ESSAYS ON THEORETICAL AND EMPIRICAL STUDIES OF COMMODITY FUTURES MARKETS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for

the Degree Doctor of Philosophy in the

Graduate School of the Ohio State University

By

Haijiang Zhou, B.A. Agricultural Economics, M.A. Economics

*****

The Ohio State University

2005

Dissertation Committee: Approved by

Dr. Matthew C. Roberts, Adviser

Dr. Carl R. Zulauf   Adviser

Dr. Cameron S. Thraen   Graduate Program, Agricultural, Environmental, and Development Economics

_____

ABSTRACT

Zhou, Haijiang. Essays on Theoretical and Empirical Studies of Commodity Futures Markets. (Under the direction of Dr. Matthew C. Roberts.)

The three essays of this thesis research several theoretical and empirical issues of the commodity futures markets, specifically, the metals markets at the London Metal Exchange (LME) and the U.S. soybean and corn markets at the Chicago Board of Trade.

Chapter two examines the cost of carry theory for five metals at the London Metal Exchange (LME). A quad-variate cointegration model is constructed and empirical results show that a long run relationship exists for cash and 3-month metals futures prices, 3-month interest rates and physical storage costs. The finding reconciles previously inconsistent findings regarding the cointegration of temporal prices in the presence of non-stationary interest rates.

Chapter three updates the measurement of the supply of storage model and develops a two-equation system model which consists of the supply of storage equation and the price spread-convenience yield equation. Three stage least squares (3SLS) estimation method and bootstrapping 3SLS are applied to the CBOT soybeans data and results reveal that convenience yield and variability of new crop futures might play key roles in making storage decisions during the crop year.

Chapter four develops a new measurement of the stock (inventory)-price relationship for commodity markets by constructing an equally weighted ending stocks-use ratio. A fully specified polynomial function is developed with consideration of three policy regimes due to the 1985 and 1996 US farm policy reforms. Model selection is conducted from both the fitting perspective and the forecast perspective. Results show that grain market analysts may benefit from using the proposed new measurement for forecasting prices.

In summary, this study contributes to the understanding of the theoretical and empirical issues of the commodity futures markets, including the cost of carry theory, the supply of storage theory and the convenience yield theory.

ACKNOWLEDGMENTS

VITA

1988-1992 … B.A., Agricultural Economics, Zhongnan University of Finance and
      Economics, Wuhan, Hubei, P.R. China.
1992-1995 … Junior Analyst, Bureau of Agriculture, Shantou, Guangdong, P.R. China.
1995-1999 … Lecturer, Hubei University of Economics, Wuhan, Hubei, P.R. China.
1997-1999 … M.A., Economics, Zhongnan University of Finance and Economics,
      Wuhan, Hubei, P.R. China.
1999-2000 … M.A., Economics, The Ohio State University.
1999-present Graduate Teaching and Research Assistant, Department of Agricultural,
      Environmental, and Development Economics, The Ohio State University.

FIELDS OF STUDY

Major Field: Agricultural, Environmental, and Development Economics
Minors: Financial Economics, Futures and Options Markets

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure **Page**

CHAPTER 1

INTRODUCTION

Working (1948, 1949)'s theory of supply of storage and the cost of carry theory have been the centerpiece of the literature of futures markets. Extensive studies have been done on theoretical discussions and empirical examinations of these theories in the past half century. Measures of variables in the supply of storage model and in the cost of carry model, including inter-temporal price spread, interest cost, physical storage cost, and convenience yield, have also been explored. This thesis updates the measurement of the supply of storage model and the cost of carry model and examines the theory for five metals markets at the London Metal Exchange (LME) and the soybeans and corn markets at the Chicago Board of Trade.

Chapter two studies the long run relationship between cash and 3-month futures prices for five metals at the London Metal Exchange (LME) and investigates the role of interest rates in this relationship. A quad-variate cointegration model is constructed and empirical results show that co-integration exists for cash and 3-month metals futures prices, 3-month interest rates and physical storage costs. These findings reconcile

previously inconsistent findings in the literature regarding the cointegration of temporal prices in the presence of non-stationary interest rates and are consistent with Working's cost of carry theory.

Chapter three updates the measurement of the supply of storage model by incorporating recent developments in the theoretical and empirical literature as well as markets. A modified supply of storage model is developed through a two period trading model. This model consists of two equations within a system, the supply of storage equation and the price spread-convenience yield equation. Analysis is conducted through three-stage least squares (3SLS) estimation method and bootstrapping 3SLS due to small sample size of data. Results reveal that in addition to net price of storage or storage costs adjusted price spread, convenience yield and price variability of new crop futures might play important roles in making storage decision for the soybean market.

Chapter four develops a new measurement of the stock (inventory)-price relationship for commodity markets. This involves constructing an equally weighted ending stock-use ratio over the crop year and utilizing it to measure grain stock (inventory)-price relationship. A fully specified polynomial function is developed with consideration of three policy regimes due to the 1985 and 1996 US farm policy reforms. Model selection is conducted for various models by comparing standard OLS results, curve fitting, and forecasting error computed using cross validation method. Empirical results show that grain market analysts may benefit from using the proposed new measurement for analyzing stock-price relationship and forecasting prices for soybean.

Chapter five summarizes each of the three essays of the thesis. This chapter also provides concluding remarks and implications of the three empirical studies in the essay.

CHAPTER 2

Testing the Cost of Carry Theory by Cointegration: A Case Study on London Metal Exchange

2.1 Introduction

The accurate anticipation of changes in the cash-futures basis is essential for a hedger to successfully use futures markets to either increase returns or reduce risk. Accurate anticipation starts with an understanding of the behavior of the basis and the long run relationship between cash and futures prices of a commodity. Because nonstationarity is a characteristic of many speculative prices, cointegration theory has emerged as an important tool for analyzing the dynamics of futures and cash prices and their long run equilibrium relationship. Cointegration theory implies that two non-stationary time series sharing the same type of stochastic trend tend to move together over the long run although deviations from the long run equilibrium can occur during the short run (Engle and Granger, 1987).

While the conceptual argument for using cointegration theory to analyze the cash-futures basis is compelling, the empirical evidence is mixed. Schwarz and Szkmary (1994), Thraen(1999), Kellard, *et al.* (1999), Haigh (2000), Yang, Bessler and Leather

(2001), and McKenzie, *et al.* (2002) find cash-futures cointegration in their studies. In contrast, Baillie and Myers (1991), Chowdury (1991), Schroeder and Goodwin (1991), and Fortenbery and Zapata (1997) find no cointegration, while Covey and Bessler (1991), Quan (1992), Fortenbery and Zapata (1993), Covey and Bessler (1995), and Sabuhoro and Larue (1997) find mixed results regarding cash-futures cointegration. Each of these studies examined the cash-futures basis. Commodities investigated include grains, livestock, cheddar cheese, coffee and cocoa, metals, petroleum products, and energy, among others.

Brenner and Kroner (1995) argue that model misspecification may explain a failure to find cash-futures cointegration. They point out that cash and futures prices are related to each other through the cost of carry model. Thus, futures-cash cointegration depends upon the time-series property of the components of the cost of carry model, which include not only the futures and cash prices, but also interest rate, physical storage cost, and convenience yield. Brenner and Kroner specifically note that interest rates generally have been found to be non-stationary. Zapata and Fortenbery (1996) provide empirical evidence supporting Brenner and Kroner's argument in their trivariate cointegration analysis of cash prices, futures prices, and interest rates for U.S. corn and soybeans. In addition, Heaney (1998) finds evidence for a single quad-variate cointegrating vector for lead traded at the London Metal Exchange (LME) involving futures price, spot price, interest rate, stock level (a proxy for convenience yield), and a constant term. Watkins and McAleer (2002) also find similar results on the quad-variate cointegration model involving LME copper cash and futures prices, interest rates, and stock level. Kellard (2002) finds cointegration in the UK wheat market for both a bivariate cash-futures

model and a trivariate model involving non-stationary cash and futures prices and interest rates and concludes that this theoretical paradox is due to the small magnitude of interest rate relative to the forecasting error.

Brenner and Kroner also argue that a finding of no futures-cash cointegration could result from using futures prices which do not have a constant time to maturity. Only the London Metals Exchange (LME) trades futures contracts that have a constant, as opposed to declining, time to maturity. This unique feature of LME futures price series allows us to avoid the decreasing time to maturity problem and the potential estimation bias. While Heaney (1998) and Watkins and McAleer (2002) used the LME lead contract and copper contract in their analyses, they did not examine the possibility of bivariate cointegration between futures and cash prices for a data set that avoids the statistical problems created by using futures prices with a declining time to maturity.

This study examines cash-futures bivariate cointegration using weekly LME futures prices for aluminum, copper, nickel, lead, and zinc. We also examine cointegration of the cost of carry model using measures of physical storage cost and convenience yield. Previous studies have treated physical storage cost as a component of the constant term in the cointegration equation. Convenience yield is measured using a recent approach suggested by Heaney (2002). These analyses extend our understanding of the cost of carry model's role in price determination by providing a detailed picture of the cointegration that exists among the components of the cost of carry model while using data that avoids an important statistical problem in examining futures-cash cointegration.

The rest of this chapter is structured as follows. Section II reviews previous studies in cash-futures cointegration for various commodity markets. Second III discusses the cost

of carry theory and cointegration theory. Discussions of econometric methodology, including unit root tests and cointegration tests, follow. Data is described in section V. Section VI contains the empirical results. Conclusions and implications are presented in the last section.

2.2 Literature Review

Cash-futures cointegration measures the long run equilibrium relationship between cash price and futures price of an asset. Cash-futures cointegration for commodities has received considerable attention in the literature because of its importance in examining other theoretical and empirical issues in the futures literatures such as market efficiency, unbiasedness, optimal hedge ratio, and futures and options pricing. Empirical studies in commodity cash-futures cointegration have covered commodities such as crude oil, metals, grains, coffee, cocoa, cheddar cheese, petroleum, among others. These studies have reported mixed evidence for cash-futures cointegration for storable commodities. Some recent studies on commodity cash-futures cointegration consider the potential non-stationarity of components of cost of carry term such as interest rates and convenience yield and construct multi-variate cointegration models. Such studies include Zapata and Fortenbery (1996), Heaney (1998), Watkin and McAleer (2002), Kellard (2002).

Schwarz and Szkmary (1994), Covey and Bessler (1995), Thraen (1999), Kellard, *et al.* (1999), Haigh (2000), Yang, Bessler and Leather (2001), and McKenzie, et al. (2002) find cash-futures cointegration for some commodity markets. Schwarz and Szakmary (1994) examined both cash-futures cointegration relationship and price discovery for

three petroleum products and their futures markets, namely, light sweet crude oil, heating oil #2, and unleaded gasoline. Cash-futures cointegration was tested under the framework of Garbade and Silber (1983) (GS) model. The futures prices were cointegrated with their deliverable spot prices for each of the three products and futures prices were less volatile than the cash prices for all three energy products. ECM and GS models were used to examine price leadership between the spot and futures markets of each product. Both approaches found that futures markets dominated spot markets for all three products. Thraen (1999) examined the existence of cointegration between CSCE Cheddar Cheese cash and futures prices using a much longer time period of data than Fortenbery and Zapata (1997). In contrast to Fortenbery and Zapata (1997)'s finding of no cointegration for the Cheddar Cheese cash and futures prices, Thraen (1999) reports empirical results showing that cheddar cheese market exhibits a cointegrating relation among the futures and spot markets, implying a flow of information between the two markets. Kellard, *et al.* (1999) adopts a cointegration technique to test unbiasedness and efficiency across a range of commodity and financial futures markets and develops a measure of relative efficiency. The findings suggest that spot and futures prices are cointegrated with a slope coefficient that is close to unity, indicating that the postulated long-run relationship is accepted. Haigh (2000) examines the long-run cointegrating relationship between freight cash and futures prices (BIFFEX freight futures market) in a forecasting model and compares the forecasting performance of this model with several alternatives. This study finds cointegration between freight cash and futures prices exists. The study also evaluates the stability of the cointegrating relationship over time using a rolling cointegration technique for the reason that there have been several changes in the freight index in the

7

history of freight index futures market. Results suggest that the long-run relationship between freight spot prices and the BIFFEX freight index futures indeed has strengthened over time. Yang, Bessler and Leather (2001) applied cointegration test to examine the price discovery role of futures markets for selected storable and non-storable commodities and to further evaluate the effect of asset storability on long run relationship between commodity cash and futures prices. Considering the possible effect of the 1996 FAIR Act on the cash-futures cointegration, the study period was divided into two sub-periods, that was, from January 1, 1992 to March 31, 1996 and from April 1,1996 to June 30, 1998. For the first period studied, results showed that cash-futures cointegration existed for five of eight storable commodities and for all non-storable commodities. Cointegration was also found for six out of eight storable commodities and all non-storable commodities for the second period studied. Based on this strong evidence, the authors concluded that the unbiasedness hypothesis for studied futures markets could not be rejected. The authors also concluded that asset storability did not affect the existence of a long-run relationship between commodity cash and futures prices. Mckenzie, *et al.* (2002) examines short-run and long-run unbiasedness within the U.S. rice futures market using standard OLS, cointegration, and error-correction models, respectively. The Johansen procedure is used to test for cointegration between cash and futures prices. The empirical results indicate that cash-futures cointegration exists.

In contrast, Baillie and Myers (1991), Chowdury (1991), Schroeder and Goodwin (1991), and Fortenbery and Zapata (1997) find no cointegration for some commodities markets. Ballie and Myers (1991) examines six different commodities using daily data over two futures contract periods by using bivariate GARCH models of cash and futures.

These commodities include beef, coffee, corn, cotton, gold, and soybeans. Tests for cointegration between cash and futures prices for these commodities are constructed by undertaking unit root tests on the residuals obtained from regressing cash prices on futures prices (Engle and Granger's cointegration test). The hypothesis of no cointegration cannot be rejected for any commodity. Chowdhury(1991) employs cointegration theory to test market efficiency for four nonferrous metals—copper, lead, tin and zinc traded in the London Metal Exchange (LME). The study tested cointegration between cash and futures prices of same metal for all four markets. Cash-futures cointegration was rejected for all metals except for copper. In the case of copper the evidence was mixed and did not permit a definitive conclusion. Hence, the author concluded that futures price was a biased predicator of cash price, implying market inefficiency for all four metals markets. Schroeder and Goodwin (1991) examine short- and long-run price relationships between Omaha cash and CME futures daily prices for live hogs for each year of the sample period from 1975 to 1989. Their empirical results show that the hog daily cash market and live hog futures market were not cointegrated for most of the years over the sample period except for the year 1980. Fortenbery and Zapata (1997) examine cash-futures cointegration and price discovery for the newly developed cheddar cheese futures markets. Sample data from June 1993 through July 1995 were used to examine the cash-futures relationship. Empirical results show that there is no cointegration between cheddar cheese cash and futures markets, suggesting that price information in one market has very little impact on price movement in the other market. To explore whether failure to find cointegration is due to market infancy, the authors further examine two newly established fertilizer products futures contracts using the same

testing procedure. It turns out that cash-futures cointegration exists for these two markets and both futures markets lead the corresponding cash markets in terms of price discovery.

More interestingly, some researchers find mixed results on their individual studies either across different sample periods for the same commodity or across different commodities. These studies include Covey and Bessler (1991), Quan (1992), Fortenbery and Zapata (1993), Covey and Bessler (1995), and Sabuhoro and Larue (1997). Covey and Bessler (1991) examines cointegration between spot slaughter cattle price and the nearby, as well as a distant live cattle futures price. The cointegration test is based on residuals from a static regression and results show marginal support for the cointegration hypothesis between cash prices and the nearby futures contract. No cointegration exists between cash prices and the distant contract. Quan (1992) investigates the price discovery role of crude oil futures using a two-step testing procedure, which involved testing cointegration between crude oil cash and futures prices as the first step. Data used in this study are crude oil cash prices, 1,3,6, and 9-month futures prices. Empirical results reveals that cointegration relationship existed for crude oil cash prices and 1-month futures prices, and for cash prices and 3-month futures prices. However, longer length futures prices (6-month and 9-month) were not cointegrated with cash prices. Fortenbery and Zapata (1993) examines cointegration relation between futures and local grain markets using the CBOT corn and soybean futures prices and cash prices for corn and soybean at Greenville, North Carolina and at Williamston, North Carolina, for the period 1980 through 1991. Specifically, they employed a full information maximum likelihood approach developed by Johansen and Juselius (1990). The soybean cash markets in both Greenville and Williamston followed the nearby futures market in only 4 out of 11 years

studied. Corn cash markets were also cointegrated with futures markets in only 4 years, coinciding with the results in soybean markets for 1983-1984 and 1988-1989 crop year. Clearly, evidence of cash-futures cointegration was limited. Sabuhoro and Larue (1997) tests the Efficient Market Hypothesis (EMH) using cointegration technique and other testing procedures for coffee and cocoa futures. Data used in this study are daily cash (spot) and futures prices from the Coffee, Sugar, and Cocoa Exchange (CSCE). Three different methods are used to examine cointegration relationship between future spot and futures markets. The Engle-Granger procedure, which involved testing unit root of residuals of the cointegration regression, provides evidence on cointegration for three of the four cases studied. The results obtained by Johansen and Juselius (1990) ML procedure provides stronger evidence since each of the four futures prices series was cointegrated with its corresponding future spot price. However, Hansen's LC test reports cointegration for the cocoa contracts but not for the coffee contracts.

Brenner and Kroner (1995) proposes two explanations for the mixed results regarding commodity cash-futures cointegration in the literature. They argue that the failure of finding cointegration between commodity cash and futures prices could be due to the following two problems. One is associated with data, i.e., futures prices with decreasing time to maturity. The other is model misspecification. They point out that cointegration between cash prices and futures prices depends upon the time-series property of the cost of carry. They further argue that interest rates, an important component of carrying costs, are potentially non-stationary and thus may play a critical role in determining the cointegration relationship between cash and futures prices. Brenner and Kroner suggest including the potentially non-stationary interest rates and/or

other random components of cost of carry in the cash-futures cointegration system. Zapata and Fortenbery (1996) found empirical support for Brenner and Kroner's argument on the importance of interest rates in their trivariate cointegration analysis of cash prices, futures prices, and interest rates for corn and soybean. Following Brenner and Kroner's argument, Heaney (1998) and Watkins and McAleer (2002) construct a four-variate cointegration system involving cash and futures prices, interest rates, and stock level for LME lead and copper markets, respectively. Both studies find cointegration exists for the studied metal markets.

Following Brenner and Kroner (1995)'s argument, Zapata and Fortenbery (1996) constructs a modified trivariate model, which incorporates stochastic interest rates into the cash-futures system. Their study used U.S. corn and soybean prices during the sample period from September 1980 through August 1995. Using Johansen's cointegration test, the authors found cointegration between cash, futures and T-Bills rates for the whole sample period. A bivariate model and a trivariate model are constructed for 15 individual crop years and are tested. In soybeans bivariate model they found cointegration for 6 of 15 individual crop years while in soybeans trivariate model cointegration existed for 10 out of 15 crop years. Thus, the authors conclude that non-stationary interest rates series are critical in determining long run cash –futures relationship for corn and soybeans, especially when interest rates were volatile. Heaney (1998) tests the relationship between futures price, cash price and two of the main components of the cost of carry term, namely, interest rate and convenience yield which takes stock level effects as a proxy. It assumes that the storage costs are a fixed proportion of the spot price and models the stock level effects including convenience yield and risk premium using stock info.

12

Empirical results show that there exists one cointegrating vector at the 5% level of significance for the examined cost of carry model. The stock level parameter is significantly different from zero, identifying stock level as a statistically significant component of the cost-of-carry model for the lead contract. Adopting a similar approach, Watkins and McAleer (2002) constructs a four-variate cointegration model for the LME copper market which involves copper cash and futures prices, interest rates and stock level. Cointegration is found to exist among these four variables. By examining the UK wheat futures contract traded at LIFFE, Kellard (2002) finds both cointegration for the bivariate cash-futures model and a trivariate model involving non-stationary cash and futures prices and interest rates, which is apparently a paradox. This cointegration paradox is examined by investigating the relative magnitudes of the forecast error and the domestic interest rate. The author concludes that the paradox is probably due to small size of interest rates relative to the forecast error and that cointegration methodology is not appropriate for evaluating commodity market efficiency.

2.3 Theory

2.3.1 The Cost of Carry Theory

The cost of carry theory, or cost of storage theory, dates to Working (1949)'s price of storage theory. Working (1949) argues that the difference between prices of the same commodity quoted on the same day for two different delivery dates can be viewed as a price of storage. The price difference is a "necessary return" in order for economic agents to store the commodity for future sale. This necessary return is determined by the supply

and demand for storage. Working's price of storage can be negative if the price for distant delivery is below the price for nearer delivery. The concept of convenience yield is used to explain why a commodity would be stored when its price is expected to decline. Convenience yield is most commonly identified as the sum of all benefits that accrue from having the cash commodity during a time of scarcity. For example, a processing firm can continue to process a commodity during times of scarcity if it has the commodity in storage.

The cost of carry theory is normally presented using cash and futures prices:

$$F_{t,t+\tau} = S_t \exp[\tau(r_t + D_t - C_t)] \tag{2.1}$$

where, $F_{t,t+\tau}$, denotes futures price as of date t for delivery at date $t + \tau$, S_t denotes cash price at date t and τ is the duration of the futures contract, r_t is the risk-free interest rate, D_t is the physical storage cost expressed as a percent of the asset cash price and includes insurance costs, and C_t is convenience yield expressed as a percent of cash price.

Equation 2.1 assumes a condition of no arbitrage. Thus, equation 2.1 describes an equilibrium situation. Specifically, the return to inventory holders is the same whether: (i) they hold inventory and sell it at some time in the future, thus earning an expected net return equal to expected future cash price plus convenience yield minus physical storage costs over the storage period; or (ii) they sell inventory at current time and invest the cash proceeds at risk-free interest rate. Equation 2.1 is also the most commonly used arbitrage approach for pricing futures on commodity and other financial assets.

Taking natural logarithms of equation 2.1 yields,

$$\ln F_{t,t+\tau} = \ln S_t + \tau(r_t + D_t - C_t) \tag{2.2}$$

14

which can be rearranged as

$$\ln F_{t,t+\tau} - \ln S_t = \tau(r_t + D_t - C_t) \qquad (2.3)$$

In equations 2.2 and 2.3, the cost of carry term, $\tau(r_t + D_t - C_t)$, provides a link between the natural logarithm of cash and futures prices. As Baillie and Myers (1991) and Brenner and Kroner (1995) point out, properties of the cost of carry term are critical in determining the long run relationship between cash and futures prices. If cost of carry is non-stationary, cash and futures prices are not cointegrated. In contrast, if cost of carry is stationary, then cash and futures prices are cointegrated.

Equations 2.1, 2.2, and 2.3 describe a special case of temporal spreads, specifically the spread between cash price and futures price. These equations apply more broadly to any temporal spread, including the spread between the prices of any two futures contracts quoted at the same time. Just as pointed out by Working (1949), the inter-temporal relationships between any prices of the same commodity can be linked by storage costs, or the cost of carrying the commodity over the time interval between the two contracts. Thus, the cost of carry model can be used to analyze the long run relationship between any pair of futures prices quoted at the same time for the same commodity, in other words, futures spread cointegration.

2.3.2 Cointegration Theory

Cointegration theory has been one of the most important new developments in time series analysis over the last two decades. The importance of cointegration theory lies in its unique treatment of non-stationary time series, or series with stochastic trend, in a multi-variate context. The presence of stochastic trends in a multi-variate system makes

standard statistic models and testing procedures inappropriate. Cointegration theory and the associated error correction model (ECM) or vector ECM allow researchers to estimate a potential linear combination of all non-stationary variables in the system and examine the dynamic co-movement of these variables in both the long run and short run. Studies have shown that many financial series are non-stationary and that many pairs of financial series and macro-economic series are cointegrated, such as foreign currency spot and forward rates, interest rates of different maturities, and consumption and income, among others.

Engle and Granger (1987) provided a formal definition of cointegration. Specifically, the components of vector x_t are said to be cointegrated of order *(d, b),* denoted as $x_t \sim CI(d,b)$, if (i) every component of x_t is *I(d)* and (ii) there exists a vector α ($\neq 0$) so that $z_t = \alpha' x_t \sim$ *I(d-b), $b \geq 0$* . Vector α is referred to as the cointegrating coefficient vector.

Most studies in cointegration theory focus on the *d=1, b=1* case since most economic series are integrated of order 1, in other words, they are unit root processes. Following Engle and Granger's definition, *CI (1,1)* means that, while each variable of vector *x* is a unit root process, some linear combination of all variables of vector x_t is stationary. Thus, even though each variable of vector x_t can wander randomly, the equilibrium error, z_t rarely drifts far away from its expected value. In other words, the cointegrated variables of vector x_t tend to move together over the long run despite the existence of short-term deviations. Cointegration can be used to test the cost of carry theory. This empirical investigation involves a cash price and a 3-month futures price. In

order to simplify the notation, the subsequent discussion will use symbols for a cash price and a futures price quoted at the same time. Equation 2.2 can be specified as an econometric equation:

$$\ln F_{t,t+\tau} = \beta \ln S_t + \tau(r_t + D_t - C_t) + e_t \tag{2.4}$$

where $F_{t,t+\tau}, S_t, \tau, r_t, D_t, C_t$ are the same as in equation 2.2 and e_t is the stationary disturbance term.

Previous studies have found that cash price and futures price of speculative assets tend to be non-stationary. However, if e_t, r_t, D_t, and C_t are all stationary, then a linear combination of non-stationary $\ln F_{t,t+\tau}$ and $\ln S_t$ should be stationary. Thus, a possible cointegrating coefficient vector for cash and futures prices could be $(1, -\beta)$, which represents a bivariate cointegration model between cash and futures prices. In other words, the difference between cash and futures prices will follow its long-run expected value, or, there exists a long run equilibrium, even though deviations from the long-run difference can exist in the short run.

Previous studies suggest that the stationarity of interest rates, *i.e.*, r_t, is debatable (for a detailed discussion, please see Brenner and Kroner, 1995). Thus, another potential scenario is that cash and futures prices and interest rates are all non-stationary but cointegrated through a trivariate system. To explore the potential implications of a unit root in the interest rates term, rewrite equation 2.4 as:

$$\ln F_{t,\tau} = \beta_1 \ln S_t + \beta_2 r_t + \tau(D_t - C_t) + \mu_t \tag{2.5}$$

Assuming the disturbance term is stationary, cointegration theory suggests that there exists a stationary linear combination of three non-stationary variables, cash price, futures

price, and interest rates. The cointegrating vector is *(1, - β₁, - β₂)*. Thus, the three non-stationary variables tend to move together in the long run although short run deviations may occur.

2.4 Econometric Methodology

While previous studies suggest that futures prices of the LME metals and interest rates are likely to follow unit root process, it is desirable to test the data being examined in the particular study for this property. Two widely used unit root tests and a stationarity test are used. Cointegration methodology is also discussed in this section.

2.4.1. Unit Root Test and Stationarity Test

The augmented Dickey-Fuller (ADF) test (Dickey & Fuller, 1979, 1981) and Phillips-Perron (PP) test (Phillips & Perron, 1988) are the most popular methods for examining whether a time series contains a unit root. The ADF model is:

$$\Delta X_t = \alpha + \beta X_{t-1} + \sum_{i=1}^{n} \lambda_i \Delta X_{t-i} + \varepsilon_t \qquad (2.6)$$

where, $\Delta X_t = X_t - X_{t-1}$. The ADF assumes that the errors are independent and have a constant variance. The coefficient of interest is β and the null hypothesis is H_0: $\beta=0$ against the alternative H_1: $\beta \leq 0$. If $\beta=0$, then equation 2.6 reduces to an equation in first difference. Such a specification implies that the variable possesses a unit root.

Phillips and Perron (1988) develop a generalization of the Dickey-Fuller test, by relaxing the assumption of independent errors with constant variance. The Phillips and Perron test (PP test) allows the error terms to be weakly dependent and heterogeneously

distributed. Standard unit root tests have been criticized for a lack of power, especially when distinguishing between unit root and weakly stationary processes (Kwiatkowski *et al.*, 1992). Consequently, we also use a stationarity test developed by Kwiatkowski, Phillips, Schmidt, and Shin (1992) (KPSS test) and use the combined results from the unit root and KPSS stationarity test to determine the stationarity of the examined series.

The null hypothesis of the KPSS test is that the examined series is stationary around a deterministic trend. It can be described as follow,

$$y_t = \phi t + \gamma_t + \varepsilon_t \tag{2.7}$$

where ϕt is a deterministic trend; γ_t is a random walk, i.e., $\gamma_t = \lambda_{t-1} + u_t$ and μ_t are *iid* $(0, \sigma_u^2)$; and ε_t is a stationary error.

The KPSS test statistic is calculated as

$$\hat{\eta} = T^{-2} \sum_{t=1}^{T} S_t^2 \Big/ s^2(l) \tag{2.8a}$$

where T is the number of observations in the tested series, $S_t = \sum_{i=1}^{t} v_i, t = 1,2,...,T$ and v_t, t=*1,2, ..., T*, are residuals from the regression of *y* on an intercept and time trend. s^2 *(l)* is a consistent estimator of the long-run variance and takes the form

$$s^2(l) = T^{-1} \sum_{t=1}^{T} v_t^2 + 2T^{-1} \sum_{s=1}^{l} (1 - s\big/(l+1)) \sum_{t=s+1}^{T} v_t v_{t-s} \tag{2.8b}$$

Critical values for the LM test statistic are available from Kwiatkowski *et al.* (1992, Table 1, p. 166). In contrast to the ADF and PP unit root tests, rejecting the null hypothesis of the KPSS test means the examined series is a unit root process because the null of KPSS test is exactly opposite to the null of the ADF and PP tests.

2.4.2 Cointegration Test

Hamilton (1994) proposed that, if economic theory suggests that a cointegrating relationship exists for the examined variables and that the cointegrating vector should be of a particular form, then a straightforward cointegration test is to evaluate the stationarity of a series constructed using the examined variables and the theoretical cointegrating vector. To implement Hamilton's approach, theory requires that each examined variable individually be a unit root process. If this requirement is met, then testing the null hypothesis that the constructed series is stationary is equivalent to testing the null hypothesis that the examined variables are cointegrated. Thus, if the constructed series is stationary, the examined variables would be cointegrated. On the other hand, if the constructed series is non-stationary, these variables would not be cointegrated. This method is similar to the two-step cointegration test proposed by Engle and Granger (1987) which involves testing the stationarity of the residuals series from regressing one examined economic series on the other one.

Johansen's full information maximum likelihood approach (Johansen, 1988; Johansen & Juselius, 1990) is widely used to conduct cointegration analysis. Following their presentation, consider a vector auto-regression (VAR) model

$$Y_t = \sum_{i=1}^{k} \Pi_i Y_{t-i} + \upsilon + \varepsilon_{t,} \tag{2.9a}$$

where, vector Y_t consists of p variables, each of which is assumed to be a unit root process and $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T$ are $IIN_p(0, \Lambda)$ and $Y_{-k+1}, \ldots, Y_0$ are fixed. The unrestricted parameters $(\upsilon, \Pi_1, \ldots, \Pi_k, \Lambda)$ are to be estimated from the vector auto-regression process. This equation can also be written as a reduced form error correction model (ECM) for a

non-stationary system,

$$\Delta Y_t = \prod Y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta Y_{t-i} + \upsilon + \varepsilon_t, t = 2,3,...T \qquad (2.9b)$$

where, $\Delta = 1 - L$ and L is the lag operator.

The long-run relationships between variables in vector Y_t are examined using the property of the coefficient matrix Π. Three cases are analyzed by Johansen and Juselius (1990): (1) matrix $\prod$ is a null matrix and thus has a rank of zero; (2) matrix $\prod$ has a full rank, which equals p, the number of variables in the system; (3) matrix $\prod$ has a rank between 0 and p. If the rank equals zero, no cointegrating relationships exist between the examined variables and a traditional VAR in differences is appropriate. If the rank equals p, variables in vector Y_t are stationary processes and a VAR in levels is the appropriate model to estimate. If the rank is between 0 and p, cointegration exists between the examined variables.

Given the three cases, the null hypothesis that there are (at most) $r(0 \le r \le p)$ cointegrating vectors can be tested by a likelihood ratio test, commonly known as trace test due to its association with the trace of the coefficient matrix. The trace test statistic is given by Johansen and Juselius (1990) as follows,

$$Trace = -T \sum_{i=r+1}^{p} \ln(1 - \lambda_i) \qquad (2.10)$$

where T is the number of observations and λ_i is the p-r smallest eigenvalues, i.e., squares of canonical correlations of Y_{t-1} with respect to ΔY_t, corrected for lagged differences. Besides the trace test, Johansen and Juselius (1990) also provide a test similar to the trace test, the $\lambda_{\max}$ test, which only evaluates the maximum eigenvalues.

2.5 The LME and Data

2.5.1 London Metal Exchange

The London Metal Exchange (LME) is the world's major non-ferrous metals market which provides producers and consumers world wide with highly liquid contracts on aluminum, aluminum alloy, copper, lead, nickel, tin, zinc, and silver. The LME has been successful since its establishment, because it has remained close to its core users in the industry by continuously providing contracts and outstanding services to meet the expectations of the industry. The LME currently maintains a turnover value of about US$ 2000 billion per annum and it is a major contributor to the UK's invisible earnings.

Three primary roles are performed by the LME. First, the exchange provides a market where market participants, primarily from base metals-related industries, can hedge against risks associated with price movement in these metals. Market participants, including producers and consumers of the non-ferrous metals can manage their price risks by taking positions in a variety of futures and options contracts and eliminate their price risks entirely or partially. The second role of the LME is to provide reference prices to global non-ferrous metals markets. Each day the exchange announces a set of official prices, which are determined from the open-outcry trading. These settlement prices officially quoted at the LME are globally accepted and are widely used by the non-ferrous metals industries world-wide as benchmark prices for contracts for the movement of physical metals. A third role performed by the LME is to provide appropriately located storage facilities and warehouses to market participants so that they can make or take physical delivery of approved brands of LME traded non-ferrous metals. All LME

contracts assume delivery of physical metals. To meet this need, large stocks of metals are held in a worldwide network of LME-approved warehouses. In reality, most contracts are settled out in the exchange and few deliveries of physical metals take place.

The LME currently trades futures and options on eight metals, including copper Grade A, primary aluminum, standard lead, primary nickel, tin, special high grade zinc, aluminum alloy, and North American Special Aluminum Alloy (NASAAC). The exchange also trades one index comprising the six primary base metals. For each metal, cash (1-day futures), 3-month, 15-month futures contracts are provided and a 27-month futures contract is also provided for some metals. Unlike other commodity markets in the world, which are usually based on monthly prompt dates, LME metal futures contracts are on a daily basis for a period of three months, a weekly basis for period between 3-month and 15-month period, and a monthly basis for 15, 27 or 63 month forward. The use of daily prompt dates is a unique feature of the LME which essentially means that the cash contract is always for delivery one day in the future and 3-month and 15-month contracts are always for delivery 3-month and 15-month in the future (The importance of this unique feature to our current study will be illustrated in the next sub-section.). Open outcry trading at the LME generates official price quotes for cash, 3-month, and 15-month contracts. In addition, quotes for any intermediate maturity are available from broker-dealers. The LME also offers options on each of these futures contracts and a traded average price options contracts (TAPOs) based on the monthly average settlement rice (MASP) for all metals futures contracts.

All LME prices are quoted in US Dollars. The LME permits contracts in sterling, Japanese yen, and Euros and provides official exchange rates from US Dollars for each of

these currencies. Metals contracts are in lots instead of tonnes and each lot of aluminum, copper, lead and zinc amount to 25 tonnes. In addition, nickel is traded in 6 tonnes lots, tin in 5 tonnes and aluminum alloy and NASAAC in 20 tonnes lots. Contract specifications are for the quality and shape which are most widely traded and demanded by industry.

2.5.2 Data

Cash prices, 3-month futures prices, and physical storage costs for aluminum, copper, nickel, lead, and zinc traded at the London Metal Exchange (LME) are used in this study. Also used is the 3-month U.S. Treasury-Bill rate as a measure of the 3-month risk-free interest rate.

Closing prices for Wednesday are collected as weekly prices. If prices are not available for Wednesday, closing prices on Thursday are collected. Prices for copper and lead begin with June 1973, while prices for aluminum, nickel, and zinc do not become available until April 1979, July 1982, and October 1986, respectively. All sample periods end in August 2004. Prices from January 1989 through August 2004 are obtained from the LME; the earlier data are collected from the Wall Street Journal. Physical storage costs for five metals for the period of 1994 through 2004 are obtained from the LME. 3-Month US Treasury-Bill rates are from the Federal Reserve Bank of St. Louis. Figure 1-5 plots weekly cash and futures prices for the studied period for aluminum, copper, nickel, lead, zinc, respectively. As can be noted from the figures, weekly cash and futures prices for each metal tend to track each other very closely over the sample period. The U.S. 3-month Treasury-Bill rates on a weekly basis are plotted in figure 6. Prices prior to March

1988 for nickel, January 1989 for aluminum, October 1989 for zinc, and January 1990 for copper and lead, were quoted in British Pounds. Subsequent prices are quoted in U.S. Dollars. The prices in British Pounds are converted into U.S. Dollars using the corresponding spot and forward exchange rates.

As mentioned earlier, in LME nomenclature the 1-day futures contract is referred to as a cash contract. However, this is not a cash price in the traditional and usual meaning of spot or immediate delivery. The 1-day futures price is used as a reference price for local spot prices quoted at 12 different locations around the world. But for simplicity, we will use the term of cash contract in the following analysis.

Also as noted earlier, in contrast to every other organized futures exchange, futures traded at the LME run on a daily basis for a period of three months and thus have a fixed time to maturity. In other words, 3-month futures has a fixed time to maturity of 3 months while 15-month contract and 27-month contract always mature 15 and 27 months in the future, respectively. In contrast, futures contracts traded at other exchanges in the world have a decreasing time to maturity as time passes because they have a fixed expiration or delivery date, not a fixed time to maturity. A fixed delivery date also means that futures contracts have to be "rolled-over" to create a continuous data series. For example, Covey and Bessler (1995) construct a nearby futures price series for fed cattle using six different delivery contracts (February, April, June, August, October, and December). Using 1988 as a specific illustration, the first observation is the settlement price on January 4, 1988 for the nearby February contract. Prices continue to be collected from the February contract until its expiration, February 19. On the next trading day, February 22, the price becomes the settlement price of the April Live Cattle contract. The same roll-over

procedure is used at the expiration of each succeeding futures contract. Such a roll-over procedure is commonly used in the literature for constructing a time series of futures contract price.

Brenner and Kroner (1995) point out that if the time to maturity (τ in equation 1) decreases, *i.e.,* expiration date is fixed, the residuals from regressing $\ln F_{t,t+\tau}$ on $\ln S_t$ must converge to zero as τ approaches zero. Thus, the variance of the residuals from the cointegrating regression changes as time to maturity converges to zero, implying that the residuals series is not covariance stationary and that cash and futures prices should not theoretically be cointegrated.

The same statistical problems may exist even if both prices are for futures contracts. The reason is that, given a fixed delivery date, futures contracts still must be rolled. It is not unusual that the new pair of futures contracts will have a different number of calendar days between expiration dates. For example, corn contracts are traded for delivery in March, May, July, September, and December. Some adjacent pairs of months are separated by two months (March-May, May-July, and July-September) while other pairs are separated by three months (September-December and December-March). Thus, when two futures prices are examined, time to maturity may vary. This time to maturity problem may create biases in estimation procedures such as cointegration and error-correction model (ECM).

Because the LME metal contracts are quoted for a fixed time to maturity, the statistical problem associated with a changing time to maturity does not exist. Hence, the LME metal contracts provide a data set that is theoretically more appropriate to assessing the cointegration of temporally differentiated prices.

2.6 Empirical Results

2.6.1 A Bivariate Cash-Futures Cointegration Model

We first examine cointegrating relationships between cash and futures prices for the studied metals. This can be done by constructing a traditional bivariate cash-futures cointegration model.

Before we conduct the cointegation analysis, the time series properties, especially the stationarity of the involved time series need to be identified. As discussed earlier, cointegration and the associated error-correction model (ECM) are designed for testing relationships between or among non-stationary time series within a multi-variate context. In other words, if the involved time series are stationary, i.e., not unit-root processes, a vector-autoregression (VAR) model can be used to examine their relationships. Thus, ADF and PP tests are applied to cash and 3-month futures price series for aluminum, copper, nickel, lead, and zinc and for 3-month U.S. T-Bill rates to test if these series are unit-root process. All tests fail to reject the null hypothesis that the examined series contains unit root at the 5 percent level of significance. The KPSS test rejects the null of stationarity for 6 out of 10 metal price series and for 3-month T-Bill rates. When combined, these results imply that all 11 series are non-stationary during the examined sample periods. Presented in table 2.1, these results indicate that cointegration technique is appropriate for examining long run relationship between cash and futures prices for LME metals and for testing the cost of carry theory for these metals markets.

Two approaches are used to conduct the bivariate cointegration tests involving cash and 3-month futures prices. The first one, following the testing procedure suggested by

Hamilton (1994), involves testing the stationarity of the series constructed as a linear combination of the studied price series with a coefficient vector suggested by the underlying theory. The other approach is the commonly used Johansen's maximum likelihood (ML) approach (Johansen, 1988; Johansen & Juselius, 1990). These tests are conducted using Eviews 5.0.

The cost of carry theory suggests that a potential cointegrating vector of natural logarithm cash and 3-month futures prices is (1,-1). Thus, testing the cointegration relationship between cash and 3-month futures prices is equivalent to testing the stationarity of a series constructed by the natural logarithm of these two price series with a theoretical cointegrating vector (1,-1). Following Hamilton (1994), the ADF, PP, and KPSS tests without trend are applied to the constructed series for five metals. Results of these tests are presented in table 2.2. Both the ADF and PP tests reject the null hypothesis that the constructed series is a unit root process for all five metals. The KPSS test fails to reject the null hypothesis that the constructed series is stationary for four out of five metals. The exception is zinc, for which KPSS test rejects the null. When combined, these results provide strong evidence showing that the constructed series are most likely to be stationary for all five metals, indicating the existence of cash-futures cointegration for these metals.

Two null hypotheses of the trace and λ_{max} tests of Johansen are tested: (1) no-cointegration ($r = 0$) and (2) at most one cointegrating relation ($r \leq 1$). The optimal lag length for these tests is chosen by minimizing the Akaike Information Criterion (AIC). Two model specifications are considered: (1) data without linear trend and (2) data with linear trend. As shown in table 2.3, the null of no-cointegration ($r = 0$) is uniformly

28

rejected for both specifications for all five metals. The null of at most one cointegration ($r \leq 1$) is not rejected for all metals in the case of data without trend. In the case of data with linear trend, the null of at most one cointegration is rejected at the 5% level for all five metals. For each of the five metals, results show that there is one cointegrating relationship between cash and 3-month futures prices. These results imply that cash price and futures price tend to move together over the long run in all five metals markets.

In summary, the weight of the evidence from this analysis is that the cash and 3-month futures prices are cointegrated for the five LME metals. This conclusion contrasts with Chowdhury's (1991) conclusions of no cointegration for lead, zinc, and weak evidence for cointegration of copper. It also is at odds with cointegration theory because interest rates were found to be non-stationary. In other words, cash price and 3-month futures price are cointegrated while an important component of the cost of carry term, the interest rates series, is unit root process.

Zapata and Fortenbery (1996) found that the highly volatile interest rates of the early 1980s played an important role in the cointegration between cash and futures prices for soybeans and corn. Hence we examine the cointegration of cash and 3-month futures prices for LME metals during the period from January 1979 to December 1984, a period with highly volatile interest rates. Table 2.4 presents a descriptive statistics summary of US 3-month Treasury-Bill rates (weekly) for two sample periods, the whole sample period from 1973 to 2004 versus the 1979-1984 period. The variability of weekly changes in interest rates during this period is nearly twice that of the period from June 1973 through August 2004, as shown in table 2.4. Coefficient variation of the T-bill rates in the 1979-1984 period is also larger than that for the whole sample period. These

descriptive statistics imply that interest rates in the truncated 1979-1984 period are much larger than the whole sample period in terms of both magnitude and variability. The 1979-1984 sample period contains more than 300 observations of weekly observations, which should be sufficient to generate acceptable power of the statistical analysis.

Same testing procedures are applied to examine the cointegration relationship between cash and 3-month futures prices for 3 LME metals, aluminum, copper, and lead during the truncated sample period (data for nickel and zinc are not available for this specific period). Results of the stationarity tests of constructed series for aluminum, copper, and lead are listed in table 2.5 while results of Johansen's tests are reported in table 2.6. Johansen's tests imply that cash and 3-month futures prices are cointegrated for aluminum and copper but not for lead. The KPSS test implies that the two futures prices are not cointegrated for all three metals. The ADF and PP tests yield mixed results. Both tests indicate that the two futures prices are cointegrated for copper but not for aluminum. The ADF and PP tests give different results for lead. In summary, when taken as a group, these results do not provide compelling evidence that volatile non-stationary interest rates explain the perplexing finding of bivariate cointegration between cash and 3-month metal futures prices when interest rates are non-stationary.

2.6.2 A Quad-variate Cointegration Model

Our empirical results from the bivariate cointegration model provide strong evidence that cash and 3-month futures prices are cointegrated for the five LME metals studied with the presence of non-stationary interest rates. This finding calls into question the important role of interest rates in determining temporal cointegration proposed by the

previous studies such as Brenner and Kroner (1995) and Zapata and Fortenbery (1996). It also seems to contradict with the cost of carry theory. As discussed earlier, the cost of carry term consists of interest rates, physical storage costs and convenience yield, among which, interest rates and physical storage are observable and convenience yield is not. Thus, physical storage costs of five metals for the period 1994 to 2004 were obtained from the LME to explore their potential impact on temporal cointegration and to examine the anomaly presented by the preceding analysis.

Physical storage cost of five metals is converted into a percentage of metals' cash prices and then the stationarity of these five series are examined by using the ADF, PP, and KPSS tests. Results show that these series are unit root processes, implying that, physical storage cost may, along with interest rates, have impact on temporal cointegration. Thus, a quad-variate cointegration model can be constructed involving metals cash and 3-month futures prices, 3-month risk-free interest rates, and physical storage cost of metals.

Prior to conducting a quad-variate cointegration analysis for the sample period 1994 to 2004, it is necessary to examine the stationarity of the involved time series and the bivariate cointegrating relationships of metals for this particular sample period. Stationarity tests show that the interest rates series and all metals price series are unit root processes. To save space, these results are not presented here but are available upon request. Bivariate cointegration is tested for five metals by using the same testing procedures as in the preceding analysis. Results from the Johansen's cointegration test are presented in table 2.7. These results indicate that cash and 3-month futures prices are cointegrated for aluminum, copper, lead, and zinc but not for nickel. Considering the non-

stationarity of interest rates series and physical storage costs series for these metals, a quad-variate cointegration model is constructed for the five metals. Results from Johansen's cointegration test are presented in table 2.10. These results provide evidence showing that cash and 3-month futures prices, 3-month interest rates, and physical storage cost are cointegrated for five metals. These findings are consistent with Working's cost of carry theory and provide a better approach than the bivariate cointegration model for explaining the theory.

Two types of convenience yield of the LME metals markets are estimated, including the residual-type convenience yield and the option-based convenience yield as suggested by Longstaff (1995) and Heaney (2002). Stationarity of these two series of convenience yield is tested using ADF, PP, and KPSS tests with results presented in table 2.8 and table 2.9, respectively. Results imply that convenience yield is stationary, consistent with results of quad-variate cointegration analysis, i.e., a cointegrating relationship exists among metals cash price, 3-month futures price, metals physical storage costs, and 3-month risk free interest rates.

2.7 Conclusion

Using simple cost of carry model, this study examines cash and 3-month futures cointegration for five metal markets at the London Metal Exchange (LME) and the role of components of cost of carry term in futures spread relationship. These components include two observable market variables, interest rates and physical storage costs, and a non-observable variable, convenience yield. Cointegration tests for a traditional bivariate model are implemented for five metals by testing the stationarity of the constructed series

32

which is a linear combination of the examined price series and by using the Johansen's maximum likelihood approach. A perfect dataset from the LME are used for this study, which enables us to eliminate any potential estimation biases and misleading conclusions in cointegration analysis that might result from the traditional data construction method on futures price series such as rolling over futures contracts with different maturities.

Empirical results from the bivariate model strongly suggest that cash and 3-month futures prices are cointegrated in these five metal markets during the examined sample periods and that non-stationary interest rates are not needed for obtaining this temporal cointegration. An analysis of interest rate variability is conducted for a truncated sample period 1979-1984 to further examine the role of interest rates in cash-futures cointegration, particularly when interest rates were highly volatile during this specific period. Results do not provide convincing evidence in support of the important effect of interest rates on the cash-3 month futures long run relationship and further confirm implications from the results of bivariate model.

These findings imply that interest rates are not needed for the existence of the metals cash-3 month futures cointegration. This calls into question the important role of interest rates in determining cash-futures cointegration and the need of a trivariate cointegrating system involving cash and futures prices and interest rates proposed by Brenner and Kronner (1995) and Zapata and Fortenbery (1996).

At the meantime, our results seem to raise a puzzle, that is, bivariate cash and 3-month futures cointegration exists with the presence of a non-stationary interest rates series. To examine this anomaly, physical storage costs for 1994 to 2004 were obtained from the LME. Using this extended data set, a quad-variate cointegration model

involving cash and 3-month metals futures prices, 3-month interest rates, and physical storage costs is found to exist for aluminum, copper, nickel, lead, and zinc. The quad-variate cointegration model provides more powerful analysis than the bivariate model on the relationship between temporal prices and the cost of carry terms, including interest rates, physical storage costs, and convenience yield. These findings are consistent with Working's cost of carry theory. Two types of convenience yield of the LME metals markets are estimated. One is traditional residual-type convenience yield and the other is an option-based convenience yield following an approximation proposed by Heaney (2002). The stationarity of these two series of convenience yield further confirms the earlier finding that a quad-variate cointegration exists for the metals markets on the LME.

This study reconciles previously inconsistent findings regarding the cointegration of temporal prices in the presence of non-stationary interest rates. It demonstrates the importance of including costs of physical storage in cointegration analyses of temporal prices. It also illustrates the importance of using data with a fixed time to maturity.
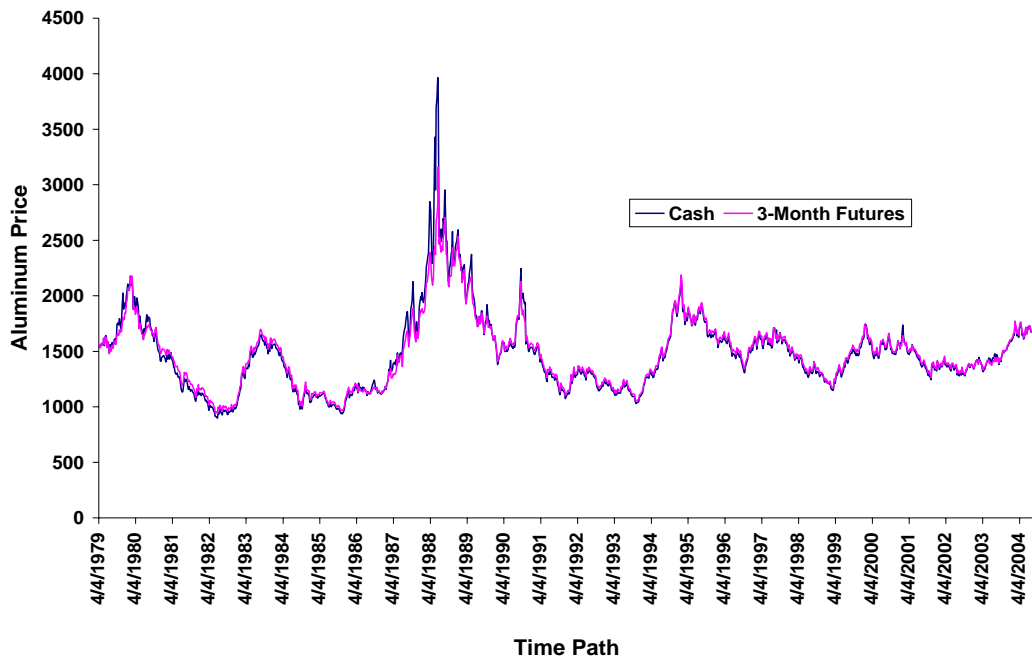
Figure 2.1: LME aluminum cash and 3-month futures prices (weekly, in U.S. dollars), 1979-2004.
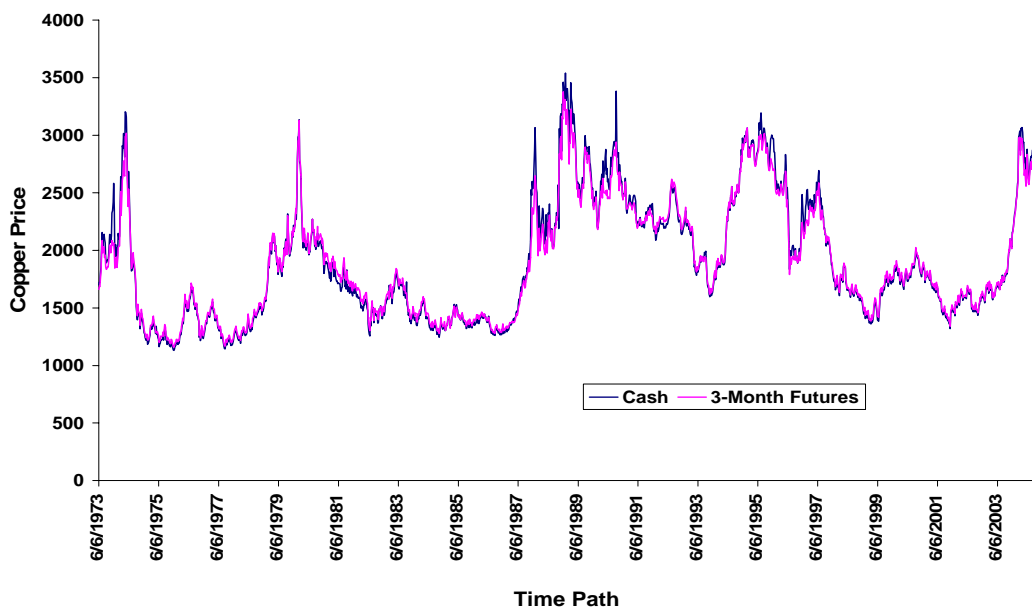
Figure 2.2: LME copper cash and 3-month futures prices (weekly, in U.S. dollars), 1973-2004.
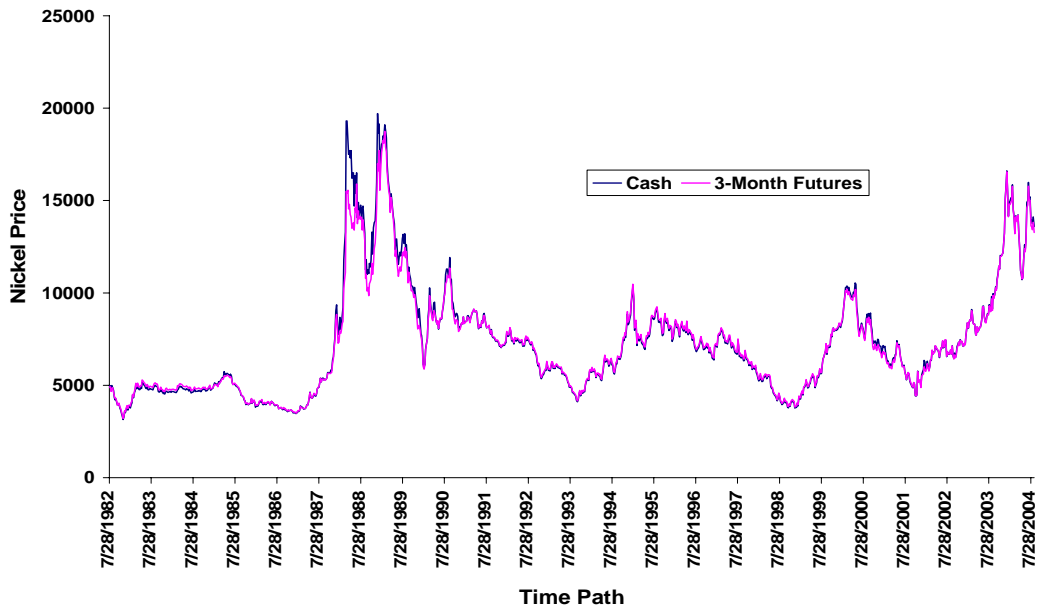
Figure 2.3: LME nickel cash and 3-month futures prices (weekly, in U.S. dollars), 1982-2004.
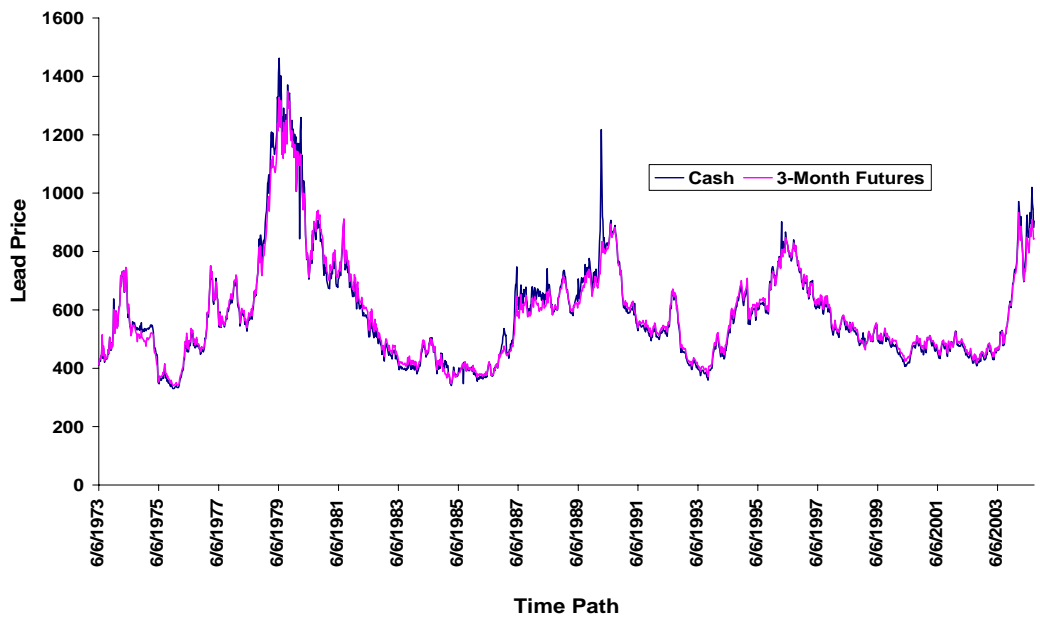
Figure 2.4: LME lead cash and 3-month futures prices (weekly, in U.S. dollars), 1973-2004.
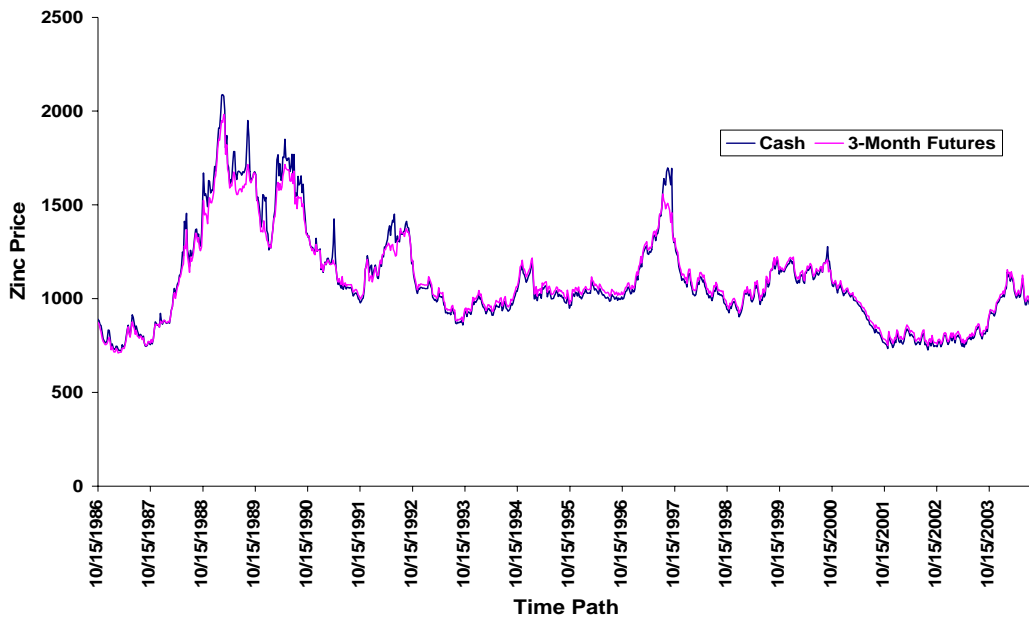
Figure 2.5: LME zinc cash and 3-month futures prices (weekly, in U.S. dollars), 1986-2004.
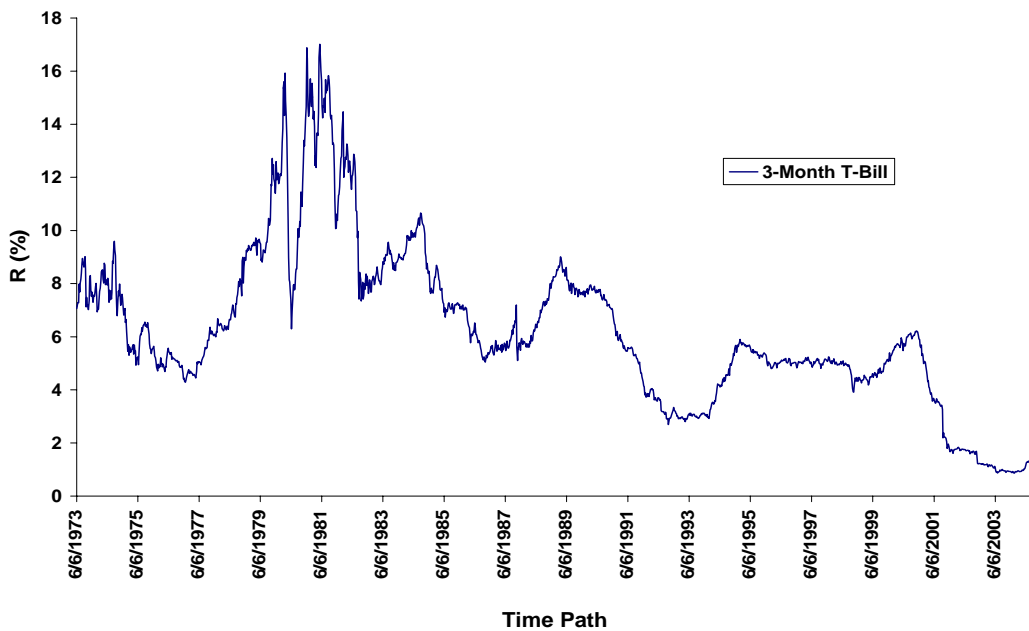
Figure 2.6: U.S. 3-month treasury-bill rates (weekly), 1973-2004.

Price Series	Statistical Test		
	ADF[a,b]	PP[a,b]	KPSS[a,b]
Aluminum Cash	-2.69	-2.76	0.17
Aluminum 3-Month	-2.35	-2.44	0.19
Copper Cash	-2.44	-2.65	0.93**[c]
Copper 3-Month	-2.18	-2.58	0.96**
Nickel Cash	-1.58	-1.94	0.59*
Nickel 3-Month	-1.92	-1.95	0.64*
Lead Cash	-2.54	-2.67	0.16
Lead 3-Month	-2.42	-2.55	0.16
Zinc Cash	-2.21	-2.22	0.82**
Zinc 3-Month	-1.97	-2.13	0.70*
3-Month T-Bill Rate	-1.64	-1.78	3.10**

Notes: (a) Null of both ADF and PP tests is that the time series is a unit root process; null of KPSS test is that the time series is stationary. (b) Critical values for ADF and PP tests at the 5% and 1% test level are -2.87 and -3.44, respectively; Critical values for KPSS test at the 5% and 1% test level are 0.15 and 0.22, respectively. (c) * and ** indicate significance at the 5% and 1% test level, respectively.

Table 2.1: Testing the stationarity of cash and 3-month futures prices for 5 metals on the LME, and 3-month US T-Bill rates, 1973-2004.

Metals	ADF[a,b]	PP[a,b]	KPSS[a,b]
Aluminum	-5.00**[c]	-7.07**	0.25
Copper	-6.66**	-10.75**	0.37
Nickel	-2.78	-5.38**	0.29
Lead	-7.16**	-10.70**	0.22
Zinc	-7.20**	-7.02**	2.08**

Note: (a) The null of both ADF test and PP test is that the examined time series is a unit root process while the null of KPSS test is that the examined time series is stationary. (b) Critical values for ADF and PP tests at the 5% and 1% test level are -2.87 and -3.44, respectively; Critical values for KPSS test at the 5% and 1% test level are 0.15 and 0.22, respectively. (c) * and ** indicate significance at the 5% and 1% test levels, respectively.

Table 2.2: Testing stationarity of the series constructed from cash and 3-month futures prices for five metals, LME, 1973-2004.

Metal by Number of Cointegrating Relationship[a]	Without Linear Trend[b,c]		With Linear Trend	
	Trace	λ_{max}	Trace	λ_{max}
Aluminum				
0	77.08**[d]	70.44**	77.07**	70.43**
1	6.64	6.64	6.64	6.64
Copper				
0	70.41**	64.14**	70.27**	64.12**
1	6.27	6.27	6.16	6.16
Nickel				
0	36.12**	31.99**	35.66**	31.91**
1	4.12	4.12	3.76	3.76
Lead				
0	80.23**	74.28**	79.96**	74.27**
1	5.95	5.95	5.69	5.69
Zinc				
0	63.33**	58.24**	63.27**	58.22**
1	5.09	5.09	5.05	5.05

Notes: (a) Null hypothesis is that r cointegrating relations exist. (b) Two cases are tested: one assumes the data have no linear trend; the other assumes the data have a linear trend. (c) The number of lag lengths used for the regression equations were one for aluminum, two for copper, and three for lead. (d) * and ** indicate significance at the 5% and 1% test levels, respectively. Source: Original Calculations.

Table 2.3: Testing the number of cointegrating relations between cash prices and 3-month futures prices (bivariate model) for selected metals, LME, 1973-2004.

Summary Statistics	June 1973 – August 2004	January 1979 – Dec. 1984
Mean	-0.0034	-0.0054
Minimum	-2.47	-2.47
Maximum	2.22	2.22
Standard Deviation	0.28	0.54

Table 2.4: A comparison of interest rates variability for two different sample periods of 1973-2004 and 1979-1984.

Metals	ADF[a]	PP[a]	KPSS[a]
Aluminum	-2.31	-2.75	0.92**[b]
Copper	-4.50**	-10.08**	0.51*
Lead	-2.38	-7.57**	0.96**

Note: (a) The null of both ADF test and PP test is that the examined time series is a unit root process while the null of KPSS test is that the examined time series is stationary; (b) * and ** indicate significance at the 5% and 1% test levels, respectively.

Table 2.5: Testing stationarity of the series constructed from cash and 3-month futures prices for three metals, LME, 1979-1984.

Metal by Number of Cointegrating Relationship[a]	Without Linear Trend[b]		With Linear Trend[b]	
	Trace	λ_{max}	Trace	λ_{max}
Aluminum				
0	22.83*[c]	21.14**	22.09**	20.93**
1	1.69	1.69	1.16	1.16
Copper				
0	37.08**	34.98**	36.93**	34.97**
1	2.10	2.10	1.96	1.96
Lead				
0	15.1	12.78	12.93	12.36
1	2.32	2.32	0.57	0.57

Note: (a) The null hypothesis is that r number of cointegrating relationships exist; (b)Two cases are tested: one assumes the data have no linear trend; the other assumes the data have a linear trend; (c) * and ** indicate significance at the 5% and 1% test levels, respectively.

Table 2.6: Testing the number of cointegrating relationships between cash and 3-month futures prices (bivariate model) for three metals, LME, 1979-1984.

Metal by Number of Cointegrating Relationship[a]	Without Linear Trend[b,c]		With Linear Trend	
	Trace	λ_{max}	Trace	λ_{max}
Aluminum				
0	35.22**[d]	28.13**	34.79**	28.13**
1	7.09	7.09	6.67**	6.67**
Copper				
0	32.12**	30.39**	31.87**	30.38**
1	1.74	1.74	1.50	1.50
Nickel				
0	12.31	10.10	11.56	9.96
1	2.21	2.21	1.59	1.59
Lead				
0	26.73**	25.02**	26.09**	25.02**
1	1.71	1.71	1.07	1.07
Zinc				
0	39.39**	35.96**	39.37**	35.96**
1	3.43	3.43	3.41	3.41

Notes: (a) Null hypothesis is that r cointegrating relations exist. (b) Two cases are examined: one assumes the data have no linear trend; the other assumes the data have a linear trend. (c) The number of lag lengths used for the regression equations were one for aluminum, two for copper, and three for lead. (d) * and ** indicate significance at the 5% and 1% test levels, respectively. Source: Original Calculations.

Table 2.7: Testing the number of cointegrating relations between cash prices and 3-month futures prices (bivariate model) for selected metals, LME, 1994-2004.

Convenience Yield	Statistical Test		
From Cost of Carry	ADF[a]	PP[a]	KPSS[a]
Aluminum	-5.56**[b]	-6.61**	0.26
Copper	-4.32**	-4.11**	1.04**
Nickel	-3.72**	-3.33*	0.92**
Lead	4.70**	-5.18**	0.29
Zinc	-5.14**	-5.67**	0.27

Notes: (a) Null of both ADF and PP tests is that the time series is a unit root process; null of KPSS test is that the time series is stationary. (b) * and ** indicate significance at the 5% and 1% test level, respectively.

Table 2.8: Testing the stationarity of weekly convenience yield calculated from the cost of carry model for selected LME metals, 1994-2004.

Convenience Yield	Statistical Test		
Heaney's Method	ADF[a]	PP[a]	KPSS[a]
Aluminum	-5.82**[b]	-5.74**	0.36
Copper	-7.44**	-8.19**	0.38
Nickel	-4.45**	-9.55**	0.54*
Lead	-8.40**	-9.05**	0.19
Zinc	-9.15**	-10.76**	0.62*

Notes: Explanations for a and b are the same as in table 8.

Table 2.9: Testing the stationarity of monthly convenience yield calculated using Longstaff-Heaney's approach for selected LME metals, 1994-2004.

Metal by Number of Cointegrating Relationship[a]	Without Linear Trend[b,c]		With Linear Trend	
	Trace	λ_{max}	Trace	λ_{max}
Aluminum				
0	62.79**[d]	38.09**	58.39**	37.92**
1	24.70	12.61	20.47	12.00
2	12.09	9.47	8.47	8.37
3	2.61	2.61	0.10	0.10
Copper				
0	67.70**	42.03**	66.95**	42.03**
1	25.66	16.65	24.92	16.39
2	9.01	7.62	8.53	7.22
3	1.38	1.38	1.31	1.31
Nickel				
0	49.25	32.10*	47.74	31.79*
1	17.15	10.76	15.95	10.06
2	6.39	5.37	5.88	5.27
3	1.02	1.02	0.61	0.61
Lead				
0	63.17**	40.55**	62.04**	40.53**
1	22.62	16.02	21.52	15.76
2	6.60	5.26	5.76	4.57
3	1.34	1.34	1.19	1.19
Zinc				
0	73.08**	49.76**	72.73**	49.69*
1	23.32	16.49	23.04	16.37
2	6.83	5.95	6.67	5.93
3	0.88	0.88	0.74	0.74

Notes: Explanations for a, b, c, and d are the same as in table 7.

Table 2.10: Testing the number of cointegrating relations between cash prices, 3-month futures prices and physical storage costs of LME metals, and US 3-Month T-Bill rates, 1994-2004.

CHAPTER 3

Updating the Estimation of the Supply of Storage Model

3.1 Introduction

Since Working's seminal articles appeared in 1948 and 1949, the supply of storage has been investigated extensively. The variables most commonly included in empirical estimations of the supply of storage models are the inter-temporal spread between a distant futures price and a nearby futures price or cash price, interest costs, physical storage costs (if available), and measures of stocks, using the stocks-to-use ratio. Furthermore, while it is commonly accepted that rational economic agents will store only when the net return to storage is positive, it is common to observe that stocks are held even when prices are expected to decline. A robust debate has ensued regarding this apparent irrational behavior. Much of this debate has focused on the idea of convenience yield, a concept introduced by Kaldor (1939). The specifics of the definition of convenience yield have varied over time, but the core idea is that a benefit accrues from having immediate access to stocks.

This study proposes to update the measurement of the supply of storage model by incorporating recent developments in the theoretical and empirical literature as well as markets. Recent theory (see Khoury and Martel (1989) for an early example) suggest

that price variability should influence the level of stockholding, yet no study could be located which has incorporated a measure of price variability into an estimation of the price of storage equation using observed data. Recent empirical work has begun to explore how to measure or at least develop an acceptable proxy for convenience yield, which historically has been unobserved variable (see Brennan (1991), Milonas and Thomadakis (1997a and 1997b) and Heaney (2002)). Last, the emergence of options trading provides the opportunity to estimate a market determined measure of volatility that is contemporaneous with the estimate of the inter-temporal price spread conventionally used in empirical estimates of the supply of storage curve. This study proposes to incorporate each of these considerations into an empirical estimation of the supply of storage curve for U.S. soybeans.

The rest of this chapter is structured as follows. Section II reviews the literature related to the supply of storage. In Section III, a supply of storage model is developed, measurement issues are discussed including Heaney's (2002) proposed proxy for the unobservable variable of convenience yield, the data is discussed, and estimation issues are addressed. Results of the analysis are presented in Section IV. The last section contains a summary, conclusions, and suggestions for further research.

3.2 Literature Review

Studies of the price of storage theory can be grouped into four categories. The first contains studies that estimate the price of storage theory empirically and often extends the theory. The second challenges the existence of convenience yield by arguing that convenience yield, i.e. the holding of stocks when the net return to storage is

negative, is an aggregation phenomenon at the market level that disappears when storage is examined at local locations. The third group of studies also challenges the existence of convenience yield, but argues that heterogeneity in a factor that affects storage, such as risk aversion, information, and transportation costs, explains the apparent existence of convenience yield at the market level. The fourth group of studies models convenience yield using option pricing theory. Each group is reviewed below.

3.2.1 Empirical Studies of the Supply of Storage

In his seminal paper, Working (1948, 1949) posited that an inter-temporal price spread, i.e., the difference between a nearby and a distant price for the same commodity, is a return to storing the commodity over the time interval. Thus, negative inter-temporal spreads (i.e., nearby price exceeds distant price) and positive inter-temporal spread both are a market determined return to storage. Working used Kaldor's (1939) idea of convenience yield to explain the holding of stocks when inter-temporal spreads were negative. Kaldor argued that convenience yield is the benefits that accrue to a stock holder from being able to continue producing during a time of scarcity and from avoiding the cost of ordering frequent deliveries and/or waiting for deliveries. Working argued that this convenience yield would be greatest when stocks were small and smallest (even zero) when stocks were large. In essence, Working argued that convenience yield offset the loss from the expected decline in price forecast by the inter-temporal spread.

Telser (1958) develops a theory of stockholding in the presence of futures markets. Demand and supply functions for storage in a two-period model are posited. Convenience yield is used to explain the holding of stocks when the inter-temporal spread

between nearby and distant futures contract is negative. As predicted by his theory, the inter-temporal spread for cotton and wheat is inversely related to the size of stocks over the 1926 -1954period.

Brennan (1958) develops theoretical demand and supply functions for storage in the context of a two-period model with uncertainty. A profit maximizing storage firm equates expected marginal revenue and marginal cost from storage. Marginal cost equals marginal outlay on physical storage plus a marginal risk premium minus marginal convenience yield. For several agricultural commodities, Brennan plots the relationship between end-of-month stocks and net marginal storage cost, which is measured as the inter-temporal price spread minus marginal outlays for physical storage. A negative relationship is found.

Weymar (1966) develops an inter-temporal pricing model which reveals that the inter-temporal spread between cash and future prices is a function of expected inventory behavior, not current inventory as Working posited. Weymar argues that Working's price of storage model is likely to hold when the expected future inventory pattern can be approximated by current inventory level. This scenario is likely to hold for agricultural commodities with a limited harvest period so that inventory declines continuously until the next harvest.

Tomek and Gray (1970) note that futures prices not only guide storage decision but also future production decisions. They compare the performance of the futures markets for corn, soybeans, and Maine potatoes in forecasting the price of the harvest futures contract at harvest during the preceding spring planting time. The authors find that, compared with the Maine potato futures market, the corn and soybean futures

markets provide a more accurate forecast but that the spring price forecast varies more from year to year. They attribute these findings to the fact that corn and soybeans are stored between crop years, whereas Maine potatoes are not. Stocks stored between crop years connect the prices in the two years.

Gray and Peck (1981) analyze the pricing performance of the Chicago Board of Trade (CBOT) wheat futures during delivery. The analysis was prompted by a Commodity Futures Trading Commission (CFTC) order that terminated trading in the CBOT 1979 March wheat futures contract. Their analysis does not support the CFTC's conclusion that a distortion existed. The inter-temporal spreads involving the 1979 March contract were similar to the historical relationship between these spreads and U.S. stocks of soft red wheat and in particular to soft wheat stocks at Chicago. However, unlike Working, they find that the March spreads are no longer related to U.S. wheat stocks. They attribute this finding to changes in the U.S. wheat market.

Using data from the U.S. wheat market from the 1970s, Sharples and Holland (1981) find that publicly-held stocks displace, at least in part, privately held stocks. Specifically, they find that a one bushel increase in wheat stocks held in the publicly-subsidized Farmer Owned Reserve increased total U.S. wheat stocks by 0.86 bushels.

Thompson (1986) estimates price of storage equations using New York and London futures prices between 1964 and 1982 for cocoa and futures prices between 1973 and 1982 for coffee. A relationship is found between world stocks of cocoa carried between crop years and the price spread involving the September (old crop) and December (new crop) contracts. However, no relationship is found between various

measures of spreads and world stocks for coffee. Although the relationship is highly variable, the best fit for a coffee price of storage curve is obtained using U.S. stocks.

Fama and French (1987) test both Kaldor-Working's theory of storage and Keynes' theory of risk premium. They use data for 21 commodities, including metals, agricultural, and wood products. To test the theory of storage, they regress the cash-futures basis against the nominal interest rate and monthly seasonal dummies. They find consistent evidence that the basis varies one-for-one with the nominal interest rate and that seasonals exist in the basis for many of the seasonally produced agricultural commodities. Both results support the theory of storage. To test for a risk premium, they regress the difference between the futures price at time t for maturity T and the cash price realized at time T against the cash-futures basis at time t. As a group, the evidence for a risk premium is mixed. The authors conclude that they find more evidence in support of the theory of storage than the risk premium theory.

Brennan (1991) posits several theoretical models, each with a different specification of convenience yield. Maximum Likelihood estimates of the models are reported for precious and commercial metals over several sample periods from January 1966 though December 1984. The estimated value of convenience yield differs significantly from zero for most of the metals and sample periods for only one of the four models. The estimates of convenience yield derived from this model are negatively related to the level of stocks, consistent with Kaldor's and Working's characterization of convenience yield.

Consistent with the price of storage theory, Heaney (1998) finds that a single cointegrating vector exists among a constant term, interest rate, three month lead futures

price at the London Metals Exchange (LME), cash LME lead price, and the total stocks held in LME-approved warehouses. Physical storage cost is assumed to be a fixed proportion of the spot price, and thus is part of the constant term. Stocks are used to proxy for variables related to the level of stocks, with the two most likely being convenience yield and risk premium. The data involves quarterly observations from March 1970 through June 1995.

Sorensen (2002) develops a pricing model that includes the seasonality of prices found in the term structure of futures prices. The model is estimated using weekly futures data for corn, soybeans, and wheat traded at the Chicago Board of Trade between January 1972 and July 1997. Consistent with Kaldor and Working, an inverse relationship is found between convenience yield and the ratio of U.S. stocks to production.

3.2.2 Convenience Yield as a Result of Mis-measurement

Recent research calls into question whether the concept of convenience yield is needed to explain the observation of storage in the presence of an expected loss from storage. One alternative explanation attributes this supposedly irrational behavior to measurement error. Measurement error occurs because stocks are aggregated across multiple local areas while price is measured at the market level. The argument implies that the use of disaggregated stock and price data should diminish the observation of storage in the presence of an expected loss.

Wright and Williams (1989) construct a model of two closely related substitute commodities linked by a production transformation technology. An example of their model is commodities stored at different locations with transportation being the

transformation function. The model utilizes a partial equilibrium theory of investment under uncertainty and assumes a competitive, risk neutral firm. The model reveals that such a firm may carry stocks even when the expected return to storage is negative. This supposedly irrational behavior will occur if the marginal cost of transforming one commodity into the other commodity is expected to be higher in the current period than in the next period. The authors show that a nonlinear production technology can result in this situation. Given that nonlinear production technologies are not unusual, Wright and Williams conclude that the incidence of stockholding in the presence of a negative inter-temporal spread will diminish sharply when stocks and inter-temporal spreads are examined at specific locations and for specific grades rather than at the aggregate level of the market. They provide empirical support for their conclusion by estimating a price of storage curve using futures prices of the New York coffee "C" contract. Consistent with their model, they find that no substantial amount of deliverable coffee stocks, a measure of local stocks, is held when the coffee futures spread is negative.

Using detailed data on the spatially dispersed marketing system of Western Australia, Brennan, Williams, and Wright (1997) construct a mathematical programming model of shipments and storage by location. Solving the mathematical program reveals that stocks are not held at their destination, the port, but instead are held at remote locations. As price for current delivery to the port increases, stocks are transported from increasingly remote areas. A plot of the port's spread against aggregate stocks reveals a typical supply-of-storage curve consistent with the existence of convenience yield. However, the first order condition of the mathematical program implies that storage will not occur at an individual site unless the net return to storage is expected to exceed

physical storage and interest expenses. Hence, the price-of-storage curve identified in other studies can result from the use of inappropriately observed price and stock location (or grade) data. In other words, like Wright and Williams, Brennan, Williams and Wright conclude that convenience yield is not required to explain the traditional price-of-storage relationship.

Benirschka and Bindley (1995) also examine the interplay between storage and different locations. Their model consists of geographically dispersed producers supplying a central market with grain of uniform quality. Because transportation cost increases as distance increases, price declines as the location of production becomes more distant from the central market. At the beginning of the marketing year, the central market draws its supply from nearer locations because of their lower transportation cost. As the marketing year continues, supply is drawn from increasingly distant locations. The interest cost of storage is lower at locations further from the central market because of their lower prices. Thus, price increases at an increasingly slower rate as the location becomes more remote. However, this observation means that prices at the central market also will increase at an increasingly slower rate as the marketing year progresses. This expected slowing in the rate of increase in prices at the central market will lead to inter-temporal spreads at the central market that will not cover the cost of storage. However, this apparent existence of convenience yield at the central market is an artifact of aggregating stocks at different locations. Benirschka and Bindley find evidence largely consistent with their model when they examine the location of storage and behavior of cash prices in the U.S. corn market between 1968 and 1994. Frechette and Fackler (1999) extends Benirschka and Bindley's analysis by assuming storing costs are additive and by

treating space as a continuum rather than a set of discrete points. Similar to Benirschka and Bindley, they show that negative inter-temporal price spreads can result when stocks are held at multiple locations and transportation is costly. However, Frechette and Fackler estimate that, compared to stocks, the impact of location was small on the inter-temporal spread between the December and March corn futures contracts over the period between 1949 and 1993. This finding casts doubt on the ability of the spatial heterogeneity of storage to explain the empirically observed magnitudes of negative inter-temporal spreads and associated negative net returns to storage.

3.2.3 Negative Inter-Temporal Spreads and Non-Convenience Yield Explanations

Williams (1987) posits a two-period model of an industry composed of risk-neutral processors who confront higher transaction costs in the cash market than in the futures markets and have a nonlinear cost of production function. The model reveals that, given these conditions, a firm will use futures markets even if the firm is not risk averse and the firm will rationally hold stocks even when inter-temporal price spreads are negative.

Chavas (1988) develops a multi-period model in which price is uncertain and competitive speculative storage agents have varying degrees of risk averseness. The model reveals that if the storage firm is sufficiently risk averse it will hold stocks in the presence of a negative inter-temporal spread. Khoury and Martel (1989) develop a three-date storage model, assuming utility maximization, risk aversion, asymmetric information with speculators being better informed than hedgers, and the existence of a futures market

in which stocks are hedged. The model reveals that under these conditions storage can occur when inter-temporal spreads are negative. The intuition is that the negative spread causes holders of stocks to expect the futures price to fall, thus generating a profit on the futures hedge that exceeds the loss from holding the cash commodity.

In an approach similar to Khoury and Martel, Frechette (1999) develops a market model in which fundamental and noise traders exist. The different expectations and risk aversion that characterize these two types of traders can lead to stocks being held when inter-temporal spreads are negative. Empirical tests are conducted using data from the US copper, corn, and wheat markets. Results are consistent with the theoretical model. Frechette (2001) extends this analysis by developing a more inclusive model of heterogeneous expectations. It also reveals that storage can occur when inter-temporal spreads are negative.

Chavas, Despins, and Fortenbery (2000) argue that Kaldor's description of convenience yield should be broadened to include all transaction costs associated with holding and handling stocks, including the opportunity cost of time and gathering information. They posit a model based on an economic agent who makes resource allocation decisions over time and manages two goods. One good is storable; the other is a consumer good. Their model reveals that the existence of transaction costs will cause stocks to be held when inter-temporal price spreads are negative. Empirical tests using data from the U.S. soybean market over the 1960 - 1995 period find that transaction costs are positive and play a significant role in storage and pricing behavior. This model suggests that convenience yield is related to the expected change in inventory, not to the level of stocks as postulated by Kaldor and Working. Similarly, the conventional measure

of convenience yield applies only to economic agents currently holding stocks. In contrast, transaction costs apply to all storage agents, whether or not they currently hold stocks.

3.2.4 Convenience Yield as an Option

Heinkel, Howe, and Hughes (1990) note that convenience yield can be recast as an option value available only to holders of stocks. The option value is derived from the ability to sell the cash commodity for a higher price should it materialize while the stock is being held. They construct a three-date theoretical model in which demand is uncertain. Storage agents are assumed to be risk neutral and sign a contract at time 0 to sell any stock remaining at time 2 for the futures price quoted for time 2 at time 0. As with the traditional view of convenience yield, the model reveals that the level of stocks is negatively related to the option value measure of convenience yield. It also reveals that the option value measure of convenience yield is positively related to the marginal cost of production and negatively related to the serial correlation in spot prices. The higher the marginal cost of production, the less likely current production will occur to meet unexpected demand. Thus, the higher the option value to sell at intermediate time 1. The more negative the serial correlation among spot prices, the more likely that low (high) futures prices at time 0 are associated with a high (low) cash price at time 1. Thus, the option value of holding stocks at time 0 is higher (lower).

Bresnahan and Spiller (1986) note that Keynes (1930) proposed two explanations for the existence of negative inter-temporal spreads. One was the commonly-investigated risk premium theory. The second was the "liquid stocks" theory. The latter argues that, if

the possibility of a stock-out exists, the cash price can exceed futures prices. In a stock-out situation, no stocks exist. In such a situation, the cash price must be high enough to postpone demand until the arrival of new supplies. Bresnahan and Spiller show that, if uncertainty about supply exists, the probability of a stock-out occurring is always positive.

Like Bresnahan and Spiller, Milonas and Thomadakis (1997a and 1997b) also construct a three-date storage model in which a storage decision is made at the intermediate date between the beginning and end of the crop cycle. They find that the decision to store or sell at the intermediate date has a payoff structure similar to a call option. This call option, which is a measure of convenience value, has value if a stock out is a possibility at the intermediate date. Their model implies that the value of the call option is positively related to the variability of cash price, and inversely related to the size of stocks, the time left until the end of the crop cycle, and the correlation between the intermediate period cash and futures prices. Milonas and Thomadakis test their model using data from the copper, corn, soybean, and wheat markets for the period 1966 to 1995. Fisher's option valuation model is used to derive the call option estimate of convenience yield. Support is found for each of the hypothesized relationships.

Heaney (2002) estimates the call option value of convenience yield by adopting a valuation technique proposed by Longstaff (1995). Longstaff used option pricing theory to estimate the upper bound on the value of liquidity in financial markets when restrictions exist on selling an asset. The upper bound equals the present value of the cash flow that could have been obtained if, during the time the asset was illiquid, a trader with perfect foresight could have sold the asset at what was known to be its highest price.

Longstaff shows that this value equals the value of a call option with a strike price equal to the price of the asset when the restriction on selling the asset in the market existed.

Heaney adopts Longstaff's technique to compute the value of profitable trading opportunities associated with holding a cash position instead of holding a futures position in an asset. The strike price of this call option is the futures price. Value of the call option is a nonlinear function of the price volatility of the underlying cash asset, price volatility of the futures contract, and the time to maturity of the futures contract. Heaney computes the call option value of convenience yield using data from cash and futures contracts traded for copper, lead, and zinc at the London Metals Exchange. He then compares the observed futures prices with theoretical futures price derived from the cost of carry model. Inclusion of the estimated convenience yield in the calculation significantly reduces the difference between the observed and theoretically derived futures prices.

Litzenberger and Rabinowitz (1995) posit that the current value of oil in a reserve can be conceptualized as the value of a call option written at a strike price equal to the extraction cost of the marginal producer. They show that the value of oil in reserve also equals the sum of the discounted difference between the futures price and the extraction cost, plus the value of the option to forego production in the future period. Both their two-period and multi-period models reveal that the existence of the call value on future production will cause the discounted futures price to be less than the current cash price at all times in the oil market. Furthermore, the futures price will be less then the current cash price if the uncertainty about future price is sufficiently large. Their model implies that, when riskiness increases, oil production is non-increasing and inter-temporal oil price spreads are non-decreasing. These implications are consistent with empirical tests

conducted using data on U.S. oil production, U.S. oil reserves, and west Texas intermediate futures and options prices over the period from December 1986 through December 1991.

Richter and Sorensen (2002) posit a model which assumes that commodities exhibit seasonality patterns in both cash price level and volatility. Price dynamics are modeled using stochastic differential equations that are heterogeneous in time and are affine asset pricing models. Their model is estimated using a quasi maximum likelihood approach and a panel data of soybean futures and options prices from the Chicago Board of Trade for October 1984 through March 1999. Seasonal patterns are found in both volatilities and convenience yields. Consistent with the price of storage theory, a negative relationship is found between stocks and convenience yield. However, in contrast to the studies discussed above, no significant correlation is found between convenience yield and volatility. This finding is inconsistent with the argument that convenience yield can be modeled as a timing option.

Fackler and Livingston (2002) examine the option value of storage from a different perspective. They argue that in most situations the grain storage and marketing decisions of farmers are irreversible because high transaction costs prohibit the replenishment of grain once it is sold. This irreversibility creates an option value similar to that found in other irreversible economic decisions, such as wilderness preservation and private investments with large sunk costs. When an investment is irreversible, the optimal decision rule is to invest if the investment's net present value exceeds the sunk investment cost plus the American option value of waiting. A model of dynamic stockholding is developed for a risk neutral farmer. The marketing problem is found to

have a number of commonalities with the optimal stopping problem of determining when to exercise an American option. The optimal sales rule reduces to the following condition based on the current price: sell everything when the current price is high; otherwise sell nothing. Numerical computation is used to calculate the cutoff between high and low prices for soybean storage in central Illinois over the period from November 1975 through October 1997. The results reveal that including the value of the American option in the marketing strategy substantially increases storage returns.

3.3 Supply of Storage

This section presents the conventional supply of storage model, as well a simplified version of recent supply of storage models that incorporate risk. Next, the critical issue of measuring the variables is discussed. Included in this discussion is a recently proposed technique for generating a proxy measure of convenience yield.

3.3.1 Supply of Storage Models

The most commonly estimated price of storage equation is:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \varepsilon_t \tag{3.1}$$

where y_t =   stock to use ratio at time t,

$x_{1,t}$ =   storage cost adjusted price spread at time t,

ε_t =   random error term, and

β_0, β_1 are coefficients.

61

Stock-to-use ratio is used instead of stock level because, everything else constant, the level of stocks carried by storage agents is expected to increase as the size of the market. Size of the market has conventionally been measured by quantity of consumption. The storage cost adjusted price spread is most often measured as an inter-temporal price spread involving a distant futures price and either a nearby futures price or a cash price, with the cost of storage adjusted.

More recent models incorporate risk. The model which follows is a simplified version of Khoury and Martel's (1989) supply of storage model. Their model is a two-period model with a risk averse representative storage firm. The firm owns quantity Q of a commodity at time 0, the first period. It must choose between selling all, part, or none of Q at time 0 and storing the remainder for sale at the cash price which prevails at time 1. A futures market is assumed to exist, thus providing information that the firm can use to predict the spot price at time 1. Unlike the model presented in this paper, Khoury and Martel assume that the firm hedges the stocks it does not sell at time 0.

Assume the storage firm has a constant (local) relative risk coefficient, γ. Thus, its utility function can be written as:

$$U(R) = (\frac{1}{\gamma})(1 - e^{-\gamma R}) \tag{3.2}$$

This representative storage firm seeks to maximize its expected utility from the revenue it expects to generate from its storage and marketing strategy by the end of time 1 as of time 0. Its utility maximization problem can thus be stated as:

$$Max_X E[U(R_{0,1})] = Max_X E[(\frac{1}{\gamma})(1 - e^{-\gamma R_{0,1}})] \tag{3.3}$$

where, $R_{0,1} = (Q_0 - X_0)S_0 \exp(1+r) + X_0(S_{0,1} - C)$ (3.4)

Q_0	=	quantity of commodity owned at time 0,
X_0	=	quantity of commodity stored at time 0,
r	=	risk free interest rate prevailing at time 0,
S_0	=	spot price of the commodity at time 0,
$S_{0,1}$	=	spot price of the commodity at time 1 expected at time 0, and
C	=	physical storage costs per unit during the storage period.

If $R_{0,1}$ is distributed normally as $N(\mu_{R_{0,1}}, \sigma^2_{R_{0,1}})$, equation (3.3) can be rewritten as:

$$\Phi(X) = \mu_{R_{0,1}} - (\frac{\gamma}{2})\sigma^2_{R_{0,1}}$$ (3.5)

where, $\mu_{R_{0,1}} = E((Q_0 - X_0)S_0 \exp(1+r) + X_0(S_{0,1} - C)$ (3.6)

and $\sigma^2_{R_{0,1}} = X^2\sigma^2_{S_{0,1}}$ (3.7)

Substituting equations 3.6 and 3.7 into equation 3.5 and taking the first order derivative with respect to stocks X_0 yields the following relationship:

$$\frac{d\Phi(X)}{dX_0} = -S_0 \exp(1+r) + E(S_{0,1}) - C - (\frac{\gamma}{2})(2X\sigma^2_{E_{S_{0,1}}})$$ (3.8)

Rearranging the terms in equation 3.8, the optimal level of stocks, X_0*, is:

$$X_0* = \frac{E(S_{0,1}) - S_0 \exp(1+r) - C}{\gamma\sigma^2_{S_{0,1}}}$$ (3.9)

If the futures market provides an unbiased estimate of the future spot price, i.e., $F_{0,1} = E(S_{0,1})$ and the futures-cash basis at contract expiration is zero, equation 3.9 can be rewritten as:

63

$$X_0{}^* = \frac{F_{0,1} - S_0 \exp(1 + r) - C}{\gamma \sigma_{F_{0,1}}^2} \tag{3.10}$$

Equation 3.10 reveals that the representative storage firm's optimal quantity of stocks is positively associated with the storage cost adjusted spread between the cash and futures price (i.e., the numerator), and inversely related to both the firm's degree of risk aversion and the current variability of the futures price for the contract for delivery at the end of inventory holding period.

3.3.2 Variable Measurement

Measurement of risk aversion is difficult. Furthermore, a time series of risk aversion measures would be needed for storage firms. No such data set exist. Thus, risk aversion is not included in this estimation of the price of storage curve.

The storage cost adjusted spread depends on a distant futures price, nearby futures price or cash price, and storage costs. Storage costs conventionally equal the sum of physical storage costs and interest, minus convenience yield. To minimize measurement errors, it is desirable that each of these variables, along with stocks or the stocks-to-use ratio and price variability be measured contemporaneously. In this context, contemporaneous means that each variable is measured as the latest measure of the variable at the time the market was determining price. Contemporaneous data reduces measurement error by aligning the information available to the market and the price at that time. Thus, variables are not measured at different times in terms of the dynamics of market price.

The advent of options trading makes it possible to extract market determined measures not only of the level of prices and inter-temporal price spreads but also the

variability of prices. Specifically, implied volatility estimates can be derived from the options price. Implied volatility also has the advantage of being determined by the market contemporaneously with the inter-temporal price spread.

Equation 3.10 implies that the optimal quantity of stocks would be less than zero if the expected return from storage is less than zero. Not only is it impossible to transport commodities backward over time but it is also economically irrational to store commodities if the expected return from storage is less than zero. As noted earlier, convenience yield has been proposed as a return to holding the cash commodity that offsets an expected loss from storing an asset. Not only is the existence of convenience yield a highly controversial topic, but also only recently have measures of convenience yield been postulated. As noted above in the literature review, option pricing theory has emerged as the most commonly used measure. This study will utilize the method proposed by Heaney (2002), which in turn is derived from a procedure proposed by Longstaff (1995).

Consider an arbitrage model in which an arbitrager buys and holds a cash asset while selling a futures contract whenever the expected net return to storing the asset is positive, i.e., futures minus cash spread exceeds the cost of storing the asset. On the other hand, if expected net return storage is negative, the arbitrager buys a futures contract and sells the asset in the cash market. The standard arbitrage model assumes that all positions are held until futures contracts mature. However, this assumption must be relaxed when investigating convenience yield because convenience yield is greater than zero only when the inventory holder has right to use the asset at any time. Thus, Heaney proposes to modify the standard arbitrage model. Specifically, convenience yield reaches a maximum

value to a trader if the trader has perfect foresight about the market and can choose to sell the asset at the highest price which will occur between the current time and the end of the storage period. Once this trader sells the asset at its highest price, he/she will invest the proceeds at the risk-free rate, and then buy the asset on the cash market at the lower price on the futures contract maturity date. This argument is similar to the one Longstaff made in deriving a model for estimating the value of marketability (liquidity) of securities.

A mathematic representation of the maximum price over the storage period from time t to time T can be expressed as follows,

$$Max(S) = \max_{t \leq \tau \leq T} \{\exp[r(T - \tau)]S_\tau\} \tag{3.11}$$

where, t = beginning of storage period

T = end of the storage period

S_τ = maximum cash price observed at time τ, where $t \prec \tau \prec T$

The convenience yield value of holding the cash commodity can be approximated as the value of an option to sell the commodity if price rises sufficiently to generate an arbitrage profit when the commodity is bought back at the end of the storage period. The value of this option, designated as $V(S_t, T)$, is:

$$V(S_t, T) = \exp[-r(T - t)]E(Max(S)) - \exp[-r(T - t)]E(S_T) \tag{3.12}$$

Heaney argues that the value of this option (i.e., convenience yield) can be proxied through the following calculations:

$$cy_{tT} = v_{tT}(S_t, T) - v_{tT}(F_{tT}, T) \tag{3.13}$$

where cy_{tT} = convenience yield of holding stock at time t, with latest sale at time T,

$$v_{tT}(S_t, T) = \ln\{[2 + \frac{\sigma_S^2(T-t)}{2}]N[\frac{\sqrt{\sigma_S^2(T-t)}}{2}] + \sqrt{\frac{\sigma_S^2(T-t)}{2\pi}} \exp[-\frac{\sigma_S^2(T-t)}{8}]\} \quad (3.14)$$

$$v_{tT}(F_{t,T}, T) = \ln\{[2 + \frac{\sigma_F^2(T-t)}{2}]N[\frac{\sqrt{\sigma_F^2(T-t)}}{2}] + \sqrt{\frac{\sigma_F^2(T-t)}{2\pi}} \exp[-\frac{\sigma_F^2(T-t)}{8}]\} \quad (3.15)$$

σ_S^2 = variance of cash prices,

σ_F^2 = variance of futures prices,

$N(\bullet)$ = cumulative normal distribution.

Equation 3.14 provides an estimated value based on the variability of the cash price. Equation 3.15 provides an estimated value based on the variability of the futures contract for delivery at the end of the storage period. Because convenience yield is the option value of potentially selling the commodity before the end of the storage period, the difference between these two values will be related to the convenience yield. In essence, the greater the variability of the cash price relative to the futures price, the greater is the value of having the potential option to sell before the futures contract matures. Note the value of convenience yield given by equation 3.13 is taken as a percentage of the cash price.

3.3.3 Simultaneous Equation System

A causal relationship exists between convenience yield and the storage cost adjusted spread. As convenience yield increases, the storage cost adjusted price spread becomes more negative, everything else held constant. Furthermore, the optimal level of

stocks is related to the storage cost adjusted price spread, among other factors. Thus, convenience yield, storage cost adjusted spread, and stocks are determined simultaneously. Hence, the following simultaneous two equation system is proposed:

$$y_{1,t} = \alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \alpha_3 x_{3,t} + \varepsilon_t \qquad (3.16)$$

$$x_{1,t} = \beta_0 + \beta_1 x_{4,t} + \upsilon_t \qquad (3.17)$$

where $y_{1,t}=$ stock-to-use ratio at time t,

   $x_{1,t}=$ storage costs adjusted price spread,

   $x_{2,t}=$ price volatility of futures contract for delivery at the end of storage period,

   $x_{3,t}=$ price volatility of futures contract squared,

   $x_{4,t}=$ Heaney's (2002) proxy measure of convenience yield.

This simultaneous equation system incorporates more information about supply of storage, including information about price volatility and the non-observable convenience yield. A quadratic term of the volatility is included into the system in order to capture possible high-order nonlinear impacts of volatility on the stock-to-use ratio. In summary, this simultaneous equation system offers the potential to provide a richer understanding of the supply of storage theory.

3.3.4 Data

The price of storage equation is estimated using data from the U.S. soybean market beginning with stocks carried out of the 1988/89 crop year and ending with stocks carried out of the 2003/2004 crop year. The soybean market is selected because among major U.S. crops it is unique in not having acreage set aside programs. Public stocks also

have been limited in size and duration. Lastly, soybean options began trading in 1985 and soybean options are generally among the most heavily traded commodity options markets.

Data used in this study are futures prices, options on futures prices, ending stocks and consumption for the current crop year, physical storage costs, and U.S. 6-month Treasury-Bill rates. Each variable is measured as of the release of the U.S. Department of Agriculture (USDA) *World Agricultural Supply and Demand Estimates* (*WASDE*). The contemporaneous nature of this data set is a unique feature of this study.

WASDE is released each month throughout the year. It contains the latest USDA forecasts of U.S. and world supply and use balance sheets for the major grains, soybeans and soybeans products, and cotton, as well as U.S. sugar and livestock products for the current crop year. Beginning with May, it contains forecasts for the upcoming crop year.

The *WASDE* reports used in this study are the ones issued in February, April, and June. These months were selected because they are non-delivery months and thus avoid potential pricing problems that can be associated with delivery contracts. Thus, empirical estimation and analysis are conducted for these three cases. Because ending stocks are analyzed, the futures prices are for the nearby contract and for the November contract. The nearby contract is March for February, May for April, and July for June. The November contract is considered the first new crop contract. Thus, the storage intervals of February-November, April-November, and June-November bridge the old and new crop years.

Prices and option premiums are the settlement values for the first non-limit trading day after the release of *WASDE*. This collection rule allows the market to

incorporate any new supply and demand data contained in the *WASDE* released reports. Option trading on soybeans did not begin until the 1984/1985 crop year. However, substantial public stocks of soybeans existed during the 1985/86 and 1986/87 crop years, but have been almost non-existent since. Studies have documented that public stocks can displace privately held stocks and thus affect the supply of storage equation (for example, see Sharples and Holland (1981)). To avoid this issue, this study uses data for 1987/88 though 2003/04 crop years.

The futures and options prices are from a data base maintained by the AgMAS project located at the University of Illinois at Champaign-Urbana. The six month Treasury-Bill rates are collected from the U.S. Federal Reserve Bank. Physical storage costs are collected from the USDA Commodity Credit Corporation. Implied volatility is calculated using Black-Schole's option pricing model for soybean option premiums and futures prices for the November contracts.

3.3.5 Estimation Issues

Because stock-to-use ratio, storage cost adjusted inter-temporal spread, and convenience yield are determined simultaneously, correlations might exist between the error terms of the two equations. A standard econometric procedure for addressing this estimation problem is three-stage least squares (3SLS). Three-stage least square is a system method that estimates all of the coefficients of the model, forms weights, and then re-estimates the model using the estimated weighting matrix. Because heteroskedasticity and autocorrelaton have been identified as potential statistical issues when using futures

price data in, heteroskadasticity and autocorrelation consistent (HAC) covariance estimation procedures are used in conjunction with the 3SLS estimation method.

Standard hypothesis tests and statistical inferences are based on strong parametric assumptions. A critical assumption in classical multiple regression analysis is that the variable have a normal distribution. However, this assumption is generally not reasonable when using data from a small sample, leading to the potential for distorted estimation results and statistical inferences. Bootstrap is a statistical technique commonly used to improve the power of statistical tests in the presence of small sample problems.

Bootstrap methods include both a nonparametric and a parametric mode. Nonparametric bootstrap, the original bootstrap, re-samples the values of variables by drawing from the empirical distribution with replacement. Parametric bootstrap re-samples residuals. Unlike parametric bootstrap, nonparametric bootstrap does not depend on a particular class of distributions. Both procedures assume that the sample's distribution is a good estimate of the population distribution. This study uses the nonparametric bootstrap because it more effectively addresses heteroskadasticity than parametric bootstrap (Wu, 1986). Nonparametric bootstrap is usually implemented as follows: (1) draw a random sample (with replacement) from the empirical distribution of the original sample with a size equal to the size of the empirical sample; (2) calculate the statistic of interest; and (3) apply a Monte Carlo-style procedure by repeating steps one and two a large number of times. A sampling distribution of the statistic of interest is generated. This distribution is used to draw inferences about the population parameter.

Estimation of Heaney's (2002) proxy for convenience yield requires only three variables, underlying commodity cash price volatility, futures price volatility, and the

futures contract time to maturity. For this study, price volatility of the nearby futures contract (March, May, and July) are used instead of cash price volatility. Volatility of the November contract is used as the measure of futures price volatility. A historical volatility is calculated using the daily returns for the 20-trading-days immediately preceding the *WASDE* report release dates for February, April, and June. These estimation parameters are the ones used by Heaney.

3.4 Empirical Results

Storage costs adjusted price spread, or net cost of storage, is constructed as: ln(November futures price) – ln(nearby futures price + interest cost + physical storage cost over the storage window). The value of this variable is plotted against the stock-to-use ratio for the February, April, and June observation dates in Figure 1. Examination of Figure 1 reveals that the relationship between these two variables is in the form of a natural logarithm. Thus, the stock-to-use variable is measured as the logarithm of the stock-to-use ratio. Previous studies have mentioned this nonlinear relationship (see Gray and Peck, 1981, for example). The usual argument made in support of this relationship is that, once the storage cost adjusted inter-temporal price spread is zero, stocks can increase over a wide range and this spread measure will not change. A positive storage cost adjusted futures price spread leads to arbitrage opportunities that will reduce the spread to zero.

Consequently, the empirical model of simultaneous two equations system is estimated as follows,

$$\ln(y_{1,t}) = \alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \alpha_3 x_{3,t} + \varepsilon_t \qquad (3.18)$$

$$x_{1,t} = \beta_0 + \beta_1 x_{4,t} + \upsilon_t \qquad\qquad (3.19)$$

where, the variables are defined in the same way as in section 3.3.3.

Panel A of table 3.1 presents results for the traditional supply of storage model estimated by the heteroskadasticity and autocorrelation consistent (HAC) covariance estimation procedure. The estimation is conducted using Eviews 5.0. The estimated coefficients for ln(storage costs adjusted price spread) for the February, April, and June observation dates are significant at the 1% test level and have a positive sign. This result is consistent with previous studies. R-squared exceeds 0.70 for all three cases, suggesting that the price spread explains a large proportion of the year to year variation in the carryout stocks of soybeans.

Panel B of table 3.1 presents the updated supply of storage system equations estimated by 3SLS method. As with the results obtained from the HAC regression of the traditional supply of storage model, the sign on the spread variable is statistically significant at the 1% test level and is positive.

The updated supply of storage model developed in this study uses the contemporaneous implied volatility for the contract month at the end of the storage period. However, the initial regression results for this variable revealed an insignificant coefficient. Therefore, a squared term was added to evaluate if a nonlinear relationship may exist between volatility and carryover stock-to-use. The coefficients on both the linear and squared volatility terms are significant at the 1% level except for the squared term in the April regression, which is significant at the 5% level.

To help examine the nonlinear relationship between stocks and price variability, a fitted stock-to-use ratio is estimated for each observation month using the parameter of

the estimated supply of storage equation within the two equation system, mean value of the observed spreads, and observed values of implied volatility. To aid in discussion, the fitted ln values of the stock-to-use ratio are converted to stock-to-use ratios by using the exponential function. The graphs that result for all three observation periods are presented in Figure 2. The graph reveals that as volatility increases the stock-to-use ratio declines until a minimum level of 9.7%, 9.3%, and 8.7% for February, April, and June, respectively. This non-linear relationship is not consistent with the theory developed in this paper and needs to be further explored.

Turning to the second equation of the simultaneous equations system, a statistically significant negative relationship is found between Heaney's proxy for convenience yield and the storage cost adjusted inter-temporal spread. This relationship is consistent with Working's argument that convenience yield and an inter-temporal spread for a storable commodity are inversely related.

The coefficient of convenience yield is between -3 and -4 for all three observation periods. Thus, each one percent point increase in Heaney's proxy for convenience yield, which is approximated as a percentage of the cash price or nearby contract price, results in a 3 to 4 percent decrease in the price spread. R^2 for the convenience yield equations lies between 0.55 and 0.61. To present a visual picture of this regression analysis, Figure 3.3 contains a scatter-graph of the data used to estimate this relationship and a fitted curve of the ln(storage costs adjusted price spread) using the estimated parameters for each of the three observation months.

As noted earlier, the Bootstrap method has become a standard procedure to use in the case of a small sample. Results from the bootstrap analysis are presented in Table 3.2.

74

They are similar to the results reported in Table 3.1 and discussed above. The higher power associated with the Bootstrap method underscores the likely robustness of the previously discussed results.

3.5 Summary and Conclusions

This study updates the estimation of the supply of storage model to reflect recent developments in the theoretical and empirical literature. Specifically, it incorporates a measure of price variability and a proxy measure of convenience yield, and measures all variables contemporaneously. The measure of price variability is implied volatility. The proxy measure of convenience yield was suggested by Heaney (2002) based on work by Longstaff (1995). Heaney argues that convenience yield is the value of an option to sell stocks before the end of the storage period should a high price occur, and shows that the value of this option is related to the variability of the cash price and the variability of price of the distant futures contract for delivery at the end of the storage period.

The model is a simultaneous two-equation system. The first equation is a supply of storage equation, which is developed through a two period trading model and a utility maximization approach which incorporates risk. The optimal quantity of stocks is shown to be a function of storage cost adjusted inter-temporal price spread and current price variability of the futures contract for delivery at the end of the storage period. A quadratic term of the price variability is added to capture potential nonlinear impacts of price variability on stocks. The second equation in the system captures the causal relationship that exists between convenience yield and storage costs adjusted price spread. The data used to estimate this simultaneous equation model are from the U.S. soybean market for

carryout stocks of the 1987/1988 crop years through the 2003/2004 crop years. The data is measured contemporaneously to the release by the U.S. Department of Agriculture of the *World Agriculture Supply and Demand Estimates*. Three cases are analyzed for the *WASDE* reports released in February, April, and June.

The variables in the model are measured as the stock-to-use ratio, storage cost adjusted futures price spread involving the nearby and November futures prices, implied volatility derived from the November options and futures contracts, and a proxy measure of convenience yield proposed by Heaney (2002). Each variable is measured as of the release date of the USDA's *World Agricultural Supply and Demand Estimates* (*WASDE*).

Results from both the 3SLS method and the bootstrapped 3SLS confirm the positive relationship between stocks and storage costs adjusted price spread conventionally recognized in the literature. The theoretical literature has proposed a negative relationship between stock level and price variability. However, this study provides the first empirical investigation of this relationship, and finds a nonlinear relationship. Initially, as price variability increases, the carryover stock-to-use ratio declines, as suggested by theory. But, as price variability increases, eventually the relationship turns positive. Last, a negative relationship is found between the storage cost adjusted price spread and the proxy measure of convenience yield. This finding is consistent with Working's argument that convenience yield is a return to storage that can offset, at least partially, some of the loss expected from storing when the storage cost adjusted inter-temporal price spread is negative.

In summary, this study provides richer understanding of the supply of storage theory and the convenience yield theory for the U.S. soybean market. It documents the

76

important role that price volatility and convenience yield have in determining carryover stocks of soybeans. It would be useful to determine if these same results can be replicated in other commodity markets. Future research could also further examine the nonlinear relationship between price variability and stock-to-use, including the development of a theoretical model to support such a relationship. Last, the relationship between the storage spread and Heaney's proxy measure, while significant, generates an explanatory power that is between 55% and 60%. Thus, additional work is needed on the measurement of convenience yield and its relationship to the storage cost adjusted inter-temporal spread.
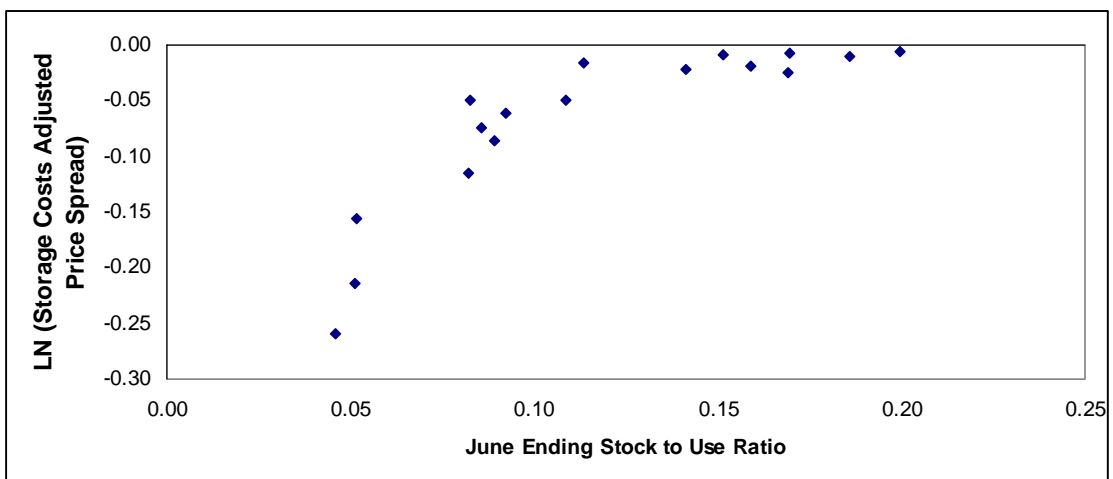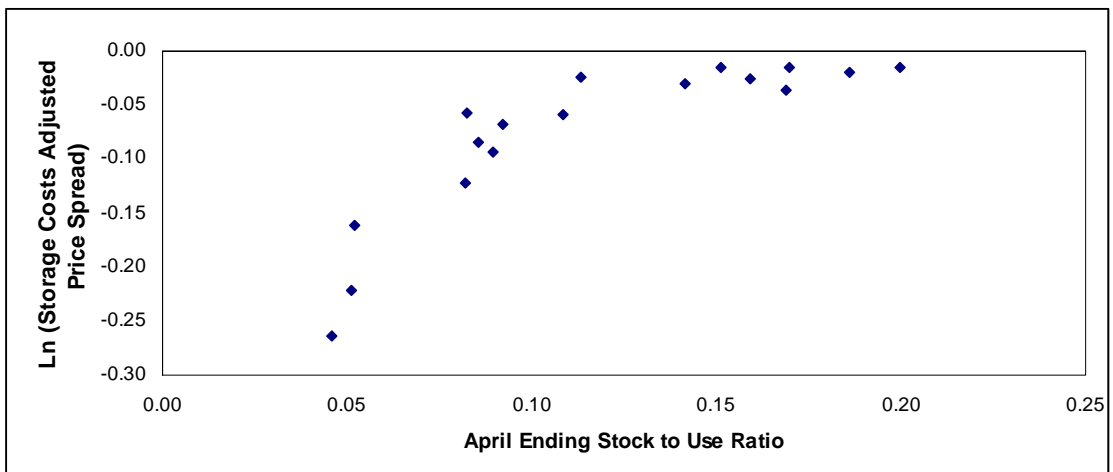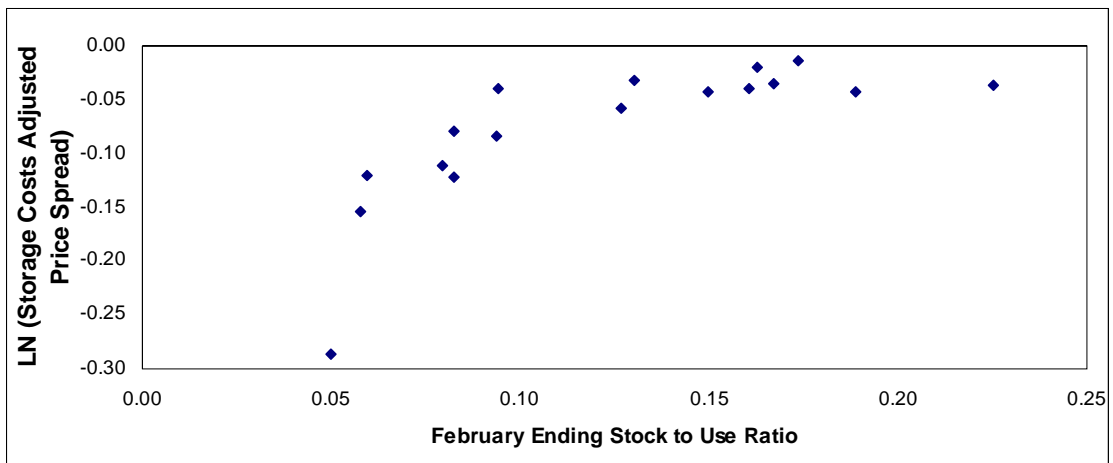
Figure 3.1: Supply of Storage Curve for Soybeans as of February, April, and June *World Agriculture Supply and Demand Estimates*, U.S., 1988-2004.

Figure 3.2: Plot of Synthetic Stock-to-Use Ratio against Implied Volatility for Three Cases, February, April, and June, respectively.
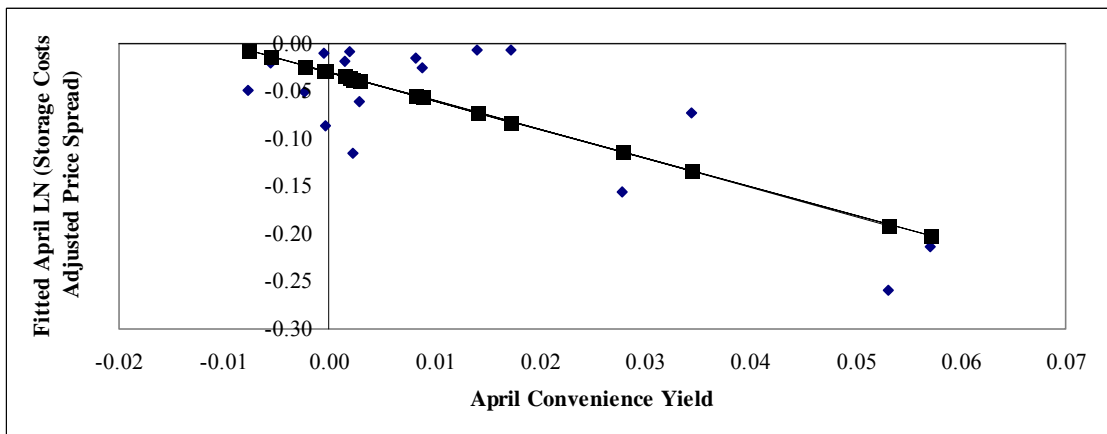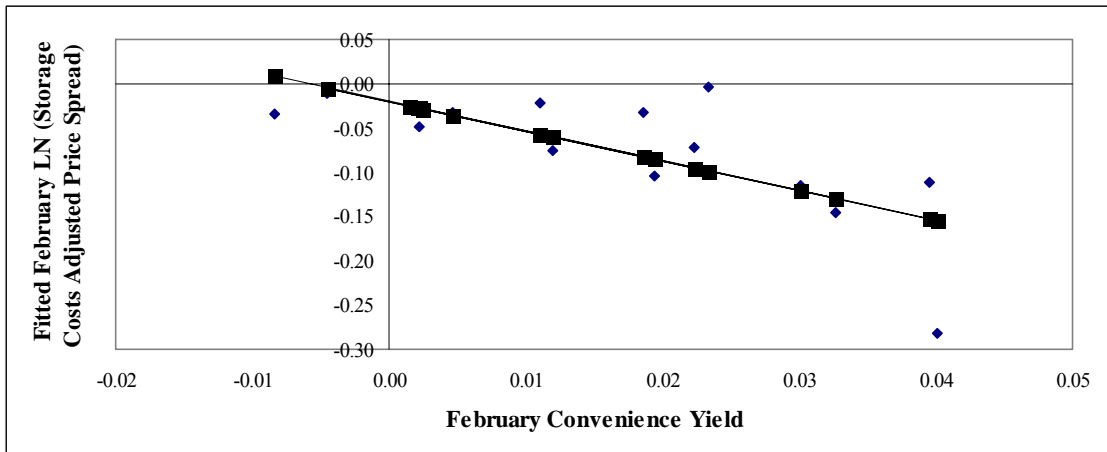
Figure 3.3: A scatter-graph and a fitted curved of Ln (Storage Costs Adjusted Price Spread) against Heaney's (2002) Proxy of Convenience Yield for February, April, and June, Respectively.

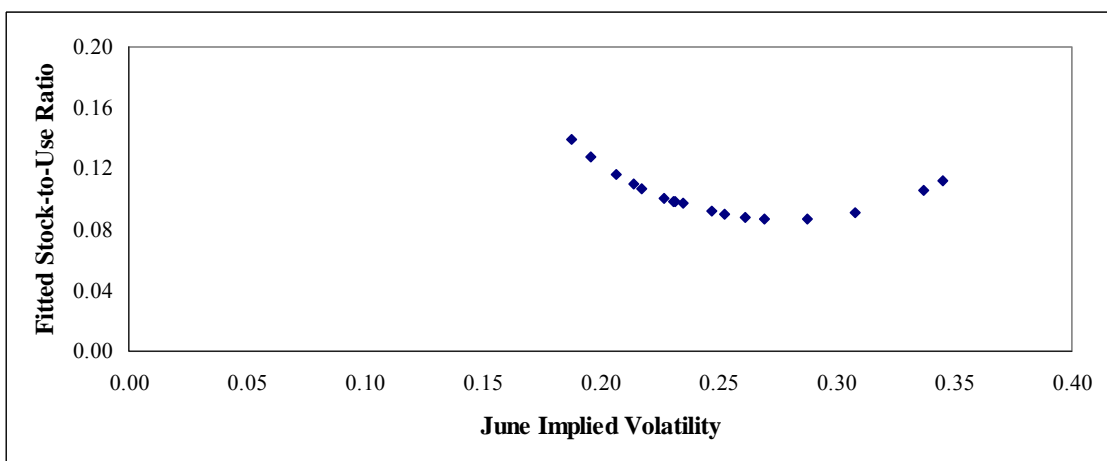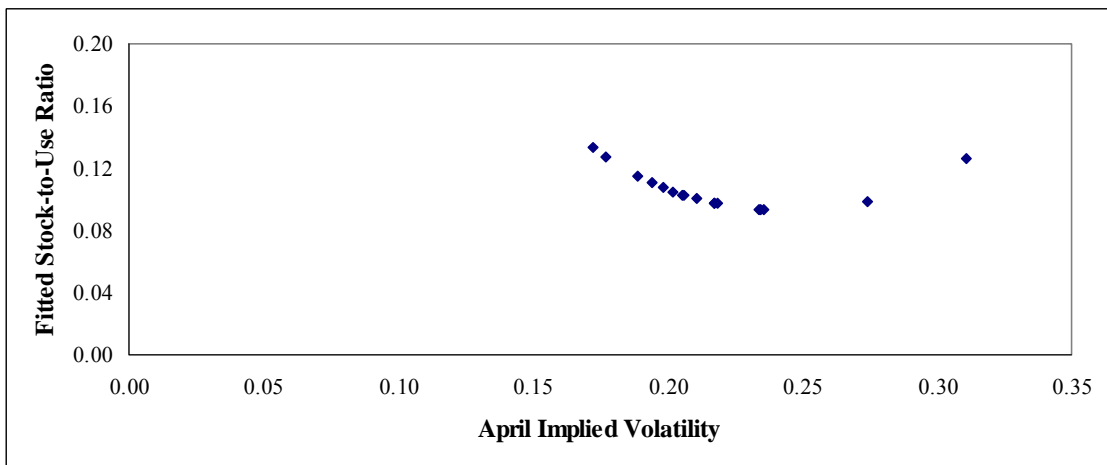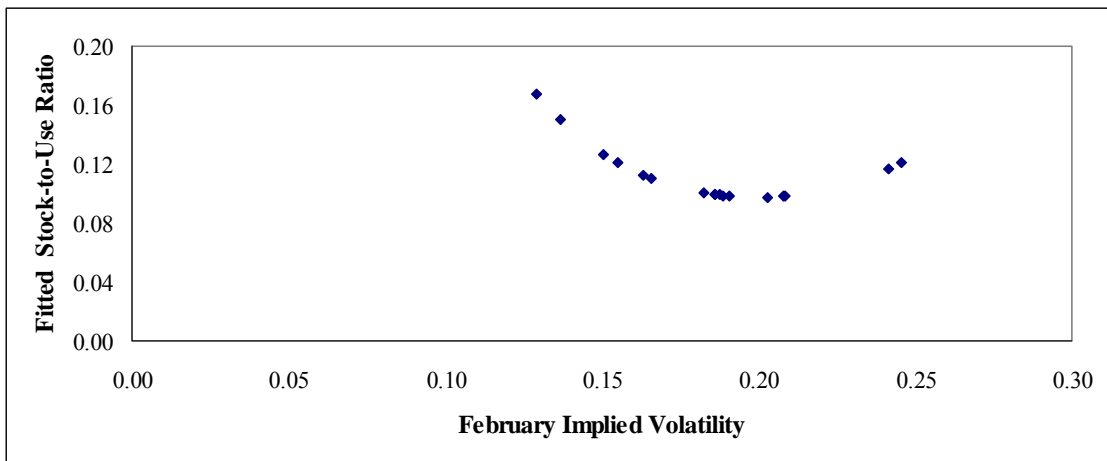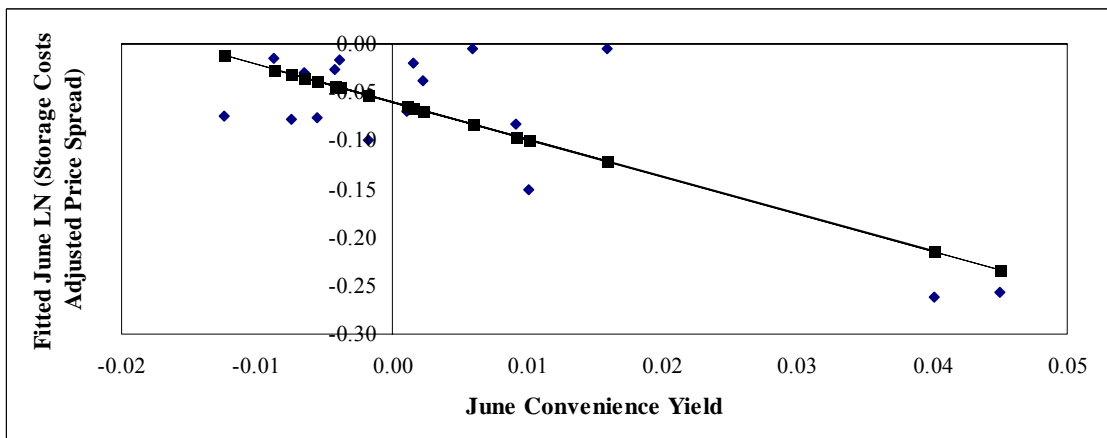| Model	World Agriculture Supply and Demand Estimates Release Date		
	February	April	June
Panel A: Traditional Supply of Storage Model (HAC) ---			
Intercept	-1.80**	-1.86**	-1.90**
	(0.09)	(0.066)	(0.08)
Ln (Spread)	5.60**	5.59**	4.98**
	(0.95)	(0.65)	(0.756)
R^2	0.70	0.83	0.74
Panel B: Updated Supply of Storage Equation (3SLS) ---			
Equation 1			
Intercept	2.44*	2.14	2.21*
	(1.05)	(1.48)	(1.04)
Ln (Spread)	6.87**	5.01**	3.48**
	(1.18)	(0.93)	(0.84)
Implied Volatility	-42.88**	-34.13**	-31.46**
	(11.34)	(12.64)	(7.96)
Volatility Squared	107.10**	69.81*	56.41**
	(30.43)	(27.09)	(14.81)
R^2	0.77	0.85	0.79
Equation 2			
Intercept	-0.02	-0.03*	-0.06**
	(0.015)	(0.0135)	(0.012)
Convenience Yield	-3.35**	-3.02**	-3.84**
	(0.74)	(0.59)	(0.75)
R^2	0.55	0.60	0.61

Notes: (a) Each variable is measured on the indicated month's release date of USDA's *World Agriculture Supply and Demand Estimates (WASDE).* (b) Estimated coefficients and corresponding standard errors are presented. (c) ** and * denote significance at 1% and 5% test levels, respectively. (d) A one-tailed test is used for all variables except the intercept. (e) Dependent variable in Panel A's equation and in equation 1 of Panel B is ln(stock-to-use ratio). The spread is measured as (ln{futures price spread adjusted for storage cost}). Convenience yield is measured using a procedure proposed by Heaney.

Table 3.1: Supply of Storage for Soybeans Estimated using HAC and 3SLS Regression, U.S., February, April, and June *WASDE* Release Dates, 1988-2004.

Model	World Agriculture Supply and Demand Estimates Release Date		
	February	April	June
Panel A: Traditional Supply of Storage Model --			
Intercept	-1.76**	-1.85**	-1.89**
	(0.08)	(0.06)	(0.074)
Ln (Spread)	6.48**	5.91**	5.42**
	(0.94)	(0.64)	(0.73)
R^2	0.74	0.84	0.76
Panel B: Updated Supply of Storage Equation (3SLS) --			
Equation 1			
Intercept	2.51*	2.29	2.30*
	(1.18)	(2.54)	(1.17)
Ln (Spread)	7.47**	5.88**	3.96**
	(1.03)	(0.98)	(0.84)
Implied Volatility	-44.11**	-35.40*	-32.08**
	(12.68)	(18.15)	(8.86)
Volatility Squared	112.52**	73.49*	57.78**
	(34.94)	(40.16)	(16.70)
R^2	0.82	0.85	0.79
Equation 2			
Intercept	-0.02	-0.03*	-0.06**
	(0.011)	(0.0115)	(0.0112)
Convenience Yield	-3.30**	-2.84**	-3.55**
	(0.65)	(0.544)	(0.65)
R^2	0.58	0.58	0.59

Notes: (a) Each variable is measured on the indicated month's release date of USDA's *World Agriculture Supply and Demand Estimates (WASDE)*. (b) Estimated coefficients and corresponding standard errors are presented. (c) ** and * denote significance at 1% and 5% test levels, respectively. (d) A one-tailed test is used for all variables except the intercept. (e) Dependent variable in Panel A's equation and in equation 1 of Panel B is ln(stock-to-use ratio). The spread is measured as (ln{futures price spread adjusted for storage cost}). Convenience yield is measured using a procedure proposed by Heaney.

Table 3.2: Supply of Storage for Soybeans Estimated using Bootstrap Method, U.S., February, April, and June *WASDE* Release Dates, 1988-2004.

Chapter 4

A New Measurement of Grain Stock-Price Relationship: An Empirical Study on Corn and Soybeans

4.1 Introduction

Working (1949)'s theory of the price of storage addresses the dominant role of stocks in determining inter-temporal price relationships, i.e., relationships between commodity prices for delivery at different points of time. It also implies a relationship between commodity stocks and commodity spot prices. As illustrated in Weymar (1966), the supply of storage theory not only can explain inter-temporal price relationships in terms of expected inventory behavior, but can also explain dynamic spot price behavior under certain further assumptions. In other words, there exists a relationship between stocks and prices. In practice, it is common for practitioners to utilize the historical relationship between stock-to-use ratio and marketing year average (MYA) price in price forecasts of the grain markets. The stock-to-use ratio, instead of the stock level, is used because everything else held constant, the level of stocks carried by storage agents is expected to increase as the size of the market. Size of the market has conventionally been

measured by quantity of consumption. The stock-to-use ratio is a better measurement and reflects the dynamics of both supply side and demand side to certain extent. It is commonly used as a proxy of stock level.

While some may view the relationship between the stock-to-use ratio and marketing year average price as analogous to a demand function, which is incorrect. The price and consumption (use) involved are determined simultaneously. The other important issue is the crop year cycle which is unique to agricultural commodities. Following harvest season, the supply for grains is highly price inelastic until the next harvest, which is a major distinguishing characteristic of agricultural products. This leads to a stronger correlation between ending stocks and price for grains than for non-agricultural markets. Thus, ending stock-to-use ratios are generally utilized to forecast grain prices.

The purpose of this study is to examine the historical relationship between ending stock-to-use ratio and marketing year average (MYA) price for U.S. corn and soybeans and to fit a forecast model for the two grains through a process of model selection and model comparison. The study also provides a new measurement for the grain stocks/price relationship. The marketing year average (MYA) price and ending stock-to-use ratio commonly used by grain market analysts are drawn from the *World Agricultural Supply and Demand Estimates* (*WASDE*) in its January report. One problem associated with this data is the mismatch between ending stock-to-use ratio and MYA. MYA is the average price received by farmers for their grain and is quoted for the whole marketing year while the stock-to-use ratio is quoted for a single point of time (the end of the crop year, i.e., August 31th of each year). We propose using an average ending stock-to-use ratio for the

marketing year consisting of equally-weighted ending stock-to-use ratios drawn from *WASDE* reports throughout the year. This new measurement enables us to reduce the potential estimation biases resulted from the data mismatch problem.

Data for MYA price, ending stocks, consumption, and exports for each crop year are collected from the *WASDE* reports for the sample period of 1981-2004 for both corn and soybeans. A model of grain prices as a function of the ending stock-to-use ratio is constructed. To control for the two farm policy reforms during the sample period, two dummies are used to measure the impacts of the 1985 and 1996 farm policy reforms. Standard ordinary least square (OLS) is applied to test the stock-price relationship using the fully specified model and several of its nested models. Leave one out cross validation (CV) is used to compare the forecasting ability of models for both traditional ending stock-to-use ratio and the proposed average ending stock-to-use ratio. The best model is chosen based on OLS R-squared, curve fitting, and the CV mean squared prediction error. Results show that there is little difference between the two cases for corn while for soybeans, the averaging significantly increases the in-sample and cross-validated results.

The rest of the paper is structured as follows: the following section discusses the policy regimes and empirical model, section III describes data and methodologies empirical results are reported and discussed in section IV and conclusions are presented in the last section.

4.2 Empirical Model

The economic significance of commodity storage has long been studied by researchers such as Working (1949), Brennan (1958), Weymar (1966), Wright and

Williams (1982 & 1984), and Williams and Wright (1991), among others. Working's (1949) theory of price of storage first recognized the critical role which stocks play in determining inter-temporal price relationships of commodities. Working (1949) also established the theory of the price of storage and the associated supply of storage curve. Weymar (1966) developed a model where inventory (or commodity stocks) helps explain the dynamic behavior of commodity spot prices. Weymar(1966) showed that under certain underlying assumptions, the spot price can be approximately expressed as a function of current inventory, the market's expectation of inventory for some finite horizon, and expected long-run equilibrium levels of price and inventory.

In practice, grain market analysts utilize the stock-price relationship in price forecasts. Specifically, historical data of *WASDE* MYA prices and ending stock-to-use ratios are used to estimate an empirical relationship between the two variables and thus to fit a forecast model. The most commonly used model can be expressed as follows,

$$\text{MYA} = a + b*\text{SU} \tag{4.1}$$

where MYA stands for commodity marketing year average price and SU for ending stock-to-use ratio, and *a* and *b* are coefficients. This can be viewed as a simplified version of the Weymar (1966)'s model.

Factors such as macroeconomic conditions and agricultural policy changes have potential influences on grain stock-price relationships and thus on the forecasting ability of the model. There were two major farm policy reforms during the sample period, namely, the 1985 Farm Policy Reform Act (S.1083) and the Federal Agricultural Improvement and Reform (FAIR) Act of 1996. These two reforms substantially altered the incentives faced by farmers, producers and other agricultural market participants and

on the historical relationship between grain stocks and prices as well. The 1985 and 1996 reforms divide the sample into three policy regimes. Due to the lag of impacts of the policy reforms, the division of the entire sample period may not exactly correspond to the years in which the reforms took place. The impacts of the two reforms on the stock-price relationship and the division of our sample period can be clearly seen in figure 4.1 for corn and figure 4.2 for soybeans. In figure 4.1, marketing year average (MYA) price of corn reported by WASDE is placed on the Y axis, and ending stock-to-use ratio on the X axis. The stock-price relationship of corn presents three different patterns over the three sub-periods. Likewise, there are three sub-periods for soybeans as shown in figure 4.2.

In order to capture potential impacts of the two policy reforms on the grain stock-price relationship, it is necessary to modify the simple stock-price equation (equation 4.1). Dummy variables are the best way to accommodate the impacts of structural change of the model. A fully specified polynomial function is proposed and expressed as follows,

$$MYA_t = \sum_{k=1}^{3}\sum_{i=-1}^{2}\beta_{k,i}D_k SU_t^i + \varepsilon \tag{4.2}$$

where MYA_t is marketing year average grain price for year t, SU_t is ending stocks-use ratio for year t, β is a coefficient, and ε is the error term. $k=1,2,3$ denotes three different policy regimes and D_k is an indicator or dummy of the k sub-period. In addition, i is the exponent of the independent variable SU. Conceptually this model can capture all possible impacts of the two farm policy reforms on grain stock-price relationship.

Since the fully specified model may not be the best model for testing the stock-price relationship or for predicting MYA price using ending stock-to-use ratio, several nested models of equation 4.2 will be selected and compared based on standard OLS,

curve fitting, and cross validation prediction errors. The specifications of these nested models are not presented here but will appear along with empirical results listed in table 4.1 and 4.2.

4.3 Data and Methodology

4.3.1 Data

Corn and soybean data, including marketing year average (MYA) prices, ending stocks, consumption, and exports are obtained from the *World Agricultural Supply and Demand Estimates* (*WASDE*) reports for the period of 1981-2004. The *WASDE* report is released by the USDA each month. It provides USDA forecasts of U.S. and world production and consumption major grains, oilseeds, cotton and U.S. production and consumption of sugar and livestock products. *WASDE* reports for corn and soybeans provide estimates of world and U.S. corn and soybean supply and demand factors, including beginning stocks, production, imports, domestic use and exports, ending stocks, and MYA price. Note that these estimates are all for certain specific crop year. Data are collected for sample period 1981 to 2004.

MYA price is average price received by farmers for their grain over the course of the marketing year and is drawn from the *WASDE* in the January report following the next year's harvest. For example, the January 2004 report would contain the final estimates of the MYA of the 2002/2003 crop year. Ending stocks are estimates of end-of-crop-year stocks (usually for the last day of August in each year) released by *WASDE* on its report release dates. Use (consumption) data includes estimates of domestic consumption and exports released by *WASDE* as of the *WASDE* release dates in each

month over the year. Figure 4.1 and figure 4.2 plot *WASDE* ending stock-to-use ratio against MYA price of three different sub-periods for corn and soybeans, respectively.

As discussed earlier, traditional analysis of stock-price relationship utilizes grain MYA prices and stocks and use data all in the January *WASDE* report for the previous crop year. However, ending stock-to-use ratio measures the ratio of the estimates of the stocks remaining as of August 31th of each crop year over the estimated total consumption plus exports throughout the year. The mismatch of quoted dates of different variables, i.e., MYA is for the whole marketing year while the ending stock-to-use ratio is quoted for a single point of time (the end of the crop year), may distort the stock-price relationship and thus the forecasting ability of the studied models. In order to avoid biases resulting from the mismatch problem, estimates of ending stocks and consumption and export are obtained for *WASDE* report of each month throughout the year and a series of 'average' ending stock-to-use ratio is created by equally weighting ending stock-to-use ratio of each month.

4.3.2 Methodology

The purpose of this study is to find the best stock-price model using the available *WASDE* data given the three policy regimes created by the1985 and 1996 farm policy reforms. Standard OLS estimation, curve fitting, and leave one out cross-validation (CV) are used to compare the different models. The comparison will be made among various models as specified in previous section.

Curve fitting is a standard tool for selecting models. The main objective of fitting data is to discriminate between different models and to test if the data is more consistent with one model relative to another. These models are related in certain ways and often

one is a simpler version of the other. Mathematically, the simpler model is called nested model. The procedure of curve fitting involves plotting observed values of studied series, regressing dependent variable against independent variable for selected model and obtaining estimated parameters, and fitting curves using estimated parameters and an appropriate range of values of the independent variable.

Cross-validation is a standard tool for measuring prediction error and is primarily used to assess forecasting performance of various models, especially in relatively small sample sizes. Thus it is often used to conduct model selection. Model selection criteria are all estimates of prediction error of the compared models, which measure how well the estimated model will perform on future (unknown) inputs. As with other forecast methodologies, the best model is the one whose estimated prediction error is smallest. If sufficient data is available then the data can be divided into two parts. One portion of the data (the 'in-sample' segment) is used to estimate the model while the remaining portion is used to assess forecast error. In this way several different models, all estimated for the in-sample, can be compared on out-of-sample for prediction (forecast) error.

The limited number of observations of the *WASDE* data makes the basic cross-validation inappropriate. According to Motulsky and Christopoulos (2003), there are possible biases introduced by the low number of observations in either in-sample estimation or out-of-sample testing and by relying on any one particular division of data into in-sample and out-of-sample. A better method, which is intended to reduce this bias, is to divide the original dataset into two samples (in- and out-of- samples) in several different ways and to compute an average prediction error over different divisions. The model with least average prediction error will be the best in terms of prediction ability.

An extreme variant of this approach is to split the dataset with *n* number of observations into a in-sample of size *n-1* and a out-of-sample (testing sample) of size 1 and average the prediction errors on the left-out sample over the *n* possible division ways. This is the so-called leave one out cross-validation (CV). The advantage of the leave one out CV is that all observations are used in the in-sample estimation and out-of-sample estimation *n* times.

4.4 Empirical Results

4.4.1 Corn

A fully specified model consistent with the polynomial function of equation 2 can be expressed as:

MYA=c+d1+d2+SU+SU1+SU2+SUInv+SUInv1+SUInv2+SUSqr+SUSqr1+SUSqr2 (3)

where MYA is the marketing average year price, *c* is the intercept, *d1* and *d2* are dummies, representing intercept of the first and the second policy regimes, respectively. *SU, SUInv*, and *SUSqr* are the ending stock-to-use ratio, its inverse, and its square. *SU1* and *SU2* represent ending stock-to-use ratio of the first and second policy regimes, and *SUInv1* and *SU Inv2*, and *SUSqr1* and *SUSqr2* are dummies for *SUInv* and *SUSqr*.

The above model and nine models nested within it are examined and compared by using OLS regression, curve fitting, and cross validation. The model best fitting the data is then selected based on its performance with the three methods discussed above. Standard OLS regression is first used to estimate the fully specified model (equation 4.2) and 9 models nested within it. Regression results are presented in table 4.1. The fully specified model (model 10), as shown in table 4.1, does not perform well since none of

the explanatory variables is statistically significant, which casts doubt on the reliability of this model. The first three models, which do not allow the different policy regimes to have individual effects, do not perform well because the R-squared of these three regressions is very low. This may also indicate it is necessary to incorporate the impacts of the two policy regimes into the model. The other six models contain dummies created for the two policy regimes and are various forms of the polynomial function specified in equation 4.2. These models provide much better regression results than the fully specified model and the models without dummies do. Thus, curve fitting and cross-validation method are applied to these six models.

In order to reduce the problem of limited number of observations (23 yearly observations for corn), leave one out cross-validation is performed for the other six models. Table 4.3 presents the mean-squared-prediction-error for these models. Model 8 has the smallest prediction error of 0.049 and the highest R-squared among six models, as shown in table 4.3. Curve fitting analysis is conducted and for each of the six models, curves for three different sub-periods are plotted using estimated parameters presented in figure 4.3. Comparison of the curves of the models shows that model 8 best fits the data. To save space, only the fitted curves of model 8 are presented in figure 4.3. Combined results of leave one out CV and curve fitting imply that model 8 is the best fitted model and the model with strongest power of prediction for the corn stock-price relationship.

As discussed earlier, one purpose of the paper is to explore whether the use of average ending stock-to-use ratios improves estimation and forecasting. Thus, the same OLS, curve fitting, and cross-validation (CV) analysis is conducted for corn MYA price and a series of average corn ending stock-to-use ratios which is obtained by equally

weighting the 12 predicted ending stock-to-use ratios during each crop year. Results of prediction error are also presented in table 4.3. The prediction error of model 8 is much higher than the one obtained for model 8 in the previous analysis and is higher than those of several models in the current set of analysis. However, model 5 has a much lower prediction error of 0.053. Model 5 also performs very well in curve fitting. The fitted curves of three different regimes for model 5 are presented in figure 4.4.

Table 4.3 presents results of leave one out CV prediction error and R-squared for 12 models, of which 6 models are for the analysis of traditional stock-price relationship and the others for the relationship between MYA price and the proposed average ending stock-to-use ratio. Models with average ending stock-to-use ratio fail to provide better estimation results or prediction results than their counterpart models with traditional ending stocks-use ratio. This indicates the use of average ending stock-to-use ratios fails to improve the model performance and predictability for corn.

4.4.2 Soybean

Following the same procedure of corn analysis as above, standard OLS regression, curve fitting, and leave one out CV are conducted for soybean stock-price relationship using both traditional ending stock-to-use ratios and the average ending stock-to-use ratios. OLS regression is run for the fully specified model as expressed in equation 4.2 and nine models nested within it, which all are some version of the polynomial function in equation 4.2. Regression results are presented in table 4.2. As in the case of corn, the first three models without dummies for the 1985 and 1996 policy reforms have very low R-squared, indicating it is necessary to incorporate the impacts of different policy

regimes into the model. Thus, curve fitting and leave one out CV are applied to the remaining models and results are compared in order to find the best forecasting model.

Table 4.4 presents the leave one out CV prediction errors and R-squared of these models for both cases, i.e., with traditional ending stock-to-use ratio and with average ending stock-to-use ratio. Model 5 with average ending stock-to-use ratio, which expresses MYA price as a function of inverse of average ending stock-to-use ratio, an intercept and two dummies of the incept, provides the smallest prediction error of 0.195 among all the examined models. Also, the fitted curve for this model, which is plotted in figure 4.5, fits the data best among all models. Thus model 5 with average ending stock-to-use ratio should be selected for examining the stock-price relationship and be used for forecasting. Specification of this model can be obtained from table 4.2 and figure 4.5. Results further show that the proposed new measurement of stock-price relationship, i.e., using average stock-to-use ratio against marketing year average (MYA) price improves both the performance of model estimation and the accuracy of prediction. This is implied by the fact that the OLS R-squared of models with average ending stock-to-use are higher than those of their counterpart models with traditional ending stock-to-use ratio and the prediction errors of models with average ending stock-to-use are much smaller than those of their counterpart models with traditional ending stock-to-use ratio.

4.5 Conclusion

The stock-price relationship has long been studied by researchers. In practice, grain market analysts frequently utilize historical relationship between stock-to-use ratios and marketing year average (MYA) prices in price forecasts of grain markets. Model

94

specification and model selection are thus critical to obtain reliable forecasts. Data of MYA prices, ending stocks, consumption, and export for grains are final estimates of a specific crop year and are usually drawn from a specific *WASDE* report for the previous crop year. However, one problem associated with such a usage of data is mismatch of MYA price and stock-to-use ratio, which may lead to estimation biases. MYA price is the average price farmers receive and is quoted for the whole marketing year while the stock-to-use ratio measures the ratio of estimate of the stocks remaining as of the end of crop year (usually August 31th for corn and soybeans) and the estimated total consumption plus exports over the marketing year. This study uses an average ending stocks-use ratio for the marketing year which can be obtained by equally weighting 12 ending stock-to-use ratios over the year.

A fully specified polynomial function is developed which takes account of various specifications of the independent variable, grain ending stock-to-use ratio, including an inverse term, a linear term, and a quadratic term. This polynomial function also incorporates dummies which represent three different policy regimes due to the U.S. 1985 and 1996 farm policy reforms. Analysis is conducted based on the fully specified model and several of its nested models for corn and soybeans.

Standard OLS regression, curve fitting, and leave-one-out cross-validation are used to choose the best model from both the fitting perspective and the forecast perspective. Empirical results show that for corn, using data with ending stock-to-use ratio, model 8, where MYA price is expressed as a function of the inverse of ending stock-to-use ratio with dummies incorporated into the model, has the smallest cross-validation prediction error and a very good fitted curve as well. When using data with the

proposed average ending stocks-use ratio, model 5, where MYA price is expressed as a function of the inverse of ending stock-to-use ratio with dummies for the intercept but without dummies for the explanatory variable, provides the smallest prediction error and performs curve fitting well. However, results imply that the use of average ending stock-to-use fails to improve the performance of selected models in either fitting historical data or predicting the future path of studied variables. For soybean, model 5 with average ending stock-to-use ratio is superior to other 13 models in terms of both CV prediction error and OLS R-squared. Furthermore, in contrast to the case of corn, for soybeans, each model with average ending stock-to-use ratio performs significantly better than their counterpart models with traditional ending stock-to-use ratio in terms of both CV prediction error and OLS R-squared. This implies that using the proposed new measurement of stock-price relationship, i.e., using average ending stock-to-use instead of traditional ending stock-to-use ratio significantly improves model performance in both estimation and prediction.

Hence, grain market analysts may benefit from using the proposed average ending stocks-use ratio for analyzing soybean stock-price relationship and forecasting soybean prices. The model should also take into account the potential impacts of the U.S. 1985 and 1996 farm policy reforms.

Corn: Ending S/U versus MYA Price
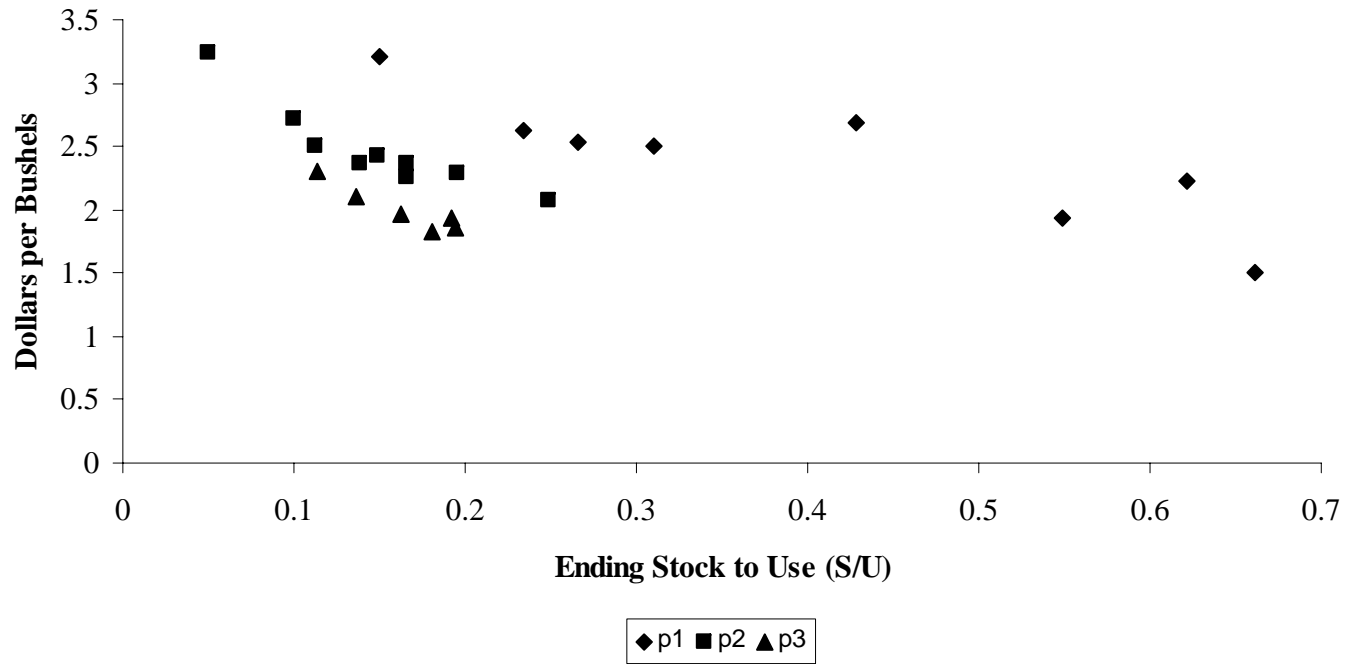
Note: p1, p2, and p3 represent period of 1981-1988, period of 1989-1997, and period of 1998-2003, respectively.

Figure 4.1: Plot of corn ending stock-to-use ratio and corn marketing year average (MYA) price for the sample period of 1981-2003.

Soybeans: Ending S/U versus MYA Price
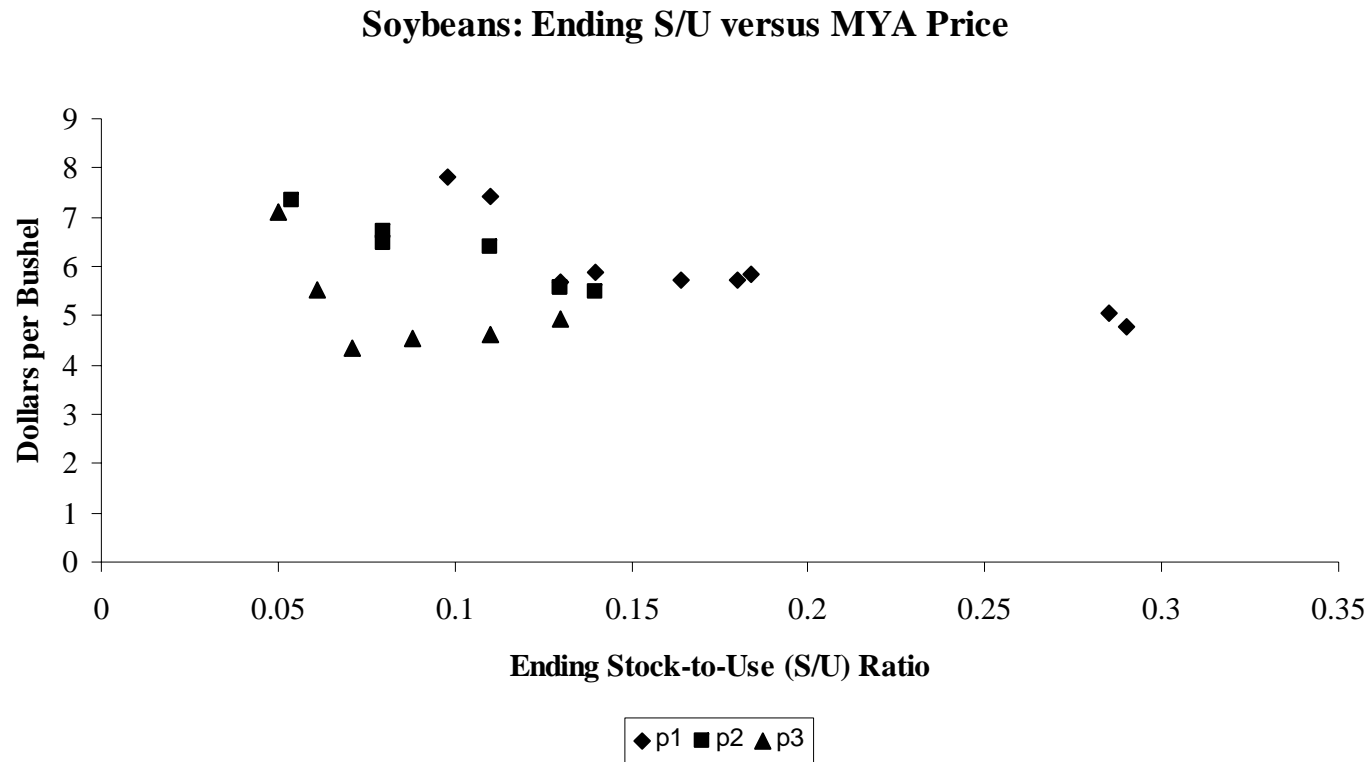
Note: p1, p2, and p3 represent period of 1981-1988, period of 1989-1997, and period of 1998-2003, respectively.

Figure 4.2: Plot of soybean ending stock-to-use ratio and soybean marketing year average (MYA) price for the sample period of 1981-2003.

Corn: S/U versus MYA Price

Note: (1) p1, p2, and p3 represent period of 1981-1988, period of 1989-1997, and period of 1998-2003, respectively. Pred1, pred2, and pred3 denote predicted values for the three sub-periods, respectively. (2) Specification of Model 8: MYA = c + d1 +d2 + SUInverse + SUInverse1 + SUInverse2, where d1, and d2 are dummies for the intercept term, SUInverse is inverse of S/U, and SUInverse1 and SUInverse2 are dummies for SUInverse, respectively.

Figure 4.3: Plot of corn predicted ending stock-to-use ratio (Model 8) and corn marketing year average (MYA) price for the three sub-periods.
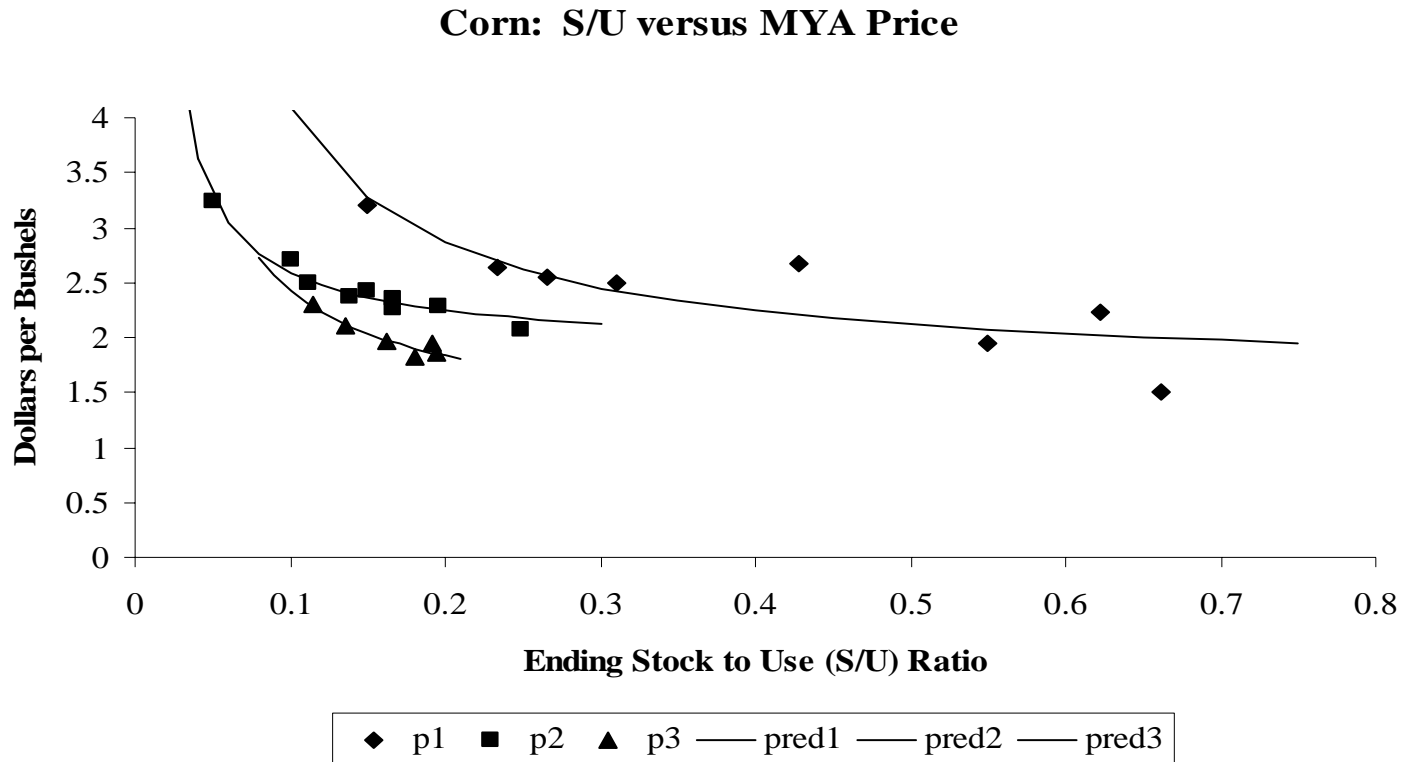
Corn: Average S/U versus MYA-Price

Note: (1) p1, p2, and p3 represent period of 1981-1988, period of 1989-1997, and period of 1998-2003, respectively. Pred1, pred2, and pred3 denote predicted values for the three sub-periods, respectively. (2) Specification of Model 5: MYA = c + d1 +d2 + AveSU, where d1, and d2 are dummies for the intercept term, and AveSU is average ending stocks-use ratio.

Figure 4.4: Plot (Model) of corn predicted average ending stock-to-use ratio and corn marketing year average (MYA) price for the three sub-periods.

Soybean: Average S/U versus MYA Price
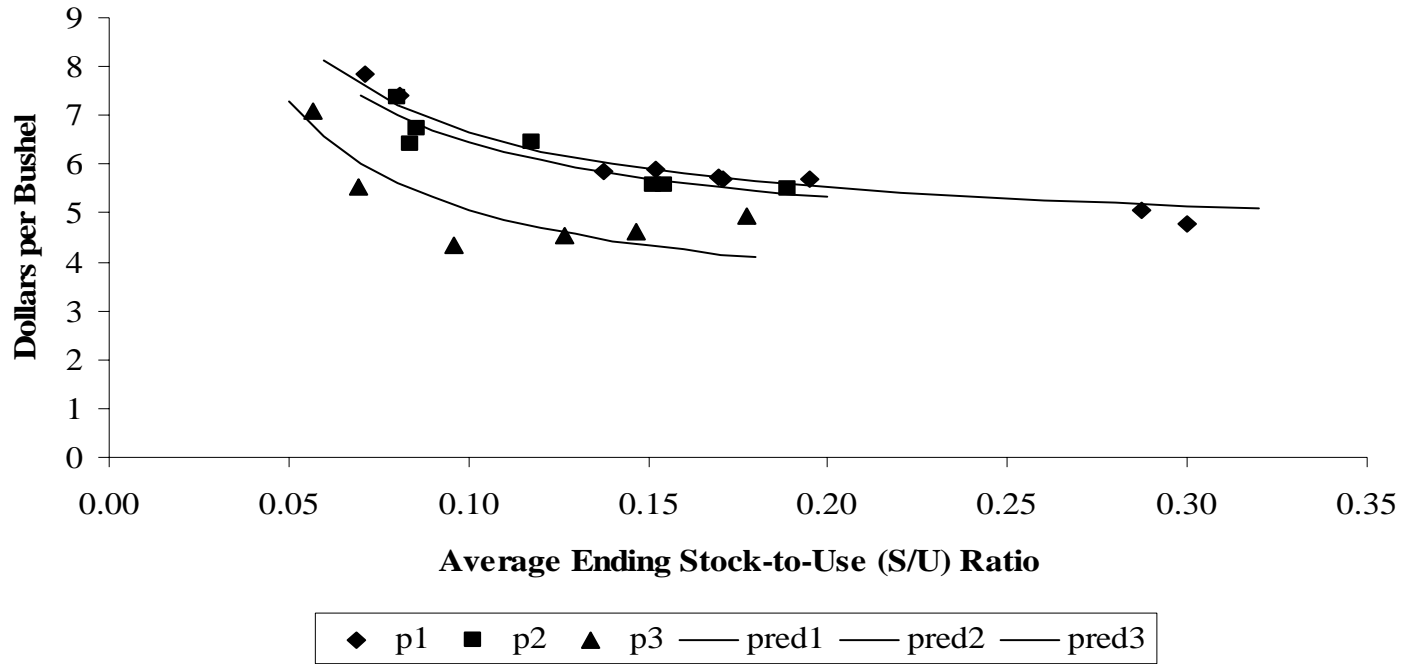
Note: (1) p1, p2, and p3 represent period of 1981-1988, period of 1989-1997, and period of 1998-2003, respectively. Pred1, pred2. and pred3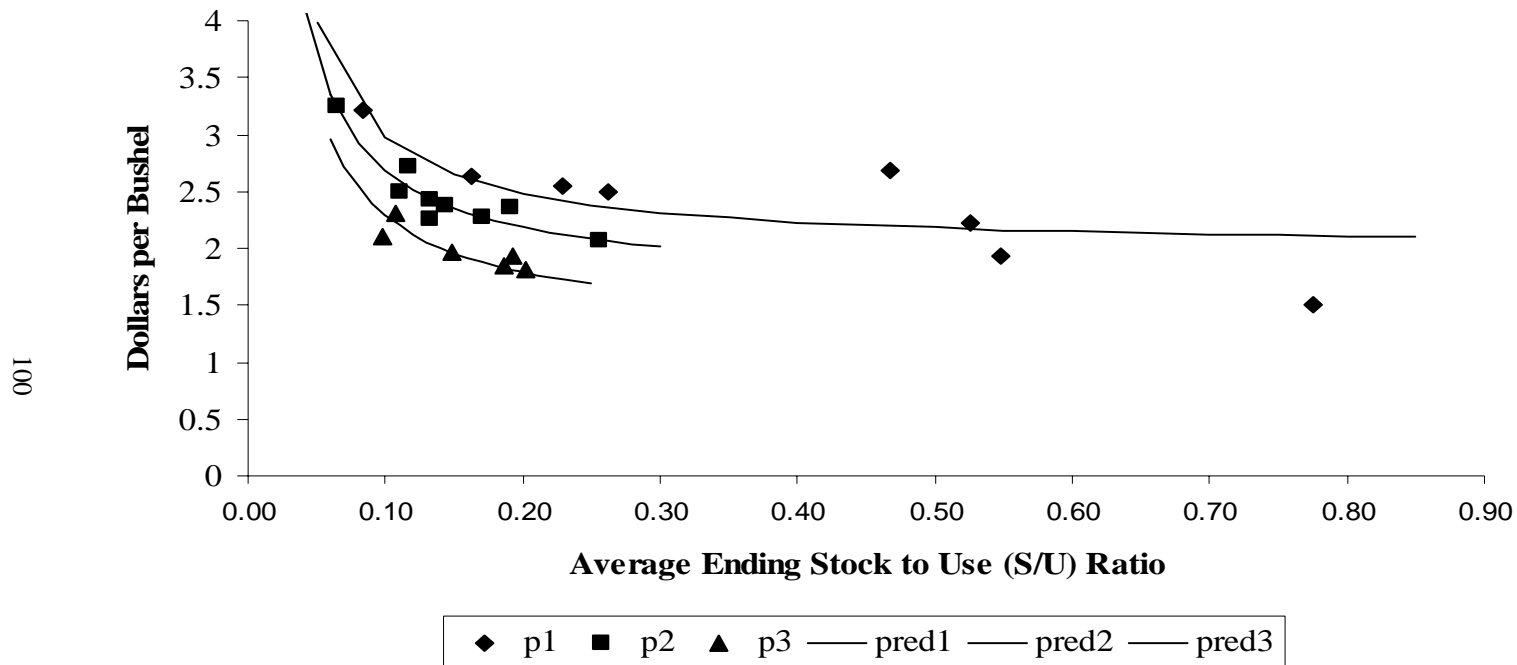 denote predicted values for the three sub-periods, respectively. (2) Specification of Model 5: MYA = c + d1 +d2 + AveSU, where d1, and d2 are dummies for the intercept term, and AveSU is average ending stocks-use ratio.

Figure 4.5: Plot (Model 5) of soybean predicted average ending stock-to-use ratio and soybean marketing year average (MYA) price for the three sub-periods.

Model[a]	Intercept	D1	D2	SU	SU1	SU2	SUIn	SUInv1	SUInv2	SUSqr	SUSqr1	SUSqr2	R-squred	Wald Test[c]
1	17.9**[b]			-2.10*									0.17	
2	14.12**						3.06**						0.31	
3	7.99**			-0.73						0.31			0.18	
4	22.24**	6.96**	3.79**	-6.68**									0.77	49.43**
5	9.43**	4.70**	2.23*				5.16**						0.68	22.08**
6	13.96**	7.92**	3.76**	-4.00**						2.19*			0.82	64.04**
7	6.47**	1.10	0.91	-1.94	1.04	-0.09							0.84	69.2**
8	3.25**	0.94	1.64				2.03*	1.80	-0.83				0.84	57.52**
9	1.51	-0.39	-0.22	-0.69	0.63	0.33				0.56	-0.56	-0.33	0.87	73.35**
10	0.23	-0.29	-0.19	-0.23	0.27	0.22	-0.15	0.38	0.17	0.24	-0.26	-0.24	0.92	109.5**

Note: (a) Model 10, the fully specified model consistent with equation 2 is expressed as follows:
MYA = c+d1+d2+SU+SU1+SU2+SUInv+SUInv1+SUInv2+SUSqr+SUSqr1+SUSqr2.
(b) Entries in the table are t-statistics of regression coefficients for two-tailed test.
(c) Wald-test is used to test joint significance of coefficients. Chi-square statistic is reported for Wald-test.

Table 4.1: Corn standard OLS regression (MYA price and ending stock-to-use ratio) results of selected models.

Model[a]	Intercept	D1	D2	SU	SU1	SU2	SUInv	SUIn1	SUIn2	SUSqr	SUSqr1	SUSqr2	R-squred	Wald Test[c]
1	14.7**			-1.90									0.15	
2	10.55**						1.93						0.16	
3	6.71**			-0.68						0.26			0.16	
4	18.29**	5.00**	3.70**	-4.92**									0.66	26.23**
5	4.80**	5.78**	4.74**				5.83**						0.72	36.06**
6	12.21**	6.07**	4.61**	-3.83**						2.42*			0.74	38.91**
7	8.31**	1.21	1.33	-2.12*	0.79	-0.13							0.69	27.36**
8	4.63**	0.55	1.80				3.65**	2.65**	-0.10				0.82	54.34**
9	6.81**	-1.16	-2.05*	-4.14**	2.96**	2.72**				3.79**	-3.39**	-2.88**	0.88	78.77**
10	-2.17*	-0.22	1.16	1.93	-0.62	-0.75	3.23**	1.75	-1.38	-1.49	0.91	0.42	0.97	272.87**

Note: (a) Model 10, the fully specified model consistent with equation 2 is expressed as follows:
MYA = c+d1+d2+SU+SU1+SU2+SUInv+SUInv1+SUInv2+SUSqr+SUSqr1+SUSqr2.
(b) Entries in the table are t-statistics of regression coefficients for two-tailed test.
(c) Wald-test is used to test joint significance of coefficients. Chi-square statistic is reported for Wald-test.

Table 4.2: Soybean standard OLS regression (MYA price and ending stock-to-use ratio) results of selected models.

	Models[a]	CV Mean-Squared-Prediction-Error	OLS R-Squared
	Model 4	0.06	0.77
	Model 5	0.13	0.68
Corn MYA & Stock-to-use Ratio	Model 6	0.07	0.82
	Model 7	0.058	0.84
	Model 8	0.049	0.84
	Model 9	0.096	0.87
	Model 4	0.058	0.75
	Model 5	0.053	0.78
Corn MYA & Average Stock-to-use Ratio	Model 6	0.143	0.76
	Model 7	0.0635	0.81
	Model 8	0.069	0.79
	Model 9	0.13	0.86

Note: (a) Model specifications can be obtained in table 1. Only models 4, 5, 6, 7, 8 are compared for leave one out mean-squared-prediction error and OLS R-squared.

Table 4.3: Model selection for corn stock-price relationship using both traditional ending stock-to-use ratio and average ending stock-to-use ratio.

	Models	CV Mean-Squared-Prediction-Error	R-Squared
	Model 4	0.488	0.66
	Model 5	0.405	0.72
	Model 6	0.396	0.74
Soybeans MYA & Stock-to-use	Model 7	0652	0.69
	Model 8	0.435	0.82
	Model 9	0.482	0.88
	Model 10	4.88	0.97
	Model 4	0.37	0.74
	Model 5	0.195	0.87
	Model 6	0.295	0.82
Soybeans MYA & Average Stock-to-use Ratio	Model 7	0.57	0.75
	Model 8	0.29	0.89
	Model 9	0.44	0.92
	Model 10	2.43	0.97

Note: (a) Model specifications can be obtained in table 1. Only models 4, 5, 6, 7, 8 are compared for leave one out mean-squared-prediction error and OLS R-squared.

Table 4.4: Model selection for soybean stock-price relationship using both traditional ending stock-to-use ratio and average ending stock-to-use ratio.

CHAPTER 5

SUMMARY AND CONCLUSION

Each of the previous three chapters explores one specific type of theoretical model of futures markets, proposes updating measurements of these models, and applies them to certain commodity futures markets, including metals markets at the London Metal Exchange (LME) and U.S. soybean and corn markets at the Chicago Board of Trade (CBOT). Empirical results provide richer understanding to the underlying theory and empirical estimation measurements of the variables in the models.

Chapter two tests the cost of carry theory using a unique dataset from the LME which has a unique price quotation system. Long run relationship between cash and 3-month futures prices for five metals at the LME is found to exist. Furthermore, a quad-variate cointegration model is constructed and empirical results show that co-integration exists for metals cash and 3-month futures prices, 3-month interest rates and physical storage costs. These findings reconcile previously inconsistent findings regarding the cointegration of temporal prices in the presence of non-stationary interest rates and are consistent with Working's cost of carry theory.

Chapter three updates the estimation of the supply of storage model to reflect recent developments in the theoretical and empirical literature. A simultaneous two

-equation system model is constructed. One equation is the supply of storage equation and the other the equation for convenience yield and storage costs adjusted inter-temporal price spread. Empirical analysis is conducted through three-stage least squares (3SLS) estimation method and bootstrapping 3SLS. Results reveal that convenience yield and variability of new crop futures might play important role in determining carryover stocks of soybean. Future studies can explore alternative measurements of convenience yield and the non-linear relationship between stock-to-use ratio and price variability.

Chapter four proposes a new measurement of the stock (inventory)-price relationship for grains markets by constructing an equally weighted ending stock-to-use ratio. A fully specified polynomial function is developed with consideration of three policy regimes due to the 1985 and 1996 US farm policy reforms. From both the fitting perspective and the forecast perspective, model selection is conducted for various models by comparing standard OLS results, curve fitting, and forecasting error computed using cross validation method. Empirical results indicate that grain market analysts may benefit from using the proposed measurement to forecast soybean prices. Unfortunately, there is no significant difference between the two measurements of stock-price relationship in terms of forecast performance or curve fitting.

In summary, this study contributes to the understanding of the cost of carry theory, the supply of storage theory and the convenience yield theory. Empirical analysis updates the measurement of variables within these models and provides results supporting the proposed theoretical arguments.

BIBLIOGRAPHY

Alexander, C. "Optimal Hedging Using Cointegration." *Philosophical Transactions of the Royal Society of London Series A-Mathematical Physical and Engineer Sciences*, 357: 2039-2058, Aug. 1999.

Baillie, R. T. and R. J. Myers. "Bivariate Garch Estimation of the Optimal Commodity Futures Hedge." *Journal of Applied Econometrics*, 6 (1991), 109-124.

Beck, S. E. "Cointegration and Market Efficiency in Commodities Futures Markets." *Applied Economics*, 1994, 26, 249-257.

Benirschka, M., and J. K.Bindley. "Optimal Storage and Marketing over Space and Time." *American Journal of Agricultural Economics*, 77 (August 1995), 512-524.

Bessler, D.A., and T. Covey. "Cointegration: Some Results on U.S. Cattle Prices." *The Journal of Futures Markets*, Vol. 11, No.4, 461-474 (1991).

Brenner, R. J. and K. F. Kroner. "Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets." *Journal of Financial and Quantitative Analysis*, 30 (1995), 23-42.

Brennan, D., J. Williams, and B. D. Wright. "Convenience Yield without the Convenience: A Spatial-Temporal Interpretation of Storage under Backwardation." *The Economic Journal*, Vol. 107, Issue 443 (Jul., 1997), 1009-1022.

Brennan, M. J. "The Supply of Storage." *American Economic Review*, Vol. 48 (1958), 50-72.

Brennan, M. J. "The Price of Convenience and the Valuation of Commodity Contingent Claims." *Stochastic Model and Option Values*, Elsevier Science Publishers B.V. (North-Holland), 1991.

Brennan, M. J. and E. Schwartz. "Evaluating natural resource investments." *Journal of Business*, Vol. 58 (1985), 135-157.

Bresnahan, T. F. and P. T. Spiller. "Futures Market Backwardation Under Risk Neutrality." *Economic Equiry*. 24(July 1986): 429-441.

Chavas, J. P. "On Competitive Speculation under Uncertainty: An Alternative View of the Inverse-Carrying Charge." *Journal of Economics and Business*, Vol. 40 (1988), 117-128.

Chavas, J. P., P. M. Desspins, and T. R. Fortenbery. "Inventory Dynamics under Transaction Costs." *American Journal of Agricultural Economics*, Vol. 82 (May 2000): 260-273.

Chowdhury, A. R. "Futures Market Efficiency: Evidence from Cointegration Tests." *Journal of Futures markets*, 11 (1991), 577-589.

Covey, T. and D. A. Bessler. "Asset Storability and the Informational Content of Inter-temporal Prices." *Journal of Empirical Finance*, 2 (1995), 103-115.

Deaton, A., and G. Laroque. "Competitive Storage and Commodity Price Dynamics." *The Journal of Political Economy*, Vol. 104, No.5 (Oct., 1996), 896-923.

Dickey, D. A. and W. A. Fuller. "Distribution of Estimates for Autoregressive Time Series with a Unit Root." *Journal of American Statistical Association*, 24 (1979): 427-431.

Engle, R. F. and C.W. J. Granger. "Co-Integration and Error Correction: Representation, Estimation, and Testing." *Econometrica*, 55 (1987), 251-276.

Fackler, P. L and M. J. Livingston. "Optimal Storage by Crop Producers." *American Journal of Agricultural Economics*. 84(August 2002): 645-659.

Fama, E.F., and K.R. French. "Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage." *Journal of Business*, 1987, Vol. 60, No.1, 55-73.

___ and ___. "Business Cycles and the Behavior of Metal Prices." *Journal of Finance*, Vol. 43, Issue 5 (1988), 1075-1093.

Fortenbery, T.R. and H.O. Zapata. "An Examination of Cointegration Relations between Futures and Local Grain Markets." *Journal of Futures Markets*, 13 (1993), 921-932.

___ and ___. "An Evaluation of Price Linkages Between Futures and Cash Markets for Cheddar Cheese", *Journal of Futures Markets*, 17(1997), 279-301.

Frechette, D. L. "The Dynamics of Convenience and the Brazilian Soybean Boom." *American Journal of Agricultural Economics,* Vol. 79 (Nov. 1997): 1108-118.

_____. "The Supply of Storage under Heterogeneous Expectations." *Journal of Agricultural and Applied Economics*, Vol. 31 (3) (December 1999): 461-474.

_____. "Aggregation and The Nature of Price Expectations." *American Journal of Agricultural Economics*, Vol. 83 (1) (February 2001): 52-63.

_____ and P. L. Fackler. "What Causes Commodity Price Backwardation?" *American Journal of Agricultural Economics*, Vol. 81 (November 1999): 761-771.

Garbade, K. D., and W. L. Silber. "Price Movements and Price Discovery in Futures and Cash Market." *The Review of Economics and Statistics*, Vol. 65, Issue 2 (May, 1983), 289-297.

Gray R. W., and A. E. Peck. "The Chicago Wheat Futures Market: Recent Problems in Historical Perspective." *Food Research Institute Studies*, Vol. XVIII, No.I, 1981.

Greene, W. H. *Econometrics Analysis*. Prentice Hall Publisher, 2002.

Hamilton, J.D. *Time Series Analysis*. Princeton University Press, 1994.

Heaney, R. "A Test of The Cost-of-Carry Relationship Using The London Metal Exchange Lead Contract." *Journal of Futures Markets*, Vol.18, No.2 (1998), 177-200.

_____. "Does Knowledge of The Cost of Carry Model Improve Commodity Futures Price Forecasting Ability? A Case Study Using The London Metal Exchange Lead Contract." *International Journal of Forecasting*, 18(2002), 45-65.

_____. "Approximation for Convenience Yield in Commodity Futures Pricing." *Journal of Futures Markets*, Vol. 22, No.10 (2002), 1005-1017.

Hein, S. F., and C. K. Ma, and S. S. MacDonald. "Testing Unbiasedness in Futures Markets: A Clarification." *The Journal of Futures Markets*, Vol. 10, No.5, 555-562 (1990).

Heinkel, R., M. E. Howe, and J. S. Hughes. "Commodity Convenience Yields as an Option Profit." *The Journal of Futures Markets*, Vol.10, No. 5, 519-533 (1990).

Hilliard, J. E., and J. Reis. "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*, Vol. 33, No. 1 (Mar., 1998), 61-86.

Hull, J. C. *Options, Futures, and Other Derivatives*, 4th edition. Prentice Hall, NJ, 2000.

Irwin, S. H., and C. R. Zulauf, and T. E. Jackson. "Monte Carlo Analysis of Mean Reversion in Commodity Futures Prices." *American Journal of Agricultural Economics*, 78 (May 1996): 387-399.

Johansen, S. "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control*, 12 (1988), 231-254.

_____. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica*, Vol. 59, Issue (Nov., 1991), 1551-1580.

___ and K. Juselius, "Maximum Likelihood Estimation and Inference on Cointegration — with Applications to the Demand for Money." *Oxford Bulletin of Economics and Statistics*, 1990, 169-210.

Kaldor, N. Speculation and Economic Stability. *The Review of Economic Studies*, Vol. 7, No.1 (Oct., 1939), 1-27.

Kellard, N. "Evaluating Commodity Market Efficiency: Are Cointegration Tests Appropriate?" *Journal of Agricultural Economics*, Vol. 53 (2002): 513-529.

_____, P. Newbold, T. Rayner, and C. Ennew. "The Relative Efficiency of Commodity Futures Markets." *The Journal of Futures Markets*, Vol. 19 (4): 413-432, June 1999.

Keynes, J. M. *A Treatise on Money*, Vol. 2: *The Applied Theory of Money*. London: Macmillan & Co., 1930.

Khoury, N. T. and J .M. Martel. "A Supply of Storage Theory with Asymmetric Information." *The Journal of Futures Markets*, Vol. 9, No. 6, 573-581 (1989).

Kofi, T.A. "A Framework for Comparing the Efficiency of Futures Markets." *American Journal of Agricultural Economics,* November 1973, 584-594.

Krehbiel, T. and L.C. Adkins. "Cointegration Tests of the Unbiasedness Expectations Hypothesis in Metals Markets." *Journal of Futures Markets*, Vol. 13 (7): 753-763, October, 1993.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin. "Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root." *Journal of Econometrics*, 54 (1992), 159-178.

Lien, D. D. "The Effect of the Cointegration Relationship on Futures Hedging: A Note." *The Journal of Futures markets*, Vol. 16, No. 7, 773-780 (1996).

Litzenberger, R. H. and N. Rabinowitz. "Backwardation in Oil Futures Markets: Theory and Empirical Evidence." *The Journal of Finance*, Vol. 50 (Dec., 1995), 1517-1545.

Longstaff, F. A. "How Much Can Marketability Affect Security." The Journal of Finance, Vol. 50, No.5 (December 1995), 1767-1774.

Low, A. H. W., and J. Muthuswamy, and R. I. Webb. "Arbitrage, Cointegration, and the Joint Dunamics of Prices Across Discrete Commodity Futures Auctions." *The Journal of Futures Markets*, Vol. 19, No.7, 799-815 (1999).

Mark, N.C. "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability." *The American Economic Review*, Vol. 85, Issue 1 (Mar., 1995), 201-218.

McKenzie, A. M., B. Jiang, H. Djunaidi, L. A. Hoffman, and E. J. Wailes. "Unbiasedness and Market Efficiency Tests of the U.S. Rice Futures Market." *Review of Agricultural Economics*, Vol. 24 (2002), 474-493.

___ and M. Holt. "Market Efficiency in Agricultural Futures Markets." *Applied Economics*, 2002, 34, 1519-1532.

Milonas, N. T. and S. B. Thomadakis. "Convenience Yield and the Option to Liquidate for Commodities with a Crop Cycle." *European Review of Agricultural Economics*, 24 (1997), 267-283.

___ and ___. "Convenience Yields As Call Options: An Empirical Analysis." *The Journal of Futures Markets*, Vol. 17, No.1, 1-15 (1997b).

Miltersen, K. R. and E. S. Schwartz. "Pricing of Options on Commodity Futures with Stochastic Term Structures of Convenience Yields and Interest Rates." *The Journal of Financial and Quantitative Analysis*, Vol. 33, No. 1 (March 1988), 33-59.

Motulsky, H. and A. Christopoulos. *Fitting Models to Biological Data Using Linear and Nonlinear Regression*. GraphPad Software, Inc., 2003.

Newey, W. and K. West. "Hypothesis Testing with Efficient Method of Moments Estimation." *International Economic Review*, 28, 777-787.

____ and ____. "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55, 703-708.

Norwood, B., M.C. Roberts and J.L. Lusk. "Ranking Crop Yield Models Using Out-of-Sample Likelihood Functions." *American Journal of Agricultural Economics*, Vol. 86, No. 4 (Nov. 2004), 1032-1043

Quan, J. "Two-Step Testing Procedure for Price Discovery Role of Futures Prices." *Journal of Futures Markets*, 12 (1992), 139-149.

Richter, M. C. and C. Sorensen. "Stochastic Volatility and Seasonlity in Commodity Futures and Options: The Case of Soybeans." Working paper, Department of Finance, Copenhagen Business School, 2002.

Routledge, B. R., D. J. Seppi, and C. S. Spatt. "Equilibrium Forward Curve for Commodities." *The Journal of Finance*, Vol. 55, No.3 (June 2000), 1297-1338.

Rudolf Wolff &Co, Limited. *Wolff's Guide to the London Metals Exchange*, 4th edition, Metal

Bulletin Books, Ltd. 1991.

Sabuhoro, J. B. and B. Larue. "The Market Efficiency Hypothesis: the Case of Coffee and Cocoa Futures." *Agricultural Economics*, 16 (1997), 171-184.

Saha, A. and J. Stroud. "A Household Model of On-Farm Storage Under Price Risk." *American Journal of Agricultural Economics*, Vol. 76 (August 1994), 522-534.

Schroeder, T. C. and B. K. Goodwin. "Price Discovery and Cointegration for Live Hogs." *The Journal of Futures Markets*, Vol. 11, No. 6, 685-696 (1991).

Schwarz, T. V. and A. C. Szakmary. "Price Discovery in Petroleum Markets: Arbitrage, Cointegration, and the Time Interval of Analysis." *Journal of Futures Markets*, 14 (1994), 147-167.

Sharples, J. A. and F. D. Holland. "Impact of the Farmer-Owned Reserve on Privately Owned Wheat Stocks." *American Journal of Agricultural Economics*, 63 (3): 538-543 1981.

Sorensen, C. "Modeling Seasonality in Agricultural Commodity Futures." *The Journal of Futures Markets*, Vol. 22, No. 5, 393-426 (2002).

Telser, L.G. "Futures Trading and the Storage of Cotton and Wheat." *The Journal of Political Economy*, Vol. 66, No. 3 (June, 1958), 233-255.

Thompson, S. "Returns to Storage in Coffee and Cocoa Futures Markets." *Journal of Futures Markets*, Vol. 6 (1986), 541-564.

Thraen, C. "A Note: The CSCE Cheddar Cheese Cash and Futures Price Long-Term Equilibrium Relationship Revisited." *The Journal of Futures Markets*, Vol. 19, No.1, 233-244 (1999).

Tomek, W. G. and R. W. Gray. "Temporal Relationships among Prices on Commodity Futures Markets: Their Allocative and Stabilizing Roles." *American Journal of Agricultural Economics,* Vol. 52, No.3, August 1970, 137-148.

Waltkins, C. and M. McAleer. "Cointegration Analysis of Metals Futures." *Mathematics and Computers Simulation*, Vol. 59 (2002): 207-221.

Weymar, H. "The supply of Storage Revisited." *American Economic Review*, Vol. 56 (1966), 1226-1234.

____. *The Dynamics of the World Cocoa Market*. MIT Press, Cambridge, Massachusetts, 1968.

White, H. "A Heteroskedasticity – Consistent Covariance Matrix and a Direct Test for

Heteroskedasticity." *Econometrica*, 48, 817-838.

Wilcox, R.R. *Applying Contemporary Statistical Techniques*. Academic Press, 2003.

Williams, J. B. "Speculation and the Carryover." *The Quarterly Journal of Economics*, Vol. 50, No.3 (May, 1936), 436-455.

Williams, J. C. "Futures Markets: A Consequence of Risk Aversion or Transactions Costs?" *Journal of Political Economy*, 1987, vol. 95, No. 51, 1000-1023.

___ and B. D. Wright. *Storage and Commodity Markets*. New York: Cambridge University Press, 1991.

Wooldridge, J. M. *Introductory Econometrics: A Modern Approach*. South-Western College Publishing, 1999.

Working, H. "Theory of the Inverse Carrying Charge in Futures Markets." *Journal of Farm Economics,* 30 (1948): 1-28.

Working, H. "The Theory of Price of Storage." *American Economic Review*, 39 (1949), 1254-1262.

Wright, B.D. and Williams, J.C. "The economic role of commodity storage", *Economic Journal*, 92 (1982), 596-614.

_____ and ____. "The welfare effects on the introduction of storage", *Quarterly Journal of Economics*, 98 (1984), 169-192.

Wright, B. D. and J. C. Williams. "A Theory of Negative Price for Storage." *The Journal of Futures Markets*, Vol. 9, No.1, 1-13 (1989).

Yang, J. and D. A. Bessler, and D. J. Leatham. "Asset Storability and Price Discovery in Commodity Futures Markets: A New Look." *Journal of Futures Markets*, 21 (2001), 279-300.

Zapata, H. O. and T. R. Fortenbery. "Stochastic Interest rates and Price Discovery in Selected Commodity Markets." *Review of Agricultural Economics*, 18 (1996), 643-654.

Zulauf, C. R., S. H. Irwin, J. E. Ropp, and A. J. Sberna. "A Reappraisal of the Forecasting Performance of Corn and Soybean New Crop Futures." *The Journal of Futures Markets*, Vol. 19, No. 5, 603-618 (1999).