THREE ESSAYS IN COMMODITY PRICE DYNAMICS

AMAL DABBOUS

A THESIS
IN
THE DEPARTMENT
OF
ECONOMICS

PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
CONCORDIA UNIVERSITY
MONTRÉAL, QUÉBEC, CANADA

JULY 2015

CONCORDIA UNIVERSITY

Economics

This is to certify that the thesis prepared

By: **Mrs. Amal Dabbous**
Entitled: **Three Essays in Commodity Price Dynamics**

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Economics)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair
Dr. Linda Dyer

______________________________ External Examiner
Dr. John W. Galbraith

______________________________ Examiner
Dr. Ian Irvine

______________________________ Examiner
Dr. Prosper Dovonon

______________________________ Supervisor
Dr. Bryan Campbell

Approved by

____
Dr. Effrosyni Diamantoudi
Graduate Program Director

Abstract

Three Essays in Commodity Price Dynamics

Amal Dabbous, Ph.D.
Concordia University, 2015

This thesis consists of three essays in commodity price dynamics. In the first essay, we embed a staggered price feature into the speculative storage model of Deaton and Laroque (1996). Intermediate goods inventory speculators are added as an additional source of intertemporal linkage which helps us to replicate the stylized facts of the observed commodity price dynamics. The staggered pricing mechanism adopted in this paper can be viewed as a parsimonious way of approximating various types of frictions that increase the degree of persistence in the first two conditional moments of commodity prices. The structural parameters of our model are estimated by simulated method of moments using actual prices for four agricultural commodities. Simulated data are then employed to assess the effects of our staggered price approach on the time series properties of commodity prices. Our results lend empirical support to the possibility of staggered prices.

The second essay investigates the determinants of the percentage change in commodity prices. We apply the dynamic Gordon growth model technique and conduct the variance decomposition for the percentage change in spot commodity prices to 6 agricultural commodities. The model explains the percentage change in spot commodity prices in terms of the expected present discounted values of interest rate, yield spread, open interest and convenience yield. Empirical results indicate that the model is successful in capturing a large proportion of the variability in the 6 agricultural commodity prices. Moreover, we show that yield spread and open interest help predicting changes in commodity prices.

Finally, the third essay evaluates different hedging strategies for eleven commodities. In addition to the traditional regression hedge ratio model (OLS) and the vector error correction model (VECM), we estimate dynamic hedge ratios using the conventional dynamic conditional correlation model (DCC) of Engle (2002) and the diagonal BEKK model (DBEKK) of Engle and Kroner (1995). Moreover, we propose two more advanced models, the DCC model and the DBEKK

model that will account for the impact of the growth rate of open interest on market's volatility and co-movements of commodity spot and futures returns. The empirical analysis shows that adding the growth rate of open interest improves the in-sample hedging effectiveness of the DCC model. Furthermore, the out-of-sample hedging exercise empirical results show that static models present the best out-of-sample hedging performance for 5 of the commodities. The DCC model presents the smallest basis variance for 4 of the commodities. The DBEKK model with the growth rate of open interest performs the best in terms of the basis variance reduction for corn and wheat. Our out-of-sample empirical findings provide important implications for futures hedging and highlight the fact that the use of static models to determine the optimal hedge ratio could be more effective than the use of dynamic hedge ratio models.

Acknowledgments

I would like to express my sincere gratitude to my supervisor Prof. Bryan Campbell for his continuous support, patience, motivation and immense knowledge. Prof. Campbell was abundantly helpful and offered invaluable assistance, insightful comments and guidance. Moreover, I would like to thank him for giving me the opportunity to present the first essay of this thesis at CIRANO[1], and helping to complete some empirical works of this part there.

I would also like to convey thanks to my then supervisor Prof. Nikolay Gospodinov and my colleague Dr. Hirbod Assa for collaborating on the first chapter of the thesis, which is a joint work with them.

Deepest gratitude are also due to the member of the supervisory committee. In particular, I would like to thank Prof. Dovonon and Prof. Irvine for their support during my studies at Concordia. I would also use this occasion to convey my thanks for Prof. LeBlanc (the economics department chairperson) and to Prof. Diamantoudi (the graduate program director) for their continuous help and guidance.

Last but not the least I would like to express my love and gratitude to my beloved family, particularly my husband (Zaher), my kids (Ziad and Mazen), my mother and my twin sister (Imane) for their understanding and endless support and patience, through the duration of my studies.

[1] Center for Interuniversity Research and Analysis of Organizations

Contents

List of Figures

List of Tables

Chapter 1

A Staggered Pricing Approach to Modeling Speculative Storage: Implications for Commodity Price Dynamics

1.1 Introduction

The last decade has witnessed a surge in commodity prices and a widespread financialization of commodity products. The upward movements and the increased volatility of the commodity prices have been largely attributed to strong demand by China and other emerging markets as well as massive capital flows into the commodity markets by institutional investors, portfolio managers and speculators. While the importance of commodity price movements for the economic policy and investors' sentiment has generated a substantial research interest, the behavior and the determination of commodity prices is not yet fully understood. The main objective of this chapter is to develop a structural model of commodity price determination that reflects the empirical properties (high persistence and conditional heteroskedasticity) of commodity prices. In order to achieve this goal and to gain further understanding into the fundamental factors that drive the observed behavior of commodity prices, we modify the structure of the speculative storage model from one where the prices adjust almost instantaneously to harvest shocks to a setup where they change slowly and infrequently. More specifically, we depart from the assumption that market prices are determined in a perfectly competitive environment and extend the basic speculative storage model by explicitly introducing intermediate goods speculators with a staggered pricing rule. One appealing aspect of this approach is its ability to mimic some important characteristics of the actual commodity prices such as high persistence and conditional heteroskedasticity, which can be generated even in the absence of correlated harvest shocks.

The speculative storage model for commodity prices can be dated back to Gustafson (1958) who defines a set of optimal storage rules that state how much grain should be carried over into the next period given the current year supply. Moreover, by introducing intertemporal storage arbitrage and supply shocks, Gustafson (1958) incorporates rational expectations. This line of research is further elaborated in Muth (1961). Samuelson (1971) develops a model for commodities which determines the behavior of the prices as the solution to a stochastic dynamic programming problem. Furthermore, Beck (1993) builds upon the work by Muth (1961) and provides a theoretical basis for treating the variance of storable commodities as serially correlated which suggests that commodity prices may exhibit conditional heteroskedasticity. The presence of storage is instrumental in ensuring that the price variance in one period directly affects inventory variance which in turn is transmitted to next period's price variation. Williams and Wright (1991) provide a comprehensive discussion of the basic storage model and its extensions and summarize the time series properties of storable commodities. Williams and Wright (1991) put an emphasis on the complex non-linear storage behavior resulting from the fact that aggregate storage cannot be negative.

Deaton and Laroque (1992, 1995, 1996) develop a partial equilibrium structural model of commodity price determination and apply numerical methods to test and estimate the model parameters, confronting for the first time the storage model with the documented behavior of actual prices. Their analysis suggests that the introduction of speculative inventories and serially correlated supply shocks do not appear to generate sufficient persistence in commodity prices although they prove to be successful in replicating the substantial volatility observed in the actual data.

More recently, numerous studies have focused on modifying the storage model in order to accommodate the persistence of commodity prices. Chambers and Bailey (1996) relax the *iid* assumption on harvest shocks, and study the price fluctuations of storable commodities, assuming that shocks are either time dependent or that the model exhibits periodic disturbances. Ng and Ruge-Murcia (2000) incorporate additional features into the storage model in an attempt to generate a higher degree of persistence in commodity prices. In particular, Ng and Ruge-Murcia (2000) allow for serially correlated shocks assuming that harvest follows a first-order moving average

(MA(1)) process. They also examine the ability of production lags and heteroskedastic supply shocks, multi-period forward contracts and convenience yields to generate an increased persistence in commodity prices. Cafiero, Bobenrieth, Bobenrieth, and Wright (2011) demonstrate that the competitive storage model can give rise to high levels of serial correlation observed in commodity prices if more precise numerical methods are employed. Moreover, estimates for seven commodities supported the specification of the speculative storage model with positive constant marginal costs and no deterioration, which is in line with Gustafson (1958).

Furthermore, Cafiero, Bobenrieth, Bobenrieth, and Wright (2011) use a maximum likelihood framework to estimate the storage model with stock-outs, which is extended to include unbounded harvests and free disposal. Their results produce more accurate small sample estimates of the structural parameters of the model compared to the previous studies based on the pseudo-maximum likelihood procedure. Miao and Funke (2011) add shocks to the trends of output and demand. Evans and Guthrie (2007) include transaction cost frictions into the speculative storage model. One important finding that emerges from their analysis is that these frictions tend to have explanatory power for the dynamic behavior of spot and futures commodity prices. In a competitive equilibrium framework, the model of Evans and Guthrie (2007) is able to capture the serial correlation and GARCH characteristics of commodity prices. Finally, Arseneau and Leduc (2012) embed the speculative storage model into a general equilibrium framework. Their main result is that the interaction between storage and interest rates in general equilibrium increases the impact of competitive storage on commodity prices and leads to higher persistence than the one observed in the storage model with fixed interest rate.

In spite of this extensive literature for understanding the determinants and the dynamic patterns of commodity prices, reproducing the documented high persistence and conditional heteroskedasticity of actual prices within a well-articulated structural model proved to be a challenging task. In this chapter, we address the issues regarding the commodity price dynamics in a unified fashion by embedding a staggered pricing mechanism into the speculative storage model. While Arseneau and Leduc (2012) also suggest to "introduce staggered price setting on the part of the final goods producing firm" in a general equilibrium framework as a possible extension for future research, our work is the first to implement this approach and assess the properties of the model-generated commodity prices against the observed data.

In an attempt to depart from the assumption of perfect competition at both the production and storage activity, Newbery (1984), Williams and Wright (1991), and McLaren (1999) investigate the effects of market power on the storage behavior. Our model differs from their work along the dimension that the final bundler does not store the good and the storage is only done by intermediate risk neutral speculators. The final bundler only bundles intermediate prices in order to set the final price. Moreover, Rust and Hall (2003a) present a model of optimal price speculation by a middleman in the steel market. A third player, the middleman is added endogenously to the model. The middlemen will purchase steel from producers and other middlemen in the market and sell it at a mark up to retail consumers in order to make profits. The firm makes a purchase only when the current level of inventory is below a purchase threshold. The authors use the simulated minimum distance estimator to estimate the structural parameters of the model since wholesale prices are endogenously sampled, that is, wholesale prices are only observable on the dates when the middleman purchases steel.

In addition, Rust and Hall (2003b) introduce a model in which producers and consumers in the steel market can choose between two intermediate market players, a market maker which typically owns or is member of a centralized exchange where bid and ask prices are publicly known, and a middlemen. In their model, the share of trade intermediated by middleman and market makers are endogenously determined. In the steel market there are no market makers, traders have to search and bargain to find the lowest price. They show that it is possible for those two intermediaries to coexist in the market. They discuss first a market equilibrium in which only middlemen exist assuming there is a continuum of middlemen. Furthermore, they define conditions under which the presence of a single market maker or monopolist is possible with the presence of middlemen, even though surviving middlemen tend to reduce their bid and ask spreads to undercut the market maker prices. They also present the range of equilibrium outcomes when the market has four players: consumers, middlemen, market maker and producers. Finally, Mitraille and Thille (2009) examine the market power in production with competitive storage by analyzing the effects that competitive storage has on the behavior of a monopolist. Using his market power, the monopolist can influence speculative activity by manipulating prices and consequently affect the distribution of prices. One of the findings of Mitraille and Thille (2009) is that stockouts occur less frequently under monopoly.

The focus of this chapter is on the improved ability of the storage model with staggered prices to account for the empirical features of commodity prices. The main impact of staggered prices in our model is to dampen the movements in prices as well as the market power of intermediate speculators to affect prices. This leads to gradual adjustments and persistent responses of prices following a harvest shock. In addition to generating sufficient persistence in commodity prices, the staggered pricing approach allows us to match other important moments in the unconditional and conditional distributions of the commodity prices.

Nominal price rigidity is often incorporated in dynamic general equilibrium models with two widely used nominal price rigidity specifications in the literature. On one hand, the partial adjustment model developed by Calvo (1983), Rotemberg (1987), and Rotemberg (1996) allows for only a randomly chosen fraction of firms to adjust their prices according to some constant hazard rate in any given period. On the other hand, the staggered price setting rule adopted by Taylor (1980) and Blanchard and Fisher (1989) permits all firms to optimize their prices after a fixed number of periods.

In this chapter, we assume that the pricing decisions are staggered as in Calvo (1983) and use a similar modeling framework as the one developed in McCandless (2008). Even though the staggered pricing is not generated endogenously within the model, it serves as a useful device to impart the inefficiencies in the agricultural commodity markets such as price floors, subsidies, import/export quotas and controls, government strategic stock reserves, collusion etc. that prevent prices to adjust instantaneously to changes in economic conditions. Note that these types of market inefficiencies allow us to depart from the typical assumption in the literature that commodities tend to be homogeneous products whose prices are fully flexible and equal to their marginal costs. Our results confirm the importance of staggered prices for commodity price dynamics and suggest that the staggered pricing mechanism appears to be consistent with the behavior of the actual data. Moreover, we show how our model can be used to analyze the response of commodity prices to harvest shocks which provides a framework for economic and policy evaluation.

The remainder of the chapter is organized as follows. The competitive storage model with staggered prices as well as the statistical characterizations of this model are presented in Section 2. Section 3 studies the practical implications of our staggered price speculative storage model using simulated data. Section 4 contains a brief description of the data and the estimation method used in the paper, and presents the main empirical results. Section 5 concludes.

1.2 Competitive Storage Model with Staggered Prices

This section introduces the model setup and characterizes the equilibrium and statistical behavior of the model-generated commodity prices.

1.2.1 Model and Equilibrium Price Behavior

The rational expectations model determines the optimal inventory decisions by risk- neutral speculators. The basic version of the model developed by Deaton and Laroque (1992, 1995, 1996)[1] incorporates competitive storage into the consumer demand and supply dynamics and establishes the concept of stationary rational expectations equilibrium (SREE). The model with serial correlation in harvest shocks is tested by Ng and Ruge-Murcia (2000). In their paper, Ng and Ruge-Murcia (2000) consider an MA(1) specification for the model harvest shocks. Our model complements and extends the original DL model by embedding a staggered price setting into the speculative storage model. Regarding the harvest shock specification, we consider both (i) *iid* harvest shocks and (ii) MA(1) harvests shocks.

Our modified model has three types of commodity market participants: final consumers, intermediate risk neutral speculators and a bundler[2] who bundles the commodities in order to set the final price. In the absence of storage, the behavior of final consumers is characterized by a linear inverse demand function

$$p_t = P(z_t) = a + bz_t,$$

where a and $b < 0$ are parameters to be estimated and z_t denotes the harvest in period t.

Let the harvest z_t be given by

$$z_t = \bar{z} + u_t,$$

where $\bar{z}$ is constant (perfectly inelastic) and u_t is a random disturbance term which is assumed either to be *iid* or to follow an MA(1) process

$$u_t = e_t + \rho e_{t-1},$$

[1] For brevity, we denote hereafter the basic speculative storage model of Deaton and Laroque by DL.

[2] In the literature, it is common to use the term "monopolist" instead of the term "bundler" that we use in this chapter. The reason that we prefer the latter is the following: in the staggered pricing literature, the final goods producer maximizes profits by setting the price. In this chapter, we do not consider any profit maximization and any type of price setting for the final goods producer. Instead, we use directly the final goods prices as set in (1.6).

where e_t is $iid(0, \sigma^2)$. If $\rho = 0$, we have the case of *iid* shocks as in DL, and when $\rho > 0$, we have MA(1) shocks as in Ng and Ruge-Murcia (2000). In this chapter, we investigate both cases and show that when we add staggered prices, the case for $\rho = 0$ gives better results compared to the case of non-staggered prices and $\rho > 0$.

Intermediate risk neutral speculators or inventory holders know the current year harvest and demand the commodity to transfer to the next period. They will do so whenever they expect to make a profit above the storage and interest cost. The depreciation rate of storage is denoted by δ. A simple form of proportional deterioration is considered which means that if in period t the speculators store I units of the commodity, they have at their disposal $(1-\delta)I$ units at the beginning of the next period. Moreover, speculators have to pay the real interest rate on the value of their storage. Let r denotes the constant exogenous real interest rate. The sum of harvest and inherited inventories, denoted by x_t, is referred to as the amount on hand and is given by

$$x_t = (1-\delta)I_{t-1} + z_t.$$

The relationship between the amount of storage and its net profit can be summarized as

$$\begin{cases} I_t > 0 \text{ if } (1-\delta)/(1+r)\mathbb{E}_t[p_{t+1}] = p_t, \\ I_t = 0 \text{ otherwise,} \end{cases}$$

where $\mathbb{E}_t$ denotes the expectation given the information at time t.

The condition for non-negative inventories is the crucial source of non-linearity in the model. This specification does not allow the market participants to borrow commodities that have not yet been grown. In addition, intermediate speculators benefit from market power that reflects their ability to affect the price. In this framework, we assume that there is a continuum of intermediate speculators (of unit mass indexed by $k \in [0,1]$) and final big players in the market. Final players collect all the commodities from intermediate speculators and bundle intermediate speculators' prices into the final price in order to sell the commodity to consumers.

For simplicity, we assume that there exists a bundler who bundles all intermediate speculators' prices into a single one. Each period t, a fraction $1-\gamma$ $(0 < 1-\gamma < 1)$ of the speculators is able to exploit their market power and to reset the prices of their commodities $P_t^*(k)$. In contrast, those who did not benefit from their market power to affect prices, retain their last period prices:

$P_t^*(k) = P_{t-1}^*(k)$. Given this staggered pricing rule, along with the assumptions that speculators are risk neutral and have rational expectations, intermediate speculators' current and expected future prices must satisfy

$$P_t^*(k) = \max\left\{p(x_t), (1-\gamma)\frac{1-\delta}{1+r}\mathbb{E}_t[P_{t+1}^*(k)] + \gamma P_t^*(k)\right\}. \tag{1.1}$$

The first term in the brackets represents the price if the harvest is sold to consumers in period t and no inventories are carried over to the next period. The second term is known as the intertemporal Euler equation. This is the value of one unit stored if $1-\gamma$ of the speculators benefit form their market power to affect the price. This, in turn, occurs if the speculators expect to cover their costs (after depreciation) from buying the commodity at time t. Since the current period bundler prices are not yet determined, it is important to stress that speculators, who do not reset their prices, use their own current prices and not the market ones in order to determine $P_t^*(k)$ in (1.1).

Finally, the bundler will bundle all intermediate prices together according to the following pricing rule (see McCandless (2008))

$$P_t^{1-\psi} = \gamma P_{t-1}^{1-\psi} + (1-\gamma)P_t^*(k)^{1-\psi},$$

where P_t denotes the bundler final price of the good, the parameter ψ is the gross markup of the intermediate goods speculators and $P_t^*(k)$ represents the price for intermediate goods speculators who can set their prices. Since all intermediate goods speculators who can fix their prices are assumed to have the same markup over the same marginal costs, $P_t^*(k)$ is the same for all intermediate risk neutral speculators who adjust their prices. Prices for intermediate speculators who cannot set their prices are the same as the previous period prices denoted by P_{t-1}.

In order to simplify the bundler's pricing rule, we use the log-linearized version of this equation so that the final price becomes

$$\tilde{p}_t = \gamma\tilde{p}_{t-1} + (1-\gamma)\tilde{p}_t^*(k), \tag{1.2}$$

where $\tilde{p}_t$ and $\tilde{p}_t^*$ denote the logarithm of P_t and P_t^*, respectively.

It is worth noting that our model bears some resemblance with the model introduced in Rust and Hall (2003a) and Rust and Hall (2003b), the idea of introducing intermediate risk neutral speculators and a bundler is very similar to middlemen and market maker introduced in their steel market model. As in their model, we also assume that there is a continuum of intermediate risk neutral speculators whose prices are not publicly known and that benefit from market power to affect the price and a bundler that will decide the final price at which the commodity is sold to consumers. Moreover, the final bundler price in our model is known and available to final consumers, which is similar to the market maker price in Rust and Hall (2003b) paper. However, the major difference resides in the fact that we add the staggered pricing rule exogenously to model intermediate risk neutral speculators ' prices, whereas wholesale prices in the steel market model are endogenously sampled.

After completing the description of our model, we elaborate on some important implications of equation (1.1). As implied by this equation, the intermediate risk neutral speculators' price follows a non-linear first-order Markov process with a kink at the price above which we do not have inventories. In the case of *iid* shocks, the kink is determined by

$$\hat{p} = (1-\gamma)\frac{1-\delta}{1+r}\mathbb{E}p(z) + \gamma\hat{p}.$$

This implies that

$$\hat{p} = \frac{1-\delta}{1+r}\mathbb{E}p(z), \tag{1.3}$$

which coincides with the kink given in DL.

However, as in Chambers and Bailey (1996), the price kink $\hat{p}$ in the case of correlated harvests shocks is no longer constant and varies with the current harvest. This is due to the fact that with serially correlated harvest shocks, speculators form their price forecasts using all the information contained in the current shock.

Under some regularity conditions, most notably $r+\delta > 0$ and that z has a compact support, DL establish the existence of a solution to (1.1) when $\gamma = 0$ and shocks are independent. Indeed, to show the existence of the demand function for non-independent shocks, it is enough to prove the independent case conditioning on time t. In our case, we proceed by following a similar approach to proving that such an equilibrium exists. Assume that the demand x_t always lies in a subset $\mathbb{X} = [\underline{z}, +\infty)$ of the real numbers and that the harvest shock z_t belongs to a compact set $\mathbb{Z} = [\underline{z}, \bar{z}]$.

Definition Assume that $\gamma \in [0,1)$. A staggered stationary rational expectation equilibrium (SSREE) is a price function $f : \mathbb{X} \times \mathbb{Z} \to \mathbb{R}$ which satisfies the following equation

$$p_t = f(x_t, z_t) = \max\left\{ p(x_t), (1-\gamma)\frac{1-\delta}{1+r}\mathbb{E}_t f(z_{t+1} + (1-\delta)I_t, z_{t+1}) + \gamma f(x_t, z_t) \right\},$$

where ,

$$I_t = x_t - p^{-1}(p_t) = x_t - p^{-1}(f(x_t, z_t)). \tag{1.4}$$

This defines the price function

$$P_t^*(k) = f(x_t, z_t),$$

where $f(x_t, z_t)$ is the unique, monotone decreasing in its first argument, solution to the functional equation. Since this price function is non-linear, numerical techniques similar to the ones adopted by DL and Michaelides and Ng (2000) are used to solve for $f(x_t, z_t)$

$$f(x_t, z_t) = \max\left\{ p(x_t), (1-\gamma)\frac{1-\delta}{1+r}\mathbb{E}_t f((z_{t+1} + (1-\delta)I_t), z_{t+1}) + \gamma f(x_t, z_t) \right\}.$$

In the case of independent shocks, we can remove the time subscript and the shocks in f.

When $\gamma = 0$ and the shocks are *iid*, we have the same model as the one considered by DL. Hence, the equilibrium is simply called SREE. In the following theorem we show that the staggered stationary rational expectation equilibrium (SSREE) coincides with the stationary rational expectation equilibrium (SREE) derived from the basic DL speculative storage model.

Theorem 1.2.1 *If shocks are iid, then SSREE=SREE.*

Proof See Appendix A. ■

Remark Theorem 1.2.1 shows that $p_t = P_t^*$. This allows us to use all of the results for the process p_t, that are available in the literature, for the process P_t^*.

We next show that the final demand for the bundler in our staggered speculative model is different form the one in DL. It proves useful to compare the price processes in the speculative storage model with and without staggered prices for the market participants who can reset their prices. In the basic speculative storage model of DL, the market participants cannot hold negative inventories. If prices are expected to increase or decrease by less than the cost of carrying the commodity from one period to another, inventories are zero. If inventories are positive, the expected price next period is equal to the current price plus the storage costs. The final price of the commodity in the basic speculative storage model satisfies

$$p_t = \max\left\{p(x_t), \frac{1-\delta}{1+r}\mathbb{E}_t p_{t+1}\right\}.$$

Hence,

$$\begin{cases} p_t = \frac{1-\delta}{1+r}\mathbb{E}_t p_{t+1} & \text{if } I_t > 0; \\ p_t = p(x_t) & \text{if } I_t = 0. \end{cases}$$

However, as stated in the description of our speculative storage model with staggered prices, the intermediate risk neutral speculators price function satisfies

$$P_t^* = \max\left\{p(x_t), (1-\gamma)\frac{1-\delta}{1+r}\mathbb{E}_t P_{t+1}^* + \gamma P_t^*\right\}.$$

In this case,

$$\begin{cases} P_t^* = (1-\gamma)\frac{1-\delta}{1+r}\mathbb{E}_t P_{t+1}^* + \gamma P_t^*, & \text{if } I_t > 0; \\ P_t^* = p(x_t) & \text{if } I_t = 0. \end{cases} \tag{1.5}$$

It can be easily seen from (1.5) that the prices for intermediate risk neutral speculators who can adjust them satisfy the same equation as the one that speculators face in the basic storage model of DL.

Since the final price process in the speculative storage model with staggered prices is given by

$$\tilde{p}_t = \gamma \tilde{p}_{t-1} + (1-\gamma)\tilde{p}_t^*(k), \tag{1.6}$$

one can infer that the demand of the bundler (the final demand) will be different from the demand presented by DL in the basic speculative storage model. We expect the final demand for speculative storage model with staggered prices to be in between the DL demand and the regular market demand. Moreover, we expect this demand to be more inelastic than the one derived from the basic speculative storage model. This is more consistent with the commodity elasticities estimated from actual data.

1.2.2 Statistical Characterization

Under the assumption of *iid* harvests shocks, the final log-price process satisfies equation (1.6). The bundler price can then be written as

$$P_t = {P_{t-1}}^{\gamma}\, {P_t^*}^{1-\gamma}. \tag{1.7}$$

The persistence of commodity prices is then simply an outcome of the staggered prices which is extensively discussed in the literature on staggered pricing. Here, we provide an alternative explanation. From the logarithmic form of the relation (1.7), we have by induction that

$$\tilde{p}_{t+1} = (1-\gamma)\sum_{i=0}^{t} \gamma^i \tilde{p}^*_{t+1-i}$$

which in turn yields

$$P_{t+1} = \left(\prod_{i=0}^{t} {P^*_{t+1-i}}^{\gamma^i}\right)^{1-\gamma}.$$

This shows that P_{t+1} shares overlapping terms prices in previous periods which gives rise to high persistence.

Next, we show that the final prices of the bundler exhibit conditional heteroskedasticity which is another salient characteristic of the observed commodity prices. Note that from (1.7), we have

$$\mathbb{E}_{t-1}(P_t^2) = {P_{t-1}}^{2\gamma}\mathbb{E}_{t-1}({P_t^*}^{2(1-\gamma)}) \tag{1.8}$$

and

$$(\mathbb{E}_{t-1}P_t)^2 = {P_{t-1}}^{2\gamma}(\mathbb{E}_{t-1}({P_t^*}^{1-\gamma}))^2. \tag{1.9}$$

Combining (1.8) and (1.9) and assuming that the shocks are *iid*, the conditional variance of the final prices is given by

$$\mathrm{Var}_{t-1}(P_t) = {P_{t-1}}^{2\gamma}\left[\mathbb{E}\left(f(z+(1-\delta)I_{t-1})^{2(1-\gamma)}\right) - (\mathbb{E}(f(z+(1-\delta)I_{t-1})^{1-\gamma})^2\right]. \tag{1.10}$$

In the absence of inventories in the previous period, $I_{t-1} = 0$, the variance reduces to

$$\mathrm{Var}_{t-1}(P_t) = P_{t-1}{}^{2\gamma}\mathrm{Var}\left(f(z)^{1-\gamma}\right). \tag{1.11}$$

From (1.10) and (1.11), we can see that the variance is time-varying and, as a result, the final commodity prices derived from our model exhibit conditional heteroskedasticity. In addition, it is worth noting that the variance also depends on the value of γ.

It is interesting to point out that the form of the conditional variance in (1.11) bears strong resemblance to modeling the conditional heteroskedasticity in interest rate models (see, for instance, Brenner, Harjes, and Kroner (1996)). In these models, there is a parameter that allows the volatility of interest rates to depend on the level of the process. Similarly, higher values of the parameter γ in equation (1.11) indicate that the volatility of commodity prices is more sensitive to their past level which generates volatility clustering.

1.3 Model Comparisons Using Simulated Data

In this section we examine the statistical properties of the simulated data from our commodity price model with staggered pricing. In order to assess the qualitative and quantitative implications of our model, we compare it to the basic speculative storage model of DL and the modified version of the speculative model of Ng and Ruge-Murcia (2000). The model of Ng and Ruge-Murcia (2000) extends the DL model by adding serially correlated harvest shocks that follow an MA(1) process, as well as gestation lags, heteroskedastic supply shocks, multi-period forward contracts and convenience yields.

In our simulations, we calibrate the models using the parameter values estimated by Deaton and Laroque (1996) for a set of 12 commodities. These parameters (a, b, δ), presented in Table 1.1, are the same as the parameters used by Ng and Ruge-Murcia (2000). The data are simulated using *iid* harvest shocks or MA(1) harvest shocks with an MA parameter $\rho = 0.8$. We denote our speculative storage model with staggered prices by ADG.

Table 1.2 presents the results for the first-order autocorrelation of the simulated prices from the different models. The first column of Table 1.2 reports the autocorrelations from the actual data used in Deaton and Laroque (1996), the second column shows the results from the basic DL model

($\rho = 0$) and the third column contains the results obtained using DL model with MA(1) shocks ($\rho = 0.8$). The highest autocorrelation for the simulated prices from the DL model is for Maize (0.413 for the basic DL model and 0.644 for the specification with MA(1) harvest shocks). For all other commodities, the serial correlation in the simulated prices is well below the persistence in the actual prices.

The last two columns of Table 1.2 report the results from our model. For all commodities, the autocorrelation coefficients of the simulated prices based on the ADG model are much higher than those of the DL model specifications and are very close to the autocorrelations obtained from actual data. Once we account for staggered pricing, the additional effect of serially correlated harvest shocks is minimal.

Furthermore, Table 1.3 lends additional support to our ADG model with staggered prices. In this table, we compare the autocorrelation coefficients for the model by Ng and Ruge-Murcia (2000) with gestation lags, overlapping contracts and convenience yields to those computed from our ADG model in columns 4 and 5 of Table 1.2.

Ng and Ruge-Murcia (2000) add gestations lags to the DL basic specification in an attempt to reduce the number of periods where the intertemporal price link between periods with and without production is severed. Consequently, this increases the serial correlation in prices. For this purpose, Ng and Ruge-Murcia (2000) assume that there are odd and even periods and that harvest takes place in the even periods. Hence, the random disturbance term of the harvest process has a variance that could differ if the period is odd (σ_1) or even (σ_2). The highest autocorrelations are reached for a value of $\frac{\sigma_2}{\sigma_1} = 1.8$. This model is denoted by GL. The results from the GL specification are reported in column 2 of Table 1.3.

Ng and Ruge-Murcia (2000) also show, in contrast to the earlier literature on storage where contracts are absent and stockholders are free to roll-over their inventories, that a model with overlapping contracts can partially explain the high serial correlation in prices. Column 3, denoted by OV in Table 1.3 reports the corresponding autocorrelation coefficients.

Finally, Ng and Ruge-Murcia (2000) add a convenience yield to the DL model. Since inventory holders might derive convenience from holding inventories, Ng and Ruge-Murcia (2000) introduce both a speculative and a convenience motive for inventory holding. Hence, since the convenience yield partially compensates inventory holders for the expected loss when the basis is below carrying charges, their model with convenience yield generates a smaller number of stock-outs and, as a result, the demand for inventories for convenience purposes strengthens the intertemporal link resulting in a higher persistence of prices. Results for $c = 50$ are reported in column 4 of Table 1.3. The model is denoted by CY.

Overall, the results in Table 1.3 suggest that the different specifications of Ng and Ruge-Murcia (2000) cannot generate autocorrelation coefficients greater than 0.640 and they are below the autocorrelation coefficients from our ADG model and the actual data across all commodities.

1.4 Empirical Application

This section presents new empirical results from estimating the structural parameters of our proposed model using monthly data for four agricultural commodities.

1.4.1 Data

The data set employed in this empirical application consists of prices for four agricultural commodities: sugar, soybeans, soybean oil and wheat. The commodity prices are obtained from the Commodity Research Bureau and are available at daily frequency for the period March 1983 – July 2008. The trading characteristics of these commodities are summarized in Table 1.4.

The spot price is approximated by the price of the nearest futures contract. Monthly commodity price series are constructed from daily data by averaging the daily prices in the corresponding month. The monthly frequency is convenient for studying the persistence and conditional heteroskedasticity in commodity prices. The real commodity prices are obtained by deflating the nominal spot prices by the CPI (seasonally adjusted) index obtained from the Bureau of Labor Statistics (BLS). Each deflated price series is then further normalized by dividing by the sample average. By performing this additional normalization, each series has a historical mean of one which allows us to conduct easier comparisons of the estimated parameters across various price series.

1.4.2 Estimation Method: Simulated Method of Moments

This section provides a brief description of the simulated method of moments (SMM) which is used for estimating the model parameters. The main advantage of SMM lies in its flexibility of the choice of moment conditions that allow us to identify the staggered pricing parameter γ. See Pakes and Pollard (1989), Lee and Ingram (1991) and Duffie and Singleton (1993) for a detailed description of the method and its asymptotic properties, and Michaelides and Ng (2000) for an investigation of its finite-sample properties in the context of the speculative storage model.

The SMM estimator requires repeatedly solving the model for given values of the structural parameters. For this reason, we present some computational details regarding the solution of the model. The function $f(x)$ is approximated using cubic splines and 100 grid of points for x. This function is calculated using an iterative procedure, starting with an initial value $f_0(x) = \max[p(x_t), 0]$. As in DL, the interest rate r is not estimated but it is fixed at 5 percent per annum or 0.41 percent ($r = 1.05^{\frac{1}{12}} - 1 = 0.0041$) per month. In addition, we calibrate the depreciation rate δ and set it equal to 0.04 per month. One reason to calibrate δ is that the SMM estimator tends to over-estimate δ as indicated by Michaelides and Ng (2000). Finally, the harvest shocks z are discretized using a discrete approximation of a standard normal random variable with z taking one of the following 10 values: $(\pm 1.755, \pm 1.045, \pm 0.677, \pm 0.386, \pm 0.126)$, with equal probability of 0.1.

It is worth noting that the prices used for estimation of ADG model parameters represent the prices of intermediate risk neutral speculators, not the final prices that are given by the data set described above. Hence, we first retrieve the prices of intermediate risk neutral speculators from the final prices given by the time series of commodity prices using the equation

$$P_t^* = \left(\frac{P_t}{P_{t-1}{}^{\gamma}} \right)^{\frac{1}{1-\gamma}}. \tag{1.12}$$

Let $\theta = (a, b, \gamma)'$ denotes the vector of structural parameters of the model. Sample paths of commodity prices can be simulated from the assumed structural model for a candidate value of θ. In what follows, we simulate one sample path of prices $\tilde{P}_t(\theta)$ of length TH, where $H = 20$ and T is the sample size of the observed prices P_t. The SMM estimator of θ is then obtained by minimizing the weighted distance (using an optimal weighting matrix) between the moments of the observed data P_t (empirical moments) and simulated data $\tilde{P}_t(\theta)$ (theoretical moments). Let $m(P_t)$ and $m(\tilde{P}_t(\theta))$ denote the set of moments from the observed and simulated data. Then, the

SMM estimator $\hat{\theta}$ is defined as

$$\hat{\theta} = \text{Argmin}_{\theta} D_T(\theta) V_T^{-1} D_T(\theta), \tag{1.13}$$

where

$$D_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} m(P_t) - \frac{1}{TH}\sum_{t=1}^{TH} m(\tilde{P}_t(\theta)),$$

and V_T denotes a consistent estimator of

$$V = \lim_{T\to\infty} \text{Var}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} m(P_t)\right).$$

The vector of moments

$$m(P_t) = [P_t, (P_t - \bar{P})^i, (P_t - \bar{P})(P_{t-1} - \bar{P})]', \text{ for } i = 2, 3, 4, \tag{1.14}$$

is chosen to capture the dynamics and the higher-order unconditional moments of actual commodity prices. The long-run variance V is estimated using the Parzen window

$$w(x) = \begin{cases} 1 - 6x^2 + 6|x|^3 & \text{if } |x| \leq 1/2, \\ 2(1 - |x|^3) & \text{if } 1/2 \leq |x| \leq 1 \end{cases} \tag{1.15}$$

with four lags.

Under some regularity conditions, Lee and Ingram (1991) and Duffie and Singleton (1993) show that the SMM estimator is asymptotically normally distributed

$$\sqrt{T}(\hat{\theta} - \theta_0) \to N(0, \Omega_H), \tag{1.16}$$

where $\Omega_H = \left(1 + \frac{1}{H}\right)\left(\mathbb{E}\left[\frac{\partial m(\tilde{P}_t(\theta_0))}{\partial \theta}\right]' V^{-1} \mathbb{E}\left[\frac{\partial m(\tilde{P}_t(\theta_0))}{\partial \theta}\right]\right)^{-1}$. The derivatives $\partial m / \partial \theta$ are computed numerically and Ω_H is replaced by a consistent estimator in constructing the standard errors of the parameter estimates.

1.4.3 Empirical Results

The estimation results for the ADG model parameters are presented in Table 1.5. The standard errors of the estimated parameters, based on the asymptotic approximation described above, are reported in parentheses below the parameter estimates. The standard errors for the staggered price parameter γ are low for all of the four commodities indicating that γ is well identified and significantly different from zero. The mean of γ for the four commodities is equal to 0.85. The parameter estimates for b satisfy the constraint $b < 0$. For most of the cases, the standard errors of the estimated parameters a and b are relatively low.

In this chapter, we argue that the high persistence and the conditional heteroskedasticity in commodity prices appear to be primarily driven by the staggered price parameter γ. To illustrate this, we simulate 200 price series, each of length of 300 observations. The set of parameters used to conduct the simulations is $(a, b, \delta) = (.7, -3, .04)$ and $r = .0041$. We compute the first-order autocorrelation for each series and then calculate the average over the Monte Carlo replications. We repeat the same exercise for four different values of γ, $\gamma = (0, 0.3, 0.6, 0.9)$. In the first three columns of Table 1.7 we report the first-order autocorrelation for the actual data, ADG and DL models, respectively. Table 1.6 shows that incorporating staggered prices into the speculative storage model does increase the first-order autocorrelation of the prices and makes it comparable to the sample autocorrelation of the actual data. More specifically, as γ increases from $\gamma = 0$ (which represents the case for the DL model) to $\gamma = 0.9$, the first-order autocorrelation increases from 0.6 to 0.9.

To visualize the differences between the two models, Figure 1.1 plots the actual price of soybeans, the simulated prices generated by our ADG model with *iid* harvest shocks and estimated parameters $(a, b, \gamma) = (0.352, -4.787, 0.909)$ and the simulated prices generated by DL model with estimated parameters $(a, b, \delta) = (0.723, -0.394, 0.130)$. It is clear from the graph that our staggered price model generates more persistent data with volatility clustering which is closer to the actual price dynamics of soybeans' prices presented in Figure 1.1. Also, in Figure 1.2 we trace the dynamic responses of the simulated commodity prices following a negative harvest shock. The gradual adjustment of the commodity prices from the ADG model stands in sharp contrast with the stronger but short-lived impact of the harvest shock on commodity prices in the DL model.

Next, in order to reveal the advantages of our ADG model in matching the dynamics in the first two conditional moments of the data, we simulate 200 series of prices, each of length of

300 observations, using the parameters estimated from ADG model (reported in Table 1.5). We repeat the same exercise, using the same values for the parameters a and b but setting $\gamma = 0$, which represents the case for the DL model. We filter the simulated prices from both the DL and ADG models using an AR(1) model and then fit a GARCH(1,1) model to each of the pre-filtered series using the following equations:

$$\begin{aligned} P_t &= a_0 + a_1 P_{t-1} + \varepsilon_t, \\ \varepsilon_t &= \sigma_t z_t, \\ \sigma_t^2 &= \kappa + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2. \end{aligned}$$

Figures 1.3 and 1.4 plot the distribution of the parameter estimates $\hat{\beta}$ and $\hat{\alpha}$ for the ADG and DL models. The figures clearly suggest that the ADG model provides an improvement over the DL model by better capturing the conditional heteroskedasticity. In fact, the medians for $\hat{\beta}$ and $\hat{\alpha}$, generated by ADG model, are much closer to the parameters (denoted by bullets) estimated from actual data. Table 1.7 summarizes the results by reporting the means of the autocorrelations and the GARCH parameters for the ADG and DL models against the statistics from the actual data. Overall, the results lend strong support to the staggered pricing feature of the modified speculative storage model of commodity price determination.

1.5 Conclusion

The main objective of this chapter is to propose a model which is able to reproduce the statistical characteristics of the actual commodity prices. Our modified speculative storage model embeds a staggered price feature into the DL storage model. The staggered pricing rule is incorporated by introducing intermediate good speculators and a final goods bundler. We examine the empirical relevance of the structural modification by comparing our model performance with several models in the literature, namely the DL and the extended DL version of Ng and Ruge-Murcia (2000). Our analysis suggests that the proposed model outperforms the existing models along several dimensions such as matching the serial correlation and GARCH dynamics of the observed commodity prices. We also estimate the vector of structural parameters for the ADG model with uncorrelated harvest shocks using monthly data for four agricultural commodity prices. The results tend to suggest that the staggered price parameter is large and it proves to be instrumental in generating the documented persistence and conditional heteroskedasticity of commodity prices.

While our chapter provides convincing evidence for the importance of integrating staggered pricing features in modeling the dynamics of commodity prices, it only serves as an initial step towards better understanding the source of the gradual adjustment of commodity prices and the role of market power and government intervention in commodity price determination. Explicitly incorporating institutional arrangements, different risk preferences as well as possible existence of financial hedges in some commodity markets might offer a more solid justification for the sluggishness of commodity prices adopted in this study. Finally, developing a full structural model in which staggered pricing is generated endogenously within the model appears to be a promising direction for future research.

1.6 Appendix: Proof of Theorem 1.2.1

First, we state the assumptions for the theorem.

Assumptions: Assume that

A.1 $r+\delta > 0$.

A.2 The harvest shocks z belong to a compact set $\mathbb{Z} = [\underline{z}, \bar{z}]$;

A.3 The function $p^{-1} : (q_0, q_1) \to \mathbb{R}$ is continuous and strictly decreasing such that

$$\lim_{q \to q_0} p^{-1}(q) = +\infty.$$

Furthermore, we have that $\underline{z} \in p^{-1}(p_0, p_1)$ and $p(\underline{z}) \in \mathbb{R}_+ \setminus \{0\}$.

Following Deaton and Laroque (1992), for any function g on the set $\mathbb{X} = [\underline{z}, +\infty)$ we introduce a function G on $\mathbb{Y} = \{(q,x) | x \in \mathbb{X},\, p(x) \leq q < q_1\}$ which has the form

$$G(q,x) = (1-\gamma)\frac{1-\delta}{1+r}\mathbb{E}g(z + (1-\delta)(x - p^{-1}(q))) + \gamma q. \tag{1.17}$$

If $\gamma = 0$, then G is the same as in Deaton and Laroque (1992). Let G^{DL} denote the function when $\gamma = 0$:

$$G^{DL}(q,x) = \frac{1-\delta}{1+r}\mathbb{E}g(z + (1-\delta)(x - p^{-1}(q))).$$

It can be seen that $G = (1-\gamma)G^{DL} + \gamma p$.

Theorem 1.2.1 aims to find a function f such that

$$f(x) = \max\{G(f(x),x), p(x)\},\ \forall x \in \mathbb{X}, \tag{1.18}$$

where we also have $f = g$. To prove the theorem, we use the following lemma.

Lemma 1.6.1 *For a given g, the unique solution $f : \mathbb{X} \to \mathbb{R}$ to* (1.18) *equals f^{DL}, where f^{DL} is the unique solution to the same problem when $\gamma = 0$.*

Proof For each x, $f(x)$ is the solution to the following equation for q

$$\max\{G(q,x)-q, p(x)-q\} = 0. \tag{1.19}$$

It can be seen that

$$G(q,x)-q = (1-\gamma)G^{DL}(q,x)+\gamma q - q = (1-\gamma)(G^{DL}(q,x)-q).$$

Thus, the solution q is a solution to

$$\max\{(1-\gamma)(G^{DL}(q,x)-q), p(x)-q\} = 0. \tag{1.20}$$

But this is equivalent to solving[3]

$$\max\{G^{DL}(q,x)-q, p(x)-q\} = 0, \tag{1.21}$$

which gives the desired result. ■

This lemma shows that for any g, there is a unique f which is the solution to (1.18). Therefore, we can introduce an operator $\mathbb{T}$ and denote f with $\mathbb{T}g$.

Proof of Theorem 1.2.1 From Lemma A.1 it follows that $\mathbb{T}$ is the same as the operator introduced in Deaton and Laroque (1992). It is shown in Deaton and Laroque (1992) that $\mathbb{T}$ is an operator from the set of non-increasing and continuous functions on $\mathbb{X}$ to itself and has a unique fixed point f, i.e., $f = \mathbb{T}f$. It then follows that this unique fixed point is the unique SSREE or SREE. This completes the proof of Theorem 1.2.1. ■

[3]For a positive number θ and two real numbers a, b, we have that $\max\{a,b\} = 0 \Leftrightarrow \max\{\theta a, b\} = 0$.

Table 1.1: Parameter Estimates from the DL Model

Commodity	a	b	δ
Cocoa	0.162	-0.221	0.116
Coffee	0.263	-0.158	0.139
Copper	0.545	-0.326	0.069
Cotton	0.642	-0.312	0.169
Jute	0.572	-0.356	0.096
Maize	0.635	-0.636	0.059
Palm oil	0.461	-0.429	0.058
Rice	0.598	-0.336	0.147
Sugar	0.643	-0.626	0.177
Tea	0.479	-0.211	0.123
Tin	0.256	-0.170	0.148
Wheat	0.723	-0.394	0.130

This table presents the parameters' values for (a, b, δ) estimated by Deaton and Laroque (1996) for a set of 12 commodities.

Table 1.2: First-Order Autocorrelations for the DL and ADG Models

Commodity	Actual	DL	DL	ADG	ADG
		$\rho = 0$	$\rho = 0.8$	$\rho = 0$	$\rho = 0.8$
		$\gamma = 0$	$\gamma = 0$	$\gamma = 0.8$	$\gamma = 0.8$
Cocoa	0.834	0.352	0.609	0.7715	0.8446
Coffee	0.804	0.219	0.576	0.7811	0.8501
Copper	0.838	0.335	0.619	0.8918	0.9074
Cotton	0.884	0.173	0.564	0.8626	0.9053
Jute	0.713	0.289	0.589	0.8817	0.9072
Maize	0.756	0.413	0.644	0.9246	0.9180
Palm oil	0.730	0.397	0.637	0.9079	0.9050
Rice	0.829	0.237	0.579	0.8700	0.9078
Sugar	0.621	0.266	0.583	0.8860	0.9184
Tea	0.778	0.213	0.571	0.8332	0.8893
Tin	0.895	0.238	0.567	0.7547	0.8462
Wheat	0.863	0.250	0.602	0.8834	0.9198

This table presents the results for the first-order autocorrelation of the simulated prices from the different models. The first column reports the autocorrelations from the actual data used in Deaton and Laroque (1996), the second column shows the results from the basic DL model ($\rho = 0$) and the third column contains the results obtained using DL model with MA(1) shocks ($\rho = 0.8$). The last two columns of this table report the results from our ADG model.

Table 1.3: First-Order Autocorrelations for the Ng and Ruge-Murcia (2000) and ADG models

Commodity	Actual	GL	OV	CY	ADG	ADG
					$\rho = 0$	$\rho = 0.8$
					$\gamma = 0.8$	$\gamma = 0.8$
Cocoa	0.834	0.511	0.462	0.522	0.7715	0.8446
Coffee	0.804	0.433	0.385	0.530	0.7811	0.8501
Copper	0.838	0.526	0.394	0.608	0.8918	0.9074
Cotton	0.884	0.365	0.337	0.473	0.8626	0.9053
Jute	0.713	0.486	0.365	0.545	0.8817	0.9072
Maize	0.756	0.620	0.418	0.623	0.9246	0.9180
Palm oil	0.730	0.640	0.438	0.625	0.9079	0.9050
Rice	0.829	0.398	0.334	0.475	0.8700	0.9078
Sugar	0.621	0.427	0.370	0.424	0.8860	0.9184
Tea	0.778	0.428	0.302	0.509	0.8332	0.8893
Tin	0.895	0.428	0.355	0.472	0.7547	0.8462
Wheat	0.863	0.411	0.368	0.505	0.8834	0.9198

This table compares the autocorrelation coefficients for the model by Ng and Ruge-Murcia (2000) with gestation lags, overlapping contracts and convenience yields to those computed from our ADG model in columns 4 and 5 of table 1.2.

Table 1.4: Description of Commodity Price Data

Description	Exchange	Contract size	Contract month
Foodstuffs			
SB : Sugar No.11/World raw	NYBOT	112,000 lbs.	H,K,N,V
Grains and Oilseeds			
S : Soybean/No.1 Yellow	CBOT	5,000 bu.	F,H,K,N,Q,U,X
BO : Soybean Oil/Crude	CBOT	60,000 lb.	F,H,K,N,Q,U,V,Z
W : Wheat/No.2 Soft red	CBOT	5,000 bu.	H,K,N,U,Z

This table provides a brief description about each commodity. The first column presents the symbol description and the second one lists the futures exchange where the commodity is traded. In this table, CBOT refers to Chicago Board of Trade, NYBOT: New York Board of Trade. The third column states the contract size and the last column provides the contract months denoted by: F = January, G = February, H = March, J = April, K = May, M = June, N= July, Q = August, U = September, V = October, X = November and Z = December.

Table 1.5: Parameter Estimations of the ADG Model Using SMM

Commodity	a	b	γ
W	0.4227	-4.6606	0.9476
	(0.0102)	(0.2929)	(0.0086)
BO	0.7860	-2.1265	0.7621
	(0.0177)	(0.1354)	(0.0237)
S	0.7209	-2.7562	0.8524
	(0.0454)	(0.3256)	(0.0343)
SB	0.2264	-5.6592	0.9474
	(0.0195)	(0.4351)	(0.0099)

This table displays the estimation results for the ADG model parameters: a, b and γ, using SMM.

Table 1.6: First-Order Autocorrelations for Simulated Price Series

	$\gamma = 0$	$\gamma = 0.3$	$\gamma = 0.6$	$\gamma = 0.9$
Auto. corr.	0.6122	0.7899	0.9172	0.9838

This table shows the first-order autocorrelation for simulated price series. In particular, we simulate 200 price series, each of length of 300 observations. The set of parameters used to conduct the simulations is $(a,b,\delta) = (.7,-3,.04)$ and $r = .0041$. We compute the first-order autocorrelation for each series and then calculate the average over the Monte Carlo replications. We repeat the same exercise for four different values of γ, $\gamma = (0,0.3,0.6,0.9)$.

Table 1.7: First-Order Autocorrelation, and β and α Parameters from a GARCH(1,1) Model

	Auto. corr.			β			α		
Com.	Actual	ADG	DL	Actual	ADG	DL	Actual	ADG	DL
W	0.9648	0.9899	0.6387	0.6977	0.6719	0.4834	0.2283	0.3006	0.5138
BO	0.9679	0.9550	0.5989	0.7903	0.5160	0.4089	0.1473	0.4709	0.5804
S	0.9697	0.9765	0.6180	0.3413	0.5674	0.4476	0.3410	0.4194	0.5483
SB	0.9620	0.9902	0.6680	0.9018	0.6781	0.4852	0.0798	0.2977	0.5126

This table reports the first-order autocorrelation, β and α parameters estimated from GARCH(1,1) model. In the first three columns we report the first-order autocorrelation for the actual data, ADG and DL models, respectively.

Figure 1.1: Actual and Simulated Data for Soybeans Prices

These figures display the actual and simulated prices for Soybeans. Simulated data is from models with staggered pricing (ADG) and without staggered pricing (DL).

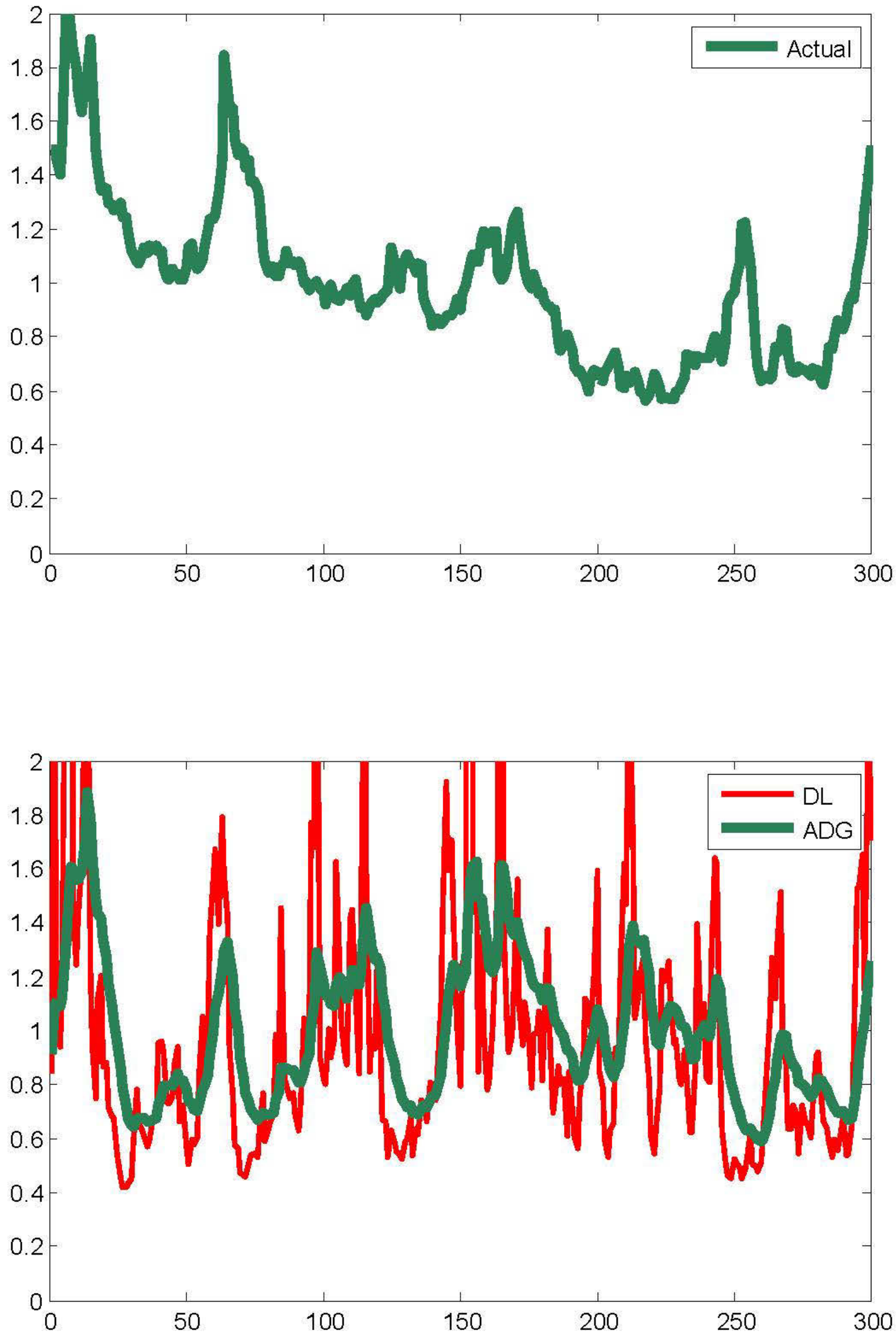


Figure 1.2: Impulse Response Function Based on Simulated Data for Soybean Prices

This figure shows the impulse response function based on simulated data for soybeans prices, following a negative harvest shock for two models, the ADG model and the DL model.

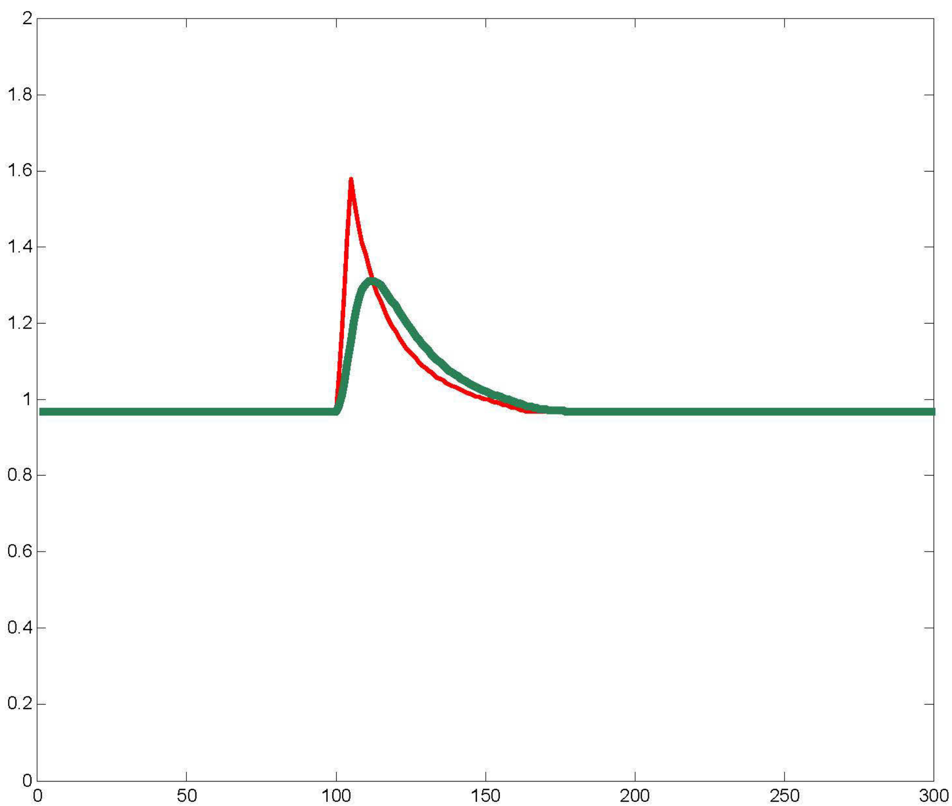


Figure 1.3: Distribution of $\hat{\beta}$

These figures display the distribution of $\hat{\beta}$ for simulated data from models with and without staggered pricing. The dashed line is based on data from the ADG model and the solid line is based on data from the DL model. The estimate of β from actual data is denoted by a circle on the horizontal axis. Simulation is conducted based on a sample of 300 periods, repeated 200 times.

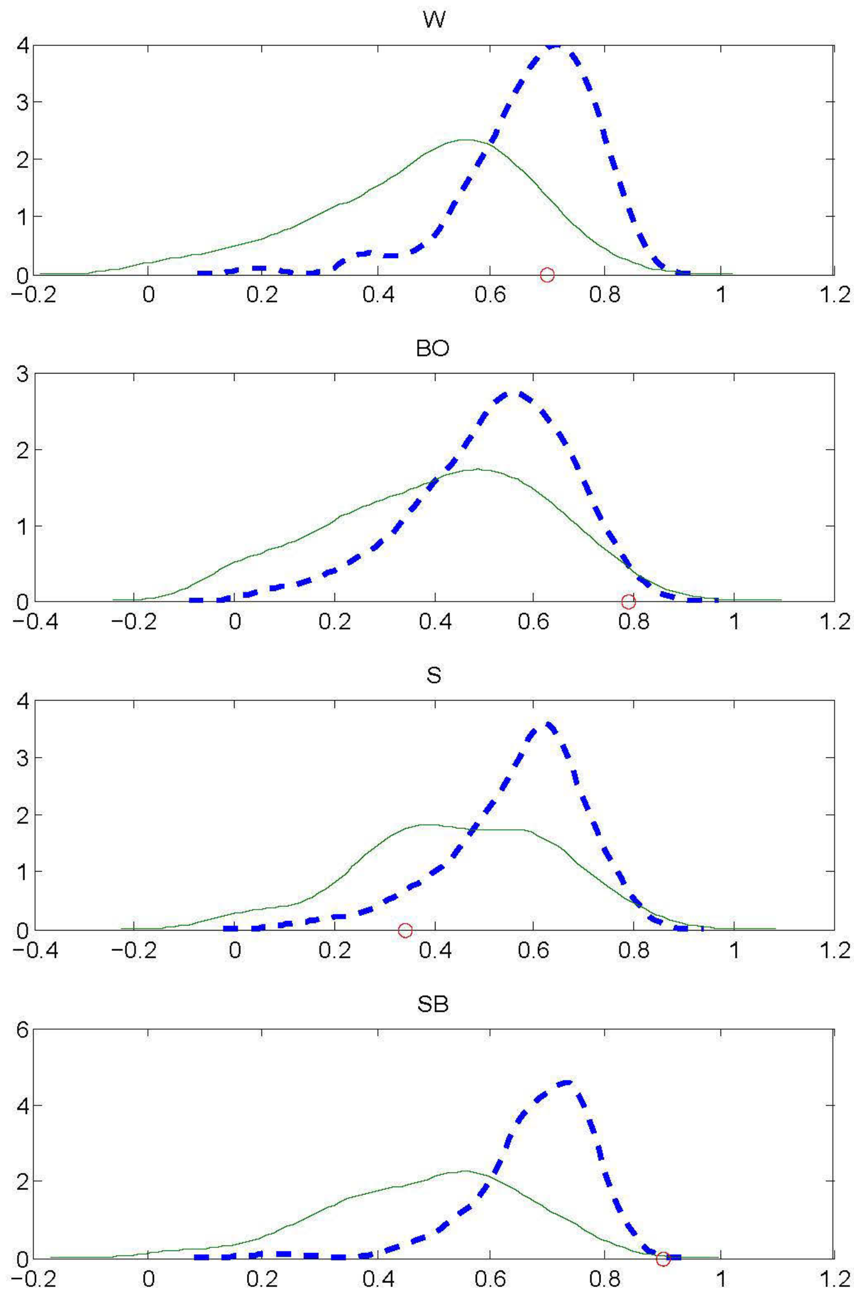
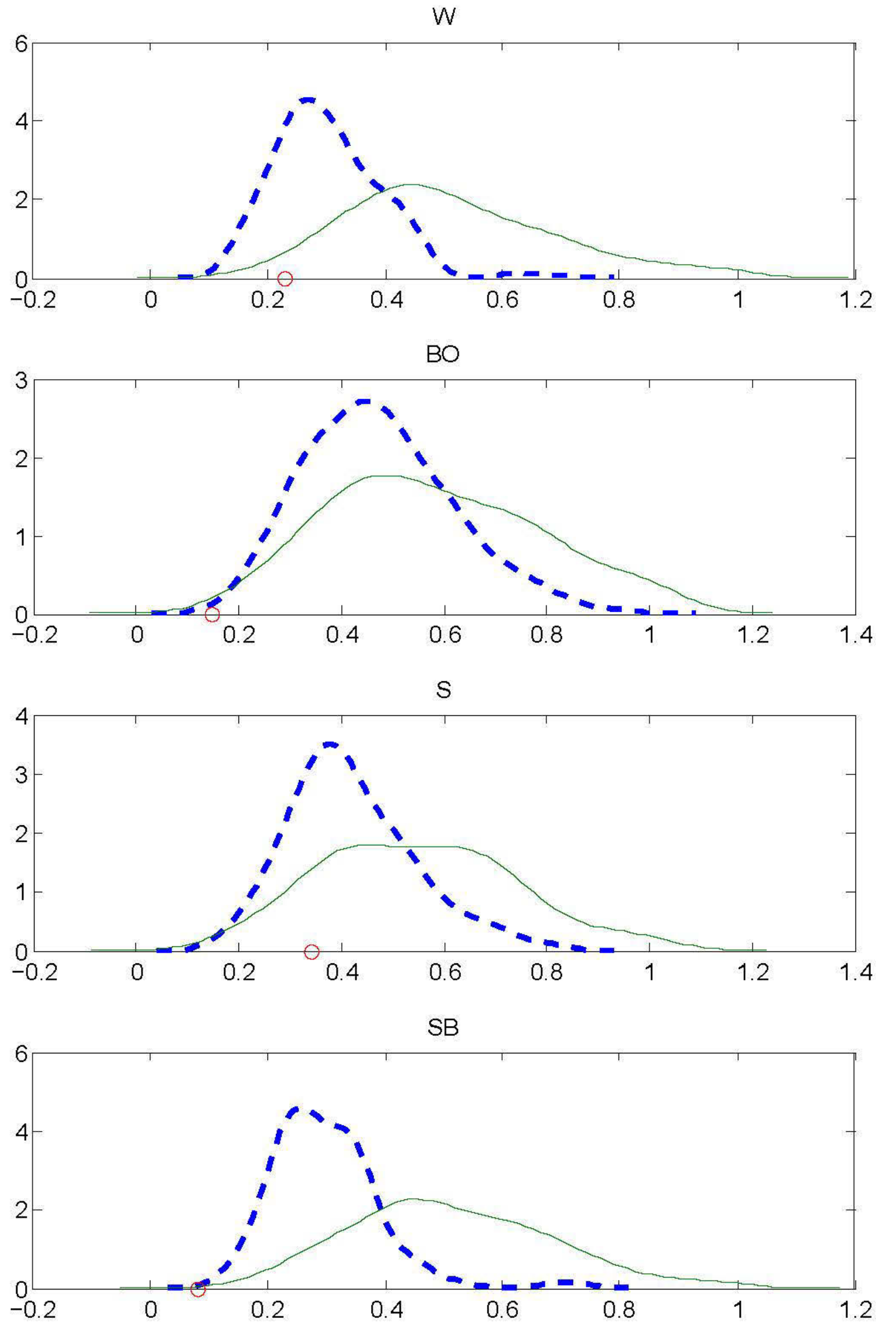


Figure 1.4: Distribution of $\hat{\alpha}$

These figures display the distribution of $\hat{\alpha}$ for simulated data from models with and without staggered pricing. The dashed line is based on data from the ADG model and the solid line is based on data from the DL model. The estimate of α from actual data is denoted by a circle on the horizontal axis. Simulation is conducted based on a sample of 300 periods, repeated 200 times.



Chapter 2

Stochastic Behavior of Commodity Prices: Variance Decomposition

2.1 Introduction

During the past decade, interest in commodities have grown substantially particularly after the remarkable increase in commodity prices and the considerable change in volatility that have accompanied these prices. Specifically, the price of major agricultural commodities such as wheat witnessed high volatility and attracted a lot of interests. Understanding the determinants of commodity prices' fluctuations is therefore of prime interest for both investors and policy makers. The reason is that commodities are considered as investment tools for many investors because they are believed to provide direct exposure to unique factors and bear special hedging characteristics. Moreover, changes in trends of commodity prices from declining to rising pose significant policy challenges for policy makers in resources based economies as well as for the international community.

The main purpose of this chapter is to understand the stochastic behavior of agricultural commodity prices and to determine the main factors that drive the volatility of commodity spot prices. Understanding the stochastic behavior of commodity prices plays a major role in the models for the valuation of commodities related investments.

This chapter permits an analysis through time for the expected percentage change in commodity prices, by relating the expected percentage change in commodity prices to the discounted present value of Treasury Bills interest rates, risk premium and convenience yield.

Several factors models have been proposed in the literature to analyze the stochastic behavior of commodity prices. The first generation of models assumed that uncertainty is summarized in one factor, the spot price of the commodity. Gibson and Schwartz (1990) develop and empirically test a two-factor model for the pricing of financial and real asset depending on the price of oil. Both the spot price of oil and the net convenience yield are assumed to follow a joint stochastic diffusion process. In his seminal work, Schwartz (1997) develops the three-factor model in which interest rates are assumed to be stochastic, in addition to convenience yield which is mean reverting and positively correlated with the spot price. His results suggest that three main factors, namely, driving spot prices, stochastic convenience yield and stochastic interest rates are necessary to capture the dynamics of commodity prices. Hilliard and Reis (1998) extended Schwartz (1997) three-factor model by adding jumps in the spot price to investigate the effect on the pricing of contingent claims. There main findings indicate that spot price process does not have an effect on futures prices but has an impact on options prices. Yan (2002) developed a multi-factor model that incorporates stochastic volatility and simultaneous jumps in the spot price and volatility in addition to the stochastic convenience yield and stochastic interest rate. He finds that stochastic volatility impact on commodity futures price changes randomly over time. Casassus and Collin-Dufresne (2005) develop a three-factor model of commodity spot price which allow for time-varying risk premium. Moreover, the convenience yield is modeled as depending on both the spot price and the risk free interest rate. In addition, they allow the risk premium to be a linear function of the state variables in contrast to previous empirical research such as Schwartz (1997). Empirical results suggest that the contribution of level dependent convenience yield and time-varying risk premium in explaining the mean reversion of commodity spot prices depends on the type of the commodity. Casassus and Collin-Dufresne (2005) also argued that ignoring time variation in risk premium may lead to a very high over-estimation for the value at risk of commodity related investments.

This literature indicates that more factors are needed to explain the stochastic behaviour of commodity prices. However, explaining the variability in commodity prices is a dynamic problem. Hence, it is important to use a framework that allows variation through time for all the factors chosen to explain the changes in spot prices. In contrast to factor models, which use a continuous time framework to analyze the dynamics of commodity prices, we conduct the analysis in a discrete framework. The novelty of the chapter is to apply the dynamic Gordon Growth model technique to the commodity market and conduct the Campbell and Shiller (1988) variance decomposition to determine the main factors that drive the volatility of commodity spot prices. This chapter investigates the ability of this dynamic modeling in explaining the variance of six agricultural commodity prices. Indeed, the model generates variances that are very close to the real ones calculated using actual prices. Stochastic variance is a plausible feature of commodity prices, determining the factors that explain this variance is the main objective of this empirical study. To explore this question, we start first by deriving a simple equation relating the expected percentage change in commodity prices to the current and discounted expected future interest rate, risk premium and net marginal convenience yield. The basic idea consists of combining equations from the theory of storage and theory of normal backwardation to find a linear relationship between expected percentage change in commodity prices and real interest rate, net marginal convenience yield and risk premium. The traditional equation for stock returns associates changes in unexpected stock returns to changes in future dividend growth, future real interest rates and future excess stock returns. As commodities share similar characteristics with stocks and bonds, it is reasonable to think that changes in commodity prices are related to changes in future convenience yield growth as the convenience yield that accrues to the owner of a commodity is directly analogous to the dividend on a stock, future real interest rate and future risk premium.

The model is then tested empirically, towards this end, we apply the dynamic Gordon growth technique of Campbell and Shiller (1988) and Campbell and Ammer (1993) to the commodity market using a set of six agricultural commodities. To implement the dynamic Gordon growth model, we identify first the factors to be included in our model and then we use a vector autoregression (VAR) approach to estimate expectations. Next, we conduct a variance decomposition to indicate how each of the factors namely real interest rate, net marginal convenience yield and risk premium contributes to the volatility of spot commodity prices. Real interest rate and net marginal convenience yield are directly calculated from the data. Risk premium is replaced by a linear function of open interest and yield spread. In addition, factors such as default spread, hedging pressure, realized volatility and Baltic Dry index are added exogenously to the model. These factors have proven to contain powerful information about commodity prices and help predicting the main variables in our model.

Next we present a brief survey of the literature that explains the choice of some of the variables in our model. As noted earlier, risk premium is unknown. We allow risk premium to be time-varying. Casassus and Collin-Dufresne (2005) modeled the risk premium as a linear function of state variables. As a proxy, we replace risk premium in our model by yield spread and trading volume. Our choice for the yield spread is in line with the strand of investigation that believes that commodity prices are driven by common predictors such as the short interest rate, the yield spread, default spread and dividend yield. These market variables are known to predict variation in stock and bonds returns. Fama and French (1989) show that dividend yield and the default spread capture similar variation in expected bond and stock returns. Hence, in our model we only focus on three factors, interest rate, yield spread and default spread. Moreover, Hong and Yogo (2012) argued that open interest is more informative and provide a better signal for higher economic activity than futures prices. They demonstrate that open interest contains information about economic activity and asset prices that are not explained by the traditional supply-demand conditions or by futures prices movements. In previous work, Hong and Yogo (2010) established a relation between open interest in commodity market and returns. They show that open interest growth is a robust predictor for commodity returns even after controlling for other known predictors such as the basis. Open interest is therefore added as an additional factor that helps predicting variation in commodity spot prices.

Furthermore, our model contains a series of factors that have proved successful in explaining and characterizing the nature of returns in stocks and bonds market (default spread, hedging pressure, realized volatility and Baltic Dry Index). Since commodity futures are financial contracts and are used as investments tools in a similar way to bonds and stocks, those factors are added exogenously to our model to help predicting changes in interest rate, risk premia and convenience yield.

Pindyck (2004) investigates the role of volatility in short run commodity market dynamics. He shows that changes in volatility have an impact on market variables such as inventories and convenience yield. Changes in volatility directly affects the marginal convenience yield and the total marginal cost of production. De Roon, Nijman, and Veld (2000) present a model showing that risk premia depends on both own market and cross market hedging pressures. They also provide empirical evidence that hedging pressure significantly affects futures returns and has an impact on the return of the underlying asset. Khan, Khoker, and Simin (2008) allow commodity returns to vary with hedging pressure and supply conditions premium for risk. Estimation results indicate a significant relation between hedging pressure and returns. Bakshi, Panayotov, and Skoulakis (2011) show that the Baltic Dry Index (BDI) growth rate has a predictive power for stock and commodity returns. In addition, the index plays an important role in predicting global economic activity. The BDI growth rate is therefore added to the model as an indicator that can forecast where the economy is heading. Kilian (2009) constructed a freight rate index to measure global economic activity, we use the Baltic Dry Index instead.

The rest of the chapter is structured as follows. Section 2 presents the theories of market basis in the literature. Section 3 introduces the model and explains the variance decomposition technique implemented to the model. Section 4 describes the data used in the empirical application. Section 5 documents our empirical results. Section 6 concludes.

2.2 Theories of Market Basis

The existing literature on commodity futures consists of two alternatives views for the basis. First, the theory of storage which implies that holders of inventories receive an implicit benefit called convenience yield. Second, the theory of normal backwardation which assumes that risk averse investors demand compensation to take position in a future contract. Let $F_{jt,T}$ denotes the futures price at time t for delivery of a commodity j at time T. Let S_t be the spot price at time t. Define the basis for each commodity j for delivery at time $t+T$ as the difference between the spot and the future price, $F_{jt,T} - S_{jT}$.

In the traditional theory of storage advanced by Kaldor (1939) and later by Brennan (1958), the spread between a future and a spot price is equal to the marginal expenditure on physical storage minus the marginal convenience yield of stocks. Since marginal convenience yield that accrues to inventory holders is negatively related to the level of inventories, the marginal convenience yield could then be larger than the marginal cost of physical storage when stocks are low. Hence, the basis could be negative. The theory of storage dictates that the return from purchasing the commodity a time t and selling it for delivery at time T is function of the opportunity cost of forgone interest from having to store the commodity and the net of insurance and storage cost convenience yield. Let $i_{t,T}$, denotes the nominal interest rate earned between period t and $t+T$ and $CY_{jt,T}$ denotes the net convenience yield. Fama and French (1987) establish the formal relation between the basis and the convenience yield:

$$F_{jt,T} - S_{jt} = S_{jt} i_{t,T} - CY_{jt,T}. \tag{2.1}$$

According to Pindyck (1993) the present value model explains changes in asset prices in terms of changes in expected future payoffs. For a storable commodity, convenience yield represents this payoff. Hence, convenience yield plays the role of dividends that help predicting unexpected spot prices movements.

Whereas the theory of storage is about spot price movements, the theory of normal backwardation is about futures price movements. The theory of normal backwardation, Keynes (1930) and Hicks (1939) presumes that commodity producers and inventory holders protect themselves against future price movements by taking short positions in the futures market. Speculators who are taking the counter position to hedgers, would demand a risk premium to take position in a future contract. When the risk premium is positive, the market is in backwardation, current futures prices are discounted relatively to expected future spot prices at maturity. The basis which picks up transitory shocks to the future prices, can be written in terms of two components: risk premium component $\Psi_{jt,T} \equiv E_t S_{jt+T} - F_{jt,T}$ and expected price change component $E_t S_{jt+T} - S_{jt}$:

$$F_{jt,T} - S_{jt} = E_t S_{jt+T} - S_{jt} - \Psi_{jt,T}. \tag{2.2}$$

Next, by combining both equations (2.1) and (2.2), we present a simple linear equation that integrates the theory of storage and the theory of normal backwardation. Together (2.1) and (2.2) imply:

$$E_t S_{jt+T} - S_{jt} = S_{jt} i_{t,T} - CY_{jt,T} + \Psi_{jt,T}. \tag{2.3}$$

Divide both sides of equation (2.3) by S_{jt}, and let:

$cy_{jt,T} = CY_{jt,T}/S_{jt}, \psi_{jt,T} = \Psi_{jt,T}/S_{jt}$, and $E_t \Delta^T s_{jt+T} = (E_t S_{jt+T} - S_{jt})/S_{jt}$. Then, we can write:

$$E_t \Delta^T s_{jt+T} = i_{t,T} + \psi_{jt,T} - cy_{jt,T}. \tag{2.4}$$

While equation (2.1) presents a definition for the convenience yield and equation (2.2) defines the risk premium, equation (2.4) decomposes the expected percentage change in commodity prices into three components:

- Interest rate or opportunity cost of buying and holding inventories, $i_{t,T}$
- Risk premium component, $\psi_{jt,T}$
- Expected marginal convenience yield component, $cy_{jt,T}$

Fama and French (1988) explain the intuition for the negative relation between the convenience yield and the expected percentage change in commodity prices. Specifically, they demonstrate that current spot prices will increase by more than expected spot prices in response to a permanent positive demand shock in commodities. They show that the relative impact of the permanent demand shock on both current and expected spot prices varies with the level of inventories. At low inventory levels, the shock will have a larger impact on spot prices. Current spot prices rise by more than expected spot prices as supply and demand conditions change to meet the increase in demand which partially offset the impact of the shock on expected spot price. Hence, when inventories are low, convenience yield rises faster and lower expected spot prices are anticipated.

2.3 The Model

2.3.1 Model Framework

According to Pindyck (1993) the present value model explains changes in asset prices in terms of changes in expected future payoffs (ψ_{t+i}) and discount rate (γ). For a storable commodity the payoff is represented by the convenience yield. The convenience yield represents the benefits that accrues to inventory holders from holding a physical stock of the commodity.

$$P_t = \gamma \sum_{i=0}^{\infty} \gamma^i E_t \psi_{t+i}. \tag{2.5}$$

In our work, equation (2.4) presents a linear relation between the expected change in prices and three factors that represent future payoffs. Following Pindyck (1993), we can say that the percentage change in price is approximately equal the sum of current and discounted expected future interest rate (i_t), risk premium (ψ_{jt}) and net marginal convenience yield (cy_{jt}).

Hence, the expected percentage change in price is written as:

$$\Delta s_t = E_t[\sum_{i=0}^{\infty} \rho^i i_{t+1+i} + \sum_{i=0}^{\infty} \rho^i \psi_{jt+1+i} - \sum_{i=0}^{\infty} \rho^i cy_{jt+1+i}], \tag{2.6}$$

where ρ represents a constant discount factor.

Campbell and Shiller (1988) introduced the dividend ratio model or dynamic Gordon model. They derived a linear equation for the model using a log linearization approximation. Their model provides an analysis through time for the dividend price ratio. Specifically, the log of the dividend price ratio (δ_t) can be written as the expected discounted value of all future rates (r_{t+i}) and dividends growth rates (Δd_{t+i}), discounted at a constant discount rate (ρ) plus a constant c.

$$\delta_t \approx E_t \sum_{i=0}^{\infty} \rho^i (r_{t+i} - \Delta d_{t+i}) + c. \tag{2.7}$$

Equation (2.6) bears strong resemblance to equation (2.7). Equation (2.6) explains the expected percentage change in price as the expected discounted value of all future payoffs. Our objective is to implement the Campbell and Shiller (1988) variance decomposition technique to our model.

Results for the autocorrelation coefficients of the 4 times series used in our empirical application show that interest rate, yield spread, open interest and convenience yield have an autocorrelation coefficient very close to 1, ranging from 0.9235 to 0.9735 as indicated in Table 2.2. Hence, the time series used in our empirical work are highly persistent. A standard problem in time series econometrics is the derivation of appropriate inferences when working with autocorrelated data. Most applied work in time series relies on the use of heteroskedasticity and autorcorrelation consistent (HAC) standard errors to make inference about the sample. These corrections perform poorly in small samples where autocorrelation is very high.

To avoid the problem of having non-standard inference, we take the percentage change for each of the variables included in our model. Table 2.3 indicates the values for autocorrelation for the 4 time series used, we can see that the autocorrelation coefficients drop dramatically.

Equation (2.6) becomes:

$$\Delta s_t = E_t[\sum_{i=0}^{\infty} \rho^i \Delta i_{t+1+i} + \sum_{i=0}^{\infty} \rho^i \Delta \psi_{jt+1+i} - \sum_{i=0}^{\infty} \rho^i \Delta cy_{jt+1+i}], \tag{2.8}$$

where, Δi_{t+1+j} denotes the percentage change in interest rate, $\Delta\psi_{jt+1+i}$ denotes the percentage change in risk premium and Δcy_{jt+1+i} denotes the percentage change in convenience yield. ρ is a discount factor. According to Pindyck (1993) for most commodities ρ is a number very close but less than one. We set $\rho = 0.9962$ in our empirical work as in Campbell and Ammer (1993) paper. Even though the parameter ρ is different across commodities, in order to check for robustness we also examined the results of our empirical work for $\rho = 0.9, 0.95$ and $\rho = 0.975$, results were very close to the ones indicated in this chapter.

2.3.2 Variance Decomposition Technique

In this section we briefly discuss our empirical methodology. Our goal is to implement the Campbell and Shiller (1988) technique to our model and conduct a variance decomposition for the expected percentage change in spot commodity prices. Following, Campbell and Shiller (1988) and Campbell and Ammer (1993), we use a first-order vector autoregressive VAR framework to compute the estimates for the expected future change in interest rates, risk premium and convenience yield.

In particular, it is assumed that expected percentage change in commodity prices may be calculated on the basis of the present value estimates of the forecasted values of those three factors. The VAR approach starts by defining a vector of state variables that helps to measure or forecast the percentage change in spot commodity prices. Denoting the state vector by Z_t we have,

$$Z_t = AZ_{t-1} + \varepsilon_t$$ [1]

We can therefore write,

$$Z_t = (\Delta i_t, \Delta\psi_t, \Delta cy_t, \Delta x_t')'. \tag{2.9}$$

Interest rate and convenience yield can be directly calculated from the data. The risk premium cannot be calculated from our data, we use instead yield spread and open interest as proxies to help forecasting risk premium. As noted earlier, the risk premium is replaced by a linear function of open interest and yield spread.

The vector x_t is a column vector that contains other variables that turn to be useful when predicting the percentage change in commodity prices. We include these variables to help forecast

[1] For simplicity, we assume that we have a first order VAR.

short interest rate, risk premium and net marginal convenience yield. Specifically, x_t is a vector of four predictors including: hedging pressure, realized volatility, default spread and Baltic Dry Index. Following Campbell, Davis, Gallin, and Martin (2009) we add the growth rate of these exogenous variables.

Hence, Equation(2.9) is written as:

$$Z_t = (\Delta i_t, \Delta ys_t, \Delta oi_t \Delta cy_t, \Delta x_t')',$$

where, $\Delta ys_t, \Delta oi_t$ denote the percentage change in yield spread and open interest respectively. Note that since risk premium was divided by spot price in our model, we also divide yield spread and open interest by spot price in our empirical application.

Given an estimate of A, which we denote by $\hat{A}$, the estimates of the present values of the growth rates of short interest rate, yield spread, open interest and net marginal convenience yield are the first 4 elements of the matrix:

$$\hat{A}(I - \rho \hat{A})^{-1} Z_t,$$

where I is an 8 by 8 identity matrix.

These four elements are denoted by $\Delta \hat{I}_t$, $\Delta \hat{ys}_t$, $\Delta \hat{oi}_t$ and $\Delta \hat{cy}_t$ respectively,

$$\Delta \hat{I}_t = E_t \sum_{i=0}^{\infty} \rho^i \Delta i_{t+1+i},$$

$$\Delta \hat{ys}_t = E_t \sum_{i=0}^{\infty} \rho^i \Delta ys_{t+1+i},$$

$$\Delta \hat{oi}_t = E_t \sum_{i=0}^{\infty} \rho^i \Delta oi_{t+1+i},$$

$$\Delta \hat{cy}_t = E_t \sum_{i=0}^{\infty} \rho^i \Delta cy_{t+1+i}.$$

The predicted percentage change in price is defined by:

$$\Delta \hat{s}_t = \Delta \hat{I}_t + \Delta \hat{ys}_t + \Delta \hat{oi}_t - \Delta \hat{cy}_t, \tag{2.10}$$

where, small letters are used to represent the variables divided by the spot price.

The present value model of commodity pricing of Pindyck (1993) indicates that changes in prices can be entirely outlined by changes in expected future payoffs and a discount rate. Furthermore, the dynamic Gordon growth model identity states that stock returns can be described by changes in short interest rates and dividend growth rate up to a certain constant of linearization. Similarly our model indicates that the expected percentage change in price can be fully explained by the four main factors included in the model up to a certain discount rate ρ. Since the predicted percentage change in price will differ from the actual percentage change in prices at each point in time, the actual percentage change in prices can be written as:

$$\Delta s_t = \hat{\Delta s_t} - e_t,$$

where e_t represents the forecast discrepancy.

Following the literature, the $-cy_t$ is treated as a residual. Hence, the present value of future growth rate in convenience yield is treated as residual. We denote $(\hat{\Delta cy_t} + e_t)$ by ε_t.

Equation (2.10) becomes:

$$\Delta s_t = \hat{\Delta I_t} + \hat{\Delta ys_t} + \hat{\Delta oi_t} - \varepsilon_t. \tag{2.11}$$

We can therefore decompose the variance of the expected percentage change in price as :

$$\begin{aligned} var(\Delta s_t) = var(\hat{\Delta I_t}) + var(\hat{\Delta ys_t}) + var(\hat{\Delta oi_t}) + var(\hat{\varepsilon_t}) + 2cov(\hat{\Delta I_t}, \hat{\Delta ys_t}) + 2cov(\hat{\Delta I_t}, \hat{\Delta oi_t}) \\ +2cov(\hat{\Delta ys_t}, \hat{\Delta oi_t}) - 2cov(\hat{\Delta I_t}, \hat{\varepsilon_t}) - 2cov(\hat{\Delta ys_t}, \hat{\varepsilon_t}) - 2cov(\hat{\Delta oi_t}, \hat{\varepsilon_t}). \end{aligned} \tag{2.12}$$

Simply, the variance of the expected percentage change in spot commodity prices is the sum of different variances and covariances. All the elements in this decomposition related to forecast discrepancy ε_t will be interpreted in terms of the contribution of the convenience yield to the variance of percentage change in commodity prices.

2.4 Data

2.4.1 Description

This section presents a description for the data that we will be using in the empirical work. The data set used in this study consists of monthly observations for commodity prices, nominal interest rate, convenience yield, yield spread, trading volume, hedging pressure, realized volatility, default spread and data for the Baltic Dry Index.

The monthly commodity prices for the six agricultural commodities are obtained from the Commodity Research Bureau and are available at a monthly frequency for the period between March 1990 and July 2008. The choice of the sample time period is dictated by the availability of the data for all the factors used in the empirical work. The data set consists of series of spot and futures prices for 6 agricultural commodities from 2 groups: grains and oilseeds (corn, oats, soybeans, soybean oil and wheat) and industrials (cotton). The trading characteristics of these commodities are summarized in Table 2.1. The spot prices series are approximated by the nearest futures contract and the futures prices are approximated by the next to the nearest futures contract. Following Svedberg and Tilton (2006) we deflate all nominal prices using Consumer Price Index. The real commodity prices series for spot and futures prices are therefore obtained by deflating the nominal spot and futures prices by the CPI (seasonally adjusted) index obtained from the Bureau of Labor Statistics (BLS).[2]

For nominal interest rates, we used data on monthly average yield on the one month Treasury Bill for US Treasury securities. As the one month treasury market yield on U.S. Treasury securities at constant maturity first appeared in July 2001. Data on monthly average yield for the one month US Treasury is retrieved from 2 sources. After 2001 data is obtained from the Federal Reserve Bank website.[3] As for data before 2001, values for earlier years are calculated using interpolation for the time period from March 1990 to July 2001 and are downloaded from the website of mortgage-x.[4]

[2] The data for CPI can be downloaded from http://www.bls.gov

[3] The data for one month nominal interest rate can be downloaded from http://www.federalreserve.gov/releases/h15/data.html

[4] The data for one month nominal interest rate before 2001 can be downloaded from http://www.mortgage-x.com/general/indexes

Following Pindyck (1993), for commodities traded on futures market, the percentage net convenience yield can be calculated using data on commodity price futures. Hence, we compute the percentage net convenience yield for commodity j as the following:

$$cy_{jt,T} = \frac{(1+i_{t,T})S_{jt} - F_{jt,T}}{S_{jt}}.$$

The data for the trading volume for the 6 commodities are downloaded from the US commodity futures trading commission website.[5] Yield spread is calculated as the difference between Moody's Aaa corporate yield and the short rate. Moody's monthly yield on seasoned corporate bonds, all industries are downloaded from the Federal Reserve Bank website.

Default spread is compiled as the difference between Moody' s BAA and Moody' s AAA corporate bond yields.

Hedging pressure is defined as the difference between the number of contracts categorized as commercial long and the number of contracts categorized as commercial short divided by the total.

$$HP = \frac{CL - CS}{CL + CS}.$$

where, CL denotes commercial long positions and CS denotes commercial short positions. The data for CL and CS is obtained from the CFTC website[6]. The CFTC classifies traders according to the size of their positions into reportable and non reportable traders. A trader's reported future position is considered as commercial if the position is used for hedging purposes as indicated by the CFTC regulations. A reportable trader, who is not engaged in business activities hedged by the use of the futures and options market, as stated by the CFTC will be categorized as non commercial.

The Baltic Dry Index monthly data is retrieved from Bloomberg. The realized volatility for spot prices is calculated using futures daily returns. Daily data for futures commodity prices are retrieved from the Commodity Research Bureau. Specifically, we collect data on all daily nearest contracts, spot returns are then computed as the difference in the logarithms of the prices. Following Amaya, Christoffersen, Jacobs, and Vasquez (2011), realized volatility is calculated using the following formula:

[5] http://www.cftc.gov

[6] Hedging pressure data as well as trading volume data are retrieved from CFTC website http://www.cftc.gov

$$Rvol = \sqrt{\frac{1}{n}\sum_{t=1}^{n} r^2}.$$

where r^2 denotes the sum of monthly square returns and n denotes the number of returns per month.

Finally, all the data series used in the empirical work are normalized by dividing each series by its corresponding sample mean. As a result, we will have a historical mean of one which allow us to conduct easier comparisons across different series.

2.4.2 Summary Statistics

This section presents the summary statistics for the variables used in our model. Tables 2.2 and 2.7 report the autocorrelation for all the series. The first order autocorrelation for all the series except hedging pressure and realized volatility are above 0.9, which is almost the same as the commodity spot prices. Realized volatility has the lowest first order autocorrelation ranging from 0.23 to 0.43 which suggests that it is a predictor variable operating at a higher frequency than the other variables included in the model.

The mean of the variables is not reported as all the time series have been normalized and have a historical mean of 1.

Tables 2.4 and 2.7 show the standard deviations of the variables used in our empirical work and the spot prices, as indicated, the standard deviation for most of the variables is around 0.4 which is higher than the spot prices standard deviation ranging between 0.2425 and 0.3243.

Tables 2.5 and 2.6 indicate that spot commodity prices are positively skewed and present fat tails with positive excess kurtosis. This is in line with the general behaviour of commodity prices that tend to crash more than we would expect them if price changes were random. Moreover they always tend to crash up. Short interest rate and net marginal convenience yield present a negative skewness. For half of the commodities (C, CT and S) hedging pressure also has a negative skewness. The remaining variables, open interest, yield spread, default spread, realized volatility and Baltic Dry Index are positively skewed as shown in Table 2.5 and Table 2.7. The excess kurtosis values for all the variables are positive as indicated by Tables 2.6 and 2.7, which means that there is a high probability for extreme values.

2.5 Results

2.5.1 Graphical Analysis

Before starting the explanation on the main sources of the commodity spot prices' volatility, we summarize the ability of our model outlined by equations (2.11) and (2.12) to explain the real changes in spot prices graphically over the sample period, 1990-2008. Figures 2.1 to 2.6, compare the real change in spot prices to the estimated percentage change in prices obtained using the factors included in our model. We plot the actual percentage change in spot prices for the 6 agricultural commodities in Figures 2.1 to 2.6, as a solid line and the estimated percentage change in spot prices $\hat{\Delta s_t}$, computed using (2.10) for the same set of commodities, as a dashed line.

The estimated percentage change in prices computed from our model captures the dynamics of the actual changes in spot commodity prices, the peaks and the troughs are almost the same. Both series demonstrate high volatility and conditional heteroskedasticity which are the main characteristics of commodity prices. However, we should note that for 3 of the commodities namely wheat, soybeans and soybean oil we can see that after 2006 the model predicts a decline in the prices while the actual prices rose. As indicated by Irwin and Good (2009), prices of corn, soybeans, and wheat started increasing in the fall of 2006 and then surged to new record highs in 2008. Price instability was mainly due to the large demand growth from developing nations, U.S. monetary policy, diversion of row crops to biofuel production and weather-related negative shocks to production.

2.5.2 Sources of the Variability in Spot Commodity Prices

This section explains our variance decomposition results which depend on the long-run dynamics of the variables as indicated by the infinite discounted sum of powers of our endogenous variables. Our objective is to use equation (2.12) to estimate the relative importance of the endogenous variables included in our model for the historical behavior of spot commodity prices. In this section we report the six numbers on the right hand side of equation (2.12).

We start by comparing the variance of the actual percentage change in spot price calculated using Δs_t and the variance estimate form our model calculated using equation (2.10). Columns 1 and 2 in Table 2.8 show that the overall variability of the actual percentage change in prices of the 6 agricultural commodities is very similar to the estimated one. The third column of Table 2.8

represents which percentage of the actual variance our estimated variance captures. The higher the percentage, the better the variables chosen in our model explain the variability in spot prices. Columns 3 of Table 2.8, indicates that the model forecasts a large share of the monthly variance of the percentage change in spot commodity prices. The variance of the estimated percentage change in spot commodity prices is at least 75 percent as large as the actual variance. The percentage of the actual variance captured by our model varies between 75 percent for oats and 95 percent for soybean oil.

Next we display the results for the variance decompositions for the 6 agricultural commodities considered in our empirical application. Table 2.9 and Table 2.10 present the shares attributable to each component of the variance of the percentage change in spot commodity prices, calculated using equation (2.12). Columns 1 to 4 in Table 2.9 report the shares of the variances of each of the endogenous variables and columns 1 to 6 of Table 2.10 refer to the shares of the variances attributable to the covariances.

Our empirical results indicate that variation in expected percentage change in yield spread, $\Delta\hat{ys}_t$, plays the largest role in accounting for the variation in the percentage change in spot commodity prices for the 6 agricultural commodities over the period between 1990 and 2008. Across all our commodities, variation in percentage change in yield spread accounts for more than 50 percent of the variation in percentage change in spot commodity prices, which represent a substantial share. Variation in expected percentage change in interest rate, $\Delta\hat{I}_t$, expected percentage change in open interest and net marginal convenience yield, $\Delta\hat{oi}_t$ and $\hat{\varepsilon}_t$, account for similar shares of the volatility of the percentage change in spot commodity prices.

The covariances shares are reported in Table 2.10. We find a negative correlation between interest rate and yield spread, interest rate and open interest, convenience yield and yield spread (positive contribution) as well as between convenience yield and open interest (positive contribution), except for soybeans. We find a positive correlation between convenience yield and interest rate (negative contribution) as well as between yield spread and open interest except for cotton and oats. Overall the covariances among different endogenous variables considered in our model serve to dampen the fluctuations in the percentage change in the spot price. Column 7 in Table 2.10, shows that the sum of the covariances of the different factors for all six commodities is negative which indicates that the correlations between the four main components dampened the variability in the spot commodity prices, reducing the variance of the spot prices by almost 15 percent for 4 of the commodities, namely, corn, cotton, soybeans and wheat. Hence, the covariance share in the total variation of spot commodity prices is fairly small.

Our empirical results suggest that most of the variation in the percentage change in spot commodity prices is attributable to variation in yield spread, convenience yield open interest and interest rate. These results are quite consistent with the fact that commodities share common characteristics with stocks and bonds and therefore commodity price variations are driven by common known predictors such as yield spread and interest rate.

These results extend the literature in several ways:

- Pindyck (1993) showed that commodity prices and convenience yield tend to move in the same direction. Looking at a set of four commodities, copper, heating oil, lumber and gold over a sample period ranging from 1971 to 1989, he finds that the present value model provides a compact explanation for changes in commodity prices and that these changes are due to changes in expected future convenience yield. Our results for the variance decomposition of the percentage change in spot commodity prices for 6 agricultural commodities for the period between 1990 and 2008 are in line with Pindyck (1993) and indicate that the net marginal convenience yield explains variations in spot commodity prices.

- Hong and Yogo (2009) investigate the determinants of aggregate commodity returns. Using data on 34 commodities for the period 1965-2008, they find that common predictors of stocks and bonds returns such as the short interest rate and the yield spread predict commodity returns. The variance decomposition conducted in our empirical exercise attributes most of the spot price variations to changes in the yield spread. The interest rate was also found to predict variations in spot commodity prices.

- Hong and Yogo (2012) argued that open interest contains information about asset prices that are not explained by the fundamentals (supply and demand shocks). Particularly, looking at a large dataset of 30 commodities over the period between 1965 and 2008, their model implies that movements in open interest predict commodity returns and are highly correlated with movements in both futures and spot prices in commodity markets. Our variance decomposition outcomes establish new findings on the relation between the variation in open interest and the volatility of spot commodity prices. The growth rate of open interest was empirically found to help predict changes in commodity prices.

2.6 Conclusion

The main objective of this chapter is to examine how the estimates of changes to expectations of the factors included in our model are related to the overall volatility of the percentage change in the spot price of six agricultural commodities, BO, C, CT, O, S and W. Using the theory of storage and theory of normal backwardation we write an expression for the percentage change in spot commodity prices in terms of short interest rate, risk premium and net marginal convenience yield. The risk premium is proxied by a linear expression of yield spread and open interest. Moreover, other variables namely, hedging pressure, realized volatility, default spread and Baltic Dry Index, are added exogenously to the the model to help predicting the endogenous variables. As part of our empirical application, we apply the Campbell and Shiller (1988) technique to our model and conduct a variance decomposition for the percentage change in spot commodity prices.

Empirical application suggests that our model captures important movements in the spot commodity prices of the 6 agricultural commodities, over the sample period between March 1990 and July 2008, capturing a percentage of at least 75 percent of the actual variance of the percentage change in spot commodity prices. At the same time, the variance decomposition conducted attributes most of the spot price variation to changes in the yield spread. Each of the three remaining variables, changes in interest rate, open interest and convenience yield were found to make relatively close contributions to the variation of the percentage change in spot commodity prices. Covariances among the 4 main factors in our model were found to dampen the variation in spot prices.

Our findings support the literature where the variables included in our model prove to play an important role in explaining changes in commodity prices. Moreover, these results are certainly consistent with the strand of literature which assumes that commodities share similar characteristics with stocks and bonds and can therefore be driven by common known factors such as yield spread, short interest rate and convenience yield which is directly analogous to dividends on a stock.

The variance decomposition conducted reveals important characteristics about the fundamental sources of the variability in spot commodity prices. Aside from providing direct insights on the main sources of variability, our work points to a new direction for forecasting models, where if factors such as yield spread and open interest are added, they may improve significantly the forecast ability.

A last important point: information about commodity prices variance structures will enhance risk management strategies in agricultural commodities. Factors that have proven to have an impact on the variability in spot commodity prices can then be added when developing simulation models of hedging strategies. Furthermore, the results of the analysis conducted will provide policy makers with a better understanding of the main driving sources of the variability in agricultural commodity prices, enabling them to anticipate the impact of the change in one of the factors on the price of an agricultural commodity.

Table 2.1: Trading Characteristics of Commodity Prices Data

Description	Exchange	Contract size	Contract month
Grains and Oilseeds			
C : Corn/No.2 Yellow	CBOT	5,000 bu.	F,H,K,N,U,X,Z
O : Oats/N0.2 White heavy	CBOT	5,000 bu.	H,K,N,U,Z
S : Soybean/No.1 Yellow	CBOT	5,000 bu.	F,H,K,N,Q,U,X
BO : Soybean Oil/Crude	CBOT	60,000 lb.	F,H,K,N,Q,U,V,Z
W : Wheat/No.2 Soft red	CBOT	5,000 bu.	H,K,N,U,Z
Industrials			
CT : Cotton/1-1/16"	NYBOT	50,000 lbs	H,K,N,V,Z

This table provides a brief description about each commodity trading characteristics. The first column presents the symbol description and the second one lists the futures exchange where the commodity is traded. In this table, CBOT refers to Chicago Board of Trade and NYBOT: New York Board of Trade. The third column reveals the contract size and the last column provides the contract months denoted by: F = January, G = February, H = March, J = April, K = May, M = June, N= July, Q = August, U = September, V = October, X = November and Z = December.

Table 2.2: Autocorrelation for the Variables Included in the Model

Commodity	Spot price	Yield spread	Open interest	Convenience yield	Hedging pressure	Realized volatility
BO	0.9324	0.9674	0.9696	0.9735	0.7532	0.4342
C	0.9300	0.9709	0.9512	0.9734	0.7781	0.3819
CT	0.9547	0.9660	0.9684	0.9716	0.7097	0.2625
O	0.9208	0.9572	0.9675	0.9690	0.6679	0.2378
S	0.9269	0.9672	0.9601	0.9734	0.8849	0.5794
W	0.9440	0.9710	0.9235	0.9706	0.7662	0.2660

The autocorrelation for the short interest rate is the same for all the commodities and it is equal to 0.974. This table reports the autocorrelation for the spot price and the predictor variables, the standardized yield spread, open interest, net marginal convenience yield, hedging pressure and realized volatility.

Table 2.3: Autocorrelation for the Percentage Change of the Variables Included in the Model

Commodity	%Δ spot price	%Δ yield spread	%Δopen interest	%Δ convenience yield
BO	-0.1748	-0.1025	-0.1464	0.1896
C	-0.0022	-0.0786	-0.0669	0.2180
CT	-0.1652	-0.0677	-0.2417	-0.0954
O	-0.0879	-0.0884	-0.0293	0.0051
S	-0.0753	-0.0968	-0.0245	0.1653
W	-0.0487	-0.0852	-0.0119	0.1679

The autocorrelation for the percentage change in short interest rate is the same for all the commodities and it is equal to 0.2302. This table reports the autocorrelation for the percentage change in the predictor variables.

Table 2.4: Standard Deviations for the Variables Included in the Model

Commodity	Spot price	Yield spread	Open interest	Convenience yield	Hedging pressure	Realized volatility
BO	0.2778	0.6162	0.5321	0.4231	1.3472	0.3321
C	0.2785	0.5832	0.4130	0.4290	75.3390	0.4523
CT	0.3243	0.7450	0.9656	0.4406	10.5353	0.4210
O	0.2784	0.5159	0.8232	0.4079	0.4143	0.4678
S	0.2425	0.5717	0.4420	0.4199	1.5693	0.4719
W	0.3078	0.5740	0.3941	0.4217	1.6990	0.4532

This table reports the standard deviations for the spot price and the predictor variables, the standardized yield spread, open interest, net marginal convenience yield, hedging pressure and realized volatility.

Table 2.5: Skewness for the Variables Included in the Model

Commodity	Spot price	Yield spread	Open interest	Convenience yield	Hedging pressure	Realized volatility
BO	1.1166	0.7882	0.3664	-0.1764	0.0358	1.0887
C	1.2787	0.4984	0.3218	-0.1527	-0.1235	2.7050
CT	0.5869	1.1766	1.5463	-0.1456	-0.0469	3.0805
O	1.2401	0.1022	0.9467	-0.2357	0.0987	2.0723
S	0.7037	0.7707	0.2636	-0.1910	-0.1904	2.0536
W	1.3959	0.4668	0.9517	-0.2182	0.3215	4.0597

This table reports the skewness for the spot price and the predictor variables, respectively the standardized yield spread, open interest and net marginal convenience yield, hedging pressure and realized volatility.

Table 2.6: Excess Kurtosis for the Variables Included in the Model

Commodity	Spot price	Yield spread	Open interest	Convenience yield	Hedging pressure	Realized volatility
BO	5.7793	3.1037	1.7935	2.3607	2.3853	4.7302
C	5.1059	2.2587	2.0679	2.3442	2.0776	20.0604
CT	2.6816	4.0349	4.5284	2.4476	2.0256	22.9868
O	5.0086	2.1754	3.0189	2.3761	2.6929	9.6417
S	3.6773	3.1042	1.7144	2.3247	2.3383	9.7081
W	5.1737	2.4792	3.6253	2.4050	2.3676	30.4685

This table reports the excess kurtosis for the spot price and the predictor variables, the standardized yield spread, open interest, net marginal convenience yield, hedging pressure and realized volatility.

Table 2.7: Data Statistics for the Common Variables

	Interest rate	Default spread	Baltic Dry Index
Autocorrelation	0.9740	0.9386	0.9618
Standard deviation	0.4251	0.2604	1.5270
Skewness	-0.1995	0.9078	0.2829
Excess Kurtosis	2.3544	3.0959	4.0834

This table reports the autocorrelation, standard deviation, skewness and excess Kurtosis for short standardized interest rate, default spread and Baltic Dry Index.

Table 2.8: Actual and Estimated Variance of the Percentage Change in Commodity Prices

Comm	Actual var	Estimated var	Percentage captured by the model
BO	0.0053	0.0050	95.23
C	0.0056	0.0043	75.99
CT	0.0073	0.0056	75.89
O	0.0076	0.0057	75.15
S	0.0048	0.0043	88.17
W	0.0062	0.0056	91.41

This table displays the variance of the actual percentage change in spot price calculated using Δs_t and the variance estimate form our model calculated using equation (2.10). Moreover, column 3 indicates which percentage of the actual variance our estimated variance captures.

Table 2.9: Variance Decomposition of the Percentage Change in Spot Price: Variances Shares

Comm	var($\Delta \hat{I}_t$)	var($\Delta \hat{ys}_t$)	var($\Delta \hat{oi}_t$)	var($\hat{\varepsilon}_t$)
BO	0.1603	0.5479	0.1583	0.1502
C	0.1949	0.6041	0.0614	0.2129
CT	0.1385	0.4804	0.1241	0.4248
O	0.1498	0.5992	0.1658	0.2052
S	0.1636	0.7431	0.1296	0.1129
W	0.1835	0.7036	0.0642	0.2017

This table presents the results for the variance decomposition of the percentage change in spot commodity prices for the 6 agricultural commodities. The columns show respectively the shares for var($\Delta \hat{I}_t$), var($\Delta \hat{ys}_t$), var($\Delta \hat{oi}_t$) and var($\Delta \hat{cy}_t$) represented by var($\hat{\varepsilon}_t$) in the overall estimated variance of the percentage change in spot price.

Table 2.10: Variance Decomposition of the Percentage Change in Spot Price: Covariances Shares

Comm	2cov($\Delta\hat{I}_t$,$\Delta\hat{y}s_t$)	2cov($\Delta\hat{I}_t$,$\Delta\hat{o}i_t$)	2cov($\Delta\hat{y}s_t$,$\Delta\hat{o}i_t$)	-2cov($\Delta\hat{I}_t$,$\hat{\varepsilon}_t$)	-2cov($\Delta\hat{y}s_t$,$\hat{\varepsilon}_t$)	-2cov($\Delta\hat{o}i_t$,$\hat{\varepsilon}_t$)	Sum of Cov
BO	-0.2879	-0.0308	0.2557	-0.3078	0.3085	0.0456	-0.0168
C	-0.4482	-0.2095	0.2913	-0.4052	0.4820	0.2164	-0.0733
CT	-0.3474	-0.0745	-0.0159	-0.3479	0.4620	0.1561	-0.1677
O	-0.3359	-0.0274	-0.0316	-0.2937	0.3868	0.1817	-0.12
S	-0.3481	0.1074	0.0674	-0.2622	0.3298	-0.0434	-0.1493
W	-0.4970	-0.1007	0.1760	-0.3701	0.5161	0.1228	-0.153

This table indicates the empirical results for the shares of the covariances of the estimated factors in the model in the overall estimated variance of the percentage change in spot price.

Figure 2.1: Actual and Forecasted Percentage Change in Price: BO

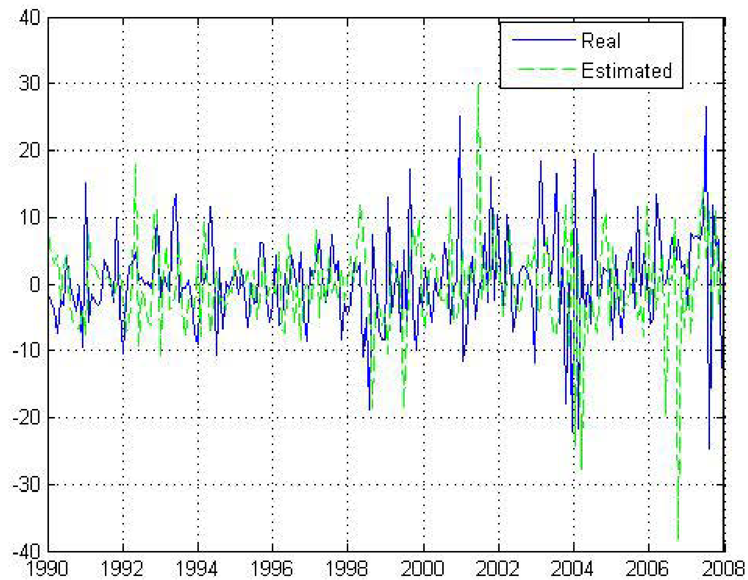


Figure 2.2: Actual and Forecasted Percentage Change in Price: C

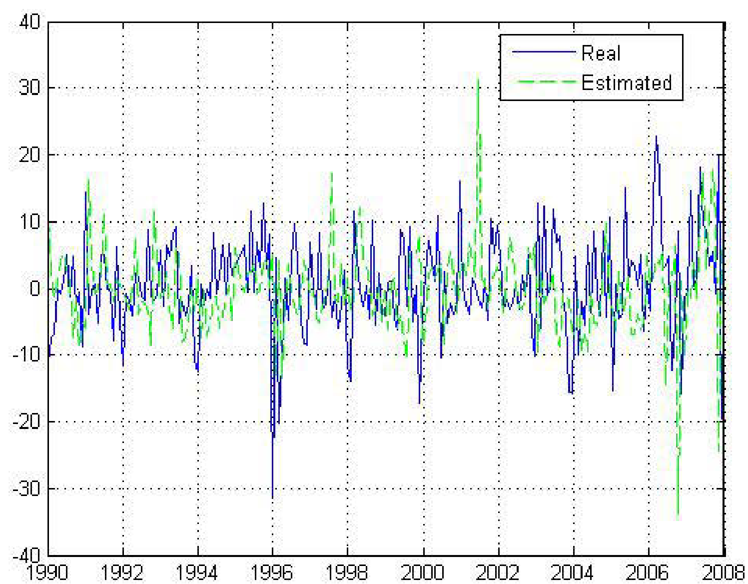


Figure 2.3: Actual and Forecasted Percentage Change in Price: CT

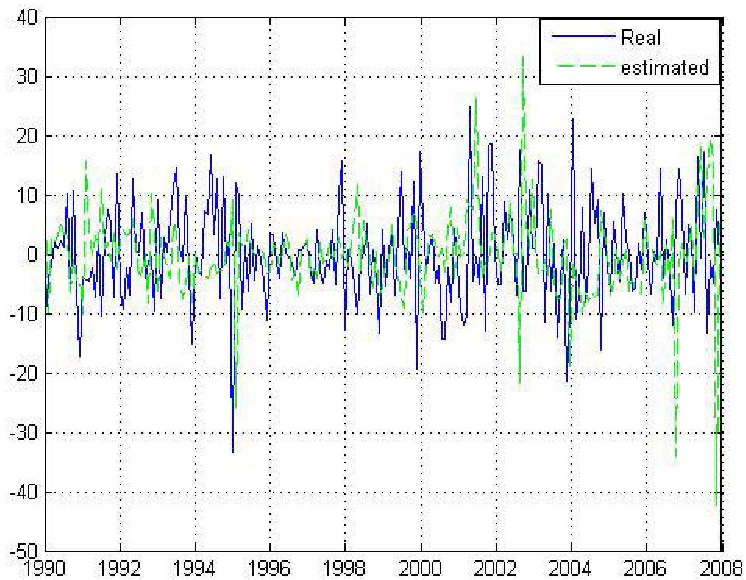


Figure 2.4: Actual and Forecasted Percentage Change in Price: O

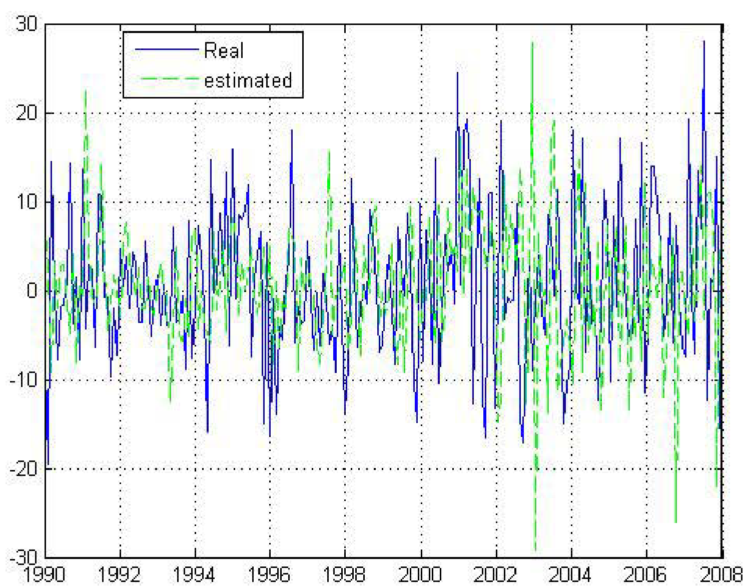


Figure 2.5: Actual and Forecasted Percentage Change in Price: S

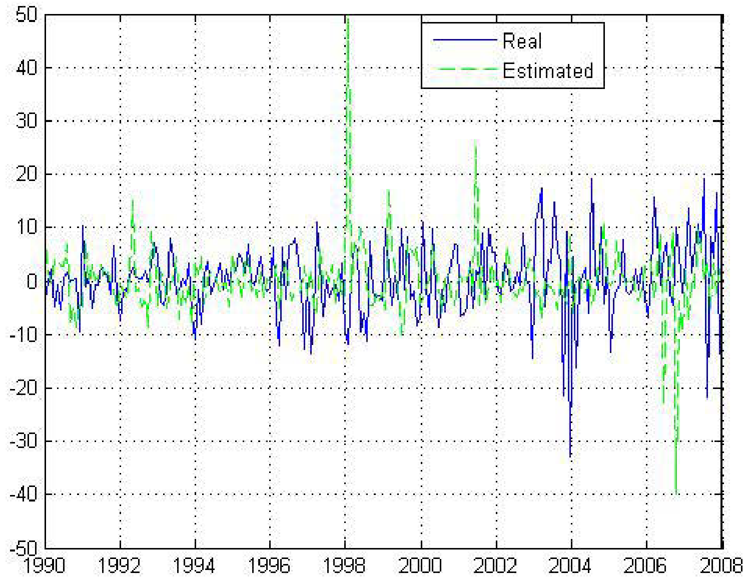
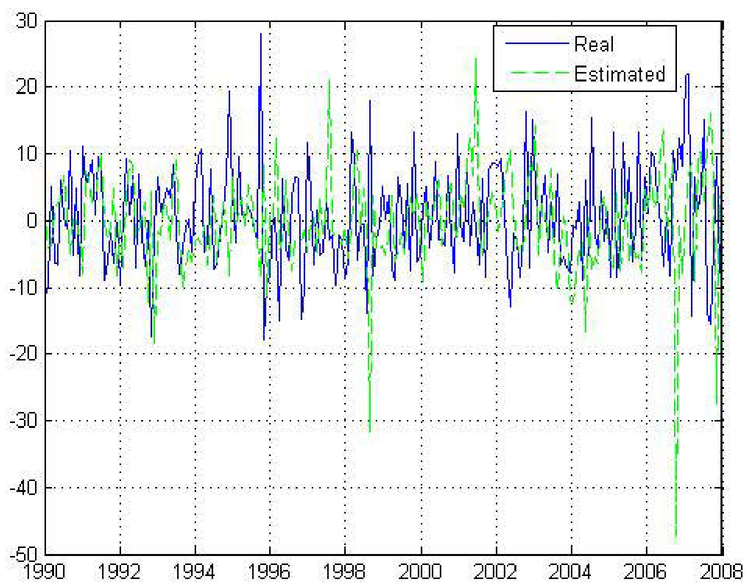


Figure 2.6: Actual and Forecasted Percentage Change in Price:W



Chapter 3

Open Interest and the Hedging Effectiveness of Time-Varying Hedge Ratio Models

3.1 Introduction

The last decades have experienced large changes in commodities futures markets. The commodity market has continuously grown and changed from a primarily physical product market into an elaborated financial market. Given the volatile nature of this market, financial risk plays a major role in this economic environment. The hedging of risk in this market has, therefore, became very important. Hedgers use instruments such as forward or futures contracts, options and derivatives, to offset any potential loss associated with adverse price changes in the corresponding cash markets. Lately, there has been increased interest towards the modeling and calculation of the optimal hedge ratio, as it has become crucial to determine the optimal quantity of instruments needed for hedging purposes. However, it is well known that optimal hedge ratio could be constant or time-varying. Hence it is necessary to test whether a dynamic parameter model performs better than static models.

There has been a significant amount of empirical research that tackles the issue of the calculation of the optimal hedge ratio. Myers (1991), Kroner and Sultan (1991) and Lien and Luo (1993) among others. However, we should note that most empirical works related to the hedge ratio and performance analysis are utilizing the stock futures contracts rather than commodity futures contracts. Despite the large existing literature on the estimation and calculation of the optimal hedge ratio, there is still no clear conclusion as to which method is better, the constant hedge ratio or the time-varying hedge ratios. In fact, there is conflicting evidence found in this literature. Lien, Tse, and Tsui (2002) found that the ordinary least squares (OLS) hedge ratio performs better than the general autoregressive conditional heteroskedasticity (GARCH) hedge ratio, based on the examination of ten spot and futures markets covering currency futures, commodity futures and stock index returns. In addition, they consider their result as an indication that the forecasts generated by the GARCH models are too variable. Thomas and Brooks (2001) do not detect a significant difference between the GARCH model, the Treshold GARCH model and traditional OLS in terms of hedging effectiveness. Myers (1991) employed a multivariate GARCH model to estimate optimal hedge ratios. Rossi and Zucca (2002) provided empirical support for the superiority of the GARCH hedge ratios over the OLS ones in their analysis on the hedging effectiveness conducted on a portfolio composed by Italian Government Bonds.

Moreover, the increasing presence of hedgers and speculators in commodity futures markets has led to many questions about the role that these players can have in these markets. In fact, many researchers suggest that the increasing presence of speculators and hedge funds in commodity futures market could provide an explanation for the increase in energy and food prices observed around the year 2008. Robles, Torero, and Von Braun (2009) argue that the large increase in the volume of globally traded futures and options could be the reason for the rise in commodity prices observed between May 2007 and May 2008. Their work analyses the role that financial speculation plays in affecting prices of agricultural commodities in recent years. Moreover, the authors found some evidence that speculative activity measured as the ratio of non commercial positions to open interest Granger-causes commodity prices of wheat, maize, soybeans and rice. Masters (2008) pointed out that the presence of speculators and institutional investors are one, if not the primary factors affecting commodity prices today. MedlockIII and Jaffe (2009) show that open interest held by speculators moved from a lagging indicator of price to a leading indicator, suggesting a possible reason for the increase in oil prices that started in 2006. Furthermore, Hong and Yogo (2012) argue that open interest is more informative and provides a better signal for higher economic activity than futures prices. They demonstrate that open interest contains information about economic activity and asset prices that are not explained by the traditional supply-demand conditions or by futures price movements. In addition, Hong and Yogo (2010) establish a relation between open interest in commodity market and returns. They show that open interest growth is a robust predictor for commodity returns even after controlling for other known predictors such as the basis. Finally, the variance decomposition for spot commodity prices conducted in our previous chapter provided empirical evidence that the growth rate of open interest contributed to the variation of the percentage change in spot commodity prices. In fact, the growth rate of open interest was empirically found to help predict changes in commodity prices. Open interest should therefore be considered as an additional factor that could help predict variation in commodity prices or commodity returns.

As a means of extending the above analyses, this study considers the case of eleven commodities where general market returns volatilities are affected by the growth rate of open interest. Towards this end, we propose two dynamic models, the diagonal BEKK model of Engle and Kroner (1995) and the the dynamic conditional correlation model of Engle (2002) for commodity returns, modified by adding the growth rate of open interest to the second moment of the commodity returns. Our goal is to test for the hedging effectiveness of these models compared to the other static and dynamic models considered in this chapter. To our knowledge, no previous study empirically investigates the hedging effectiveness of the time-varying hedge ratios for the commodities calculated using the dynamic conditional correlation model and diagonal Bekk model with the growth rate of open interest added as an additional factor influencing the variance and the correlation of spot and futures returns.

The aim of this chapter is to shed the light on these questions. More precisely, we focus on two research questions. First, does the dynamic conditional correlation and the diagonal BEKK models with the growth rate of open interest added to the second moment of the returns provide better hedging effectiveness than the traditional dynamic conditional correlation model of Engle (2002) and the diagonal BEKK model of Engle and Kroner (1995). Second, we examine the out-of-sample performance of various hedging models, and we test if the dynamic hedge models proposed can outperform the standard static hedge models in terms of the basis variance reduction.

The remainder of this chapter is organized as follows. The next section discusses the research methodology and includes a literature review and modeling background for the models used in this empirical study to calculate the optimal hedge ratio and evaluate in-sample and out-of-sample hedging effectiveness. Section 3 describes the data characteristics. Section 4 presents the empirical results. Section 5 concludes the paper.

3.2 Research Methodology

In the present study, several models are employed to estimate the optimal hedge ratios. The performance of these hedge ratios is then compared by checking the in-sample hedging effectiveness and the out-of-sample hedging performance in terms of the basis variance reduction of each model. To estimate the constant hedge ratio, we use two static models, the ordinary least squares (OLS) model and the vector error correction model (VECM). Time-varying hedge ratios are calculated using 4 models: the diagonal Bekk model (DBEKK) and the (DBEKKoi) model modified by adding the growth rate of open interest into the variance and the correlation equations of spot and futures returns. As well, we use the dynamic conditional correlation model (DCC) and the (DCCoi) model modified by adding the growth rate of open interest to the variance and the correlation equations of the commodity spot and futures returns.

3.2.1 Optimal Hedge Strategy and Hedging Effectiveness

The basic idea of the traditional hedging theory is that for each unit of a spot position, one unit in the future market is taken in the opposite position. This is referred to as the naive hedging strategy. However, spot and futures prices do not move in the same direction and by the same amount, thus the value of the hedged portfolio will change over time. Johnson (1960) and Stein (1961) used a portfolio approach to determine the optimal hedging strategy adopting an expected utility maximization approach. A simpler version of their framework, would be the minimum variance hedging strategy, widely used in the literature of hedging as in Ederington (1979), Kroner and Sultan (1991) and Brooks, Henry, and Persand (2002), among many others.

Assume we have a simple hedge model with one period. Let h_t refers to the quantity of futures contract of a commodity sold at time t, h_t will then denotes the hedge ratio. Let R_p denotes the return from a hedged portfolio, R_s and R_f denote spot and futures returns respectively.

$$R_{p,t+1} = R_{s,t+1} - h_t R_{f,t+1}.$$

The variance of the unhedged and the hedged portfolios are calculated as:

$$Var_{unhedged} = \sigma_s^2,$$
$$Var_{hedged} = \sigma_s^2 + h^2\sigma_f^2 - 2h\sigma_{s,f},$$

where, σ_s and σ_f represent the standard deviations of the spot and futures returns, and $\sigma_{s,f}$ denotes the covariance of the two series. The optimal hedge ratio will minimize the variance of the hedged portfolio. The minimum variance constant hedge ratio is:

$$h^* = \frac{\sigma_{s,f}}{\sigma_f^2}. \tag{3.1}$$

In addition, if the variances and covariances vary over time within a GARCH framework, so will the optimal hedge ratio. The time-varying optimal hedge ratio is:

$$h_t^* = \frac{\sigma_{sf,t}}{\sigma_{f,t}^2}. \tag{3.2}$$

Another way of looking at the risk faced by the hedger is to analyze the basis risk faced when taking a position in the futures market, the basis risk is:

$$S_2 - hF_2,$$

where, h represents the quantity of futures contract of a commodity sold at time 1, S_2 denotes the spot price of the commodity to be hedged at time 2, F_1 refers to the future price of the contract used at time 1 and F_2 refers to the future price of the contract used at time 2. For a more extensive analysis of hedging strategies using futures and the corresponding basis risk see Hull (2012).

Let $Var(S - hF)$ denotes the variance of the basis risk. The optimal h is chosen to minimize the basis risk variance, leading to:

$$h^* = \frac{\sigma_{s,f}}{\sigma_f^2}.$$

In-sample Hedging Effectiveness

In order to evaluate the in-sample hedging effectiveness of the various models considered in this research, the variances of the hedged and the unhedged portfolios are calculated for all the models. Ederington (1979) used portfolio theory to propose a better measure for the hedging effectiveness of a model. Specifically, the hedging effectiveness of a model could be measured by the percentage reduction in the variance of the hedged portfolio to the unhedged portfolio. Hence, hedging effectiveness is computed using:

$$HE = \frac{(Var_{unhedged} - Var_{hedged})}{Var_{unhedged}} * 100 \tag{3.3}$$

For a dynamic hedge model, (3.3) becomes:

$$HE_t = \frac{(Var_{unhedged,t} - Var_{hedged,t})}{Var_{unhedged,t}} * 100 \tag{3.4}$$

The in-sample hedging effectiveness for static and dynamic models in our empirical work will be determined using equations (3.3) (3.4) which frequently serve as a measure of hedging effectiveness in the body of research on optimal hedge estimation (see among others, Lien and Luo (1993), Dewally and Marriott (2008), Santhosh (2012) and Caporin, Jimenez-Martin, and Gonzalez-Serrano (2014).)

Forecasting Optimal Hedge Ratios Using Dynamic GARCH Models

Many authors employed GARCH family models to forecast the variance and covariance for volatility (see Engle and Ng (1993) and YU (2002) among others). However, a few studies exist for hedge ratio forecasting (see Kroner and Sultan (1993) and Yang and Allen (2005) among others).

The one-step ahead forecast of conditional variance for a standard GARCH(1,1) model is

$$\hat{\sigma}^2_{t+1} = \alpha_0 + \alpha_1 E(u_t^2 \mid I_{t-1}) + \beta_1 \sigma_t^2,$$

where u_t^2 denotes the squared residuals, $E(u_t^2 \mid I_{t-1}) = \sigma_t^2$. For a more extensive analysis see Bollerslev (1987) and Hamilton (1994). The one-step ahead forecast of conditional variance could be written as:

$$\hat{\sigma}^2_{t+1} = \alpha_0 + (\alpha_1 + \beta_1)\sigma_t^2.$$

Since the unconditional variance of returns $E[\sigma^2] = \frac{\alpha_0}{1-\alpha_1-\beta_1}$,

$$\hat{\sigma}_{t+1}^2 = \sigma^2 + (\alpha_1 + \beta_1)(\sigma_t^2 - \sigma^2).$$

Similarly, the two-step ahead variance forecast is

$$\begin{aligned}
\hat{\sigma}_{t+2}^2 &= \alpha_0 + \alpha_1 E(u_{t+1}^2 \mid I_{t-1}) + \beta_1 E[\sigma_{t+1}^2 \mid I_{t-1}], \\
&= \alpha_0 + (\alpha_1 + \beta_1)\hat{\sigma}_{t+1}^2, \\
&= \sigma^2 + (\alpha_1 + \beta_1)^2(\sigma_t^2 - \sigma^2).
\end{aligned}$$

In a such way, the l-step ahead forecast can be written as:

$$\hat{\sigma}_{t+l}^2 = \sigma^2 + (\alpha_1 + \beta_1)^l(\sigma_t^2 - \sigma^2). \tag{3.5}$$

From the l-step ahead variance forecast, we can see that $(\alpha_1 + \beta_1)$ determines how fast the variance forecast converges to the unconditional variance. Equation (3.5) is used to forecast the conditional variances and covariances over a period of 24 months, for dynamic models, DBEKK and DCC.

For the DBEKKoi and DCCoi models we add the growth rate of open interest to the conditional variance and correlation equations of the commodity spot and futures returns. By adding the lagged growth rate of open interest to the second moment of the commodity returns, we model the time-varying conditional variances and covariances within a GARCHX framework. The GARCHX model was proposed by Hwang and Satchell (2005) for modeling aggregate stock market return volatility. The model includes a measure of the lagged cross-sectional return variation as an explanatory variable in the GARCH conditional variance equation. Using 10 years UK and US daily data, Hwang and Satchell (2005) show that the return volatility of an individual stock can be better specified with GARCHX models.

We start by assuming that the growth rate of open interest follows an AR(q) process.The optimal lag length q is determined using the multivariate version of Akaike's information criteria. [1]

[1] Results for the AR(q) order selection can be provided upon request.

Let X_t denotes ΔOI_t,

$$X_t = c + \sum_{i=1}^{q} \rho_q X_{t-q} + \varepsilon_t.$$
$$E[X_t] = c + \sum_{i=1}^{q} \rho_q E[X_{t-q}].$$

Let $E[X_t] = \mu$. We assume that X_t follows a stationary process, $E[X_t] = E[X_{t-1}] = ...E[X_{t-q}] = \mu$ (see Lutkepohl (1991) and Hamilton (1994) among others), $E[X_t]$ can be written as:

$$\begin{aligned} E[X_t] &= c + \mu \sum_{i=1}^{q} \rho_q, \\ \implies \mu &= \frac{c}{1 - \sum_{i=1}^{q} \rho_q}. \end{aligned} \tag{3.6}$$

The GARCH(1,1) conditional variance with growth rate of open interest is

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_t^2 + \beta_1 \sigma_{t-1}^2 + \gamma_1 X_{t-1}. \tag{3.7}$$

The one-step ahead forecast of conditional variance for a GARCH(1,1) model with growth rate of open interest is:

$$\begin{aligned} \hat{\sigma}_{t+1}^2 &= \alpha_0 + \alpha_1 E(u_t^2 \mid I_{t-1}) + \beta_1 \sigma_t^2 + \gamma_1 E(X_t \mid I_{t-1}), \\ \hat{\sigma}_{t+1}^2 &= \alpha_0 + (\alpha_1 + \beta_1)\sigma_t^2 + \gamma_1 \mu, \\ \hat{\sigma}_{t+1}^2 &= \alpha_0 + (\alpha_1 + \beta_1)\sigma_t^2 + \frac{\gamma_1 c}{1 - \sum_{i=1}^{q} \rho_q}. \end{aligned} \tag{3.8}$$

Let σ_{oi}^2 denotes the unconditional variance of returns. Since $\sigma_{oi}^2 = \frac{\alpha_0 + \frac{\gamma_1 c}{1 - \sum_{i=1}^{q} \rho_q}}{1 - \alpha_1 - \beta_1}$, equation (3.8) could be written as:

$$\hat{\sigma}_{t+1}^2 = \sigma_{oi}^2 + (\alpha_1 + \beta_1)(\sigma_t^2 - \sigma_{oi}^2).$$

The l-step ahead forecast for the conditional variance for the GARCH(1,1) model with growth rate of open interest is:

$$\hat{\sigma}_{t+l}^2 = \sigma_{oi}^2 + (\alpha_1 + \beta_1)^l (\sigma_t^2 - \sigma_{oi}^2). \tag{3.9}$$

Equation (3.9) is used to forecast the conditional variances and covariances over a period of 24 months for dynamic models, DBEKKoi and DCCoi.

Finally, the forecasts obtained for conditional variances and covariances for the 4 dynamic models are used to calculate the optimal out-of-sample hedge ratios over the 24 months time horizon.

Out-of-Sample Hedging Exercise

In order to determine whether a static or a dynamic model is more appropriate to calculate the optimal hedge ratio, we conduct an out-of-sample hedging exercise. To evaluate the out-of-sample hedging performance of the various models, we adopt the following procedure. First we split the sample into two periods, period A covers observations from first month up to period $t-24$ months, where t represents the sample size. Period B covers the last 24 months in the sample. We then use observations from the sample in period A to estimate the parameters of the various models, namely the OLS, the VECM, the DCC, the DCCoi, the DBEKK and the DBEKKoi models. Once we obtain the estimated parameters, we use them to forecast the optimal hedge ratios for the four dynamic models for the last 24 months of the sample in period B. For the static models, OLS and VECM, the hedge ratios obtained from the estimation of the sample in period A are used to conduct the hedging exercise in period B.

Finally, we compute the basis for each commodity and for each model over the period of the last 24 months of the sample. Specifically, we use the forecasted optimal hedge ratios for the last 24 months in the sample for the 4 dynamic models and the two constant hedge ratios calculated from observations in sample period A for OLS and VECM to calculate the basis. Spot and futures prices used in the calculation of the basis for the last 24 months in the sample are the actual prices not the forecasted ones. Since the basis changes, the model that minimizes the variance of the basis over the period of the last 24 months of the sample will be chosen as the best model to calculate the optimal hedge ratio. The risk reduction is therefore associated with a reduction in the basis.

Let $\hat{h}^*$ denotes the optimal hedge ratio, $\hat{h}^*$ could be less than 1 or greater than one, in which case the hedger takes a greater position in the futures than in the cash market. Assuming that the value of the commodity contract to be hedged and the value of the underlying future contract used in hedging are the same, the basis is:

$$Basis = S_2 - \hat{h}^* F_2. \tag{3.10}$$

We use equation (3.10) to calculate the basis for all six models and all 11 commodities for the last

24 months in the sample. In addition, we calculate the variance of the basis for each commodity and each model over the last 24 months time horizon.

We also consider two extreme cases. First, we set the optimal hedge ratio to be equal to 1, which represents the perfect hedge situation and second we set it to zero, which represents the naked position. Results for the basis variance computed over the last 24 months time period for these two cases are than compared to the ones obtained from OLS and to the variance of the best out-of-sample performing model for the last 24 months in the sample.

3.2.2 Optimal Hedge Ratio Estimates of Static Models

The Conventional Regression Method for a Constant Hedge Ratio

The most traditional way to compute a constant hedge ratio is to use linear regression of changes in spot prices on changes in future prices. Johnson (1960), Stein (1961) and Ederington (1979) adopted the methodology of regressing spot prices on futures prices using ordinary least squares (OLS). Let S_t and F_t denote the natural logarithm of spot and futures prices at time t respectively. The one- period minimum variance hedge ratio is estimated using the following expression:

$$\Delta S_t = \alpha + h^* \Delta F_t + \varepsilon_t, \tag{3.11}$$

where, ΔS_t and ΔF_t represent changes in the logarithm of spot and futures prices or spot and futures prices returns, ε_t is the error term from OLS regression and $\hat{h}^*$ is the estimated minimum variance constant hedge ratio, it is the estimated slope coefficient from the OLS regression.

Even though the conventional OLS method is easy to apply, this approach suffers from a number of aspects that could invalidate the hedge ratio estimates. It does not account for any possible autocorrelation between the residual series of spot and futures returns, as noted at the outset by Herbst, Kare, and Marshall (1989).

The VECM Model

To overcome the fact that there might be serial correlation between the residuals series and to account for the possibility of cointegration between return series, Lien and Luo (1994) and Lien (1996) proposed to use the VECM model instead of the VAR model. In the Bivariate VECM model we have two variables, spot prices returns ΔS_t and futures prices returns ΔF_t. Each variable current value will depend on a combination of previous values of both variables along with an error correction term which accounts for the long-run equilibrium between spot and futures price movements.

$$\Delta S_t = \lambda_s + \sum_{i=1}^{p} \alpha_{si} R_{s,t-i} + \sum_{j=1}^{p} \beta_{sj} R_{f,t-j} + \kappa_s Z_{t-1} + e_{s,t}, \tag{3.12}$$

$$\Delta F_t = \lambda_f + \sum_{i=1}^{p} \alpha_{fi} R_{s,t-i} + \sum_{j=1}^{p} \beta_{fj} R_{f,t-j} + \kappa_f Z_{t-1} + e_{f,t}, \tag{3.13}$$

where, λ_s and λ_f are constant terms, $\alpha_s, \alpha_f, \beta_s$, β_f, κ_s and κ_f are positive parameters and e_{st} and e_{ft} are white noise residuals. The optimal lag length p is determined using the multivariate version of Akaike's information criteria. [2]

Z_{t-1} represents the error correction term, which indicates how the dependent variables adjusts to the previous period's deviation from the long run equilibrium.

$$Z_{t-1} = S_{t-1} - C - \eta F_{t-1}, \tag{3.14}$$

where, C represents a constant and η is the cointegrating vector. The coefficients κ_s and κ_f explain the speed of adjustment of the dependent variables. A higher value for κ_s is interpreted as a larger response of S_t to the previous period's deviation from the long-run equilibrium.

Let $Var(e_{st}) = \sigma_{ss}$, $Var(e_{ft}) = \sigma_{ff}$ and $cov(e_{st}, e_{ft}) = \sigma_{sf}$. The minimum variance optimal hedge ratio is computed as:

$$\hat{h}^* = \frac{\hat{\sigma}_{sf}}{\hat{\sigma}_{ff}}. \tag{3.15}$$

The VECM model generates a constant hedge ratio, which implicitly assumes that the risk in spot and futures market is constant over time. Bollerslev (1990) and Kroner and Sultan (1991) argue that the availability of new information will change the degree of risk associated with various assets. The optimal hedge ratio calculated should therefore account for the conditional distribution of spot and futures returns and as result we need to consider the choice of dynamic models to estimate time-varying optimal hedge ratios.

[2] Results for the VECM order selection can be provided upon request.

3.2.3 Optimal Hedge Ratio Estimates of Dynamic Models

Diagonal BEKK Model (DBEKK)

Most financial series are characterized by the presence of conditional heteroskedasticity. As the presence of conditional heteroskedasticity partly affects the hedge ratio estimates obtained using OLS or Bivariate VECM methods, dynamic models are therefore employed. Specifically, multivariate GARCH models have been developed and applied in financial econometrics in an attempt to get a more accurate computation of the optimal hedge ratio. The univariate GARCH model has been generalized to N-variable multivariate GARCH models in many different ways. Bollerslev, Engle, and Wooldridge (1988) suggest a general representation of the GARCH model, called the VECH model. However, the VECH model suffers from a disadvantage that only a sufficient condition for the positive definiteness of the variance-covariance matrix H_t is known. To reduce the disadvantages of the VECH model, Bollerslev, Engle, and Wooldridge (1988) proposed the diagonal VECH model. Even though the multivariate diagonal VECH model sets constraints on the VECH model to successfully decrease the number of parameters, this model framework still presents an important drawback. The conditional covariance might be negative if other restrictions are not added to the model. To ensure a positive definite conditional covariance matrix H_t, Engle and Kroner (1995) proposed the model known as Baba-Engle-Kraft-Kroner (BEKK) model and also the diagonal BEKK model.

In order to reduce the number of parameters and ensure positive definitiveness of the covariance matrix, we finally use the Diagonal BEKK model given by Engle and Kroner (1995). It is worth noting that our empirical exercise shows that for most of the commodities we were not able to obtain a positive definite conditional covariance matrix when using the VECH model, the diagonal VECH model or even the BEKK model, particularly when adding the growth rate of open interest to the models' equations. The DBEKK model GARCH (1,1) used in our empirical work is written as:

$$h_{11,t} = c_{11}^2 + a_{11}^2 u_{1,t-1}^2 + b_{11}^2 h_{11,t-1}, \tag{3.16}$$

$$h_{22,t} = c_{12}^2 + c_{22}^2 + a_{22}^2 u_{2,t-1}^2 + b_{22}^2 h_{22,t-1}, \tag{3.17}$$

$$h_{12,t} = c_{11}c_{12} + a_{11}a_{22}u_{1,t-1}u_{2,t-1} + b_{11}b_{22}h_{12,t-1}, \tag{3.18}$$

where, $h_{i,j,t}$ in these variance covariance equations only depend on their own lagged values $h_{i,j,t-1}$. As for the model specifications for the returns equations, we use the VECM specification described above.

Using the DBEKK model, the time-varying hedge ratio can be computed as

$$\hat{h}_t^* = \hat{h}_{12,t}/\hat{h}_{22,t}, \tag{3.19}$$

where, $\hat{h}_{12,t}$ is the estimated conditional covariance between the cash and futures returns, and $\hat{h}_{22,t}$ represents the conditional variance of futures returns. Since the conditional covariance is time-varying, the optimal hedge ratio would be time- varying too.

The Dynamic Conditional Correlation Model (DCC)

The heteroscedastic nature of the price series was first recognized by Engle (1982) who introduced the Autoregressive Conditional Heteroscedasticity (ARCH) model, later, the Generalized Autoregressive Conditional Hetroscedasticity (GARCH) was proposed by Bollerslev (1986). Engle (2002) developed the Constant Correlation-GARCH model and proposed Dynamic Conditional Correlation GARCH (DCC-GARCH) model. He argued that this model is characterized by the flexibility of univariate GARCH models and provides at the same time parsimonious correlation specifications without the computational difficulties of multivariate GARCH models. DCC estimators proposed by Engle (2002) are estimated in two steps: a series of univariate GARCH estimates and the correlation estimate. Specifically, let $R_t = (R_{st}, R_{ft})'$ denotes spot and futures prices returns and let $e_t = (e_{st}, e_{ft})'$ denotes the errors. To model the long-run equilibrium relationship, the VAR equations are replaced by VECM equations. Hence, the spot and futures prices are specified using VECM:

$$R_{s,t} = \lambda_s + \sum_{i=1}^{p} \alpha_{si} R_{s,t-i} + \sum_{j=1}^{p} \beta_{sj} R_{f,t-j} + \kappa_s Z_{t-1} + e_{s,t}, \tag{3.20}$$

$$R_{f,t} = \lambda_f + \sum_{i=1}^{p} \alpha_{fi} R_{s,t-i} + \sum_{j=1}^{p} \beta_{fj} R_{f,t-j} + \kappa_f Z_{t-1} + e_{f,t}. \tag{3.21}$$

Assuming that returns are normally distributed with zero mean, the conditional variance- covariance matrix of the returns and the residuals series is denoted as follows:

$$Var(R_t \mid I_{t-1}) = Var(e_t \mid I_{t-1}) \equiv H_t = \begin{bmatrix} h_{s,t} & h_{sf,t} \\ h_{sf,t} & h_{f,t} \end{bmatrix}. \tag{3.22}$$

The DCC model specifies the equations for time-varying variances and covariance as follows:

$$h_{s,t} = \omega_s + \theta_s e^2_{s,t-1} + \gamma_s h_{s,t-1}, \tag{3.23}$$

$$h_{f,t} = \omega_f + \theta_f e^2_{f,t-1} + \gamma_f h_{f,t-1}, \tag{3.24}$$

$$\rho_t = (1 - \delta_1 - \delta_2)\bar{\rho} + \delta_1 \varepsilon_{t-1} + \delta_2 \rho_{t-1}. \tag{3.25}$$

Following Johnson (1960), the optimal hedge ratio $\hat{h}^*_t$, is determined as:

$$\hat{h}^*_t = \hat{\rho}_t \sqrt{\frac{\hat{h}_{s,t}}{\hat{h}_{f,t}}}, \tag{3.26}$$

where, $h_{s,t} = Var(R_{s,t} \mid I_t)$, $h_{f,t} = Var(R_{f,t} \mid I_t)$ and ρ_t is the conditional correlation between $R_{s,t}$ and $R_{f,t}$.

The Diagonal BEKK Model with the Growth Rate of Open Interest (DBEKKoi)

In addition to the above DBEKK model for determining the optimal hedge ratio, we propose a modified DBEKK model that accounts for the impact of the lagged growth rate of open interest on commodity spot and futures returns' volatility and covariance.

The model proposed represents an attempt to provide an empirical evidence of the ability of the growth rate of open interest to improve the hedging effectiveness when added to the variances and covariance equations of the DBEKK model.

The DBEKKoi model equations for conditional variances and covariance are therefore given by:

$$h_{11,t} = c^2_{11} + a^2_{11} u^2_{1,t-1} + b^2_{11} h_{11,t-1} + d^2_{11} \Delta OI_{t-1}, \tag{3.27}$$

$$h_{22,t} = c^2_{12} + c^2_{22} + a^2_{22} u^2_{2,t-1} + b^2_{22} h_{22,t-1} + d^2_{22} \Delta OI_{t-1}, \tag{3.28}$$

$$h_{12,t} = c_{11} c_{12} + a_{11} a_{22} u_{1,t-1} u_{2,t-1} + b_{11} b_{22} h_{12,t-1} + d_{11} d_{22} \Delta OI_{t-1}. \tag{3.29}$$

The optimal hedge ratio $\hat{h}_t^*$ for a DBEKKoi model is computed using equation (3.19).

The Dynamic Conditional Correlation Model with Growth Rate of Open Interest (DCCoi)

In addition to the above usual models for determining the optimal hedge ratio, we propose a more advanced DCC model to account for the impact of the growth rate of open interest on commodity spot and futures returns' volatility and co-movements. In an attempt to provide empirical evidence of the ability of the growth rate of open interest to improve the hedging effectiveness of the DCC model, this factor is added to the second moment of the model. In our previous chapter, empirical results for the variance decomposition of commodity prices support the claim that the growth rate of open interest partly explains the variation in commodity prices. Specifically, we incorporate the lagged growth rate of open interest into the DCC equations for conditional variances and conditional correlation. Little research to date has focused on how dynamic optimal hedge ratios perform when time-varying conditional variances and correlations modeled as in the traditional DCC model of Engle (2002) are affected by an additional factor. Among others, Sim and Zurbruegg (2001) examined the hedging effectiveness of KOSPI 200 futures contract by adding the lagged basis to the returns equations and the absolute value of the lagged basis to the variances and covariances equations. Zhong, Darrat, and Otero (2004) adopted a similar methodology using EGARCH instead of GARCH formulation and adding the square of the basis to the variances and covariances equations, in an attempt to test the hedging effectiveness of Mexico IPC index futures contract.

The model specifications for the returns equations are the same as in the DCC model. The DCCoi model specifies the equations for time-varying variances and covariance as follows:

$$h_{s,t} = \omega_s + \theta_s e_{s,t-1}^2 + \gamma_s h_{s,t-1} + \tau_s \Delta OI_{t-1}, \tag{3.30}$$

$$h_{f,t} = \omega_f + \theta_f e_{f,t-1}^2 + \gamma_f h_{f,t-1} + \tau_f \Delta OI_{t-1}, \tag{3.31}$$

$$\rho_t = (1-\delta_1-\delta_2)\bar{\rho} + \delta_1 \varepsilon_{t-1} + \delta_2 \rho_{t-1} + \delta_3 \Delta OI_{t-1}. \tag{3.32}$$

3.3 Data Description

This section presents a description for the data that we will be using in the empirical work. Our empirical exercise considers monthly observations for spot and futures commodity prices as well as trading volume data.

We demonstrate the application of the above procedures using data sets for eleven commodities. This application is interesting due to the fact that for six of the commodities considered in this research, changes in open interest were empirically found to make relatively close contributions to the variation of the percentage change in spot commodity prices, as in the previous chapter. Furthermore, as commodity markets are fairly active, we enlarge our data set to include five additional commodities in an attempt to gain wider representation.

Data for spot and futures commodity prices are retrieved from two different sources, the Commodity Research Bureau (CRB) and Bloomberg. Data obtained from CRB are available at a monthly frequency for the period between March 1990 and July 2008. The data set consists of series of spot and futures prices for seven commodities from 3 groups: grains and oilseeds (corn, oats, soybeans, soybean oil and wheat), industrials (cotton) and foodstuffs (coffee). The trading characteristics of these commodities are summarized in Table 3.1. Data retrieved from Bloomberg are available at a monthly frequency for the period between October 1992 and December 2012. We collect data for four commodities from 3 different groups: metals (gold), foodstuffs (cocoa and sugar), and live stock and meat (feeder cattle). The trading characteristics of these commodities are summarized in Table 3.2. The spot prices series are approximated by the nearest futures contract and the futures prices are approximated by the next to the nearest futures contract.

Trading volume data for the 7 commodities futures contracts retrieved from CRB, are downloaded from the US commodity futures trading commission website.[3] Monthly data for the trading volume for the remaining 4 commodities futures contracts are retrieved from Bloomberg.

To estimate the in-sample optimal hedge ratios and the in-sample hedging effectiveness, we use the whole sample period. To conduct the out-of-sample hedging exercise we split our sample into two samples A and B. Parameter estimates from sample A are used to forecast optimal hedge ratios for dynamic models for sample B and to estimate the constant hedge ratios used to calculate the basis variance for sample B. For the CRB data, sample A covers the period between March 1990

[3]Trading volume data retrieved from CFTC website are available from September 1992. Website: http://www.cftc.gov

and August 2006. Sample B covers the period between September 2006 and July 2008. For the Bloomberg data, sample A covers the period between October 1992 and December 2010. Sample B covers the period between January 2011 and December 2012.

3.4 Empirical Results

3.4.1 Cointegration Test Results for VAR Model

In order to avoid spurious regression results, it is necessary to first test for stationarity of the variables used in our empirical research. We then conduct a cointegration test.

The Augmented Dickey Fuller (ADF) tests results for all eleven commodities indicate that all log prices have a unit root. However, unit root tests results for spot and futures prices returns reject the presence of unit root at 95 % confidence level.[4]

To test for cointegration between spot and futures prices series, we conduct the Johansen (1988) cointegration test on logged spot and futures prices. Results for Johansen's cointegration test are reported in tables 3.3 and 3.4. Two tests are used, the first one is designed to test for the presence of r cointegrating vectors, the trace test. The second will test the null hypothesis of r cointegrating vectors versus the alternative of $r+1$ cointegrating vectors, the maximum eigenvalue test. The Johansen's cointegration test is a cointegration likelihood ratio test based on the maximal eigenvalue and the trace of the stochastic matrix. Table 3.3 presents the results for the trace statistic and table 3.4 indicates the maximal eigenvalue statistic. Results indicate that both the trace and eigenvalue statistics strongly reject the null that there is no cointegrating vector existing for all commodities. However, when the null is that there exists a single cointegrating vector, both statistics tend not to reject it for all commodities considered in this research.

Our cointegration test results indicate that logged spot and futures price series are cointegrated. Since the Bivariate VAR model ignores the possibility that the two returns series might be cointegrated, which in that case could affect the estimation results, the VECM model is used in our empirical application to model the commodity spot and futures returns.

[4] ADF unit root test results are available upon request.

3.4.2 In-Sample Estimation Results for DCC and DCCoi Models

The parameter estimates for the mean equations for the DCC and the DCCoi models are presented in Table 3.5. In this table, we only report estimates for the cointegrating equation parameters and the cointegrating coefficient estimate.[5] The standard errors and t-ratios statistics are presented below the coefficients to show each coefficient's relative significance at the 95% significance level. The error correction term coefficients κ_s are statistically significant in all spot prices returns equations except for wheat coffee and cocoa. The error correction term coefficients κ_f are statistically significant for four of the commodities. It is noted that for all commodities the coefficient κ_f is bigger than κ_s, indicating that the futures prices returns adjust more rapidly than spot prices returns to the previous period 's deviation from the log-run equilibrium. For the whole sample period, commodity returns react positively to the spot returns from the previous periods and negatively to the futures returns from the previous periods.

Estimation results for equations (3.23)and (3.24) of the DCC model are summarized in Tables 3.6 and 3.7. The GARCH effects parameters, γ_s and γ_f, vary between 0.65 and 0.96 for all commodities except for soybean oil and wheat where it is around 0.1. A large coefficient of the GARCH term indicates that shocks to conditional variances take a long time to die out and volatility persists. The ARCH effects parameters, θ_s and θ_f are around 0.2. These estimates are in line with the empirical properties of commodity prices which are characterized by high persistence and conditional heteroskedasticity.

Tables 3.8 and 3.9 show the estimation results for equations (3.30) and (3.31) of the DCCoi model. Similar observations for GARCH and ARCH effects parameters are observed. Volatilities of the spot and futures commodity returns are affected positively by the growth rate of open interest. As the growth rate of open interest increases, both markets become more volatile. All the parameters τ_s are statistically significant at the 95% significance level except for corn as indicated by the t-test ratios in Table 3.8. The parameters τ_f are significant at the 95% significance level for all commodities except for corn, cotton and coffee as indicated by the t-test ratios in Table 3.9.

Tables 3.10 and 3.11 present the estimation results for equations (3.25) and (3.32). The GARCH effect parameters δ_2 are all ranging between 0.7 and 0.97 except for cotton and corn. The ARCH effect parameters δ_1 are around 0.1 for all commodities. The growth rate of open interest parameter δ_3 is significant for all commodities at the 95% significant level, except for soybean oil, wheat

[5]Results for the estimates of the coefficients of the spot and futures returns lags are available upon request.

and cocoa. In addition, the growth rate of open interest has a positive impact on the correlation between the spot and futures commodity returns.

The in-sample parameter estimations results for the variances and the covariance equations of the DCC and the DCCoi models indicate that the GARCH parameter estimates are relatively high for most of the commodities. These results reflect the empirical properties (high persistence and conditional heteroskedasticity) of commodity prices. Furthermore, the growth rate of open interest parameter is statistically significant for most commodities and was empirically found to have a positive effect on the variance and the correlation between the spot and futures commodity returns.

3.4.3 In-Sample Estimation Results for DBEKK and DBEKKoi Models

To estimate the parameters for the mean equations for spot and futures commodity returns we also used the VECM model. Hence, the results are the same as the one discussed for the DCC and DCCoi models presented in Table 3.5.

Estimation results for equations (3.16),(3.17) and (3.18) of the DBEKK model are presented in Tables 3.12 and 3.13.

Tables 3.14 and 3.15 display the parameter estimates for equations (3.27),(3.28) and (3.29) of the DBEKKoi model.

For both models, the GARCH effects parameters b_{11} and b_{22} are ranging from 0.5 to 0.94. Large GARCH parameters are in line with the fact that commodity prices are highly persistent. a_{11} and a_{22} parameters represent the effect of the squared residuals of the two returns series on the corresponding conditional variance and covariance series.

The growth rate of open interest parameters d_{11} and d_{22} added to the DBEKKoi model, presented in Tables 3.14 and 3.15, are all statistically significant at the 95% significance level, except for corn, soybeans and coffee.

Results for the parameter estimates for the DBEKKoi model are very similar to the ones obtained from the DCCoi model where the growth rate of open interest parameters were also not statistically significant for corn and coffee in the conditional variance equation. Furthermore, we can notice that the growth rate of open interest has a positive impact on the conditional variance and covariance between the return series for all the commodities except for cotton where the impact is negative. As the market becomes more volatile, the conditional variance and the conditional correlation of the commodity returns increase.

3.4.4 Hedging Effectiveness Comparisons

Comparing In- Sample Hedging Effectiveness for the DCC and the DCCoi Models

Based upon the above estimation results, we calculate the dynamic optimal hedge ratios generated from the conventional DCC model and the DCCoi model which accounts for the impact of the growth rate of open interest, for the eleven commodities considered in this empirical exercise. The averages of the dynamic hedge ratios are presented in Table 3.16. The DCCoi model generates higher optimal hedge ratios for all the commodities, as indicated in Table 3.16. A hedger that does not account for the impact that the growth rate of open interest has on the volatility and correlation of spot and futures returns is likely to take a smaller than optimal futures position.

Table 3.17 displays the in-sample hedge performance comparison results for the DCC and the DCCoi models. The DCCoi model outperforms the DCC model in terms of the in-sample hedging effectiveness. The variance reduction of the unhedged portfolio is either marginally or substantially higher for the DCCoi model for nine of the commodities except for soybean oil and wheat. This result is excepted since for these two commodities, the GARCH effect parameters in the variance equations, γ_s and γ_f, are very low and the growth rate of open interest is not significant in the correlation equation.

Our empirical results indicate that the dynamic hedge ratios generated by the DCCoi model are higher than the ones estimated using the DCC model. Results for the in-sample hedging effectiveness show that the DCCoi model produces higher in-sample hedging effectiveness for most of the commodities considered.

Comparing In- Sample Hedging Effectiveness for the DBEKK and the DBEKKoi Models

The dynamic minimum optimal hedge ratios generated from the DBEKK model and the DBEKKoi model which adds the impact of the growth rate of open interest on the conditional variance and covariance of the returns are calculated using equation (3.19) for the eleven commodities considered in our empirical work.

Table 3.18 presents the means for the dynamic hedge ratios calculated using DBEKK and DBEKKoi models. As indicated by the results, the hedge ratios for all the commodities are almost the same for both models.

In-sample hedging effectiveness results for both models are summarized in Table 3.19. As expected and since there is no large differences in the hedge ratios obtained from the two models, variance reductions of the unhedged portfolios are either the same or slightly higher for the DBEKK model.

Our estimation results for the in-sample dynamic hedge ratios and the in-sample hedging effectiveness indicate that there no sufficient empirical evidence to support the claim that the DBEKKoi model outperforms the DBEKK model in terms of increasing the in-sample hedging effectiveness.

Out-of-Sample Hedging Exercise Results for Static and Dynamic Models

In order to determine whether a static or a dynamic model is more appropriate to calculate the optimal hedge ratio, we compute two static optimal hedge ratios using the OLS and VECM models.

Even though most empirical results for in-sample hedging effectiveness in the literature are in favor of using dynamic models to calculate optimal hedge ratios, nonetheless the highly volatile nature of the dynamic hedge ratios presents a difficulty in terms of choosing the appropriate hedging strategy, due to the fact that the hedged portfolio must be rebalanced on a period by period basis. This procedure might incur high transactions costs and therefore the hedger may search for another hedge method such as the conventional OLS.

To visualize this highly volatile nature of dynamic optimal hedge ratio, we plot the dynamic hedge ratios obtained from DCC and DCCoi models as well as the static optimal hedge ratio estimated using OLS, for each commodity. Optimal hedge ratios are estimated using all observations in the sample. Figures 3.1 to 3.3, represent the hedge ratios for the eleven commodities. We notice that the difference between the two dynamic hedge ratios is not very high, with the hedge ratio of the DCCoi model being always higher than the one estimated from the DCC model. Moreover, compared with the static OLS optimal hedge ratio, the various hedge ratios are close. In addition, Figures 3.1 to 3.3 provide an insight for the high rate of fluctuation of the dynamic hedging.

Finally, we check the performance of each of the models by conducting an out-of-sample hedging exercise over the last 24 months time horizon of the sample.

Results for each commodity and for all models are shown in Table 3.20 . Each row in the table indicates for each commodity the variance of the basis for all six models over the last 24 months of the sample. Table 3.21, presents a summary for the out-of-sample hedging exercise results. For each commodity we provide a ranking for the models in terms of the basis variance reduction. The model that gives the smallest variance for the basis over the last 24 months period of the sample will have a ranking of 1.

As indicated by our empirical results in Tables 3.20 and 3.21, for 5 commodities the static models are the best performing models in terms of the basis variance reduction. The OLS model provides the smallest basis variance for oats, gold and cocoa. It is followed by the VECM model which reduces the variance of the basis for cotton and coffee. Dynamic models are providing the best out-of-sample performance for the remaining 6 commodities. Particularly, the DCC model is the best performing model for 4 commodities, soybeans, soybean oil, sugar and feeder cattle. Finally, our DBEKK model with growth rate of open interest added to the variances and covariance equations provides the smallest variance for the basis for 2 commodities, corn and wheat. This result is in line with our results in the previous chapter were the growth rate of open interest played a role in determining the variance of spot commodity prices. In addition, it highlights the fact that adding the growth rate of open interest to our dynamic models might improve the hedging performance of these models for some commodities.

Moreover, in an attempt to explain our empirical results, we plot the out-of-sample optimal hedge ratios for each commodity and for all 6 models considered in this research. Figures 3.4 to 3.7 show the two static optimal hedge ratios and the forecasted optimal hedge ratios calculated using the 4 dynamic models for the last 24 months of the sample. By analyzing these figures we were able to depict some common trends which help us to understand parts of the results.

The DCCoi model does not outperform the DCC model in terms of basis variance reduction because it overestimates the optimal hedge ratios. Optimal hedge ratios calculated using DCCoi model are higher than the ones calculated using DCC model for all commodities except cotton.

The DBEKKoi model is not outperforming the DBEKK model because the forecasted hedge ratios tend to be lower than the ones obtained from DBEKK model for 8 of the commodities. However, for corn and wheat, DBEKKoi model gives higher optimal hedge ratios than DBEKK model as shown in Figure 3.4, this explains why it outperforms the DBEKK model.

The DBEKK model forecasted hedge ratios tend to be lower than the ones obtained from the DCC model for all commodities except sugar, soybean oil and coffee.

Furthermore, Figures 3.4 to 3.7 show that optimal hedge ratios calculated using static or dynamic models are very close to 1 for all commodities. Hence, we extend our analysis by calculating the out-of-sample variance of the basis for the case where optimal hedge ratio is set to be equal to 1. Table 3.22 compares the results for the basis variance for the perfect hedge situation, where

$h^* = 1$, the naked position where the variance is simply the variance of the spot prices for the last 24 months in the sample, the basis variance calculated using optimal hedge ratios estimated from OLS and the basis variance calculated using the estimated optimal hedge ratios of the best performing model in terms of basis variance reduction. As indicated by the results, it is always better to hedge than to stay naked and support the risk of the variance in spot prices. Moreover, we notice that setting the optimal hedge ratio to be equal to 1 provides the smallest basis variance for 6 commodities, soybean oil, corn, oats, soybeans, wheat and coffee. For gold and cocoa, the OLS model is the best performing model, for cotton it is the VECM model. However, when looking at the estimated optimal hedge ratios they are equal to 0.996 for gold, 1.03 for cocoa and 1.02 [6] for cotton. Hence they are almost equal to one. Finally, for sugar and feeder cattle, $h^* = 1$ does not outperform the optimal hedge ratios calculated using the DCC model. The reason is that for those 2 commodities the optimal hedge ratio is smaller than 1, it is on average equal to 0.91 for FC and 0.83 for SB.

Our out-of-sample hedging exercise empirical results are summarized as follows:

- Static models present the best out-of-sample hedging performance for 5 of the commodities.
- The DCC model presents the smallest basis variance for 4 of the commodities.
- The DBEKKoi model performs the best in terms of the basis variance reduction for corn and wheat.
- The DCCoi optimal hedge ratios are higher than the ones estimated using the DCC model.
- The DBEKKoi model generates lower optimal hedge ratios compared to the DCC model for most of the commodities.
- The DBEKKoi estimated hedge ratios are lower than the ones calculated using the DBEKK model except for 4 commodities.
- Our empirical results comparing the variance of the basis for the two extreme cases ($h^* = 1$ and $h^* = 0$) to the basis variance of the best performing model and the one calculated using OLS suggest that it is worth hedging and that the samples for the commodities considered in this research have an optimal hedge ratio very close to 1.

[6] Results for optimal hedge ratios calculated for the last 24 months for all the models are available upon request

3.5 Conclusion

The present study is concerned with estimating optimal hedge ratios and evaluating the hedging effectiveness of six alternative modeling frameworks. The OLS based model, the VECM model, the DCC model, the DBEKK model and two more sophisticated models, the DCCoi model and DBEKKoi model that incorporate the effects of the growth rate of open interest on the commodity market's volatility and co-movements.

Estimation results of the conditional variance equations for the DCCoi model indicate that volatilities of the spot and futures commodity returns are positively affected by the growth rate of open interest. The estimated parameters for the growth rate of open interest are all statistically significant at the 95% significance level except for spot price variances of corn and futures price variances for corn, cotton and coffee. In addition, the growth rate of open interest parameter added to the conditional correlation equation of the DCCoi model is significant for all commodities at the 95% significant level, except for soybean oil, wheat and cocoa. Furthermore, it has a positive impact on the correlation between the spot and futures commodity returns. Similar results were obtained for the DBEKKoi model where the estimated parameters for the growth rate of open interest are all statistically significant at the 95% significance level except for corn, soybeans and coffee. These results support our results in the previous chapter were the growth rate of open interest played a role in determining the variance of spot commodity prices.

Next, our in-sample hedging empirical results demonstrate the superiority of the in-sample dynamic time-varying hedge ratios calculated using DCCoi model compared to the alternative DCC model. Furthermore, the DCCoi model provides better in-sample hedging effectiveness for nine of the commodities compared to the traditional DCC model. However, when comparing DBEKK and DBEKKoi models we could not find an empirical evidence to support the claim that adding the growth rate of open interest to both the conditional variance and conditional covariance equations will improve the in-sample hedging effectiveness of the model. These results highlight the fact that adding the growth rate of open interest to our dynamic models could improve the hedging performance of these models and hence might be considered when deciding on the choice of the dynamic model used to calculate optimal hedge ratios.

Moreover, results for out-of-sample hedging exercise indicate that static models present the best results in terms of the reduction of the basis variance for 5 commodities. The DCC model gives the smallest variance for 4 of the commodities. Finally, for dynamic models with growth rate of open interest, the DBEKKoi model performs the best in terms of basis variance reduction for only 2 commodities corn and wheat. This result indicates that the growth rate of open interest could be considered as a factor affecting the volatility and co-movement of spot and futures commodity returns when searching for the most suitable dynamic hedging strategy.

In addition, a graphical comparison of the out-of-sample optimal hedge ratios of our dynamic and static models reveals some common trends which help us to interpret our empirical results. The optimal hedge ratios calculated using the DCCoi model are always higher than the ones calculated using the DCC model except for cotton. The DBEKKoi model tends to underestimate the optimal hedge ratios compared to the DBEKK model. Furthermore, optimal hedge ratios calculated using the DBEKK model are lower than the ones calculated using the DCC model for most of the commodities.

Finally, our out-of sample hedging exercise comparing the variance of the basis for the two extreme cases, where the optimal hedge ratio is equal to 1 and 0 consecutively to the basis variance of the best performing model and the one calculated using OLS optimal hedge ratio suggests that it is worth hedging. The naked position presented a very high variance, higher than the variance of all the other models considered in this research. Moreover, setting the optimal hedge ratio to be equal to 1 shows that the perfect hedge presented the smallest basis variance for 6 of the commodities. Static models performs the best in terms of basis variance reduction for 3 commodities, however, the optimal hedge ratios calculated using these static models are very close to 1. Even though these results are based on a sample, they open the door for a more elaborated hedging exercise where the possibility of fixing the optimal hedge ratio for most commodities to be equal to 1 should be considered when deciding on the choice of the best optimal hedge ratio.

These findings provide important implications for futures hedging and highlight the fact that the use of static models to determine the optimal hedge ratio is more effective than the use of dynamic hedge ratio models and provides better results in terms of the basis variance reduction for most of the commodities. Furthermore, the highly volatile nature of the dynamic hedge ratios and the high transactions costs associated with the fact that the hedged portfolio must be rebalanced on a period by period basis have significant implications when deciding on the choice of the model to estimate the optimal hedge ratio.

Table 3.1: Trading Characteristics of Commodity Prices Data Obtained form CRB

Description	Exchange	Contract size	Contract month
Grains and Oilseeds			
C : Corn/No.2 Yellow	CBOT	5,000 bu.	F,H,K,N,U,X,Z
O : Oats/N0.2 White heavy	CBOT	5,000 bu.	H,K,N,U,Z
S : Soybean/No.1 Yellow	CBOT	5,000 bu.	F,H,K,N,Q,U,X
BO : Soybean Oil/Crude	CBOT	60,000 lb.	F,H,K,N,Q,U,V,Z
W : Wheat/No.2 Soft red	CBOT	5,000 bu.	H,K,N,U,Z
Industrials			
CT : Cotton/1-1/16”	NYBOT	50,000 lbs	H,K,N,V,Z
Foodstuffs			
KC : Coffee (C)/ Columbian	NYBOT	37,500 lbs.	H,K,N,U,Z

Table 3.2: Trading Characteristics of Commodity Prices Data Retrieved from Bloomberg

Description	Exchange	Contract size	Contract month
Metals			
GC : Gold	CMX	100 troy ounces	G,J,M,Q,V,Z
Foodstuffs			
CC : Cocoa/world benchmark for CC contracts	NYB-ICE	10 metric tons	H,K,N,U,Z
SB : Sugar No.11/World raw	NYB-ICE	112,000 lbs.	H,K,N,V
Live stock and Meats			
FC : Cattle, feeder	CME	50,000 lbs.	F,H,J,K,Q,U,V,X

These tables provide a brief description about each commodity trading characteristics. The first column presents the symbol description and the second one lists the futures exchange where the commodity is traded. CME refers to Chicago Mercantile Exchange and Chicago Board of Trade, CMX is an abbreviation for COMEX, a commodities exchange in the CME Group Inc., and NYBOT or NYB-ICE: New York Board of Trade, renamed ICE Futures US in September 2007. The third column reveals the contract size and the last column provides the contract months denoted by: F = January, G = February, H = March, J = April, K = May, M = June, N= July, Q = August, U = September, V = October, X = November and Z = December.

Table 3.3: Johansen MLE Estimates, Trace statistic

Commodity	Trace Statistic	Crit 95%
BO	17.174 *	15.494
	0.021	3.841
C	31.812 *	15.494
	0.803	3.841
CT	38.377 *	15.494
	5.317	3.841
O	24.668*	15.494
	0.449	3.841
S	43.271 *	15.494
	1.090	3.841
W	21.350 *	15.494
	0.630	3.841
KC	26.477 *	15.494
	3.389	3.841
GC	20.102 *	15.494
	3.134	3.841
CC	21.542*	15.494
	2.639	3.841
C	26.671*	15.494
	0.513	3.841
SB	20.916 *	15.494
	3.510	3.841
FC	39.794 *	15.494
	0.321	3.841

Table 3.4: Johansen MLE Estimates, Maximum Eigenvalue Statistic

Commodity	Eigenvalue statistic	Crit 95%
BO	17.153*	14.264
	0.021	3.841
C	31.009 *	14.264
	0.803	3.841
CT	33.060 *	14.264
	5.317	3.841
O	24.219*	14.264
	0.449	3.841
S	42.181 *	14.264
	1.090	3.841
W	20.720*	14.264
	0.630	3.841
KC	23.088 *	14.264
	3.389	3.841
GC	19.967 *	14.264
	3.134	3.841
CC	18.902 *	14.264
	2.639	3.841
CL	26.671 *	14.264
	0.513	3.841
SB	17.406 *	14.264
	3.510	3.841
FC	39.794 *	14.264
	0.321	3.841

These tables summarize the results for the trace statistic and the maximal eigenvalue statistic for the Johansen's cointegration test. The * denotes that the Null hypothesis of no cointegrating vector is rejected. $H_0 : r <= 0$ versus $r <= 1$.

Table 3.5: Parameter Estimates of the VECM Model of the Mean Equations

Commodity	η	C	κ_s	κ_f
BO	-1.009	9.53E-05	-6.38	-4.75
	(-96.24*)		(-2.70*)	(-2.05*)
C	-1.098	0.00063	-2.207	-0.52
	(-47.27*)		(-3.906*)	(-0.99)
CT	-1.006	0.000759	-3.59	-1.38
	(-32.94*)		(-7.78*)	(-3.16*)
O	-1.035	-8.25E-05	-1.68	0.17
	(-24.68*)		(-3.59*)	(0.41)
S	-1.0385	0.00018	-3.84	-1.86
	(-67.85*)		(-3.65*)	(-1.78)
W	-1.11	0.00039	-1.06	0.54
	(-38.30*)		(-1.62)	(0.9)
KC	-1.101	-4.6E-06	-0.759	1.137
	(-89.98*)		(-0.74)	(1.13)
GC	-1.007	3.98E-05	30.64	31.93
	(-303.32*)		(3.65*)	(3.8*)
CC	-1.102	-0.00014	-4.89	-1.99
	(-83.84*)		(-1.5)	(-0.83)
SB	-1.001	-.00016	-3.12	-0.84
	(-37.44*)		(-4.34*)	(-1.20)
FC	-1.00113	0.00013	-5.14	-2.48
	(-59.60*)		(-7.21*)	(-3.32*)

This table shows the estimates of the parameters of the cointegration equation Z_{t-1}, η and C from (3.14). It also shows the results of the estimates of parameters κ_s and κ_f from (3.12) and (3.13). The t-ratios are presented in parenthesis and show the relative significance at the 95% level. The statistically significant coefficients are marked with a *.

Table 3.6: Spot Variance Returns Equation Parameters and t-test Values of the DCC Model

Commodity	ω_s	γ_s	θ_s
BO	0.0029	0.2702	0.1927
	(46643.03)	(23415.27)	(357258.88)
C	0.0005	0.9002	0.0304
	(2.86)	(23.38)	(18.63)
CT	0.0006	0.8269	0.0900
	(4.63)	(90.60)	(7.24)
O	0.0019	0.7074	0.0405
	(10.34)	(36.85)	(3.18)
S	0.0000	0.9533	0.0467
	(35.01)	(1070.82)	(37.67)
W	0.0049	0.1487	0.1225
	(9487.40)	(88.45)	(67.95)
KC	0.0018	0.7712	0.1233
	(239.91)	(9535.44)	(289.35)
GC	0.0001	0.8147	0.1523
	(1.43)	(29.77)	(3.26)
CC	0.0017	0.6502	0.1648
	(91976.74)	(371844.21)	(444839.53)
SB	0.0011	0.6690	0.2011
	(13.10)	(88.03)	(143.35)
FC	0.0003	0.8487	0.0079
	(1.22)	(25.24)	(1.51)

Table 3.7: Future Variance Returns Equation Parameters and t-test Values of the DCC Model

Commodity	ω_f	γ_f	θ_f
BO	0.0031	0.2148	0.2009
	(10852.08)	(3881.49)	(91400.26)
C	0.0006	0.8382	0.0573
	(3.10)	(22.62)	(7.01)
CT	0.0006	0.8098	0.1012
	(0.59)	(4.49)	(4.51)
O	0.0012	0.7389	0.0695
	(7.88)	(41.04)	(5.02)
S	0.0000	0.9677	0.0323
	(34.58)	(1318.35)	(30.96)
W	0.0037	0.1796	0.1632
	(1089.58)	(299.59)	(115.21)
KC	0.0016	0.7730	0.1334
	(238.16)	(229652.02)	(285.31)
GC	0.0001	0.8160	0.1505
	(2.94)	(60.24)	(10.73)
CC	0.0007	0.7970	0.1231
	(209.53)	(1170.33)	(424.08)
SB	0.0006	0.7507	0.1825
	(45.70)	(568.94)	(63.25)
FC	0.0003	0.8501	0.0224
	(1.27)	(25.04)	(5.96)

These tables summarize the parameter estimations results for the spot and future variance equations of the commodity returns for the DCC model over the whole sample period, equations (3.23) and (3.24). The numbers in parenthesis indicate the t-score values calculated using HAC variance covariance estimation method.

Table 3.8: Spot Variance Returns Equation Parameters and t-test Values of the DCCoi Model

Commodity	ω_s	γ_s	θ_s	τ_s
BO	0.0033	0.2019	0.1933	0.0054
	(361.47)	(164.63)	(816.66)	(24.41)*
C	0.0005	0.9031	0.0287	0.0000
	(0.32)	(3.07)	(0.53)	(0.02)
CT	0.0007	0.7795	0.1166	0.0038
	(5.45)	(104.24)	(8.80)	(1384.00)*
O	0.0010	0.8019	0.0674	0.0047
	(27.06)	(124.70)	(31.50)	(1921.91)*
S	0.0000	0.9441	0.0547	0.0010
	(0.13)	(16.37)	(19.03)	(5.24)*
W	0.0043	0.2023	0.1473	0.0070
	(936.59)	(16.40)	(49.26)	(38.15)*
KC	0.0018	0.7717	0.1233	0.0001
	(8.79)	(43.20)	(40.24)	(2.19)*
GC	0.0001	0.6865	0.2026	0.0001
	(0.33)	(24.51)	(6.81)	(3.18) *
CC	0.0014	0.6783	0.1552	0.0004
	(1498.76)	(3785.58)	(2691.32)	(494.35)*
SB	0.0012	0.6497	0.2005	0.0005
	(1.51)	(7.62)	(50.06)	(2.17)*
FC	0.0002	0.8609	0.0152	0.0003
	(1.71)	(26.93)	(0.29)	(18.03)*

This table summarizes the in-sample parameter estimations results for the spot variance equation of the commodity returns for the DCCoi model, equation (3.30). The numbers in parenthesis indicate the t-score values calculated using HAC variance covariance estimation method. The * indicates that the parameters for the growth rate of open interest are statistically significant at the 95 % significance level. The * is used only to indicate the statistical significance of the growth rate of open interest parameters.

Table 3.9: Future Variance Returns Equation Parameters and t-test Values of the DCCoi Model

Commodity	ω_f	γ_f	θ_s	τ_f
BO	0.0033	0.1643	0.1993	0.0062
	(77.87)	(19.49)	(270.80)	(788.05)*
C	0.0006	0.8388	0.0570	0.0000
	(1.85)	(21.97)	(2.41)	(0.00)
CT	0.0006	0.7687	0.1270	0.0028
	(0.37)	(2.54)	(1.64)	(0.86)
O	0.0007	0.8169	0.0720	0.0034
	(2.48)	(10.47)	(9.73)	(24.75)*
S	0.0000	0.9636	0.0357	0.0005
	(0.46)	(53.45)	(4.54)	(14.74)*
W	0.0031	0.2360	0.2205	0.0062
	(55.52)	(19.67)	(59.27)	(322.82)*
KC	0.0016	0.7733	0.1334	0.0001
	(6.96)	(25.99)	(193.28)	(0.43)
GC	0.0001	0.6787	0.2007	0.0001
	(0.46)	(27.72)	(12.55)	(10.06)*
CC	0.0006	0.7914	0.1273	0.0005
	(233547.26)	(47849.19)	(13296.50)	(1425.65)*
SB	0.0005	0.7446	0.1678	0.0009
	(3.47)	(37.73)	(19.06)	(2.42)*
FC	0.0002	0.8139	0.0283	0.0004
	(1.90)	(31.77)	(1.19)	(3.59)*

This table summarizes the in-sample parameter estimations results for the future variance equation of the commodity returns for the DCCoi model, equation (3.31). The numbers in parenthesis indicate the t-score values calculated using HAC variance covariance estimation method. The * indicates that the parameters for the growth rate of open interest are statistically significant at the 95 % significance level. The * is used only to indicate the statistical significance of the growth rate of open interest parameters.

Table 3.10: Correlation Equation parameters and t-test Values of the DCC Model

Commodity	δ_1	δ_2
BO	0.0255	0.7348
	(0.66)	(4.57)
C	0.1208	0.8629
	(5.21)	(29.27)
CT	0.2030	0.3508
	(1.58)	(1.71)
O	0.1764	0.2783
	(2.68)	(1.80)
S	0.0351	0.7774
	(0.89)	(6.06)
W	0.1025	0.8862
	(3.97)	(27.88)
KC	0.1481	0.7662
	(3.57)	(9.05)
GC	0.1208	0.8619
	(2.18)	(13.23)
CC	0.0356	0.7208
	(1.56)	(3.72)
SB	0.0134	0.9766
	(1.41)	(55.19)
FC	0.0457	0.9498
	(3.50)	(42.24)

Table 3.11: Correlation Equation Parameters and t-test Values of the DCCoi Model

Commodity	δ_1	δ_2	δ_3
BO	0.0308	0.6722	0.0000
	(5.39)	(12.42)	(0.00)
C	0.1217	0.8607	0.0176
	(75.73)	(414.46)	(4.8536)*
CT	0.2251	0.4226	0.3523
	(27.82)	(57.93)	(6.03)*
O	0.1790	0.3524	0.4685
	(40.54)	(26.39)	(14.74)*
S	0.0396	0.7081	0.2523
	(18.01)	(21.90)	(3.75)*
W	0.1030	0.8841	0.0129
	(59.95)	(414.18)	(1.71)
KC	0.1503	0.7588	0.0909
	(49.21)	(104.12)	(2.58)*
GC	0.1340	0.8219	0.0137
	(23.30)	(120.67)	(12.57)*
CC	0.0369	0.7256	0.0000
	(22.74)	(43.58)	(0.00)
SB	0.0101	0.9824	0.0075
	(13.93)	(423.64)	(13.04)*
FC	0.0490	0.9428	0.0082
	(37.17)	(317.05)	(3.64)*

These tables summarize the in-sample parameter estimations results for equations (3.25) and (3.32). The t-score values in parenthesis are calculated using HAC variance covariance estimation method. The * indicates that the parameters for the growth rate of open interest are statistically significant at the 95% significance level. The * is used only to indicate the statistical significance of the growth rate of open interest parameters.

Table 3.12: Parameter Estimates of the DBEKK Model, CRB Data

Parameters	BO	C	CT	O	S	W	KC
c_{11}	0.0398	0.0584	0.0602	0.0320	0.0316	0.0175	0.0416
	(948.23)	(1737.93)	(1076.88)	(38.39)	(507.35)	(588.34)	(641.10)
c_{12}	0.0393	0.0572	0.0425	0.0399	0.0316	0.0171	0.0368
	(1070.01)	(1719.15)	(834.29)	(31.69)	(511.05)	(706.85)	(662.94)
c_{22}	0.0000	0.0047	0.0000	-0.0000	0.0032	-0.0026	-0.0000
	(0.98)	(139.58)	(3957.05)	(-4209.04)	(494.62)	(-2322.79)	(-118.52)
a_{11}	0.1474	0.4010	0.3592	0.3866	0.2792	0.2781	0.3447
	(26.68)	(84.93)	(46.11)	(13.10)	(15.87)	(69.71)	(125.43)
a_{22}	0.0978	0.5339	0.4351	0.3899	0.3324	0.2890	0.3285
	(19.26)	(68.41)	(46.63)	(41.98)	(31.01)	(51.24)	(126.82)
b_{11}	0.8265	0.5916	0.6107	0.8668	0.8706	0.9451	0.8900
	(225.01)	(134.42)	(47.75)	(23.68)	(232.01)	(1925.27)	(1357.22)
b_{22}	0.8313	0.4978	0.7547	0.7884	0.8578	0.9403	0.9061
	(289.35)	(70.80)	(213.78)	(8.63)	(281.14)	(1452.31)	(1916.02)

Table 3.13: Parameter Estimates of the DBEKK Model, Bloomberg Data

Parameters	GC	CC	SB	FC
c_{11}	0.0123	0.0411	0.0554	0.0135
	(621.17)	(100.77)	(248.49)	(616.69)
c_{12}	0.0124	0.0419	0.0398	0.0142
	(630.19)	(47.53)	(230.82)	(778.34)
c_{22}	-0.0002	0.0029	0.0115	0.0044
	(-15861.86)	(9.12)	(342.55)	(2121.78)
a_{11}	0.3560	0.2453	0.5542	0.2097
	(43.88)	(3.00)	(74.00)	(77.17)
a_{22}	0.3581	0.2869	0.5035	0.2482
	(43.36)	(6.98)	(68.86)	(86.41)
b_{11}	0.9039	0.8602	0.6001	0.9261
	(337.05)	(261.92)	(22.41)	(662.89)
b_{22}	0.9029	0.8304	0.7271	0.9143
	(332.42)	(49.11)	(54.38)	(634.84)

The tables above show the parameter estimates for the DBEKK model for equations (3.16), (3.17) and (3.18) over the whole sample period.

Table 3.14: Parameter Estimates of the DBEKKoi model, CRB Data

Parameters	BO	C	CT	O	S	W	KC
c_{11}	0.0399	0.0584	0.0600	0.0272	0.0316	0.0160	0.0420
	(964.81)	(1737.87)	(1103.04)	(14.46)	(502.81)	(322.47)	(279.19)
c_{12}	0.0393	0.0572	0.0423	0.0340	0.0316	0.0163	0.0371
	(1071.41)	(1716.97)	(855.22)	(11.01)	(506.93)	(567.49)	(336.91)
c_{22}	-0.0000	0.0047	0.0000	0.0000	0.0032	0.0021	0.0000
	(-0.33)	(138.59)	(49.32)	(1.31)	(495.42)	(279.72)	(0.37)
a_{11}	0.1499	0.4011	0.3580	0.3579	0.2792	0.2731	0.3480
	(27.35)	(84.71)	(44.18)	(4.82)	(15.90)	(45.47)	(38.94)
a_{22}	0.1001	0.5339	0.4345	0.3735	0.3324	0.2915	0.3322
	(19.28)	(68.40)	(46.80)	(10.71)	(31.08)	(44.58)	(34.07)
b_{11}	0.8252	0.5916	0.6140	0.8955	0.8706	0.9471	0.8877
	(249.51)	(134.49)	(51.27)	(14.32)	(231.12)	(1185.11)	(206.33)
b_{22}	0.8302	0.4978	0.7561	0.8359	0.8578	0.9393	0.9039
	(311.70)	(70.84)	(223.49)	(4.88)	(279.51)	(1185.22)	(250.03)
d_{11}	0.0228	0.0000	-0.0061	0.0407	0.0000	0.0572	0.0217
	(4.38)*	(4.20)*	(-6.71)*	(9.29)*	(0.33)	(63.01)*	(0.24)
d_{22}	0.0220	0.0000	-0.0131	0.0348	0.0000	0.0535	0.0255
	(4.49) *	* (3.92)*	(-2.80)*	(10.40)*	(0.33)	(63.88)*	(0.22)

Table 3.15: Parameter Estimates of the DBEKKoi Model, Bloomberg Data

Parameters	GC	CC	SB	FC
c_{11}	-0.0007	0.0763	0.0595	0.0131
	(-80.44)	(1342.93)	(212.59)	(312.85)
c_{12}	-0.0003	0.0689	0.0411	0.0135
	(-44.02)	(1366.38)	(197.19)	(377.25)
c_{22}	-0.0001	0.0054	0.0076	0.0046
	(-9757.51)	(343.72)	(65.13)	(1657.03)
a_{11}	0.3607	0.2959	0.5536	0.2136
	(81.64)	(74.56)	(73.32)	(65.80)
a_{22}	0.3631	0.3236	0.5040	0.2503
	(79.87)	(76.75)	(90.68)	(62.58)
b_{11}	0.8802	0.4524	0.5022	0.9141
	(485.05)	(50.67)	(14.44)	(319.75)
b_{22}	0.8792	0.5183	0.6737	0.9083
	(478.34)	(92.05)	(55.18)	(335.45)
d_{11}	0.0106	0.0295	0.0527	0.0141
	(1829.32)*	(314.50)*	(189.05)*	(443.16)*
d_{22}	0.0107	0.0164	0.0545	0.0128
	(1817.63) *	(183.39)*	(214.31)*	(402.98)*

The tables above show the in-sample parameter estimates for equations (3.27), (3.28) and (3.29). The * indicates that the parameters for the growth rate of open interest are statistically significant at the 95 % level. The * is used only to indicate the statistical significance of the growth rate of open interest parameters.

Table 3.16: In-Sample Hedge Ratio Estimates of the DCC and the DCCoi Models

Commodity	HR DCC	HR DCCoi H
BO	1.0167	1.0167
C	1.0232	1.0233*
CT	0.9381	0.9461 *
O	1.0113	1.0143 *
S	0.9702	0.9703*
W	1.0464	1.0544*
KC	1.0019	1.0020*
GC	0.9955	0.9976*
CC	1.0409	1.0410*
SB	0.9649	0.9736*
FC	0.8990	0.9046*

This table shows the estimates of the in-sample optimal hedge ratios of the DCC and the DCCoi models. The numbers presented in the table represent the averages of the estimated hedge ratios over the whole sample period. Hedge ratios are calculated using equation (3.26).

Table 3.17: In-Sample Hedging Effectiveness of the DCC and the DCCoi Models

Commodity	HE DCC	HE DCCoi
BO	99.2990*	99.2627
C	92.8064	92.8126 *
CT	79.5396	80.6189*
O	81.4760	81.5195 *
S	96.5407	96.5572 *
W	91.8363*	91.4854
KC	97.4525	97.4778*
GC	99.9161	99.9162 *
CC	97.7632	97.8255*
SB	82.6995	83.6952*
FC	88.3708	88.9770*

This table displays the estimated in-sample hedging effectiveness of the DCC and the DCCoi models. Hedging effectiveness is calculated using equation (3.4) over the whole sample period. The numbers presented in this table denotes the mean of the hedging effectiveness calculated using the DCC and the DCCoi models. The * shows the model with the highest hedging effectiveness.

Table 3.18: In-Sample Hedge Ratio Estimates of the DBEKK and the DBEKKoi Models

Commodity	HR DBEKKoi	HR DBEKK
BO	1.015	1.0149
C	1.024	1.024
CT	0.9582	0.9581
O	1.0418	1.0402
S	0.9829	0.9829
W	1.0465	1.0481
KC	1.007	1.007
GC	0.9982	0.998
CC	1.0387	1.0383
SB	0.99	0.9884
FC	0.8897	0.8926

This table shows the in-sample optimal hedge ratios estimates of the DBEKK and DBEKKoi models. Optimal hedge ratios are calculated using equation (3.19). The numbers in the table represent the means of the optimal hedge ratios calculated over the whole sample period.

Table 3.19: In-Sample Hedging Effectiveness of the DBEKK and the DBEKKoi Models

Commodity	HE DBEKKoi	HE DBEKK
BO	99.2495*	99.2486
C	95.8048	95.8048
CT	84.7775	84.7842*
O	83.8821	83.936*
S	97.5998	97.5998
W	94.0833	94.1721*
KC	98.0141	98.0182*
GC	99.9243	99.9321*
CC	97.9727*	97.9411
SB	83.8215	83.5636*
FC	89.6487	89.6686*

This table indicates the in-sample hedging effectiveness of the DBEKK and DBEKKoi models. Hedging effectiveness is calculated using equation (3.4) over the whole sample period. The numbers in the table denote the averages of the hedging effectiveness. The * indicates the model with the highest mean.

Table 3.20: Comparing Out-of-Sample Basis Variance

Commodity	OLS	VECM	DBEKK	DBEKKoi	DCC	DCCoi
BO	0.22	0.11	0.14	11.87	0.09*	0.39
C	43.16	53.66	38.86	18.17*	81.70	105.40
CT	2.82	2.33*	9.62	15.27	248.91	2.60
O	43.25 *	44.02	197.57	221.79	156.78	147.21
S	64.71	87.23	965.86	6394.42	42.44*	55.74
W	118.79	182.30	64.80	26.28 *	351.57	433.57
KC	1.21	0.97 *	1.45	65.09	4.33	1.22
GC	0.36*	0.41	0.86	1100.79	6.47	70628.68
CC	1646.61*	1654.04	1825.84	46455.01	2187.13	2204.47
SB	1.31	1.39	0.53	2.51	0.48*	0.89
FC	1.29	1.77	3.01	318.72	1.25 *	4.61

This table shows the basis variance calculated for all six models and for each commodity over the last 24 months of the sample. The * indicates the smallest variance for the basis.

Table 3.21: Ranking the Models in Terms of Basis Variance Reduction

Commodity	OLS	VECM	DBEKK	DBEKKoi	DCC	DCCoi
BO	5	4	2	6	1	3
C	3	4	2	1	5	6
CT	3	1	4	5	6	2
O	1	2	5	6	4	3
S	3	4	5	6	1	2
W	3	4	2	1	5	6
KC	2	1	4	6	5	3
GC	1	2	3	5	4	6
CC	1	2	3	6	4	5
SB	4	5	2	6	1	3
FC	2	3	4	6	1	5

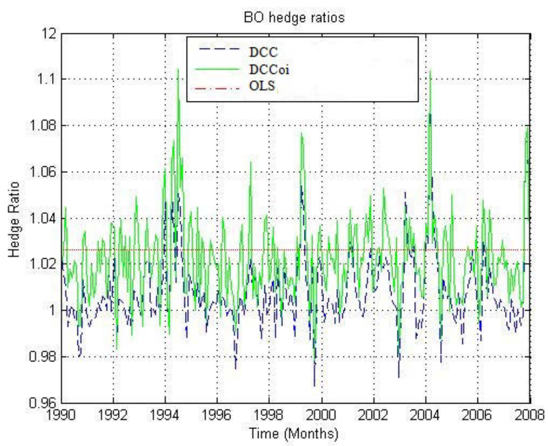
This table shows for each commodity the ranking of the models in terms of the basis variance reduction.

Table 3.22: Comparing Out-of-Sample Basis Variance for OLS, Best Performing Model and Extreme Cases Hedge Ratios

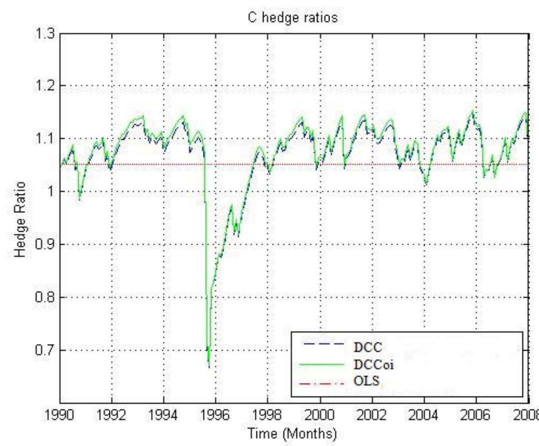
Commodity	S2	Model with Min Variance	OLS	h*=1
BO	182.01	DCC: 0.09	0.22	0.03*
C	14852.50	DBEKKoi:18.17	43.16	14.41*
CT	76.71	VECM:2.33*	2.82	2.59
O	4466.79	OLS:43.25	43.25	42.12*
S	99048.46	DCC:42.44	64.71	39.30 *
W	41725.53	DBEKKoi:26.28	118.79	22.40*
KC	235.38	VECM:0.97	1.21	0.95*
GC	14564.67	OLS:0.36*	0.36*	0.72
CC	214018.65	OLS:1646.61*	1646.61*	1812.29
SB	17.96	DCC:0.48*	1.31	0.96
FC	84.96	DCC:1.25*	1.29	1.38

This table shows the basis variance calculated for all the commodities over the last 24 months time horizon using consecutively optimal hedge ratios equal to zero (basis=S_2), optimal hedge ratios estimated from OLS and optimal hedge ratio set to 1. Moreover, results are compared to the variance of the best performing model in terms of basis variance reduction. The * indicates the smallest variance for the basis realized over the last 24 months time horizon.

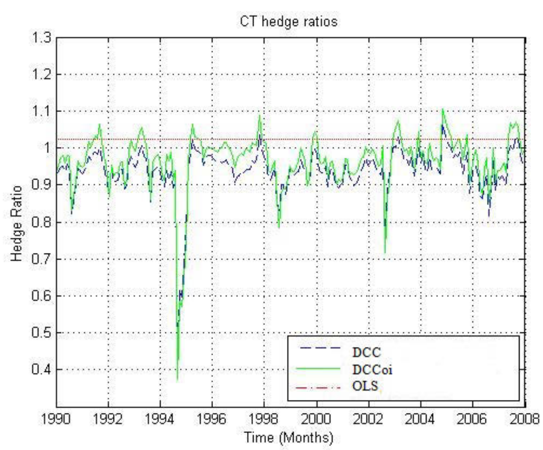
Figure 3.1: In-Sample Hedge Ratios for BO, C, CT and O for OLS, DCC and DCCoi Models



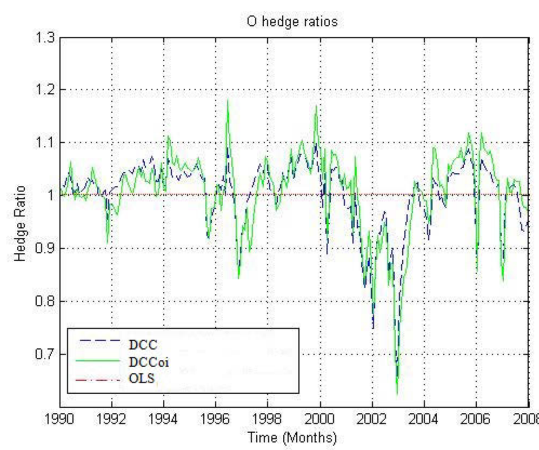
(a) BO hedge ratios



(b) C hedge ratios

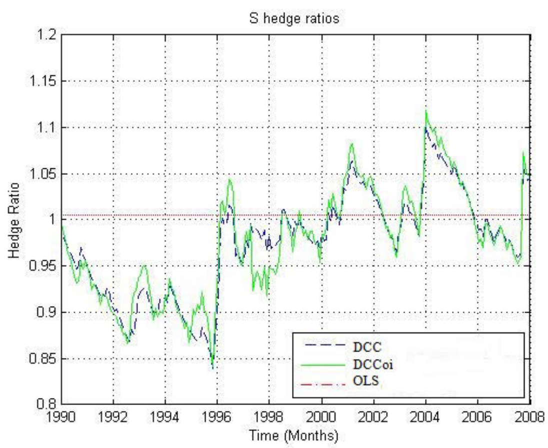


(c) CT hedge ratios

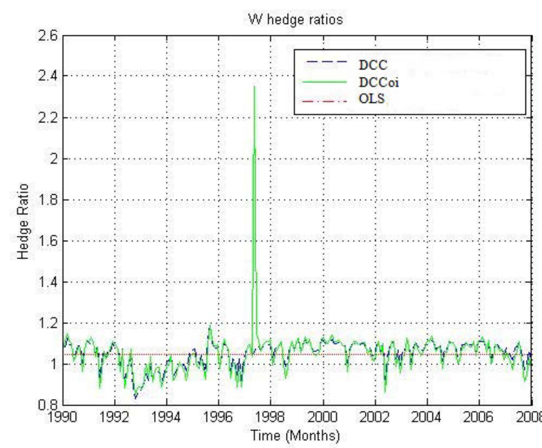


(d) O hedge ratios

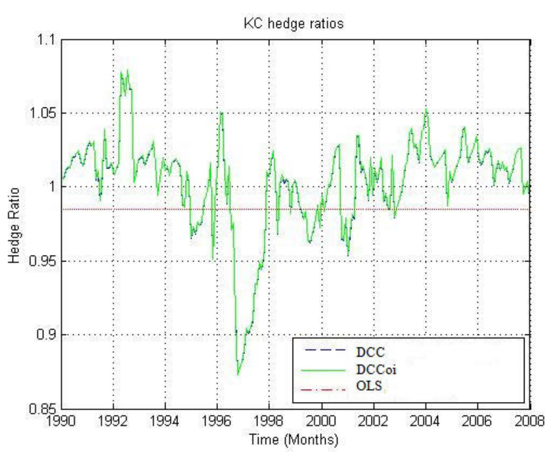
Figure 3.2: In-Sample Hedge Ratios for S, W, KC and GC for OLS, DCC and DCCoi Models



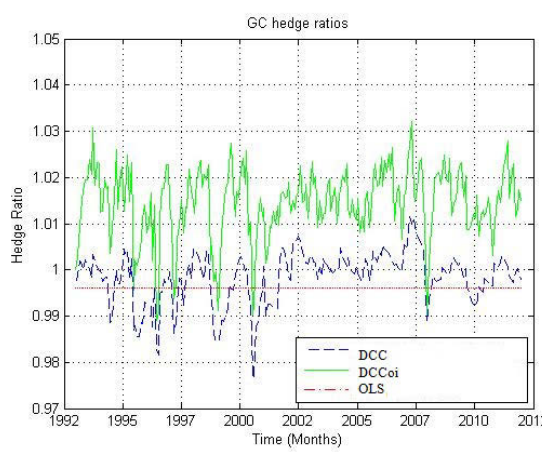
(a) S hedge ratios



(b) W hedge ratios

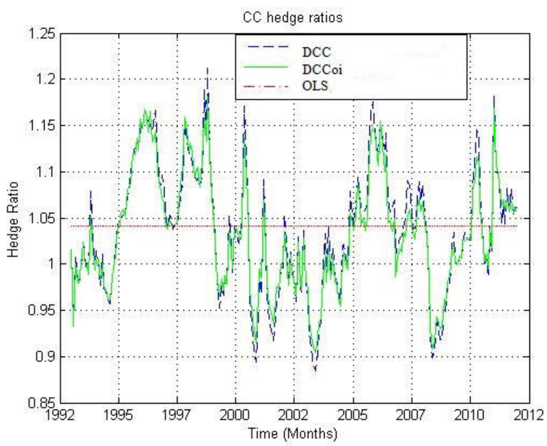


(c) KC hedge ratios

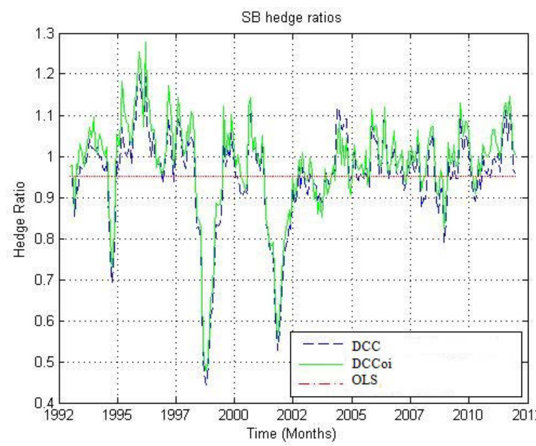


(d) GC hedge ratios

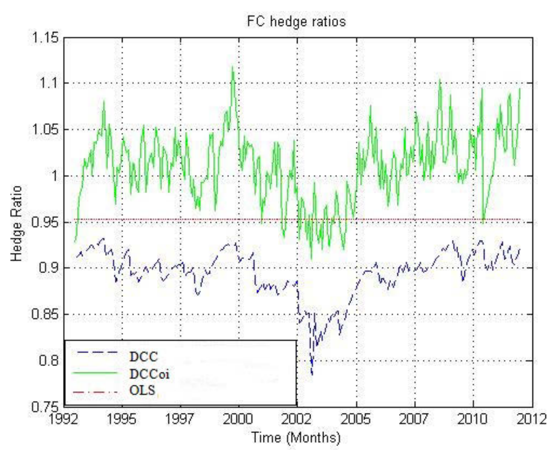
Figure 3.3: In-Sample Hedge Ratios for CC, SB and FC for OLS, DCC and DCCoi Models



(a) CC hedge ratios

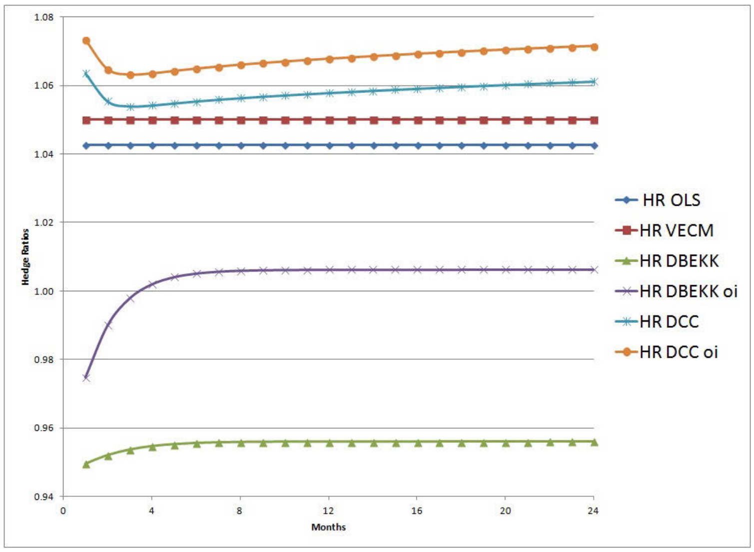


(b) SB hedge ratios

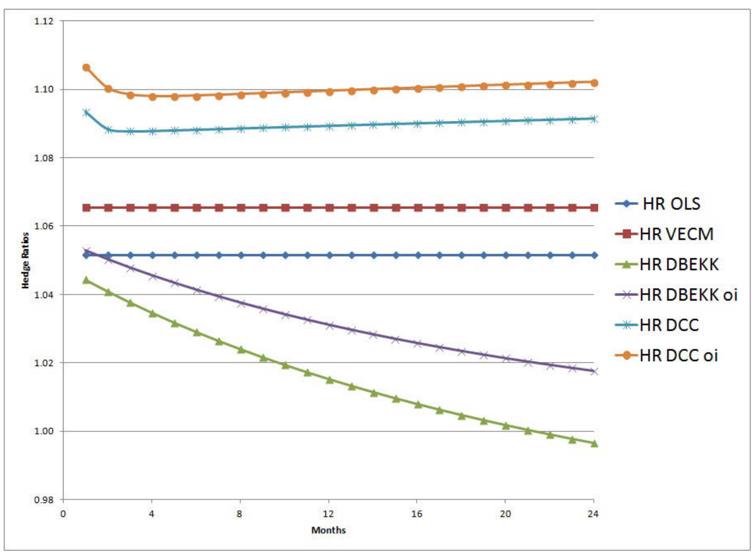


(c) FC hedge ratios

Figure 3.4: Out-of-Sample Hedge Ratios for C and W

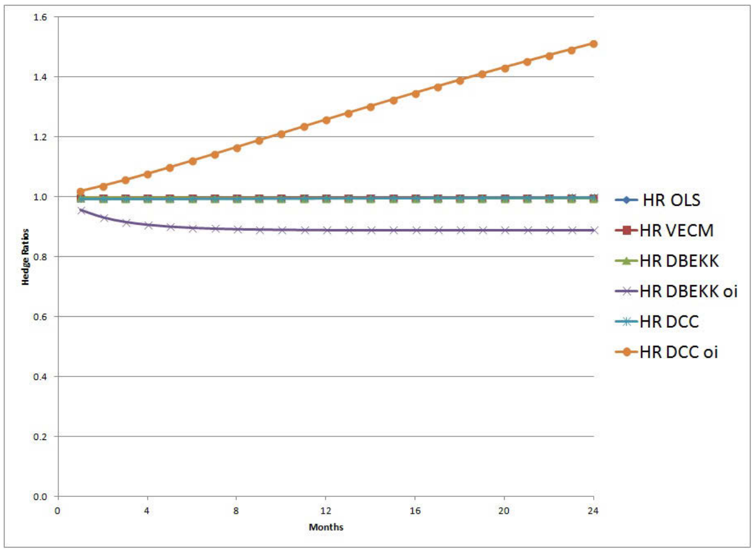


(a) C hedge ratios

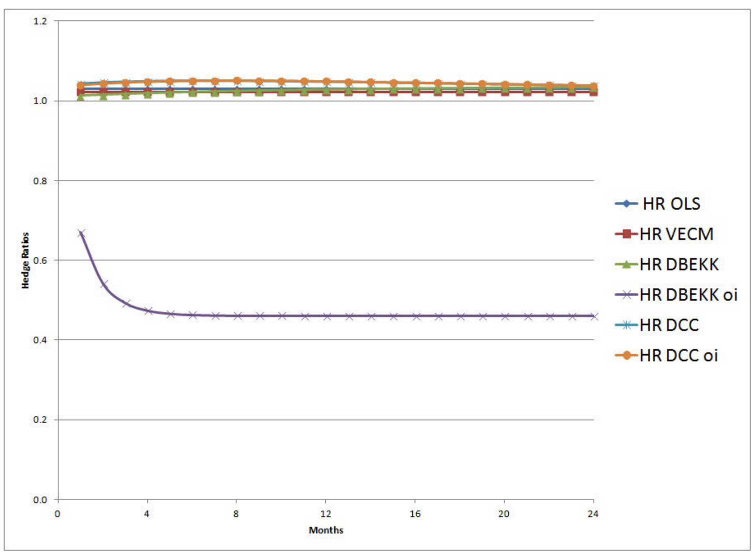


(b) W hedge ratios

Figure 3.5: Out-of-Sample Hedge Ratios for GC and CC

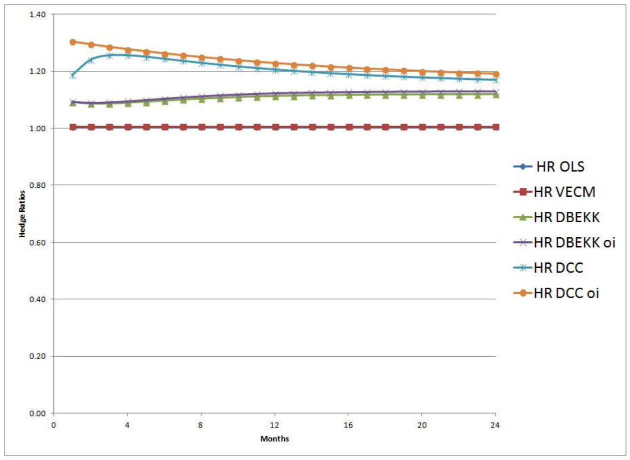


(a) GC hedge ratios

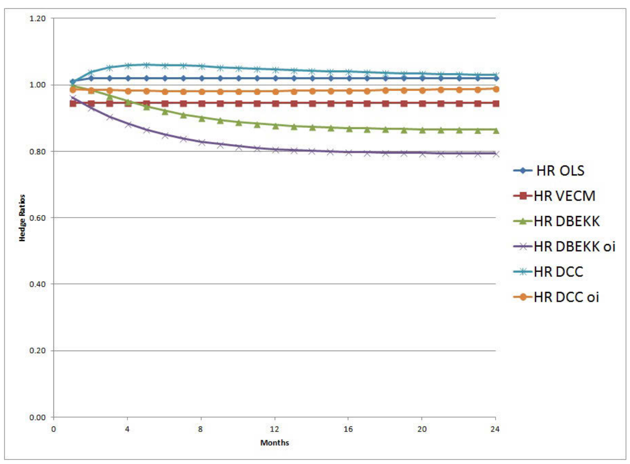


(b) CC hedge ratios

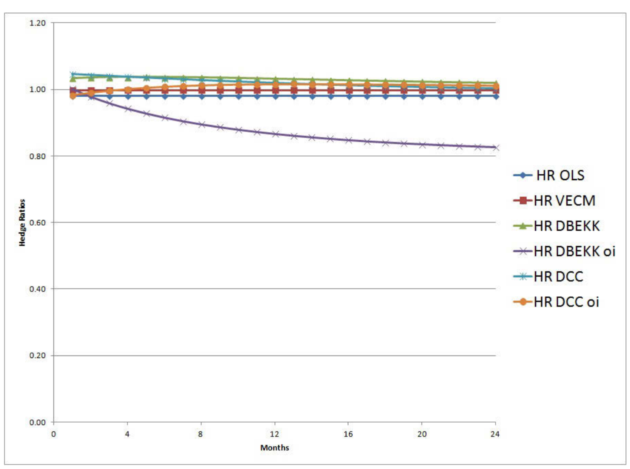
Figure 3.6: Out-of-Sample Hedge Ratios for O, CT and KC



(a) O hedge ratios

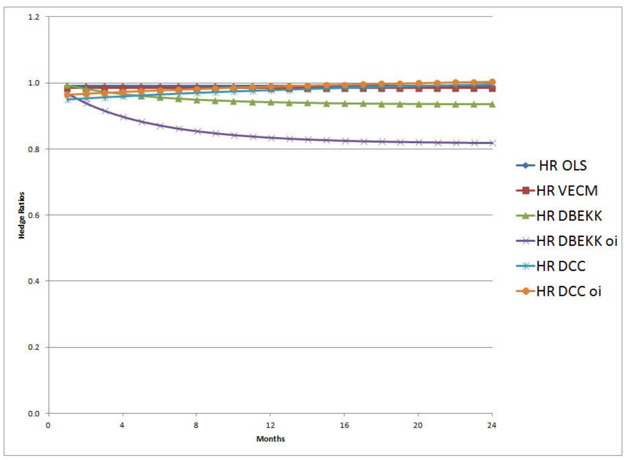


(b) CT hedge ratios

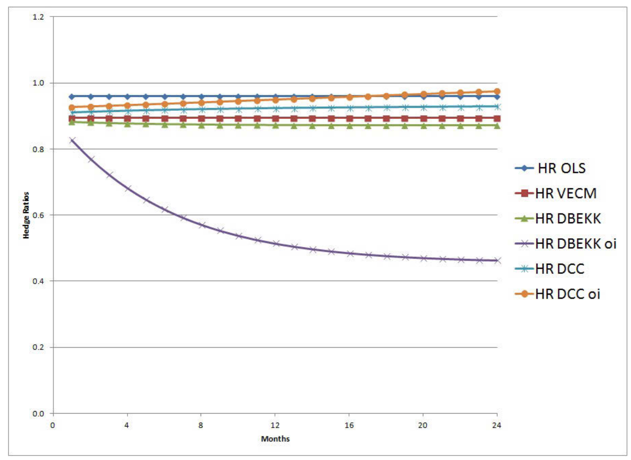


(c) KC hedge ratios

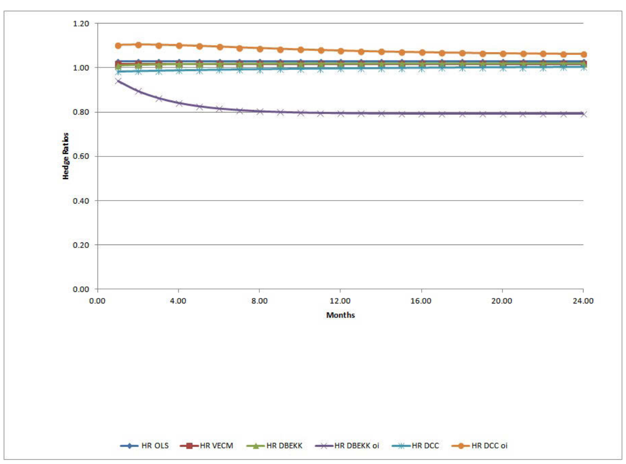
Figure 3.7: Out-of-Sample Hedge Ratios for S,FC, BO and SB



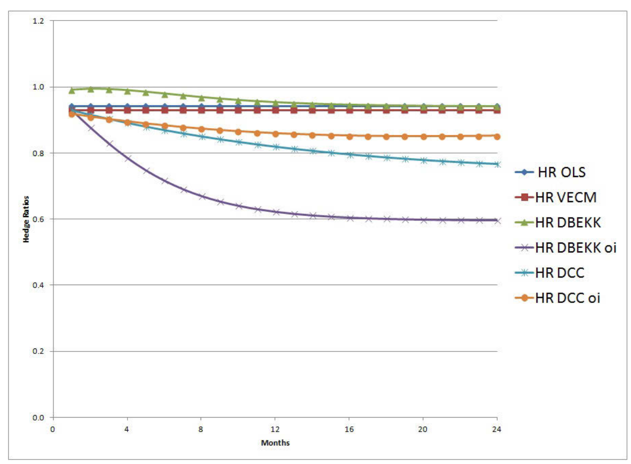
(a) S hedge ratios



(b) FC hedge ratios



(c) BO hedge ratios



(d) SB hedge ratios

References

AMAYA, D., P. CHRISTOFFERSEN, K. JACOBS, AND A. VASQUEZ (2011): "Do Realized Skewness and Kurtosis Predict the Cross-Section of Equity Returns?," CREATES Research Papers 2011-44, School of Economics and Management, University of Aarhus, `http://ideas.repec.org/p/aah/create/2011-44.html`.

ARSENEAU, D., AND S. LEDUC (2012): "Commodity Price Movements in a General Equilibrium Model of Storage," *preprint*.

BAKSHI, G., G. PANAYOTOV, AND G. SKOULAKIS (2011): "The Baltic Dry Index as a Predictor of Global Stock Returns, Commodity Returns, and Global Economic Activity," AFA 2012 Chicago Meetings Paper, `http://ssrn.com/abstract=1787757`.

BECK, S. E. (1993): "A Rational Expectations Model of Time Varying Risk Premia in Commodities Futures Markets: Theory and Evidence," *International Economic Review*, 34(1), 149–168.

BLANCHARD, O., AND S. FISHER (1989): *Lectures on Macroeconomics*, Lipsey Lectures. MIT Press, Cambridge.

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31(3), 307–327.

——— (1987): "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *The Review of Economics and Statistics*, 69(3), 542-547.

——— (1990): "Modelling the Coherence in Short-run Nominal Exchange Rates: A Multivariate Generalized ARCH Model," *The Review of Economics and Statistics*, 72(3), 498–505.

BOLLERSLEV, T., R. F. ENGLE, AND J. M. WOOLDRIDGE (1988): "A Capital Asset Pricing Model with Time-Varying Covariances," *Journal of Political Economy*, 96(1), 116–31.

BRENNAN, M. J. (1958): "The Supply of Storage," *The American Economic Review*, 48(1), 50–72.

BRENNER, R. J., R. H. HARJES, AND K. F. KRONER (1996): "Another Look at Models of the Short-Term Interest Rate," *The Journal of Financial and Quantitative Analysis*, 31(1), 85–107.

BROOKS, C., O. T. HENRY, AND G. PERSAND (2002): "The Effect of Asymmetries on Optimal Hedge Ratios," *The Journal of Business*, 75(2).

CAFIERO, C., E. S. BOBENRIETH, J. R. BOBENRIETH, AND B. D. WRIGHT (2011): "The empirical relevance of the competitive storage model," *Journal of Econometrics*, 162(1), 44 – 54.

CALVO, G. A. (1983): "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics*, 12(3), 383 – 398.

CAMPBELL, J. Y., AND J. AMMER (1993): "What Moves the Stock and Bond Markets? A Variance Decomposition for Long-Term Asset Returns," *The Journal of Finance*, 48(1), 3–37.

CAMPBELL, J. Y., AND R. J. SHILLER (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *The Review of Financial Studies*, 1(3), 195–228.

CAMPBELL, S. D., M. A. DAVIS, J. GALLIN, AND R. F. MARTIN (2009): "What moves housing markets: A variance decomposition of the rentprice ratio," *Journal of Urban Economics*, 66(2), 90 – 102.

CAPORIN, M., J.-A. JIMENEZ-MARTIN, AND L. GONZALEZ-SERRANO (2014): "Currency hedging strategies in strategic benchmarks and the global and Euro sovereign financial crises," *Journal of International Financial Markets, Institutions and Money*, 31(0), 159 – 177.

CASASSUS, J., AND P. COLLIN-DUFRESNE (2005): "Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates," *The Journal of Finance*, 60(5), 2283–2331.

CHAMBERS, M. J., AND R. E. BAILEY (1996): "A Theory of Commodity Price Fluctuations," *Journal of Political Economy*, 104(5), 924–957.

DEATON, A., AND G. LAROQUE (1992): "On the Behaviour of Commodity Prices," *The Review of Economic Studies*, 59(1), 1–23.

——— (1995): "Estimating a Nonlinear Rational Expectations Commodity Price Model with Unobservable State Variables," *Journal of Applied Econometrics*, 10(S), S9–40.

——— (1996): "Competitive Storage and Commodity Price Dynamics," *Journal of Political Economy*, 104(5), 896–923.

DE ROON, F. A., NIJMAN, T. E., AND C. VELD (2000): "Hedging Pressure Effects in Futures Markets," *The Journal of Finance*, 55(3), 1437-1456.

DEWALLY, M., AND L. MARRIOTT (2008): "Effective Basemetal Hedging: The Optimal Hedge Ratio and Hedging Horizon," *Journal of Risk and Financial Management*, 1(1), 41–76.

DUFFIE, D., AND K. J. SINGLETON (1993): "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 61(4), 929–952.

EDERINGTON, LOUIS, H. (1979): "The Hedging Performance of the New Futures Markets," *The Journal of Finance*, 34(1), 157–170.

ENGLE, R. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50(4), 9871007.

——— (2002): "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models," *Journal of Business and Economic Statistics*, 20(3), 339–350.

ENGLE, R. F., AND K. F. KRONER (1995): "Multivariate Simultaneous Generalized Arch," *Econometric Theory*, 11(1), 122–150.

ENGLE, R. F., AND V. K. NG (1993): "Measuring and Testing the Impact of News on Volatility," *The Journal of Finance*, 48(5), 1749–1778.

EVANS, L., AND G. GUTHRIE (2007): "Commodity Price Behavior With Storage Frictions," Working paper, Victoria University of Wellington, The New Zealand Institute for the Study of Competition and Regulation.

FAMA, E. F., AND K. R. FRENCH (1987): "Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage," *The Journal of Business*, 60(1), 55–73.

——— (1988): "Business Cycles and the Behavior of Metals Prices," *The Journal of Finance*, 43(5), 1075–1093.

——— (1989): "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics*, 25(1), 23 – 49.

GIBSON, R., AND E. S. SCHWARTZ (1990): "Stochastic Convenience Yield and the Pricing of Oil Contingent Claims," *The Journal of Finance*, 45(3), 959–976.

GUSTAFSON, R. L. (1958): "Implications of Recent Research on Optimal Storage Rules," *Journal of Farm Economics*, 40(2), 290–300.

HAMILTON, J. D. (1994): *Time Series Analysis*. Princeton University Press.

HERBST, A., D. KARE, AND J. MARSHALL (1989): "A time varying, convergence adjusted,minimum risk futures hedge ratio," *Advances in Futures and Options Research*, 6, 137–155.

HICKS, T. R. (1939): *Value and Capital*. Oxford University Press.

HILLIARD, J. E., AND J. REIS (1998): "Valuation of Commodity Futures and Options Under Stochastic Convenience Yields, Interest Rates, and Jump Diffusions in the Spot," *The Journal of Financial and Quantitative Analysis*, 33(1), 61–86.

HONG, H., AND M. YOGO (2009): "Digging into commodities," *Presented at the USC FBE Finance Seminar*

——— (2010): "Commodity market interest and asset return predictability," *Unpublished Working paper*.

——— (2012): "What does futures market interest tell us about the macroeconomy and asset prices?," *Journal of Financial Economics*, 105(3), 473 – 490.

HULL, J. C. (2012): *Options, Futures, and Other Derivatives*, chap. Hedging Strategies Using Futures. Prentice-Hall.

HWANG, S. AND S.E. SATCHELL(2005): "GARCH Model with Cross-Sectional Volatility: GARCHX Models," *Applied Financial Economics*, 15, 203-216.

IRWIN, S. H., AND D. L. GOOD (2009): "Market Instability in a New Era of Corn, Soybean, and Wheat Prices," *Choices*, 24(1).

JOHANSEN, S. (1988): "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control*, 12(2-3), 231–254.

JOHNSON, LELAND, L. (1960): "The Theory of Hedging and Speculation in Commodity Futures," *The Review of Economic Studies*, 27(3), 139–151.

KALDOR, N. (1939): "Speculation and Economic Stability," *The Review of Economic Studies*, 7(1), 1–27.

KEYNES, J. M. (1930): *A treatise on Money*, vol. 2. London: Macmillan.

KHAN, S. A., Z. KHOKER, AND T. T. SIMIN (2008): "Expected Commodity Futures Returns," Working paper, `http://ssrn.com/abstract=1107377`.

KILIAN, L. (2009): "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review*, 99(3), 1053–69.

KRONER, K., AND J. SULTAN (1991): "Exchange Rate Volatility and Time Varying Hedge Ratios," Pacific-Basin Capital Market Research.

KRONER, K. F., AND J. SULTAN (1993): "Time-Varying Distributions and Dynamic Hedging with Foreign Currency Futures," *The Journal of Financial and Quantitative Analysis*, 28, 535–551.

LEE, B.-S., AND B. F. INGRAM (1991): "Simulation estimation of time-series models," *Journal of Econometrics*, 47(23), 197 – 205.

LIEN, D., AND X. LUO (1993): "Estimating multiperiod hedge ratios in cointegrated markets," *Journal of Futures Markets*, 13(8), 909–920.

——— (1994): "Multiperiod hedging in the presence of conditional heteroskedasticity," *Journal of Futures Markets*, 14(8), 927–955.

LIEN, D., Y. K. TSE, AND A. K. C. TSUI (2002): "Evaluating the hedging performance of the constant-correlation GARCH model," *Applied Financial Economics*, 12(11), 791–798.

LIEN, D.-H. D. (1996): "The effect of the cointegration relationship on futures hedging: A note," *Journal of Futures Markets*, 16(7), 773–780.

LUTKEPOHL, H. (1991): *Introduction to Multiple Time Series Analysis*. Springer-Verlag, Berlin.

MASTERS, M. W. (2008): "Testimony of Michael W. Masters before the Committee on Homeland Security and Governmental Affairs United States Senate," `www.hsgac.senate.gov/public/_files/052008Masters.pdf`.

MCCANDLESS, G. T. (2008): *The ABCs of RBCs: An Introduction to Dynamic Macroeconomic Models*. Harvard University Press, Cambridge, Massachusetts, and London, England.

MCLAREN, J. (1999): "Speculation on Primary Commodities: The Effects of Restricted Entry," *The Review of Economic Studies*, 66(4), 853–871.

MEDLOCKIII, K. B., AND A. M. JAFFE (2009): "Who Is In the Oil Futures Market and How Has It Changed?," *James A. Baker III Institute For Public Policy, Rice University*.

MIAO, Y. W. W., AND N. FUNKE (2011): "Reviving the Competitive Storage Model: A Holistic Approach to Food Commodity Prices," *IMF Working Paper*.

MICHAELIDES, A., AND S. NG (2000): "Estimating the rational expectations model of speculative storage: A Monte Carlo comparison of three simulation estimators," *Journal of Econometrics*, 96(2), 231 – 266.

MITRAILLE, S., AND H. THILLE (2009): "Monopoly behaviour with speculative storage," *Journal of Economic Dynamics and Control*, 33(7), 1451 – 1468.

MUTH, J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29(3), 315–335.

MYERS, ROBERT, J. (1991): "Estimating time-varying optimal hedge ratios on futures markets," *Journal of Futures Markets*, 11(1), 39–53.

NEWBERY, D. M. (1984): "Commodity Price Stabilization in Imperfect or Cartelized Markets," *Econometrica*, 52(3), 563–78.

NG, S., AND F. J. RUGE-MURCIA (2000): "Explaining the Persistence of Commodity Prices," *Computational Economics*, 16, 149–171, 10.1023/A:1008713823410.

PAKES, A., AND D. POLLARD (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57(5), 1027–1057.

PINDYCK, R. S. (1993): "The Present Value Model of Rational Commodity Pricing," *The Economic Journal*, 103(418), 511–530.

——— (2004): "Volatility and commodity price dynamics," *Journal of Futures Markets*, 24(11), 1029–1047.

ROBLES, M., M. TORERO, AND J. VON BRAUN (2009): "When Speculation Matters," No. 57. International Food Policy Research Institute (IFPRI).

ROSSI, E., AND C. ZUCCA (2002): "Hedging interest rate risk with multivariate GARCH," *Applied Financial Economics*, 12(4), 241–251.

ROTEMBERG, J. (1987): *The New Keynesian Microfoundations*, 69–116. The MIT Press.

ROTEMBERG, J. J. (1996): "Prices, output, and hours: An empirical analysis based on a sticky price model," *Journal of Monetary Economics*, 37(3), 505 – 533.

RUST, J., AND G. HALL (2003a): "Minimum Distance Estimation of a Model of Optimal Commodity Price Speculation with Endogenously Sampled Prices," Working paper, Yale University.

——— (2003b): "Middleman versus Market Makers: A theory of Competitive exchange," *Journal of Political Economy*, 111(2), 353–403.

SAMUELSON, P. A. (1971): "Stochastic Speculative Price," *Proceedings of the National Academy of Sciences of the United States of America*, 68(2), 335–337.

SANTHOSH, K. (2012): "Commodity Price Volatility and Time varying Hedge Ratios Evidence from the Notional Commodity Futures Indices of India," *World Economics Association (WEA), Conferences, 2012, Rethinking Financial Markets*.

SCHWARTZ, E. S. (1997): "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging," *The Journal of Finance*, 52(3), 923–973.

SIM, A. B., AND R. ZURBRUEGG (2001): "Dynamic Hedging Effectiveness in South Korean Index Futures and the Impact of the Asian Financial Crisis," *Asia-Pacific Financial Markets*, 8(3), 237–258.

STEIN, J, L. (1961): "The Simultaneous Determination of Spot and Future Prices," *American Economic Review*, 59(5), 1012–1025.

SVEDBERG, P., AND J. E. TILTON (2006): "The real, real price of nonrenewable resources: copper 1870-2000," *World Development*, 34(3), 501–519.

TAYLOR, J. B. (1980): "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, 88(1), 1–23.

THOMAS, S., AND R. BROOKS (2001):"GARCH based hedge ratios for Australian share price index futures: does asymmetry matter?," *Accounting, Accountability and Performance*,7, 61–76.

WANG, P. (2008): *Financial Econometrics*. Routledge,Taylor and Francis group.

WILLIAMS, J. C., AND B. D. WRIGHT (1991): *Storage and Commodity Markets*. Cambridge University Press.

YAN, X. (2002): "Valuation of commodity derivatives in a new multi-factor model," *Review of Derivatives Research*, 5(3), 251–271.

YANG, W., AND D. E. ALLEN (2005): "Multivariate GARCH hedge ratios and hedging effectiveness in Australian futures markets," *Accounting and Finance*, 45(2), 301–321.

YU, J. (2002): " Forecasting volatility in the New Zealand stock market," *Applied Financial Economics*, 12, 193–202.

ZHONG, M., A. F. DARRAT, AND R. OTERO (2004): "Price discovery and volatility spillovers in index futures markets: Some evidence from Mexico," *Journal of Banking and Finance*, 28(12), 3037 – 3054.