0155	0.0155	0.0155	0.0155	0.0155	0.0155
Intercept	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
AR1	-0.002 (0.202)	0.409 (0.111)	0.148 (0.176)	0.767 (0.073)	0.564 (0.036)	-0.630 (0.062)
AR2	0.634 (0.153)	-0.014 (0.126)	-0.455 (0.219)	0.918 (0.087)	0.083 (0.021)	0.336 (0.056)
AR3	0.190 (0.151)	0.649 (0.125)	-0.236 (0.104)	-0.65 (0.059)		
AR4	0.093 (0.028)	-0.125 (0.065)	0.536 (0.13)	-0.063 (0.024)		
AR5			0.042 (0.228)			
AR6			0.615 (0.126)			
AR7			0.140 (0.018)			
MA1	-0.146 (0.203)	-0.631 (0.109)	-0.265 (0.176)	-0.931 (0.07)	-0.834 (0.032)	0.392 (0.056)
MA2	-0.661 (0.169)	0.09 (0.136)	0.463 (0.239)	-0.805 (0.094)		-0.539 (0.048)
MA3	-0.073 (0.015)	-0.682 (0.130)	0.204 (0.120)	0.774 (0.067)		
MA4		0.341 (0.07)	-0.531 (0.134)			
MA5			0.044 (0.244)			
MA6			-0.598 (0.148)			

Seasonal (sinusoidal) coefficients are very small and not significantly different from zero, they are therefore omitted.

One approach that holds much promise but is not considered here follows the literature on continuous-time asset pricing models, which suggests that jump-diffusion models (Merton 1980) are more plausible on theoretical grounds than are long memory models (e.g. Granger 2003, 2005). Jump-diffusion models are increasingly used and particularly useful to link futures with options-on-futures (see e.g. Koekebakker and Lien 2004; Saphores, Khalaf and Pelletier 2002).

Table 3.3 (continued).

Commodity futures contract	CBOT corn	CBOT soybeans	CME lean hogs	CME live cattle	KCBOT wheat	WCE canola
Log-likelihood	14906.4	15144.8	14979.1	16885.3	8112.46	15035.5
Wald test Ho:d=0, model I(d)	24.95	26.35	26.27	25.36	23.15	35.79
Wald test Ho: d=0, model ARFIMA (p,d,q)	24.86	26.25	26.17	25.27	23.06	35.66
Shimotsu's adjusted KPSS test, Ho: d=0	0.14*	0.61***	0.46***	1.04***	0.04	0.02
Shimotsu's Phillips-Perron Z test, Ho: d=1	-1.46	-0.30	-0.60	0.52	-3.63**	-4.20***
Ohanissian-Russell-Tsay test, Ho: true long memory	4.42**	4.55**	4.06**	3.92**	2.82*	3.10*
Long memory true?	**No(#)**	**No**	**No**	**No**	**Yes**	**Yes**

Notes: Critical test values (exact values were computed and used in the analysis but approximate values are included here for convenience, source: Shimotsu 2006 Table 2): adjusted KPSS test 0.135 (10%), 0.17 (5%), 0.26 (1%) and adjusted Phillips-Perron Z test -3.09 (10%), -3.36 (5%), -3.90 (1%); and Chi-Square(1) for Ohanissian-Russell-Tsay Wald-type test 2.706 (10%), 3.841 (5%) and 6.635 (1%).
(#): Test results for corn futures are weaker: Shimotsu’s test null can only be rejected at the 10% level of significance.

However, these models are substantially more difficult to work with and hypothesis testing is complicated by the presence of nuisance parameters that must be dealt with through simulation methods (e.g. Khalaf, Saphores and Bilodeau 2003).

If the true long memory (fractional integration) model and the alternative (e.g. Markov-switching) model are nested, with or without ARMA or GARCH short memory parameters, Likelihood Ratio tests could be used to evaluate claims of spurious long memory. In the absence of a clear alternative model specification, or if the two models are non-nested as is the case here, Shimotsu’s test is appropriate and is therefore used.

Table 3.4: Log-range volatility ARFIMA (p,d,q) model estimates, standard errors and hypothesis test results, for NYBOT commodities

Commodity futures contract	NYBOT cocoa	NYBOT coffee	NYBOT FCOJ	NYBOT cotton	NYBOT sugar#11
Long memory *d*	0.271	0.258	0.194	0.29	0.279
Correct standard error for *d*	0.0190	0.0191	0.0192	0.0190	0.0191
Naïve standard error for *d*	0.0155	0.0155	0.0155	0.0155	0.0155
Intercept	<0.001	<0.001	<0.001	<0.001	<0.001
AR1	0.624	0.001	0.308	0.337	0.529
	(0.145)	(0.110)	(0.541)	(0.033)	(0.073)
AR2	0.649	0.067	0.677	0.429	-0.109
	(0.383)	(0.059)	(0.539)	(0.037)	(0.031)
AR3	-0.058	0.141	-0.848	0.967	
	(0.236)	(0.031)	(0.045)	(0.032)	
AR4	-0.367	0.365	-0.105	-0.425	
	(0.172)	(0.032)	(0.017)	(0.076)	
AR5		0.634			
		(0.062)			
AR6		-0.526			
		(0.113)			
AR7					
MA1	-0.717	-0.065	-0.309	-0.442	-0.649
	(0.158)	(0.104)	(0.541)	(0.031)	(0.067)
MA2	-0.611	-0.065	-0.672	-0.412	0.104
	(0.400)	(0.057)	(0.526)	(0.037)	(0.031)
MA3	0.136	-0.122	0.013	0.887	-0.970
	(0.256)	(0.036)	(0.018)	(0.050)	(0.032)
MA4	0.352	-0.344		0.564	
	(0.184)	(0.033)		(0.069)	
MA5		-0.572			
		(0.058)			
MA6		0.592			
		(0.104)			
Seasonal (sinusoidal) coefficients are very small and not significantly different from zero, they are therefore omitted.					

To evaluate the hypothesis of true long memory against unknown forms of spurious long memory, adjusted KPSS and Phillips-Perron tests are applied to the fractionally differenced and appropriately demeaned data. Critical test values are provided by Shimotsu (2006, Table 2). Suppose the data generating process appears to be fractionally integrated I(d) but is in fact I(1), e.g. a unit root, mean plus noise or stochastic break process. Then taking the d^{th} difference will result in a new process that is I(1-d) where (1-d)≠0 while we believe it is I(0).

Table 3.4 (continued).

Commodity futures contract	NYBOT cocoa	NYBOT coffee	NYBOT FCOJ	NYBOT cotton	NYBOT sugar#11
Log-likelihood	12902.4	11677.9	12427.04	13733.3	11791.5
Wald test Ho:d=0, model I(d)	22.24	21.18	15.92	23.80	22.90
Wald test Ho: d=0, model ARFIMA (p,d,q)	22.16	21.10	15.87	23.72	22.82
Shimotsu's adjusted KPSS test, Ho: d=0	0.48***	0.75***	0.76***	0.40***	0.28***
Shimotsu's Phillips-Perron Z test, Ho: d=1	-0.61	-0.58	-0.61	-0.80	-0.95
Ohanissian-Russell-Tsay test, Ho: true long memory	3.80*	5.21**	3.91**	4.50**	4.27**
Long memory true?	**No**	**No**	**No**	**No**	**No**

Notes: Critical test values (approximate, source: Shimotsu 2006 Table 2): adjusted KPSS test 0.135 (10%), 0.17 (5%), 0.26 (1%) and adjusted Phillips-Perron Z test -3.09 (10%), -3.36 (5%), -3.9 (1%); and Chi-Square(1) for Ohanissian-Russell-Tsay Wald-type test 2.706 (10%), 3.841 (5%) and 6.635 (1%).

The KPSS test will correctly reject the null that this new, fractionally differenced process is I(0) but the Phillips-Perron test, which has low power, will fail to reject the null that the new process is I(1). We therefore learn that the process is not true long memory. If the process is true long memory I(d), then the fractionally differenced process will be I(0), the KPSS test will correctly fail to reject its null of I(0) and the Phillips-Perron will also correctly reject its null of I(1). We then confirm the process is true long memory.

The results, presented in Table 3.3 and 3.4, are summarized as follows. There is strong evidence that long memory is only true for two out of eleven commodities, namely wheat and canola futures. There is strong evidence of spurious long memory for eight commodities and weaker but reasonable evidence for corn futures. For most commodities, therefore, the results suggest the data are better explained by a short memory but nearly non-stationary model such as Engle and Smith's (1999) stochastic break process or a Markov-switching process.

Ohanissian, Russell and Tsay (2005) suggest a simple test of spurious long memory that is based on Hausman's (1978) result that an efficient estimator must have zero asymptotic covariance with any other consistent, asymptotically normal estimator. Under the null of true long memory, the covariance of two estimates of long memory for the same data but aggregated two different ways will asymptotically equal the variance of the long memory estimator for the less-aggregated data.

The test is however limited because it relies on the GPH long memory estimator, which as explained earlier, is markedly inferior to wavelet and Whittle-type estimators. Moreover, the Ohanissian et al. test is best suited for large datasets such as ultra high-frequency financial data ("tick" observations), and for optimal size and power requires a large number of aggregation levels is high. This test has a Chi-Squared(*M*) asymptotic distribution where *M* is the number of aggregation levels.

In this case, the less aggregated data are the daily observations, and each aggregation level *m* results in a number of ordinates *l(m)* generally chosen to be $(n/m)^{1/2}$ then the test asymptotically converges to:

$$\lim_{n\to\infty} 4l^{(m_i)}\left(Cov(\hat{d}^{(m_i)},\hat{d}^{(m_j)})-Var(\hat{d}^{(m_i)})\right)=o(1) \qquad (3.12)$$

For our data, n=4096, m=8, $l(m)$=22.627. Computation of the test statistic is detailed in Ohanissian et al. (2005).

The results in Table 3.3 and 3.4 imply rejection of the null of true long memory for all commodities at the 5% level of significance except for wheat and canola as well as rejection of the null at the 10% level for cocoa. These test results agree with the Shimotsu test approach and provide strong evidence that Kansas City wheat and Winnipeg canola futures volatility are characterized by true long memory, while for all other commodities the long memory is spurious.

In conclusion, since the wavelet-based estimator is robust to the presence of short memory dynamics, findings of spurious memory for most of the commodities suggests other dynamics must be responsible for the illusion of persistence in volatility. One leading candidate addressed in the next section is a Markov-switching model that generates spurious long memory.

3.12 An Alternative Model of Futures Price Volatility

The evidence only weakly supports rejecting the true long memory model for the CBOT corn futures contract. We consider estimating for these data an alternative model, a Markov-switching process that has been found to generate spurious long memory (see e.g. Hamilton 1994; Shimotsu 2006). The idea is to obtain state-dependent means, e.g. "low" and "high" volatility states, along with the probabilities associated with each state. To determine whether the true long memory or Markov-switching model better describes the data, a non-nested test can be constructed following Pesaran and Ulloa (2006) and Gourieroux and Monfort (1994).

The model is a simple two-state Markov-switching process augmented by AR and MA terms (Hamilton 1990, 1994), defined as follows: $\phi(L)(Y_t - \mu) \sim N(\xi_0, \sigma^2)$ under state 0 and $\phi(L)(Y_t - \mu) \sim N(\xi_1, \sigma^2)$ under state 1. The Markov transition probability matrix determines from one time period to the next what states occur. Given state 0 at time *t*, the probability of staying in state 0 the next period is defined by p_{00} which necessarily lies between 0 and 1. It follows that the probability of going from state 0 to state 1 is p_{01} and so forth. Estimating a two-state Markov-switching model therefore produces values for the four transition matrix probabilities as well as the two state means, which may be interpreted as "low" and "high" volatility states in our case.

As suggested by Hamilton (1990, 1994), the EM algorithm is used to help the likelihood converge (Dempster 1977). This algorithm will improve the likelihood with every step but is not guaranteed to converge to the best estimates (Wornell and Oppenheim 1992). The resulting estimates are state dependent means ξ_0=0.0133 and ξ_1 =0.0355 and a Markovian transition matrix: {p_{00}=0.955, p_{01}=0.045, p_{10}=0.608 and p_{11}=0.392}.

This means the daily price volatility process, if in a low volatility state, is more than 95% likely to remain in this low volatility state, but if in a high volatility state, is about 60% likely to switch to the low volatility state. The ARMA parameters are statistically significant and are: ϕ=(0.676, 0.701, -0.394) and θ=(-0.587, -0.652, 0.327) with White covariance robust standard errors: se(ϕ)=(0.0320, 0.0133, 0.0236) and se(θ)=(0.0314, 0.0134, 0.0177).

The difficulty of implementing a non-nested hypothesis test in this context concerns how to apply the encompassing principle (Deaton 1982). That is, since neither model can be written as a special case of the other, the test relies on defining a third model that will serve as the alternative for two tests, each of which involves only one of the two estimated models. The test is said to be non-informative because if we reject both nulls or fail to reject both nulls, we still cannot decide which of the two estimated models is more plausible. Construction of this test is left for an extension of this work.

3.13 Conclusion

This chapter investigates one important cause and consequence of bias in commodity futures option pricing and contributes to the active literature on the robust estimation of long memory in commodity futures price volatility using a novel empirical strategy that also enables the computation of efficient standard errors for the long memory parameter jointly with the unbiased estimation of short memory parameters.

There is evidence of long memory for all commodity futures contracts in the log-range volatility of prices. The estimates are, however, smaller in magnitude than those found in previous research and, based on the evidence from carefully designed tests, the results appear to be spurious for all commodity futures except for Kansas City Board of Trade wheat and Winnipeg Commodity Exchange canola. Further support for this interpretation comes from results from a wavelet-based semi-parametric estimation of long memory. We find that for all commodities the Hurst parameter H is not significantly different from 0.5, which implies that increments of the data generating process are consistent with white noise and not long memory (fractional white noise). Although the time series model used is relatively simple, Lordkipandize (2004) estimated a much larger, stochastic volatility model of commodity derivative prices and concluded that once breaks and seasonality are properly accounted for, the effect of long memory is inconsequential.

The results are weaker for Chicago Board of Trade corn futures, so a Markov-switching process that generates spurious long memory is estimated and found to fit the data well. Although standard asymptotic nests cannot be applied, it is suggested a non-nested hypothesis test could be constructed to evaluate the null of true long memory against regime-switching.

The implication of this chapter's research is that true long memory is unlikely to be a good description of the data generating process underlying agricultural commodity futures prices and volatility. Since spurious long memory is often found, however, models of commodity futures should be selected to reproduce the illusion of long memory that is observed in the data. Many such candidate models exist, including stochastic break, regime-switching, and stochastic unit root.

Option pricing in agribusiness is therefore unlikely to gain much by using fractional Brownian motion and fractional noise as building blocks instead of relying on the classic Black-Scholes-Merton model. The results in this chapter do, however, provide support for the jump-diffusion models of option pricing and related econometric procedures, for which the volume of research has greatly expanded in recent years.

CHAPTER 4

REVEALING THE IMPACT OF INDEX TRADERS ON COMMODITY FUTURES MARKETS

4.1 Introduction

This chapter presents original results on two timely questions on the relationship between trader type heterogeneity and futures price volatility. First, should the Commodity Futures Trading Commission make permanent its pilot project whereby positions of Index Traders (defined later in this section) are reported separately from other large traders? Second, has the time horizon of trading (short run or long run) changed over the last two decades across commodity markets? An approximate measure of the impact of Index Traders on commodity futures price volatility is revealed by estimating the long-run trade volume process using an application of wavelet transforms. Similarly, using wavelet transforms allows us to obtain, for a given commodity and time period, the approximate distribution of trade volume across time horizons, from which an interpretation of trader types can be made.

The Commodity Futures Trading Commission in Washington D.C. is a federally-mandated regulatory agency responsible for helping commodity derivatives markets run smoothly, free of market cornering attempts and insider trading. It also produces, since 1924, widely read reports on the Commitment of Traders (CoT). These CoT reports, published weekly for a number of years, provide information on the futures and options positions of large traders in all markets regulated by the CFTC.

In recent years, the demand for commodity derivatives has substantially increased, as commodities are now considered a vital class of assets to help diversity a financial investment portfolio (Gorton and Rouwenhorst 2006). The recent and growing participation by a non-traditional class of large traders, defined by the CFTC as Index Traders, explains much of the expansion in commodity derivatives. Index Traders consist of large investment funds such as commodity pools, pension funds and swaps dealers that are involved neither in production, delivery or ownership of the underlying asset. The CFTC evaluates that: "On the Chicago exchanges, for example, the [Index] funds make up 47 percent of long-term contracts for live hog futures, 40 percent in wheat, 36 percent in live cattle and 21 percent in corn" (The New York Times, January 19, 2007).

In 2006, the CFTC conducted a large-scale survey to learn about the perceptions of commodity futures market participants regarding Index Traders. The outcome was the largest number of responses ever for a CFTC survey. Most respondents were concerned that Index Traders (also called index funds) are responsible for increasing market volatility, with consequences for price volatility along the distribution chain. As of January 2007 and on a two-year pilot basis, the CFTC will publish a Supplemental Commitment of Traders report for twelve selected major agricultural commodities. This supplementary report defines and analyzes Index Traders separately from Commercials and Non-Commercials.

Two main questions on the relationship between futures trade volume and price volatility are asked and answered in this chapter. First, have Index Traders caused greater price volatility through an increased volume of trade? Second, how has the time horizon of commodity futures trading (e.g. short run, long run) changed in the last two decades? The results are made possible by a wavelet transform decomposition of the time series data into mutually orthogonal "artificial" time series.

Each artificial time series corresponds to the fraction of the original price series that is explained by a specific time horizon, for example, weekly variation. As a consequence, we remove variation associated with time horizons smaller than two weeks. The selection of a two-week threshold beneath which all trade volume fluctuations are filtered out is arbitrary to some extent but is supported by the available empirical evidence (Haigh, Hranaiova and Overdahl 2005). Indeed, confidential CFTC position-level data show that Index Funds seldom engage in short-term trading.

In summary, there are two findings in this chapter. (1) Index Traders may have caused greater price volatility in the only two non-storable commodity futures markets considered (live/lean hogs and live cattle contracts), but not in the storable commodity markets (grains). The empirical results may prove timely and of directly relevance to the CFTC's pilot project on Index Traders. In addition, the methodology may prove useful to evaluate the impact of specific trader types in futures markets for which no position-level reports are produced or available. (2) The distribution across time horizons of trade volume reveals that storable commodity market participants trade at a more distant horizon than do non-storable commodity market participants, and also that in recent years intermediate time horizons have gained in importance, which may be well explained by the rising participation of Index Traders.

4.2 Index Funds and the Commitment of Traders Report

Participation in futures markets is traditionally explained in terms of hedging (largely commercial) and speculation (largely non-commercial) motives. It is well understood, however, that large commercial institutions are sometimes involved in speculation while non-commercials may hedge. Commitment of Traders reports classify the positions of large traders into commercials and non-commercials and are used, for example, by hedgers to evaluate future demand and by speculators to design technical trading rules (Park and Irwin 2006, 2007; Roberts 2005). A technical trader may try to assess how long a bullish uptrend will last by looking at how the large noncommercial traders bid futures prices forward such that there is a premium over normal carrying charges.

The rising importance of Index Traders in commodity markets can be explained by the business cycle behavior of commodity prices as an asset class over holding periods of one month to five years (Erb and Harvey 2006; Gorton and Rouwenhorst 2006). Commodity futures are seen as highly desirable because they are positively correlated with inflation (actual and unexpected) and negatively correlated with stock and bond returns. A report by Ibbotson Associates (2006), for example, finds that "commodities have low correlations to traditional stocks and bonds, produce high returns, hedge against inflation, and provide diversification through superior returns when they are needed most" (p.iii).

Index Traders are not allowed to physically own the underlying commodities (CFTC 2007). The Index Traders category contains swap dealers, who hold long futures positions to hedge short OTC commodity index risk against long positions taken by institutional traders such as pension funds. Metals and energy commodities are not included because there exist for these many alternative exchanges that are not regulated by the CFTC, such as Over-the-Counter markets and derivative instruments. It would be difficult to get meaningful results from their inclusion in the pilot project.

Before carrying out its pilot project, the CFTC collected thousands of survey responses on questions about the usefulness of its weekly reports and perceptions about the impact of Index Traders. To explain why most market participants are concerned by the impact of Index Traders, consider the following example. A farmer wants to lock in a crop price using a futures hedge. He uses the CoT reports to assess expected consumer demand based on commercial positions. If commercials increase their long futures positions, the farmer believes it reflects sales of cash commodities and suggests a strong demand for cash grain. In that case the farmer postpones the short hedge in anticipation of a bullish market. But suppose instead the increased long open interest reflects swap and pension fund institutional trader positions. Then the farmer waits but finds that demand does not increase, so he must form a hedge with a less favorable basis. Index Traders may therefore lead market participants to wrongly infer greater export activity and end use buys.

A second concern is the risk of a sudden and large exit of the (mainly long) index fund positions in commodity futures if one day in the future commodities cease to be as desirable as asset class as they are today. Such a risk is, however, unlikely (Gorton and Rouwenhorst 2006).

In contrast, the International Swaps and Derivatives Association opposed the plan to create a new reportable class for Index Traders. It argued that Index Trader long positions do not increase volatility because they are passive, predictable and instead contribute to increased liquidity. On the contrary, it claimed, disclosing Index Trader positions would encourage speculation and increase volatility, because rolling index positions are recurring and can be anticipated.

The empirical evidence tends to support the Swaps and Derivatives Association's claims that Index Traders do not increase market volatility. For example, Chatrath and

Song (1999) find that both the number and the commitment of speculators are negatively correlated with the underlying cash market volatility. On the contrary, it is hedger positions that are found to be positively correlated with market volatility. Irwin and Yoshimaru (1999) find that managed commodity pools do not appear to contribute to market volatility.

Most studies, of necessity, use aggregated data from the CFTC's weekly Commitment of Traders reports. Recent work by Haigh, Hranaiova and Overdahl (2005, 2007) and by Haigh, Harris, Overdahl and Robe (2007), however, use confidential position data at the level of the participants and examine the direction of causal relationships to evaluate the hypothesis that price changes are caused by large trader speculation (i.e. changes in futures positions). Their results for oil and natural gas futures show that, on the contrary, large traders provide liquidity for the markets and change positions less often than do other traders.

4.3 Futures Market Volatility and the Long-Run Volume of Trade

The distribution of trader type heterogeneity and whether trading causes volatility has attracted a large volume of research (French and Roll 1986). At least since Friedman (1953), it has been suggested that while informed traders should reduce volatility, uninformed traders are likely to increase volatility. Avramov, Chordia and Goyal (2006) show, for example, that a model with both informed (contrarian) traders and uninformed (herding) traders explains well observed empirical patterns of volatility, including asymmetry, at daily and lower frequencies and is far more robust to model specification issues than are alternative explanations such as the leverage effect (Black 1976) or time-varying expected returns.

A useful approach is to consider how trade volume affects volatility, because a trader's level of information is proportional to the size of his or her trades (Easley and O'Hara 1987). Copeland (1976) suggested the role of volume as an information arrival process and proposed a model of sequential information. Blume, Easley and O'Hara (1994) develop a model of the informational role of volume based on differing qualities of signals.

It has long been a known stylized fact that "a small (large) volume is usually accompanied by a price fall (increase)" (Ying 1966). At least since Godfrey, Granger and Morgenstern (1964), researchers have examined how trade volume contains information on the unknown process that drives asset (e.g. futures) prices. The relationship of volume with different functions of price has been studied, including price changes, absolute or squared price changes, or the direction of price changes. Relatively little work has been done, however, to better understand the trade volume process itself (Lo and Wang 2001). Yet the relationship between trade volume and price volatility remains an active area of research, as for example, Pan and Poteshman (2006) show that option pricing volume contains useful information about future stock prices.

In an early survey of the literature, Karpoff (1987) gives four reasons for the importance of the price-volume relation. These are: to learn about the information structure of financial markets, to improve the quality of event studies that use both price and volume data, to better estimate the empirical joint distribution of asset prices, and lastly to examine implications for futures prices. Most of the early studies characterized volume as an exogenous variable. Lamoureux and Lastrapes (1990), for example, find that a volume variable is significant when included in a GARCH model of price volatility. Much subsequent work has gone into modeling the endogenous, joint determination of volume and price changes (Cornell 1981; Foster and Visnawathan 1995; Grammatikos and Saunders 1986; Lamoureux and Lastrapes 1994).

The mixture of distributions hypothesis has been proposed for price returns and volume and has led to a large literature on the joint process estimation (Clark 1973; Epps and Epps 1976; Tauchen and Pitts 1983). Andersen (1996) finds that volatility persistence is substantially reduced in a model where volume and returns are jointly estimated. Bollerslev and Jubinski (1999) find that the volume-volatility relationship associated with a “news arrival” process is characterized by long memory and in particular, that the hyperbolic rate of decay described by long memory is the same for both variables. Wang and Yau (2000) examine the two most actively traded financial and metals futures contracts and estimate a three-equation structural model of volume, price volatility and bid-ask spread (computed from CFTC intraday data adjusted for microstructure effects).

4.4 Data, Estimation and Identification Strategies

This chapter makes two contributions to the literature on trader heterogeneity, price volatility and the volume of trade. First, the relationship between trader type and time horizon is used to help answer the question whether the increasing participation of large Index Traders in commodity futures markets has increased the volatility of futures prices through increased long-run trade volume. A joint model of trade volume and price volatility is considered, with contemporaneous and lagged volume and volatility variables as the regressors. To remove short-run futures trade volume variation from the data, a Discrete Wavelet Transform is applied to the data, which produces wavelet coefficients defined over time and timescale (time horizon), as described in Chapter 2 (e.g. section 2.7-2.8). All wavelet coefficients associated with time horizons of two weeks and less are set to 0, then the appropriate Inverse Discrete Wavelet Transform is applied to the wavelet coefficients. This results in a new trade volume time series of the same length as the original data, but which excludes all variation caused by the short run (i.e., time horizons of two weeks or less).

Second, to determine whether commodity futures trading has focused on the short-run or the long-run over the years, a wavelet transform-based method is applied to trade volume data for a dozen leading agricultural commodities. This approach provides a revealed measure of the time horizon of trading and, indirectly, an aggregate measure of trader heterogeneity and proportions of trader types in different markets over time. Lastly, to evaluate whether the volatility of trade volume has experienced structural breaks, two tests are used. First, a wavelet-based Monte Carlo is conducted to detect change-points over the entire time period and recover both the precise date of the break and the time horizon at which it occurred. This test has the advantage of being robust to the presence of long memory, which appears to characterize trade volume time series data (Lobato and Velasco 2000). Second, a sup-Wald type test in the Andrews-Ploberger-Hansen class is applied to a wavelet-based linear regression of daily volume differences over variations due to different time horizons. This test provides direct evidence of changes in the influence of specific time horizons.

The data consist of business daily observations on settlement price and trade volume (total from all maturities) from the Chicago Board of Trade soybeans and corn futures contracts, Kansas City Board of Trade wheat futures contracts, Winnipeg Commodity Exchange canola futures contracts, and Chicago Mercantile Exchange live cattle and lean hogs futures contracts. This allows us to consider differences between storable (grain) and non-storable (meat) commodities. Volume data for contracts traded at the Chicago and Kansas City Boards of Trade (corn, soybeans, wheat) are adjusted for consistency because on January 1st 1998, the reported measurement unit changed from 1000 bushels to one contract (5000 bushels).

Commodities may be categorized as non-storable, storable with large inventories ("overhangs") and storable with small inventories. These categories also lead to testable predictions of futures forecasting accuracy. Futures provide an unbiased forecasting measure for non-storable commodities (as well as other instruments such as Federal Funds). For storable commodities with large inventories, futures prices incorporate a cost of carry (storage plus interest), and perhaps a convenience yield (this need not be the case, however, see e.g., Brennan, Williams and Wright 1997). Storable commodities with small inventories can be described by two cases. If futures prices are higher than spot prices ("contango") then the analysis follows the large-inventory case. But if futures prices are lower than spot prices ("backwardation"), we can use apply the analysis as if it were non-storable.

Figure 4.1 shows the evolution of soybean futures trade volume over the time period 1988-2004. A recurring pattern can be identified, where volume rises and falls over the lifetime of a single maturity. Figure 4.2 compares actual with wavelet-filtered trade volume over a short time period, 3/23/1988 to 5/12/1988, also for the soybeans futures contract. The wavelet-filtered trade volume removes all variation that is explained by time horizons of less than two weeks. There is a visible difference between the actual and wavelet-filtered data.

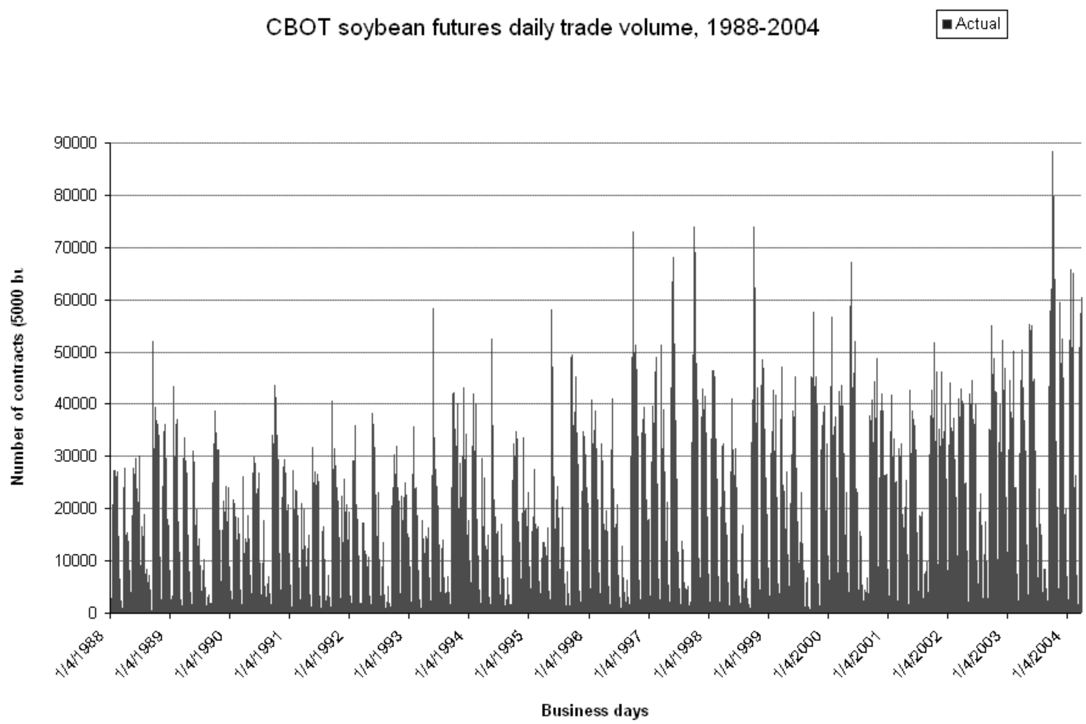


Figure 4.1: Daily trade volume, Chicago Board of Trade soybean futures 2/1988-1/2005.

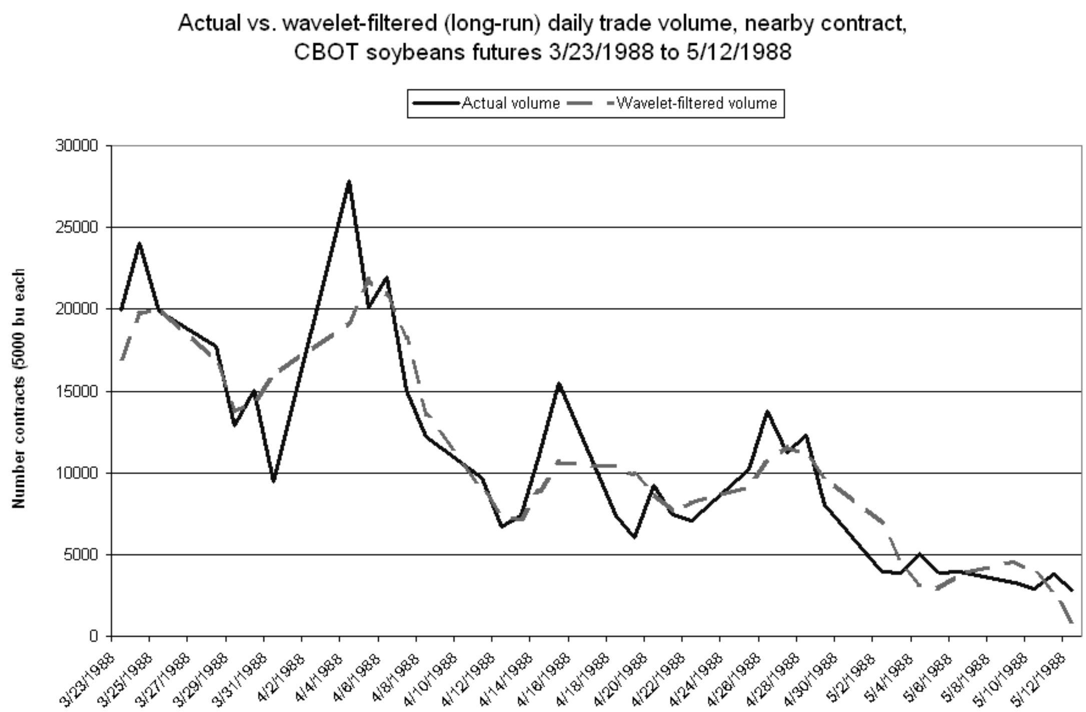


Figure 4.2: Actual and wavelet-filtered (i.e., no short run variation) daily trade volume, Chicago Board of Trade soybean futures, 3/23/1988 to 5/12/1988.

Instead of the commonly-used but noisy logreturn volatility measure, a daily log-range volatility measure is computed following Alizadeh, Brandt and Diebold (2002) and Yang and Zhang (2002):

$$\mathrm{h_t} = \ln(\sup F(t,T) - \inf F(t,T)) \tag{4.1}$$

The daily log-range measure of volatility is asymptotically superior to absolute or squared logreturns and appropriate given the large number of observations used. The measure of volume used is the natural logarithm of daily total volume for all maturities traded on a given day, expressed in thousands of contracts (each contract equals 5000 bushels).

4.5 Endogenously Biased Model and Hausman-Wu Test

Three estimation procedures are applied to data for all commodities under scrutiny to determine whether trade volume explains price volatility, and specifically whether Index Traders have an adverse influence. The first approach is endogenously biased and used to provide benchmark estimates. The second and third are unbiased, and the third moreover uses wavelet filtering to focus only on the likely impact of Index Traders.

To present the different approaches, we begin with Chicago Board of Trade corn futures data, and provide the results for the other commodities in the following section. The first model specification is an autoregressive moving average with exogenous term (ARMAX) estimated by maximum likelihood:

$$\phi(L)h_t = \theta(L)\varepsilon_t + \ln V_t \tag{4.2}$$

where the price volatility is defined as the daily log range h_t , $\phi(L)$ and $\theta(L)$ are the AR and MA lag polynomials, and $\ln(V_t)$ is the natural logarithm of the daily trade volume time series. ARMA lag length is selected on the basis of Likelihood Ratio tests, comparing pairwise a larger unrestricted model with a smaller restricted model and considering one to twenty-one lags (i.e., one month in business days). Akaike and Schwarz Information Criteria are also computed for consistency.

Similarly, we can estimate an ARMAX model of the effect of volatility on volume:

$$\phi(L)\ln V_t = \theta(L)\varepsilon_t + h_t \tag{4.3}$$

To improve the computational convergence of the likelihood function, volatility h_t is expressed as one hundred times the log-range and volume V_t is expressed in thousands of contracts, where each contract is, e.g., 5000 bushels of corn. Taking natural logarithms of all variables, in addition to making variances more symmetrical, also conveniently allows the estimated coefficients to be interpreted as elasticities. For both the price and volume data, the first and last ten observations are deleted to avoid possible boundary effects caused by the data transformation.

Before estimating the ARMAX model, diagnostic tests are computed to establish the stationarity of the sample data. Since unit root tests are well-known to have low power (Cochrane 1987) and as the existence of both a time trend (deterministic) and a unit root (stochastic) is unlikely in economic time series (Perron 1988), a two-step test procedure is used. First is computed a unit root test assuming no time trend (Augmented Dickey-Fuller and Phillips-Perron). If we fail to reject the null of a unit root, we compute a t-test of the regression of the differenced series on an intercept.

This evaluates the presence of a time trend (drift term). If we reject the null of a unit root, we are confident the data are stationary. We can evaluate the presence of a deterministic trend by computing a t-test of the data in levels on a time trend vector t=(1, 2, 3, 4, …, T}. The ADF test on the trade volume data a value of -3.11, which falls between the 1% and 5% critical values of -3.46 and -2.87. We reject the null hypothesis of a unit root at the 5% level of significance.

Estimates using CBOT corn futures data for the baseline equations and obtained independently are:

$$\mathrm{h_t} = -0.79 + 0.74 \ln V_t - 0.376 \ln V_{t-1} + 0.422 h_{t-1} \quad (4.4)$$

$$\mathrm{lnV_t} = 1.167 + 0.298 h_t + 0.385 \ln V_{t-1} + 0.243 \ln V_{t-2} \quad (4.5)$$

The results suggest that daily price volatility is serially correlated and affected positively by contemporaneous volume but negatively by lagged volume. All coefficient estimates are individually statistically significant at the 1% level. Standard errors are provided in Table 4.1. The adjusted R^2 are 0.428 for the volatility equation and 0.606 for the volume equation. The ARMAX values are presented as naïve baseline estimates against which are compared the unbiased Two Stage Least Squares estimates in the next section. Standard errors are computed using a Newey-West heteroskedasticity and autocorrelation consistent (HAC) covariance.

Volatility dynamics are better captured by the ARCH-GARCH family of models (Engle 1982; Bollerslev 1986). A large number of studies use GARCH models to describe the volume-price volatility relationship, based on the mixture of distributions hypothesis that provides an explanation why price returns are heteroskedastic

(Tauchen and Pitts 1983; Lamoureux and Lastrapes 1990). Nelson (1992) shows that even if significantly mis-specified, GARCH models may provide an acceptable fit to the data and to short-term forecasts. Both GARCH and ARMAX estimates are, however, biased in this model because theory suggests the volatility and volume variables are jointly determined and therefore endogenous to each other. Moreover, including volume as an exogenous variable in a GARCH model is likely to introduce a simultaneity bias.

To verify empirically the endogeneity bias between contemporaneous volume and price volatility, a Hausman-Wu test is computed for both the volume and the volatility equations. The test statistic has a null hypothesis of no correlation between the potentially endogenous regressor and the error term and is distributed as F with degrees of freedom being the number of restrictions and the adjusted number of observations. For trade volume, the statistic of 188.28 is much larger than the value of the F test statistic which is 6.63. Likewise, for price volatility, the statistic of 27.89 is larger than 6.63. Therefore, the null hypothesis of no endogeneity bias is rejected for both equations and we may conclude that joint estimation is preferable.

4.6 Full Sample Unbiased Structural Model Estimates

To solve the problem of simultaneity bias and the endogeneity of volume and volatility we use once-lagged volume instead of contemporaneous volume. This model better describes information arrival flows in the Copeland (1976) sense rather than the actual price-volume relationship. A Generalized Method of Moments framework such as the one used by Foster (1995) for oil futures contracts is better suited for this problem to recover the structural model parameters.

The structural model is the following:

$$\ln V_t = \alpha_0 + \alpha_1 h_t + \alpha_2 \ln V_{t-1} + \alpha_3 \ln V_{t-2} + \nu_{1,t} \tag{4.6}$$

$$h_t = \gamma_0 + \gamma_1 \ln V_t + \gamma_2 \ln V_{t-1} + \gamma_3 h_{t-1} + \nu_{2,t} \tag{4.7}$$

where $\ln V_t$ is the natural logarithm of trade volume (normalized as thousands of contracts), h_t is price volatility measured as the natural logarithm of the daily price range, and $\nu_{1,t}$ and $\nu_{2,t}$ are assumed to be mean zero white noise innovations. In this structural model volume and volatility are endogenous and OLS estimates are inconsistent. The instruments needed for GMM estimation are lags of both variables. As there is one excluded variable in each equation, one lag may be used in each equation and then the model is exactly identified such that unique, consistent estimates of all parameters can be recovered.

The moment conditions reduce to a 2SLS problem (see Hamilton 1994, pp. 233-247 for time series 2SLS estimation) where we solve the reduced form equations and recover the structural parameters, which are then compared to the biased, benchmark OLS estimates. A greater number of instruments (GMM moment conditions) could be used to provide over-identifying restrictions. The simulation results of Tauchen (1986) and Kocherlakota (1990) show, however, that a parsimonious selection of instruments is often preferable, particularly in a time series context where lagged variables provide a very large number of potential instruments that are likely to be weak. The potential weakness of instruments is evaluated using the Hausman-Wu specification test.

By adding two instruments, the twice-lagged volatility in the volume structural equation and the twice-lagged volume in the volatility structural equation, the reduced form model has a total of ten equations in ten coefficients, equivalently, just identification. In this case, the 2SLS estimator is simply the Instrumental Variables (IV) estimator, which is consistent but does not provide heteroskedasticity and autocorrelation consistent (HAC) standard errors. The GMM literature suggests various HAC standard errors, including the Newey-West (1987, 1991) correction which we use. The simple decision rule for lag truncation is to choose a number of lags equal to 0.45 $T^{1/3}$ where T is the number of time series observations. Alternatively, full-information estimation methods such as 3SLS, GMM-3SLS or simultaneous equations FIML may be considered to estimate the joint system of equations (see Hamilton 1994, pp. 247-253). Such methods are asymptotically superior but there is, in limited size samples, a risk of a specification error propagating to the entire system of equations. Monte Carlo evidence suggests it is not clear whether one approach is preferable to the other (Judge et al. 1985, pp. 646-53).

Structural model coefficients can be estimated using the IV estimator as follows. Let *y* be the dependent variable, let *X* be the matrix of original regressors including the endogenous variable and let *Z* be the matrix of instruments, excluding the pre-determined variables but including other original regressors as well as additional lags as necessary to attain exact identification.

$$\theta_{2SLS} = (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y \qquad (4.8)$$

The 2SLS estimation procedure yields the following results for the original structural equation (not including instruments) for the CBOT corn futures data:

$$h_t = -1.06 + 0.97 \ln V_t - 0.55 V_{t-1} + 0.447 h_{t-1} \tag{4.9}$$

$$\ln \mathrm{V_t} = 0.879 - 0.163 h_t + 0.487 \ln V_{t-1} + 0.347 \ln V_{t-2} \tag{4.10}$$

For the volatility equation, the results are qualitatively the same as for the biased model. All variables are individually statistically significant on the basis of t-tests. Standard errors are provided in Table 4.1. The adjusted R^2 is 0.406 for the volatility equation and 0.376 for the volume equation. Price volatility is positively associated with contemporaneous volume, but negatively with lagged volume, and also that volatility is positively serially correlated as expected. Trade volume is positively serially correlated, but negatively associated with contemporaneous volatility.

4.7 Wavelet-Filtered Sample Unbiased Structural Model Estimates

In this section, the approximate impact of Index Traders on price volatility is estimated using an indirect, revealed methodology. 2SLS estimates are obtained for the structural model using wavelet-filtered data that excludes variation associated with time horizons of less than one month. As explained earlier, research by the CFTC has found that Index Traders do not engage in short-run trading and selecting a one-month time horizon as threshold is most likely conservative.

As it is not meaningful to compare the statistical significance of two coefficient estimates (e.g. Gelman 2006), a qualitative interpretation is provided in this section and appropriate hypothesis tests are presented in a later section.

Estimation results using the 2SLS method and wavelet-filtered data Chicago Board of Trade corn futures are then:

$$h_t = -0.756 - 0.136\ln V_t - 0.525\ln V_{t-1} + 0.308 h_{t-1} \quad (4.11)$$

$$\ln \mathrm{V_t} = 0.0559 - 0.0061 h_t + 1.926\ln V_{t-1} - 0.941\ln V_{t-2} \quad (4.12)$$

Using wavelet-filtered data, price volatility is negatively associated with both contemporaneous and lagged volume (long-term horizon). All coefficients are individually statistically significant except for V_t in the h_t equation (3.23). Standard errors are provided in Table 4.1.

Table 4.1: Volume-Price Volatility Model for Chicago Board of Trade corn futures contract: biased (individual) model estimates, full-sample 2SLS estimates and wavelet-filtered (no short run variation in volume) estimates

Corn futures			
Volatility equation	Biased individual estimates	2SLS, full sample	2SLS, wavelet-filtered data
Intercept	-0.79	-1.06	-0.756
Std error	(0.089)	(0.069)	(0.074)
volume(t)	0.74	0.97	-0.136
Std error	(0.024)	(0.0994)	(0.17)
volume(t-1)	-0.376	-0.55	-0.525
Std error	(0.0287)	(0.0548)	(0.171)
volatility(t-1)	0.422	0.447	0.308
Std error	(0.0328)	(0.0164)	(0.015)
Volume equation	Biased individual estimates	2SLS, full sample	2SLS, wavelet-filtered data
Intercept	1.167	0.879	0.0559
Std error	(0.07)	(0.0374)	(0.0032)
volatility(t)	0.298	-0.163	0.0061
Std error	(0.0166)	(0.0593)	(0.0017)
volume(t-1)	0.385	0.487	1.926
Std error	(0.014)	(0.02)	(0.0055)
volume(t-2)	0.243	0.347	-0.941
Std error	(0.014)	(0.02)	(0.0056)
Notes: All coefficients are statistically significant except volume(t) in the wavelet-filtered 2SLS volatility equation. Standard errors are computed using the Newey-West HAC covariance.			

4.8 Estimation Results for Other Commodities

A similar three-part analysis, the results of which are presented in Tables 4.2 through 4.6, is conducted for four other commodities: soybeans, canola, lean hogs and live cattle. The estimation procedures are the biased, independent ARMAX model, the 2SLS joint model and the 2SLS joint model using wavelet-filtered data. The data used for the analysis consists of daily observations from 2/1988 to 1/2005 with the exception of Chicago Board of Trade soybean futures, for which the data used run from 4/19/1990 to 7/21/2006.

Results for soybeans, presented in Table 4.2, are generally similar to the results from corn futures data. Indeed, while correcting for the endogeneity bias does not change the sign of the volume-volatility relationship, wavelet-filtering does. The 2SLS estimates using wavelet-filtered data suggest long-run trade volume reduces price volatility, although the estimate for contemporaneous volume is not significantly different from zero.

To consider a major commodity that is not under the CFTC's jurisdiction, and for which there is less position-level data available, we include canola futures traded at the Winnipeg Commodity Exchange in Canada. The results are presented in Table 4.3 and both ARMAX and full sample 2SLS estimates are qualitatively similar to the results for corn and soybeans futures, namely that contemporaneous volume has a positive effect on volatility but lagged volume has a negative effect and moreover that correcting for the endogeneity bias does not change the signs in the structural equation. However, 2SLS results from using wavelet-filtered data suggest both the present and lagged volume variables have no effect on volatility. Indeed, neither point estimate is significantly different from zero on the basis of a t-test.

The Chicago Mercantile Exchange hogs contract changed on January 1st 1998 from a "live" (animal) specification to a "lean" (carcass) one. To avoid spurious effects from this structural change, only data beginning in 1998 are used for this analysis. Indeed, Carter and Mohapatra (2006) find that the new hogs contract has led to a substantial increase in trade volume that is plausibly independent of the role played by Index Traders. Results for Chicago Mercantile Exchange lean hogs are presented in Table 4.4 and show that the effect of present and lagged volume on volatility is qualitatively the same and always significant at the 1% level whether we use biased, correct, or wavelet-filtered correct estimates. In all cases, contemporaneous volume has a positive effect and lagged volume has a negative effect.

Results for Chicago Mercantile Exchange live cattle futures are presented in Table 4.5. Once again, correcting for the endogeneity bias does not change the sign of the current and lagged volume coefficients, respectively positive and negative. Estimates using only wavelet-filtered data suggest, however, that both current and lagged volume have a positive effect on price volatility, though the coefficient for lagged volume is not significantly different from zero.

4.8 Do Index Traders Increase Futures Price Volatility?

To summarize the results obtained in the first part of this chapter: we first provide benchmark estimates for a simple futures volume-price volatility model without accounting for the endogeneity bias, using an ARMAX maximum likelihood approach with Newey-West HAC covariance. We note that using a GARCH approach would also be biased because theoretical research suggests volume and volatility are jointly determined.

Table 4.2: Volume-Price Volatility relationship for Chicago Board of Trade soybean futures contract: biased (individual) model estimates, full-sample 2SLS estimates and wavelet-filtered (no short run variation in volume) estimates

Soybean futures			
Volatility equation	Biased individual estimates	2SLS, full sample	2SLS, wavelet-filtered data
Intercept	-0.449	-0.947	-0.651
se	(0.07)	(0.10)	(0.083)
volume(t)	1.035	1.456	-0.0219
se	(0.023)	(0.067)	(0.198)
volume(t-1)	-0.643	-0.952	-0.564
se	(0.027)	(0.053)	(0.199)
volatility(t-1)	0.489	0.519	0.307
se	(0.021)	(0.015)	(0.015)
Volume equation	Biased individual estimates	2SLS, full sample	2SLS, wavelet-filtered data
Intercept	0.937	0.882	0.0495
se	(0.058)	(0.05)	(0.0027)
volatility(t)	0.328	-0.06	0.0018
se	(0.0117)	(0.026)	(0.0013)
volume(t-1)	0.393	0.485	1.925
se	(0.0127)	(0.017)	(0.0054)
volume(t-2)	0.189	0.321	-0.938
se	(0.013)	(0.018)	(0.0054)
Notes: All coefficients are statistically significant except volume(t) in the wavelet-filtered 2SLS volatility equation. Standard errors are computed using the Newey-West HAC covariance.			

The endogeneity of volume and volatility is clearly supported by Hausman-Wu test results. For the five major agricultural commodity futures examined in this chapter, ARMAX estimates suggest that volatility is positively correlated with contemporaneous volume but negatively with lagged volume, in addition to being autocorrelated. Adjusting for the endogeneity bias by using a Two Stage Least Squares estimator does not qualitatively change the results as the regressor signs remain the same. To evaluate the impact of large Index Traders on market volatility, 2SLS estimates are obtained from filtered volume data where wavelet transform analysis is used to remove all variation associated with time horizons shorter than one month. This threshold is supported by the CFTC's research on Index Trader activity.

Table 4.3: Volume-Price Volatility Relationship for Winnipeg Commodity Exchange canola futures contract: biased (individual) model estimates, full-sample 2SLS estimates and wavelet-filtered (no short run variation in volume) estimates

Canola futures			
Volatility equation	Biased individual estimates	2SLS estimates, full sample	2SLS estimates, wavelet-filtered data
Intercept	0.137	0.115	0.142
se	(0.045)	(0.065)	(0.044)
volume(t)	0.462	0.496	0.207
se	(0.037)	(0.085)	(0.227)
volume(t-1)	-0.336	-0.357	-0.069
se	(0.033)	(0.055)	(0.227)
volatility(t-1)	0.681	0.682	0.659
se	(0.036)	(0.0115)	(0.012)
Volume equation	Biased individual estimates	2SLS, full sample	2SLS, wavelet-filtered data
Intercept	0.414	0.491	0.022
se	(0.03)	(0.025)	(0.001)
volatility(t)	0.094	-0.0255	-0.00052
se	(0.012)	(0.0118)	(0.0005)
volume(t-1)	0.428	0.447	1.924
se	(0.017)	(0.015)	(0.0055)
volume(t-2)	0.265	0.278	-0.937
se	(0.015)	(0.015)	(0.0055)

Notes: All coefficients are statistically significant except volume(t) and volume(t-1) in the wavelet-filtered 2SLS volatility equation. Standard errors are computed using the Newey-West HAC covariance.

The results using wavelet-filtered data suggest that for Chicago corn and soybean futures, volatility falls when current and lagged long-run volume rises. For Winnipeg canola, volatility is not affected by current or lagged volume as the estimates are not significantly different from zero. Results for the two non-storable commodities are qualitatively different. Volatility in live cattle futures is positively affected by both current and lagged volume, while volatility in lean hogs futures is positively affected by current volume but negatively by lagged volume. The results suggest that the impact of Index Traders, approximated using the long-run volume of trade, is beneficial to futures markets for storable commodities because it reduces price volatility.

Table 4.4: Volume-Price Volatility Relationship for Chicago Mercantile Exchange lean hogs futures contract: biased (individual) model estimates, full-sample 2SLS estimates and wavelet-filtered (no short run variation in volume) estimates

Lean hogs futures			
Volatility equation	Biased individual estimates	2SLS, full sample	2SLS, wavelet-filtered data
Intercept	2.747	2.8	2.91
se	(0.105)	(0.14)	(0.10)
volume(t)	0.553	0.47	0.80
se	(0.033)	(0.176)	(0.136)
volume(t-1)	-0.495	-0.427	-0.716
se	(0.0355)	(0.145)	(0.127)
volatility(t-1)	0.368	0.364	0.319
se	(0.024)	(0.02)	(0.021)
Volume equation	Biased individual estimates	2SLS, full sample	2SLS, wavelet-filtered data
Intercept	-0.508	0.75	0.748
se	(0.07)	(0.31)	(0.20)
volatility(t)	0.19	-0.0088	-0.113
se	(0.014)	(0.069)	(0.044)
volume(t-1)	0.639	0.657	1.354
se	(0.02)	(0.023)	(0.032)
volume(t-2)	0.205	0.191	-0.456
se	(0.02)	(0.023)	(0.031)

Notes: All coefficients are statistically significant except volatility(t) in the 2SLS full sample volume equation. Standard errors are computed using the Newey-West HAC covariance.

However, the evidence also lends support to the claim that Index Traders increase volatility for non-storable commodity futures markets such as live cattle and lean hogs.

4.10 The Distribution of Trader Time Horizons

The second contribution of this chapter is to provide a measure of the distribution of trader types across time horizons over the past two decades across all major agricultural commodities.

Table 4.5: Volume-Price Volatility Relationship for Chicago Mercantile Exchange live cattle futures contract: biased (individual) model estimates, full-sample 2SLS estimates and wavelet-filtered (no short run variation in volume) estimates

Live cattle futures			
Volatility equation	Biased individual estimates	2SLS, full sample	2SLS, wavelet-filtered data
Intercept	-1.626	0.409	1.807
se	(0.074)	(0.254)	(0.078)
volume(t)	0.86	1.836	0.383
se	(0.022)	(0.197)	(0.16)
volume(t-1)	-0.489	-1.05	0.10
se	(0.0256)	(0.116)	(0.16)
volatility(t-1)	0.381	0.404	0.267
se	(0.023)	(0.019)	(0.015)
Volume equation	Biased individual estimates	2SLS, full sample	2SLS, wavelet-filtered data
Intercept	-0.181	1.264	0.0539
se	(0.067)	(0.123)	(0.0057)
volatility(t)	0.367	-0.065	0.0079
se	(0.144)	(0.035)	(0.002)
volume(t-1)	0.419	0.501	1.902
se	(0.016)	(0.0177)	(0.006)
volume(t-2)	0.07	0.133	-0.935
se	(0.14)	(0.0172)	(0.0056)
Notes: All coefficients are statistically significant except volume(t-1) in the wavelet-filtered 2SLS volatility equation. Standard errors are computed using the Newey-West HAC covariance.			

The principal questions asked in this part of the chapter are: Can we identify the influence of Index Traders in recent years on the aggregate shape of trading time horizons? Has the time horizon of trading become longer as futures markets have matured and deepened? Do we find that markets for storable commodities have longer time horizons because inventories provide inter-temporal smoothing?

The heterogeneity of traders has been advanced as an explanation for several stylized facts observed in financial and commodity markets (Bessembinder and Seguin 1993; Daigler and Wiley 1999). Trader types have been characterized in terms of their access to information (e.g. herd or well-informed), motivation (e.g. hedger or speculator), risk aversion and prudence, or time horizon (e.g. short-run or long-run).

Trader type is defined in this chapter by the decision-making time horizon of trading, which is itself estimated by attributing to each time horizon (i.e. wavelet timescale) a proportion of the variation in trade volume. The variable used is daily trade volume aggregated for all maturities. Volume, a flow variable, is better suited to this problem than open interest, a stock variable. The goal is to measure the contribution of each distinct time horizon to variation in trade volume. The approximate distribution of trader heterogeneity as it has evolved over time is inferred, separately for each commodity, from an estimate of the distribution of trade volume across time horizons.

To determine whether differences exist among commodities in the time horizon of trading, we consider a simple linear model of daily trade volume regressed on a matrix consisting of vectors each of which is defined as variation associated with different time horizons, from daily to greater than annual. To provide correct estimates and hypothesis test results, the data is differenced because Augmented Dickey-Fuller tests suggest the data may be non-stationary. The model may written as follows:

$$\begin{aligned} \Delta V_t &= \beta_0 + \beta_1 x_{t,daily} + \beta_2 x_{t,semiweekly} + \beta_3 x_{t,weekly} + \beta_4 x_{t,biweekly} + \beta_5 x_{t,monthly} \\ &+ \beta_6 x_{t,bimonthly} + \beta_7 x_{t,trimestrial} + \beta_8 x_{t,semestrial} + \beta_9 x_{t,annual} + \beta_{10} x_{t,>annual} + \varepsilon_t \end{aligned} \tag{4.13}$$

where each x_t is associated with a specific time horizon and is orthogonal to the other time horizon vectors as described in Chapter 2. Estimation results together with White robust standard errors and individual coefficient t-tests are presented in Tables 4.6 and 4.7. The results suggest that the time horizon of trading is, with a few exceptions, similar across commodities: the three shortest time horizons are highly significant while all others are not at all significant. Moreover, the shorter is the time horizon, the more important the effect on daily volume in differences. The fourth time horizon (two weeks) is significant for live cattle, lean hogs, wheat and sugar #11. This is unexpected because it is often assumed non-storable markets have shorter time horizons than do storable markets. For cocoa and coffee, time horizons have essentially no explanatory power, which suggests trade volume is mostly driven by the long-run trend, which is not included in the matrix of regressors.

Since we know that the CME hogs contract specification changed from 1997 to 1998, we can test for a change in the parameter values associated with different time horizons. For example, the biweekly time horizon coefficient is not significant for either the 1988-1997 or 1998-2004 time periods but is qualitatively higher in the later period. A simple t-test computed to evaluate the hypothesis that β_5 (the biweekly horizon coefficient) is the same before and after the contract specification change suggests we cannot reject the null and therefore the coefficient difference is not statistically significant.

4.11 Testing for Changes over Time in Trader Heterogeneity

The evidence presented in the last section suggests that, cocoa and coffee aside, the time horizon of trading does not differ much between commodities. As these results are point estimates computed from a sixteen year sample, we would like to determine individually for each commodity whether the distribution of trading across time horizons 1988 and 2005. For example, is trading increasingly focused on the short run, on the long run, or has it not changed?

Table 4.6: Regression of daily trade volume (in differences) on wavelet-computed time horizon factors using White's robust covariance, Chicago Board of Trade, Chicago Mercantile Exchange, Kansas City Board of Trade and Winnipeg Commodity Exchange commodities

Commodity futures contract	**WCE canola R^2=0.662**		**CBOT corn R^2=0.639**		**CBOT soybean R^2=0.663**		**CME live cattle R^2=0.592**	
Time horizon factor	Coef. value	Std. error	Coef. value	Std. error	Coef. value	Std. error	Coef. value	Std. error
daily	1.59***	0.031	1.58***	0.033	1.59***	0.024	1.50***	0.027
semiweekly	0.64***	0.029	0.61***	0.029	0.64***	0.025	0.60***	0.024
weekly	0.19***	0.035	0.20***	0.032	0.19***	0.03	0.21***	0.031
biweekly	0.034	0.037	0.01	0.034	0.022	0.031	0.064***	0.032
monthly	0.018	0.039	0.005	0.028	0.018	0.03	-0.002	0.031
bimonthly	0.004	0.046	0.009	0.036	0.003	0.028	0.011	0.055
trimestrial	0.019	0.046	0.005	0.036	0.009	0.034	0.001	0.042
semestrial	-0.008	0.057	0.002	0.037	-0.003	0.034	-0.004	0.049
annual	0.002	0.053	-0.004	0.034	0.001	0.04	-0.005	0.066
greater than annual	-0.004	0.083	0	0.031	-0.001	0.025	-0.004	0.069
Notes: statistical levels of significance are *** (1%), ** (5%) and * (10%). Intercept term is not significantly different from zero (p>0.9).								

Table 4.6 (continued).

Commodity futures contract	**CME lean hogs R^2=0.587**		**CME lean hogs (1989-1997)**		**CME lean hogs (1998-2004)**		**KCBOT wheat R^2=0.628**	
Time horizon factor	Coef. value	Std. error	Coef. value	Std. error	Coef. value	Std. error	Coef. value	Std. error
daily	1.498***	0.030	1.49***	0.042	1.50***	0.042	1.535***	0.052
semiweekly	0.604***	0.026	0.63***	0.034	0.57***	0.039	0.651***	0.050
weekly	0.178***	0.031	0.18***	0.042	0.18***	0.045	0.159***	0.053
biweekly	0.046*	0.028	0.038	0.047	0.051	0.035	0.078*	0.047
monthly	0.011	0.034	0.005	0.053	0.014	0.043	0.005	0.042
bimonthly	0.004	0.041	-0.001	0.063	0.006	0.053	0.006	0.039
trimestrial	-0.002	0.040	-0.003	0.050	-0.003	0.066	0.000	0.036
semestrial	0.003	0.058	0.001	0.087	0.004	0.077	-0.006	0.015
annual	0.004	0.038	0.002	0.061	0.007	0.049	0.011	0.022
greater than annual	0.007	0.031	0.007	0.045	0.006	0.040	-0.002	0.022
Notes: statistical levels of significance are *** (1%), ** (5%) and * (10%). Intercept term is not significantly different from zero (p>0.9).								

Two test approaches are used and contrasted. First, a wavelet-based Monte Carlo test for the presence and date of change-points in the variance process. Second, a sup-Wald test of endogenous structural breaks in the Andrews-Ploberger-Hansen class.

Table 4.7: Regression of daily trade volume (in differences) on wavelet-computed time horizon factors using White's robust covariance, New York Board of Trade commodities

Commodity futures contract	**NYBOT cocoa R^2<0.01**		**NYBOT coffee R^2<0.01**		**NYBOT sugar#11 R^2=0.627**	
Time horizon factor	Coef. value	Std. error	Coef. value	Std. error	Coef. value	Std. error
daily	0.107***	0.034	0.012	0.061	1.541***	0.031
semiweekly	-0.006	0.035	0.057	0.061	0.627***	0.029
weekly	-0.08***	0.036	0.026	0.067	0.149***	0.038
biweekly	-0.068**	0.040	0.028	0.099	0.064**	0.035
monthly	-0.020	0.040	0.059	0.077	0.001	0.029
bimonthly	-0.018	0.078	-0.008	0.047	0.014	0.048
trimestrial	0.009	0.060	-0.001	0.043	0.005	0.048
semestrial	0.001	0.083	0.000	0.034	0.001	0.049
annual	0.001	0.057	0.000	0.023	-0.006	0.068
greater than annual	-0.002	0.039	0.001	0.017	0.000	0.045

Commodity futures contract	**NYBOT cotton R^2=0.588**		**NYBOT orange juice R^2=0.606**	
Time horizon factor	Coef. value	Std. error	Coef. value	Std. error
daily	1.528***	0.067	1.525***	0.042
semiweekly	0.609***	0.040	0.653***	0.033
weekly	0.221***	0.042	0.229***	0.043
biweekly	0.045	0.036	0.020	0.039
monthly	-0.006	0.022	0.014	0.030
bimonthly	0.011	0.011	0.015	0.087
trimestrial	-0.004	0.007	0.001	0.070
semestrial	0.006	0.011	-0.002	0.072
annual	-0.001	0.006	0.000	0.053
greater than annual	-0.002	0.006	0.002	0.042

Notes: statistical levels of significance are *** (1%), ** (5%) and * (10%). Intercept term is not significantly different from zero (p>0.9).

The Monte Carlo wavelet-based test is related to the cumulative sum of squares (CuSum) test of Brown, Durbin and Evans (1975). The null hypothesis is that the variance of wavelet coefficients is homogeneous, against a null of one or several change-points at specific time horizons (timescales). Because change-points in the wavelet coefficients imply breaks in the actual time series data, rejecting the null means we can identify not only the date of the breaks but also the time horizon at which they occur. For example, we may expect that trade volume would be smooth in the long run but not in the short run.

This test has more power than the Quandt-LR (sup-Wald) class of tests in the presence of long memory (Banerjee and Urga 2005). This is helpful in light of Lobato and Velasco's (2000) findings that trade volume exhibits long memory. The wavelet transform's orthogonality property provides robustness against long-range dependence (Teyssiere and Abry 2006). Another advantage of the wavelet-based test is that it identifies precisely the time horizons at which the change-points occur. For instance, trade volume for a commodity could have increased at the daily horizon, decreased at the annual horizon, and remained approximately the same for all other horizons. The null hypothesis is that the wavelet variance is homogeneous over time, which implies no change-points. If we reject the null, we can precisely identify the date of the change-point (structural break).

An approximate test statistic is constructed by Monte Carlo simulation (Dufour and Khalaf 2004). The statistic relies on uniformly-distributed pseudo-random numbers that are consistent with the sample moments of the wavelet coefficients. These wavelet coefficients are obtained from an application of the Discrete Wavelet Transform using the Daubechies(10) wavelet as described in Chapter 2 (Daubechies 1992).

The test statistic is specific to the data sample and must be computed separately for each dataset. 10,000 simulated sequences are used and the standard 1% and 5% levels of significance are saved. To minimize the computational burden, ten sets of a thousand simulations are iteratively saved, put aside and deleted.

The test results, presented in Table 4.8, show that for all commodities studied and across all time horizons, we fail to reject the null of a homogeneous wavelet variance. Equivalently, this implies there is no change-point found in the time series and therefore the volatility of futures trade volume has not changed across time horizons. This says nothing however about the mean trend in futures trade volume, which evidently has gone up in most commodity markets over the years and in particular with the increased participation of Index Traders.

An important class of hypothesis tests considers the possibility of sudden parameter changes in a time series model. Since the pioneering work of Chow (1960) and Quandt (1960) for single structural breaks at pre-determined points in time, the literature has considered the presence of multiple breaks at unknown points in time (Andrews 1993; Andrews and Ploberger 1994; Hansen 1990, 1992, 1997, 2000).

In this section are presented the results from an application of a test from the sup-Wald (sup-LM) class (Andrews 1993). The traditional Chow-Quandt *F* test has been criticized by Hansen (1990) and Zivot and Andrews (1992) because the researcher's selection of potential break points is likely to be a source of data mining. A large family of asymptotic tests for endogenous structural break points was developed among others by Andrews (1993), Andrews and Ploberger (1994), and Hansen (1991, 1992, 2000). The tests use *sup*, *exp* and *ave* functionals for LM, LR and Wald tests. Evidence suggests the *ave* has the most power against standard alternatives while the *exp* functional has most power against distant alternatives. Diebold and Chen (1996) show using bootstrapped critical test values instead of asymptotic values reduce the test size distortion substantially. In effect, structural breaks reject too often the null (e.g. Alston and Chalfant 1988, 1991). The tests are computed in R based on code by Zeileis (2006) and in Matlab based on code by Hansen (2006).

The null hypothesis is that the coefficients associated with the wavelet explanatory variables (equation 3.25) are constant over the entire sample. The results (see Appendix) show that for the *sup*, *exp* and *ave* functional tests, we cannot reject the null of no structural change in the wavelet factor model for any commodity. Consider for example Figure 4.3, which plots the empirical process and critical value for the exp-LM test using Chicago Board of Trade corn futures data. This shows the empirical process does not come close to the critical value at any point in the time series. Results for the other commodities are qualitatively the same.

A reality check is provided by applying the test to data for the Chicago Mercantile Exchange lean hogs contract, where conventional wisdom suggests a structural break occurred on January 1st 1998 when the contract specification changed from live animals to carcasses. Yet all three tests fail to reject the null of no structural change, which forces us to reconsider the true size and power of the test in this context. It is however plausible that a smooth transition occurred due to the Chicago Mercantile Exchange's efforts.

More generally, as Alston and Chalfant (1988, 1991) argue, apparent structural breaks reported in the economics literature are often explained by a model specification error, which suggests this chapter's findings of no breaks or change-points are sensible.

Table 4.8: Monte Carlo Wavelet-Based Test for Breaks in the Variance of Daily Total Futures Volume. The null hypothesis is a homogeneous variance, or equivalently no structural break or change-point in the variance. The test results imply that for all commodities and for all time horizons, we cannot reject the null (either at the 1% or 5% level of significance).

Winnipeg Commodity Exchange Canola Futures Contract, T=6534

Time horizon	Daily	Semiweekly	Weekly	Biweekly	Monthly	Bimonthly
Test value	0.3384	0.3249	0.3068	0.3714	0.3592	0.309
Critical value, 5% level	1.1755	1.1291	1.0237	1.0147	1.0684	1.01
Critical value, 1% level	1.3852	1.2534	1.0635	1.0849	1.0811	1.0375

Chicago Board of Trade Corn Futures Contract, T=7080

Time horizon	Daily	Semiweekly	Weekly	Biweekly	Monthly	Bimonthly
Test value	0.4331	0.432	0.4554	0.5002	0.5096	0.3984
Critical value, 5% level	1.0653	1.0425	1.0101	1.0151	1.0159	1.1122
Critical value, 1% level	1.1599	1.1197	1.0465	1.0284	1.1601	1.2669

Chicago Board of Trade Soybeans Futures Contract, T=8192

Time horizon	Daily	Semiweekly	Weekly	Biweekly	Monthly	Bimonthly
Test value	0.3139	0.3206	0.3368	0.3403	0.3593	0.3773
Critical value, 5% level	1.1146	1.0875	1.0714	1.0217	1.0443	1.0043
Critical value, 1% level	1.337	1.258	1.1305	1.0487	1.2081	1.1368

Kansas City Board of Trade Wheat Futures Contract, T=8192

Time horizon	Daily	Semiweekly	Weekly	Biweekly	Monthly	Bimonthly
Test value	0.5185	0.5324	0.5393	0.5489	0.5448	0.5481
Critical value, 5% level	1.0672	1.0989	1.0354	1.0313	1.0575	1.0734
Critical value, 1% level	1.1942	1.2344	1.0566	1.072	1.146	1.1995

Chicago Mercantile Exchange Lean Hogs Futures Contract, T=5944

Time horizon	Daily	Semiweekly	Weekly	Biweekly	Monthly	Bimonthly
Test value	0.1263	0.1302	0.3013	0.5033	0.4575	0.5464
Critical value, 5% level	1.1757	1.0799	1.0325	0.9947	1.0813	1.0392
Critical value, 1% level	1.2469	1.4817	1.0953	1.0383	1.2292	1.0855

Chicago Mercantile Exchange Live Cattle Futures Contract, T=4550

Time horizon	Daily	Semiweekly	Weekly	Biweekly	Monthly	Bimonthly
Test value	0.097	0.0731	0.1707	0.1586	0.2593	NA
Critical value, 5% level	1.0274	1.0246	1.0035	1.0369	1.0333	NA
Critical value, 1% level	1.1616	1.0423	1.0525	1.1414	1.0908	NA

4.12 The Distribution of Trade Volume Volatility

Lastly, we estimate, for all eleven major commodities examined in this thesis, the changing distribution of a measure of trade volume volatility (here, variance) over the time period 2/1988-1/2005. This provides an aggregated measure of the distribution of traders where trader type is defined by the time horizon of decision-making.

The variance of daily futures trade volume is decomposed across wavelet-estimated time horizons, to attribute to each time horizon its explanatory power. We examine data using sub-samples of 1024 observations each, which corresponds approximately to four years given 252 business days in one year. This method allows us to identify the contribution of each time horizon, as a factor, to the volatility of futures trade volume. The results, normalized to sum to one, are summarized in Table 4.10. The table presents, for all commodities studied in this work, the proportions of variance explained by each time horizon for each four-year time period.

Figures 4.4 to 4.9 display, for five major commodities and over sub-sample periods of four years, the variance of trade volume decomposed across distinct time horizons from daily to greater than annual. Two questions are answered by these plots. First, for a given time period, say 1989-1992, is trade volume concentrated in only one or two time horizons or rather is it uniformly, or normally, distributed? Second, has this distribution changed over the years or has it remained approximately the same?

Our empirical strategy consists of applying a wavelet transform to the volume data for a four-year period to compute wavelet coefficients and then applying an inverse wavelet transform to subsets of the wavelet coefficients.

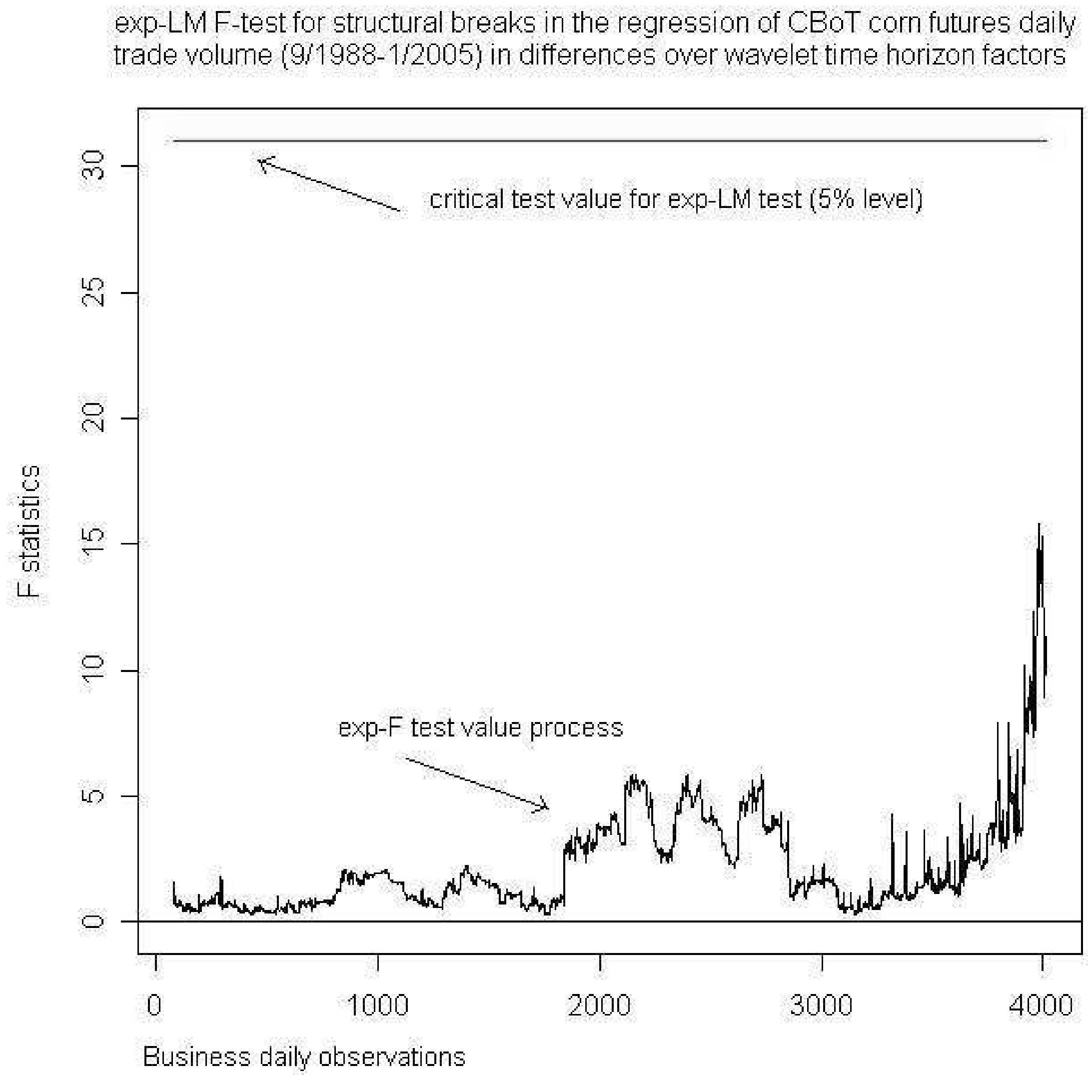


Figure 4.3: Plot of exp-LM F-test process for Chicago Board of Trade corn futures trade volume 2/1988 to 1/2005, using wavelet time horizon model

This produces a number of artificial, orthogonal time series, each of which represents a proportion of trade volume associated with a distinct time horizon. The sum of all these artificial time series yields precisely the original volume data.

An interpretation of the results on the distribution of trade volume across time horizons and over time follows. All commodities are discussed except those traded at the New York Board of Trade, for which the results and descriptive statistics are broadly consistent. Chicago hogs futures (Figure 4.4) have been, until recently, traded mostly over short time horizons, but the evidence shows that longer time horizons appear to have gained importance since the contract specification changed from “live” to “lean”.

The distribution for Kansas City wheat futures (Figure 4.6) has changed back and forth over the years but has been generally more uniform than for the two non-storable commodities, which implies there is more explanatory power found in the longer time horizons. Exceptions are the years 1993-1996, during which the longest run explained most of the variance, and 1997-2000, during which the shortest run contained most explanatory power. In contrast, the distribution for Winnipeg canola futures (Figure 4.7) has been nearly constant over the years 1981-2006, with a downward-sloping shape that implies the shortest time horizons explain more than do longer time horizons.

For Chicago live cattle futures (Figure 4.5), however, the distribution did not change over time and has been downward-sloping. This implies the shorter a time horizon, the more explanatory power it has. The exception is the sub-sample time period 1989-1992, when all time horizons contributed roughly the same to the variance of trade volume.

For Chicago corn futures (Figure 4.8), three phases are visible. From 1979 to 1986, variance was explained by the very long run, that is, time horizons greater than one year. But from 1991 to 2003, a downward-sloping shape characterized the distribution; implying shorter time horizons contributed most of the variance. Since 2003, it appears intermediate and longer-term horizons have gained in importance.

Lastly, the distribution over time for Chicago soybean futures (4.9) is generally similar to that of corn futures. From 1979 to 1990 the longest time horizons explained most of the variance, but from 1990 to 2002 the familiar downward-sloping shape was visible. Since 2003, the intermediate and long run has become more important such that all time horizons appear to contribute significantly.

Table 4.9: Variance of futures trade volume: proportion explained by time horizon

Chicago Mercantile Exchange Live Hogs Futures Contract

	1983-86	1987-90	1991-94	1995-98	1999-2002	2003-06
daily	0.18	0.21	0.26	0.25	0.19	0.06
semiweekly	0.18	0.21	0.23	0.19	0.17	0.06
weekly	0.12	0.14	0.17	0.15	0.14	0.07
biweekly	0.06	0.10	0.12	0.11	0.16	0.17
monthly	0.06	0.08	0.07	0.07	0.10	0.08
bimonthly	0.06	0.05	0.04	0.07	0.08	0.05
quarterly	0.05	0.13	0.01	0.07	0.07	0.08
semestrial	0.03	0.04	0.08	0.06	0.02	0.12
annual	0.10	0.04	0.01	0.02	0.03	0.09
longer than annual	0.16	0.00	0.01	0.01	0.05	0.22

Chicago Mercantile Exchange Live Cattle Futures Contract

	1989-92	1993-96	1997-2000	2001-04	2003-06
daily	0.18	0.25	0.25	0.21	0.18
semiweekly	0.17	0.20	0.21	0.19	0.16
weekly	0.12	0.15	0.19	0.14	0.16
biweekly	0.08	0.10	0.15	0.10	0.14
monthly	0.11	0.11	0.10	0.14	0.16
bimonthly	0.07	0.04	0.04	0.07	0.06
quarterly	0.07	0.08	0.04	0.08	0.04
semestrial	0.15	0.05	0.01	0.02	0.03
annual	0.04	0.01	0.01	0.05	0.02
longer than annual	0.00	0.02	0.00	0.00	0.05

Table 4.9 (continued).

Kansas City Board of Trade Wheat Futures Contract							
	1979-82	1983-86	1987-90	1991-94	1995-98	1999-2002	2003-06
daily	0.20	0.23	0.21	0.26	0.07	0.23	0.18
semiweekly	0.14	0.15	0.13	0.14	0.05	0.17	0.13
weekly	0.10	0.09	0.11	0.07	0.03	0.15	0.16
biweekly	0.07	0.09	0.06	0.13	0.04	0.14	0.06
monthly	0.11	0.12	0.10	0.09	0.05	0.15	0.10
bimonthly	0.07	0.03	0.07	0.11	0.05	0.06	0.15
quarterly	0.06	0.15	0.06	0.08	0.06	0.06	0.04
semestrial	0.09	0.06	0.01	0.09	0.05	0.01	0.09
annual	0.03	0.06	0.02	0.03	0.58	0.01	0.06
longer than annual	0.13	0.01	0.23	0.00	0.02	0.02	0.02
Chicago Board of Trade Corn Futures Contract							
	1979-82	1983-86	1987-90	1991-94	1995-98	1999-2002	2003-06
daily	0.07	0.13	0.16	0.20	0.22	0.16	0.14
semiweekly	0.05	0.12	0.11	0.11	0.14	0.14	0.08
weekly	0.03	0.09	0.09	0.11	0.11	0.14	0.09
biweekly	0.03	0.06	0.05	0.10	0.09	0.12	0.09
monthly	0.09	0.09	0.10	0.12	0.14	0.17	0.10
bimonthly	0.05	0.07	0.11	0.14	0.10	0.07	0.08
quarterly	0.02	0.06	0.17	0.06	0.03	0.07	0.03
semestrial	0.06	0.04	0.03	0.07	0.10	0.04	0.17
annual	0.60	0.03	0.15	0.03	0.06	0.06	0.05
longer than annual	0.00	0.32	0.03	0.06	0.00	0.03	0.17
Chicago Board of Trade Soybeans Futures Contract							
	1979-82	1983-86	1987-90	1991-94	1995-98	1999-2002	2003-06
daily	0.12	0.10	0.15	0.19	0.19	0.21	0.15
semiweekly	0.07	0.07	0.09	0.12	0.12	0.13	0.10
weekly	0.05	0.07	0.06	0.08	0.11	0.13	0.07
biweekly	0.08	0.05	0.05	0.06	0.07	0.09	0.07
monthly	0.12	0.06	0.04	0.10	0.13	0.13	0.14
bimonthly	0.09	0.06	0.08	0.18	0.12	0.09	0.18
quarterly	0.07	0.07	0.05	0.06	0.07	0.10	0.02
semestrial	0.29	0.21	0.02	0.09	0.08	0.08	0.19
annual	0.01	0.28	0.12	0.07	0.02	0.03	0.02
longer than annual	0.10	0.03	0.33	0.05	0.09	0.00	0.05

The evidence presented in this section suggests two stylized facts and testable hypotheses: (1) The time horizon of trading for non-storable commodities is shorter than it is for storable commodities, and (2) In the last five to ten years, intermediate time horizons have gained in importance for nearly all commodities, which may reflect the increased role played by Index Traders.

Table 4.9 (continued).

Winnipeg Commodity Exchange Canola Futures Contract							
	1981-84	1985-88	1989-92	1993-96	1997-2000	2001-2004	2003-06
daily	0.32	0.22	0.25	0.27	0.27	0.25	0.25
semiweekly	0.23	0.16	0.18	0.19	0.19	0.16	0.14
weekly	0.12	0.12	0.14	0.14	0.13	0.10	0.12
biweekly	0.07	0.07	0.08	0.12	0.11	0.13	0.09
monthly	0.06	0.09	0.13	0.09	0.08	0.08	0.11
bimonthly	0.06	0.10	0.10	0.06	0.07	0.05	0.06
quarterly	0.07	0.05	0.04	0.02	0.09	0.07	0.09
semestrial	0.03	0.06	0.06	0.04	0.01	0.07	0.01
annual	0.00	0.04	0.01	0.06	0.01	0.05	0.07
longer than annual	0.03	0.09	0.01	0.01	0.04	0.04	0.06

New York Board of Trade Sugar #11 Futures Contract							
	1979-82	1983-86	1987-90	1991-94	1995-98	1999-2002	2003-06
daily	0.09	0.20	0.25	0.28	0.30	0.22	0.13
semiweekly	0.10	0.16	0.15	0.18	0.19	0.20	0.11
weekly	0.04	0.10	0.11	0.12	0.15	0.16	0.12
biweekly	0.05	0.09	0.05	0.10	0.13	0.13	0.06
monthly	0.05	0.11	0.06	0.08	0.09	0.14	0.15
bimonthly	0.03	0.12	0.07	0.05	0.05	0.08	0.09
quarterly	0.02	0.02	0.08	0.05	0.04	0.03	0.09
semestrial	0.10	0.07	0.07	0.10	0.02	0.03	0.03
annual	0.22	0.13	0.00	0.03	0.02	0.00	0.06
longer than annual	0.29	0.00	0.16	0.02	0.00	0.00	0.17
New York Board of Trade Cotton Futures Contract							
	1979-82	1983-86	1987-90	1991-94	1995-98	1999-2002	2003-06
daily	0.13	0.17	0.22	0.27	0.30	0.01	0.00
semiweekly	0.14	0.15	0.23	0.25	0.23	0.01	0.00
weekly	0.07	0.13	0.15	0.17	0.16	0.01	0.01
biweekly	0.06	0.06	0.09	0.09	0.07	0.01	0.02
monthly	0.05	0.03	0.08	0.10	0.09	0.03	0.06
bimonthly	0.04	0.04	0.06	0.05	0.06	0.04	0.17
quarterly	0.14	0.03	0.01	0.03	0.03	0.06	0.23
semestrial	0.02	0.06	0.07	0.04	0.04	0.05	0.09
annual	0.31	0.12	0.06	0.00	0.02	0.51	0.23
longer than annual	0.05	0.20	0.04	0.00	0.00	0.26	0.19

4.12 Conclusion

This chapter asks two main questions about the diversity of traders in commodity futures markets. First, has the increased participation by large Index Traders led to higher futures price volatility? Should the Commodity Futures Trading Commission's pilot project where Index Trader positions are reported separately from those of other

Table 4.9 (continued).

New York Board of Trade Coffee Futures Contract

	1989-92	1993-96	1997-2000	2001-2004	2003-06
daily	0.22	0.23	0.20	0.15	0.15
semiweekly	0.14	0.15	0.19	0.14	0.15
weekly	0.14	0.13	0.17	0.13	0.11
biweekly	0.12	0.09	0.12	0.11	0.14
monthly	0.10	0.13	0.20	0.23	0.34
bimonthly	0.06	0.11	0.03	0.06	0.04
quarterly	0.09	0.06	0.03	0.03	0.01
semestrial	0.04	0.04	0.04	0.00	0.02
annual	0.04	0.00	0.01	0.03	0.03
longer than annual	0.04	0.05	0.01	0.11	0.01

New York Board of Trade Cocoa Futures Contract

	1989-92	1993-96	1997-2000	2001-04	2003-06
daily	0.22	0.32	0.22	0.04	0.19
semiweekly	0.22	0.22	0.21	0.04	0.17
weekly	0.14	0.17	0.17	0.04	0.14
biweekly	0.11	0.09	0.14	0.06	0.15
monthly	0.12	0.07	0.14	0.05	0.19
bimonthly	0.04	0.02	0.05	0.24	0.08
quarterly	0.03	0.03	0.01	0.25	0.01
semestrial	0.08	0.03	0.02	0.16	0.00
annual	0.03	0.02	0.04	0.11	0.05
longer than annual	0.01	0.01	0.00	0.01	0.02

New York Board of Trade FCOJ Futures Contract

	1979-82	1983-86	1987-90	1991-94	1995-98	1999-2002	2003-06
daily	0.16	0.15	0.18	0.18	0.22	0.29	0.23
semiweekly	0.17	0.14	0.18	0.16	0.19	0.19	0.11
weekly	0.13	0.07	0.15	0.15	0.18	0.12	0.12
biweekly	0.10	0.05	0.06	0.09	0.12	0.11	0.14
monthly	0.06	0.11	0.05	0.06	0.12	0.19	0.23
bimonthly	0.03	0.07	0.04	0.04	0.02	0.02	0.02
quarterly	0.06	0.15	0.21	0.03	0.04	0.03	0.03
semestrial	0.08	0.13	0.05	0.01	0.03	0.01	0.04
annual	0.19	0.09	0.00	0.16	0.06	0.01	0.00
longer than annual	0.03	0.03	0.07	0.12	0.01	0.03	0.07

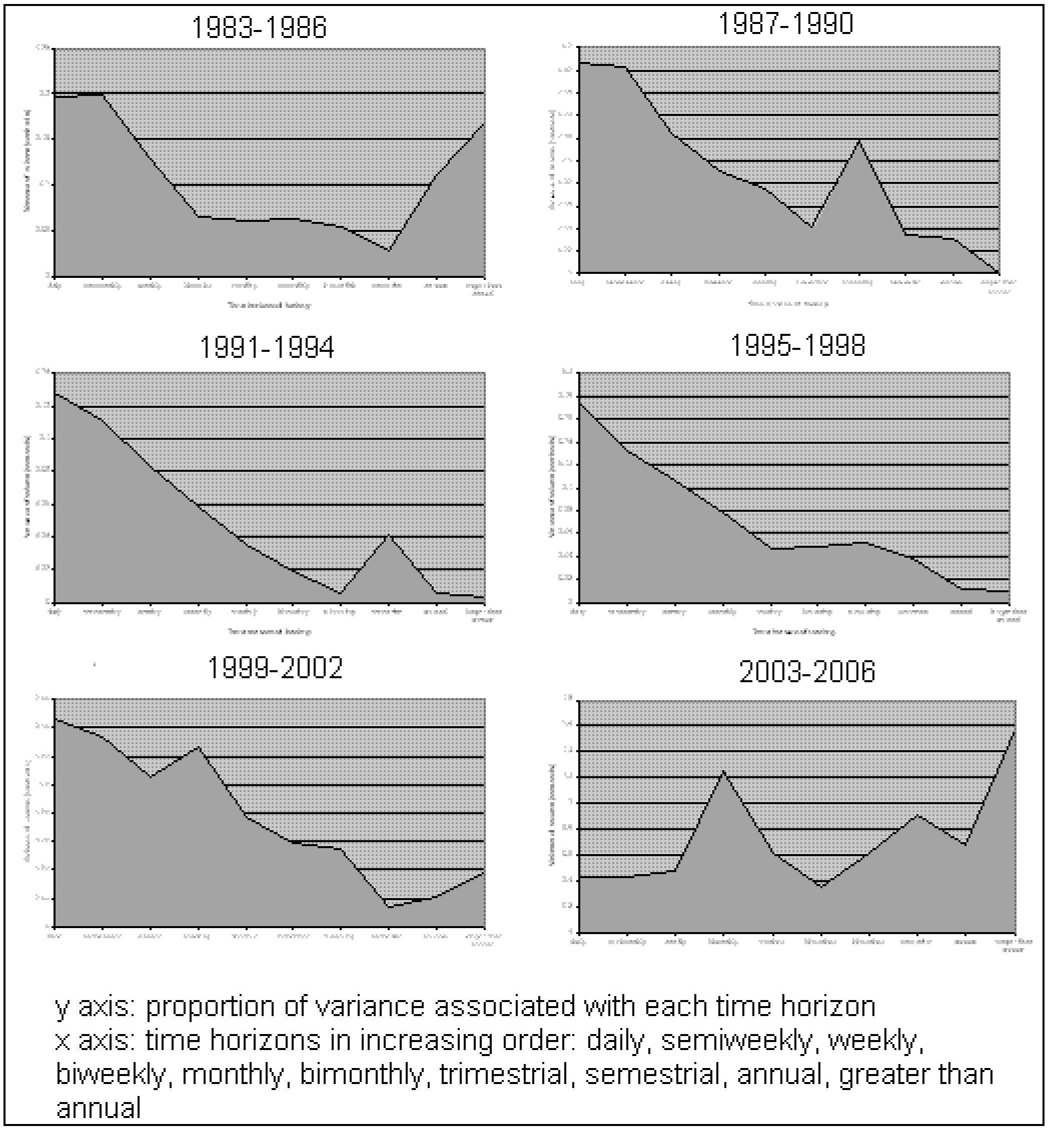


Figure 4.4: Distribution of futures trade volume variance, Chicago Mercantile Exchange live/lean hogs contract, 1983-2006

large traders become permanent? The empirical evidence presented in this chapter suggests that the impact of Index Traders may be adverse for non-storable commodity markets but is neutral or beneficial to storable commodity markets. Second, how has the time horizon of trading changed over the past two decades? Do traders increasingly trade with a shorter or longer time horizon? Are there systematic differences between storable and non-storable commodities?

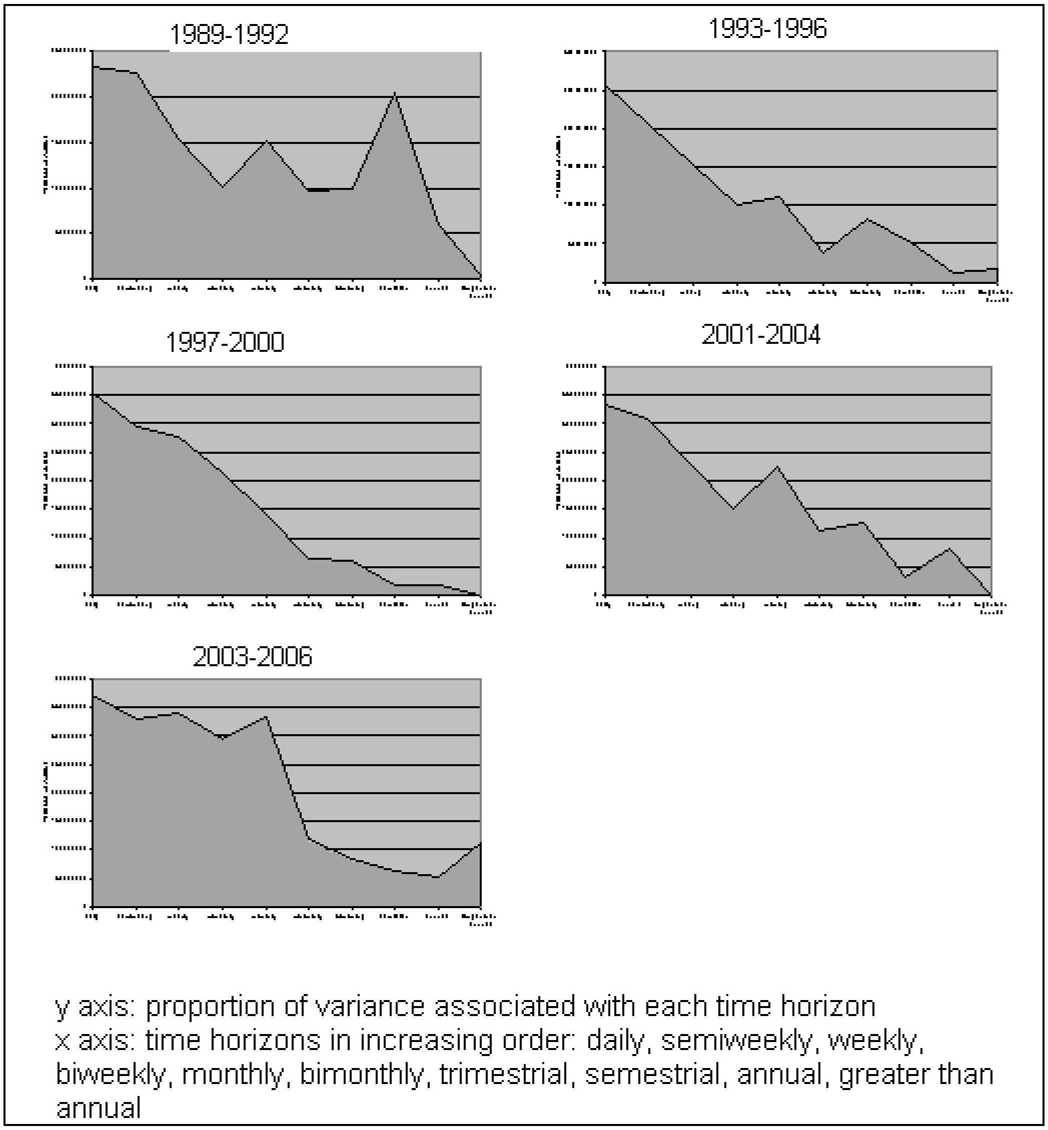


Figure 4.5: Distribution of futures trade volume variance, Chicago Mercantile Exchange live cattle contract, 1983-2006

The evidence from a wavelet transform-based decomposition of the data shows that, in the last five to ten years, intermediate and long-run time horizons have gained in importance, which may coincide with the greater role played by Index Traders. There is also some evidence to support the claim that storable commodity markets have longer time horizons.

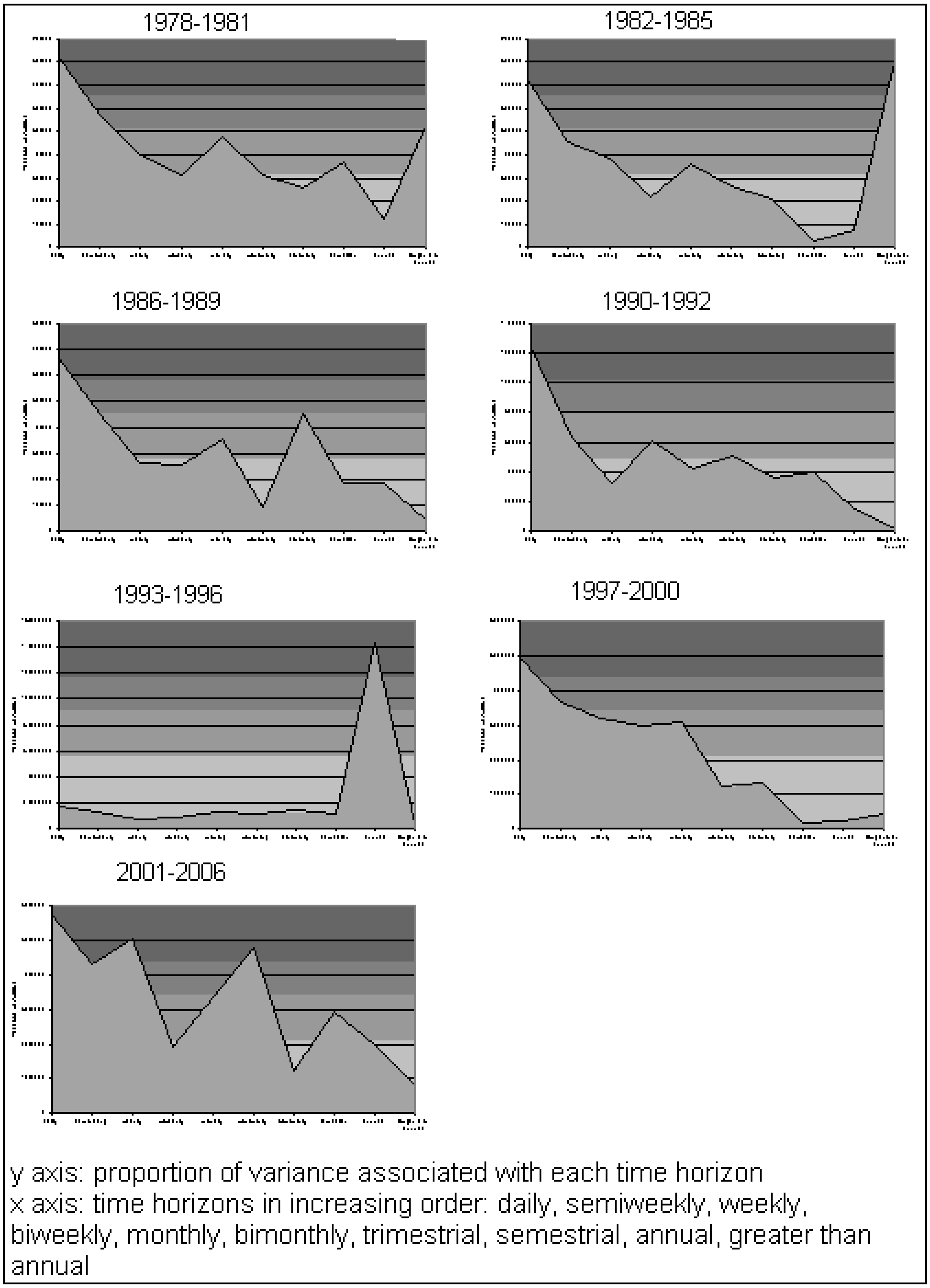


Figure 4.6: Distribution of futures trade volume variance, Kansas City Board of Trade wheat contract, 1978-2006

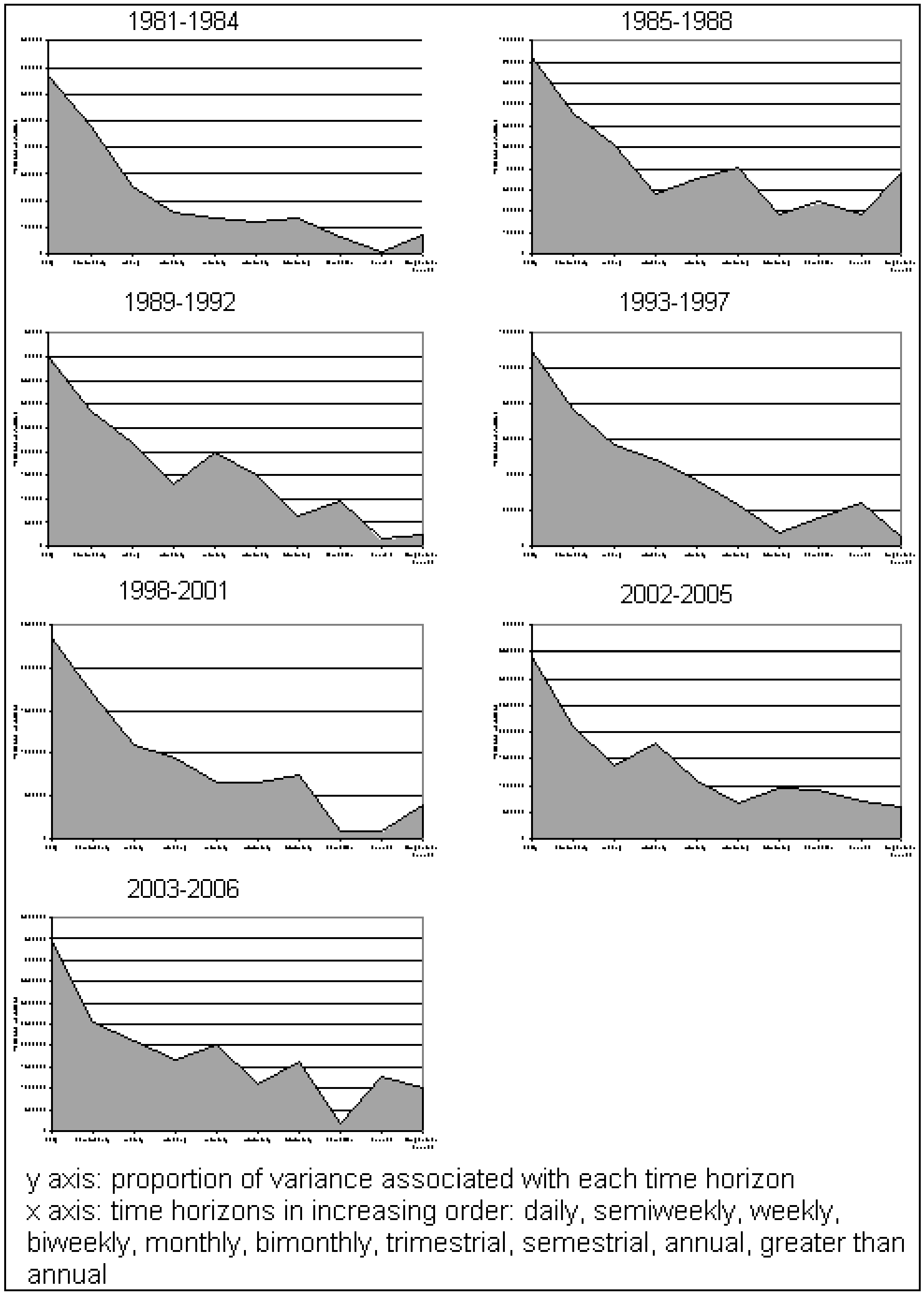


Figure 4.7: Distribution of futures trade volume variance, Winnipeg Commodity Exchange canola contract, 1981-2006

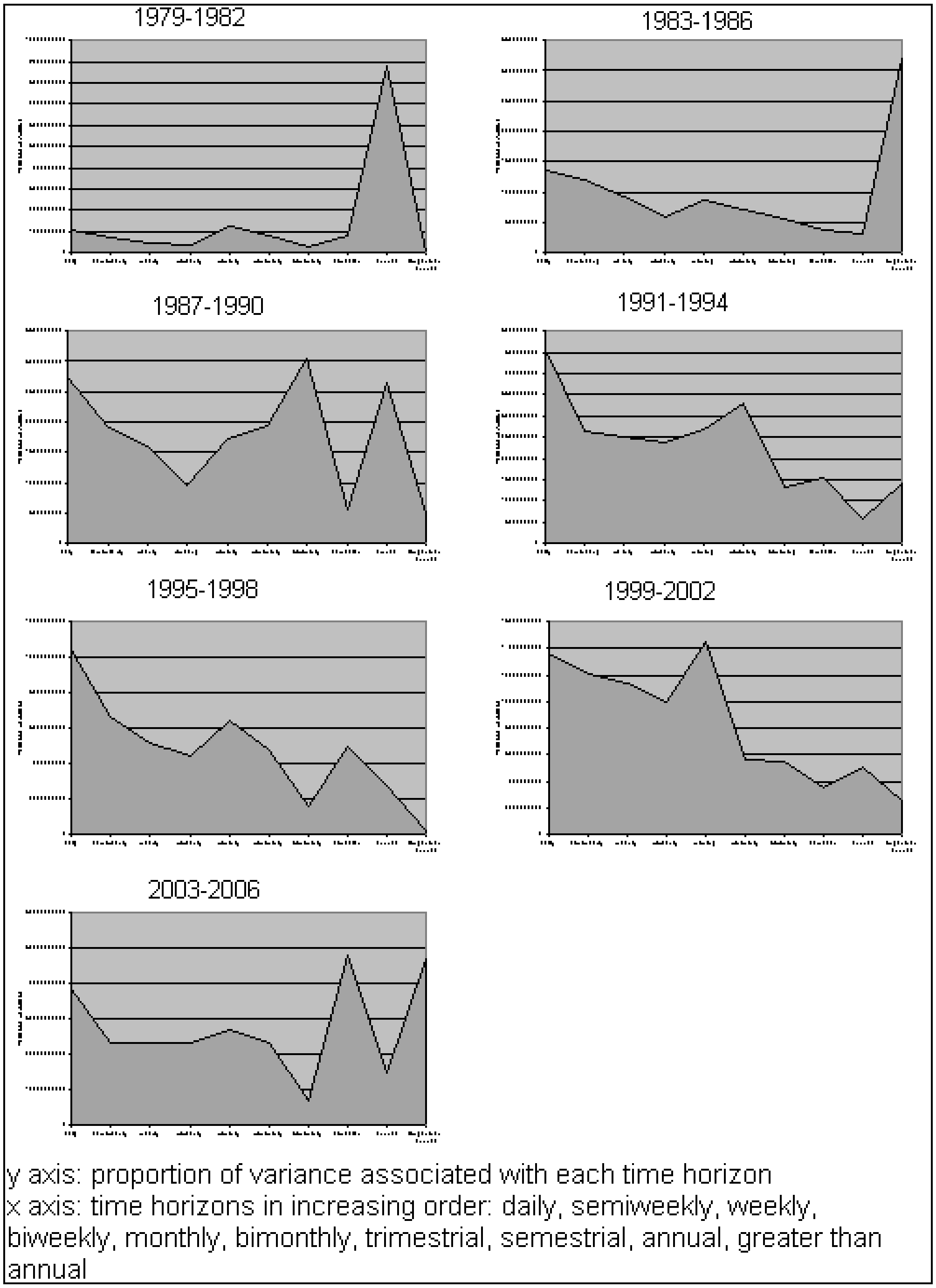
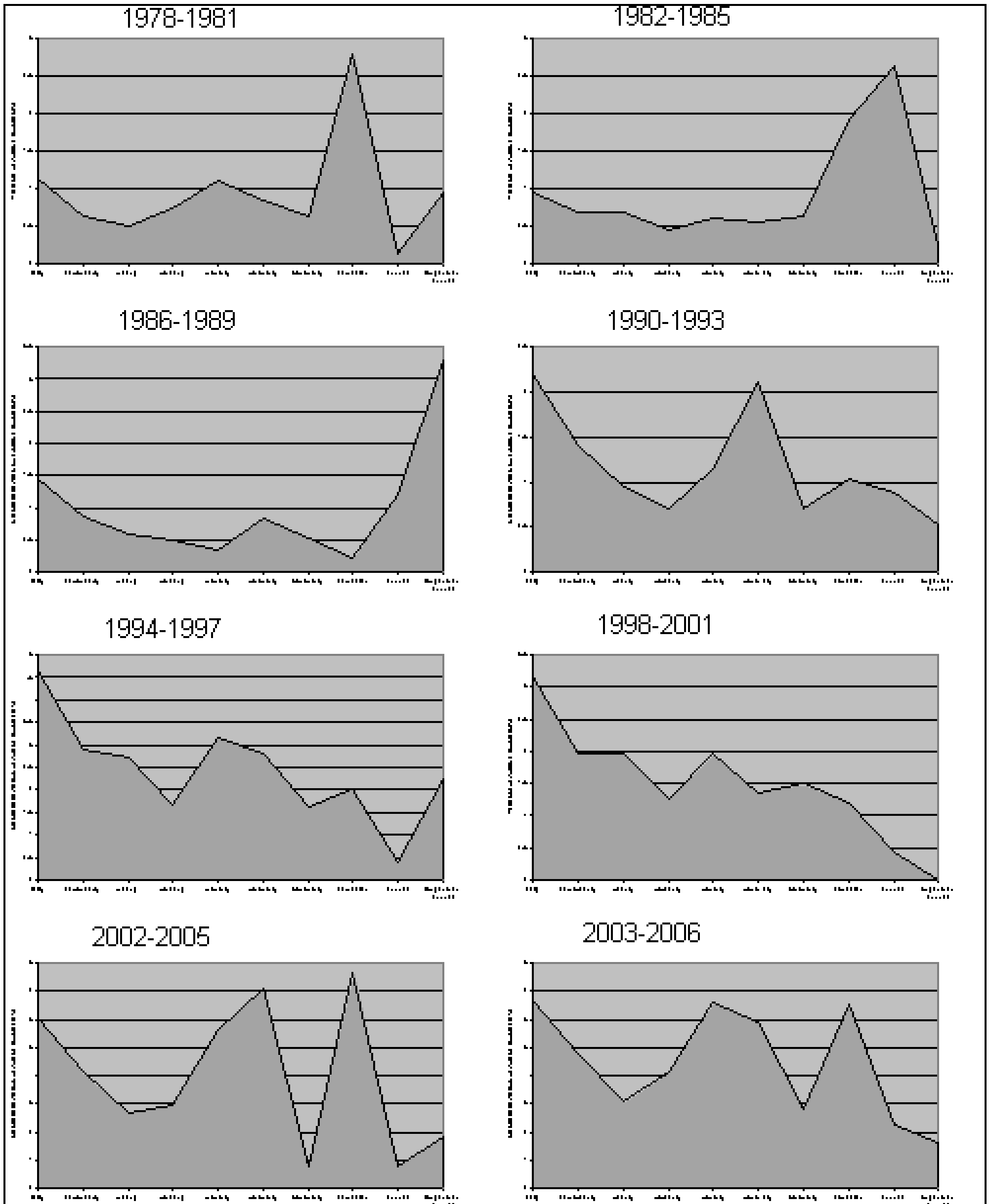


Figure 4.8: Distribution of futures trade volume variance, Chicago Board of Trade corn contract, 1979-2006



y axis: proportion of variance associated with each time horizon
x axis: time horizons in increasing order: daily, semiweekly, weekly, biweekly, monthly, bimonthly, trimestrial, semestrial, annual, greater than annual

Figure 4.9: Distribution of futures trade volume variance, Chicago Board of Trade soybeans contract, 1979-2006

Lastly, by regressing differenced trade volume over wavelet-estimated time horizon factors, we find that only the shortest three or four time horizons have coefficients that are statistically different from zero. Two exceptions among the eleven commodities are cocoa and coffee traded at the New York Board of Trade, for which the time horizons have no explanatory power in this simple regression. This suggests cocoa and coffee are mostly driven by very long-run factors.

The theoretical structure assumed in this chapter is simple and robustness of the results should be evaluated using other plausible model specifications. In particular, it should be possible to derive model testable implications based on differences between storable and non-storable commodities that are supported by theory. Furthermore, the estimates on the distribution of trader time horizons would benefit from substantial refinements to better explain why changes appear to have occurred over time. It is encouraging, however, to find that wavelet-based methods contribute new insights into persistent economic problems.

CHAPTER 5

ESTIMATING THE TERM STRUCTURE OF COMMODITY FUTURES PRICES USING WAVELET THRESHOLDING

5.1 Introduction

The term structure of futures prices approach considers how to use information from an unbalanced panel dataset, namely the constellation of futures prices traded at every business day, to extract estimates of latent (stochastic) variables such as convenience yield, cost of carry and risk premium. The literature has found that in many cases, only two or three latent factors is sufficient to track and forecast futures prices and one additional factor allows good volatility forecasting (e.g. Korn 2005; Lautier 2005; Schwartz 1997; Sorensen 2002). Motivated by theoretical advances such as Dai and Singleton (2000), recent work has considered the usefulness of models with an arbitrarily large number of latent variables (e.g. Cortazar and Naranjo 2006). For example, while a three-factor model explains 97% of the interest rate forward curve, ten factors are needed to explain 95% of the Nordic electricity term structure (Koekkebakker and Ollmar 2005).

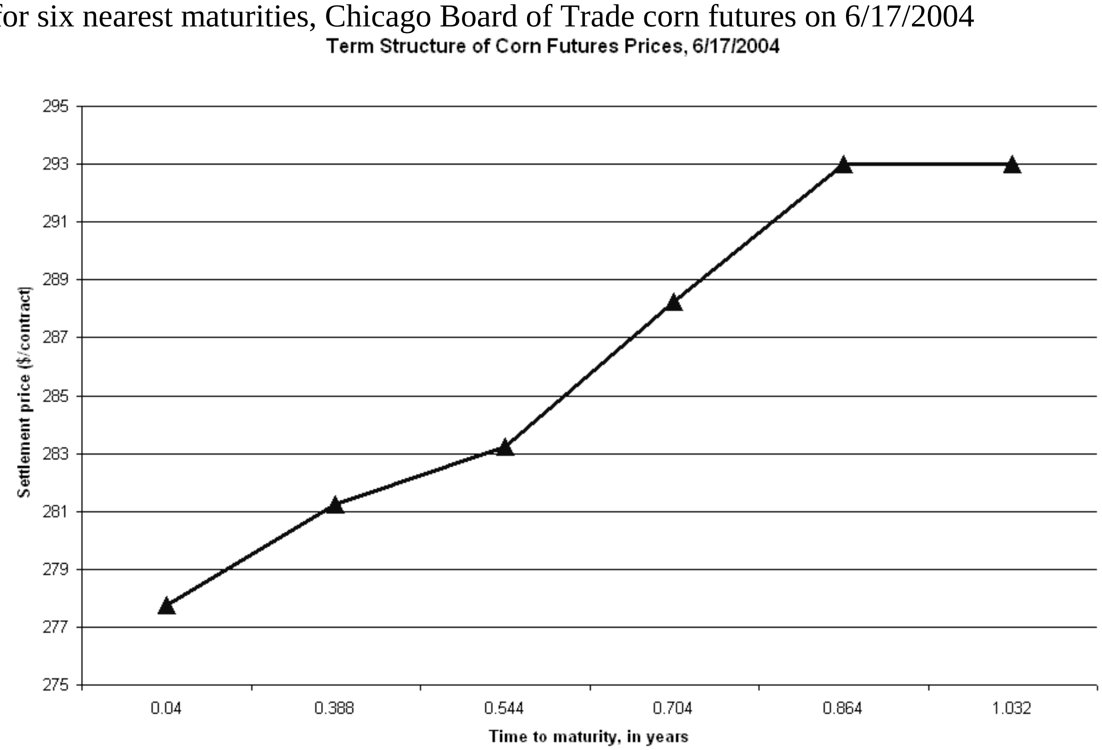
In this chapter, a new approach is suggested for the estimation of the term structure of commodity futures prices, with an application to data on one of the most traded agricultural commodities. This work follows in the literature on the stochastic behavior of commodity prices, where the Kalman filter is used to solve a multi-variate state-space time series model of observed and unobserved variables (Schwartz 1997; Schwartz and Smith 2000). The model is tractable because, following Cox, Ingersoll, Ross (1981), the futures log prices are solved as affine functions of the state variables. This chapter makes two contributions to the literature. It is, to our knowledge, the first time seasonal state variables have been combined with a large number of latent state variables to study agricultural commodity futures prices. It is also the first work to use the statistical method of wavelet thresholding (Donoho and Johnstone 1994, 1995) to improve estimation efficiency by pre-filtering the data using a data-tailored loss function. The economic interpretation of wavelet thresholding is that beneath some threshold that is unknown but can be estimated, any mean zero variation is only measurement noise of no economic significance. Filtering out this noise must necessarily reduce the process variance and therefore improve the efficiency of estimation. The purpose of the chapter is therefore, more generally, to compare the improvement in forward curve fit accuracy from using much larger models with the improvement from filtering out what appears to be noise of no economic significance.

5.2 The Term Structure of Commodity Futures Prices

Before presenting the state-space model and estimation procedure, we examine historical data on daily settlement prices for two major agricultural commodity futures contracts, Chicago Board of Trade corn and soybeans. The first nearby to sixth nearby maturities are examined. For corn futures, the forward curve since 1997 has been generally in contango, which means distant futures contracts are priced higher. The conventional explanation is that there is a positive net convenience yield which is a benefit from holding stocks into the future. From 1993 to 1997 and during a few brief additional periods of time, the forward curve was generally in backwardation, which means distant futures prices are lower. In this case, the net convenience yield is negative, which may be explained by a relatively large cost of carry, which is interpreted as the price of storing inventories. It is well understood that for agricultural commodities much of the shape is explained by seasonality (Tomek 1994; Fackler and Roberts 1999). An example of an actual commodity futures price term structure is presented in Figure 5.1 for Chicago Board of Trade corn futures on 6/17/2004. This figure shows how futures prices increased as the time to maturity increased, a pattern that is called contango.

Two general approaches to model the term structure of contingent claim prices have been used in the literature. The first, pioneered by Brennan and Schwartz (1985) and by Gibson and Schwartz (1990), estimates the unobservable convenience yield of a real or financial asset. The second, developed among others by Schwartz and Smith (2000), is based on the results of Duffie, Pan and Singleton (2000) and Dai and Singleton (2000) and models the asset price as an affine function of state variables, which are usually unobservable. This second approach nests the first and is more general. Therefore convenience yield can generally be recovered from the affine model.

Figure 5.1: An example of the term structure of futures prices: daily settlement prices for six nearest maturities, Chicago Board of Trade corn futures on 6/17/2004



The fundamental approach (Black 1976; Harrison and Kreps 1979; Cox, Ingersoll and Ross 1981; Cox, Ross and Rubinstein 1979) considers that the futures price F_t for a given date t and maturity T equals the risk-adjusted expectation of the spot price S_T at maturity under the risk-neutral probability measure Q:

$$F(x_t, t, T) = E_t^Q(S_T) \tag{5.1}$$

and it is assumed that the log of the spot price is an affine function of N different state variables as well as a deterministic seasonal function and parameters that characterize the state variable dynamics. The dynamics of each state variable is described by a stochastic differential equation (see e.g. Shreve 2004) that is solved the traditional Feynman-Kac partial differential equation approach following Black and Scholes (1973), Merton (1973), Black (1976) and Cox, Ingersoll and Ross (1981). The general multi-variate stochastic differential equation may be written as follows, where x_t is the state variable, K is a matrix of drift terms (such as mean-reverting parameters), Σ is a matrix of diffusion terms and w_t is a Brownian motion (Wiener process).

$$\mathrm{dx}_\mathrm{t} = -Kx_t dt + \Sigma dw_t \tag{5.2}$$

The canonical seasonal function is time-varying but deterministic and is identical every year for any given day.

$$s_t = \sum_{k=1}^{K} \gamma_k \cos(2\pi kt) + \tilde{\gamma}_k \sin(2\pi kt) \tag{5.3}$$

Sorensen (2002) has found that K=2 appears to provide a good and parsimonious fit.

The model to be estimated by Quasi-MLE using the Kalman filter is similar to the ones used by Roberts and Fackler (1999), Sorensen (2002) and Tien and Fackler (2003), where the logarithm of the spot price, possibly unobserved, is:

$$\ln(P_t) = s(t) + x_t + z_t \tag{5.4}$$

and where s(t) is the seasonal function and x_t and z_t are two state variables the dynamics of which are governed by a stochastic differential equation for each. Solving the spot-futures price relationship by no-arbitrage (Black 1976; Cox, Ingersoll and Ross 1981) provides the solution to the futures price as an affine function of the seasonal variable, the state variables and the time to maturity.

We follow Cortazar and Naranjo’s (2006) generalization of Schwartz and Smith (2000) because it is flexible and is designed to accommodate small changes in the model’s assumptions. This N-factor Gaussian model nests most term structure models with the notable exception of models that assume non-Gaussian Normal innovations, for example to allow a heavy-tailed error distribution. The affine transformation results of Dai and Singleton (2000) enable any model in this literature that satisfies some basic assumptions to be written in this canonical Gaussian form.

Before presenting formally the different models to be estimated, we explain briefly the economic meaning associated with each parameter. Although agricultural commodity price data are mean-reverting over long periods of time, we are also interested in testing the hypothesis of slow, gradual permanent changes caused by commodity demand or technological improvement. Therefore, the first state variable is defined as geometric Brownian motion, which is non-stationary and represents permanent changes caused for example by economic shocks in technology and preferences.

$$dx_1(t) = \mu x_1(t)dt + \sigma_1 x_1(t)dw_1(t) \tag{5.5}$$

The geometric Brownian motion state variable is associated with a long run drift term μ, a risk premium λ_1 and a diffusion σ_1 the latter which determines the degree of randomness by multiplying a Brownian motion process. The effect of time-to-maturity is captured by a risk-adjusted drift defined as:

$$\alpha = \mu - \lambda_1 + \frac{\sigma^2}{2} \tag{5.6}$$

Additional state variables x_2 through x_N are defined as Ornstein-Uhlenbeck, i.e. mean-reverting, processes where the speed of mean-reversion is captured by κ and the long-run mean to which the process is drawn is *C* (Cox and Miller 1965):

$$dx_n(t) = -\kappa_n(x_n(t) - C)dt + \sigma_n dw_n(t) \tag{5.7}$$

One-factor models universally do poorly, whether the state variable is geometric Brownian motion or Ornstein-Uhlenbeck. Multiple factor models have also considered stochastic interest rates or convenience yields as additional state variables and we return later to the definition of our factors. The Brownian motions are assumed to be pairwise correlated through a coefficient ρ_{ij}. The term structure of futures price volatility is obtained from the estimated diffusion and correlation parameters:

$$\sigma_F^2(T-t) = \sum_{i=1}^{N}\sum_{j=1}^{N} \sigma_i \sigma_j \rho_{ij} \exp^{-(\kappa_i+\kappa_j)(T-t)} \tag{5.8}$$

For the simple one-factor model, the term structure of volatility reduces to σ^2 and is the same constant regardless of time to maturity, a characteristic that is generally seen as a poor description of observed data. But for two or more state variables, volatility has a term structure that is dependent on time to maturity. The literature finds that three factors usually provide an acceptable fit, and we investigate in this chapter the gains from considering larger models.

The second approach considered to help improve estimation is a statistical filtering method called wavelet thresholding. Filters have been widely used in some areas of economics, for example, two popular macroeconomic filters are the Hodrick-Prescott filter (1980, 1997) and the Baxter-King (1999) bandpass filter. Guay and St-Amant (1997) find, however, that both filters perform poorly in recovering the business cycle component from macroeconomic time series because these data are characterized by the typical Granger spectral shape and as a result, low frequencies (long run cycles) dominate and create bias.

Wavelet thresholding is used to filter out variation beneath a precise threshold, under the assumption that it is noise of no economic significance. To evaluate the claim that this noise is of no consequence, we fit several term structure models to the data and compare both the in-sample tracking ability and out-of-sample forecasting ability of models with and without the noise. In theory, as wavelets provide an orthogonal decomposition of variance, filtering out mean zero unbiased variation must result in better (more efficient) model fitting, unless the noise is economically meaningful.

To account for backwardation and contango, that is the shape of the term structure, convenience yield is best modeled as asymmetric, because inventories cannot be negative, and time-varying, also because inventories fluctuate significantly over time. An additional source of data, commodity inventory stocks, is therefore necessary to model the asymmetry of convenience yields. Routledge, Seppi and Spatt (2000) develop such a term structure model and apply it to crude oil futures data. Casassus and Collin-Dufresne (2005) further enrich this model by incorporating stochastic interest rates and time-varying risk premia. This chapter does not adopt their model because previous research has found that, at least for agricultural commodity futures, interest rate risk is of little consequence and risk premia are small and often not significantly different from zero.

An entirely different approach which is not pursued in this chapter is to use the information contained in options to model the term structure of futures prices and volatility. For example, Egelkraut, Garcia and Sherrick (2007) use the implied volatility from commodity options on futures to estimate the term structure of volatility. They find that, at least for the nearby interval, implied volatility leads to better forecasts than do methods that use historical volatility, but the forecasting power of option implied volatility is limited when the derivative has a small trading volume.

5.3 Recovering the Net Convenience Yield

A long standing question in the literature on commodity markets, fiercely debated since the days of Keynes, Kaldor and Hicks, concerns the existence of a convenience yield. Simply stated, the convenience yield is a value to holding commodity stocks, explained for example by the benefits of positive inventories to maintain a smooth running commercial operation. This concept motivates much analysis on the shape of commodity prices at different maturities (contango and backwardation).

The net convenience yield is the difference between the convenience yield (a positive return) and the cost of carry (a negative return) the latter which is incurred through inventory expenses for bulky commodities. Williams (1989, 2001) provides a detailed treatment and critique of these concepts. Brennan, Williams and Wright (1997) argue that convenience yield is an artifact of data aggregation.

In the simplest model of the forward price curve for commodities, the following relationship holds at all times:

$$F(t,T) = T(t,t)\exp^{(r+c-\delta)(T-t)} \tag{5.9}$$

where F(t,t) is the futures price for a contract expiring "today" (i.e. the spot price notwithstanding basis risk), *r* is the risk-free rate of interest (e.g. 3-month U.S. Treasury bill), *c* is the cost of carry and δ is the convenience yield. In this simple model, the shape of the forward curve (futures prices over time to maturity) depends only on the net convenience yield: $r+c-\delta$. If $r+c>\delta$, contango results, and if $r+c<\delta$, backwardation results.

The existence of a convenience yield is not a question addressed in this chapter but, to provide a link to the vast literature on the topic, a simple identity is presented to recover the convenience yield from the model parameters estimated in this chapter. As explained by, e.g., Fackler and Roberts (1999), under the risk-neutral measure, asset price dynamics imply the following relationship using the same parameters as found in the stochastic differential equation model:

$$\mu + \delta = r + \sigma\lambda \tag{5.10}$$

where μ is the actual drift term, δ is the convenience yield, r is the risk-free rate of interest, σ is the diffusion term, and λ is the market price of risk for the state variable in question. The equation may be rearranged to give:

$$\mu - \sigma\lambda = r - \delta \tag{5.11}$$

which implies the risk-adjusted drift in the process equals the risk-free rate minus the convenience yield. Convenience yield can be recovered because the left-hand side parameters are estimated from the data using the above model and the 3-month US Treasury bill provides a good proxy for the risk-free rate of interest. For multi-factor models, additional parameters must be incorporated in the equation but the approach is the same. If reliable inventory data are available, better estimates of the cost of carry and convenience yield can be obtained, in particular accounting for asymmetry in the yield.

5.4 Wavelet Thresholding

Wavelet thresholding or shrinkage (Donoho and Johnstone 1994, 1995, 1998) has proven to be in engineering and physical sciences applications a remarkably efficient and accurate method to remove noise from data and recover the true signal. It consists of applying a filtering rule not to the actual data but rather to the wavelet coefficients computed from the data. After applying the thresholding rule, the filtered time series data are recovered from the thresholded wavelet coefficients. While the algorithm is most powerful against IID white noise, properly adjusted it provides excellent results when the noise is a dependent and non-IID stochastic process. The hypothesis is that wavelet thresholding, by filtering out short-run noise, will enable us to obtain better out-of-sample forecasts when combined with traditional time series methods.

There exist a wide variety of filtering methods other than wavelet-based. Two important class of filters are sinusoidal (Fourier) and polynomial knot (spline) smoothers. These methods, however, have been found to systematically either remove too little or too much noise. The outcome is a recovered signal that is either over-smoothed or still too noisy to be informed on the true data generating process. In contrast, wavelet thresholding has been found to provide a powerful signal recovery without oversmoothing. In particular, features of the data that are sharp remain so after wavelet thresholding, while previously existing methods tend to dull such sharp features. This is because wavelets have been designed to provide optimal information compression and efficient transformation. Formal proofs of these results are found in Donoho and Johnstone (1994, 1995, 1998).

The objective of wavelet thresholding is to determine an optimal value (threshold) using a clear criterion, such as a loss function or minimum risk value (Stein 1981). Both a threshold choice and a thresholding rule must be carefully selected. Before using the threshold, a Discrete Wavelet Transform is applied to the data to produce a vector or matrix of wavelet coefficients. The threshold is then used with the wavelet coefficients. Applying an Inverse Discrete Wavelet Transform to the filtered wavelet coefficients yields a filtered version of the original time series with no loss of information other than from filtering. Donoho and Johnstone show that a so-called universal threshold, together with a soft thresholding rule, are both asymptotically optimal and also remarkably robust when used in empirical applications. The universal threshold, assuming a variance of innovations (errors) σ^2_ε and a number of observations *T* is given by:

$$\delta = \sqrt{2\sigma_e^2 \ln(T)} \tag{5.12}$$

and the soft thresholding rule applied to wavelet coefficients *w* is:

$$\mathrm{w}^{\mathrm{soft}} = \mathrm{sgn}(w)\left(\frac{1}{2}\left(|w| - \delta + |(|w| - \delta)|\right)\right) \tag{5.13}$$

Since the true variance of the innovations is unknown, a mean absolute deviation estimate can be computed as the ratio of the median of wavelet coefficients at the finest timescale over a normalization factor that has been found to be optimal:

$$\hat{\sigma}_{MAD} = \frac{median(w^{j=1})}{0.6745} \tag{5.14}$$

5.5 State-Space Estimation with Wavelet Thresholding

Hidden component models are increasingly used and particularly well suited to estimation by the state-space approach (Durbin and Koopman 2001). In this class of models, potentially unobservable (latent) state variables are estimated together with the model parameters using available data. The standard method is to first derive a reduced form of the theoretical relationship that is to be estimated in a state-space framework. This reduced form is estimated using the Kalman filter that relates the measurement equation, for which the dependent variable is observable, to the transition equation, for which the dependent variable is typically unobservable.

Use of the Kalman filter follows previous work in this area by Schwartz 1997, Schwartz and Smith 2000, Fackler and Roberts 1999, Sorensen 2002, Korn 2005, and Fackler and Tian 2003. Although the state space approach using the Kalman filter is powerful and enlightening, it is computationally difficult to ensure that global rather than local optima are attained. In fact, the developers of the R programming language explain that: "Optimization of structural models is a lot harder than many of the references admit. For example, the Air Passengers data are considered in Brockwell and Davis (1996): their solution appears to be a local maximum, but nowhere near as good as that produced by [R procedure] *StructTS*. It is quite common to find fits with one or more variances zero…" (R Development Team 2006, pp. 1220).

We follow most closely Sorensen's (2002) estimation structure but with two significant differences. First, we consider not just a two-state variable model but several models with a number of state variables ranging from one to four. Second, we pre-filter the price data using wavelet thresholding to remove very short term noise that may obscure meaningful economic parameters. Where Sorensen lets the number of traded maturities on any given day vary within the sample, we use only the five nearby contracts. Our justification is that trade volume for more distant maturities is very low and these data points may not be entirely reliable. We have considered imposing parametric identifying restrictions based on previous findings in the literature. However, as this literature is still young and previous results are not always in agreement, it was decided to only use model restrictions such as cross-term covariance restrictions to ensure identification. For example, although empirical evidence suggests the market prices of risk λ are small and sometimes not significantly different from zero, we nonetheless include these parameters because theory suggests they are economically meaningful. We also allow correlation between state variables rather than impose a zero correlation restriction. Sorensen (2002) finds that for his two-state variable model, correlation is small but significant. Lastly, identifying restrictions may be obtained by using exogenous (more accurately, pre-determined) variables such as using the daily log-range of prices to estimate the diffusion terms σ or using the 3-month US Treasury bill to provide a measure of the risk-free rate of interest for the drift term. For the objectives of this chapter, however, these do not appear necessary.

The Kalman filter is used to estimate the maximum likelihood parameters of the state-space model of futures prices. The two most important issues in this estimation problem are solving the reduced form identification problem and providing the Kalman filter with sensible starting values. For the latter, we initialize the procedure using the estimates found by Sorensen (2002). The identification problem in this case is the recovery of structural model parameters from the estimated reduced form model. As explained by Roberts and Fackler (1999), the complete model of the term structure of futures prices for agricultural commodities is over-parameterized, equivalently, under-identified. This implies there is not a unique solution to the estimation problem.

The state-space model is based on a measurement equation and a transition (state) equation. For each time series date t={1,2,3,…,T}, the transition equation is:

$$X_{t+1} = a + AX_t + \eta_{t+1} \tag{5.15}$$

where, for the case of three state variables we have:

$$a=(\mu - 0.5\sigma^2,0,0,0)^{\mathrm{T}}$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\kappa_2\Delta} & 0 \\ 0 & 0 & e^{-\kappa_3\Delta} \end{pmatrix} \tag{5.16}$$

and the covariance matrix of the state variable innovations, from which are derived parameter identifying restrictions, is:

$$\Omega = \begin{pmatrix} \sigma_1^2\Delta & \frac{\rho_{12}\sigma_1\sigma_2}{\kappa_2}\left(1-e^{-\kappa_2\Delta}\right) & \frac{\rho_{13}\sigma_1\sigma_3}{\kappa_3}\left(1-e^{-\kappa_3\Delta}\right) \\ \frac{\rho_{12}\sigma_1\sigma_2}{\kappa_2}\left(1-e^{-\kappa_2\Delta}\right) & \frac{\sigma_2^2}{2\kappa_2}\left(1-e^{-2\kappa_2\Delta}\right) & \frac{\rho_{23}\sigma_2\sigma_3}{\kappa_2+\kappa_3}\left(1-e^{-(\kappa_2+\kappa_3)\Delta}\right) \\ \frac{\rho_{13}\sigma_1\sigma_3}{\kappa_3}\left(1-e^{-\kappa_3\Delta}\right) & \frac{\rho_{23}\sigma_2\sigma_3}{\kappa_2+\kappa_3}\left(1-e^{-(\kappa_2+\kappa_3)\Delta}\right) & \frac{\sigma_3^2}{2\kappa_3}\left(1-e^{-2\kappa_3\Delta}\right) \end{pmatrix} \tag{5.17}$$

where Δ is an increment in the unit of time, here 0.04 which is the ratio of one business day over one year (250 business days). The covariance matrix for the case of four state variables follows naturally from the above three-variable matrix.

The measurement equation for five maturities, such that Y_t is a vector of length five at each point in time, is:

$$Y_t = c_t + C_t X_t + \varepsilon_t \tag{5.18}$$

where:

$$\begin{aligned} c_t &= s(t) + (\mu + \lambda_1 - 0.5\sigma^2)(T^{(1)} - t), \\ & s(t) + (\mu + \lambda_1 - 0.5\sigma^2)(T^{(2)} - t), \ldots, s(t) + (\mu + \lambda_1 - 0.5\sigma^2)(T^{(5)} - t) \end{aligned} \tag{5.19}$$

$$C_t = \begin{pmatrix} 1 & e^{-\kappa_2(T^{(1)}-t)} & e^{-\kappa_3(T^{(1)}-t)} \\ \vdots & \vdots & \vdots \\ 1 & e^{-\kappa_2(T^{(5)}-t)} & e^{-\kappa_3(T^{(5)}-t)} \end{pmatrix} \tag{5.20}$$

and ε_t is distributed IID Normal with mean zero and covariance $\sigma_\varepsilon^2 \mathbf{I}_t$. The Kalman filter is initialized with starting values for the state variables and covariance, then computes one-step ahead forecast errors between forecast and actual observations. The exact diffuse prior of Durbin and Koopman (2001) is used to improve the behavior of the transition covariance matrix.

5.6 Estimation Results for One-Factor to Four-Factor Models

Table 5.1 presents estimated parameter values for the one, two, three and four-factor models using both the original (full sample) data and the wavelet filtered data using Donoho and Johnstone's threshold criterion. For one to four factors, the number of estimated parameters is, respectively, 3, 7, 12 and 18. This implies the computational burden grows substantially as the number of factors increases. The simplest model nested in the Gaussian N-factor framework considers the log of futures prices to be an affine function of one non-stationary state variable in addition to parametric terms:

$$\log F(t,T) = \mu t + \left(\mu - \lambda + \frac{1}{2}\sigma^2 \right)(T-t) + s(t) + x_t + \varepsilon_t \tag{5.21}$$

$$x_t = \left(\mu - \frac{1}{2}\sigma^2 \right) + x_{t-1} + \eta_t \tag{5.22}$$

where *s(t)* is the seasonal, deterministic function described earlier and (T-t) is the time to maturity expressed as a fraction of one year.

The parameter estimates suggest that both the non-stationary long-run drift and the risk premium are small, as expected from theory, although all are significant at the 1% level assuming sensible convergence of the numerical derivatives. The diffusion term is consistent with previous estimates found in the literature. Looking at wavelet-filtered one-factor model estimates, the main difference is that the risk premium parameter is now nearly zero. This may be interpreted as evidence that very short-run variation consists of a non-zero risk premium rather than noise. As expected, estimation convergence also improves because the filtered data variance is smaller.

Table 5.1: Estimation results from one- to four-factor models of the term structure of futures prices, Chicago Board of Trade corn futures five nearby maturities, from 2/1988 to 1/2005. Results provided for both full sample and wavelet-filtered sample data.

	One factor		Two factor		Three factor		Four factor	
	original	filtered	original	Filtered	original	filtered	original	filtered
μ	0.0058	0.0051	0.0049	0.00405	0.00947	0.003478	0.00509	0.000469
$\kappa 2$	.	.	0.162	0.00034	1.1361	0.0144	1.31392	2.2581
$\kappa 3$	.	.	.	.	1.1357	2.9974	0.49811	2.257
$\kappa 4$	.	.	.	.	.	.	0.45184	2.2701
$\sigma 1$	0.1077	0.101	0.1995	0.0891	0.1378	0.1668	0.096	0.1934
$\sigma 2$	.	.	0.0297	0.0488	0.0242	0.0707	0.1242	0.2964
$\sigma 3$	.	.	.	.	0.0686	0.0294	0.0782	0.2055
$\sigma 4$	.	.	.	.	.	.	0.0186	0.198
$\lambda 1$	0.011	-0.0046	-0.119	-0.1775	-0.2661	-0.2225	-0.077	0.0261
$\lambda 2$	.	.	0.0872	0.1821	0.15996	0.2049	0.0803	0.0103
$\lambda 3$	.	.	.	.	0.15363	0.1637	0.0506	0.0234
$\lambda 4$	.	.	.	.	.	.	0.1033	-0.0631
$\rho 12$	.	.	-0.2128	0.270	-0.0232	-0.0103	0.680	0.99
$\rho 13$	.	.	.	.	-0.8559	-0.5335	0.324	0.99
$\rho 14$	.	.	.	.	0.1454	0.99	0.199	-0.99
$\rho 23$	.	.	.	.	.	.	-0.99	0.99
$\rho 24$	.	.	.	.	.	.	-0.7488	0.99
$\rho 34$	.	.	.	.	.	.	0.458	0.99

Note: all parameter estimates are individually significant at least at the 5% level.

The second and additional factors are mean-reverting state variables. Although a clear economic meaning is elusive, these factors help explain the shape of the forward curve (e.g. contango or backwardation) through their interaction with the remaining time to maturity, and can be used to recover estimates of convenience yield and cost of carry. Note that if two or more mean-reverting state variables are used, mixed shapes can be captured, for example if the curve is in contango for the three nearest maturities but in backwardation for the most distant three maturities.

The two-factor model is:

$$\log F(t,T) = s(t) + \mu t + \left(\mu - \lambda_1 + 0.5\sigma^2\right)(T-t)$$
$$+x_{1,t} + e^{-\kappa_2 (T-t)} x_{2,t} - \frac{\lambda_2}{\kappa_2}\left(1 - e^{-\kappa_2 (T-t)}\right) + 0.5\sigma_1\sigma_2\rho_{12}\left(\frac{1-e^{-\kappa_2 (T-t)}}{\kappa_2}\right) + \varepsilon_t \tag{5.23}$$

$$x_t = (\mu - 0.5\sigma^2, 0)^T + Ax_{t-1} + \eta_t \tag{5.24}$$

where the matrix A is:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & e^{-\kappa_2 \Delta} \end{pmatrix} \tag{5.25}$$

Recall that the first state variable is non-stationary geometric Brownian motion so implicitly we have imposed the restriction κ_1=0. For both the full sample data and the wavelet-filtered data, parameters are statistically significant at least at the 5% level. The wavelet-filtered sample estimates are less plausible than those obtained from the full sample, in particular the small value of the mean-reversion parameter κ.

The three-factor model incorporates a second mean-reverting state variable and provides a superior fit to the data on days when the curve is not smooth but rather kinked. The results suggest once more that the non-stationary variable has a negligible but nonzero drift and significant diffusion, while the mean-reverting speed for the other two state variables is fast and consistent with previous findings—larger than Sorensen's (2002) but smaller than Fackler and Roberts's (1999). The mean-reverting state variables have diffusion parameters that are smaller than those found in the literature but not unreasonable. The three risk premium parameters are sizable but, crucially, add up to only 0.048, which confirms the literature's findings that the overall impact of risk premia for agricultural commodities is small. The non-stationary state variable is essentially uncorrelated with the first mean-reverting state variable but strongly negatively correlated with the second. The two stationary state variables are only weakly correlated. It appears pre-filtering the data using wavelet thresholding fails to improve estimation and some of the resulting parameter estimates are less plausible. In particular, the estimates for the two mean-reverting parameters are poor. It seems that wavelet filtering makes it difficult to separate the influence of the two stationary state variables as their correlation coefficient nearly equals 1.

Lastly, we consider the results from estimating a four factor model, which is characterized by one non-stationary state variable, three stationary state variables and 17 constant parameters to be estimated. The wavelet-filtered estimates are better overall. In particular, the mean-reverting speed parameters κ and the diffusions σ take far more sensible values and the market prices of risk are smaller and more consistent with the literature's previous findings. However, the correlation coefficients are unreasonable.

5.7 Interpretation of the Results and In-Sample Tracking

To evaluate the tracking ability of each model, the dates 3/17/2004 and 6/17/2004 are selected. On the first date a clear backwardation pattern is visible, and on the second date it is contango. It is assumed no economic structural change has taken place between the two dates as they are only three months apart and seasonality is controlled by the deterministic sinusoidal term. Parameter estimates and the Kalman filter estimated state variable (latent) time series are used to compute in-sample predictions of futures prices for all maturities on the given dates. Futures prices predicted from all four models, with and without wavelet thresholding, are compared with the actual prices on those days.

Figures 5.2 to 5.5 provide examples of how well each model tracks the data, with and without wavelet thresholding. Figure 5.2 shows that for a typical contango pattern of futures prices the one factor model performs poorly and, with or without wavelet thresholding, substantially under-estimates the prices. In Figure 5.3, which also displays a contango pattern, the two factor model greatly over-estimates the prices again whether or not a wavelet threshold is used. In Figures 5.4 and 5.5 (backwardation and contango, respectively), three- and four-factor models using wavelet thresholding perform well, but these need not be representative.

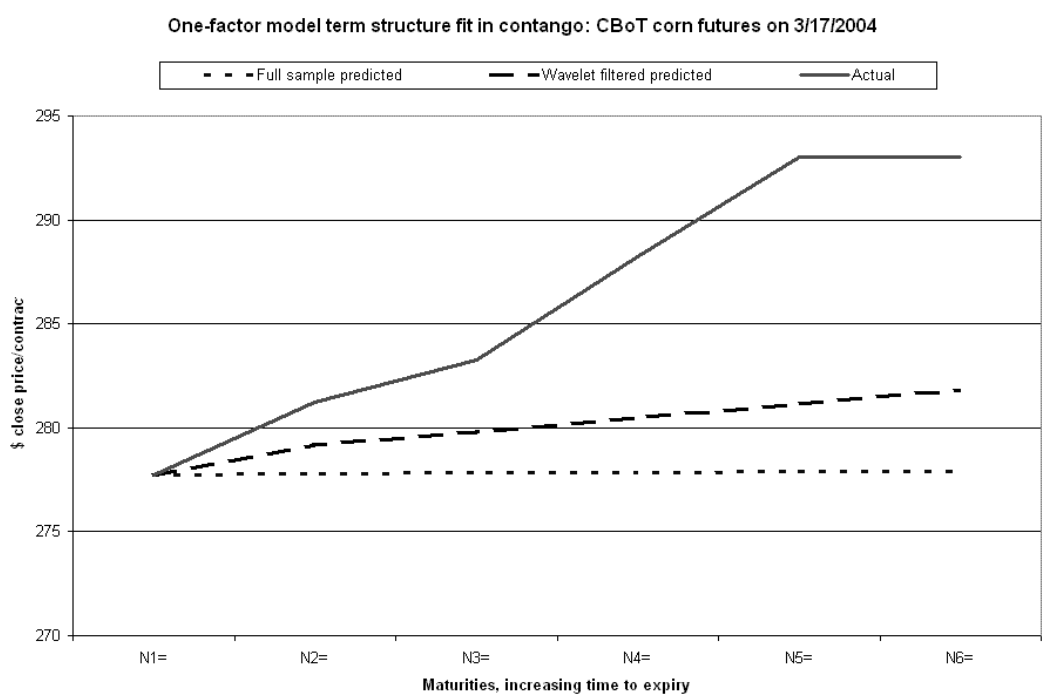


Figure 5.2: One factor model estimation with and without wavelet thresholding, in-sample tracking for Chicago Board of Trade corn futures on 3/17/2004 (contango).

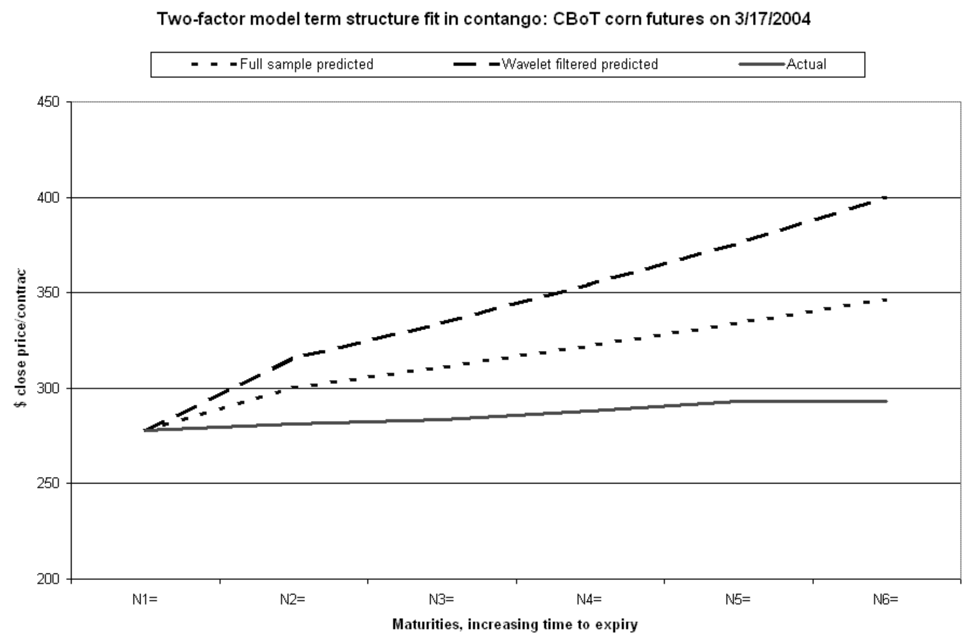


Figure 5.3: Two factor model estimation with and without wavelet thresholding, in-sample tracking for Chicago Board of Trade corn futures on 3/17/2004 (contango)

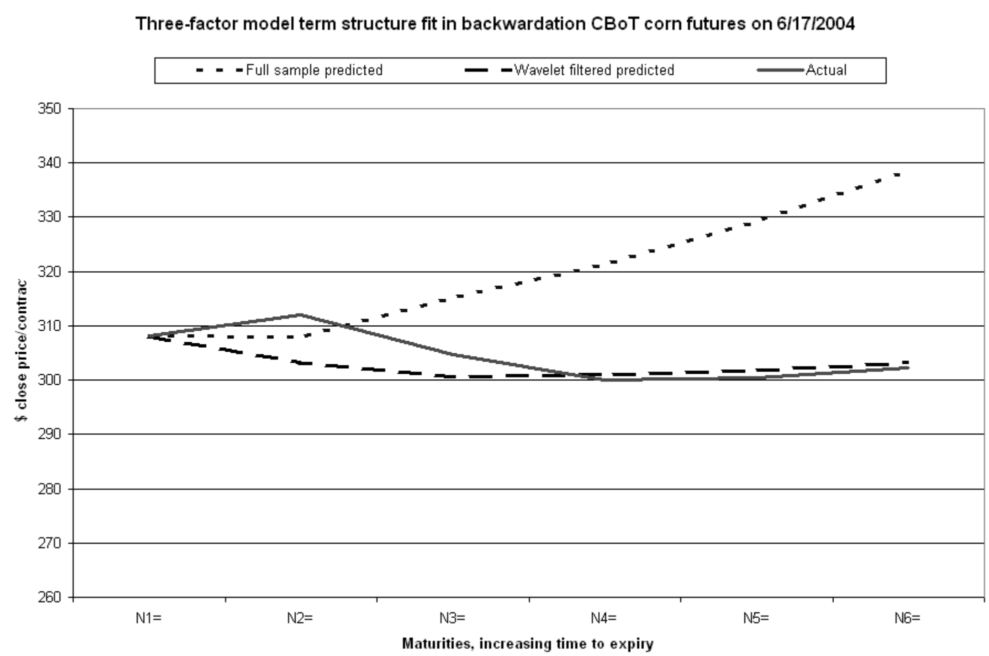


Figure 5.4: Three factor model estimation with and without wavelet thresholding, in-sample tracking for Chicago Board of Trade corn futures on 6/17/2004 (backwardation)

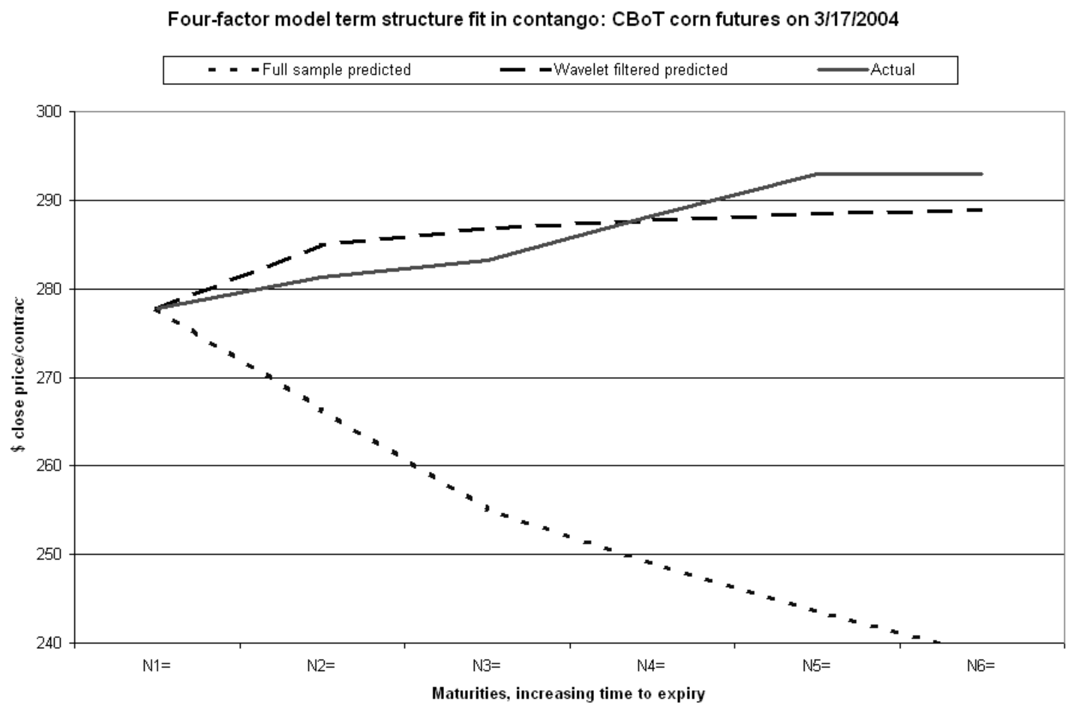


Figure 5.5: Four factor model estimation with and without wavelet thresholding, in-sample tracking for Chicago Board of Trade corn futures on 3/17/2004 (contango)

The fit of the different models ranges from good to very poor and varies substantially. One unexpected result is that filtering using the wavelet threshold does not improve estimation. It may imply that short-run variation that appears to be noise is in reality economically meaningful. Alternatively, it may be that wavelet-based filtering only improves the estimation of models that are already robust and stable, which is not the case here.

To explain the difficulty of obtaining sensible estimates, a likely cause is the combination of a non-stationary variable and one or more stationary variables, which creates instability in the Kalman filter estimation. To improve convergence, we used Durbin and Koopmans's (1997) exact diffuse prior (initial condition) for the Kalman transition variance and we excluded the first few observations from the variance calculation. These steps appear insufficient. One interpretation of the results is that, given the very small value taken by the long-run drift term μ, it may be better to impose the restriction that there is no non-stationary component in agricultural commodity futures prices (see more generally e.g. Korn 2005).

Another explanation for the poor convergence of the models is that daily observations were used rather than the more traditional weekly sampling. Alternatively, it may be beneficial to conduct the analysis on two or more sub-samples of the entire dataset, as Cortazar and Naranjo (2006) have done. The strategy has the added benefit of providing evidence on whether any of the parameters have changed over time.

Previous research suggests the market prices of risk associated with state variables are negligible, and inclusion of these parameters substantially complicates the estimation procedure. Crucially, it appears that incorrect estimation of the market prices of risk contaminates the accuracy of mean reversion parameters, which are essential to capturing the shape of the forward curve. Yet the market prices of risk contain information on whether the shape is in contango or in backwardation. A potential extension of this work is to test the hypothesis that wavelet-filtered noise is a good estimator of the time-varying market price of risk.

5.8 Conclusion

Risk management in commodity markets depends on an understanding the constellation of futures prices. A powerful framework to model the relationship between futures maturities is the term structure of futures prices, also called the forward curve. Recent theoretical advances show that the term structure can be described by a convenient, affine model specification that lends itself well to state space econometric estimation using the Kalman filter. This chapter asks: Can we better track and forecast the term structure of commodity futures prices with the help of carefully designed filters? Is variation at the very short run only measurement noise or is it economically meaningful? And, given improvements in computing power, how much accuracy is gained by modeling substantially larger, more complicated models?

The evidence presented in this chapter suggests that wavelet thresholding, a class of filtering methods that has been found to be optimal and highly successful in the natural sciences, does not help us understand futures prices. A plausible interpretation is that what appears to be noise in economic data, unlike experimental data, is likely to be meaningful. As a result, larger models may, despite the loss of parsimony, a better approach than filtering to obtain accurate estimates of the term structure of futures prices.

The results also show that while three-factor models are superior to one and two-factor models, it is not clear including a fourth factor improves the results. This finding has practical implications because the number of parameters to be estimated increases faster than does the number of factors in the model. The results also confirm that a non-stationary state variable does not appear warranted, and elimination of this variable is likely to improve convergence of the model. It is difficult to evaluate the significance of market prices of risk. Individually, each parameter is found to be significant, but the sum of all market prices of risk is only weakly different from zero. An potential extension of this work concerns the hypothesis that noise filtered out using wavelet thresholding provides good estimates of the time-varying market prices of risk.

CHAPTER 6

CONCLUSION

This thesis addresses three problems in the literature on commodity futures markets and provides new insights by combining empirical time series analysis with statistical methods derived from wavelet transforms. The common theme to all three essays is the identification of effects that are specifically explained by distinct time horizons of decision-making, from the short-run to the long-run. Although this thesis adopts a particular hierarchy of time horizons (i.e. daily, semiweekly, weekly,…), wavelets allow the researcher to define any hierarchy of time horizons, subject to some conditions, to provide the best analysis for the economic problem under scrutiny.

An introduction to wavelets is presented in Chapter 2 using the lifting scheme approach of Sweldens (1994). After providing an intuitive demonstration of wavelets as building blocks for transformations of the data, we define and explain the most important wavelet properties for time series analysis. Illustrations are provided using the Haar and Daubechies wavelets, which are the two most widely used in this area of research. We show, using results of a simulation study on two typical economic time series, that applying wavelet transforms to the data does not cause loss of statistical information beyond a trivial level of machine precision and moreover does not alter the stationarity of the data.

In Chapter 3, we ask whether findings of long memory in commodity futures prices and price volatility are spurious, and, more generally, test Granger's conjecture that economic and financial time series are not characterized by true long memory. Using a robust wavelet-based estimator, we find that long memory appears to be significant for all commodities and is not overly sensitive to either choice of estimator or to the bias caused by the presence of short memory (e.g. ARMA, GARCH effects). Because standard asymptotic tests have been found to over-reject the null of long memory, we use three recently developed tests of spurious long memory and find that only two out of eleven commodities (wheat and canola) are characterized by true long memory. Certain stochastic break models are known to generate spurious long memory, so we fit the data to a Markov-switching model and show that it provides a good fit.

Several extensions to the chapter appear promising. The long memory models estimated in this work are fractionally integrated ARMA (ARFIMA), but a wavelet-based fractionally integrated GARCH model could be estimated instead and may better capture short memory volatility dynamics. Also, a large number of models can in theory generate spurious long memory. The difficulty of finding out which model provides the best fit is that the different alternatives are generally non-nested, so that traditional Likelihood Ratio, Wald and Score tests are not appropriate. A systematic study of competing models of true and spurious long memory appears warranted.

Chapter 4 asks: Have large Index Traders increased volatility in commodity markets? Should the Commodity Futures Trading Commission consider making permanent its pilot project whereby the positions of Index Traders are reported separately from the positions of other large traders? Without access to confidential CFTC data, we adopt a “revealed” methodology and infer the effect of Index Traders in a joint model of volume-price volatility. Wavelets allow us to filter out all variation in trade volume that is associated with shorter time horizons at which it is known Index Traders are not active. The evidence suggests large Index Traders may have increased price volatility for non-storable commodities (live cattle and lean hogs contracts), but not for storable commodities (grains). In contrast, most of the previous literature has only examined the effect of speculators and found that there is no evidence their trading increases market volatility. The results should be particularly useful in light of the CFTC's current actions and may provide a suitable methodology to examine markets for which confidential trader-level data are not available. A worthwhile extension may be to test structural hypotheses on the theory-motivated differences in production dynamics between storable and non-storable commodities. Production dynamics may well explain why non-storable commodities are influenced by Index Traders.

Chapter 5 considers the problem of modeling the dynamics that explain, each day, the pattern or constellation of futures prices expiring at different maturities. Adopting a recently developed affine term structure model, we ask: Can we better track and forecast the term structure of commodity futures prices and volatility by carefully designing filters to remove from the data what ought to be noise? Are substantially larger and more complex state-space models warranted to obtain a superior fit to the data? The evidence found in this chapter suggests that even wavelet thresholding filters, found to be optimal and highly successful in the natural sciences, do not appear to help in the case of futures data. As a result, using a greater number of factors or state variables appears to be still the best way to improve the results, despite the loss of parsimony and identification difficulties associated with having a very large number of unobservable parameters. Yet it is not clear that including four or more state variables pays off its higher computational cost.

This chapter provides several possible extensions. If noise removed by wavelet thresholding is economically meaningful, it may provide a method to estimate a time-varying market price of risk. Indeed, constant estimates of the market prices of risk are often not significantly different from zero. Difficulties encountered with model convergence suggest that the non-stationary state variable, included to capture permanent economic shocks, should be excluded, particularly since the long-run drift parameter is consistently found to be of negligible size.

In conclusion, this thesis contributes several new findings on timely and persistent questions in commodity derivatives markets, and combines well-established time series analysis with statistical methods based on wavelet transforms to better identify and measure the economic importance of various distinct time horizons in different problems. The thesis shows that using wavelets allows new economic hypotheses to be formally tested and contributes to a better understanding of existing results in the literature.

APPENDIX

A.1 Data Cleaning

Observations for 19 October 1987 ("Black Monday") and 11 September 2001 are considered outliers and removed from the sample.

Soybean prices were allegedly manipulated by the agribusiness giant Feruzzi over the time period May-July 1989 (e.g. Kolb and Overdahl, 2006, p.83-84; Pirrong 2004). For most of the analyses, the data used begin with August 1989.

The measurement unit of corn and soybeans futures contract positions at the Chicago Board of Trade and wheat futures at the Kansas City Board of Trade changed on January 1st 1998 from thousands of bushels to number of contracts, each of which equals five thousand bushels. To ensure consistency in the time series, observations before January 1st 1998 are divided by five, so the unit of measurement throughout is the number of contracts.

The Chicago Mercantile Exchange replaced in 1997 the live hog futures contract (live animal weight-based) with a lean hogs futures contract (carcass weight-based), as a result of which trade volume has increased substantially (Ditsch and Leuthold 1996; Carter and Mohapatra 2006). The new contract is cash settled using a daily price index (weighted average) provided by the USDA and excludes prices from terminal markets. Ditsch and Leuthold (1996) predicted the new contract would provide a better hedge and Carter and Mohapatra (2006) found empirical evidence that the futures contract during its first six years (1998-2004) indeed provided good forecast power

New York Board of Trade cocoa data contained a mistake: the volume and open interest columns were inverted for all observations in 9/2002 and 10/2002. This was corrected before estimation.

Chicago Mercantile Exchange live hogs futures total volume data records data entry errors all for the year 2001: 2/9/2001, 6/22/2001, 7/13/2001, 7/25/2001, 7/25/2001, 8/17/2001, 9/25/2001, 10/16/2001, 10/25/2001, 11/1/2001, 11/5/2001, 11/13/2001, 12/5/2001, 12/14/2001.

KCBOT wheat futures prices are reported in dollars and fractions of a dollar, not cents. Before using these data in any way, they were adjusted into dollars and decimal values (cents).

A.2 Additional Estimation and Test Details

In Chapter 2, the ADF test on the typical futures price time series (corn futures), with no time trend, returns values ranging from -0.76 to -0.93 (one to eight lags), all of which are far smaller (in absolute value) than the critical values (-2.57 to -3.45, 10% to 1% levels of significance). The ADF test including a time trend returns test values ranging from -1.92 to -2.25, all of which are smaller (in absolute value) than the critical values (-3.13 to -3.99, 10% to 1% levels of significance). This version of the test is nearly equivalent to computing the detrended price time series and applying a unit root test (no time trend) on the detrended time series (test values are instead -1.93 to -2.27).

Alternatively, a Variance ratio test (Lo and MacKinlay, 1988, 1989) can be computed to evaluate the null hypothesis of no random walk. This specification test considers, for different levels of time aggregation, the ratio of sample variances, under the assumption that a random walk will display increasing variance as the level of aggregation increases. The test results suggest we cannot reject the null.

The ADF test applied to each wavelet-computed time horizon data provides the following results. For daily variation D_1, test results range from -40.31 to -30.82 (preferred lag selection of six leads to a test value of -37.13), all of which exceed the critical values of -3.99 to -3.14 (10% to 1% levels of significance), and there is no doubt the null of a unit root is rejected. For semiweekly variation D_2, the test results range from -6.48 to 12.76 (9.46 for preferred choice, six lags). The null can only be rejected if the number of lags specified is one or two. Therefore, for a plausible lag specification, we cannot reject the null hypothesis. For weekly variation D_3, the test results range from -7.92 to 5.63 (1.06 for preferred choice, six lags). For biweekly variation D_4, the test results range from -10.25 to 2.09 (-2.41 for preferred choice, six lags). For monthly variation D_5, the test results range from -12.05 to 1.17 (-6.18 for preferred choice, six lags).

In chapter section 4.7, ADF tests show that canola futures trade volume is stationary (test value = -38.456, p<0.01).

In Chapter 5, Augmented Dickey-Fuller tests on the log-price corn futures data return the values: -2.7373^{*}, -2.9595^{**}, -3.1235^{***}, -3.6002^{***}, -3.9984^{***} for each of the six closest maturities, from nearest to most distant. The levels of significance are 10% (*), 5% (**) and 1% (***). The test was computed using an intercept, no time trend, and one month of daily business day lags (20 lags). For the first nearby futures data, we consider the possibility of a time trend and regress the once-differenced log-prices on an intercept, but cannot reject the null hypothesis that this intercept (time trend in levels) is zero.

Estimation of state-space models is done using different procedures in Matlab, R and RATS depending on the desired objective. Linear ARMA full-information estimation by state space is done in R. Constrained optimization procedures are generally done in Matlab. Hidden component state space model estimation using the Kalman filter is done mainly in RATS using the DLM procedure with NONLIN parameter description and constraints and optimization criteria set by NLPAR. Optimization routines are SIMPLEX for the first approximation and BFGS for the actual solution in order to obtain standard errors for the parameters. 200 iterations and 100 sub-iterations are allowed for the BFGS, and up to 5000 trials for the SIMPLEX method. The EXACT diffuse initial conditions of Durbin and Koopmans (2001) are used to control the behavior of the non-stationary component of variance in the Kalman filter procedure. The Kalman gain matrix variance is assumed scaled proportional to the system variances.

The wavelet threshold filtered data contain 16 unfiltered observations at the beginning and end of the sample because the initial and final filtered observations are likely to suffer from boundary effects caused by the wavelet transform.

REFERENCES

Abramovich F., T. Sapatinas and B.W. Silverman (1998). Wavelet thresholding via a Bayesian approach. Journal of the Royal Statistical Society, Ser. B 60, 725-749.

Abramovich, F. and B.W. Silverman (1998). Wavelet decomposition approaches to statistical inverse problems. Biometrika, 85: 115-129.

Abramovich, F. and T. Sapatinas (1999). Bayesian approach to wavelet decomposition and shrinkage. In: Bayesian Inference in Wavelet Based Models (Eds. Muller P. & Vidakovic B.), Lecture Notes in Statistics 141, Springer-Verlag, 33-50.

Abramovich, F., T. Bailey and T. Sapatinas (2000). Wavelet analysis and its statistical applications. Journal of the Royal Statistical Society Series D 49: 1-29.

Abramowitz, M. and I.A. Stegun, eds. (1972). Handbook of mathematical functions with formulas, graphs, and mathematical tables. New York: Dover.

Abry, P. (1997). Ondelettes et turbulence. Multirésolutions, algorithmes de décomposition, invariance d'échelles, Paris: Diderot Editeur.

Abry, P. and D. Veitch (1998). Wavelet analysis of long-range-dependent traffic, Information Theory, IEEE Transactions on, 44:1, 2-15, Jan 1998.

Abry, P., P. Flandrin, M.S. Taqqu and D. Veitch (2002). Self-similarity and long-range dependence through the wavelet lens. In: P. Doukhan, G. Oppenheim and M.S. Taqqu, Eds., Long-Range Dependence Theory and Applications, Boston: Birkhauser, 527-556.

Abry, P., P. Flandrin, M.S. Taqqu and D. Veitch, (2000). Wavelets for the analysis, estimation, and synthesis of scaling data. In: K. Park and W. Willinger, Eds. Self Similar Network Traffic and Performance Evaluation, Wiley: New York, 39-88.

Adelman, I. (1965). Long cycles: fact or artifact? American Economic Review, 60, 443-63.

Adenstedt, R.K. (1974). On large sample estimation for the mean of a stationary random sequence. Annals of Mathematical Statistics, 2, 1095-1107.

Agiakloglou, C., P. Newbold and M. Wohar (1992). Bias in an estimator of the fractional difference parameter. Journal of Time Series Analysis, 14, 235-246.

Ahn, C.M. and H.E. Thompson. (1988). Jump-diffusion processes and the term structure of interest. Journal of Finance, 43, 155-74.

Ait-Sahalia, Y. (2002). Maximum-likelihood estimation of discrete-sampled diffusions: a closed form approximation approach. Econometrica, 70, 223-262.

Ait-Sahalia, Y. (2004). Disentangling diffusion from jumps. Journal of Financial Economics, 74: 487-528.

Akgiray, V. and G. Booth (1988). Mixed diffusion-jump process modelling of exchange rate movements. Review of Economics and Statistics, 70, 631-37.

Akgiray, V., G. Booth, J.J. Hatem and M. Chowdhury (1991). Conditional dependence in precious metal prices. Financial Review, 26, 367-86.

Alizadeh, S., M.W. Brandt and F.X. Diebold (2002). Range-based estimation of stochastic volatility models. Journal of Finance, 57:3, 1047-1091.

Allen, F., S. Morris and H.S. Shin (2003). Beauty contests, bubbles and iterated expectations in asset markets. Preprint.

Alston, J.M. and J.A. Chalfant (1988). Accounting for changes in tastes. Journal of Political Economy, 96(2): 391-410.

Andel, J. (1986). Long memory time series models. Kybernetika, 22, 105-123.

Anderson, D., S. Hamori and N. Hamori (2001). An empirical analysis of the Osaka rice market during Japan's Tokugawa era. Journal of Futures Markets, 21:9, 861-74.

Anderson, K. and B.W. Brorsen (2005). Marketing performance of Oklahoma farmers. American Journal of Agricultural Economics, 87:5, 1265-1270(6).

Anderson, R.W. (1985). Some determinants of the volatility of futures prices. Journal of Futures Markets, 5, 331-48.

Anderson, R.W. and J.-P. Danthine (1983). Hedger diversity. The Economic Journal, 93, 370-389.

Anderson, R.W. and J.-P. Danthine (1983). The Time pattern of hedging and the volatility of futures prices. Review of Economic Studies, 50, 249-266.

Andreou, E. and E. Ghysels (2002). Detecting multiple breaks in financial market volatility dynamics. Journal of Applied Econometrics, 17, 579-600.

Andrews, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. Econometrica, 61, 821-856.

Andrews, D.W.K. (2001). Testing when a parameter is on the boundary of the maintained hypothesis. Econometrica, 69: 683-733.

Andrews, D.W.K. and O. Lieberman (2002). Higher-order improvements of the parametric bootstrap for long memory Gaussian processes. Cowles Foundation, Discussion Paper No. 1378, Yale University.

Andrews, D.W.K. and W. Ploberger (1994). Optimal tests when a nuisance parameter is present only under the alternative. Econometrica, 62, 1383-1414.

Antoniadis A. and D.T. Pham (1998). Wavelet regression for random or irregular design, Comp. Stat. and Data Analysis, 28, 353-369.

Antoniadis, A. (1994), Smoothing noisy data with coiflets, Statistica Sinica, 4:2, 651-678.

Antoniadis, A. and G. Gregoire (1999). Density and hazard rate estimation for right-censored data using wavelet methods, J. R. Statist. Soc. B, 61: 1, 63-84.

Antoniadis, A. and G. Oppenheim, Eds. (1995). Wavelets and statistics. Lecture Notes in Statistics, 103, Springer Verlag.

Antoniadis, A., A. Feuerverger and P. Goncalves (2006). Wavelet based estimation for univariate stable laws. Forthcoming, Annals of the Institute of Statistical Mathematics.

Aoki, M. (1996). A New approach to macroeconomic modeling, New York: Cambridge University Press.

Arneodo, A., F. Argoul, E. Bacry, J. Elezgaray and J.F. Muzy (1995). Ondelettes, multifractales et turbulence. Diderot Editeur: Paris.

Arteche, J. (2002). Semiparametric robust tests on seasonal or cyclical long memory time series. Journal of Time Series Analysis, 23, 251-285.

Arthur, W.B., Yu.M. Ermoliev and Yu.M. Kaniovski (1986). Strong laws for a class of path-dependent urn processes. In: V.I. Arkin, A. Shiryaev and R. Wets, Eds, Proceedings of International Conference on Stochastic Optimisation, Kiev, 1984. Lecture Notes in Control and Information Sciences Vol. 81, Springer: New York.

Avery, C. and P. Zemsky (1998). Multidimensional uncertainty and herd behaviour in financial markets, American Economic Review, 88, 724-748.

Aydogan, K. and G.G. Booth (1988). Are there long cycles in common stock returns? Southern Economic Journal, 55, 141-149.

Bacchetta, P. and E. van Wincoop (2003). Can information heterogeneity explain the exchange rate determination puzzle? Preprint.

Bai, J. (1999). Likelihood ratio tests for multiple structural change. Journal of Econometrics, 91, 299-323.

Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. Econometrica, 66, 47-78.

Bai, J. and P. Perron (2003a). Computation and analysis of multiple structural change models. Journal of Applied Econometrics, 18, 1-22.

Bai, J. and P. Perron (2003b). Critical values for multiple structural change tests. The Econometrics Journal, 6, 72-78.

Baillie, R. T., Y.-W. Han, R. J. Myers and J. Song (2007). Long memory models for daily and high frequency commodity futures returns. Journal of Futures Markets, 27:7, 643-68.

Baillie, R.T. (1996). Long memory processes and fractional integration in econometrics. Journal of Econometrics, 73(1): 5-59.

Baillie, R.T. and C.F. Chung (1993). Small sample bias in conditional sum of squares estimators of fractionally integrated ARMA models. Empirical Economics, 18, 791-806.

Baillie, R.T. and T. Bollerslev (1994). The Long memory of the forward premium. Journal of International Money and Finance, 13, 565-71.

Baillie, R.T., C.-F. Chung and M.A. Tieslau (1996). Analysing inflation by the fractionally integrated ARFIMA-GARCH model. Journal of Applied Econometrics, 11:1, 23-40.

Baillie, R.T., T. Bollerslev, and H.-O. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 73:1, 5-59.

Bakshi, G., C. Cao and Z. Chen (1997). Empirical performance of alternative option pricing models. Journal of Finance, 52: 2003-49.

Bakshi, G., C. Cao and Z. Chen (2000). Pricing and hedging long-term options models. Journal of Econometrics, 94, 277-318.

Ball, C. and W. Torous (1985). On jumps in common stock prices and their impact on call option pricing. Journal of Finance, 40, 155-73.

Banerjee, A. and G. Urga (2005). Modeling structural breaks, long memory and stock market volatility: an overview. Journal of Econometrics, 129, 1-34

Banerjee, A., R.L. Lumsdaine and J.H. Stock (1992). Recursive and sequential tests of the unit root and trend break hypotheses: theory and international evidence. Journal of Business and Economic Statistics, 10, 271-287.

Banerjee, A., S. Lazarova, and G. Urga (1998). Bootstrapping sequential tests for multiple structural breaks, Discussion Paper No.17-98. London Business School.

Bardet, J.M., G. Lang, E. Moulines and P. Soulier (2000). Wavelet estimator of long-range dependence processes. Statistical Inference for Stochastic Processes, 3, 85-99.

Bardet, J.-M., G. Lang, G. Oppenheim, A. Philippe, S. Stoev, and M.S. Taqqu (2003). Generators of long-range dependence processes: a survey Theory and applications of long-range dependence. Birkhäuser.

Barkoulas, J. and C.F. Baum (1996). Long-term dependence in stock returns. Economics Letters, 53, 253-259.

Barkoulas, J., W.C. Labys and J. Onochie (1997). Fractional dynamics in international commodity price. Journal of Futures Markets, 17, 161-89.

Barkoulas, J.T., C.B. Baum and M. Caglayan (1999). Long memory or structural breaks: can either explain nonstationary real exchange rates under the current float? Journal of International Financial Markets, Institutions and Money, 9, 359-376.

Barksdale, H.C., J.E. Hilliard and M.C. Ahlund (1975). A cross-spectral analysis of beef prices. American Journal of Agricultural Economics, 57:2, 309-315.

Barnett, W., E. Offenbacher and P. Spindt (1984). The new Divisia monetary aggregates. Journal of Political Economy, 92:6, 1049-1085.

Barnhill, T. M., J. V. Jordan and W. E. Seal (1987). Maturity and refunding effects on treasury-bond futures price variance. Journal of Financial Research, 10, 121-131.

Barrow, S. (1994). Fundamental currency forecasting an agnostic view. Chemical Working Papers in Financial Economics, 3, 8-12.

Basseville, M. and I.V. Nikiforov (1993). Detection of abrupt changes: theory and applications. Englewood Cliffs, NJ: Prentice-Hall.

Bassingthwaighte, J. B. and G. M. Raymond (1994). Evaluating rescaled range analysis for time series. Ann. Biomed. Eng., 22, 432-444.

Bates, D.S. (1996a). Dollar jump fears, 1984-1992. Distributional abnormalities implicit in currency futures options. Journal of International Money and Finance, 15, 65-93.

Bates, D.S. (1996b). Jumps and stochastic volatility. Exchange rate processes implicit in Deutsche mark options. Review of Financial Studies, 9, 69-107.

Bates, D.S. (2000). Post-87 crash fears in S&P 500 futures options. Journal of Econometrics, 94, 181-238.

Baum, C.F., J. T. Barkoulas and M. Caglayan (1997). Persistence in international inflation rates, working paper No. 333, Boston College.

Beck, S. (2001). Autoregressive conditional heteroscedasticity in commodity spot prices. Journal of Applied Econometrics, 16, 115-32.

Benedetto J. and M. W. Frazier, Eds. (1994). Wavelets: mathematics and applications. CRC Press.

Beran, J. (1994a). Statistics for long-memory processes. New York: Chapman & Hall.

Beran, J. (1994b). On a class of M-estimators for Gaussian long memory models. Biometrika, 81, 755-766.

Beran, J. and N. Terrin (1996). Testing for a change of the long-memoryparameter. Biometrika, 83, 627-638.

Berkes, I. and L. Horváth (2003). Limit results of the empirical process of squared residuals in GARCH models. Stochastic Processes and their Applications, 105, 271-298.

Berkes, I., E. Gombay, L. Horváth and P.S. Kokoszka (2002). Sequential change-point detection in GARCH(p,q) models. Economic Theory, forthcoming.

Berkes, I., L. Horváth and P.S. Kokoszka (2003). Testing for parameter constancy in GARCH (p, q) models. Preprint.

Bernard, J.-T., L. Khalaf, M. Kichian and S. McMahon (2006). Forecasting commodity prices: GARCH, jumps, and mean reversion. Bank of Canada working paper 2006-14.

Bessembinder, H. (1992). Systematic risk, hedging pressure, and risk premiums in futures markets. Review of Financial Studies, 5, 637-67.

Bessembinder, H., J. F. Coughenour, P. J. Seguin and M. M. Smoller (1996). Is there a term structure of futures volatilities? Reevaluating the Samuelson hypothesis. Journal of Derivatives, Winter, 45-58.

Bessembinder, H., J. F. Coughenour, P. J. Seguin and M. M. Smoller (1995). Mean reversion in equilibrium assets prices: evidence from the futures term structure. Journal of Finance, 50, 361-375.

Beylkin, G. (1993). Wavelets and fast numerical algorithms. In I. Daubechies, Ed., Different perspectives on wavelets (San Antonio, TX), Proc. Sympos. Providence, RI: Amer. Math. Soc. Appl. Math., 47, 89-117.

Beylkin, G., R. Coifman and V. Rokhlin (1991). Fast wavelet transforms and numerical algorithms. I. Comm. Pure Appl. Math., 44:2, 141-183.

Bhansali, R.J. and P.S. Kokoszka (2001). Estimation of the long-memory parameter: a review of recent developments and an extension. In: Basawa, I.V., Heyde, C.C., Taylor, R.L. (Eds.), Proceedings of the Symposium on Inference for Stochastic Processes, 125-150.

Bhansali, R.J., M.P. Holland and P.S. Kokoszka (2003). Chaotic maps with slowly decaying correlations and intermittency. Fields Institute Communications, forthcoming.

Bhansali, R.J., M.P. Holland and P.S. Kokoszka (2004). Intermittency, long-memory and financial returns. In: Teyssière, G., Kirman, A. (Eds.), Long-Memory in Economics. Berlin: Springer, in press.

Bilodeau, J.F. (1998). Analyse de processus de sauts dans les prix du petrole brut. Masters Thesis, Universite Laval.

Bird, A. (1990). Price determination in the market for aluminium. Primary commodity prices: Economic models and policy. Cambridge, New York and Melbourne: Cambridge University Press.

Birgé, L. and P. Massart (1997). From model selection to adaptive estimation, in D. Pollard (ed), Festchrift for L. Le Cam, Springer, 55-88.

Black, D. (1986). Success and failure of futures contracts: Theory and empirical evidence, monograph series in finance and economics, Monograph #1986-1, Salomon Brothers Center, Graduate School of Business, New York University.

Black, F. (1976). The pricing of commodity contracts. Journal of Financial Economics, 3, 167-179.

Black, F. (1986). Noise. Journal of Finance, 41, 529-543.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81, 637-659

Blanchard, O. and M.W. Watson (1982). Bubbles, rational expectations and financial markets. In: P. Wachtel, Ed., Crisis in the Economic and Financial System, Lexington, MA: Lexington Books, 295-315.

Blank, S.C. (1991). Chaos in futures markets? A nonlinear dynamical analysis. Journal of Futures Markets, 11, 718-728.

Bloomfield, P. (1976). Fourier analysis of time series: An Introduction, New York: John Wiley and Sons.

Blume, M.E. and R. F. Stambaugh (1983). Biases in computed returns: An application to the size effect. Journal of Financial Economics, 12, 13-32.

Board, J.L.G. and C.M.S. Sutcliffe (1990). Information, volatility, volume and maturity: An investigation of stock index futures. Review of Futures Markets, 9, 532-549.

Bodie, Z. and V. Rosansky (1980). Risk and return in commodity futures. Financial Analysts Journal, May/June, 27-39.

Boggess, A. and F.J. Narcowich (2001). A first course in wavelets with Fourier analysis, Upper Saddle River, New Jersey: Prentice-Hall.

Böhringer, C., A. Löschel and T. F. Rutherford (2007). Decomposing the integrated assessment of climate change. Journal of Economic Dynamics and Control, 31:2, 683-702.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31, 307-27.

Bollerslev, T. and J.W. Wright (2000). Semiparametric estimation of long-memory volatility dependencies: the role of high-frequency data. Journal of Econometrics, 98, 81-106.

Bollerslev, T., and H.O. Mikkelsen (1996). Modeling and pricing long memory in stock market volatility. Journal of Econometrics, 73, 151-84.

Bollerslev, T., and J.M. Wooldridge (1992). Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances. Econometric Reviews, 11, 143-72.

Bollerslev, T., R.F. Engle and D.B. Nelson (1994). ARCH models. In: R.F. Engle and D. McFadden, Eds. Handbook of Econometrics, vol. IV. Amsterdam: Elsevier Science B. V., 2959-3038.

Bollerslev, T., R.Y. Chou and K.F. Kroner (1992). ARCH modeling in finance: A review of the theory and empirical evidence. Journal of Econometrics, 52:1, 5-59.

Booth, G.G. and Y. Tse (1995). Long memory in interest rate futures markets: A fractional cointegration analysis. Journal of Futures Markets, 15, 573-84.

Booth, G.G., F.R. Kaen and P.F. Koveos (1982a). Persistent dependence in gold prices. J. Finance Research, Spring, 85-93.

Booth, G.G., F.R. Kaen and P.F. Koveos (1982b). R/S analysis of foreign exchange rates under two international monetary regimes. J. Monetary Economics, 10, 407-415.

Bos, C., S.P.H. Franses and M. Ooms (1999). Long-memory and level shifts: re-analyzing inflation rates. Empirical Economics, 24, 427-449.

Bowden, R. and J. Zhu (2006). The agribusiness cycle and its wavelets. Victoria University of Wellington School of Economics and Finance, working paper.

Box, G.E.P. and G.M. Jenkins (1976). Time series analysis: Forecasting and control, Revised Edition, San Francisco: Holden-Day.

Boyle, P.P. and T. Wang (1999). The valuation of new securities in an incomplete market: The catch-22 of derivative pricing Working Paper, University of Waterloo.

Boyle, P.P., M. Broadie and P. Glasserman (1997). Monte Carlo methods in security pricing. Journal of Economics, Dynamics and Control, 21, 1267-1327.

Brandt, M.W. and C. Jones (2006). Forecasting volatility with range-based EGARCH models. Journal of Business and Economic Statistics, 24: 470-486.

Breidt, J.F., N. Crato and P.J.F. de Lima (1998). On the detection and estimation of long-memory in stochastic volatility. Journal of Econometrics, 83, 325-348.

Brislawn, C.M. (1995). Fingerprints to digital, Notices of the AMS, 42, 1278-1283.

Brock, W.A. and C.H. Hommes (1997). A rational route to randomness. Econometrica, 65, 1059-1095. MathSciNet | Full Text via CrossRef.

Brock, W.A. and C.H. Hommes (1999). Rational animal spirits. In: P.J.J. Herings, G. van der Laan and A.J.J. Talman, Eds, The Theory of Markets, Amsterdam: North-Holland, 109-137.

Brockwell, P.J. and R.A. Davis (1991). Time series: Theory and methods, 2nd ed. New York: Springer.

Brodsky, B. and B. Darkhovsky (1993). Nonparametric methods in change-point problems. The Netherlands: Kluwer Academic Publishers.

Brorsen, B.W. and S.R. Yang (1994). Nonlinear dynamics and the distribution of daily stock index returns. Journal of Financial Research, 17, 187-203.

Burke Hubbard, B. (1996). The world according to wavelets, AK Peters, Wellesley. The French original version is titled Ondes et Ondelettes. La saga d'un outil mathématique, Pour la Science, (1995).

Burrus, C.S, R.A. Gopinath and H. Guo (1998). Introduction to wavelets and wavelet transforms: A Primer, Upper Saddle River, New Jersey: Prentice-Hall.

Burton, M. (1993). Some illustrations of chaos in commodity models. Journal of Agricultural Economics, 44, 38-49.

Busetti, F. and A. Harvey (2001). Testing for the presence of a random walk in series with structural breaks. Journal of Time Series Analysis, 22, 127-150.

Busetti, F. and A. Harvey (2003). Further comments on stationarity tests in series with structural breaks. Journal of Time Series Analysis, 24, 137-140.

Caginalp, G., D. Porter and V.L. Smith (1998). Experimental asset markets. In: P.E. Earl and S. Kemp, Eds, The Elgar Companion to Consumer Research and Economic Psychology, Edwar Elgar, Cheltenham, UK, 215-222.

Campbell, J.Y. and R.J. Shiller (1987). Cointegration and tests of present value models. Journal of Political Economy, 95, 1062-1088. Full Text via CrossRef.

Cannon, M.J., D.B. Percival, D.C. Caccia, G.M. Raymond and J.B. Bassingthwaighte (1996). Evaluating scaled windowed variance methods for estimating the Hurst coefficient in time series. Physica A, 241, 606-626.

Carlton, D. (1984). Futures markets: Their purpose, their history, their growth, their successes and failures. Journal of Futures Markets, 4, 237-271.

Carmona R., W.L. Hwang and B. Torresani (1998). Practical time-frequency analysis: Gabor and wavelet transforms with an implementation, San Diego, CA: S. Academic Press.

Carter, C. A. (1999). Commodity futures markets: A survey. Australian Journal of Agricultural and Resource Economics, 43, 209-247.

Cashin, P. and C. J. McDermott (2002). The long-run behavior of commodity prices: Small trends and big variability. IMF Staff Papers, 49, 175-99.

Castelino, M.G. and J. C. Francis (1982). Basis speculation in commodity futures: The maturity effect. Journal of Futures Markets 2, 195-206.

Chamberlain, T.W. (1989). Maturity effects in futures markets: Some evidence from the city of London. Scottish Journal of Political Economy, 90-95.

Chambers, M. J. (1998). Long memory and aggregation in macroeconomic time series. International Economic Review, 39, 1053-1072.

Chan, N.H. and L.T. Tran (1989). On the first order autoregressive process with infinite variance. Econometric Theory, 5, 354-362. MathSciNet.

Chan, N.H., Terrin, N. (1995). Inference for unstable long-memoryprocesses with applications to fractional unit root autoregression. Annals of Statistics, 23, 1662-1683.

Chatrath, A. and F. Song (1999). Futures commitments and commodity price jumps. Financial Review, 34, 95-112.

Chavas, J.-P. and M.T. Holt (1991). Market instability and nonlinear dynamics. American Journal of Agricultural Economics, 73, 819-28.

Chavas, J.-P. and M.T. Holt (1993). On nonlinear dynamics: The case of the pork cycle. American Journal of Agricultural Economics, 75, 113-20.

Chen, W.W. and C.M. Hurvich (2003). Estimating fractional cointegration in the presence of polynomial trends. Journal of Econometrics, 117, 95-121.

Chernov, M., A.R. Gallant, E. Ghysels and G. Tauchen (2003). Alternative models for stock price dynamics. Journal of Econometrics, 116, 225-57.

Cheung, Y.W. and F.X. Diebold (1994). On maximum likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean. Journal of Econometrics, 62:2, 301-316.

Cheung, Y.-W. and K.S. Lai (1995). A search for long memory in international stock market returns. Journal of International Money and Finance, 14, 597-615.

Cheung, Y.-W., M.D. Chinn and A.G. Pascual (2002). Empirical exchange rate models in the nineties: are they fit to survive. NBER. Working Paper, 9393.

Chionis, D. and R. MacDonald (2002). Aggregate and disaggregate measures of the foreign exchange risk premium, International Review of Economics and Finance, 11, 57-84.

Christensen, B.J. and M.O. Nielsen (2002). Semiparametric analysis of stationary fractional cointegration and the implied-realized volatility relation. University of Aarhus Working Paper No. 2001-4.

Chui, C.K. (1992a), Wavelets: a tutorial in theory and applications, Academic Press.

Chui, C.K. (1992b), An introduction to wavelets, Academic Press.

Chung, C. (1996a). Estimating a generalized long memory process. Journal of Econometrics, 73, 237-259.

Chung, C. (1996b). A generalized fractionally integrated autoregressive moving-average process. Journal of Time Series Analysis, 17, 111-140.

Clements, M.P. and D.F. Hendry (1996). Intercept corrections and structural change. Journal of Applied Econometrics, 11, 475-494.

Clements, M.P. and D.F. Hendry (1998). Forecasting economic time series. Cambridge: Cambridge University Press.

Cohen, A. (1992), Ondelettes, analyses multirésolution et traitement numérique du signal, Ph.D. thesis, University of Paris IX, Dauphine.

Cohen, A. (1995), Wavelets and multiscale signal processing, Chapman and Hall.

Cohen, A., I. Daubechies and J.C. Feauveau (1992). Biorthogonal basis of compactly supported wavelets. Communications in Pure and Applied Mathematics, 45, 485-560.

Cohen, A., I. Daubechies, B. Jawerth and P. Vial (1993). Multiresolution analysis, wavelets and fast wavelet transform on an interval, CRAS Paris, Ser. A, t. 316, 417-421.

Coifman, R., Y. Meyer, S. Quaker and V. Wickerhauser (1990). Signal processing and compression with wave packets. Numerical Algorithms Research Group, New Haven, CT: Yale University.

Coifman, R.R. and D.L. Donoho (1995). Translation invariant de-noising, Lecture Notes in Statistics, 103, 125-150.

Coifman, R.R. and M.V Wickerhauser (1992). Entropy-based algorithms for best basis selection. IEEE Trans. on Inf. Theory, 38:2, 713-718.

Coifman, R.R.; Y. Meyer, M.V. Wickerhauser (1992). Wavelet analysis and signal processing, in Wavelets and their applications, M.B. Ruskai et al. (Eds.), 153-178, Jones and Bartlett.

Comte, F., Coutin, L., Renault, E.E. (2003). Affine fractional stochastic volatility models with application to option pricing, Discussion paper no. 2003-10, Universite de Montreal.

Comte, F., Renault, E.G. (1996). Long memory continuous time models. Journal of Econometrics, 73, 101-149.

Comte, F., Renault, E.G. (1998). Long memory in continuous-time stochastic volatility models. Mathematical Finance, 8, 291-323.

Copeland, T.E. (1976). Model of asset trading under the assumption of sequential information arrival. Journal of Finance, 31, 1149-68.

Corazza, M., A.G. Malliaris and C. Nardelli (1997). Searching for fractal structure in agricultural futures markets. Journal of Futures Markets, 17:4, 433-473.

Cornew, R.W., D.E. Town and L.D. Crowson (1984). Stable distributions, futures prices, and the measurement of trading performance. Journal of Futures Markets, 4, 531-57.

Cortazar, G. and E.S. Schwartz (2003). Implementing a stochastic model for oil future prices. Energy Economics 25, 215-38.

Cox, J.C., J.E. Ingersoll and S.A. Ross (1985). An intertemporal general equilibrium model of asset prices. Econometrica, 53, 363-384.

Craigmile, P.F. (2003). Simulating a class of stationary Gaussian processes using the Davies-Harte algorithm, with application to long memory processes. Journal of Time Series Analysis, 24:5, 505-511.

Craigmile, P.F. (2005). Approximate wavelet based simulation of long memory processes. Journal of Statistical Computation and Simulation, 75:5, 363-80.

Craigmile, P.F. and D. B. Percival (2002). Wavelet-based trend detection and estimation Entry in the Encyclopedia of Environmetrics, edited by A. El-Shaarawi and W. W. Piegorsch, Chichester, UK: John Wiley & Sons.

Craigmile, P.F. and D.B. Percival (2005). Asymptotic decorrelation of between-scale wavelet coefficients. IEEE Transactions on Information Theory, 51:3, 1039-1048.

Craigmile, P.F. and D.M. Titterington (1997). Parameter estimation for finite mixtures of uniform distributions. Communications in Statistics: Theory and Methods, 26:8, 1981-95.

Craigmile, P.F. and W.M. King (2004). Periodogram based tests for distortion products otoacoustic emission. The Journal of the Acoustical Society of America, 116:1, 442-451.

Craigmile, P.F., D.B. Percival and P. Guttorp (2001). The impact of wavelet coefficient correlations on fractionally differenced process estimation. European Congress of Mathematics (Barcelona, July 10-14, 2000), Volume II, edited by C. Casacuberta, R. M. Miro-Roig, J. Verdera and S. Xambo-Descamps, Basel, Switzerland: Birkhauser Verlag, 591-99.

Craigmile, P.F., P. Guttorp and D.B. Percival (2004). Trend assessment in a long memory dependence model using the discrete wavelet transform. Environmetrics, 15:4, 313-335.

Craigmile, P.F., P. Guttorp and D.B. Percival (2005). Wavelet based estimation for polynomial contaminated fractionally differenced processes. IEEE Transactions on Signal Processing, 53:8, 3151-3161.

Crato, N. and B.K. Ray (2000). Memory in returns and volatilities of futures contracts. Journal of Futures Markets, 20, 525-43.

Cressie, N., B.E. Buxton, C.A. Calder, P.F. Craigmile, C. Dong, N.J. McMillan, M. Morara, T.J. Santner, K. Wang, G. Young and J. Zhang (2006). From sources to biomarkers: A hierarchical Bayesian approach for human exposure modeling, forthcoming in the Journal of Statistical Planning and Inference.

Cressie, N., J. Zhang and P. F. Craigmile (2005). Geostatistical prediction of spatial extremes and their extent. Geostatistics for Environmental Applications, Proceedings of the Fifth European Conference on Geostatistics for Environmental Applications, edited by P. Renard, H. Demougeot-Renard, R. Froidevaux, Springer, 27-37.

Cromwell, J.B., W.C. Labys and E. Kouassi (2000). What color are commodity prices?: A fractal analysis. Empirical Economics, 25, 563-580.

Crowley, P. (2007). An Intuitive guide to wavelets. Journal of Economic Surveys, forthcoming.

Crownover, R.M. (1995). Introduction to fractals and chaos. London, U.K: Tones and Bartlett Publishers.

Csörgö, M. and L. Horváth (1997). Limit theorems in change-point analysis, New York: Wiley.

Cutland, N.J., P.E. Kopp and W. Willinger (1995). Stock price returns and the Joseph Effect: A Fractional Version of the Black-Scholes Model. Progress in Probabilities, 36, 27-351.

Daal, E., J.B. Farhat and P.P. Wei (2003). Reexamining the maturity effect using extensive futures data. Working paper 2003-06, University of New Orleans, Department of Economics and Finance.

Dalla, V. and J. Hidalgo (2005). A parametric bootstrap test for cycles. Journal of Econometrics, 129, 219-261.

Danthine, J.-P. (1978). Information, futures prices and stabilizing speculation. Journal of Economic Theory, 17, 79-98.

Das, S.R. (2002). The surprise element: Jumps in interest rates. Journal of Econometrics, 106, 27-65.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. Communications on Pure and Applied Mathematics, 41, 909-96.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. IEEE Transactions on Information Theory, 36, 961-1005.

Daubechies, I. (1992). Ten lectures on wavelets. Society for Industrial and Applied Mathematics, Philadelphia.

Daubechies, I. (1993). Orthonormal bases of compactly supported wavelets II. Variations on a theme. SIAM J. Math. Anal., 24:2, 499-519.

Davidson, J. (2002). A model of fractional cointegration, and tests for cointegration using the bootstrap. Journal of Econometrics, 110, 187-212.

Davidson, J. and T.T. Terasvirta, Eds. (2002). Long memory and nonlinear time series. Journal of Econometrics, 110:2, 105-437.

Davidson, J.E. (2004). Moment and memory properties of linear conditional heteroskedasticity models, and a new model. Journal of Business and Economics Statistics, 22:1, 16-29.

Davidson, R., W.C. Labys, J.-B. Lesourd (1997). Wavelet analysis of commodity price behavior. Computational Economics, 11:1-2, 103 - 128,

Davies, R.B. and D.S. Harte (1987). Tests for the Hurst effect. Biometrika, 74, 95-102.

De Long, J.B., A. Schleifer, L.H. Summers and R.J. Waldmann (1989). The size and incidence of the losses from noise trading. Journal of Finance 44, 681-699.

de Peretti, C. (2003). Bilateral bootstrap tests for long memory: an application to the silver market. Computational Economics, 22, 187-212.

de Peretti, C. and G. Urga (2004). Stopping tests in the sequential estimation for multiple structural breaks. Working Paper CEA-06-2004, Centre for Econometric Analysis, London: Cass Business School.

Deaton, A. and G. Laroque (1992). On the behavior of commodity prices. Review of Economic Studies, 59, 1-23.

Deaton, A. and G. Laroque (1995). Estimating nonlinear rational expectations commodity price models. Journal of Applied Econometrics, 10, 510-540.

Decoster, G.P., W.C. Labys and D.W. Mitchell (1992). Evidence of chaos in commodity futures prices. Journal of Futures Markets, 12, 291-305.

DeLong, J.B., A. Shleifer, L. Summers and R.J. Waldman (1990). Noise trader risk in financial markets. Journal of Political Economy, 98, 703-38.

Den Haan, W.J. and A. Levin (1997). A practitioners guide to robust covariance matrix estimation. In: G.S. Maddala and C.R. Rao, Editors, Amsterdam: North-Holland, Handbook of Statistics, 15, 291-341.

DeVore, R.A., B. Jawerth and B.J. Lucier (1992). Image compression through wavelet transform coding. IEEE Trans. on Inf. Theory, 38: 2, 719-746.

Diba, B.T. and H.I. Grossman (1988). Explosive rational bubbles in stock prices? American Economic Review, 78, 520-530.

Dickey, D.A. and W.A. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74, 427-431.

Diebold, F. X. and G. D. Rudebusch (1989). Long memory and persistence in aggregate output. Journal of Monetary Economics, 24, 189-209.

Diebold, F. X. and G. D. Rudebusch (1991). Is consumption too smooth? Long memory and the Deaton paradox. Review of Economics and Statistics, 71, 1-9.

Diebold, F.X. and A. Inoue (2001). Long memory and regime switching. Journal of Econometrics, 105, 131-159.

Diebold, F.X. and G.D. Rudebusch (1991). On the power of Dickey-Fuller tests against fractional alternatives. Economics Letters, 35, 155-160.

Ding, Z. and C.W.J. Granger (1996). Modelling volatility persistence of speculative returns. A new approach. Journal of Econometrics, 73, 185-215.

Dirac, P.A.M. (1958). The Principles of quantum mechanics, Fourth Edition, Oxford: Oxford University Press.

Dittmann, I. (2000). Residual-based tests for fractional cointegration: a Monte Carlo study. Journal of Time Series Analysis, 21, 614-647.

Dittmann, I. (2004). Error correction models for fractionally cointegrated time series. Journal of Time Series Analysis, 25, 27-32.

Dolado, J.J., J. Gonzalo and L. Mayoral (2002). A Fractional Dickey-Fuller test for unit toots. Econometrica, 70, 1963-2006.

Donaldson, R.G. and M. Kamstra (1996). A New dividend procedure that rejects bubbles in asset prices the case of the 1929's stock crash. The Review of Financial Studies, 9, 333-383.

Donoho, D.L. (1993). Progress in wavelet analysis and WVD: a ten minute tour, in Progress in wavelet analysis and applications, Y. Meyer, S. Roques, 109-128. Frontières Ed.

Donoho, D.L. (1995). De-Noising by soft-thresholding. IEEE Trans. on Inf. Theory, 41:3, 613-627.

Donoho, D.L. and I.M. Johnstone (1994a). Ideal de-noising in an orthonormal basis chosen from a library of bases, CRAS Paris, Ser I, t. 319, 1317-1322.

Donoho, D.L. and I.M. Johnstone (1994b). Ideal spatial adaptation by wavelet shrinkage. Biometrika, 81, 425-455.

Donoho, D.L., I.M. Johnstone, G. Kerkyacharian and D. Picard (1995). Wavelet shrinkage: asymptopia, Jour. Roy. Stat. Soc., series B, 57:2, 301-369.

Donoho, D.L., I.M. Johnstone, G. Kerkyacharian and D. Picard (1996). Density estimation by wavelet thesholding, Annals of Stat., 24, 508-539.

Doornik, J.A. and D.F. Hendry (2001). Econometric modelling using PcGive, vol. III. London: Timberlake Consultants Press.

Doornik, J.A. and M. Ooms (2003). Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. Computational Statistics and Data Analysis, 41, 333-348.

Doornik, J.A. and M. Ooms (2004). Inference and forecasting for ARFIMA models with an application to US and UK inflation. Studies in Nonlinear Dynamics and Econometrics, 8:2, Article 14. http:// www.bepress.com/snde/vol8/iss2/art14.

Doukhan, P., G. Oppenheim and M.S. Taqqu, Eds. (2003). Theory and applications of long-range dependence. Berlin: Birkhauser.

Drost, F.C., T.E. Nijman and B.J.M. Werker (1998). Estimation and testing in models containing both jump and conditional Heteroskedasticity. Journal of Business and Economic Statistics, 16, 237-43.

Dueker, M. and R. Startz (1998). Maximum-likelihood estimation of fractional cointegration with an application to US and Canadian bond rates. Review of Economics and Statistics, 83, 420-426.

Dufour, J-M. and E. Ghysels, Eds. (1996). Recent developments in the econometrics of structural change. Journal of Econometrics, 70, 1-316.

Dusak, K. (1973). Futures trading and investor returns: An investigation of commodity market Risk Premiums. Journal of Political Economy, 81, 1387-406.

Dusak-Miller, K. (1979). The relation between volatility and maturity in futures contracts, Leuthold, R.M (ed.) Commodity Markets and Futures Prices, Chicago Mercantile Exchange, 25-36.

Elder, J. (2001). F versus t tests for unit roots (with Peter E. Kennedy). Economics Bulletin, 3:3, 1-6.

Elder, J. (2004). More on F versus t tests for unit roots when there is no trend (with Peter E. Kennedy), Economics Bulletin, 3:37, 1-6.

Elder, J. and H.J. Jin (2006). Fractional integration in commodity futures, forthcoming Journal of Futures Markets.

Elliott, G. (2001). Confidence intervals for autoregressive coefficients near one. Journal of Econometrics, 103, 155-81.

Elliott, G. (2003). Testing for unit roots with stationary covariates. Journal of Econometrics, 115, 75-89.

Elliott, G. and U. Muller (2003). Tests for unit roots and the initial condition. Econometrica, 71, 1269-1286.

Elliott, G. and U. Muller (2004). Minimizing the impact of the initial condition on testing for unit roots, mimeo. Forthcoming in Journal of Econometrics.

Elliott, G. and U. Muller (2004).Confidence sets for the date of a single break in linear time series regressions, mimeo.

Elliott, G., T.J. Rothemberg and J.H. Stock (1996). Efficient tests for an autoregressive unit root. Econometrica, 64:4, 813-836.

Ellison, G. and D. Fudenberg (1993). Rules of thumb for social learning. Journal of Political Economy, 101, 612-643. Full Text via CrossRef.

Embrechts, P., C. Klüppelberg and T. Mikosch (1997). Modelling extremal events for insurance and finance, Berlin: Springer.

Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50, 987-1007. MathSciNet | Full Text via CrossRef.

Engle, R.F. and A.D. Smith (1999). Stochastic permanent breaks. Review of Economics and Statistics, 81, 553-574.

Engle, R.F. and C.W.J. Granger (1987). Co-integration and error correction: representation, estimation and testing. Econometrica, 55, 251-276.

Engle, R.F. and T. Bollerslev (1986). Modeling the persistence of conditional variance. Econometric Reviews, 5, 1-50.

Erb, C. B. and C. R. Harvey (2005). The tactical and strategic value of commodity futures, forthcoming Financial Analysts Journal.

Evans, G.W. (1986). A test for speculative bubbles in the sterling dollar exchange rate. American Economic Review, 76, 621-636.

Evans, G.W. (1991). Pitfalls in testing for explosive bubbles in asset prices. American Economic Review, 81, 922-930.

Fama, E. and K. French (1987). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. Journal of Business

Fama, E. and K. French, (1988). Business cycles and behavior of metals prices. Journal of Finance, 43, 1075-1093.

Fama, E.F. (1981). Stock returns, real activities, inflation and money. American Economic Review, 71, 545-65.

Fama, E.F. and G.W. Schwert (1977). Asset returns and inflation. Journal of Financial Economics, 5, 115-146.

Fama, E.F. and K.R. French (1988). Permanent and temporary components of stock prices. Journal of Political Economy, 96:2, 246-273.

Fan, Y and R. Gencay (2006). Unit root and cointegration tests with wavelets. Vanderbilt University Department of Economics working paper.

Feder, J. (1988). Fractals. New York, NY: Plenum Publishing Corp.

Figlewski, S. (1997). Forecasting volatility. Financial Markets, Institutions, and Instruments, 6, 2-87.

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional Brownian motion. IEEE Transactions on Information Theory, 38, 910-917. MathSciNet | Full Text via CrossRef.

Flood, R.P. and P.M. Garber (1980). Market fundamentals versus price level bubbles the first tests. Journal of Political Economy, 88, 745-770. Full Text via CrossRef.

Flood, R.P. and R.J. Hodrick (1990). On testing for speculative bubbles. Journal of Economic Perspectives, 4, 85-102.

Fox, R. and M. S. Taqqu (1986). Large sample properties of parameter estimates for strongly dependent Gaussian time-series. Annals of Statistics, 14, 517-532.

Frankel, J. and A.K. Rose (1995). Empirical research on nominal exchange rates. In: G.M. Grossman and K. Rogoff, Eds, Amsterdam: North-Holland. Handbook of International Economics, 3, 1680-1729.

Frankel, J.A. and K.A. Froot (1986). The dollar as an irrational speculative bubble the tale of fundamentalists and chartists. The Marcus Wallenberg Papers on International Finance, 1, 27-55.

Franses, P.H., M. van der Leij and R. Paap (2002). Modelling and forecasting level shifts in absolute returns. Journal of Applied Econometrics, 17, 601-616.

Frechette, D.L. (2001). Aggregation and the nature of price expectations. American Journal of Agricultural Economics, 83:1, 52-63

French, A.P. and E.F. Taylor (1978). An introduction to quantum physics, Wokingham: Van Nostrand Reinhold.

French, K.R. (1986). Detecting spot price forecasts in futures prices. Journal of Business, 59, S39-S54.

Friedman, B. and K. Kuttner (1992). Money, income, prices, and interest rates. American Economic Review, 472-492.

Fung, H.-G. and W.-C. Lo. (1993). Memory in interest rate futures. Journal of Futures Markets 13:8, 865-72.

Gagliardini, P., F. Trojani and G. Urga (2004). Robust GMM tests for structural breaks. Journal of Econometrics, 10.1016/j.jeconom.2004.09.006.

Galloway, T. M. and R. W. Kolb (1996). Futures prices and the maturity effect. Journal of Futures Markets, 7, 809-828.

Gao, A.H. and G.H.K. Wang (1999). Modeling nonlinear dynamics of daily futures price changes. Journal of Futures Markets, 19:3, 325-351.

Garber, P.M. (2000). Famous first bubbles, Cambridge, MA: MIT Press.

Garman, M.B. (1977). A General theory of asset valuation under diffusion state processes Working Paper, Berkeley: University of California.

Garman, M.B. and M.J. Klass (1980). On the estimation of security price volatilities from historical data. Journal of Business, 53:1, 67-78.

Gelb, A. H. (1979). A Spectral analysis of coffee price oscillations. International Economic Review, 20, 495-514.

Gencay, R., F. Selcuk and B. Whitcher (2001). Scaling properties of foreign exchange volatility. Physica A: Statistical Mechanics, 289:1-2, 249-266.

Gencay, R., F. Selcuk and B. Whitcher (2003). Systematic risk and timescales. Quantitative Finance, 3:2, 108-116.

Gencay, R., F. Selcuk and B. Whitcher (2004). Asymmetry of information flow between volatilities across time scales. Simon Fraser University, Department of Economics working paper.

Geweke J. and S. Porter-Hudak (1983). The estimation and application of long memory time series models. Journal of Time Series Analysis, 4, 221-238.

Gibson, R. and E.S. Schwartz (1990). Stochastic convenience yield and the pricing of oil contingent claims. Journal of Finance, 45, 959-76.

Gil-Alana, L.A. (2003). Testing of fractional cointegration in macroeconomic time series. Oxford Bulletin of Economics and Statistics, 65, 517-529.

Gil-Alana, L.A. (2004). A joint test of fractional integration and structural breaks at a known period of time. Journal of Times Series Analysis, 25, 681-700.

Gilbert, C.L. (1995). Modelling market fundamentals: A model of the aluminium market. Journal of Applied Econometrics, 10, 385-410.

Giraitis, L., and P.M. Robinson (2001). Whittle estimation of ARCH models. Econometric Theory, 17, 608-631.

Giraitis, L., P.S. Kokoszka and R. Leipus (2000a). Stationary ARCH models dependence structure and central limit theorem. Econometric Theory, 16, 3-22. Full Text via CrossRef.

Giraitis, L., P.S. Kokoszka, R. Leipus and G. Teyssière (2000b). Semiparametric estimation of the intensity of long-memory in conditional heteroskedasticity. Statistical Inference for Stochastic Processes, 3, 113-128. MathSciNet | Full Text via CrossRef.

Giraitis, L., P.S. Kokoszka, R. Leipus and G. Teyssière (2003a). Rescaled variance and related tests for long-memory in volatility and levels. Journal of Econometrics, 112, 265-294.

Giraitis, L., P.S. Kokoszka, R. Leipus and G. Teyssière (2003b). On the power of R/S-type tests under contiguous and semi long-memory alternatives. Acta Applicandae Mathematicae, 78, 285-299. Full Text via CrossRef.

Giraitis, L., R. Leipus and D. Surgailis (2004). Recent advances in ARCH modelling. In: Teyssière, G., Kirman, A. (Eds.), Long-Memory in Economics. Berlin: Springer, in press.

Goodwin, B.K. and V.H. Smith (2005). Harvest-time protein shocks and price adjustments in the U.S. wheat markets. Agricultural Marketing Policy Paper No.5

Gopinath, R.A. and C.S. Burrus (1993). Wavelet Transforms and Filter Banks. In C.K. Chui, Ed., Wavelets: A Tutorial in Theory and Applications, 603-655, San Diego: Academic Press, Volume 2 of Wavelet Analysis and its Applications.

Gould, J.P. and C.R. Nelson (1974). The Stochastic structure of the velocity of money. American Economic Review, 64, 405-18.

Gourieroux, C. and J. Jasiak (2001). Memory and infrequent breaks. Economics Letters, 70, 29-41.

Grammatikos, T. and A. Saunders (1986). Futures price variability: A test of maturity and volume effects. Journal of Business, 59, 319-330.

Granger, C.W. (1980). Long memory relationships and the aggregation of dynamic models. Journal of Econometrics, 14, 227-38.

Granger, C.W.J. (1981). Some properties of time series data and their use in econometric specification. Journal of Econometrics, 16, 121-130.

Granger, C.W.J. (1983). Cointegrated variables and error correction models. Unpublished Manuscript, San Diego: University of California.

Granger, C.W.J. (1986). Developments in the study of cointegrated economic variables. Oxford Bulletin of Economics and Statistics, 48, 213-228.

Granger, C.W.J. (2000). Current perspectives on long memory processes. Chung-Hua Series of Lectures, No. 26, Taiwan: Institute of Economics, Academia Sinica.

Granger, C.W.J. (2002). Some comments on risk. Journal of Applied Econometrics, 17, 447-456.

Granger, C.W.J. (2005). The past and future of empirical finance: some personal comments. Journal of Econometrics, 129, 35-40.

Granger, C.W.J. and C.-Y. Sin (2000). Modeling the absolute returns of different stock indices exploring the forecastability of alternative measures of risk. Journal of Forecasting, 19, 277-298.

Granger, C.W.J. and N. Hyung (1999). Occasional structural breaks and long-memory. Preprint.

Granger, C.W.J. and N. Hyung (2004). Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. Journal of Empirical Finance, 11, 399-421.

Granger, C.W.J. and N. Swanson (1997). An introduction to stochastic unit root processes. Journal of Econometrics, 80, 35-61 (originally UCSD Economics Department Discussion Paper 1992).

Granger, C.W.J. and O. Morgenstern (1970). Predictability of stock market prices. Lexington, MA: Heath Lexington Books.

Granger, C.W.J. and R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis, 1, 15-29.

Granger, C.W.J. and T. Terasvirta (1999). A simple nonlinear time series model with misleading linear properties. Economics Letters, 62, 161-165.

Granger, C.W.J. and Z. Ding (1995). Some properties of absolute returns an alternative measure of risk. Annales d'Économie et de Statistique, 40, 67-91. MathSciNet.

Granger, C.W.J. and Z. Ding (1996). Varieties of long memory models. Journal of Econometrics, 73, 61-77.

Granger, C.W.J., (2002). Long memory, volatility, risk and distribution. UCSD Department of Economics working paper.

Gray, H.L., W. Woodward and N. Zhang (1989). On generalized fractional processes. Journal of Time Series Analysis, 10, 233-257.

Gray, H.L., W. Woodward and N. Zhang (1994). On generalized fractional processes: a correction. Journal of Time Series Analysis, 15, 561-562.

Gray, R. (1961). The search for a risk premium. Journal of Political Economy, 70:3, 250-60.

Greene, M.T. and B.O. Fielitz (1997). Long-term dependence in common stock returns. J. Financial Economics, 4, 339-349.

Greer, R. J. (2000). The nature of commodity index returns. Journal of Alternative Investments, 3:1, 45-52.

Grilli, E. R. and M.C. Yang (1988). Primary commodity prices, manufactured goods prices, and the terms of trade in developing countries: What the long run shows. World Bank Economic Review, 2, 1-47.

Grossman, S.J., and J.E. Stiglitz (1980). On the impossibility of informationally efficient markets. American Economic Review, 70, 393-408.

Haigh, M. S., J. H. Harris, J. A. Overdahl, M. A. Robe (2007). Market Growth, Trader Participation and Derivative Pricing. Commodity Futures Trading Commission working paper.

Haigh, M. S., J. Hranaiova and J. A. Overdahl (2007). Price volatility, Liquidity provision and the role of hedge funds in energy futures markets. Forthcoming: Journal of Alternative Investments.

Haken, H. (1977). Synergetics. Berlin: Springer.

Hall, J.A., B.W. Brorsen and S.H. Irwin (1989). The distribution of futures prices: A test of the stable paretian and mixture of normal hypotheses. Journal of Financial and Quantitative Analysis, 21:1, 105-16.

Hall, P., S. Penev, G. Kerkyacharian and D. Picard (1997). Numerical performance of block thresholded wavelet estimators. Stat. and Computing, 7, 115-124.

Hall, R.E. (1979). Stochastic implication of the life cycle-permanent income hypothesis: theory and evidence. Journal of Political Economy, 86, 971-987.

Hamilton, J.D. (1994). Time series analysis. Princeton, NJ: Princeton University Press.

Hamilton, J.D. and C.H. Whiteman (1985). The Observable implications of self-fulfilling expectations. Journal of Monetary Economics, 16, 353-373.

Hansen, B.E. (1992). Tests for parameter instability in regressions with I(1) processes. Journal of Business and Economic Statistics, 10, 321-335.

Hansen, B.E. (2000). Testing for structural change in conditional models. Journal of Econometrics, 97, 93-115.

Hansen, H. and S. Johansen (1999). Some tests for parameter constancy in cointegrated VAR-models. Econometrics Journal, 2, 306-333.

Hansen, P.R. (2003). Structural breaks in the cointegrated vector autoregressive model. Journal of Econometrics, 114, 261-295.

Haraf, W. S. (1986). Monetary velocity and monetary rules. Cato Journal, 6, 641-62.

Hardle, W., G. Kerkyacharian, D. Picard and A. Tsybakov (1998). Wavelets, approximation and statistical applications. Lecture Notes in Statistics, 129, Springer Verlag.

Hartman, D. G. (1974). The egg cycle and the ability of recursive models to explain it. American Journal of Agricultural Economics, 56:2, 254-262.

Harvey, D.I., S.J. Leybourne and P. Newbold (2001). Innovational outlier unit root tests with an endogenously determined break in level. Oxford Bulletin of Economics and Statistics, 63, 559-575.

Haslett, J. and A.E. Raftery (1989). Space-time modelling with long-memory dependence: assessing Ireland's wind power resource. Journal of the Royal Statistical Society, 38 (Series C), 1-50.

Hassler, U. (1993). Regression of spectral estimators with fractionally integrated time series. Journal of Time Series Analysis, 14, 369-380.

Hassler, U. and J. Wolters (1994). On the power of unit roots against fractional alternatives. Economics Letters, 45, 1-5.

Hassler, U. and J. Wolters (1995). Long memory in inflation rates: International evidence. Journal of Business and Economic Statistics, 13, 37-45.

Haubrich, J.G. and A.W. Lo (1992).The sources and nature of long-term dependence in the business cycle. Working paper, Dept. Finance, MIT.

Hauser, M.A. and E. Reschenhofer (1995). Estimation of the fractionally difference in g parameter with the R/S method. Computational Statistics and Data Analysis, 20, 569-579.

Heiner, R.A. (1983). The origin of predictable behavior. American Economic Review, 73, 560-95.

Helms, B.P., F.R. Kaen and R.E. Rosenman (1984). Memory in commodity futures contracts. J. Futures Markets, 4:4, 559-567.

Helson, J. and Y. Sarason (1967). Past and future. Mathematica Scandanavia, 21, 5-16.

Henry, M. (2001). Averaged periodogram spectral estimation with long memory conditional heteroscedasticity. Journal of Time Series Analysis, 22, 431-459.

Henry, M. (2001). Robust automatic bandwidth for long-memory. Journal of Time Series Analysis, 22, 293-316. MathSciNet

Henry, M. (2004). Bandwidth choice, optimal rates and adaptivity in semiparametric estimation of long memory. In: Teyssière, G., Kirman, A. (Eds.), Long-Memory in Economics, Berlin: Springer, in press.

Henry, M. and P.M. Robinson (1996). Bandwidth choice in Gaussian semiparametric estimation of long range dependence. In: P.M. Robinson and M. Rosenblatt, Editors, Athens Conference on Applied Probability and Time Series Analysis, Time Series

Analysis In Memory of E.J. Hannan. New York: Springer. Lecture Notes in Statistics, 115, 220-232.

Hicks, J. R. (1939). Value and capital. Cambridge: Oxford University Press.

Hidalgo, J. and P.M. Robinson (1996). Testing for structural change in a long-memory environment. Journal of Econometrics, 70, 159-174.

Higuchi, T. (1990). Approach to an irregular time series on the basis of fractal theory. Physica D, 31:277-283

Higuchi, T. (1990). Relationship between the fractal dimension and the power law index for a time series: A numerical investigation. Physica D, 46:254-264.

Hillebrand, E. (2004). Neglecting parameter changes in GARCH models. Journal of Econometrics, 10.1016/j.jeconom.2004.09.005.

Hodrick, R. J. and E.C. Prescott (1980). Post-war U.S. business cycles: An empirical investigation. Discussion paper 451, Carnegie-Mellon University.

Hodrick, R. J. and E.C. Prescott (1997). Postwar U.S. business cycles: An empirical investigation. Journal of Money, Credit and Banking, 29:1, 1-16.

Holt, M.T. and L.A. Craig (2006). Nonlinear dynamics and structural change in the U.S. Hog-corn cycle: A time-varying star approach. American Journal of Agricultural Economics, 88:1, 215-233.

Hommes, C.H. (2001). Financial markets as nonlinear adaptive evolutionary systems. Quantitative Finance, 1, 149-167

Horváth, L. and P.S. Kokoszka (2003). A bootstrap approximation to a unit root test statistic for heavy tailed observations. Statistics and Probability Letters, 62, 163-173.

Horváth, L., Kokoszka, P.S., Teyssière, G. (2004). Bootstrap misspecification tests for ARCH based on the empirical process of squared residuals. Journal of Statistical Computation and Simulation, 74, 469-485.

Horváth, L., P.S. Kokoszka and G. Teyssière (2001). Empirical process of the squared residuals of an ARCH sequence. The Annals of Statistics 29, 445-469.

Hosking, J.R.M. (1981). Fractional differencing. Biometrika, 68, 165-176.

Hsieh, D.A. (1989). Testing for nonlinear dependence in daily foreign exchange rate changes. Journal of Business, 62, 339-68.

Hull, J.C. (2002). Options, futures and other derivatives, 5th edition. Prentice Hall.

Hurst, H.E. (1951). Long-term storage capacity of reservoirs. Transactions of the American Society of Civil Engineers, 116,770-799.

Hurst, H.E. (1957). A suggested statistical model of some time series that occur in nature. Nature, 180, 494.

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.

Hyung, N. and P.H. Franses (2001). Structural breaks and long memory in US inflation rates: do they matters for forecasting? Erasmus University Rotterdam, Econometric Institute Research Report EI2001-13.

Hyung, N. and P.H. Franses (2002). Inflation rates long-memory, level, shifts, or both? Econometric Institute Report 2002-08, Erasmus University.

Igloi, E. and G. Terdik (1999). Bilinear Stochastic Systems with Fractional Brownian Motion Input. Annals of Applied Probability, 9:1, 46-77.

Inclán, C. and G.C. Tiao (1994). Use of cumulative sums of squares for retrospective detection of changes in variance. Journal of the American Statistical Association, 89, 913-923.

Irwin, S.H., C.R. Zulauf and T.E. Jackson (1996). Monte Carlo analysis of mean reversion in commodity futures prices. American Journal of Agricultural Economics, 78:2, 387-99

Istas, J. and G. Lang (1994). Quadratic variations and estimation of the local Hölder index of a Gaussian process. Ann. Inst. Poincaré, 33, 407-436.

J.R. Blum, J. Kiefer and M. Rosenblatt (1961). Distribution free tests of independence based on the sample distribution function. The Annals of Mathematical Statistics, 32, 485-498. MathSciNet

Jach, A. and P.S. Kokoszka (2004). Subsampling unit root tests for heavy-tailed observations. Methodology and Computing in Applied Probability, 6, 73-97. Full Text via CrossRef.

Jagannathan, R. (1985). An investigation of commodity futures prices using the consumption-based intertemporal capital asset pricing model. Journal of Finance, 40, 175-91.

Jarrow, R.A. and E.R. Rosenfeld (1984). Jump risks and the intertemporal capital asset pricing model. Journal of Business, 57, 337-51.

Jegadeesh, N. and S. Titman (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance, 48, 65-92.

Jeganathan, P. (1999). On asymptotic inference in cointegration time series with fractionally integrated errors. Econometric Theory, 15, 583-621.

Jensen G. R., J. M. Mercer and R. R. Johnson (2002). Tactical asset allocation and commodity futures. Journal of Portfolio Management, Summer, 100-111.

Jensen, M. J. and B. Whitcher (2000). Time-varying long-memory in volatility: Detection and estimation with wavelets. University of South Carolina working paper.

Jensen, M.J. (1999). An approximate wavelet MLE of short and long memory parameters. Studies in Nonlinear Dynamics and Econometrics, 3, 239-253.

Jensen, M.J. (1999). Using wavelets to obtain a consistent ordinary least squares estimator of the long memory parameter. Journal of Forecasting, 18, 17-32.

Jensen, M.J. (2000). An alternative maximum likelihood estimator of long-memory processes using compactly supported wavelets. Journal of Economic Dynamics and Control, 24, 361-387.

Jensen, M.J. and B. Whitcher (2000). Wavelet estimation of a local long-memory parameter. Exploration Geophysics, 31, 89-98.

Jin, H.J. and D. L. Frechette (2004). Fractional integration in agricultural futures price volatilities. American Journal of Agricultural Economics, 86:2, 432-433.

Johansen, S., R. Mosconi and B. Nielsen (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. The Econometrics Journal, 3, 216-249.

Jonker, J. and O. Gelderblom (2005). The origins of derivatives markets in: William N. Goetzmann and K. Geert Rouwenhorst, The Origins of Value, Oxford University Press, in press.

Jorion, P. (1988). On jump processes in the foreign exchange and stock markets. Review of Financial Studies, 1, 427-45.

Just, R.E. (1976). Spectral analysis of stochastic properties in regression: An application in supply response identification. American Journal of Agricultural Economics, 58:4, 712-720.

Kaen, F.R. and R.E. Roseman (1986). Predictable behavior in financial markets: Some evidence in support of Heiner’s hypothesis. American Economic Review, 76, 212-20.

Kahane, J.P. and P.G Lemarié (1995). Fourier series and wavelets, Gordon and Research Publishers. Studies in the Development of Modern Mathematics, 3.

Kaiser, G. (1994). A friendly guide to wavelets. Birkhauser.

Kao, G.W. and C.K. Ma (1992). Memories, heteroscedasticity, and price limit in currency futures markets. Journal of Futures Markets, 12, 679-92.

Kazakevicius, V. and R. Leipus (2002). On stationarity in the ARCH(?) model. Econometric Theory, 18, 1-16. Full Text via CrossRef.

Kazakevicius, V. and R. Leipus (2003). A new theorem on existence of invariant distributions with applications to ARCH processes. Journal of Applied Probability, 40, 147-162.

Kenyon, D. and Jones, E. (1993). Forecasting performance of corn and soybean harvest futures contracts. American Journal of Agricultural Economics, 0002-9092, 75:2.

Keynes, J.M. (1930). A treatise on money, Vol. 2. London: Macmillan.

Khalaf, L., J.D. Saphores and J.F. Bilodeau (2003). Simulation-based exact jump tests in models with conditional heteroskedasticity. Journal of Economic Dynamic and Control, 28, 531-53.

Kim, T-H., S.J. Leybourne and P. Newbold (2000). Spurious rejections by Perron tests in the presence of a break. Oxford Bulletin of Economics and Statistics, 62, 433-444.

Kinav, O. (1999). Fractionally integrated generalized autoregressive conditionally heteroskedastic (FIGARCH) processes and currency risk hedging. Ph.D. thesis, North Carolina State University.

King, R., C. Plosser, J. Stock and M. Watson (1991). Stochastic trends and economic fluctuations. American Economic Review, 81, 819-840.

Kirman, A. (1991). Epidemics of opinion and speculative bubbles in financial markets. In: M. Taylor, Ed., Money and Financial Markets, London: Macmillan, 354-368.

Kirman, A. (1993). Ants, rationality and recruitment. The Quarterly Journal of Economics, 108, 137-156. Full Text via CrossRef.

Kirman, A. and G. Teyssière (2002a). Microeconomic models for long-memory in the volatility of financial time series. Studies in Nonlinear Dynamics and Econometrics, 5, 281-302.

Kirman, A. and G. Teyssière (2002b). Bubbles and long-range dependence in asset prices volatilities. In: C.H. Hommes, R. Ramer and C. Withagen, Editors, Equilibrium, Markets and Dynamics. Essays in Honour of Claus Weddepohl, Berlin: Springer, 307-327.

Koekebakker, S. and G. Lien (2004). Volatility and price jumps in agricultural futures prices: Evidence from wheat options. American Journal of Agricultural Economics, 86:4, 1018-1031.

Kokoszka, P.S. and A. Parfionovas (2004). Bootstrap unit root tests for heavy tailed time series. In: S.T. Rachev, Ed., Handbook of Computational and Numerical Methods in Finance, Berlin: Springer, 175-197.

Kokoszka, P.S. and G. Teyssière (2002). Change-point detection in GARCH models: asymptotic and bootstrap tests. Journal of Business and Economic Statistics, Under revision.

Kokoszka, P.S. and R. Leipus (1999). Testing for parameter changes in ARCH models. Lithuanian Mathematical Journal, 39, 231-247.

Kokoszka, P.S. and R. Leipus (2000). Change-point estimation in ARCH models. Bernoulli, 6, 513-539. Full Text via CrossRef.

Kokoszka, P.S. and R. Leipus (2002). Detection and estimation of changes in regime. In: P. Doukhan, G. Oppenheim and M.S. Taqqu, Eds., Long-Range Dependence: Theory and Applications, Boston: Birkhauser, 325-337.

Kolb, R.W. (1992). Is normal backwardation normal? Journal of Futures Markets, 12, 75-91.

Kolb, R.W. (1996). The systematic risk of futures contracts. Journal of Futures Markets, 16, 631-654.

Kouri, P.J.K. (1983). Balance of payments and the foreign exchange market a dynamic partial equilibrium model. In: J.S. Bhandari and B.H. Putnam, Eds, Economic Interdependence and Flexible Exchange Rates, Cambridge, Massachussets: MIT Press.

Kramer, W. and P. Sibbertsen (2002). Testing for structural change in the presence of long memory. International Journal of Business and Economics, 1, 235-242.

Kuan, C.M. and C.C. Hsu (1998). Change point estimation of fractionally Integrated Processes. Journal of Time Series Analysis, 19, 693-708.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. How sure are we that economic series have a unit root? Journal of Econometrics, 54, 159-178

Labys W., J.B. Cromwell and E. Kouassi (2000). What color are commodity prices? A fractal analysis. Empirical Economics, 25:4.

Labys, W.C., A. Achouch and M. Terraza (1999). Metal prices and the business cycle. Resources Policy, 25, 229-38.

Labys, W.C., J.B. Lesourd and D. Badillo (1998). The existence of metal price cycles. Resources Policy, 24, 147-55.

Lampart,C. and J.B. Ramsey (1998). The decomposition of economic relationships by the scale using wavelets: Money and income. Macroeconomic Dynamics, 2, 49-71.

Larson, A.B. (1964). The hog cycle as harmonic motion. Journal of Farm Economics, 46, 375-86.

Lavielle, M. (1998). Optimal segmentation of random processes. IEEE Transactions on Signal Processing, 46, 1365-1373. Full Text via CrossRef.

Lavielle, M. (1999) Detection of multiple changes in a sequence of dependent variables. Stochastic Processes and their Applications, 83, 79-102.

Lavielle, M. and E. Moulines (2000). Least squares estimation of an unknown number of shifts in a time series. Journal of Time Series Analysis, 21, 33-59.

Lazarova, S. (2004). Testing for structural change in regression with long memory processes. Journal of Econometrics. 10.1016/j.jeconom.2004.09.011.

Lazarová, S. (2005). Testing for structural change in regression with long memory processes. Journal of Econometrics, 129, 329-372.

Le Roy, S.F. and R.D. Porter (1981). The present value relation tests based on implied variance bounds. Econometrica, 49, 555-574. Full Text via CrossRef.

LeBaron, B. (1992). Some relations between volatility and serial correlation in stock market returns. Journal of Business, 65, 199-220.

LeBaron, B. (2001) Empirical regularities from interacting long- and short-memory investors in an agent-based stock market. IEEE Transactions on Evolutionary Computation, 5, 442-455. Full Text via CrossRef.

Lee, D. and P. Schmidt (1996). On the power of the KPSS Test of stationarity against fractionally integrated alternatives. Journal of Econometrics, 73, 285-302.

Lee, J. and M. Strazicich (2001). Break point estimation and spurious rejections with endogenous unit root tests. Oxford Bulletin of Economics and Statistics, 62, 535-558.

Leipus, R. and D. Surgailis (2003). Random coefficient autoregression, regime switching and long memory. Advances in Applied Probability, 35, 737-754.

Leipus, R. and M.C. Viano (2003). Long memory and stochastic trend. Statistics and Probability Letters, 61, 177-190.

Leipus, R., V. Paulauskas and D. Surgailis (2004). Renewal regime switching and stable limit laws. Journal of Econometrics, 10.1016/j.jeconom.2004.09.010.

Leland, W.E., M.S. Taqqu, W. Willinger and D.V. Willson (1994). On the self-similar nature of Ethernet traffic (Extended Version). IEEE/ACM Transactions on Networking, 2:1, 1-15.

Lemarié, P.G., Ed. (1990). Les ondelettes en 1989. Lecture Notes in Mathematics, Springer Verlag.

Leuthold, R. M. (1972). Random walk and price trends: the live cattle futures market. Journal of Finance, 27, 879-889.

Levy, H. (1992). Stochastic dominance expected utility: survey and analysis. Management Science, 38, 555-593.

Leybourne, S., M. McCabe and J. Tremayne (1996b). Can economic time series be differenced to stationarity? Journal of Business and Economic Statistics, 14, 435-446.

Leybourne, S., M. McCabe and M. Mills (1996a). Randomized unit root processes for modeling and forecasting financial time series: theory and applications. Journal of Forecasting, 15, 153-270.

Leybourne, S.J., T.C. Mills and P. Newbold. (1998). Spurious rejections by Dickey-Fuller tests in the presence of a break under the null. Journal of Econometrics, 87, 191-203.

Lin, C. and M.C. Roberts (2006). A term structure model for commodity prices: Does storability matter? Proceedings of the 2006 NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting and Market Risk Management.

Lin, C., and M.C. Roberts (2006). A term structure model for commodity prices: Does storability matter? Proceedings of the 2006 NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting and Market Risk Management.

Lin, C.-F. and T. Terasvirta (1994). Testing the constancy of regression parameters against continuous structural change. Journal of Econometrics, 62, 211-228.

Lin, C.-F. and T. Terasvirta (1999). Testing parameter constancy in linear models against stochastic stationary parameters. Journal of Econometrics, 90, 193-213.

Ling, F. and C.L. Nikias (1992). Advanced digital signal processing. New York: Macmillan Publishing Co.

Liu, M. (2000). Modelling long memoryin stock market volatility. Journal of Econometrics, 99, 139-171.

Ljung, G.M. and G.E.P. Box (1978). On a measure of lack of fit in time series models. Biometrika, 65, 297-303.

Lo, A.W. (1991). Long-term memory in stock market prices. Econometrica, 59, 1279-1313.

Lo, A.W. and A.C. Mackinnlay (1999). A non-random walk down Wall Street. Princeton, N.J.: Princeton Press.

Lo, A.W. and J. Wang (2001). Stock market trading volume. In: Yacine Ait-Sahalia and Lars Peter Hansen, Ed., Handbook of financial econometrics, forthcoming.

Lobato, I.N. (1999). A semiparametric two-step estimator in a multivariate long memory model. Journal of Econometrics, 90, 29-153.

Lobato, I.N. and P. Robinson (1998). A nonparametric test for I(0). Review of Economic Studies, 68, 475-495.

Lobato, I.N. and P.M. Robinson (1996). Averaged periodogram estimation of long memory. Journal of Econometrics, 73, 303-324.

Lordkipandize, N. (2004). Modeling and estimation of long memory in stochastic volatility: Application to options on futures contracts. PhD Dissertation, Cornell University, January 2004.

Lyons, R. (2001). The microstructure approach to exchange rates, Cambridge MA: MIT Press.

Ma, C.K., J.M. Mercer and M.A. Walker (1992). Rolling over futures contracts: A note. Journal of Futures Markets, 12, 203-217.

Mackey, M.C. (1989). Commodity price fluctuations: Price dependent delays and nonlinearities as explanatory factors. Journal of Economic Theory, 48: 2, 497-509.

MacKinnon, J.G. (1991). Critical values for cointegration tests. In: C.W.J. Granger and G. Mizon, Eds., Long Run Economic Relations: Readings in Cointegration, Oxford: Oxford University Press, 267-276.

MacKinnon, J.G. (1996). Numerical distribution functions for unit-root and cointegration tests. Journal of Applied Econometrics, 11, 601-618. Full Text via CrossRef.

Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. IEEE Pattern Anal. and Machine Intell., 11: 7, 674-693.

Mallat, S. (1998). A wavelet tour of signal processing. Academic Press.

Mallat, S.G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11, 674-693.

Mandelbrot, B. (1963a). The variation of certain speculative prices. Journal of Business, 36, 394-419.

Mandelbrot, B.B. (1977). Fractals: Form, chance and dimension. San Francisco: W.H. Freeman and Co.

Mandelbrot, B.B. (1963b). New methods in statistical economics. Journal of Political Economy, 71, 421-40.

Mandelbrot, B.B. (1972). Statistical methodology for non-periodic cycles: From covariance to R/S analysis. Annals of Economic and Social Measurement, 1, 259-290.

Mandelbrot, B.B. (1975). A fast fractional Gaussian noise generator. Water Resources Research, 7, 543-553.

Mandelbrot, B.B. (1982). The fractal geometry of nature. New York: W. H. Freeman.

Mandelbrot, B.B. and J. Wallis (1968). Noah Joseph and operational hydrology. Water Resources Research, 4, 909-918.

Mandelbrot, B.B. and J.R. Wallis (1969). Robustness of the rescaled range R/S in the measurement of non-cyclic long-run statistical dependence. Water Resources Research, 5, 321-340.

Mandelbrot, B.B. and J.W. Van Ness (1968). Fractional Brownian motions, fractional noises and applications. SIAM Review, 10, 422-437.

Mandelbrot, B.B. and M.S. Taqqu (1979). Robust R/S analysis of long run serial correlation. Proceedings of 42nd Session ISI, 69-99

Mankiw, N.G., D. Romer and M. Shapiro (1985). An unbiased reexamination of stock market volatility. Journal of Finance, 40, 677-689. Full Text via CrossRef.

Marinucci, D. (2000). Spectral regression for cointegrated time series with long-memory innovations. Journal of Time Series Analysis, 21, 685-706.

Marsaglia, G. (1996). DIEHARD: A battery of tests of randomness. Click to view the MathML source.

Marsh, T.A. and R.C. Merten (1986). Dividend variability and variance bounds tests for the rationality of stock market prices. American Economic Review, 76, 483-498.

McCallum, B. (1983). On non-uniqueness in rational equilibrium models an attempt at perspective. Journal of Monetary Economics, 11, 139-168.

McCoy, E. J., and A. T. Walden (1996). Wavelet analysis and synthesis of stationary long-memory processes. Journal of Computational and Graphical Statistics, 5, 1-31.

McCurdy, T.H. and I. Morgan (1988). Testing the Martingale hypothesis in Deutsche mark futures with models specifying the form of heteroskedasticity. Journal of Applied Econometrics, 3, 187-202.

McDonald, R.L. (2002). Derivatives markets. Addison Wesley.

McLeod, A.I. and K.W. Hipel (1978). Preservation of the rescaled adjusted range, 1. A reassessment of the Hurst phenomenon. Water Resources Research, 14, 159-178.

Meadows, D. H., D. L. Meadows, J. Randers and W. W. Behrens III (1972). The Limits to Growth. New York, NY: University Books.

Meese, R.A. (1986). Testing for bubbles in exchange markets the case of sparkling rates. Journal of Political Economy, 94, 345-373. Full Text via CrossRef.

Meese, R.A. and K. Rogoff (1983). Empirical exchange rate models of the seventies do they fit out-of-sample? Journal of International Economics, 14, 3-24

Merton, R.C. (1973). The theory of rational options pricing. Bell Journal of Economics, 4, 141-183.

Merton, R.C. (1976). Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics, 3:1, 125-44.

Meyer, Y. (1990). Ondelettes et opérateurs, Tome 1, Hermann Ed. (English translation: Wavelets and operators, Cambridge Univ. Press, 1993).

Meyer, Y. (1992). Wavelets and operators. Cambridge University Press.

Meyer, Y. (1993). Les ondelettes. Algorithmes et applications, 2nd edition. Paris: Colin Ed. (English translation: Wavelets: algorithms and applications, SIAM).

Meyer, Y. (1993). Wavelets: algorithms and applications. SIAM.

Meyer, Y. and S. Roques, Eds. (1993). Progress in wavelet analysis and applications, Frontières Ed.

Mikosch, T. and C. Starica (2004). Nonstationarities in financial time series the long, range dependence and the IGARCH effects. Review of Economics and Statistics, 86, 378-390.

Miller, M.H. (1991). Financial markets and market volatility, Cambridge, MA: Basil Blackwell.

Milonas, N.T. (1986). Price variability and the maturity effect in futures markets. Journal of Futures Markets, 3, 433-460.

Miltersen, K.R. (2003). Commodity price modeling that matches current observables: a new approach. Quantitative Finance, 3:1, 51-58.

Misiti, M., Y. Misiti, G. Oppenheim and J.M. Poggi (1993a). Analyse de signaux classiques par décomposition en ondelettes. Revue de Statistique Appliquée, XLI:4, 5-32.

Misiti, M., Y. Misiti, G. Oppenheim and J.M. Poggi (1993b). Ondelettes en statistique et traitement du signal. Revue de Statistique Appliquée, XLI: 4, 33-43.

Misiti, M., Y. Misiti, G. Oppenheim and J.M. Poggi (1994). Décomposition en ondelettes et méthodes comparatives: étude d'une courbe de charge électrique. Revue de Statistique Appliquée, XLII:2, 57-77.

Misiti, M., Y. Misiti, G. Oppenheim and J.-M. Poggi (2003). Les ondelettes et leurs applications, Hermes.

Misiti, M., Y. Misiti, G. Oppenheim and J-M. Poggi (1997). Wavelet toolbox for use with MATLAB, Natick, Massachusetts: The MathWorks Inc.

Montañés, A., I. Olloqui and E. Calvo (2004). Selection of the break in the Perron-type tests. Journal of Econometrics, 129, 41-64

Moore, M.J. and U. Cullen (1995). Speculative e±ciency on the London metal exchange. The Manchester School, 63, 235-56.

Muller, U. (2004). A theory of robust long-run variance estimation, mimeo.

Muller, U. (2004). The impossibility of consistent discrimination between I(0) and I(1) processes, mimeo.

Muller, U. (2005). Efficient tests for general persistent time variation in regression coefficients, mimeo. (Joint with Graham Elliott.) Forthcoming in Review of Economic Studies.

Muller, U. (2005). Size and power of tests for stationarity in highly autocorrelated time series. Journal of Econometrics, 128, 195-213.

Mundlak, Y. and H. Huang (1996). International comparisons of cattle cycles. American Journal of Agricultural Economics, 78: 4, 855-869.

Murphy, J.A. (1987). The Seasonality of risk and return on agricultural futures positions. American Journal of Agricultural Economics, 69: 3, 640-648, 8p; (AN 4595097).

Myers, R.J. (1994). Time series econometrics and commodity price analysis: a review. Review of Marketing and Agricultural Economics 62(2): 167-181.

Nash, D.J. (2001). Long-term investing in commodities, global pensions quarterly, Morgan Stanley Global Pensions Group, 4:1, 25-31.

Nason, G.P. and B.W. Silverman (1995). The Stationary wavelet transform and some statistical applications. Lecture Notes in Statistics, 103, 281-299.

Nason, G.P. and R. von Sachs. (1999). Wavelets in time series analysis. Philosophical Transactions of the Royal Society of London, Series A, 357, 2511-2526.

Nason, G.P. and T. Sapatinas (2002). Wavelet packet transfer function modelling of nonstationary time series. Statistics and Computing, 12, 45-56.

Nelson, C.R. and C.I. Plosser (1982). Trends and random walks in macroeconomic time series. Journal of Monetary Economics, 10, 139-162.

Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: a new approach. Econometrica, 59, 347-70.

Newey, W. and K.D. West (1987). A Simple positive definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55, 703-708.

Newland, D.E. (1993). An introduction to random vibrations, spectral analysis and wavelets, 3rd edition, Harlow, Essex: Longman Scientific and Technical.

Ng, S. and P. Perron (2001). Lag length selection and the construction of unit root tests with good size and power. Econometrica, 69:6, 1519-54.

Nielsen, M.O. (2004). Optimal residual based tests for fractional cointegration and exchange rate dynamics. Journal of Business and Economic Statistics, 22, 331-345.

Nyblom, J. (1989). Testing for the constancy of parameters over time. Journal of the American Statistical Association, 84, 223-230.

Ogden, R.T. (1997). Essential wavelets for statistical applications and data analysis. Birkhauser.

Ohanissian, A., J.R. Russell and R. Tsay (2004). True or spurious long memory? A new test. University of Chicago Graduate School of Business working paper.

Ohanissian, A., Russell, J.R. and R.S. Tsay (2003). Using temporal disaggregation to distinguish between true and spurious long memory. Mimeo, University of Chicago.

Ojanen, H. (1991). Remarks on the Sobolev regularity of wavelets and interpolation schemes. Research Reports A305, Helsinki Univ. of Technology, Inst. of Math.

Ojanen, H. (1998). Orthonormal compactly supported wavelets with optimal Sobolev regularity. Rutgers Univ. Dept. Math., preprint.

P. F. Craigmile, N. Cressie, T. J. Santner and Y. Rao (2006). A Loss function approach to identifying environmental exceedances. Extremes, 8:3, 143-159.

Pai, J. and S.N. Ravishanker (1998). Bayesian analysis of autoregressive fractionally integrated moving average processes. Journal of Time Series Analysis, 19, 99-112.

Pan, J. (2002). The Jump-risk premia implicit in options: Evidence from an integrated time-series study. Journal of Financial Economics, 63, 3-50.

Panas, E. (2001). Estimating fractal dimension using stable distributions and exploring long memory through ARFIMA models in Athens stock exchange. Applied Financial Economics, 11, 395-402.

Paparoditis, E. and D. Politis (2003). Residual-based block bootstrap for unit root testing, Econometrica, 71, 813-855. Full Text via CrossRef.

Park, C. and S. H. Irwin (2004). The profitability of technical analysis: a review. University of Illinois AgMAS Project Research Report 2004-04.

Parke, W.R. (1999). What is fractional integration? The Review of Economics and Statistics, 81, 632-638.

Parkinson, M. (1980). The Extreme value method for estimating the variance of the rate of return. Journal of Business, 53: 1, 61-65.

Percival, D. and A.G. Bruce (1998). Wavelet-based approximate maximum likelihood estimation for trend-contaminated fractional difference. University of Washington mimeo.

Percival, D.B. and A.T. Walden (2000). Wavelet methods for time series snalysis. Cambridge: Cambridge University Press.

Percival, D.B. and H.O. Mofjeld (1997). Analysis of subtidal coastal sea level fluctuations using wavelets. Journal of the American Statistical Association, 92:439, 868-80.

Perron, P. (1989). The Great crash, the oil price shock, and the unit root hypothesis. Econometrica, 57, 1361-1401.

Perron, P. (1993). Erratum: The Great crash, the oil price shock, and the unit root hypothesis. Econometrica, 61, 248-49.

Perron, P. (1997). Further evidence from breaking trend functions in macroeconomic variables. Journal of Econometrics, 80, 355-385.

Perron, P. (2006). Dealing with structural breaks, in: Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory, K. Patterson and T.C. Mills (Eds.), Palgrave Macmillan, 278-352.

Perron, P. and T.J. Vogelsang (1992). Testing for a unit root in a time series with a shift in mean, corrections and extensions. Journal of Business and Economic Statistics, 10, 467-470.

Perron, P. and X. Zhu (2005). Structural breaks with deterministic and stochastic trends. Journal of Econometrics, 129, 65-119.

Pesaran, H. and A. Timmerman (2004). Small sample properties of forecasts from autoregressive models under structural breaks. Journal of Econometrics, 10.1016/j.jeconom.2004.09.007.

Pesaran, M.H. and A. Timmermann (2005). Small sample properties of forecasts from autoregressive models under structural breaks. Journal of Econometrics, 129, 183-217.

Pesquet, J.C., H. Krim and H. Carfatan (1996). Time-invariant orthonormal wavelet representations. IEEE Trans. Sign. Proc., 44:8, 1964-1970.

Peters, A. and P. Sibbertsen (2001). Robust tests on fractional cointegration. SFB 475 Technical Report 29/ 2001, University of Dortmund.

Peters, E. (1996). Chaos and order in the capital markets, 2nd Edition. New York: John Wiley and Sons.

Peterson, H.H. and W.G. Tomek (2005). How much of commodity price behavior can a rational expectations storage model explain? Agricultural Economics 33: 289-303.

Phillips, P.C.B. and K. Shimotsu (2004b). Exact local Whittle estimation of fractional integration. Cowles Foundation Discussion Paper N. 1367 (Revised).

Phillips, P.C.B. and K. Shimotsu, K. (2004a). Local Whittle estimation in nonstationary and unit root cases. Annals of Statistics, 32, 656-692.

Phillips, P.C.B. and P. Perron (1988). Testing for a unit root in time series regression. Biometrika, 75, 335-346..

Phillips, P.C.B. and S.J. Durlauf (1986). Multiple time series regression with integrated processes. Review of Economic Studies, 53, 473-495.

Pindyck, R.S. (1993).The Present value model of rational commodity pricing. The Economic Journal.

Pindyck, R.S. (1994). Inventories and the short-run dynamics of commodity prices. RAND Journal of Economics.

Pindyck, R.S. (1999). The Long run evolution of energy prices. Energy Journal, 20, 1-27.

Pindyck, R.S. (2004). Volatility and commodity price dynamics, J Futures Markets.

Pindyck, R.S. and J. Rotemberg (1990). The Excess co-movement of commodity prices

Pollock, D.S.G. (1999). Time-series analysis, signal processing and dynamics. London: Academic Press.

Pollock, D.S.G. (2002). Circulant matrices and time-series analysis. The International Journal of Mathematical Education in Science and Technology, 33, 213-230.

Porter-Hudak, S. (1990). An Application of the seasonal fractionally differenced model to the monetary aggregates. Journal of the American Statistical Association, 85, 338-344.

Prebisch, R. (1950). The Economic development of Latin America and its principal problems. New York: United Nations.

Quandt, R.E. (1960). Tests of the hypothesis that a linear regression obeys two separate regimes. Journal of American Statistical Association, 55, 320-330.

Rachev, S.T. and S. Mittnik (2000). Stable Paretian models in finance, New York: Wiley.

Ramsey, J. and Z. Zhang (1997). The Analysis of foreign exchange rates using waveform dictionaries. Journal of Empirical Finance, 4, 341-372.

Ramsey, J.B. (1999). The Contribution of wavelets to the analysis of economics and financial data. Philosophical Transactions of the Royal Society of London, Series A, 357, 2593-2606.

Rangarajan, G. and M. Ding (2000). Integrated approach to the assessment of long run correlation in times series data. Physical Review A, 61:5, 4991-5001.

Rappoport, P. and L. Reichlin (1989). Segmented trends and nonstationary time series. The Economic Journal, 99 (Conference Volume), 168-177.

Rausser, G.C. and T.F. Cargill (1970). The Existence of broiler cycles: An application of spectral analysis. American Journal of Agricultural Economics, 52:1, 109-122

Readinger, C.R., E. Mosley-Thompson, P. Craigmile, L.G. Thompson, and C.A. Calder (2005). Regional sensitivity of Greenland precipitation to NAO variability. Geophysical Research Letters, 32, L24707, doi:10.1029/2005GL024776.

Resnick, S.I. (1987). Extreme values, regular variation and point processes. Berlin: Springer.

Richter, M. and C. Sorensen (2002). Stochastic volatility and seasonality in commodity futures and options: the case of soybeans. Working paper. University of Copenhagen Business School.

Roberts, M. C. (2002). Homogenizing heterogeneous data: inducing ARCH effects in futures data. Selected Paper presented at the annual meeting of the American Agricultural Economics Association. Tampa, Florida.

Roberts, M. C. and P. L. Fackler (1999). State-space modeling of commodity price term structure. Selected Paper presented at the annual meeting of the American

Robinson, P. (1995a). Log-periodogram regression of time series with long range dependence. Annals of Statistics, 23, 1048-1072.

Robinson, P.M. (1978). Statistical inference for a random coefficient autoregressive model. Scandinavian Journal of Statistics, 5, 163-168.

Robinson, P.M. (1991). Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. Journal of Econometrics, 47, 67-84.

Robinson, P.M. (1994). Time series with strong dependence. In: C.A. Sims, Ed, Advances in Econometrics, Sixth World Congress, Cambridge: Cambridge University Press, 47-95.

Robinson, P.M. (1994b). Semiparametric analysis of long-memory time series. The Annals of Statistics, 22, 515-539.

Robinson, P.M. (1994c). Efficient tests of nonstationary hypotheses. Journal of the American Statistical Association, 89, 1420-1437.

Robinson, P.M. (1995b). Gaussian semiparametric estimation of long range dependence. The Annals of Statistics, 23, 1630-1661.

Robinson, P.M. (2001). The memory of stochastic volatility models. Journal of Econometrics, 101, 195-218.

Robinson, P.M. (2003). Long memory time series. In: Robinson, P.M. (Ed.), Time Series with Long Memory. Oxford: Oxford University Press.

Robinson, P.M. and D. Marinucci (2000). The averaged periodogram for nonstationary vector time series. Statistical Inference for Stochastic Processes, 3, 149-160.

Robinson, P.M. and D. Marinucci (2001). Narrow-band analysis of nonstationary processes. The Annals of Statistics, 29, 947-986.

Robinson, P.M. and F. Iacone (2004). Cointegration in fractional systems with deterministic trends. Journal of Econometrics, 10.1016/j.jeconom.2004.09.009.

Robinson, P.M. and J. Hualde (2003). Cointegration in fractional systems with unknown integration orders. Econometrica, 71, 1727-1766.

Robinson, P.M. and P. Zaffaroni (1997). Modelling nonlinearityand long memory in time series. Fields Institute Communications, 11, 161-170.

Robinson, P.M. and P. Zaffaroni (1998). Nonlinear time series with long memory: a model for stochastic volatility. Journal of Statistical Planning and Inference, 68, 359-371.

Robinson, P.M. and Y. Yajima (2002). Determination of cointegrating rank in fractional systems. Journal of Econometrics, 106, 217-241.

Rogers, L.C.G. (1997) Arbitrage with fractional Brownian motion. Mathematical Finance, 7, 95-105.

Rogers, L.C.G. and S.E. Satchell (1991). Estimating variance from high, low and closing prices. Annals of Applied Probability, 1:4, 504-512.

Roll, R. (1983). On computing mean returns and the small firm effect. Journal of Financial Economics, 12, 371-86.

Rosen, S. (1987). Dynamic animal economics. American Journal of Agricultural Economics, 69, 547-57.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. Proceedings of the National Academy of Sciences, 42, 43-47.

Rubinstein, M. (1979). The pricing of uncertain income streams and the pricing of options. Bell Journal of Economics, 7, 407-424.

Rutledge, D.J.S. (1976). A note on the variability of futures prices. Review of Economics and Statistics, 58, 120-123.

Samuelson, P.A. (1965). Proof that properly anticipated prices fluctuate randomly. Industrial Management Review, 6, 41-49.

Samuelson, P.A. (1973). Proof that properly discounted present values of assets vibrate randomly. The Bell Journal of Economics and Management Science, 4, 369-374.

Sanders, D.R. and M. R. Manfredo (2005). Forecast Encompassing as the Necessary Condition to Reject Futures Market Efficiency: Fluid Milk Futures. American Journal of Agricultural Economics, 87:3, 610-620.

Sanders, D.R., S.H. Irwin and R.M. Leuthold (1996). Noise trader demand in futures markets. OFOR paper number 96-02, University of Illinois.

Sanders, D.R., S.H. Irwin and R.M. Leuthold (1997). Noise traders, market sentiment, and futures price behavior, working paper 9702 University of Illinois

Saphores J.-D., L. Khalaf and D. Pelletier (2002). On Jumps and ARCH Effects in Natural Resource Prices: An application To Pacific Northwest stumpage prices. American Journal of Agricultural Economics, 84:2, 387-400.

Schleicher, C. (2002). An Introduction to wavelets for economists, Working Papers 02-03, Bank of Canada.

Schmidt, C.M. and R. Tschernig (1993). Identification of fractional ARMA models in the presence of long memory, Munchener Wirschaftswissenschaftliche Beitrage 93-04.

Schroeder, M.N. (1991). Fractals, chaos and power laws. New York: W.H. Freeman and Company.

Schwartz, E.S. (1997). The Stochastic behavior of commodity prices: implications for valuation and hedging. Journal of Finance, 52, 923-73.

Schwartz, E.S. and J.E. Smith (2000). Short-term variations and long-term dynamics in commodity prices. Management Science, 46: 893-911.

Schwert, W. (1981). The Adjustment of stock prices to information about inflation. Journal of Finance, 36, 15-29.

Sephton, P. (2002). Fractional cointegration: Monte Carlo estimates of critical values, with an application. Applied Financial Economics, 12, 331-335.

Serletis, A. (1995). Random walks, breaking trend functions, and the chaotic structure of the velocity of money. Journal of Business and Economic Statistics, 13:4, 453-58.

Shannon, C.E. and W. Weaver (1964). The Mathematical theory of communication, Urbana, Illinois: The University of Illinois Press.

Shea, G.S. (1991). Uncertainty and implied variance bounds in long-memory models of the interest rate term structure. Empirical Economics, 16, 287-312.

Shiller, R.J. (1981). Do stock prices move by too much to be justified by subsequent changes in dividends? American Economic Review, 71, 421-436.

Shiller, R.J. (2000). Irrational exuberance, Princeton, NJ: Princeton University Press.

Shimotsu, K. (2006). Simple but effective tests of spurious long memory, QED working paper #1101, Queen's University Department of Economics.

Singer, H.W. (1950). The Distribution of gains between investing and borrowing countries. American Economic Review, Papers and Proceedings, 40, 473-485.

Smith, A. (2005). Level shifts and the illusion of long memory in economic time series. Journal of Business and Economic Statistics, 23:3, 321-335.

Smith, A. (2005). Partially overlapping time series: A new model for volatility dynamics in commodity futures. Journal of Applied Econometrics, 20:3, 405-422.

Smith, J.O. (2003). Mathematics of the discrete fourier transform (DFT), W3K Publishing, http://books.w3k.org/, 2003, ISBN 0-9745607-0-7.

Sorensen, C. (2002). Modeling seasonality in agricultural commodity futures. Journal of Futures Markets, 22, 393-426.

Sottinen, T. (2001). Fractional Brownian motion, random walks and binary market models. Finance and Stochastics, 5, 343-355.

Sowell, F. (1992a). Modeling long-run behavior with the fractional ARIMA model. Journal of Monetary Economics, 29, 277-302.

Sowell, F. (1992b). Maximum likelihood estimation of stationary univariate fractionally-integrated time-series models. Journal of Econometrics, 53, 165-188.

Sowell, F.B. (1990). The Fractional unit root distribution. Econometrica, 58, 495-505.

Starica, C. and C.W.J. Granger (2001). Non-stationarities in stock returns. Preprint.

Steffen, P., P.N. Heller, R.A. Gopinath and C.S. Burrus (1993). Theory of regular m-band wavelets. IEEE Transactions on Signal Processing, 41, 3497-3511.

Stiglitz, J.E. (1990). Symposium on bubbles. Journal of Economic Perspectives 4, 13-18.

Stock, J.H. and M.W. Watson (1988). Testing for common trends. Journal of the American Statistical Association (Theory and Methods), 83, 1097-1107.

Stock, J.H. and M.W. Watson (2002). Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics, 20, 147-162.

Strang, G. and T. Nguyen (1997). Wavelets and filter banks, Upper Saddle River, New Jersey: Prentice-Hall.

Sweldens, W. (1995). The Lifting Scheme: A new philosophy in biorthogonal wavelet constructions. In A. F. Laine and M. Unser, editors, Wavelet Applications in Signal and Image Processing III, pp. 68-79, Proc. SPIE 2569, 1995.

Sweldens, W. (1998). The Lifting scheme: A construction of second generation of wavelets. SIAM J. Math. Anal., 29:2, 511-546.

Szego, G. (1975). Orthogonal polynomials, 4th Edition. Colloquium Publications, XXIII, Rhode Island: American Mathematical Society.

Tanaka, K. (1999). The Nonstationary fractional unit root. Econometric Theory, 15, 549-582.

Tanaka, K. (2004). Frequency domain and wavelet-based estimation for long memory plus noise models, in: Harvey, A.C., S.J. Koopman and N. Shephard, Ed., State Space and Unobserved Component Models: Theory and Applications, Cambridge, U.K: Cambridge University Press.

Taqqu, M.S. (1975). Weak convergence to fractional Brownian motion and to the Rosenblatt process. Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete, 31, 287-302.

Taqqu, M.S., V. Teverovsky and W. Willinger (1995). Estimators for long range dependence: an empirical study. Fractals 3, 785-798.

Telser, L.G. (2000). Classic futures: lessons from the past for the electronic age (Risk Books).

Teolis, A. (1998). Computational signal processing with wavelets, Birkhauser.

Teyssière, G. (1996). Double long-memory financial time series. Preprint.

Teyssière, G. (1997). Modelling exchange rates volatility with multivariate long-memory ARCH processes. Journal of Business and Economic Statistics, under revision.

Teyssière, G. (1998). Multivariate long-memory ARCH modelling for high frequency foreign exchange rates. In: Proceedings of the HFDF-II Conference, Zurich 1998. Zurich: Olsen & Associates.

Teyssière, G. (2003). Interaction models for common long-range dependence in asset price volatility. In: G. Rangarajan and M. Ding, Eds, Processes with Long Range Correlations: Theory and Applications Lecture Notes in Physics Vol. 621, Berlin: Springer, 251-269.

Teyssière, G. and P. Abry (2006). Wavelet analysis of nonlinear long-range dependent processes with applications to financial time series. In: Teyssière, G. and A. Kirman, ed. (2006), Long memory in economics, Berlin: Springer, pp. 173-238.

Thaler, R.H. (1991). The Winner's curse. Princeton NJ: Princeton University Press.

Theil, H. (1961). Economic forecast and policy. Amsterdam: North-Holland.

Thornton, D. and P. Yue (1992). An Extended series of Divisia monetary aggregates, Federal Reserve Bank of St. Louis Review, 74:6, 35-52.

Tirole, J. (1985). Asset bubbles and overlapping generations. Econometrica, 53, 1499-1528.

Tomek, W. G., and H. H. Peterson (2001). Risk management in agricultural markets: a review. The Journal of Futures Markets, 21: 953-985.

Tomek, W.G. (1997). Commodity futures prices as forecasts. Review of Agricultural Economics, 19(1): 23-44.

Tomek, W.G. (2000). Commodity prices revisited. Agricultural and Resource Economics Review 29(2): 125-137.

Townsend, R.M. (1983). Forecasting the forecasts of others. Journal of Political Economy, 91, 546-588.

Tsay, W. and C. Chung (1995). The Spurious regression of fractionally integrated processes, working paper 9503, Michigan State University.

Turvey, C.G. (2007). A Note on scaled variance ratio estimation of the Hurst exponent with application to agricultural commodity prices. Physica A, 377:1, 155-165.

Vaidyanathan, P.P. (1990). Multirate digital filters, filter banks, polyphase networks and applications: a tutorial. Proceedings of the IEEE, 78, 56-93.

Vaidyanathan, P.P. and P.Q. Huong (1988). Lattice structures for optimal design and robust implementation of two-channel perfect-reconstruction QMF banks. IEEE Trans. on Acoustics, Speech, and Signal Processing, 36:1, 81-94.

Veitch, D. and P. Abry (1999). A wavelet based joint estimator of the parameters of long-range dependence. IEEE Transactions on Information Theory, 45, .878-897.

Velasco, C. (1999). Gaussian semiparametric estimation of non-stationary time series. Journal of Time Series Analysis, 20, 87-127.

Velasco, C. (2003). Gaussian semi-parametric estimation of fractional cointegration. Journal of Time Series Analysis, 24, 345-378.

Velasco, C. and P.M. Robinson (2000). Whittle pseudo-maximum likelihood estimation for nonstationary time series. Journal of the American Statistical Association, 95, 1229-1243.

Vetterli, M. and J. Kovacevi´c (1995). Wavelets and subband coding, Upper Saddle River, New Jersey: Prentice-Hall.

Vogelsang, T.J. and P. Perron (1998). Additional tests for unit root allowing the possibilityof breaks in the trend function. International Economic Review, 39, 1073-1110.

Vostrikova, L.Ju. (1981). Detection of disorder in multidimensional random processes. Soviet Mathematics Doklady, 24, 55-59.

Vrugt, E.B. (2003). Tactical commodity strategies in reality: An economic theory aproach, ABP Working Paper.

Wang, D. and W. Tomek (2007). Commodity prices and unit root tests. American Journal of Agricultural Economics, forthcoming.

Wärneryd, K.-E. (2001). Stock market psychology, Cheltenham, UK: Edwar Elgar.

Wei, A. and R.M. Leuthold (2000). Agricultural futures prices and long memory processes. Working paper, University of Illinois Office for Futures and Options Research, Paper Number 00-04.

Weiser, S. (2003). The strategic case for commodities in portfolio diversification. Commodities Now, 7-11.

Weiss, J. (1970). A Spectral analysis of world cocoa prices. American Journal of Agricultural Economics, 52:1, 122-127.

West, B.J. and L. Griffin (1998). Allometric control of human gait. Fractals, 6:2, 101-108.

West, B.J., P. Hamilton and D.J. West (1999). Fractal scaling and the teen birth phenomenon. Fractals, 7:2, 113-122.

Whitcher, B. (2004). Wavelet-based estimation for seasonal long-memory processes. Technometrics, 46, 225-238.

Whitcher, B. (2005). Wavelet analysis. In: B. S. Everitt and C. Palmer (Eds.) Encyclopaedic Companion to Medical Statistics, Hodder Arnold, 367-368.

Whitcher, B. and P. F. Craigmile (2004). Multivariate spectral analysis using Hilbert wavelet pairs. International Journal of Wavelets, Multiresolution and Information Processing, 2:4, 567-587.

Whitcher, B., P. F. Craigmile and P. Brown (2005). Time-varying spectral analysis in neurophysiological time series using Hilbert wavelet pairs. Signal Processing, 85:11, 2065-2081.

Whitcher, B., P. Guttorp and D. B. Percival (2000). Multiscale detection and location of multiple variance changes in the presence of long memory. Journal of Statistical Computation and Simulation, 68:1, 65-88.

Whitcher, B., P. Guttorp and D. B. Percival (2000). Wavelet analysis of covariance with application to atmospheric time series. Journal of Geophysical Research, 105:D11, 14941-62.

Whittle, P. (1951). Hypothesis testing in time series analysis. Almquist and Wiksells, Uppsala.

Wickerhauser, M.V. (1991). INRIA lectures on wavelet packet algorithms. Proceedings ondelettes et paquets d'ondes, 17-21 june, Rocquencourt, France, 31-99.

Wickerhauser, V.M. (1994). Adapted wavelet analysis from theory to software, Natick, Massachusetts: A.K. Peters.

Williams, J. C. (2001). Commodity futures and options, in Handbook of Agricultural Economics, vol. 1, B.Gardner and G. Rausser, Eds. Elsevier Science, Chapter 13.

Woo, W.T. (1987). Some evidence of speculative bubbles in the foreign exchange markets. Journal of Money, Credit and Banking, 19, 499-514.

Woodward, P.M. (1953). Probability and information theory with applications to radar, New York: McGraw-Hill Book Company.

Wright, J.H. (1998). Testing for a structural break at unknown date with long-memory disturbances. Journal of Time Series Analysis, 19, 369-376.

Wright, J.H. (2000). Log-periodogram estimation of long memory volatility dependencies with conditionally heavy-tailed returns. FRB International Finance Discussion Paper No. 2000-685.

Wu, G. and Z. Xiao (2002). Are there speculative bubbles in stock markets? Evidence from an alternative approach. Preprint.

Xiao, Z. and P.C.B. Phillips (2002). A CUSUM test for cointegration using regression residuals. Journal of Econometrics, 108, 43-61.

Yang, S. and B.W. Brorsen (1993). Nonlinear dynamics of daily futures prices: Conditional heteroskedasticity or chaos? Journal of Futures Markets, 13, 175-192.

Yang, S.-R. and B.W. Brorsen (1992). Nonlinear dynamics of daily cash prices. American Journal of Agricultural Economics, 74, 706-15.

Yao, Y. (1988). Estimating the number of change-points via Schwarz criterion. Statistics and Probability Letters 6 (1988), pp. 181-189.

Yoon, G. (2003). A simple model that generates stylized facts of returns. Korea: Pusan National University, UCSD Working Paper, San Diego, CA.

Young, C.H. (1974). Asymptotic behaviour of trigonometric series. Hong Kong: Chinese University.

Zaffaroni, P. (2003). Gaussian inference on certain long-range dependent volatility models. Journal of Econometrics, 115, 199-258.

Zarepour, M. and K. Knight (1991). Bootstrapping unstable first order autoregressive process with errors in the domain of attraction of a stable law. Communications in Statistics-Stochastic Models, 15, 11-27.

Zeeuw, P.M. (1998). Wavelet and image fusion, CWI, Amsterdam. http:/www.cwi.nl/~pauldz/

Zivot, E., and D.W.K. Andrews (1992). Further evidence on the great crash, the oil-price shock, and the unit root hypothesis. Journal of Business and Economic Statistics, 10, 251-270.