Iowa State University
Digital Repository

Retrospective Theses and Dissertations | Iowa State University Capstones, Theses and Dissertations

1997

Corn and soybean basis behavior and forecasting: fundamental and alternative approaches

Bingrong Jiang
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Agricultural and Resource Economics Commons, Agricultural Economics Commons, Finance Commons, and the Finance and Financial Management Commons

Recommended Citation
Jiang, Bingrong, "Corn and soybean basis behavior and forecasting: fundamental and alternative approaches " (1997). *Retrospective Theses and Dissertations*. 12214.
https://lib.dr.iastate.edu/rtd/12214

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

المنارة للاستشارات

www.manaraa.com

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Corn and soybean basis behavior and forecasting: fundamental and alternative approaches

by

Bingrong Jiang

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Major Professor: Marvin Hayenga

Iowa State University

Ames, Iowa

1997

UMI Number: 9737727

UMI Microform 9737727
Copyright 1997, by UMI Company. All rights reserved.

This microform edition is protected against unauthorized
copying under Title 17, United States Code.

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of

Bingrong Jiang

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

The Importance of Basis in Hedging

Basis is defined as the difference between cash and futures prices. In grain merchandising, basis is usually defined as the difference between the local cash price and the nearby futures price, i.e. the current price of the nearest futures delivery contract. It has been argued by many researchers that the key to successful hedging is understanding the basis (e.g., Garcia and Good 1983, Hieronymous 1978, Leuthold et al. 1989, Marshall 1989, Karlson et al. 1993, Tomek 1996). This is because most hedging involves two opposite positions: one in the cash market and another in the futures market. It is the difference between cash and futures prices, together with the futures price, that determines the return from hedging and hedging's effectiveness in reducing risk.

Grain merchandisers and processors routinely need to accurately forecast basis to offer forward purchase or sales contracts. Corn and soybean producers need to know the basis to evaluate contracts offered to them, or in making hedging decisions. The Chicago Board of Trade (CBOT) asserts "Without a knowledge of the usual basis and basis patterns for your particular commodity, it is impossible to make fully informed decisions about, for example, whether to accept or reject a given price; whether and when to store your crop; whether, when, and in what delivery month to hedge; when to close (or 'lift') a hedge; or when and how to turn an unusual basis situation into a possible profit opportunity" (p. 23).

Further, Karlson et al. (1993) stated that changes in the basis are more predictable than changes in the cash price because cash and futures prices converge and become equal at the delivery point. The more predictable the basis is, the more valuable hedging is to the market participants. They show that storage hedging can be a profitable corn marketing strategy which may be overlooked by producers. But in order to practice this strategy effectively, producers or elevator managers must acquire and understand the basis history.

The Objectives of This Study

Though basis is extremely important, there have been only few basis behavior studies published, and even fewer basis forecasting studies (with the exception of forecasts using simple moving averages of historical basis). The objective of this study to investigate corn and soybean basis behavior and to improve the accuracy of basis forecasts.

This paper first reviews the theory of storage and basis studies on grain. The theory of storage is related closely to the temporal and spatial price relationships in the grain market. The basis of storable commodities is actually a temporal price relationship, the difference between current spot price and futures price. Local basis has another component of spatial price relationship, which is the difference between a local cash price and the cash price at Chicago Board of Trade delivery points. All the available basis studies on grain are reviewed to assess the current state of knowledge, and identify where further contributions could be useful.

Since basis patterns differ from location to location, what is important to hedgers is the basis at their location. To provide some diversity in locations and to better represent the U.S. corn and soybean market, several local markets are studied. These markets include Chicago, St. Louis, Toledo, Gulf Coast, NE Iowa, Central Illinois, Richmond, and the Pacific

Northwest market (corn only).[1]

Several approaches are utilized in this study to explain and subsequently forecast the local grain basis of these markets. A fundamental structural model incorporating storage cost, transportation costs, and regional supply and demand variables is developed to explain basis behavior. Several forecasting techniques are used in forecasting corn and soybean basis. These include including traditional methods such as a simple three-year-average forecasts, structural econometric model, time series methods such as seasonal ARIMA and State Space models, and other approaches such as a modified three-year average model, artificial neural networks, and composite forecasts. The ability of the structural model to explain past basis behavior is examined, and out-of-sample forecast performance of these alternative basis forecasting approaches is evaluated.

[1] Iowa and Illinois are major grain producing states; the Gulf, a major grain export port; Pacific Northwest, a fast growing export port; Richmond, an East coast market; and Chicago, Toledo and St. Louis, delivery locations for futures contracts (St. Louis is not a delivery location for soybean futures contract). In 1998, the CBOT is proposing dropping Chicago due to closure of major elevators there, and replacing it with Northern Illinois River elevators (Burns Harbor to Pekin). Prices are highly correlated with Chicago, but basis may drop slightly.

CHAPTER 2. LITERATURE REVIEW

Theory of Storage

The efficient market hypothesis states that prices reflect information to the point where the marginal benefits of acting on information do not exceed the marginal costs. Correspondingly, the theory of storage suggests that basis--the difference between contemporaneous spot and futures prices--should equal the cost of storage. Otherwise, there will be opportunities for profitable arbitrage between the spot and futures markets. This suggests that basis has a predictable temporal pattern. As the cost of storage decreases as one gets closer to delivery (maturity), the cash price will gain in relation to futures price. On the first day of delivery at the par delivery point, cash and futures prices should be equal theoretically, and the basis should be zero (but the basis at non-par delivery points is usually not zero because of transportation cost differences). Similarly, the price spread of two futures contracts within a given crop year, at least when it is positive, will not exceed cost of storage between the two delivery months. These are the clear aspects of the theory of price of storage as stated by Working (1949).

Empirical studies have found that it is not uncommon to have the basis to be less than the full cost of storage, and even negative returns for storage are possible. Several studies have been offered to explain this discrepancy between theory and empirical evidence. The popular explanations for the failure of the theory are convenience yield and risk premium. The most recent studies show that the theory should not have been rejected, because the analyst's misconception or mis-measurement of the concept led to its rejection (Wright and Williams, Brennan et al., and Benirschka and Binkley), which will be discussed in detail later.

Convenience yield, first introduced by Kaldor (1939, 1940), exists because it takes time to acquire commodities. According to Kaldor, stock has a yield because stock holders can "lay hands on them the moment they are wanted and thus saving the cost and trouble of ordering frequent deliveries, or of waiting for deliveries" (p. 4). Then the net carrying cost depends on the yield of stocks, in the sense that convenience yield is deducted in the calculating of net carrying costs. That is the reason that the net carrying cost could be either positive or negative. He argued that the convenience yield is negatively related to the stock level. It rises or falls with the decrease or increase in stock, and the yield is zero when redundant stock exists. He concluded that the futures price can not exceed the current spot price by more than the sum of interest (short term rate of interest) plus carrying costs due to arbitrage, "while there is no limit, apart from expectations, to the extent to which the futures price may fall short of the current price" (Kaldor 1940, p. 198).

The Working (1948, 1949) theories of carrying and inverse carrying charges state that the basis for the same commodity reflects a return to storage or price of storage which is determined by the demand for and supply of storage. Negative return to storage is possible, reflecting a relative scarce currently supply of the commodity. He used the famous supply curve of storage to illustrate the situation. This storage supply curve showed that the amount of wheat storage supplied is an increasing function of price of wheat storage, even when there were negative returns for wheat storage. Supply shortages put pressure on market participants to sell now and avoid holding stocks. It suggests that they have to pay for storing a commodity if they choose to store. He concluded that "they (hedgers) are willing to risk loss on a fraction of the stocks for the sake of assurance against having their merchandising or manufacturing activities handicapped by shortage of supplies." (p. 24). He agreed with Kaldor that convenience yield rationalizes what appears as a loss from storage. Following Kaldor's idea of "speculative stock" and "normal requirement" stocks, Working argued that stocks may be divided into two categories: surplus stocks and necessary working stocks. Surplus stocks are defined as the stocks a holder will not carry without assurance or expectation of a positive return, while necessary working stocks will be carried for reasons of convenience yield. This distinction may help in the understanding of "negative carrying charge" or "less than full carrying charge." Besides the convenience yield, Working also offered other possible explanations for large amount of storage with zero return from storage. One is that most costs of storage are fixed costs. Another is that costs of storage are joint, that is the owners may be in the business of merchandising and processing, so that storage may incur losses while other sectors' profits may offset those losses.

Telser (1958), in his study of cotton and wheat storage, specified the net marginal cost of holding a given amount of stocks as the remainder of the marginal cost of storage minus the marginal convenience yield. He further explored the factors that affect marginal convenience yield. More current consumption reduces the stock level, so it is positively related to convenience yield because lower stock level is associated with higher convenience yield. Another factor that has a positive effect on convenience yield is the percentage of total stocks under government price-support program. The real cost of communication and transport, supply of the commodity and average distance of stocks from consuming centers are negatively related to the convenience yield. There are also other supporters of the convenience yield concept such as Brennan (1958).

The risk premium explanation of the failure of price spreads in futures markets to cover commodity holding costs can be traced back to Keynes (1930). He believed that futures prices are downward biased estimates of the forthcoming spot prices. Hedgers use futures market to avoid risk. They have to pay a risk premium to speculators to get rid of their risk. The premium is realized by speculators refusing to purchase the hedger's contract except at a price lower than the futures price is expected to approach. Therefore, futures prices customarily rise as the contracts approach maturity. According to Benirschka and Binkley (1995), there are some studies that supported this argument, while more studies found little or no evidence of risk premiums.

Fama and French (1987) examined and empirically tested the theory of storage [convenience yield approach] and an alternative model that views basis as the sum of an expected risk premium (the bias of the futures price as a forecast of future spot price) and an expected change in spot price [risk premium approach]. In explaining negative basis, especially before a harvest when the futures price is for delivery after harvest, the theory of storage states that a negative basis exists because inventories are low and the convenience yield is larger than interest and storage costs; while the alternative model will argue that it is because the spot price is expected to fall when a harvest will substantially increase inventories. The theory of storage predicts the basis equal to interest foregone in storing a commodity, warehousing costs, and a convenience yield on inventory. After dividing the basis by the spot price, a well-known implication of this model is shown: basis (adjusted by spot price) for any commodity should vary one-for-one with interest rate after controlling for variation in marginal storage cost and the marginal convenience yield.[2] Farm and French predict high basis variability for seasonal, high-storage-cost commodities. They argued that one source of basis variation is seasonal supply and demand, and seasonal variation in the basis should be an increasing function of storage costs. Since storage costs deter inventory holding, the effects of demand and supply shocks on the variability of the basis should be an increasing function of storage costs. They estimated these two models, convenience yield, and risk premium approaches, for 21 commodities (10 agricultural products, 2 wood products, 5 animal products, and 4 metals[3]). They used futures prices on maturing contracts as a measure of spot price in estimating the theory of storage model. Seasonal dummies were used to crudely capture variation in the marginal convenience yield, which is due to seasonal variations in production or demand. They found that the results could not provide convincing evidence of one-for-one variation in the basis in response to nominal interest rate for agricultural products. Basis variation must be explained primarily in terms of economic conditions that generate variation in marginal storage cost and marginal convenience yields (p. 61). Reliable seasonal patterns in the basis are found, as expected, for many of seasonally produced agricultural commodities (based on the F statistics on the seasonal dummies). They claimed that the theory of storage view of futures price is not controversial, but the alternative model is, which

---

[2] According to the theory of storage, their basis equation is

$$F(t, T) - S(t) = S(t)R(t, T) + W(t, T) - C(t, T)$$

where $S(t)R(t, T)$ is interest foregone, $W(t, T)$ is marginal storage cost, and $C(t, T)$ the marginal convenience yield. Dividing both sides by the current spot price, then the equation becomes:

$$(F(t, T) - S(t))/S(t) = R(t, T) + (W(t, T) - C(t, T))/S(t)$$

where the coefficient of $R(t, T)$ is one.

[3] Including: cocoa, coffee, corn, cotton, oats, orange juice, soybeans, soy meal, soy oil and wheat, lumber and plywood, broilers, cattle, eggs, hogs, pork bellies, copper, gold, platinum and silver.

expresses the difference between the futures price and current spot price as the sum of an expected premium and an expected change in the spot price. Consistent with their claim, they found in their empirical analysis that "the tracks of the storage-cost variables in the basis are identified more easily than variation in the basis due to expected premiums and forecasts of future spot prices" (p. 72).

Though Fama and French (1987) stated that convenience yield version of the theory of storage is not controversial, recent studies by Wright and Williams (1989); Brennan, Williams and Wright (1997); and Benirschka and Binkley (1995) have found that convenience yield has no appeal in explaining storage behavior. Wright and Williams (1989) showed that the mystery of the supply-of-storage curve can be explained as a phenomenon of aggregation. Based on a partial equilibrium theory of investment, they set up a two-commodity model to analyze the problem. The two commodities are closely related substitutes linked by some necessary transformations such as transportation, cleaning, processing and so on. Three major findings of their study are: (1) "total transformation cost minimization" is a more descriptive rationale than "convenience yield." During the time when spreads are below full carrying charges, stocks are held, not because of convenience of having them in that location or in that form, but because of the inconvenience of transforming them into the commodity for which there is a premium for immediate delivery. (2) Sub-aggregate measures of prices and stocks differed from aggregate measures. Evidence from the coffee market indicates that with greater precision in the definition of relevant prices and stocks, the significant storage in times of backwardation (when cash price is high than futures price) was much less. (3) Distortions in commodity markets change the relationship between the amount stored and the cash-futures price backwardation (futures price less than cash price). For example, public stockpiles distort not only the total quantity of private storage but also the location and grades of private stocks. Other possible distortions can be export subsidies or noneconomic inventories such as the strategic petroleum reserve.

Brennan, Williams and Wright (1997) studied storage behavior and returns to storage in a wheat marketing system in Western Australia. A mathematical programming model of shipments and storage was constructed using detailed, engineering-based costs of transport networks and storage facilities. They argued that "Working's supply-of-storage curve relating some aggregate measure of stocks to the spot/futures spread at a pricing center has been found for every storable commodity investigated. Working's explanation [convenience yield] for the supply-of-storage curves has been widely accepted as well. Despite the many empirical studies of aggregate storage behavior, none has traced that behavior to an indirect yield earned by individual firms from the stocks that they own" (p. 2). From their spatial-temporal model, they found that no convenience yield is received by storers at any location, and convenience yield is not required to explain Working's supply-of-storage curve. The capacity constraints on linear technologies were sufficient to induce a typical aggregate supply-of-storage curve. "The confusion comes from supposing the price relationships at a pricing center apply to the commodity at the location or in the form in which marginal storage is taking place" (p. 19). Moreover, the conditions often thought necessary for convenience yield by previous studies, namely risk and nonlinear cost functions for the merchandising and processing by individual firms, were absent from the model. They concluded that any apparent loss was an illusion from spatial aggregation of stocks and improper attribution of incentives at one locality to storage decisions at all locations.

Benirschka and Binkley (1995) developed a mechanism guiding the storage and marketing over space and time as an analogy of the theory of resource extraction. Their model is based on the theory of efficient commodity markets in which the rate of return from holding commodity stocks must equal the rate of return from holding financial assets. This implies that the price difference between contemporaneous spot and futures markets should equal the cost of storage (which includes warehousing cost, insurance cost and interest foregone). Given few simplifying assumptions, the only variable cost for storing grain is the opportunity cost (interest income foregone from financial assets). Due to transportation costs, grain prices in surplus areas decline as distance to market increases, and interest foregone would also be less where grain prices are lowest. This suggests that storage costs are minimized when storage occurs at the point of production. Because of the difference in transportation cost, storage firms close to the market can offer the grain to the market at lower prices than more remote firms.[4] As a result, the authors concluded that firms supply the market sequentially. The firm farther away from the market supplies only after the firm closer to the market does. Because of this sequential marketing pattern, the transport cost for the grain received at the market increases over time. This results in two important implications: market price grows at a rate smaller than the rate of interest, and its rate of growth decreases over time.[5] They concluded that "this 'storage at a loss' illusion exists because prices at the

[4] This assertion implicitly assumes the costs of gathering grain are the same for storage firms with different distances from the market. If closer firm has higher cost of obtaining grain, it may not be able to offer the lower price.

[5] I believe the authors made mistakes in deriving the formulas supporting these arguments. They argued that the sequential marketing pattern, then the changing transportation cost of the grain

market are used to compute the opportunity cost of storage. Correct computation requires use of grain prices where storage firms are located" (p. 515). Because opportunity costs based on market prices will be biased upward due to the transport costs, the bias will be significant when large portion of storage occurs far from the market. To check their argument that firms supply the market sequentially, they tested the hypothesis indirectly by examining whether the grain storage capacity increases with distance to the market. Two models of grain storage capacity were set up, one for total storage capacity, another for on-farm and off-farm (commercial) grain storage capacity separately. The corn loan rate was chosen as the proxy variable for the distance to the market. Results showed that the loan rate coefficients were significant for on-farm equations but not for off-farm equations. For the total grain storage capacity model, the loan rate coefficient was significant for the OLS estimate, but not for the ML estimate which does not correct for the problem of spatial autocorrelation. They concluded "grain storage capacity increases with distance to market, especially capacity located on farms" (p. 522). As further evidence, descriptive analysis was used to show that the length of storage varies directly with distance to market and also that locations with higher prices at harvest tended to have lower rates of price growth.

---

supplying to the market over time results in these two implications. These two implications were obtained after differentiating the market price equation with respect to time and transportation cost. However, in differentiating, they treated transportation cost as fixed. If this changing transportation cost is incorporated into their derivation, these two implications may not hold (cannot sign the two derived equations from which they draw the implications).

Grain Basis Studies

Besides various studies dealing with the theory of storage, there have been few empirical studies related to the basis behavior and basis forecasting of grain. Heifner (1966) set up a prediction equation for both cash price change and basis change, which stated that the change variables over a particular interval are a function of the basis at the beginning of the interval. He showed that much better predictions were made for basis changes than for cash price changes (in terms of reductions in the residual variation). He then incorporated the basis change prediction equations into a set of conditional storage rules in order to estimate the potential gains from this forecast. A Monte Carlo procedure was employed in the storage decision simulation. The average revenue from a conditional storage rule utilizing the basis change forecasts was compared to revenue from automatic storage every year. The results showed that the conditional storage rule which makes use of the forecasts was most useful for storage decisions in January and March, but not in November.

Kenyon and Kingsley (1973) predicted the harvest time basis at planting time and compared the performance of this harvest time forecasted basis with other historical average basis estimates in the hedging effectiveness. They specified a basis change model as a function of initial basis, Chicago cash price, and the residual of open interest in the futures market with respect to a linear trend. Both cash price and futures price were used separately in place of initial basis for a better model fit. The model was estimated for three time periods, with three alternative planting times and a single harvest time, and two crops (corn and soybeans) in Richmond, Virginia area. Most of the variables were significant and all the equations had high R-Squares. They concluded that initial basis at planting time can be used to predict harvest time basis with a considerable degree of accuracy. This predicted harvest basis produced estimates of expected net price (net hedging returns) superior to those based on mean basis for some historical time periods.

Martin, Groenewegen and Pidgeon (1980) modeled the factors affecting corn basis in southwestern Ontario over crop years from 1962 to 1976. They were concerned with basis when corn is marketed (termed "ending basis"). Based on previous studies, they summarized the theoretical factors expected to affect basis. These were marginal net storage cost, market liquidity, and market characteristics unique to a specific commodity. For the nondelivery point basis, local market conditions were also important factors. The basis had two components, net spatial costs and a basis residual, at a given point of time. The factors in net spatial costs were usually loading, tariff, transfer and handling charges. Among the net spatial costs, rail costs, the tariff, and average annual vessel rates were known. The unknown component of the basis was termed the basis residual. The independent variables in the empirical model of corn basis residual in southwestern Ontario were the ratio of eastern Canadian production to consumption, Canadian production, U.S. production, availability of western feed grains to eastern Canada, and dummies representing short-run pricing aberrations. When the ratio of eastern Canadian production to consumption exceeded unity, its corn was being exported to the U.S. Another measure of competition was the availability of western feed grains shipped to eastern Canada. Canadian corn production determined the demand for storage while U.S. production affected futures price level and demand for transportation services in Great Lakes-St. Lawrence system. The seasonality of basis was captured when the model was estimated for each month. The results suggested that a substantial amount of variation in the southwestern Ontario basis was explained by variables reflecting local market conditions. For the fall months (Nov., Dec.), the size of crops in both U.S. and Canada was the most important factor; for winter (Jan. through Apr.), only the Canada crop relative to demand mattered; during spring season (May to July), local production relative to demand and western feed grain production were significant; for months of August and September, no significant factors were found.

Kahl (1982) focused on the change in corn basis patterns from the sixties to seventies in Chicago. Though the theoretical seasonal pattern prevailed, both basis plots and t-tests indicated differences in the basis between the '60s and '70s for harvest and summer months. A basis regression model included three variables: one-month lagged basis, the percentage of current bin space capacity available for corn, and private corn inventory minus corn consumption multiplied by 1 minus the price elasticity of demand. The empirical results confirmed the hypothesis that the increase in the harvest basis in the seventies was partially caused by increased demand for available storage space. No clear evidence supported or refuted the other hypotheses. A change in price expectations from one decade to the next may have caused the basis to change, and the decreased basis in the summer months of the seventies could be attributed to increased convenience yield. Though the lagged basis term, included to capture price expectations, was significant for all the months, it was difficult to interpret its meaning and implications.

Garcia and Good (1983), based on the theory of carrying charge, analyzed the factors influencing the Illinois corn basis for the period from 1971 to 1981. The observation of historical prices showed that cash prices did not always increase at a rate consistent with that theory. This, as they argued, suggested that the magnitude of the basis may be influenced by factors other than storage and transportation costs (for non-delivery point basis). It had to be modified by supply and demand of storage. They included three sets of factors (cost, stock and flow factors) in the basis model, where stock and flow factors determined the supply and demand for storage. Both production and ending stocks of corn and soybeans relative to commercial storage capacity were used to represent stocks. Cost variables were represented by barge rates, interest rates, regional and monthly dummies. Regional dummies accounted for regional transportation cost differences within the state, and monthly dummies accounted for monthly transportation marketing costs not explicitly accounted for by the barge rates. The monthly average cash price was included to reflect grain flow volumes, as a high price induces heavy marketing. They estimated their model in three seasonal time periods: harvest (Oct.-Dec.), post-harvest (Jan.-Apr.), and distant-harvest (May-Jul.). The results showed that the variation of Illinois local basis was explained by the model quite well. The importance of variables in the model varied among these three periods. Stock variables had strong effects during harvest period. The importance of cost variables varied across time periods. Barge rate only mattered during distant-harvest period. Monthly dummies captured some transportation change effects on basis. The interest rate was significant for all three periods and its coefficient was not different from one. The flow variable, cash price, was significant in explaining the local basis. A dummy variable included to examine the effect of the Russian Grain Embargo in 1979 was significant for post-harvest and distant harvest periods.

Powers and Johnson (1983) studied Wisconsin corn basis during the storage season for three years from 1978-1980. They specified a monthly basis model to explain the effect of spatial, temporal, and local market conditions on the Wisconsin basis. In their analysis, the local market was assumed as a small market so that it had no effect on futures price. The basis was defined as the difference between local cash price and July corn futures price. The prices were calculated as average of second and third mid-week prices or of second and fourth mid-week prices if the month had five mid-week prices. They argued that three factors would influence Wisconsin corn basis: spatial price relationships (transportation cost), local market conditions, and temporal factors (storage cost). A monthly trend variable was used as a proxy for transportation cost from Wisconsin to Chicago. Interest rates combining the length of storage were used to account for storage cost. Several variables (expected production, corn harvested and corn stocks) expressed as percent of off-farm storage capacity were used to represent local market conditions. Results showed that the monthly trend dummies, interest rate and stock variables were significant with the right sign. The expected production was insignificant. The equation was re-estimated for a storage season defined as the October-July period. The results were very similar. No significant change was found for three storage seasons. In general, this model explained a high level of the variation of Wisconsin corn basis, with the R square higher than 0.8.

Taylor and Tomek (1984) developed a simple model to forecast the November corn basis in Batavia, New York. Their model specification was based on two general concepts: (1) regional cash price difference between New York and Chicago and (2) the difference between Chicago cash and futures prices. It was argued that the regional differences in price depend on local production adjusted for local consumption (production minus estimated consumption by dairy cows) and on national production. The relationship between cash and futures prices in Chicago was related to the costs associated with making and taking delivery. The open interest relative to the deliverable supply in Chicago would determine the delivery cost, since the "squeeze" is more likely when open interest is larger relative to the deliverable supply. The variables included in the empirical model were U.S. corn production, feed deficit in New York State, Chicago corn stock at first Friday of November and the open interest in December corn futures on first trading day of November. The model was fitted to data from 1972 to 1982. All the variables except Chicago corn stock were significant with expected signs. In-sample forecasts of this econometric model were compared to naïve forecast models (the forecast of this year's basis is equal to last year's basis or the average of last three years' bases). The results showed that their econometric model outperformed naive models in turning point error and goodness-of-fit measures (RMSE and Theil's U_2). Since the values of the regressors were unknown until early November for the November basis forecast, ancillary forecasts were needed for these variables. They showed that the basis forecasts based on ancillary forecasts differed a lot from the in-sample forecasts. It was concluded that the difficulty of making precise ancillary forecasts of the regressors was a serious limitation of their basis behavior models.

Brorsen et al. (1985) investigated dynamic price relationships of corn, sorghum, and soybeans in different locations. Both price relationships across space and among commodities in the same location were studied. Three markets: Kansas City, Houston and the Triangle Area of Texas were selected to represent a major terminal market, an export market, and a country point market, respectively. They first developed a static general equilibrium model based on market equilibrium conditions. They argued that this structural equation was not appropriate if the markets were in disequilibrium; instead, a dynamic structural equation that incorporated the past structural shifters was more appropriate. Because of the difficulty in measuring relevant supply/demand shifters, time series modeling was used as an alternative approach. In their time series model, prices of corn and grain sorghum were functions of a deterministic component and a stochastic process. The deterministic part, the trend, was removed by first differencing before the time series modeling. Then bivariate autoregressive (AR) models were used to model each pair of prefiltered series. Granger causality and dynamic multipliers were used to investigate the price relationships. The results showed that prices across space were slower to reach an equilibrium than prices in one location. Kansas City prices led the other two prices. The null hypothesis of efficient markets was rejected since price adjustments were not instantaneous.

Kahl and Curtis (1986) analyzed various factors that influence the magnitude and variation of a grain surplus area (Illinois) and a grain deficit area (North Carolina). They observed that the cash price generally remains below the nearby futures price throughout the marketing year in major production areas; in contrast, the cash price typically rises above the nearby futures price in the later months of the marketing year in grain deficit areas. In the empirical model, seven factors were hypothesized to determine the corn basis. Those were: corn stocks, soybean stocks, storage costs, cash price level, local corn consumption, transportation costs and quarterly dummy variables. Seemingly unrelated regression was used to take into account the regional interdependence in cash markets in Illinois and South Carolina. The results indicated that soybean stocks, storage costs, price level and transportation costs were significant in the Illinois corn basis model. In the South Carolina model, corn stocks, storage costs, grain consuming animal units, transportation cost and the dummy variable for the August through October period[6] were significant with the expected signs. The coefficient of transportation cost for Illinois was positive, though they expected a negative sign. Their calculated elasticities at the mean showed that the South Carolina corn basis tended to be more responsive to changes in the independent variables than the basis in Illinois. They concluded that the seemingly unrelated regression seemed to be superior to separate models for each state or one aggregate model for the two regions.

Hauser, Garcia and Tumblin (1990) evaluated alternative soybean basis expectation (forecast) models and how basis expectations play a role in measuring hedge effectiveness. One focus of their paper was to develop a measure of hedging effectiveness which explicitly incorporated basis expectations into its formulation. Using data from ten elevators across Illinois from 1966 to 1983, they compared forecast performance of naive models, market-based models and regression models. Three naive formulations used current basis, last year's expiration basis, or an average of the expiration basis over the last three years as the expected basis, respectively. Two market-based formulations used the return to storage implied by the prices of the two nearest futures contracts. Four other regression models were also analyzed.[7] They selected closing futures prices for the November, March, and July soybean contracts for the analysis. They focused on the three month period prior to expiration for each contract,

---

[6] Other dummy variables are for periods: November to January, and February to April.

[7] These regression models were discussed in a previous version of this paper by Garcia, Hauser and Tumblin (1986). They regressed current basis during the pre-hedging period on time to expiration (TTE). The TTE coefficient was used during the hedging period to extrapolate the current basis to expiration. One other regression model used basis in the previous year's hedging period. Two other alternatives used the intercept terms of these two regressions as the expected basis.

and two hedging periods of approximately 1.5 months each were considered. A Theil's coefficient of inequality was calculated for each basis expectation model by contract, year, and period. Comparison of the Theil's coefficients suggested that simple models provided reliable basis forecasts. The market-based models, based on observable price spreads, worked well for the old-crop contracts (March and July) and for the second period of the November contract. For the first period of the November contract, it was argued that "the return of storage implied by the spread between the November and January futures prices does not reflect the high convenience yield usually present during this pre-harvest period. As the new crop is harvested during period 2, the convenience yield is presumably less important in determining the current change in basis; consequently, the futures-price spread model produces a more accurate forecast" (p. 130). It was found that the best forecasts for the first period of the November contract were from the models using either the previous year's expiration basis or the average of past three years' expiration basis.

Thompson, Eales and Hauser (1990) investigated the relationship of spatial grain (corn and soybeans) basis changes between cash market locations in the North Central region and used the result to explain how cash prices are linked to other cash market prices and futures price. Two methods were used in their analysis: comparisons of cash and futures price changes and comparisons of spatial basis changes. Basis here was defined as the nearby futures price minus cash price. In analyzing pairwise basis relationship between cash market locations, they assumed that the cash price change at one location was a weighted sum of the futures price change and the cash price change at the other location. They studied basis at delivery locations for corn and soybeans futures contacts,[8] Minneapolis, and six Illinois county elevators for the periods 1978-1986 and 1979-1986 for corn and soybeans, respectively. The results showed that cash and futures price behavior differed between delivery and non-delivery months. The results of the analysis of basis relationships indicated cash price changes were linked to futures price changes rather than to other cash market price changes, and there were seasonal difference in the extent of this relationship. Corn cash prices were more closely related to futures in non-delivery months than during delivery months, while there was no such contrast for soybeans. In general, soybean basis changes were weakly related across markets in most months. They stated that Chicago's weaker relationship to other cash markets "more likely indicate that except for as a delivery location for futures contracts, Chicago is no longer well integrated with the remainder of the grain marketing channel" (p. 254).

Naik and Leuthold (1991) developed a theoretical basis relationship model based on expected utility theory. Their basis relationship consisted of storage cost, opportunity cost, expected basis at maturity, speculation and basis risk premium. This paper empirically tested the components of the corn basis using cash prices for East Central Illinois elevators and CBOT futures prices. They showed that the basis risk premium was determined by the correlation coefficient between cash and futures prices at maturity. If the absolute value of the coefficient is not equal to one, then a basis risk premium exists. The speculative component of the basis was tested by examining the regression coefficient of cash price on

---

[8] Chicago, Toledo and St. Louis for corn futures, and the Chicago and Toledo for soybean contracts.

futures price during the maturity month. The coefficient of one indicated there was no speculative component. The testing of expected maturity basis as a component of the basis model was more complicated. It was argued that if basis could be predicted by appropriate factors, then expected basis exists at maturity. A basis prediction model was derived using market equilibrium conditions of cash and futures markets. Their correlation coefficients showed that a maturity basis risk premium existed in approximately one-half of the contract tested. Slightly more than 50% of the regression coefficients (regressing cash price on futures price) were significantly different from one, which suggested a speculative component also existed for approximately one-half of the corn contracts tested. The basis prediction equation indicated that expected maturity basis could be predicted one to three months before maturity, which in turn suggested that expected maturity basis exists.

Karlson, Anderson and Dahl (1993) analyzed the role of futures markets in corn marketing decisions. They studied the seasonality of a local Minnesota July corn basis during 1982-1992 and demonstrated how the basis information can be used to earn returns on stored corn. "Futures prices are the product of a considerable volume of trading and are very sensitive to new market information on supply and demand affecting the general level of market prices. Local cash prices reflect these changes in the futures price, but they also reflect local economic factors such as transportation costs and availability; local supply and demand conditions; and the availability of local storage" (p. 9-10). The authors calculated the July basis (not a nearby basis) for corn at Clarkfield, Minnesota to examine its seasonality and the sources of the seasonal pattern. They argued that basis patterns differed across different marketing years because market fundamentals were different from year to year. They specified three market fundamentals that explain basis pattern: prime rate of interest, stocks relative to storage capacity, and weather. As the prime rate of interest went down, or as grain supply relative to the availability of storage went down, or as weather conditions got worse (such as drought), the basis got stronger (cash price gains relative to futures price).

In a recent paper, Tomek (1996) outlined a simple model of price level and basis behavior based on supply and demand of storage. A two-period (no production in the second period) framework was used for the model. A supply of storage equation was obtained from the profit maximization rule over the two periods. The equation showed that a basis or a price of storage is equal to the opportunity cost of storage, the direct costs of storage, and convenience yield. The demand equation for storage was derived assuming production was consumed over two periods (i.e. second period consumption equals inventory). The equation stated that the demand for storage depends on expected demand for consumption in period two relative to the demand in period one for current consumption. The supply and demand equations for storage were plotted to show that the demand shifts have different impacts on both price level and basis at different storage levels. In terms of basis forecasting, Tomek distinguished two types of models: interyear and intrayear. He argued that it is difficult to explain the basis variability from year to year in the interyear model, and the model also depends on ancillary forecasts of variables. He then concluded that it is not surprising the simple time-series or naïve models have been often used for the basis forecasts. For the intrayear basis model, he set up a revenue equation for the hedger. It showed that store and hedge decisions depend on expected basis convergence (basis change). A simple model could be utilized to forecast the basis convergence using the initial basis as the explanatory variable.

The model is potentially useful though, as some argued, the basis has relatively small ability to forecast price changes.

CHAPTER 3. BASIS BEHAVIOR MODEL

The first objective of this study is to study basis behavior of selected grain markets. In this chapter, a basis behavior model is established to analyze past basis behavior. An overview of the grain and oilseeds markets is presented first, so that the models and their relationships to the grain market economic environment will be in clearer perspective.

Grain and Oilseeds Markets

Production and Disappearances

The grain industry is an important part of U.S. economy. Grains are roughly classified into three categories as food grain (wheat and rice), feed grains (corn, barley, sorghum, oats, and rye), and oilseeds (soybeans, sunflowers, rapeseed, and flax). Grain production in the U.S. has been increasing since the 1960s. Corn production increased from around 4.5 billion bushels in later '60s to an average of 8.2 billion bushels in the first half of '90s (see Figure 3.1). The Northern Plains states (CO, KS, MT, NE, ND, SD, and WY) had the largest increase while the production in Cornbelt states (IL, IN, IA, MI, MN, MO, OH, and WI) is the largest share in total U.S. corn production. The Cornbelt states' share had dropped from around 75 percent to less than 70 percent from later '60s to early '90s and the Northern Plains states share has increased from 13 percent to 20 percent. Over the last 30 years, soybean production has increased from about one billion bushels to more than two billion bushels (see Figure 3.2). Both the Cornbelt states and Northern Plains states (only KS, NE, ND, and SD) increased their shares of the total U.S. soybean production from about 65 percent to more than 70 percent for the Cornbelt and 5 percent to 10 percent for the Northern Plains, respectively.

Domestic use accounts for more than half of the total supply for both corn and soybeans. Corn domestic use increased from less than 4 to more than 6 billion bushels over the last 30 years; it increased from 0.6 to 1.5 billion bushels for soybeans. The domestic use of soybeans ranges from 50-60 percent most of the years, with slightly lower percentage from later '70s to later '80s.

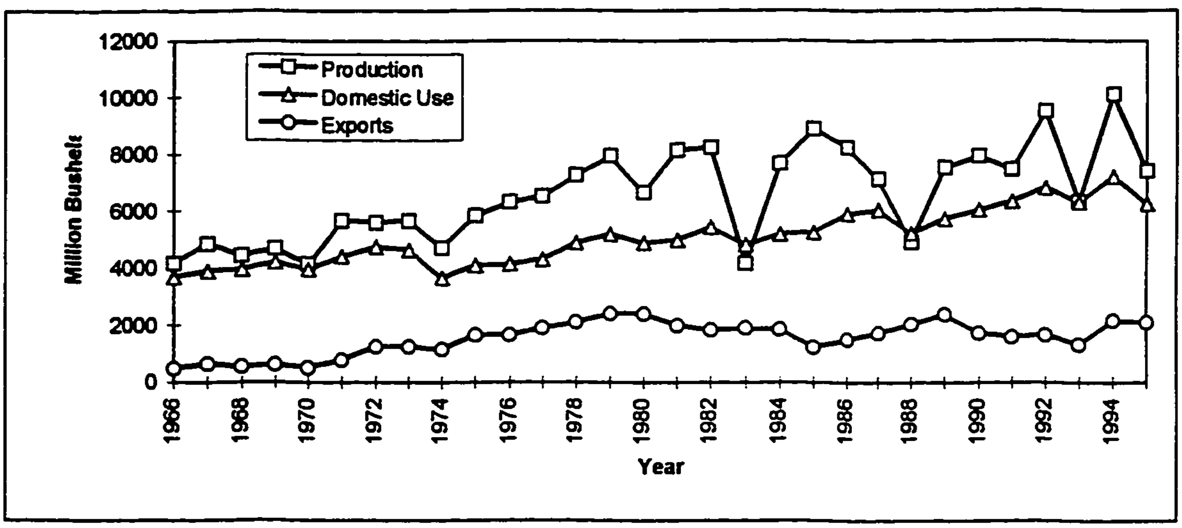

Figure 3.1. U.S. corn production, domestic use and exports, 1966-1995

U.S. corn exports have been increasing over the last 30 years from around 0.5 to about 2 billion bushels, or from around 10 percent of total supply to around 20 percent. The highest share was in late '70s, around a 25 percent share. Soybean exports have increased from around 0.3 to 0.8 billion bushels over the last 30 years. Though the export share of soybean use has been quite stable around 30 percent recently, it peaked in late '70s to early '80s with more than 36 percent share. Imports have been negligible.

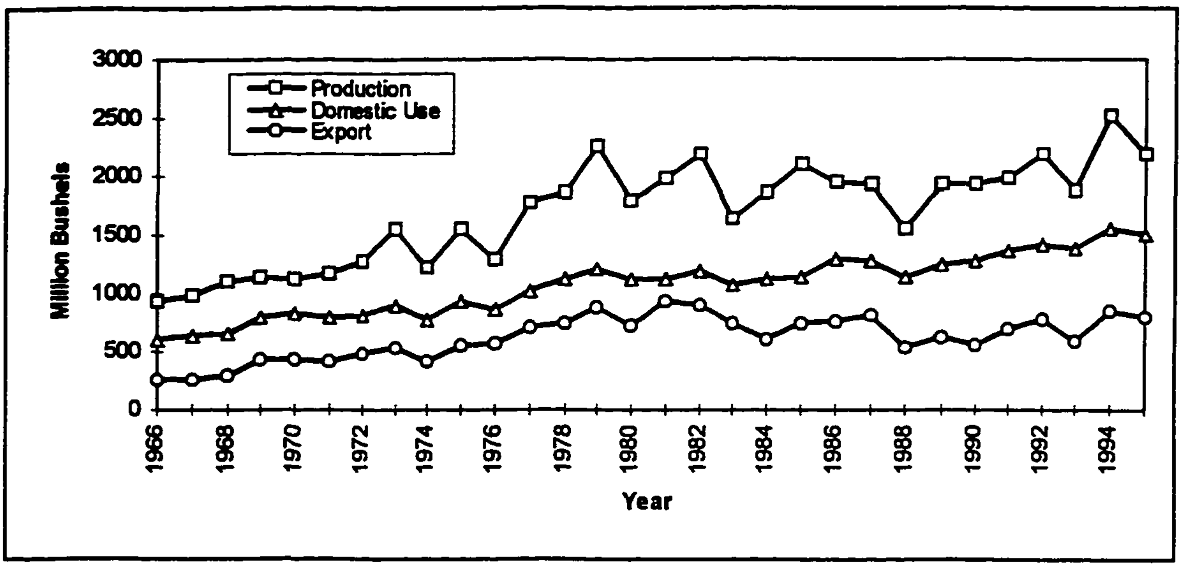

Figure 3.2. U.S. soybean production, domestic use and exports, 1966-1995

<u>Grain Storage and Movement</u>

Production of grain occurs once a year. The harvest season is a relatively short period of time, while consumption of grain is evenly dispersed through out the year. Storage is necessary to smooth the consumption over the year. Storage occurs at each stage of the grain marketing system: farm, country elevator, terminal and port elevator, and processing plant. These are usually grouped into on-farm and off-farm storage facilities. Corn is stored on farm for use as feed or later sale, while soybeans are stored mostly for later sale. Grain storage capacity has expanded at all stages on the U.S. grain complex over the years, with the greatest growth near the point of production.

The major grain production areas (Cornbelt and Plains states) are quite distant from major domestic population centers on the East Coast and West Coast and from the export ports along the East Coast, Gulf Coast, and West Coast (except Great Lakes). Because of this, transportation plays a major role in grain marketing. There are several means to move the grain across the country, mainly barge, rail and truck (on waterway, railway and highway, respectively). Table 3.1 summarized from the studies of Fruin et al. (1990) and Larson et al. (1990), shows the percentage of these three modes of transportation used to move corn and soybeans out of state and to export ports in 1985. Different means of transportation are used for different market outlets. Exported grain often moves by barge to the Gulf, or rail to the East and West coast, and domestic users rely mostly on truck and rail transportation.

Table 3.1. 1985 Grain Shipment and Mode of Transportation (%)

		Truck	Rail	Barge
Corn				
	Interstate Movement	21.12	46.84	32.04
	Shipments to Export	5.58	38.76	55.66
Soybean				
	Interstate Movement	88.00	11.60	0.40
	Shipments to Export	10.30	18.60	71.10

Source: Fruin et al. and Larson et al.

Among these three transportation industries for grain, the railroad industry seems to be the least competitive with only a small number of firms. Many grain shippers are served by only one railroad. On the other hand, the barge and truck industries have a competitive structure that approaches pure competition (Sorenson 1993). Since the supply of barge services is substantially less elastic than truck service, the short-run volatility of barge rates is higher than that of truck rates.

There have been dramatic grain marketing changes over the last few decades. One major change is the general decline of exports through the Great Lakes and the increased importance of exports through Mississippi River ports. According to U.S. Department of Agriculture, "nearly one-fourth of the United States' agricultural exports move down the

Mississippi River through New Orleans, which has reigned for the past two centuries as queen of the U.S. ports for bulk commodities" (p. 4). Along with rail rate deregulation, and grain processing facility relocation, the export route shift has resulted in the structural change in grain movement over the country. In general the grain receipts at all terminal markets decreased while export elevators along the Mississippi River increased their share of grain receipts over the years.

Modeling Basis Behavior

Conceptual Model

Basis at a location involves two price relationships: (1) delivery point cash and futures price relationship, and (2) local and delivery point cash prices relationship. The first component is a temporal price difference which, according to the theory of storage, should equal the storage return (or price of storage). The second component represents spatial price difference. The law of one price suggests that it will be equal to transportation cost between two locations. That is, the basis can be viewed as the sum of storage return and transportation cost. The price of storage (par delivery point basis) is determined by both the supply of storage and the demand for storage. Tomek's (1996) supply of storage equation expressed the price of storage as a function of opportunity cost, direct storage cost and convenience yield, and the demand for storage equation had inventory over a time interval is a function of expected demand for consumption in period two relative to the demand in period one for current consumption, production and the price of storage. A reduced form for the price of storage, derived from these two equations, shows that the price for storage is a function of production, opportunity cost, relative demand in two periods, direct storage cost, and convenience yield. The basis--the sum of storage and transportation costs (assuming the transportation cost as given)--is affected conceptually by storage and transportation costs, production and stocks, and local economic conditions such as local grain consumption, and constraints of storage and transportation capacities.[9]

(EQ 3.1) $$BASIS_i = f(SCost_i, TCost_i, Supply_i, Demand_i, Storage_i, Others_i)$$

where the subscript stands for i^{th} region or market. The first two terms are storage cost and transportation cost, respectively. The remaining terms stand for variables representing supply, demand, storage constraints and other relevant factors.

Basis

Basis is defined here is the nearby basis, the difference between the price of nearby futures contract and local cash price, as the term is used in the grain business. The basis corresponding a particular futures contract is usually used in academic studies. Since this study is practically oriented, the nearby basis is the focus here. Basis always refers to the nearby basis in the present study.

Storage costs

There are on-farm storage and off-farm (commercial) storage facilities. On-farm storage accounts for the majority of storage space, 58 percent of the total storage facilities.

[9] A sketch of formal derivation is included in appendix A.

Storage costs, including warehouse charges, interest and insurance, are the major components of temporal price difference. Most of these costs are fixed once the storage facility is built, whether grain is stored or not. The variable costs are quite small, except interest on the capital tied up in grain (the opportunity cost of holding grain stocks), which is the interest rate times the cash price. Time is also a factor in this interest cost, since less interest is forgone with a shorter time of storage. The variable storage costs will differ when the grain producers or merchandisers use their own storage facilities or rent storage spaces. As storage costs increase, the basis will be wider.

Transportation costs

Grain transportation takes the forms of rail, barge and truck. The transportation costs account for the spatial price differences according to the law of one price. So theoretically, spatial price differences should vary one-to-one with transportation costs, unless deficit areas switch to surplus areas. For example, if a grain elevator ships grain to the river port, the rail rates are simply subtracted from the corresponding river port prices in the manager's calculation of local grain prices. Since not much grain is shipped to Chicago for delivery, the transportation cost of shipping grain from local markets to Chicago may not be relevant to study local basis. As most of the grain in Cornbelt states moves down the Mississippi River, barge rates from ports along the Mississippi River to Gulf Ports are used to represent the transportation costs for various local markets. This proxy will be a good approximation if the barge rates effects on basis in different regions are highly correlated with each other. Rail and truck rates could also affect grain cash prices, especially those where much grain is used domestically. Because of the potentially important role of transportation on local basis behavior over time, an effort will be made to investigate how barge and rail rates affect local basis behavior. For example, during the time when the river is too low for barge in the summer or too cold (ice) in the winter, Iowa price will decline, resulting in a wider basis. Rail may be used to move the grain to the Gulf coast relieve the short supply.

Supply and demand

Local cash prices of grain will be affected by local production, stocks and consumption in addition to overall national or international supply and demand conditions for grain. Some local price changes may be at different rates than overall price change due to differences in local supply and demand conditions, which affect storage and basis. In general, the local production and stock variables relative to local storage capacity are believed to capture the demand and supply for storage.

In theory, higher export demand puts pressure on immediate delivery and bids up current price, and this will result a strengthening of the basis. On the other hand, a large supply will put downward pressure on current price and weaken the basis.

Storage constraint

Storage capacity limits the amount of grain which can be stored under roof. When the demand for storage exceeds this amount, it not only will bid up the storage price (if more demand for higher cost commercial storage), but also will put more pressure on transportation services (more has to be marketed now, it puts a downward pressure on current market price).

This happens usually during harvest time. Some corn and soybeans are stored outside at the peak of harvest in large crop years, which results in extra shrink and handling costs. All these effects will push the local basis wider. On the other hand, when demand for storage is within the capacity, grain merchandisers have better (less costly and less constrained) storage options.

Others

"Others"—other variables which are candidates for inclusion in the model may include supply and demand factors for other products competing for storage or transportation with the product under study. There may be possible arbitrage across products and across regions to be taken into account. Historical basis plots usually show seasonal patterns for both corn and soybeans, reflecting seasonal production and demand patterns. The basis is normally weak at harvest months as large supply puts downward pressure on cash price, and generally strengthens over the marketing year as less grain supply available. The degree of strengthening over the marketing year may vary from year to year. It depends on market fundamentals: prime rate of interest, availability of storage capacity relative to grain supply, and weather conditions (Karlson et al. 1993).

All previous basis studies (Martin et al. 1980, Kahl 1982, Garcia and Good 1983, Powers and Johnson 1983, Kahl and Curtis 1986) used models based on this general approach, though the actual data used in estimation were quite different. These studies were conducted in the early '80s, and they were confined to one or two markets. They studied basis in Ontario, Chicago, Illinois, Wisconsin, and Illinois and North Carolina, respectively.

There have been significant changes in grain marketing since then. Figure 3.3 shows that Chicago nearby corn basis pattern seems changed after 1986 (a similar pattern is found in basis plots of other markets). Besides, there is also a need to study basis behavior in a broader spectrum of locations which represent large volumes of grain in the changed market structure. Basis behavior may differ among different markets such as production areas, port markets and terminal markets.

The cash price relationships among different markets may change as the marketing pattern changes. The question then is how the local basis is affected by the changes in marketing pattern. For example, if Iowa delivers more of its grain to the Gulf port than before, its cash price may be more closely related to Gulf price than to Chicago; Iowa's price relationship with Chicago may be affected. According to the law of one price, one major component in spatial price difference is transportation cost, yet this is not well captured in previous studies. Garcia and Good (1983) used barge rates on the Mississippi River, Powers and Johnson (1983) used a trend variable in the place of transportation cost, Kahl and Curtis (1986) took a U.S. rail rate index as the transportation cost variable, and Martin et al. (1980) did not include a transportation cost variable. As discussed earlier, three modes of transportation are used in moving corn around the country. Barge cost is a component of price difference between local markets along the river and the Gulf ports. But rail rates may be more influential at some times of the year, and between other markets. Improved specification of an appropriate transportation cost variable may be helpful in explaining basis behavior.

Four out of the five fundamental basis models discussed above used a single equation approach, and OLS in estimating the basis equation. Grain markets may be well integrated, and cash prices in different locations may interact with each other. A system of equations approach may capture market behavior better. Kahl and Curtis (1986) found that Seemingly Unrelated Regression outperformed Ordinary Least Squares. In this study, interregional price effects will be taken into account which requires a system approach to model the basis behavior rather than the single equation approach.

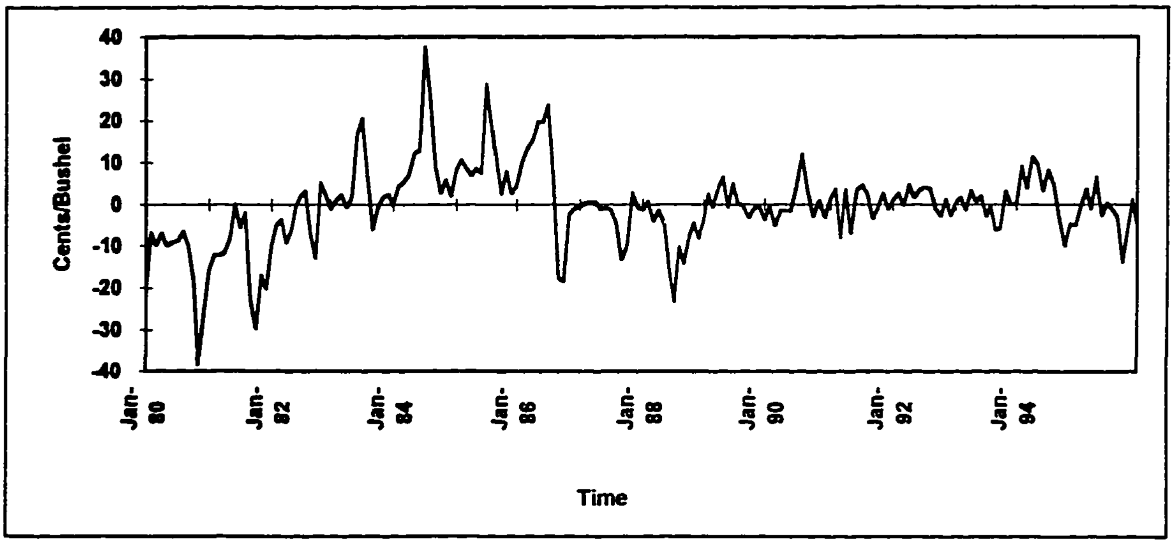

Figure 3.3. Chicago nearby corn basis, 1980-1995

In summary, this model studies several grain markets across the U.S. at the same time, taking into account the regional price interrelationships. It also studies both corn and soybean basis at each market location, which takes into account the inter-product price relationships between them. Since transportation is the single most important factor in basis relationships, it will be emphasized in the basis behavior analysis. In addition, seasonal variation in the basis behavior is a major component in this study. This study can provide an updated analysis on corn and soybean basis, to determine whether basis patterns and factors influencing basis have changed since most of the previous studies were done in the 1980s.

Empirical Model

The basis behavior model is estimated contract by contract for both corn (EQ 3.2) and soybean (EQ 3.3). Table 3.2 lists the variable definitions for the following models.

The corn basis behavior structural model is

(EQ 3.2) $$BS\_i_j = \alpha_j + \beta_{1j}*RP\_i_j + \beta_{2j}*BR\_i_j + \beta_{3j}*CN\_i_j + \beta_{4j}*EX\_i_j + \beta_{5j}*AUC\_G_j + \beta_{6j}*TTM_j,$$

for ith market and jth contract, where i = nia, cil, chg, stl, tol, gul, pnw, and rmd, and j = March, May, July, September, and December. The soybean basis behavior structural model is

(EQ 3.3) $$BS\_i_j = \alpha_j + \beta_{1j}*RP\_i_j + \beta_{2j}*BR\_i_j + \beta_{3j}*SB\_i_j + \beta_{4j}*EX\_i_j + \beta_{5j}*CR\_i_j + \beta_{6j}*AUC\_G_j + \beta_{7j}*TTM_j,$$

for ith market and jth contract, where i = nia, cil, chg, stl, tol, gul, and rmd, and j = January, March, May, July, September, and November.

Variables without location subscript, such as AUC_G (animal units consuming grain) is U.S. data, TTM (time to maturity) is the same for all locations. The following equation is an example of the empirical corn basis model being estimated for Northeast Iowa:

(EQ 3.4) $$BS\_nia_j = \alpha_j + \beta_{1j}*RP\_nia_j + \beta_{2j}*BR\_nia_j + \beta_{3j}*CN\_nia_j + \beta_{4j}*EX\_nia_j + \beta_{5j}*AUC\_G_j + \beta_{6j}*TTM_j,$$

for jth contract, where $_j$ = March, May, July, September, and December.

Table 3.2. Variable Definitions

		Definitions
Variable Abbreviations	BS	Nearby basis
	RP	Opportunity cost (prime interest rate times monthly cash price)
	BR	Barge rate
	SB	Soybean production
	CN	Corn production
	EX	Corn or soybean export
	CR	Soybean crushing
	DM	Dummy variable
	PIR	Prime Interest Rate
Location Abbreviations	CIL	Central Illinois
	NIA	Northeast Iowa
	STL	St. Louis
	CHG	Chicago
	TOL	Toledo
	GUL	Gulf Ports
	RMD	Richmond, Virginia
	PNW	Pacific Northwest
Variables for Iowa Model[a]	RP_NIA	Storage cost of Northeast Iowa
	BR_NIA	The barge rate of Upper Mississippi River
	CN_NIA	Corn production of Iowa and its surrounding states
	EX_NIA	Export through Mississippi River ports
	AUC_G	Animal unit consuming grain,
	TTM	Time (month) to maturity of its nearby contract for corn or soybean contracts.

[a] Variables for other location models are defined in similar fashion, only differs in location abbreviation. For soybean models, SB is used in the place of CN. Variables without CN or SB are defined the same in both corn and soybean models.

The product of cash price and prime interest rate is the opportunity cost of holding corn and soybean stocks; it is expected to have a negative effect on local basis, since a higher storage cost should result in a larger price difference between current and later prices. When transportation costs (barge and rail rates) get higher, the spatial price difference becomes larger, therefore barge and rail rates are expected to have negative coefficients. The local supply variable is represented by ratio of annual production level to storage capacity.

Production relative to storage capacity is expected to be the best index to represent supply, as basis will widen as outside storage is employed. This means that all commodities competing for storage space ideally should be jointly considered, both beginning inventory and production, with the primary impact expected to be found only at or shortly after harvest. The production and storage capacity variables in the model are regional variables (which sums the productions and capacities of surrounding states), though localized state variables are also estimated. The demand factors include animal units consuming grain and U.S. monthly export volume at the port most relevant to the location being analyzed. For some export variables, the total exports in several ports are used instead of exports by a single port; for example, the export variable in the Chicago model is the sum of both Chicago port and Gulf port exports.

Data and Estimation Methods

Data

The study utilizes monthly time series data from January, 1980 to December, 1995.[10] Futures prices for corn and soybeans from January, 1980 to April, 1996 were generously provided by the Chicago Board of Trade. Monthly average settlement prices are selected as the futures prices. The nearby future prices (NBFPs) are compiled from the futures prices of all the contracts. For any month (either delivery or non-delivery month), the current futures prices of next closest futures contract are taken as this month's nearby futures prices. For example, there are five contract months for corn: March, May, July, September and December. The daily nearby futures prices in January and February are the March contract, and March and April's daily NBFPs are the May contract, etc. The NBFPs for soybeans are

[10] Monthly data is used in this study because weekly cash price time series are not available for some locations or back to the 1980s.

constructed in the same fashion though there are seven contract months: January, March, May, July, August, September and November (see Table 3.3). Monthly NBFPs are arithmetic monthly averages of the daily prices.

The monthly grain cash prices are from USDA for various locations across United States. Most of the price series start in the 1974/1975 crop season and end with the 1995/1996 season. The corn cash prices for the following locations will be used: Chicago, St. Louis, Toledo, Gulf Port (Baton Rouge), Pacific Northwest (Portland), Central Illinois, Northeast Iowa, and Richmond. The selected soybean cash prices locations are: Chicago, St. Louis, Toledo, Gulf Coast (Baton Rouge), Central Illinois, Northeast Iowa, and Richmond. The nearby basis series are obtained by subtracting the nearby futures price from the cash price. Table 3.4 has some basic statistics of the basis series. It shows that bases are negative for major production areas and positive for port market locations. Richmond soybean average basis is negative because it is a soybean surplus production area, while it is positive for corn. Richmond is a surplus production region at harvest and a few months afterward, but it is a

Table 3.3. Corn and soybean nearby contracts

	Futures Contracts	
Month	Corn	Soybean
Jan	March	March
Feb	March	March
Mar	May	May
Apr	May	May
May	July	July
Jun	July	July
Jul	September	August
Aug	September	September
Sep	December	November
Oct	December	November
Nov	December	January
Dec	March	January

Table 3.4. Basic statistics of the nearby corn and soybean basis

	Locations							
	Chicago	St. Louis	Toledo	Gulf	NE Iowa	C Illinois	Richmond	Pacific NW
	Corn							
MEAN	-0.55	2.33	-6.99	22.34	-24.94	-12.49	9.49	51.76
STD DEV	9.68	11.38	11.38	10.65	15.10	11.76	14.04	10.66
MIN	-38.40	-37.71	-43.40	-9.40	-83.40	-50.71	-34.23	18.24
MAX	37.50	35.65	33.50	54.65	12.50	18.72	53.50	75.95
	Soybean							
MEAN	-6.94	1.57	-9.95	24.60	-32.29	-15.03	-16.47	
STD DEV	10.72	12.39	11.14	12.52	17.17	12.26	12.83	
MIN	-54.43	-37.43	-54.23	-19.43	-85.43	-53.43	-61.43	
MAX	28.14	34.46	26.46	54.46	2.48	20.15	27.48	

deficit area the rest of year, importing corn for the poultry industry. A surprising negative basis for Chicago and Toledo soybean, even though they are delivery points, possibly reflecting average carrying charges for 2-3 months and perceived delivery costs.

Crop production data are available annually from 1966 to 1995 for major production states, mainly in the Cornbelt and Northern Plains, and for the entire United States. Stock variables are available quarterly. Monthly corn and soybean exports by port area are taken from the USDA Grain and Feed Market News. Grain storage capacity for each state is obtained from various USDA publications. These are capacity of off-farm storage facilities. On-farm storage capacity is not available for the entire time from 1980-1995. Grain consuming animal unit numbers are taken from USDA publications Feed Yearbook and Feed Outlook.

Weekly barge rates data series (from January, 1980 to December, 1995) are provided by the USDA. It is published by the St. Louis Merchant Exchange. The series are available for eight segments of Mississippi River and its tributaries (Ohio River and Illinois River) to Louisiana (see Table 3.5). Markets at different locations will use barge rates closest to them. For example, the appropriate barge rate for Central Illinois is the Illinois River segment.

The barge rates series are in the form of percentage of tariff. The tariff is the benchmark barge freight rates used as the basis for quoting barge rates. The tariff is quoted for shipping points on Mississippi, Illinois and Ohio Rivers to New Orleans, LA. The actual barge rates of a particular port can be calculated by multiplying the percentage of tariff of the port by the corresponding tariff. The tariff is quoted in cents per ton. This series has lots of typos in the dates, so there may be also typos in the barge rates. Monthly barge rates are calculated as the arithmetic average of the weekly series.

Table 3.5. River Segments and Barge Rates

Segment		Locations	Closest Markets[a]
TWC	Twin Cities (Upper Miss.)	McGregor, IA through Savage, MN	NE Iowa
MM	Mid Miss.	Winfield, MO to McGregor, IA	
ILL	Illinois River	Grafton, IL to Lockport, IL	Chicago, Central IL
STL	St. Louis	Cairo, IL to Winfield, MO	St. Louis
CINC	Upper Ohio	Louisville, KY through Cincinnati, OH	Toledo
LOH	Lower Ohio	Cairo, IL to Louisville, KY	
CAR-MEM	Lower Miss.	Cairo, IL through Memphis, TN	
MEM-NO	Lower Miss.	Memphis, TN to New Orleans, LA	

[a] The location used barge rate of this segment. Other locations such as Richmond, Pacific NW and Gulf Ports use St. Louis barge rate as a proxy for transportation cost.

A major component of storage cost is the opportunity cost of storing grain. The potential interest earned from selling grain now is foregone; the prime interest rate could be a good index for this opportunity cost. The prime rate charged by banks on short-term business loans is taken from various issues of Annual Statistical Digest and Federal Reserve Bulletin published by the Board of Governors of the Federal Reserve System. All other costs of physical storage facility, such as warehouse charges, interest and insurance, are mostly fixed once the facility is built.

Soybean monthly crushing data by states are from U.S. Bureau of Census publication Fats and Oil: Oilseed Crushing. There are three months' data gap in October, November and December of 1990; the average of previous three years' crushing is taken for these missing data. The crushing data in 1991 is a quarterly series; monthly crushing is taken as one third of the corresponding quarterly data.[11]

Stationarity

Stationarity properties of all the data series are tested. A stochastic process, Y_t, is said to be weakly stationary (or covariance stationary) if neither the mean nor the autocovariance depend on time. Both simple and augmented Dicky-Fuller unit root tests are conducted for all the time series.

(EQ 3.5) $$DY_t = m + g^{*}Y_{t-1} + e_t$$

(EQ 3.6) $$DY_t = m + g^{*}Y_{t-1} + S_j f_j DY_{t-j} + e_t \qquad (j=1, \ldots, r-1)$$

where the first one is the simple D-F test equation and the second is the augmented D-F test. The null hypothesis of $g^{*}=0$ (unit root) is tested for both equations. The tests for stationarity with a time trend can be obtained by adding a time trend variable in each of the above equations. The methodology of the tests is the same as the conventional t test, but the critical

[11] Historical crushing data shows that monthly crush is evenly distributed within each quarter.

values have to be revised. These critical values are derived by Dicky and Fuller through a Monte Carlo study. These appropriate critical values depend on the form of regression, such as whether intercept term and time trend are included, and sample size. The significance level used in all of the following tests is 5% unless otherwise specified.

For the monthly barge rates series, the null hypothesis of a unit root is rejected in six out of eight monthly series by augmented D-F tests. Another series is rejected by the augmented test but not the simple test. The grain production data from 1966 to 1995 are tested, and the results show that they are nonstationary except Cornbelt and North Plains corn productions using the simple D-F test. A close look at the production plots reveal that the production series have an upward time trend. All the corn production data series tested (U.S., Cornbelt and North Plains productions) are stationary with a time trend in both simple and augmented D-F tests, except U.S. production in augmented test. The corresponding soybean production series are stationary in simple D-F test with a time trend, but not with the augmented test. The beginning stocks of corn and soybeans are nonstationary (except the soybean beginning stock in augmented D-F test with a time trend).

In summary, the above unit root tests for all the independent variables show mixed results with most of the series being stationary. The basis time series, the dependent variables, are all stationary by both simple and augmented D-F tests. This may due to their being derived from a combination of several nearby futures contract time series segments, or due to the stationarity properties of transportation cost and storage cost (the interest rate in this case), since these costs are components of local basis. Stationary dependent variables mean that the OLS estimator is consistent but has a non-standard asymptotic distribution (Tomek and Myers 1993).

Estimation Methods

A stacked model is constructed to test the seasonality and to test whether SUR can improve the model fit over OLS. The stacked model pools all the contract models together, essentially the same model as in equation 3.2 and 3.3 but without subscript j and with additional dummy variables. Equation 3.7 is a stacked corn basis model for Northeast Iowa.

$$\text{(EQ 3.7)}\quad BS\_nia = \alpha + \beta_1{*}RP\_nia + \beta_2{*}BR\_nia + \beta_3{*}CN\_nia + \beta_4{*}CN\_nia + \beta_5{*}AUC\_G + \beta_6{*}TTM + \Sigma_s(\beta_s DM_s) + \Sigma_s(\gamma_{1s}DM_s{*}RP\_nia) + \Sigma_s(\gamma_{2s}DM_s{*}BR\_nia) + \Sigma_s(\gamma_{3s}DM_s{*}CN\_nia) + \Sigma_s(\gamma_{4s}DM_s{*}CN\_nia) + \Sigma_s(\gamma_{5s}DM_s{*}AUC\_G) + \Sigma_s(\gamma_{6s}DM_s{*}TTM),\quad \text{s = month or contract.}$$

In order to test whether the seasonal patterns are significantly different from zero, three stacked model specifications are estimated: (1) with monthly intercept dummies ($\beta_6=\gamma_s=0$); (2) with contract dummies and a time to maturity variable, which is defined as the months left before the contract expires ($\gamma_s=0$); and (3) specification (2) plus contract slope dummies.

This null hypothesis of no seasonal difference in basis is tested, using the nested F-test, which tests joint significance of all the seasonal dummies in each equation. Table 3.6 shows that the F statistics of monthly dummy models are significant in six of eight corn market equations (six at 10%, five at 5% and three at 1% significance levels). The numbers of significant nested F statistics for the contract dummy stacked models are seven, six and five at these three significance levels, respectively. For model specifications with both intercept and slope contract dummies, all the F statistics are significant at the 1% significance level.

In summary, these nested F-test statistics reject the null hypothesis that seasonal dummies, either contract or month, are zero in most of the market equations. Similar results are found for the soybean equations: we reject the null hypothesis that all monthly dummies are jointly not different from zero in six out of seven market equations, in all seven market equations with contract intercept dummies, and all seven equations with both intercept and slope contract dummies.

The root mean square error (RMSE) and R^2 of different model specifications show the differences in the model fit. For both corn and soybean stacked models, except in few cases,

Table 3.6. Stacked models F statistics of nested F tests on seasonal dummies

Market Locations	(1) Month Intercept Dummies	(2) Contract Intercept Dummies	(3) Contract Intercept and Slope Dummies
Corn			
Northeast IA	1.28	2.10*	2.07***
Central IL	1.96**	4.42***	2.49***
Chicago	2.25**	4.10***	2.21***
St. Louis	1.26	1.95	2.16***
Toledo	7.97***	19.73***	4.62***
Gulf Ports	1.76*	3.29**	2.07***
Pacific NW	2.40***	5.65***	2.84***
Richmond	3.12***	4.15***	2.14***
Soybean			
Northeast IA	1.50	2.25**	2.43***
Central IL	2.63***	3.66***	2.39***
Chicago	2.69***	3.56***	2.37***
St. Louis	2.86***	3.53***	2.08***
Toledo	7.52***	10.04***	2.46***
Gulf	2.67***	2.77***	2.38***
Richmond	4.13***	5.34***	1.61*

Note: single star (*), double stars (**) and triple stars (***) indicate significance level of 10%, 5% and 1%, respectively.

the adjusted R squares are the highest for model specification with contract intercept and slope dummy variables, second highest is model specification with month intercept dummies, and the model specification with contract intercept dummies has lowest adjusted R-squares. The RMSE statistics show similar results (for detailed comparison of R^2 and RMSE statistics among these models, see appendix Table B.1).

In choosing between model specifications with different forms of seasonal dummies (contract and monthly dummies), non-nested tests have to be used. However, because contract dummies are linearly related to monthly dummies, both the non-nested F test and J test are not practical. Non-nested tests fail in this special case. But from the nested F test statistics in table 3.6, model specifications with both intercept and slope contract dummies may be the most reliable and most appropriate to use because it has more significant F statistics than other model specifications. Though the stacked model with seasonal dummies can capture the seasonal basis variation, estimation by contract[12] will be more appropriate. The Goldfeld-Quandt tests show that the variances of the contract models are not all equal. It is not surprising that this contract-by-contract modeling approach results in a higher R-squares and lower root mean squared errors than the stacked model with seasonal dummies.

Since basis behavior differs among different market locations, there will be five separate regressions for the corn model in each of the eight market locations, and six regressions for each of the six soybean futures contracts (August and September contracts, each containing only one month of nearby basis, are combined into one model) in each of the

[12] The reduced number of observations is also a problem for estimating the basis model by month; therefore, we only focus on the contract model in this study.

seven market locations. There are a total of 40 and 42 structural basis models for corn and soybeans, respectively.

In estimation, both Ordinary Least Square (OLS) and Seemingly Unrelated Regression (SUR) are used. The market integration tests[13] show that corn and soybean markets are integrated. Therefore, the error terms of each equation may be correlated and SUR can take this cross region interrelation of error terms into account. There is another inter-price relationship, the cross products correlation at the same location. Corn and soybean prices could be correlated because they compete for limited storage space and transportation services; they also are competitive products in their uses in some degree. So, one corn equation and one soybean equation at the same location are grouped together as a system and estimated by SUR to account for the cross products correlation. However the stacked model estimations show that OLS estimators and SUR estimators are very similar. One possible reason is that the correlations among independent variables are high across equations. Several variables, such as prime interest rate and grain consuming animal units, are the same for these equations. Other variables are highly correlated (e.g., the barge rates for different segments of the river) in different equations are highly correlated. This high correlation of independent variables across equations makes SUR less efficient. Therefore, SUR estimation results are not reported.

Durbin-Watson (D-W) statistics indicate that both corn and soybean contract models usually initially exhibite autocorrelation, mostly first order, though a few are second order. To

[13] Ravallion's market integration tests are conducted for a sample of markets. The results show that, at 1% significance level, the hypothesis of market segmentation is rejected for all these markets, while the hypothesis of long-run market integration failed to reject. This suggests that these markets are integrated with the Gulf market.

correct for that, autoregressive error model's, which take into account residual autocorrelations, are estimated by utilizing the Yule-Walker estimation method. When the hypothesis of autocorrelation for contract models is rejected, Ordinary Least Squares is used.

The variance inflation factor (VIF) values for all independent variables in corn contract models are all below 10, indicating that collinearity is not a problem for the corn models. For soybean models, only two variables have VIF values larger than 10 in three equations (out of total of 42 equations) in two locations. This suggests harmful collinearity for these equations. In addition, the condition number, another statistic to check multicollinearity, is also calculated. It generally is consistent with the findings of VIF statistics, though the condition number is not significant for the variables that have VIF larger than 10. In conclusion, multicollinearity is not a problem for both corn and soybean contract models.

Heteroscedsticity is checked by both a portmanteau test (Q*) and a Engle's Lagrange Multiplier (LM) test for the basis models by contract. Most of the test statistics fail to reject the null hypothesis of homoscedasticity.

Possible structural change, as noted from Figure 3.1, of the basis behavior model before and after 1986 is tested. The standard F test (or Chow test) is generally used to test for structural change, but the test is based on the important assumption that the disturbance variance is the same in both regressions. Therefore, first a test of equal variance is conducted using the Goldfeld-Quandt (G-Q) test.[14] The whole data set from 1980 to 1995 is divided into two subsets: 1980-85 and 1987-95. The observations in 1986 are dropped to increase the power of the G-Q test. The tests show that the null hypothesis of homoscedasticity is rejected

[14] The test statistic is $F\ [n_1\text{-}K, n_2\text{-}K] = SSE_1/SSE_2$, with larger value as the numerator. Under the null hypothesis of homoscedasticity (equal variance), it has an F distribution.

at 5% significance level in six out of eight market locations for corn models. The null hypothesis of equal variance for soybean models is rejected for only two out of seven market locations. When the disturbance variance is not the same for several markets, the standard F test can not be used to test the structural change for these markets. Instead, the Wald test is used, which is valid whether or not the disturbance variances are the same. Test results show that the null hypothesis of no structural change is rejected for 22 corn contract models, but not rejected in 18 corn contract models (for soybean model, the null hypothsis is rejected in 27 models and not rejected for 15 contract models). One problem with the Wald test is that it is valid for large samples. For our contract models, each data subset has about 10 and 20 observations, respectively. The W test has a persistently large type 1 error in small and moderate-sized samples, so a larger critical value should be used (Greene 1990, p. 216). If a larger critical value is used, we could reject the null hypothesis of equal variance less frequently. We conclude that the W test results for structural change are not conclusive.[15] However, several alternative model specifications are estimated trying to capture the possible structure change. These model estimation results are discussed in the last section of this chapter.

Estimation Results

Table 3.7 shows the general model fit of the corn contract model. The R-squares show that these contract models have a good fit; most explain between 50-80% of basis variation. There is seasonal variation in the estimated coefficients throughout the year. Table

[15] Of cause, more complicated tests, like the "bounds" test that gives a partial remedy for the problem (Greene), should be used for a more precise result.

3.8 shows the number of significant coefficients with expected sign in each contract model for eight corn markets. Opportunity costs, which represent the storage costs, are usually found negative and significant in early storage season, with seven and five significant coefficients out of eight for contracts March and May, respectively. In Northeast IA, Central IL and Gulf Port, storage cost is also marginally significant (at 10% level) for July and September contract models. Barge rates[16] are mostly significant in the months of May, June, September, October and November. Out of eight markets, it is significant with negative sign in seven and six markets for July and December contract models, respectively. That makes sense because Upper Mississippi River is closed in some of the winter months due to ice and occasionally in few summer months due to low water level or flooding. As expected, corn production relative to storage capacity[17] is an important factor only for the months right after harvest (from December to February) and it does not affect corn basis significantly thereafter. Corn exports only have significant coefficients in two of the port markets: Pacific NW and Richmond. Animal Units Consuming Grain seems to be more important in Pacific NW than other markets. December contract models has the most significant AUC_G coefficients. The last variable, time to maturity, matters from December to February for all eight markets. The significance level also decreases over time for storage cost, as the lower portion of Table 3.8 shows. All other significant coefficients are significant mostly at 1% level (for detailed estimation results of the corn basis models, see appendix Tables B.2-4).

[16] We were able to get a rail rate series from a local Iowa cooperative. The estimation results using this rail rate are not satisfactory (mostly not significant). A rail rate index from the USDA is also used in estimating both corn and soybean models, and is also dropped because of the lack of significance.

[17] It is found that regional production relative to regional storage capacity performs better than state production relative to state capacity

Table 3.7. R^2 and RMSE of corn contract basis models

Contract	NE IA	Central IL	Chicago	St. Louis	Toledo	Gulf Port	Richmond	PNW
				Markets				
				R_square				
March	0.80	0.82	0.70	0.66	0.70	0.72	0.58	0.84
May	0.78	0.82	0.64	0.78	0.66	0.62	0.66	0.57
July	0.70	0.62	0.56	0.55	0.43	0.42	0.45	0.81
September	0.82	0.72	0.62	0.60	0.61	0.60	0.49	0.69
December	0.64	0.61	0.55	0.57	0.60	0.47	0.24	0.80
				Root MSE				
March	6.37	4.44	4.40	5.82	4.96	5.70	8.49	4.13
May	6.11	4.47	4.75	4.35	4.44	4.99	5.91	8.20
July	6.68	6.72	6.33	7.01	7.26	6.92	12.97	4.80
September	8.48	8.70	8.70	9.69	9.76	8.56	13.75	7.39
December	10.68	8.61	7.15	10.31	7.83	10.20	12.15	5.99

Table 3.8. Number of significant coefficients with expected sign in each contract model for eight corn markets

Variables	March	May	July	September	December
			Contracts		
		Total number of significant coefficients			
Opportunity cost	7	5	3	3	1
Transport cost		1	7	2	6
Production	7	1		2	
Export	1		1		
AUC_G	1		1	1	4
TTM	8	2			
		Number of significant coefficient at different significance levels			
Storage cost	6,6,7	4,5,5	0,0,3	0,0,3	0,0,1
Transport cost		0,0,1	3,6,7	1,1,2	6,6,6
Production	3,7,7	0,0,1		0,1,2	
Export	0,1,1		0,1,1		
AUC_G	1,1,1		1,1,1	1,1,1	2,4,4
TTM	8,8,8	0,2,2			

Note: The three numbers represents the number of coefficients significant at 1%, 5% and 10% significance levels, respectively.

In general, the soybean contract models explain 50-75% of the variation of soybean basis for all the locations (see Table 3.9, the detailed estimation results are included in appendix Table B.5-7). There are also noticeable differences in significance and magnitude of the coefficients for the same variables across the soybean contracts (see Table 3.10). Opportunity (or storage) costs are found to be frequently significant in the November, January and March contract models and not significant for the July contract model. Storage costs affect the soybean basis at every market right after harvest season (November and December) and infrequently influences the basis in May through August when soybean stocks are low. Barge rate coefficients are significantly different from zero mostly in May, July and November contract models. Most of the significant production (relative to storage capacity) coefficients showed up in March and May contract models, which suggests that the size of the new crop has some impact on soybean basis from January to April.

Table 3.9. R^2 and RMSE of soybean contract basis models

Contract	Markets						
	NE Iowa	C Illinois	Chicago	St. Louis	Toledo	Gulf Port	Richmond
				Total Rsq			
January	0.79	0.77	0.75	0.78	0.73	0.73	0.51
March	0.86	0.84	0.82	0.82	0.38	0.72	0.49
May	0.60	0.83	0.83	0.77	0.66	0.14	0.55
July	0.65	0.67	0.62	0.61	0.60	0.42	0.55
September	0.75	0.66	0.45	0.40	0.51	0.46	0.56
November	0.84	0.73	0.54	0.79	0.59	0.55	0.31
				Root MSE			
January	7.99	6.66	6.29	6.55	6.28	8.33	10.19
March	7.74	5.75	4.86	5.03	7.62	6.23	7.83
May	11.11	4.57	3.11	4.82	5.50	7.61	7.07
July	9.69	7.26	5.53	8.29	5.80	10.79	9.39
September	11.48	8.52	9.96	13.44	10.49	13.27	12.98
November	7.93	8.03	11.71	6.71	9.58	10.50	11.37

Table 3.10. Number of significant coefficients with expected sign in each contract model for seven soybean markets

	Contracts					
Variables	January	March	May	July	Aug&Sep	November
	Total number of significant coefficients					
Opportunity cost	7	5	2		3	4
Transport cost	1		4	6	3	3
Production	1	4	4		2	
Export						
Crushing			2	2		1
AUC_G	1	1				
TTM	6	2	2	1		2
	Number of significant coefficient at different significance levels					
Storage cost	5,6,7	5,5,5	1,2,2		0,2,3	4,4,4
Transport cost	0,0,1		2,3,4	6,6,6	2,2,3	1,3,3
Production	0,0,1	0,1,4	2,4,4		1,2,2	
Export						
Crushing			1,2,2	1,2,2		1,1,1
AUC_G	0,0,1	0,1,1				
TTM	6,6,6	2,2,2	0,2,2	0,1,1		0,1,2

Note: The three numbers represents the number of coefficients significant at 1%, 5% and 10% significance levels, respectively.

The results are mixed for demand factors. Soybean crushing generally has coefficients with signs inconsistent with expectations. Animal units consuming grain,[18] not significant in general, exhibits both positive and negative significant coefficients. Export volume does not affect soybean basis significantly throughout the year. Time to maturity variables are significant in the January contract model for all markets (except Northeast Iowa) and also in other contract models for small number of markets.

[18] Animal units consuming protein maybe more appropriate for soybean models, but it is highly correlated with AUC_G (correlation coefficient equals 0.96) and may not make much difference in estimation.

Corn Basis Models

Opportunity costs

All of the estimated coefficients for storage costs are negative as expected (with few exceptions in Pacific NW and Richmond market models). For Northeast Iowa, the storage (opportunity) cost coefficient declines in magnitude for the contracts March, May, and July, respectively (see Table 3.11). The estimated coefficients suggest that, for a given change in the storage cost, corn basis in Northeast Iowa is affected more in early storage season (for months of December through April) than in late storage season (May through August). Assuming the corn price of 250 cents/bushel and prime interest rate of 10%, a 1 cent per bushel increase in the opportunity cost equals a 10 cents/bushel increase in corn price (given the interest rate), or a 0.4% increase in interest rate (given the cash price). For example, the opportunity cost coefficient of 0.8 for Northeast Iowa in December, January and February suggests that its corn basis will be 0.8 cent per bushel weaker with a 10 cents/bushel increase in corn price or 0.4% increase in interest rate (2 cents/bushel wider with 1% increase in interest rate). The coefficients are smaller in March and April (0.56), and in May and June (0.3). Most of the significant opportunity cost coefficients are in the range of 0.2-0.6 for other market locations. This coefficient pattern makes sense because more storage is demanded in the early part of the crop marketing year. Production may exceed storage constraints at the harvest time, and transportation service may not meet the demand. More commercial (off-farm) storage facilities will be used. Some corn or soybeans may have to be stored outside. These all lead to a larger cash price depression in the face of storage cost increase. As the marketing year progresses, there is less corn demand for storage facilities and transport services, and local cash prices are affected less by storage cost increases.

Similar patterns exist in Central Illinois model as expected since it is also a major corn producing state. Corn basis in Central Illinois for December, January, February, March, and April will be about a half cent weaker if the annual storage cost increases by a cent. The coefficient is not significant for December contract which covers the nearby basis for October and November.

Table 3.11. Estimated opportunity cost coefficients of corn contract models

Market	Contracts				
Locations	March	May	July	September	December
Northeast IA	-0.800***	-0.558***	-0.296*	-0.423*	-0.501*
Central IL	-0.502***	-0.550***	-0.280*	-0.367*	0.002
Chicago	-0.306***	-0.150	-0.042	-0.139	-0.022
St. Louis	-0.469***	-0.409***	-0.214	-0.336	0.084
Toledo	-0.385***	-0.272**	0.010	-0.162	-0.103
Gulf Ports	-0.536***	-0.383***	-0.274*	-0.384*	0.011
Pacific NW	0.217	0.303	0.319	1.387***	0.732***
Richmond	-0.266*	0.270**	0.298	-0.395	-0.172

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Storage (opportunity) cost affects Chicago corn basis only in December, January and February. In St. Louis and Toledo, the variable is significant at 5% level in March and May contract models, which cover months from December to April. Statistically, storage costs do not matter in all other months. Economically, later in the marketing year, there is less corn in the market and less corn stored in the commercial storage facilities. As storage cost increases, corn basis in these delivery markets (or terminal markets) may not be affected significantly.

In the Gulf Port model, storage cost is important for months from December to August, with decreased importance from March contract to July contract models when demand for storage is low. There is more corn shipped down the Mississippi River early in the marketing year. Again, like in the Northeast IA and Central IL market models, the coefficient is higher in the September contract model than the July contract model, and is not significant in the December contract model. For Richmond, an increase in storage cost pushes corn basis wider for months from December to February, but the basis gets narrower in months of March and April as storage cost get higher. This may be related to fact that Richmond is a corn deficit area; it may first use stocks from its own local production (a temporary surplus), then corn is shipped from major corn producing areas like the Cornbelt. This shift from surplus area to deficit area could be one possible explanation for the storage cost coefficient change from negative in March contract model to positive in May contract model. The meaning for the positive coefficients in Pacific NW market for September and December contract models is unclear.

Transportation cost

Most of the significant barge rate coefficients are negative as expected (see Table 3.12); that is, as the transportation cost increases, corn basis gets wider. Overall, the significant coefficients are mostly in July and December contract models, which indicates that barge rates affect corn basis in the months of May, June, September, October and November. The grain shipments on the Illinois waterway and the Mississippi River are relatively high in these months of the year. The Upper Mississippi River is closed due to ice in the winter months from December to February every year, and low water level or flooding sometimes causes the river to close during the summer season. That may be a partial contributor to the barge rate coefficients not being significant for some of the months. Most of the coefficients are between -0.1 and -0.2. Barge rates are quoted as percentage over bench mark tariff. As the barge rate increases one percentage point, corn basis in these months will be wider by 0.1 to 0.2 cents per bushel.

Table 3.12. Estimated barge rate coefficients of corn contract models

Market Locations	Contracts				
	March	May	July	September	December
Northeast IA	-0.003	-0.104*	-0.210***	-0.162***	-0.184***
Central IL	0.036*	-0.024	-0.107**	-0.090*	-0.145***
Chicago	0.015	-0.056	-0.170***	-0.084	-0.085***
St. Louis	0.090**	-0.019	-0.139**	-0.065	-0.186***
Toledo	0.052	-0.050	-0.188**	-0.001	-0.107***
Gulf Ports	0.214***	0.074	-0.026	0.045	-0.108***
Pacific NW	0.036	0.046	-0.163***	-0.063	-0.033
Richmond	0.058	-0.059	-0.174*	-0.031	0.018

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

For Northeast Iowa, the barge rate coefficient is significantly different from zero throughout the year, except the March contract. This is because the upper Mississippi River is closed from December to February every year due to ice.

One interesting case is the March contract model, where barge rate coefficients are significant but positive for three markets: Gulf Port, St. Louis, and Central IL at the significance level of 1%, 5% and 10%, respectively. These positive coefficients probably are related to the closed upper Mississippi River from December to February, and greater local grain (and barge) demand in the lower Mississippi River.

Production

As expected, corn production relative to storage capacity is quite important in the months right after harvest (from December to February (see Table 3.13)). The coefficient is not significant for all the other months in general. It seems that the pressure from the new crop is relieved after February, and is not a factor determining the local corn basis any more. Production/storage capacity ratio affects corn basis negatively. If the ratio of crop size to storage capacity increases by 10% (that is, the ratio by 0.1),[19] the basis will be wider by approximately 0.7 cents in most market locations in the months of December, January and February. Given the same amount of change in the production/storage capacity ratio, the basis change is 2.5 cents for Richmond, Virginia. This higher than average basis response in Richmond may due to the fact that it does not have much local production and or storage facilities,[20] so small changes may cause a larger price response. The Northern Plains corn production and grain storage capacity is used in Pacific NW market model. The coefficient

Table 3.13. Estimated corn production coefficients of corn contract models

Market	Contracts				
Locations	March	May	July	September	December
Northeast IA	-7.691**	-0.285	-3.099	-13.020**	5.692
Central IL	-7.623***	-1.092	-1.794	-6.272	0.695
Chicago	-4.862**	-4.473	-2.564	-6.253	-0.860
St. Louis	-8.745**	-1.855	-3.053	-10.218	-0.210
Toledo	-11.606***	-5.780*	-4.281	-6.800	0.886
Gulf Ports	-6.663	2.403	0.640	-15.820*	6.118
Pacific NW	23.631**	33.856*	26.811**	48.034*	23.689*
Richmond	-25.042***	-4.265	5.609	-6.258	-13.369

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

[19] A 10% increase in production to storage ratio equals 0.27, 0.26, 0.27, 0.23, 0.19, 0.53, 0.21, and 0.19 million bushels of corn production in Northeast IA, Central IL, Chicago, St. Louis, Toledo, Gulf Ports, Pacific NW and Richmond market models, respectively.

[20] Its off-farm storage capacity, around 30 million bushels, is about 3% of Iowa capacity.

for Pacific NW is unexpectedly positive and is significant at 10% significance level throughout the year.

Demand factors

The export volume and grain consuming animal units, two variables representing demand shifts, are generally not significant except in two port markets: Pacific NW and Richmond. The export volume is significant at 5 % level with the expected positive sign in March and July Pacific NW contract models. The coefficients of 2.01 and 2.17 in March and July contract models, respectively, suggest that corn basis in Pacific NW for December, January, February, May and June will be a little more than two cents narrower if the Pacific ports monthly corn exports increase by 10 million bushels. But the export coefficient in Richmond is negative and significant in May and July contract models, at 1% and 10% significance level, respectively. The negative sign is inconsistent with what is expected.

The other demand variable is the animal units consuming grain (AUC_G). Since there are no localized AUC_G statistics available, national AUC_G data are used. The animal numbers are important in corn basis determination mostly in December contract models (see Table 3.14). It suggests that if animal units consuming grain increase one million units (about 1.25%), the corn basis in September, October and November will strengthen by about 1.7 cents per bushel in Central IL, St. Louis, Gulf Ports and Pacific NW. This variable is also significant for other contract models in Pacific NW (except May). Again, Richmond models in May and September contracts have coefficients with signs inconsistent with what theory suggests.

Table 3.14. Estimated AUC_G coefficients of corn contract models

Market	Contracts				
Locations	March	May	July	September	December
Northeast IA	-0.097	-0.567	-0.356	-1.017	1.054
Central IL	0.277	-0.346	-0.103	-0.620	1.654***
Chicago	0.075	-0.047	0.105	-0.272	0.525
St. Louis	-0.324	-0.393	-0.252	-1.436	1.706**
Toledo	-0.457	-0.573	-0.351	-1.517	0.302
Gulf Ports	-0.394	-0.505	-0.445	-0.951	1.649**
Pacific NW	1.087***	1.034	1.593***	1.131*	1.651***
Richmond	-0.584	-1.066*	-1.053	-2.210*	-0.569

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Time to maturity

The last variable in the model is the time to maturity (TTM). Basis theory suggests that basis will converge as contract maturity approaches. Here time to maturity captures how basis converges in the two or three months before contract maturity. Basically only the March contract models demonstrate this convergence (see Table 3.15). It is significant at 1% level for all the markets studied. In the May contract models, time to maturity is significant at 5% level for Northeast IA and Chicago markets. These significant coefficients are negative, which means the farther from contract maturity, the wider the basis will be. There is no evidence indicating basis convergence for any of the other contracts.

Table 3.15. Estimated time to maturity coefficients of corn contract models

Market	Contracts				
Locations	March	May	July	September	December
Northeast IA	-3.422***	-3.733**	-1.207	0.261	1.225
Central IL	-2.865***	-2.092	-0.307	1.084	-0.032
Chicago	-2.531***	-3.220**	-2.905	1.657	-1.603
St. Louis	-2.893***	-1.709	-0.955	0.554	-2.184
Toledo	-3.344***	-2.418	-3.529	5.095	2.187
Gulf Ports	-3.255***	-1.507	-1.285	2.532	0.799
Pacific NW	-2.166***	-3.636	-0.540	-3.617	-1.884
Richmond	-5.484***	2.029	-2.820	9.478**	-2.030

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Soybean Basis Models

Storage (opportunity) costs

All the significant storage cost coefficients (except Richmond models) are negative (see Table 3.16), which indicate soybean basis gets wider as storage cost increases. Assuming soybean price is 600 cents/bushel and prime interest rate is 10%, a 1 cent change in storage cost equals 10 cents/bushel change in soybean price (given interest rate) or 0.167% change in interest rate (given cash price). Most of these coefficients range from 0.2 to 0.4, which suggests that soybean basis is 0.2-0.4 cents/bushel wider with a 10 cents per bushel increase in cash price or 0.167% increase in prime interest rate (the basis is weaker by 1.2-2.4 cents/bushel with 1% increase in interest rate). Soybean basis is affected by the variation in storage costs mostly from November to February for most of the markets. In the July contract model, corresponding to the months of May and June, soybean basis is not affected by changes in storage cost, possibly because there is not much storage demand during these two months.

For the Central Illinois soybean contract model, the magnitude of the storage cost coefficients gradually declines as time from harvest increases: it is -0.324, -0.280 and -0.104 for January, March and May contracts, respectively. That is, for a given amount of change in storage cost, the basis is affected more in November, December than in January and February, and basis in March and April is affected the least among these months. For the Toledo market, storage cost is only important statistically in the January contract model. Again, the Richmond contract model often shows the opposite effect of storage cost; it is significant and negative as expected for January contract, but it is significantly positive for May and Aug&Sep contract models.

Table 3.16. Estimated storage cost coefficients of soybean contract models

Market	Contract					
Locations	January	March	May	July	Aug&Sep	November
Northeast IA	-0.322***	-0.441***	-0.105	-0.043	-0.208**	-0.452***
Central IL	-0.324***	-0.280***	-0.104**	-0.096	-0.102	-0.287***
Chicago	-0.254***	-0.334***	-0.133***	-0.063	0.038	-0.161
St. Louis	-0.206**	-0.217***	-0.021	-0.032	-0.252**	-0.289***
Toledo	-0.275***	-0.050	0.018	0.027	-0.140	-0.200
Gulf Ports	-0.404***	-0.437***	-0.107	-0.028	-0.255*	-0.347***
Richmond	-0.238*	0.011	0.270***	0.040	0.242**	-0.109

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Transportation cost

Out of seven July contract models, the barge rate is significant at 1% significance level in six models, with negative coefficients as expected (see Table 3.17). That means the soybean basis for May and June in all the markets (except Gulf Ports) is affected by barge rate changes. The magnitude of the basis widening is around 0.2 cents per bushel (ranges from 0.14 to 0.35) for barge rate increases of one point in the percentage over tariff. In both May and November contract models, four out of seven markets have significant negative barge rate coefficients. Changes in barge rates affect the soybean basis more in months of May and June (July contract) than in months of September, October, March and April (November and May contracts). Few markets have significant barge rate coefficients for Aug&Sep, March and January contract models. The river is closed for some of the months covered by these contracts: ice in December to February for upper Mississippi River every year, or infrequently due to flooding or low water level in summer months. Similar to the corn contract models, the barge rates are significantly positive in March contract models for St. Louis and Gulf Ports. The closed upper Mississippi River may cause this sign shift.

Table 3.17. Estimated barge rate coefficients of soybean contract models

Market Locations	January	March	May	July	Aug&Sep	November
	Contract					
Northeast IA	-0.111*	0.093	-0.168***	-0.347***	-0.262***	-0.141***
Central IL	0.007	0.027	-0.047	-0.199***	-0.012	-0.062
Chicago	-0.037	0.038	-0.054**	-0.144***	-0.140***	-0.131**
St. Louis	-0.059	0.096**	-0.056	-0.269***	-0.119*	-0.092**
Toledo	0.025	-0.012	-0.103*	-0.183***	-0.068	-0.058
Gulf Ports	0.089	0.268***	0.047	-0.171	0.078	0.026***
Richmond	0.013	-0.002	-0.178***	-0.202***	-0.032	-0.058

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Production

The ratio of soybean production relative to storage capacity is significant mostly in contracts of March and May, which cover January to April. The coefficients show negative relationships between soybean production and soybean basis. Basis will be wider by 10 cents per bushel for Northeast Iowa in January and February if the regional production/capacity ratio increases by 10%, which equals about 0.27 million bushels of soybean.[21] For these same two months, the basis change is about four to five cents for St. Louis and Toledo for an 10% increase in soybean production/capacity ratio (see Table 3.18). For the months of March and April, the same increase in the ratio will cause soybean basis to be weaker by five to seven cents per bushel for Central IL, St. Louis, Toledo and Richmond.

Demand factor

Demand factors in the soybean contract model include export volume by ports, animal units consuming grain, and soybean crushing. Export volume is not significantly different

[21] An 10% increase of production/storage ratio equals 0.27, 0.26, 0.27, 0.23, 0.19, 0.53, and 0.19 million bushels of soybean production in Northeast IA, Central IL, Chicago, St. Louis, Toledo, Gulf Ports, and Richmond market models, respectively.

Table 3.18. Estimated soybean production coefficients of soybean contract models

Market Locations	Contract January	March	May	July	Aug&Sep	November
Northeast IA	-42.725	-98.589**	-24.042	5.011	-6.676	15.135
Central IL	-3.233	-18.247	-63.969***	-3.982	-97.321***	-10.051
Chicago	2.325	7.469	1.859	10.970	-24.556	1.574
St. Louis	-2.871	-44.167*	-74.138***	-16.232	48.343	11.692
Toledo	-19.019	-46.500*	-61.905**	-19.898	-44.418	-0.544
Gulf Ports	12.699	-24.699	-11.462	-16.421	-60.165	17.339
Richmond	-72.564*	-71.184*	-52.791**	43.517	-120.306**	35.951

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

from zero throughout the year for all seven markets. Animal units consuming grain are significant only in few contract models at the significance level at 5% (three coefficients) and 10% (eight coefficients) (see Table 3.19). They are positive in Illinois model and negative for Chicago, St. Louis and Toledo models. As a demand factor, a positive coefficient is expected. Since the animal unit data are a U.S. totals, a state or regional animal number might be more appropriate to represent local soybean demand.

Most of the soybean crushing coefficients are not significant, and the significant coefficients are usually at 5% or 10% significance levels. The higher the demand from soybean crushing is, it is expected that cash soybean price will be higher. However, this

Table 3.19. Estimated AUC_G coefficients of soybean contract models

Market Locations	Contract January	March	May	July	Aug&Sep	November
Northeast IA	0.084	1.091	1.274	0.719	-0.527	0.077
Central IL	0.520	0.863*	0.954**	0.816	-0.740	0.164
Chicago	-0.399	-0.789*	-0.261	-0.210	1.029	0.720
St. Louis	-0.505	-0.499	0.297	0.077	-1.377*	-0.533
Toledo	-0.546	-1.111	-1.456*	-1.112**	-1.648*	-4.309**
Gulf Ports	0.521	0.529	1.077	1.132	-0.909	0.837
Richmond	0.019	-0.684	-1.149	-1.466	0.190	-1.760

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

positive relationship is only confirmed in Toledo and Richmond market models (see Table 3.20). The estimated coefficients of soybean crushing for all the other markets show that local soybean basis has a negative relationship with soybean crushing. One possible explanation is that most processors buy large portion of their annual soybean purchases during the harvest time. Therefore monthly crushing data may not reflect impact of the level of current demand for soybeans from crushers. This may also due to the inadequate local soybean crushing data used in some models. Because of its availability and completeness, Illinois soybean crushing is used as an index of crush volumes in most of the market models, including Northeast Iowa, Central IL, Chicago and St. Louis. Gulf Ports model uses total U.S. soybean crushing. The Ohio soybean crush data is used in the other two markets: Toledo and Richmond. Alternatively, lower cash prices may stimulate more crush, so there may be a simultaneous equation problem here that has not been adequately modeled.

Time to maturity

As a measure of basis convergence, time to maturity shows up to be important in six out of seven January soybean contract models (the exception is Northeast IA, see Table 3.21).

Table3.20. Estimated soybean crushing coefficients of soybean contract models

Market	Contract					
Locations	January	March	May	July	Aug&Sep	November
Northeast IA	-0.026	-0.073*	-0.111**	-0.090**	-0.003	-0.047*
Central IL	-0.022	-0.087***	-0.032	-0.056	0.002	-0.053*
Chicago	-0.047	-0.057**	-0.025	-0.032	-0.006	-0.042
St. Louis	-0.022	-0.072**	-0.016	-0.056	-0.010	-0.054**
Toledo	-0.027	0.142	0.192**	0.137**	0.068	0.400***
Gulf Ports	-0.013	-0.027***	-0.016	-0.003	0.007	-0.017
Richmond	0.041	0.131	0.363***	0.268***	0.008	0.092

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

These are also two significant time to maturity coefficients in March, May and November contract models, and one in the July contract model. All the significant coefficients have expected negative signs--basis gets narrower as contract maturity approaches. Other things being equal, soybean basis in all the markets except Northeast IA converges by 7 to 12 cents from November to December. Soybean basis from January to February in Toledo and Richmond converges by 9.5 and 11 cents, respectively. The convergence is about 3 to 5 cents for the May and July contracts. Chicago and St. Louis soybean basis also indicate significant basis convergence from September to October by 13 cents and 7.6 cents, respectively.

Table 3.21. Estimated time to maturity coefficients of soybean contract models

Market	Contract					
Locations	January	March	May	July	Aug&Sep	November
Northeast IA	-3.051	-3.915	-1.350	-1.125	6.145	-6.140
Central IL	-7.689***	-1.115	-1.534	-1.121	1.884	-7.029
Chicago	-8.619***	-3.686	-2.930**	-2.197	-0.543	-13.164**
St. Louis	-9.975***	-1.165	-1.705	-1.158	6.131	-7.665*
Toledo	-7.919***	-9.468***	-4.832**	-4.835**	1.689	8.385
Gulf Ports	-12.160***	3.085	1.678	-3.089	4.588	-6.679
Richmond	-11.852***	-11.088***	-3.422	1.071	1.167	-14.571

Note: *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Alternate Model Specifications

Though the tests of structural change are not conclusive, we attempt to capture the possible change in structure via alternate model specifications. We first incorporated into the basis model a trend variable that had an upward trend before 1986, and flat thereafter. The estimated coefficient is significant, but it results in more insignificant coefficients for other independent variables. Secondly, separate models are estimated for these two periods for each contract at all locations. There are only few significant coefficients due to the small sample size (6 and 10 years observations for each period, respectively).

Finally, intercept and slope dummy variables are employed for all variables for the post 1986 period. The nested F tests show that set of dummy variables are significant in most contracts (at 5% significance level, the null hypothesis of the set of dummy variables is equal to zero is rejected in 35 out of 40 contract models for corn, and is rejected only in 18 out of 42 coybean contract models). Tables 3.21 and 3.22 show the numbers of significant coefficients with expected sign for different model specifications (for 8 corn markets and 7 soybean markets). Model specification (1) is the original structural basis behavior model discussed in the previous sections. Specifications (2) and (3) are alternate specifications, where (2) is the structural model with intercept and slope dummies for the period after 1986, and (3) differs from (2) only in one variable--the total U.S. corn and soybean supply[22] relative

Table 3.21. Number of significant coefficients with expected sign for three corn model specifications

	Model	Contract				
Variables	Specifications	March	May	July	September	December
Opportunity Cost						
	1	7	5	3	3	1
	2	4	3	2	3	1
	3	6	3		4	2
Barge rate						
	1		1	7	2	6
	2		1	4	2	6
	3			2	2	4
Production						
	1	7	1		2	
	2	4	3	2	2	
	3	4	1		2	4

[22] Production plus beginning stock in the harvest period and beginning stock for other quarters.

Table 3.22. Number of significant coefficients with expected sign for three soybean model specifications

Variables	Model specifications	Contract January	March	May	July	September	November
Opportunity Cost							
	1	7	5	2		3	4
	2	7	4	2	1	1	1
	3	7	5	1		1	
Barge rate							
	1	1		4	6	3	3
	2				1	1	1
	3				4	1	3
Production							
	1	1	4	4		2	
	2	1	2	1	1		
	3	1	5	3	2	1	

to U.S. storage capacity is used instead of regional corn or soybean production relative to regional storage capacity.

In general, the results of three model specifications are consistent with each other. The general pattern does not change. For example, opportunity cost is mostly significant in the early storage season and barge rate is important in July and December contract models. The size of the estimated coefficients are also similar, on average. The number of significant coefficients with expected signs do not increase with these alternate specifications. The original structural basis behavior results (1) presented earlier have the most significant negative opportunity cost and barge rate coefficients for both corn and soybean models. Specification 3 using total U.S. corn and soybean supply relative to storage capacity (model 3) seems more significant than regional production relative to supply (model 2).

Though the general pattern is confirmed, the coefficients in some contracts and some locations, after incorporating the slope dummy for the second period, change signs after 1986 (e.g. the expected negative coefficient of opportunity cost switches to positive after 1986), especially for storage cost variables. Barge rate coefficients are similar for both periods. With model specification (3), the sign changes occur less frequently, and the magnitude of some coefficients change less.

The results of these alternative model estimates do not seem to be appreciably better than the original structural basis behavior model (1) discussed earlier. Further effort to identify more appropriate variables to be used in basis behavior model for the period after 1986 may still be worthwhile. A good index for variables such as government program changes and or rail rate deregulation may improve the structural model.

CHAPTER 4. BASIS FORECASTING MODELS

Several studies on basis forecasting (Heifner 1966, Kenyon and Kingsley 1973, Taylor and Tomek 1984, Hauser et al. 1990, Tomek 1996, and Strobl et al. 1996) have been done on corn and soybeans. However, basis forecasting was not the main subject of some of these papers; it was treated as a means to improve hedge effectiveness (Kenyon and Kingsley 1973, and Hauser et al. 1990) or storage decisions (Heifner 1966). Several forecast techniques were utilized in these studies: a basis change model[23] (Heifner 1966, and Kenyon and Kingsley 1973), which considers that basis changes over a time interval as a function of initial basis; a structural econometric forecasting model (Taylor and Tomek 1984, and Strobl et al. 1996); seasonal ARIMA technique (Strobl et al. 1996); and a simple naive forecast models (Hauser et al. 1990). The naive models usually forecast basis as a function of the previous year basis, the average of last three years basis, the price spread between the two nearest futures contracts, or the time to expiration of the futures contract (Hauser et al. 1990).

Forecasting Grain Basis

This paper studies several of these standard forecasting models, and some additional forecast methods such as state space modeling approach and artificial neural networks. Since the naive model based on the previous three year average basis is often used in the grain trade,

[23] This model, also called basis convergence model, is not used in this forecasting exercise as it is more appropriate to use in forecasting longer term changes in basis versus a single contract, rather than when periodic shifts to another contract are considered, as in this study.

it will serve as the standard for comparison for the more sophisticated forecast models outlined below. A composite forecast is also constructed as a simple average of the other model forecasts.

Three-Year-Average Forecasts

This benchmark model calculates basis in any month t (BS_t) as a function of the average of previous three years basis for that month:

$$BS_t = (1/3)(BS_{t-12} + BS_{t-24} + BS_{t-36}).$$

In a regression framework, this is equivalent to:

$$BS_t = \alpha + \beta*3yr\_avg_t + u_t, \text{ with } \alpha=0 \text{ and } \beta=1$$

where $3yr\_avg_t$ is the simple average of previous three years basis at month t. This regression model has a poor fit. Frequently, we observe low R^2 and beta coefficients far different from one. A nested F test is performed for the hypothesis of $\alpha=0$ and $\beta=1$, which will be discussed in the performance section later.

Three-Year-Average-Plus Forecasting Model

If you assume that basis is stable over the years, the last three years' average basis for a particular month will be a good forecast of basis for the same month this year. Though it is easy to use, this approach does not take into account any current market information. In addition to previous basis, current corn and soybean basis could be affected by current corn and soybean market supply and demand conditions such as production and exports, especially when this year differs a lot from previous years. If a larger than usual crop is harvested this year, other things being constant, the basis will be weaker than usual; and a higher consumption or export demand this year will result in a smaller basis than previous years. Therefore, it is hypothesized that the naive model forecasts could be improved if current market information is added to that model. That is a three-year-average plus model:[24]

$$\text{(EQ 4.1)} \qquad BS_t = f(3yr\_avg_t, Total\ Supply_t,\ Exports_t,\ DM_i)\ , \quad i = 1, \ldots, 12$$

where 3yr_avg is the average basis of previous three year at the same month, DM_i are monthly dummies variables from January through December (except October), Total Supply and Exports are USDA supply and export estimates.

Basis forecasts for a month are calculated with the estimated coefficients, projections of total supply and export for the next crop year available in October (USDA WASDE predictions), and the previous three years' average basis for that month. All the 1-12 ahead basis forecasts are calculated by that procedure.

Fundamental Forecast Model

The traditional approach is to forecast the independent variables individually and then insert the forecasted independent variables into the estimated model to derive the basis forecast. EQ 4.2 is an example for Northeast Iowa corn basis forecasting model:

$$\text{(EQ 4.2)} \qquad BS\_nia_j^f = \alpha_j + \beta_{1j}{*}RP\_nia_j^f + \beta_{2j}{*}BR\_nia_j^f + \beta_{3j}{*}CN\_nia_j^f + \beta_{4j}{*}EX\_nia_j^f + \beta_{5j}{*}AUC\_G_j^f + \beta_{6j}{*}TTM_j^f \qquad j = \text{Mar., May, Jul., Sep., and Dec.}$$

[24] The October basis, as an additional variable, is also included to check whether the October basis, the most recent basis available when making forecast in the end of October, has predictive power for November up to October next year basis forecasts. The joint F test shows that it is not significant for 1-12 month ahead forecasts of both corn and soybean basis.

where *i* represents market, the superscript f stand for forecasted value, and the Greek characters are estimated coefficients from the behavior model. The independent variables with superscript f are the ancillary forecasts which usually introduce additional errors into the forecasts.

The estimated coefficients are obtained from fundamental basis behavior model in chapter 3.[25] The ancillary forecasts have to be obtained before calculating basis forecasts. The following table (Table 4.1) outlines the methods used to get the ancillary forecasts of one to 12 month ahead for these relevant variables. Those ancillary forecasts, together with the estimated coefficients, are used to derive the 1-12 month ahead basis forecasts.

Seasonal ARIMA Model

Though the fundamental approach to basis forecasting has appeal, ARIMA time series methods may produce better forecasts. Hauser et al. (1990) set up several simple ARIMA basis expectation models and compared them to four simple regression models. They found that simple moving average models provided better forecasts than the simple regression models. A seasonal ARIMA (SARIMA) was used by Strobl et al. (1996) in their study of forecasting local soybean basis in North Carolina.

Theoretically, time series techniques have some advantages over econometric modeling. Jenkins (1979) suggested econometric models usually rely on static least squares methods, but economic phenomena are dynamic in nature. Others, such as Strobl et al.

[25] The behavior model is estimated by OLS, though autocorrelation exists. This is because OLS estimates with autocorrelated disturbances are still unbiased and consistent, therefore the forecasts based on these unbiased coefficients are not affected.

(1996), cited practical reasons for preferring time series forecasting methods over econometrics forecasting model. They argued that econometrics models can be very difficult for market agents to employ because of the data requirements and statistical skills necessary for implementation. In addition, econometric models involve simultaneously determined independent variables and basis. The forecasting of these independent variables introduces additional sources of error, therefore reducing the accuracy of the basis forecast.

Table 4.1. Methods of ancillary forecasting

Variables	Methods used to obtain ancillary forecasts
Prime interest rate	Naive model (equals t-1 PIR)
Cash prices	Previous three year average
Barge rates	ARIMA model
Production	Regression model and USDA WASDE (the relationship between state and US productions is identified by simple regression, then the coefficients from the regression and WASDE of US production are used to calculate the state corn and soybean forecasts)[a]
Storage capacity	Naive model (equals to last year's capacity)
Soybean crushing	Previous three year average
Animal units consuming grain	Uses the animal units consuming grain of last crop year
Exports by port	Previous three year average

[a] USDA state production estimates are also considered. These state production estimates and the estimates obtained through this regression approach are highly correlated and are compared in their out-of-sample forecasting performance. The estimates from the regression approach perform little better in the out-of-sample forecast performance.

Standard Box-Jenkins time series methods are used in this study to forecast the spatial grain basis. The appropriate models depend on the stationarity of the time series and whether any differencing is needed to achieve stationarity. Model identification is based on the examination of the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. The general form of a SARIMA(p,d,q)(P,D,Q)$_s$ model is:

(EQ 4.3) $$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B)^D BS_t = \Theta_Q(B^s)\theta_q(B)a_t$$

where B is the backshift notation, s is the seasonal period (observations per period, here it is 12 months), the (p,d,q) are orders of autoregressive, differencing and moving average terms, respectively, and (P,D,Q) are the orders of seasonal counterparts corresponding to (p,d,q). The Greek letters, ϕ, θ, Φ, and Θ, are the coefficients for regular autoregressive terms, moving average terms, seasonal regular autoregressive terms, and seasonal moving average terms, respectively.

A practical three-stage procedure is proposed by Box and Jenkins (1976) to find an adequate model: identification, estimation and diagnostic checking. This procedure can be iterative until an statistically adequate model is reached. First of all, at the identification stage, the visual inspection of a plot of the original realization is an important preliminary. This is mainly focused on the stationarity of both the mean and variance of the realization. As a practical rule (Pankratz, 1991), the number of useful estimated autocorrelations is about n/4. Since the five data sets have 130, 142, 154, 166 and 178 observations, respectively, the number of estimated autocorrelations will be 32, 35, 38, 41 and 44 for each of the corresponding in-sample data set. Since the model identification is based on the properties of sample data set and each of the five in-sample data sets cut off at different points, this may result in a different model for each in-sample data set studied for the same market.

There is a seasonal pattern in the prices of many agricultural commodities. Therefore, all the models we are going to consider are mixed models consisting of both seasonal and non-seasonal elements. They are multiplicative models. The ACF and PACF of all the realizations are examined first; if the ACF does not drop quickly to zero, then first order differencing is done to achieve a stationary series. The first order differencing can also help in identifying seasonal patterns even if the original series are stationary. The practical warning levels for absolute t-values in estimated ACFs (suggested by Pankratz) are used in the identification and diagnostic-checking stages. These warning levels are listed in Table 4.2.

The basis series is monthly data, therefore the seasonal period is 12. Seasonal differencing is decided in the same fashion as non-seasonal difference. The only difference is that ACFs are examine at lags of 12, 24, 36, ... for seasonal stationarity instead of lags at 1, 2, 3,... for non-seasonal stationarity. If the ACFs drop quickly to zero at these seasonal lags, then it is seasonal stationary; otherwise, seasonal differencing is used.

Table 4.2. Warning levels for ARIMA modeling

acf lag	Identification Stage (initial estimated acf)	Diagnostic-checking Stage (residual acf)
Short (1, 2, perhaps 3)	1.60	1.25
Seasonal (s, 2s, ...)	1.25	1.25
Near-Seasonal (s-1, s+1, 2s-1, 2s+1,...)	-	1.25
and half season (.05s, 1.5s. ...)	1.60	1.60

Note: the numbers in the table are t-values.

After stationary series are identified with appropriate regular and/or seasonal differencing, tentative multiplicative models are selected based on the autocorrelation function (ACF) and partial autocorrelation function (PACF). As a rule of thumb, if ACFs cut off and PACFs decay, it suggests a moving average model; if ACFs die down slowly and PACFs cut off, then it indicates an autoregressive model. This rule is also used to identify the seasonal terms, but we look at the seasonal lags' ACFs and PACFs.

These tentative models are estimated. Stationarity and invertibility conditions have to be met by the estimated models. Residual ACF plots show whether the selected model is adequate or not. If no residual ACFs exceed the warning levels listed in Table 4.2, that means random shocks are independent. Modified Box-Pierce (Ljung-Box-Pierce) statistics can also be used to test the residual autocorrelation. On the other hand, if residual ACFs exceed the warning levels, the residual ACFs may indicate how the model could be improved; otherwise the three stage circle of identification, estimation and diagnostic checking has to be repeated until an adequate model is found. The final satisfactory model is used for forecasting one to 12 months ahead forecasts. An example of the SARIMA modeling process is included in appendix C.

State Space Model

Mehra (1982) stated that the state space modeling technique is closely related to the Box-Jenkins methodology, but with different criteria for order selection and parameter estimation. "State space approach uses objective information criteria for order determination, whereas the Box-Jenkins approach relies on subjective judgment on the part of the user" (p. 279) There are several important differences in the multivariate case. "The state space approach considers all time series simultaneously and develops a minimal canonical representation for the system"(p. 279) The ARIMA model is not canonical without special parameter restrictions. The order of the state space model will be generally less than the modeling approach recommended by Jenkins (1979) in which univariate models are developed for each series and then a multivariate model is used for residuals. In a forecasting tournament[26] of developing univariate models for four time series, the state space out-performed other time series methods such as Box-Jenkins, AEP, and Adaptive Filtering. It also had lower order.

The terminology of state space modeling and filtering was originally developed by control engineers. It has been successfully applied in physical science in diverse fields. Its applications in economic times series were started with Akaike's (1976) canonical correlation method and Mehra's (1982) study. More recently, Aoki (1987), Aoki and Havennner (1991), and others proposed an alternative method in estimating state space models. Diebold reviewed both approaches in his review essay. He termed the Aoki's method as "top down" approach in which the state space representation is obtained directly from the data. Other analyses were characterized as "bottom up" approaches, which means a model is first specified and then converted to state space form. In their empirical studies, Koehler and Murphree (1988), McCarthy and Najand (1993), and Patterson (1994, 1995) used Akaike's approach. The studies by Mittnik (1990a, 1990b), and Vinod and Basu (1995) used Aoki's method.

In this paper, the approach proposed by Akaike (1976) will be used. One reason for this selection is that a procedure called STATESPACE, which follows Akaike's method, is readily available in the SAS program. This approach utilizes Kalman filters to compute the optimal estimates. It also allows for the maximum likelihood estimation of the unknown parameters in the model, which is done by prediction error decomposition.

---

[26] More details are in Granger, C.W.J. and G. McCollister (1979).

The general state space form (SSF)[27] applies to a multivariate time series Y_t.

(EQ 4.4) $Y_t = G_t X_t + W_t + d_t$ (Observation or measurement equation)
$X_{t+1} = F_t X_t + V_t + c_t$ (State or transition equation)

where all the variables are in matrices or vector forms. Y_t, W_t and d_t are w x 1 vectors; X_t, V_t and c_t are v x 1 vectors; G_t is a matrix with dimension of w x v and F_t is v x v. W_t and V_t have means of zero and variances of R_t and Q_t, respectively, and covariance of S_t, W_t, d_t, V_t, c_t, R_t, Q_t and S_t are system matrices. If these matrices do not depend on time, the model is said to be time-invariant or time-homogeneous.

As Harvey (1989) put it, the aim of the state space formulation is to set up X_t, the state vector, in such a way that it contains all the relevant information on the system at time t and that it does so by having as small a number of elements as possible. A state space form which minimizes the length of the state vector is said to be minimal realization. Once a model has been put in a state space form, a number of algorithms can be applied. The Kalman filter is the dominant one used in state space modeling. The Kalman filter is a recursive procedure to compute the optimal estimate of the state vector at time t, based on the information available at time t. One advantage of Kalman filter is that it enables the likelihood function to be calculated when the disturbances and initial state vector are normally distributed; this is done via prediction error decomposition. In the Gaussian case, the resulting estimate is the best mean square error prediction. It is the best linear prediction in non-Gaussian case.

One critical point in estimation is the initial state vector. It has to be assumed to be

[27] This section is primarily based on Harvey's discussion of state space modeling.

certain values. The estimator $X_{t|k}$ of X_t is obtained by projecting each component of X_t onto all components of Y_o, Y_1, ..., Y_k. The covariance matrix of the estimation error is $W_{t|k} = E(X_t - X_{t|k})(X_t - X_{t|k})'$. The purpose is to compute $X_{t|t-1} = X^*_t$, and $W_{t|t-1} = W_t$.

Estimators are obtained in the following prediction equations:

$$X^*_{t+1} = F_t\, X_{t|t} + c_t$$

$$W_{t+1} = F_t\, W_{t|t}\, F_t' + Q_t$$

where $X_{t|t}$ and $W_{t|t}$ are derived from the filter equations:

$$X_{t|t} = X^*_t + W_t\, G_t' D_t^{-1} I_t$$

$$W_{t|t} = W_t + W_t\, G_t' D_t^{-1} G_t\, W_t$$

where I_t and D_t are derived from the innovation equations:

$$I_t = Y_t - G_t\, X^*_t - d_t$$

$$D_t = G_t\, W_{t|t} G_t' + R_t.$$

The starting values for this process may be specified as X_o and W_o. Each new observation is incorporated into the process in the innovation equations, and an optimal estimator is obtained. When all the observations have been processed, the filter yields the optimal estimator of the current state vector.

The state space model used in the SAS procedure is the following model:

(EQ 4.5)

$$Y_t = G_t\, X_t \quad \text{(Observation or measurement equation)}$$

$$X_{t+1} = F_t\, X_t + K_t V_{t+1} \quad \text{(State or transition equation)}$$

where F and K are the transition and input matrices, respectively. The state vector X_t will be determined automatically within the estimation, with no subjective judgment needed. As stated in the SAS help menu, "the procedure fits a sequence of vector autoregressive models using Yule-Walker equations and selects the order for which Akaike's information criterion is minimized. This order is then taken as the number of lags into the past to use in a canonical correlation analysis" (p.774). In this analysis, we estimate both an univariate and two multivariate state space models for each basis series. In the multivariate framework, we group basis by locations so that both corn and soybean basis at the same location are fitted in the same model. In addition, a multivariate state space model including all seven corn market equations[28] and seven soybean market equations is estimated and used to forecast basis.

Because STATESPACE analyzes and forecasts stationary multivariate time series data or data that can be made stationary after differencing, it is necessary to first identify the order of differencing necessary to achieve stationarity for each basis series. These same differencing for both regular and seasonal orders in the SARIMA model is used here. Forecasts for one to 12-ahead forecasts are made based on the estimates.

Artificial Neural Networks (ANN or NN)

Since the pioneering efforts in the 1940s, researchers from many scientific disciplines have designed artificial neural networks to solve a variety of problems in pattern recognition, prediction, optimization, associative memory and control, etc. Neural networks are nonlinear mapping structures based on the study of human brain, because human brains have many desirable characteristics not present in Von Neumann or modern parallel computers. The

[28] Pacific Northwest market equation is excluded because it has 48 observations less than all other markets. Including it causes problems in estimating this multivariate state space model.

brain consists of 10^{11} processing units of different types called neurons. Neurons are massively connected, each neuron is connected to 10^3 to 10^4 other neurons. A neuron (or nerve cell) is a special biological cell that processes information, receiving information from other neurons and sending its outputs to other neurons after processing the signals. Neural networks try to model this complex neuron structure to solve real world problems.

Neural network consists of a collection of inputs, processing and output neurons. The connections between neurons are directional and with weights. Based on the connection pattern (architecture), neural networks can be grouped into two categories: feed-forward networks and recurrent (bi-directional or feedback) networks. There are various networks under each of these two categories. Feed-forward networks include single layer perceptron, multilayer perceptron, and radial basis function nets (RBF); the recurrent networks consist of competitive networks, Kohonen's self-organizing maps (SOM), Hopfield network, and adaptive resonance theory (APT) models.

Feed forward networks are most popular; in this class neurons are organized into layers that have unidirectional connections between them. There are three types of layers: input, hidden and output layers. The number of hidden layers and hidden neurons are arbitrarily selected by users. Usually one hidden layer is sufficient for most time series forecasting models and there is no theoretically supported rule for selecting the correct number of hidden neurons (some suggest 75% of input neurons). Figure 4.1 shows a typical two-layer feed forward network (bias inputs are not shown).

In this network, each neuron sums its weighted input, with an activation (transfer) function for mapping the inputs to output:

(EQ 4.6) $O_j = f(SW_{ij} X_i)$

where f (.) is the activation function, i and j are weights associated with input i and neuron j, respectively. O, W, and X represent output, weight and input, respectively. The activation function usually is nonlinear. It can take various forms, but the most frequently used is the sigmoid function. Others also used include tanh, Gaussian (pdf), threshold, piecewise linear, and so on. The sigmoid transfer function is as follows:

(EQ 4.7) $$f(x) = \frac{1}{1 + e^{(-x/T)}}$$

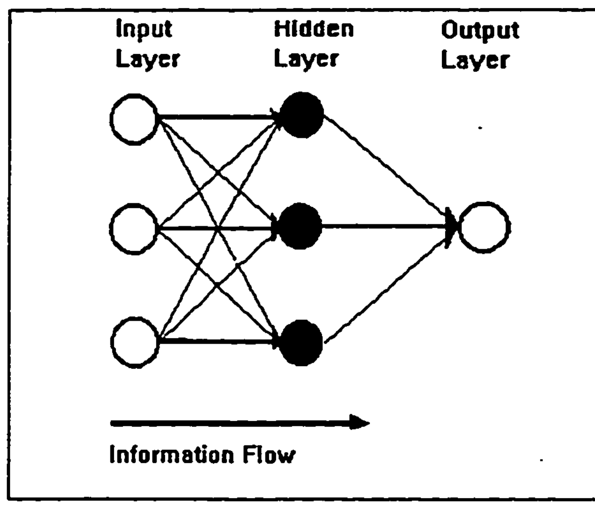

Figure 4.1 A feed forward network

One property that artificial neural network has that makes it attractive and exciting is its learning ability. The learning process is a process of adjusting connection weights so that a network can efficiently perform a specific task. There are three main learning paradigms which specify the procedure of how learning is used to adjust the weights--supervised, unsupervised, and hybrid. In supervised learning, output (a correct answer) is provided for every input pattern in the training data set. Under unsupervised learning, no correct answer is provided. The hybrid is the combination of the above two paradigms. There are different learning rules in updating the weights: error correction, Boltzmann, Hebbian, competitive learning and so on. The error correction learning is basically to reduce the error (the difference between the network produced output and desired output) by modifying the connection weights. The perceptron learning algorithm and back propagation algorithm are based on this principle.

Each learning algorithm is designed for training a specific architecture, though some learning algorithms can be used in several types of network architectures. Each algorithm can perform few tasks well. According to Jain and Mao (1996), error-correction learning rules in the feed-forward perceptron and RBF network, with supervised learning or hybrid, are suitable for prediction. In supporting this view, Kohzadi et al. (1995) stated that back propagation feed-forward networks with supervised learning rules are the most popular and useful for time series forecasting.

ANN Studies in Forecasting

Artificial neural networks have been applied in various scientific fields with success. However, it is not until recently that it began to be applied in financial and economic studies. In order to make statisticians and economists more comfortable with artificial neural networks, ANN models were compared to traditional statistical models. Azoff (1994) stated that neural networks "can be considered as a 'multivariate nonlinear nonparametric inference technique that is data driven and model free'" (p. 1). Cheng and Titterington (1994) reviewed the neural networks from a statistical perspective and pointed out some of the links of ANN with statistical methodology. Perceptrons were shown to have strong associations with discriminant analysis and regression, and unsupervised networks with cluster analysis. Some other statistical procedures can also be given a neural network expression. Kohzadi et al. (1995) stated that regression models may be viewed as a feed-forward network with no hidden layers and linear transform functions in the output neurons, and networks with hidden layers resemble nonlinear regression models. They noticed that several studies showed that feed-forward networks with at least one hidden layer belong to a class of flexible functional forms that do not make any assumptions about the distribution of variables concerned or the underlying functional form in the data generating process. In regression models, certain assumptions regarding the distribution of error terms and functional forms must be made, while in the neural networks model, these assumptions are not required.

Hill et al. (1994) listed some advantages of the artificial neural network model over statistical methods: ANN can be mathematically shown to be universal function approximators; ANN can capture the nonlinearity in time series since the network is nonlinear in nature; and the estimation of a ANN model can be automated, whereas the estimation of many kinds of statistical time series models requires human interaction and evaluation. The disadvantages of the ANNs include difficulty in interpreting the model,[29] potentials for overfitting because of more parameters to estimate, and time-consuming computation (Kahzadi et al. 1995, Gorr 1994, and Hill et al. 1994). The architecture also lacks explanation and complete theory, the software lags behind developments in the field, and requires a large number of observations.

Artificial neural networks' applications in financial and economics studies are mostly forecasting models. Kohzadi et al. (1995) argued that the neural networks have been shown to be universal and highly flexible function approximators to any data-generating process. Therefore, they are powerful models for forecasting purpose, especially when the underlying

[29] Hypothesis testing for individual coefficients is not possible as in regression models, that's one reason why it is more useful in forecasting than in policy analysis.

data generating process are unknown. Gorr et al. (1994) stated ANNs are appropriate for complex phenomena for which we have good measures but a poor understanding of the relationships within these phenomena. Because of the self-adaptive, automatic modeling properties, ANNs are ideally suited for predicting and forecasting. Gorr (1994) further argued that ANNs are more appropriate for multivariate extrapolation cases with high data collection rates (shorter time intervals), while for univariate forecasting the tough competition from conventional time series methods may result in ANNs not making a major improvement.

Most of the empirical studies using neural networks in forecasting compare their performance with that of traditional statistical methods. Empirically Hill et al. (1994) compared ANN with several time series forecasting models[30] using 111 data series from the well known "M-Competition". The ANN model was found to be significantly better than statistical and human judgment methods in the monthly time series. They found that ANN has comparable performance with regression models. Comparisons were also made with other models such as logistic regression, discriminant analysis and ARCH models. In addition, the authors noticed from other studies and confirmed in their own study that ANN was better at forecasting monthly and quarterly than annual series, and ANN was superior in the later periods of the forecast horizon. Kohzadi et al. (1995) forecasted corn futures prices with ANN and ARIMA and found ANN outperformed ARIMA by the criterion of both mean absolute percentage error and mean square error. Dasgupta et al. (1994), using two real-world individual-level cross-sectional data sets relating to the marketing of financial services,

[30] The statistical time series models include Box-Jenkins methods, several exponential smoothing methods and a filter method.

compared the performance of ANN model with two statistical market response models--logistic regression and discriminant analysis models. They found that "ANN model performs better than the market response models, however, this superiority is not statistically significant based on a Chi-Square test for equality of proportions." There are more neural networks applications in forecasting financial and economic time series. Kuan and Liu (1995) used feed-forward and recurrent neural networks in forecasting exchanges rates, Kaastra and Boyd (1995) forecasted futures trading volume, Grundnitski and Osburn (1993) forecasted S&P and gold futures prices, Uhrig et al. (1992) applied neural networks in predicting corn yields, Hamm et al. (1993) built a futures trading model which uses a neural network to produce trading signals, Claussen and Uhrig (1994) used neural networks to predict the directional movements in the central Illinois cash soybean price, and Kohzadi et al. (1994) compared the performance of neural network models with traditional ARIMA models in forecasting commodity prices (monthly U.S. cattle prices).

The ANN could be an appropriate model to forecast local grain basis. The true model generating the observed price series is uncertain. The data-driven property of the ANN could find the hidden patterns in the series. The interactions among several markets across the country and between cash and futures markets are probably nonlinear in nature, which ANN can handle with its nonlinear mapping. As Kohzadi et al. (1995) stated, many price series were found to be non-random and nonlinear, the opposite of what traditional expectations are.

A feed-forward neural network architecture with supervised learning is chosen in this study. A standard Scaled Conjugate Gradient and Noise-Feedback Descent training algorithm (a new algorithm developed by Craig Carmichael, Department of Mechnical Engineering, Iowa State University) are selected as the error correction method in adjusting the weights. The data set is divided into three subsets: training, testing and forecasting sets, the first two are in-sample and the last one is out-sample sets. Artificial neural network estimation and forecasting is in close cooperation with Craig Carmichael and Dr. Eric Bartlett, Department of Mechanical Engineering, Iowa State University. A neural network tool kit developed by their Adaptive Computing Laboratory is used in analyzing and forecasting corn and soybean basis, and the estimation and forecasting is done in the Laboratory by Craig Carmichael.

A neural network model has to be trained separately for each forecast, so there are 12 models for each of five sets of 1-12 month ahead forecasts, with a total of 60 models for one crop and one market location. Each forecast is generated by the standard Scaled Conjugate Gradient and Noise-Feedback Descent training algorithms. Therefore, eight corn market locations and seven soybean market locations need 960 and 840 neural network models for corn and soybean basis forecasts, respectively. In each neural network model, there are more than 100 variables (e.g. 8 barge rates; 26 corn or soybean production, stock, and export variables; 19 export volumes by ports; 45 storage capacity variables; 9 measures of animal units consuming grain or high protein). All these variables are included because it is believed that the neural network model can pick the important variables automatically. Due to this large number of models and the large number of variables in each model, the training time for each model is limited to one minute, although the forecasts probably could be improved if more time is allowed in the training stage of the modeling procedure.

Composite Forecast

Composite forecasts often outperform individual model forecasts, because the errors involved in these separate forecasts will tend to be canceled out when they are averaged, especially if the principles on which these different forecasts are based are sufficiently different from one another (Kennedy 1989). Different weighting systems can be employed for different forecasting models, and the choice of models to be included in the composite forecast also matters.

Data and Estimation

Data

Most of the data series are discussed in chapter 3. USDA world agricultural supply and demand estimates (WASDE) are used as the predictions for total supply and exports in the three-year-average-plus forecasting equations. These estimates are published once a month. The estimates for the current marketing year are available from December to November, the projections for the next marketing year are available in issues starting in May each year.[31] Therefore for the months May-November, both estimates for current (old crop) marketing year and the next (new crop) marketing year are available. Total supply and export estimates for the marketing year 1979/80 are used for January-November 1980; in December

[31] For example, WASDE for marketing year 1979/80 are available in months from December 1979 to November 1980; while the WASDE for next market year 1980/81 available started in May 1980.

of 1980, the estimates for the 1980/81 marketing year are used.[32] The estimates for all other years, 1981-1995, are generated in the same manner.

Probably the largest volume of corn and soybean storage and hedge decisions are made in October, the harvest time with the weakest basis in the year. About 50 percent of the corn and soybeans are sold in the period from October to February. This study is going to base the forecasting performance tests in October.[33] Usually market participants are interested in basis forecasts for next month, two-months ahead, up to a year ahead. Therefore, basis forecasts are made in October for basis in November, December, January through October of the following year. Forecasting performance is tested with out-of-sample data. All the models are estimated on in-sample data sets and forecasts are made out-of-sample. Estimated models based on each in-sample data set are used to forecast 1-12 months ahead. Data from Nov. 1990 to Oct., 1995 are set aside for out-of-sample forecast testing. Therefore, there are five in-sample data sets, from January 1980 to October of 1990, 1991, 1992, 1993, and 1994, respectively. The in-sample data is updated each year until October 1994. Each time in-sample data are updated and new model estimates are made, the basis forecast for one-month ahead to 12-months ahead are obtained. For example, using in-sample data of January 1980--October 1990, basis forecasts are made for November 1990 (one-month ahead), December 1990 (two-month ahead), ..., October 1991 (12-months ahead). This procedure results in a total of five out-of-sample 1-12 months ahead basis forecasts.

---

[32] This data is set up in this way because most of the 1-12 months ahead forecasts involve old crop. The production estimates of old crop is used as much as possible.

[33] Forecasting based on other months can be extended easily from the models based in October.

Estimations

Before estimating all of the forecasting models, the simple three-year-average forecasts are tested by the following equation:

$$BS_t = \alpha + \beta * 3AVG_t + u_t, \text{ with } \alpha=0 \text{ and } \beta=1$$

The nested F test of the null hypothesis (α=0 and β=1) is rejected for most locations. Table 4.3 shows the F statistics of the test for two periods: one for the whole data set (1983-1995) and another for 1989-1995. The three-year-average forecast model fits well only in St. Louis soybean basis forecast for the period of 1983-1995, and for St. Louis and Toledo markets for 1989-1995. The test results are also indications of forecast performance.

The three-year-average-plus model has a better fit, compared to the simple three-year-average. For example, for the Northeast Iowa corn model, the R^2 increases from 0.03 to 0.22. At least one of these two additional variables is statistically significant (5% significance

Table 4.3. Nested F test of three-year-average basis forecasts

Product	Market	F value	Prob > F	F value	Prob > F
		Data Set: 1983-95		Data Set: 1989-95	
Corn	Chicago	42.51	0.0001	18.05	0.0001
	St. Louis	55.15	0.0001	13.34	0.0001
	Toledo	30.75	0.0001	4.46	0.0144
	NE Iowa	57.45	0.0001	13.56	0.0001
	C Illinois	71.64	0.0001	29.94	0.0001
	Gulf ports	60.80	0.0001	22.95	0.0001
	Richmond	30.16	0.0001	7.87	0.0007
	Pacific NW	18.69	0.0001	35.29	0.0001
Soybean	Chicago	55.91	0.0001	21.76	0.0001
	St. Louis	2.34	0.0995	3.56	0.0329
	Toledo	18.37	0.0001	1.71	0.1866
	NE Iowa	50.59	0.0001	10.71	0.0001
	C Illinois	46.97	0.0001	12.05	0.0001
	Gulf ports	51.86	0.0001	23.29	0.0001
	Richmond	22.00	0.0001	8.23	0.0005

level) in every market model. Similar results are found for the soybean models. In addition, when monthly dummies are added to the model, the nested F test shows that these dummies are significant. Thus a model with three-year-average basis, total supply, exports and monthly dummies is estimated. The mean R^2s over 5 equations, each for one of the 5 in-sample data sets, range from 0.33 to 0.57 for corn models, and from 0.24 to 0.40 for soybean models. A sample output is included in appendix B (Table B.8. Parameter estimates of three-year-average-plus forecasting model for Chicago).

Ordinary Least Squares is used in estimating structural basis behavior forecasting models for each contract. The model estimation is similar to the basis behavior models discussed in chapter 3 (see appendix Table B.9 for an example). As table 4.1 shows, different approaches are utilized to make ancillary forecasts for independent variables. The details of these ancillary forecasts are not listed here; examples are included in appendix B to illustrate the ancillary forecasts (Table B.10. Estimated coefficients of St. Louis barge rate SARIMA forecasting, and Table B.8. Northeast Iowa production regression estimates). Seasonal ARIMA captures the seasonal variation for barge rates. For example, the St. Louis barge rates seasonal ARIMA models have an R^2 range of 0.68 to 0.73 for all 5 data sets. The regional production for Northeast Iowa is the sum of the production of Iowa, Minnesota, Wisconsin and Illinois. It is this regional production data that is regressed on U.S. production. As expected, the simple regressions of state production on U.S. production have good fit, usually higher than 0.95 (see appendix Table B.11 for an example). The ancillary forecasts for other variables are just simple naive approaches, either equaling last period value (prime interest rate and animal units consuming grain), or the average of last three years values for the same month (cash prices, port export volumes and soybean crushing).

Table 4.4 shows the selected seasonal ARIMA models and their fits for corn and soybean at all the locations. The final models are different for each location and for each crop, and, occasionally, for different data sets. Though the R^2 is not very high, these selected models satisfy model selection criteria, such as convertibility, stationarity, uncorrelated disturbance, no significance disturbance ACF, etc., for ARIMA models.

State space model selection is basically automatically done, with no subjective judgment involved. The model is selected for which the Akaike's information criterion is minimized. In the 2-crop multivariate framework, corn and soybean basis at the same location are fitted in the same model. The corn basis multivariate state space model for all locations is estimated without Pacific NW since it has much fewer observations than all seven other markets (see appendix Table B.12 for a sample output from a 2-crop multivariate state space model).

Neural network estimation is another automatic procedure. It uses root mean squared error (RMSE) as the criterion for model selection in the training data sets. The model fit for the training set depends on the learning algorithm chosen and time used in the data training. In this study, because of the large number of models involved, an one minute training time is allowed for each model. In general these models do not have a good fit.

Table 4.4. Selected seasonal ARIMA models and R^2

		Com		Soybean	
	DATA SET	Model	R-Square	Model	R-Square
Chicago	Jan.80—Oct. 90	$(0,1,3)(0,1,2)_{12}$	0.34	$(3,1,1)(0,1,1)_{12}$	0.40
	Jan.80—Oct. 91	$(0,1,3)(0,1,2)_{12}$	0.34	$(3,1,1)(0,1,1)_{12}$	0.34
	Jan.80—Oct. 92	$(0,1,3)(0,1,2)_{12}$	0.39	$(3,1,1)(0,1,1)_{12}$	0.38
	Jan.80—Oct. 93	$(0,1,3)(0,1,2)_{12}$	0.37	$(3,1,1)(0,1,1)_{12}$	0.43
	Jan.80—Oct. 94	$(0,1,3)(0,1,2)_{12}$	0.39	$(3,1,1)(0,1,1)_{12}$	0.43
Gulf Ports	Jan.80—Oct. 90	$(1,1,1)(0,0,1)_{12}$	0.23	$(2,1,5)(1,0,1)_{12}$	0.42
	Jan.80—Oct. 91	$(1,1,1)(0,0,1)_{12}$	0.23	$(2,1,5)(1,0,1)_{12}$	0.35
	Jan.80—Oct. 92	$(1,1,1)(0,0,1)_{12}$	0.22	$(2,1,5)(1,0,1)_{12}$	0.37
	Jan.80—Oct. 93	$(1,1,1)(0,0,1)_{12}$	0.22	$(2,1,5)(1,0,1)_{12}$	0.39
	Jan.80—Oct. 94	$(1,1,1)(0,0,1)_{12}$	0.22	$(2,1,5)(1,0,1)_{12}$	0.39
Pacific NW	Jan.80—Oct. 90	$(2,1,1)(0,0,2)_{12}$	0.30		
	Jan.80—Oct. 91	$(1,1,2)(0,0,2)_{12}$	0.31		
	Jan.80—Oct. 92	$(2,1,1)(0,0,2)_{12}$	0.31		
	Jan.80—Oct. 93	$(2,1,1)(0,0,2)_{12}$	0.31		
	Jan.80—Oct. 94	$(2,1,1)(0,0,2)_{12}$	0.31		
Richmond	Jan.80—Oct. 90	$(1,1,3)(2,1,1)_{12}$	0.52	$(0,1,2)(1,0,1)_{12}$	0.38
	Jan.80—Oct. 91	$(1,1,3)(2,1,1)_{12}$	0.52	$(0,1,2)(1,0,1)_{12}$	0.42
	Jan.80—Oct. 92	$(1,1,3)(2,1,1)_{12}$	0.50	$(0,1,2)(1,0,1)_{12}$	0.37
	Jan.80—Oct. 93	$(3,1,1)(2,1,1)_{12}$	0.52	$(0,1,2)(1,0,1)_{12}$	0.36
	Jan.80—Oct. 94	$(1,1,3)(2,1,1)_{12}$	0.58	$(0,1,2)(1,0,1)_{12}$	0.33
St. Louis	Jan.80—Oct. 90	$(1,1,3)(0,0,3)_{12}$	0.39	$(3,1,1)(1,1,1)_{12}$	0.58
	Jan.80—Oct. 91	$(1,1,3)(0,0,3)_{12}$	0.38	$(3,1,1)(1,1,1)_{12}$	0.59
	Jan.80—Oct. 92	$(1,1,3)(0,0,3)_{12}$	0.34	$(3,1,1)(1,1,1)_{12}$	0.61
	Jan.80—Oct. 93	$(1,1,3)(0,0,3)_{12}$	0.33	$(3,1,1)(1,1,1)_{12}$	0.58
	Jan.80—Oct. 94	$(1,1,3)(0,0,3)_{12}$	0.34	$(1,1,1)(1,1,1)_{12}$	0.61
Toledo	Jan.80—Oct. 90	$(1,1,3)(0,1,2)_{12}$	0.39	$(0,1,3)(1,1,1)_{12}$	0.51
	Jan.80—Oct. 91	$(1,1,3)(0,1,2)_{12}$	0.40	$(0,1,3)(1,1,1)_{12}$	0.49
	Jan.80—Oct. 92	$(1,1,3)(0,1,2)_{12}$	0.43	$(0,1,3)(1,1,1)_{12}$	0.52
	Jan.80—Oct. 93	$(1,1,3)(0,1,2)_{12}$	0.40	$(0,1,3)(1,1,1)_{12}$	0.51
	Jan.80—Oct. 94	$(1,1,3)(0,1,2)_{12}$	0.41	$(0,1,3)(1,1,1)_{12}$	0.58
Northeast IA	Jan.80—Oct. 90	$(3,1,3)(1,0,0)_{12}$	0.29	$(1,1,1)(2,1,1)_{12}$	0.60
	Jan.80—Oct. 91	$(3,1,3)(1,0,0)_{12}$	0.27	$(1,1,1)(2,1,1)_{12}$	0.60
	Jan.80—Oct. 92	$(3,1,3)(1,0,0)_{12}$	0.27	$(1,1,1)(2,1,1)_{12}$	0.61
	Jan.80—Oct. 93	$(3,1,3)(1,0,0)_{12}$	0.26	$(1,1,1)(2,1,1)_{12}$	0.60
	Jan.80—Oct. 94	$(3,1,3)(1,0,0)_{12}$	0.25	$(1,1,1)(2,1,1)_{12}$	0.59
Central IL	Jan.80—Oct. 90	$(0,1,3)(1,0,0)_{12}$	0.30	$(1,1,2)(1,0,0)_{12}$	0.35
	Jan.80—Oct. 91	$(0,1,3)(1,0,0)_{12}$	0.31	$(1,1,2)(1,0,0)_{12}$	0.36
	Jan.80—Oct. 92	$(0,1,3)(1,0,0)_{12}$	0.31	$(1,1,2)(1,0,0)_{12}$	0.35
	Jan.80—Oct. 93	$(0,1,3)(1,0,0)_{12}$	0.30	$(1,1,2)(1,0,0)_{12}$	0.35
	Jan.80—Oct. 94	$(0,1,3)(1,0,0)_{12}$	0.30	$(1,1,2)(1,0,0)_{12}$	0.35

Basis Forecasts and Performance Comparison

This section will first discuss the criteria to measure the forecasting accuracy. The properties of the three-year-average forecasts are examined first, then compared to the performance of all the forecasting models. A more detailed discussion of each forecasting models' performance in Northeast Iowa follows as a case study.

Measures of Forecast Accuracy

All of the forecasts by various forecasting models discussed above are compared to forecasts of the bench mark model, the simple 3-year-average model, and to each other. Five different criteria measuring the accuracy of forecasts are used: mean absolute error (MAE), root mean squared errors (RMSE), two Theil's U statistics (or Theil's inequality coefficients), and the Henriksson-Merton test. Mean absolute error (MAE) calculates the average of absolute values of the forecast errors, while root mean square error (RMSE) is the square of the average squared values of forecast errors. Two Theil's U statistics (defined as the square root of ratio of the mean square error of the predicted (or percentage) change to the average squared actual (or percentage) change) are also used. The formulae for the first four criteria are:

$$MAE = \frac{1}{n^0}\Sigma_i|A_i - P_i|$$

$$RMSE = \sqrt{\frac{1}{n^0}\Sigma_i(A_i - P_i)^2}$$

$$U_{\Delta} = \sqrt{\frac{(1/n^0)\Sigma_i(\Delta A_i - \Delta P_i)^2}{(1/n^0)\Sigma_i \Delta A_i^2}}$$

where P and A represent predicted and actual values, respectively, and n^0 is the number of periods being forecasted. The U statistics can be calculated in two different ways: (1) $\Delta A_i = A_i - A_{i-1}$ and $\Delta P_i = P_i - A_{i-1}$, or (2) $\Delta A_i = (A_i - A_{i-1})/A_{i-1}$ and $\Delta P_i = (P_i - A_{i-1})/A_{i-1}$,

These four measures will be zero for perfect forecasts, larger values indicate poor forecasts. There are few differences among these forecasting accuracy measures. The RMSE penalizes models with large prediction error more than MAE does. Therefore, MAE is more appropriate when the cost of forecast errors is proportional to the absolute size of the forecast error, while RMSE is more appropriate to situation in which the cost of the error increases in line with the square of the forecast error. These two measures have scaling problems which should not affect the comparison among our forecasting models. The U statistic, calculated in absolute change or percentage change, will measure the model's ability to track tuning points in the data (Greene 1993). When $U = 1$, the forecast is as good as no-change forecast ($\Delta P = 0$). For $U > 1$, the forecast is less accurate than the simple forecast of no change.

The Henriksson-Merton test[34] is a probability-based measure of turning points forecasting performance. A confidence level, c, of rejecting the null hypothesis of no information value (i.e. can not predict the direction of revision in the series) can be constructed based on the test:

[34] For more detail of the test, see Henriksson and Merton (1981), Cumby and Modest (1987), and McIntosh and Dorfman (1992).

$$c = 1 - \sum_{x=n_1}^{\min(N_1,n)} \left(\frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}} \right)$$

where N_1 = number of observations with downward movement
N_2 = number of observations with nondownward movement
n_1 = number of correct forecasts of downward movement
n_2 = number of incorrect forecasts of nondownward movement
$N = N_1 + N_2$
$n = n_1 + n_2$ = number of forecasts of downward movement

As suggested by McIntosh and Dorfman (1992), using a one-tailed test with a confidence level of c, the null hypothesis of no information value would be rejected at any significance level greater than 1-c.

Performance Comparison

All five performance measures are calculated for the forecasts generated by alternative models. The results will be little different with different error or turning point criteria. The MAE and RMSE give similar results and two Theil's U coefficients do not differ much. The Henriksson-Merton test result is similar to U statistics. In general, the results of all the measures are consistent with each other. For brevity, we select RMSE, one U coefficient with percentage change and Henriksson-Merton tests to illustrate the forecast performance.

Forecast errors

Table 4.5 shows the number of times the particular forecasting method has the lowest RMSE. There are 5 sets of 1-12 months ahead forecasts for each market and 8 markets for corn models (7 for soybean models), therefore there will be total of 40 RMSEs for corn and 35 for soybean. The RMSE averages are calculated for forecast period of 1-12 month ahead, and also for three shorter periods: 1-4, 5-8, and 9-12 months ahead (defined roughly as short term, intermediate term, and long term forecasts, respectively).

Table 4.5. Number of time the forecasting models have the lowest RMSE

Forecast	Forecasting Models [a]									
Periods	3YR	3YR+	ARIMA	USS	MSS2	MSS7	SBBM	NFD	SCG	COMP
					Corn					
1-12	6	4	6	2	4	4	1	8	0	5
1-4	7	8	0	3	0	2	8	4	1	7
5-8	6	5	1	10	3.5[b]	1.5	2	5	0	6
9-12	8	0	3	4	3	3	1	7	6	5
					Soybean					
1-12	3	11	9	1	1	2	1	1	1	5
1-4	5	5	7	1	3	1	3	1	2	7
5-8	6	12	2	1	2	0	4	2	3	3
9-12	2	9	8	1	2	4	0	3	1	5

[a] The forecasting models are simple 3-year-average, 3-year-average-plus, seasonal ARIMA, univariate state space, 2-crop multivariate state space, 7-market multivariate state space, structural basis behavior, neural network with NFD algorithm, neural network with SCG algorithm, and a composite forecast, respectively.
[b] The number with 0.5 indicates the lowest RMSE is shared by two models.

In general, three-year-average-plus (3YR+), artificial neural network with noise-feedback descent (NFD) learning algorithm, composite forecasts (COMP), and simple three-year-average (3YR) models perform better than other models in forecasting corn basis for eight locations. For 1-12 month ahead corn basis forecasts, artificial neural networks with NFD has the largest number of lowest RMSEs; the second best models are the simple three-year-average and seasonal ARIMA model (each has the lowest RMSE 6 times), closely followed by the composite model. The structural basis behavior forecasting model (SBBM) and three-year-average-plus models outperform other models (each with the lowest RMSE 8 times) in the short term (1-4 months) forecasts. The univariate state space model is best for the intermediate term forecasts, and no models beat the simple three-year-average model in long term forecasts (though both neural network models, NFD and with scaled conjugate gradient (SCG), come close). The multivariate state space models, both 2-crop (MSS2) and 7-markets (MSS7), do not perform well in corn basis forecasting.

For soybean basis forecasts, the three-year-average-plus model, seasonal ARIMA model and the composite forecasts generally outperform simple three-year-average forecasts. The three-year-average-plus model is the best according this criterion for 1-12, 5-8 and 9-12 month ahead forecasts, while seasonal ARIMA and composite forecasting models outperform all other models in the short-term soybean basis forecasts, and also perform well for other forecasting periods.

Tables 4.6 and 4.7 list the mean RMSE of 5-sets of 1-12 month ahead basis forecasts in each market location for corn and soybean, respectively. The artificial neural network with NFD algorithm is good at forecasting 1-12 month ahead corn basis for three markets (Toledo, Gulf Ports and Pacific NW), while composite forecasts outperform simple three-year-average forecasts in two markets. The composite is even better in short-term forecasts (it has the lowest value in three markets). Structural basis behavior forecasting model does a good job in forecasting short-term corn basis for port markets of Pacific NW and Richmond. For relatively long-term corn basis forecasts, the neural networks with noise-feedback learning

Table 4.6. Mean of the RMSE of 5-sets of corn basis forecasts

Markets	3YR	3YR+	ARIMA	USS	MSS2	MSS7	SBBM	NFD	SCG	COMP
				1-12 month ahead						
NE Iowa	6.88	6.83	7.24	7.72	6.25 [a]	7.55	15.35	7.26	8.56	8.40
C Illinois	5.05	5.77	4.87 [a]	9.33	5.79	6.15	9.42	5.61	6.72	5.00
Chicago	4.32	4.97	4.24	4.70	4.59	4.51	6.82	4.46	7.10	3.72 [a]
St. Louis	4.49 [a]	5.50	6.51	6.38	6.47	9.79	11.85	7.82	8.04	5.55
Toledo	6.58	8.45	6.64	8.61	8.37	8.39	9.25	5.84 [a]	8.88	5.86
Gulf Port	6.94	6.69	6.27	6.42	6.68	6.16	10.71	5.78 [a]	7.55	6.91
Richmond	11.63	13.88	11.72	12.34	15.37	18.57	10.91	12.41	12.89	10.28 [a]
Pacific NW	8.54	9.89	7.90	7.78	[b]		10.22	7.70 [a]	10.62	13.36
				1-4 month ahead						
NE Iowa	4.35	4.03 [a]	4.10	4.37	4.86	5.59	8.31	7.53	7.69	5.53
C Illinois	3.88	3.34	2.93	6.66	3.05	3.24	5.52	4.72	6.18	2.79 [a]
Chicago	2.79	2.43	2.68	2.82	3.26	3.06	3.54	4.04	4.74	2.10 [a]
St. Louis	3.06 [a]	3.73	4.36	6.19	7.00	10.94	8.80	10.12	9.81	4.47
Toledo	5.22	4.92	4.81	6.38	6.23	6.86	5.76	4.90	7.37	3.99 [a]
Gulf Port	6.83	6.55	6.21	6.23	6.74	5.92 [a]	10.34	6.64	8.14	6.83
Richmond	11.34	12.29	12.14	12.91	16.39	18.43	9.32 [a]	11.81	11.97	10.12
Pacific NW	7.47	8.47	7.56	7.92			6.08 [a]	8.72	12.94	13.55
				5-8 month ahead						
NE Iowa	4.83	4.50	3.81	3.66 [a]	4.66	5.73	14.26	4.98	7.34	5.93
C Illinois	3.06	2.04 [a]	2.57	7.02	3.18	3.75	6.35	3.91	3.64	2.09
Chicago	4.63	4.46	4.32	4.96	4.87	4.92	6.69	4.57	5.30	3.73 [a]
St. Louis	2.48 [a]	3.06	4.85	3.47	4.64	9.45	9.98	4.02	4.31	3.56
Toledo	5.46 [a]	6.40	7.05	10.03	8.25	8.88	7.08	5.66	10.75	5.54
Gulf Port	4.87	5.02	5.34	6.21	6.10	4.42 [a]	9.38	4.76	6.11	5.16
Richmond	8.92	11.33	9.42	8.05	13.25	17.89	9.11	9.36	9.77	7.09 [a]
Pacific NW	8.48	9.99	7.45	7.06			10.67	6.31 [a]	8.21	11.50
				9-12 month ahead						
NE Iowa	9.07	9.53	10.80	11.74	8.18	9.98	18.61	8.01 [a]	9.56	11.11
C Illinois	6.28 [a]	8.92	7.28	11.98	8.68	9.21	12.62	7.24	9.03	7.58
Chicago	4.57	6.56	4.87	5.23	5.18	4.91	8.63	3.97 [a]	9.70	4.25
St. Louis	6.55 [a]	8.01	8.99	7.83	6.85	8.38	15.13	7.26	7.47	7.55
Toledo	7.62	12.12	6.73	8.36	8.96	8.15	11.79	5.99 [a]	6.97	6.60
Gulf Port	7.76	7.76	6.54	5.89	6.58	7.32	10.87	5.43 [a]	7.76	7.98
Richmond	12.73	16.77	11.21 [a]	13.74	14.66	18.34	11.76	14.29	14.39	11.25
Pacific NW	9.02	10.63	8.38	7.61			11.81	7.43 [a]	9.07	14.09

[a] indicate the lowest value among the forecasting methods.
[b] Multivariate state space model is not estimated for Pacific NW model.

Table 4.7. Mean of the RMSE of 5-sets of soybean basis forecasts

Markets	3YR	3YR+	ARIMA	USS	MSS2	MSS7	SBBM	NFD	SCG	COMP
				1-12 month ahead						
NE Iowa	9.81	7.27 [a]	8.79	8.28	8.00	8.18	17.22	9.60	12.98	8.96
C Illinois	7.28	6.83	7.22	9.76	9.86	10.90	8.26	7.35	8.42	6.16 [a]
Chicago	6.79	4.63 [a]	8.33	16.48	18.00	16.47	9.00	6.93	6.54	6.71
St. Louis	6.72	5.73	5.82	7.94	10.54	10.54	10.23	7.96	9.31	5.62 [a]
Toledo	5.99	6.63	7.46	10.99	11.75	11.30	8.95	7.28	9.37	5.58 [a]
Gulf Port	9.03	7.37 [a]	8.69	8.69	8.86	8.40	15.80	8.61	8.06	8.68
Richmond	12.59	10.73	11.60	17.29	12.83	14.44	15.56	10.69 [a]	13.52	11.16
				1-4 month ahead						
NE Iowa	7.96	6.36	7.34	5.41 [a]	6.71	6.52	8.70	7.94	10.16	6.03
C Illinois	8.07	8.13	6.98	10.32	10.52	11.46	8.22	7.85	10.10	6.23 [a]
Chicago	4.30	3.39 [a]	7.29	12.37	13.81	12.35	4.58	6.93	4.95	4.91
St. Louis	4.64	4.60	4.78	5.64	5.56	6.43	5.40	5.96	9.17	4.06 [a]
Toledo	4.87	5.07	5.05	7.36	8.18	8.17	7.02	5.77	6.92	4.04 [a]
Gulf Port	6.79	6.12	5.83 [a]	7.02	7.35	7.49	10.05	7.92	8.80	7.33
Richmond	9.51	6.77	6.06 [a]	14.04	9.26	11.25	11.68	9.56	12.37	8.56
				5-8 month ahead						
NE Iowa	9.44	5.91 [a]	9.34	7.02	9.16	9.60	18.22	9.43	14.81	7.93
C Illinois	5.88	4.24 [a]	6.10	9.93	9.89	11.58	7.08	5.75	7.13	4.91
Chicago	3.47	2.21 [a]	5.81	14.22	15.74	14.25	4.58	4.43	3.90	4.19
St. Louis	5.14	3.38 [a]	5.36	6.56	9.85	11.78	13.99	5.59	7.56	4.50
Toledo	2.95 [a]	2.98	6.90	10.04	11.63	11.22	5.98	5.22	7.42	3.34
Gulf Port	9.85	6.87 [a]	9.93	9.52	9.69	9.00	19.32	8.44	7.01	8.81
Richmond	10.59	10.29	11.47	18.23	12.23	15.21	14.38	9.17	12.99	10.18 [a]
				9-12 month ahead						
NE Iowa	9.21	8.14	7.92	10.32	6.77 [a]	7.17	18.43	10.23	11.11	10.39
C Illinois	6.69	6.03 [a]	7.24	7.58	7.48	8.22	7.82	6.87	7.28	6.23
Chicago	9.94	6.52 [a]	9.70	20.30	21.90	20.32	14.00	7.94	8.89	8.83
St. Louis	8.30	7.49	5.88 [a]	10.48	12.71	11.72	8.31	8.34	9.09	6.53
Toledo	7.78	9.00	8.57	13.52	12.75	12.28	11.95	8.59	11.21	7.12 [a]
Gulf Port	8.48	7.42	8.16	7.93	8.31	7.28 [a]	12.87	8.22	7.49	8.52
Richmond	14.91	11.63	14.01	16.90	14.53	14.76	17.51	11.45 [a]	13.52	11.60

[a] indicate the lowest value among the forecasting methods.

algorithm beats other forecasts in 5 out of 8 markets. Generally, the three-year-average forecasts have RMSEs of 4-7 cents/bushel for major corn producing areas. Relative to the size of the basis, that error appears large, and potentially important to large volume grain merchandisers. Improved forecasts shave a fraction of a cent up to less than 2 cents per bushel off the RMSE, which is approximately equivalent to the standard error of the forecast at the means. For some large volume grain merchandisers, that may be worth the added effort to develop and update their forecast procedures, but for many it will not be worth while.

For soybean basis forecasts, three-year-average-plus model and composite forecast stand out as the best basis forecast methods 1-12 months ahead; each has lowest mean value of RMSE in three market out of seven markets. The three-year-average-plus model forecasts outperform other 5-8 month ahead forecasts in 5 markets. Seasonal ARIMA technique forecasts have the lowest mean value of RMSE in two port markets for short-term soybean basis forecasts. The artificial neural network approach enjoys less success in forecasting soybean basis than in forecasting corn basis. The improvement in lowest forecast RMSEs are slightly larger for soybeans than corn, but the three-year-average RMSE is also larger, as expected for a commodity with higher price levels.

Turning point accuracy

The Theil's U coefficients (see Table 4.8) generally are consistent with the findings of RMSEs. For corn basis forecasting, the seasonal ARIMA model outperforms all others for 1-12 months ahead forecasts. The three-year-average-plus and structural basis behavior models are the best two in short term (1-4 months) forecasts, while the univariate state space model leads in intermediate term forecasts. The simple three-year-average is good for 9-12 months ahead forecasts, followed by neural network approach with NFD algorithm. For soybean basis forecasting, it is very clear that three-year-average-plus model is the best for all forecasting periods, seasonal ARIMA comes second, which is followed by composite forecasts (all of these three models beat simple three-year-average forecasts).

Table 4.8. Number of time the forecasting models have the lowest U coefficients

Forecast	Forecasting Models									
Period	3YR	3YR+	ARIMA	USS	MSS2	MSS7	SBBM	NFD	SCG	COMP
					Corn					
1-12	6	5	8	3	3	4	3	3	2	3
1-4	5	9	2	3	0	3	7	4	2	5
5-8	7	5	1	11	3	1	2	5	1	4
9-12	9	1	4	4	1	4	2	7	4	4
					Soybean					
1-12	4	11	8	2	1	1	0	2	2	4
1-4	5	8	6	2	2	0	1	1	1	9
5-8	4	12	1	2	1	1	3	3	3	5
9-12	2	9	7	3	3	4	1	3	1	2

Though Theil's U coefficient of inequality gives some indication in this aspect, theHenriksson-Merton test can provide additional information about model's ability to forecasting directional change. The Henriksson-Merton test is a probability-based measure and its statistics has a known distribution, so the test can provide exact confidence and significance levels for the rejection of the null hypothesis of no information (can not predict the direction of revision in the series) (McIntosh and Dorfman 1992). For example, for a confidence level of 0.95, the null hypothesis can be rejected at the 5% significance level. Table 4.9 shows that simple three-year-average, three-year-average-plus model, seasonal ARIMA model, and composite forecasts have high confidence levels, which indicate that these

Table 4.9. Confidence level, c, of Henriksson-Merton test

Location	3YR	3YR+	ARIMA	USS	MSS2	MSS7	SBBM	NFD	SCG	COMP
					Corn					
NE Iowa	0.892	0.905	0.996	0.996	0.998	0.783	0.663	0.663	0.317	0.933
C Illinois	0.990	0.984	1.000	0.990	0.998	0.999	0.781	0.849	0.219	0.998
Chicago	1.000	1.000	1.000	1.000	1.000	1.000	0.981	0.990	0.783	1.000
St. Louis	0.999	0.998	0.982	0.995	0.996	0.997	0.612	0.949	0.901	0.987
Toledo	0.979	0.997	0.938	0.984	0.916	0.968	0.771	0.932	0.888	0.998
Gulf Ports	0.998	0.964	0.977	0.999	0.987	0.947	0.947	0.987	0.977	0.996
Richmond	0.997	0.927	0.996	0.592	0.617	0.557	0.986	0.924	0.952	0.929
Pacific NW	0.993	0.995	0.981	0.924	(b)	(b)	0.996	0.775	0.954	0.986
					Soybean					
NE Iowa	1.000	0.991	0.995	0.992	0.996	0.989	0.980	0.981	0.696	1.000
C Illinois	0.996	0.999	0.998	0.999	0.999	0.970	0.998	0.982	0.849	1.000
Chicago	0.991	1.000	0.998	0.990	0.964	0.964	0.980	0.999	0.711	0.990
St. Louis	1.000	0.998	1.000	0.989	0.951	0.960	0.984	0.998	0.982	1.000
Toledo	1.000	1.000	0.999	0.999	0.984	0.918	0.990	0.977	0.679	1.000
Gulf Ports	0.990	0.999	0.927	0.805	0.562	0.941	0.876	0.882	0.882	0.992
Richmond	0.995	0.992	0.998	0.298	0.853	0.945	0.992	0.965	0.604	0.991

(b) Multivariate state space model is not estimated for Pacific Northwest.

forecasts contain information and perform well in predicting directional movement of basis change. The worst are state space models and neural network approaches.

Table 4.10 shows the number of locations the null hypothesis of no information is rejected at different significance levels. At the 5% level, simple three-year-average, three-year-average-plus, seasonal ARIMA models and composite forecasts predict the directional moves of corn basis well. For soybean basis forecast, the null hypothesis is rejected for every model (except neural network model with scaled conjugate gradient algorithm) in more than 4 out of 7 locations at 5% significance level, with the simple three-year-average, three-year-average-plus, seasonal ARIMA and composite forecasts significant at the 1% significance level. That suggests turning points in the soybean basis are captured fairly well by most models in most locations.

Table 4.10. Number of locations the null hypothesis of no information is rejected at 5% and 1% significance level

	3YR	3YR+	ARIMA	USS	MSS2	MSS7	SBBM	NFD	SCG	COMP
5% Corn	7	6	6	5	4	4	3	2	3	6
Soybean	7	7	6	5	5	4	6	6	1	7
1% Corn	6	4	3	4	3	3	1	1	0	4
Soybean	6	7	6	3	2	0	2	2	0	6

Case Study: Northeast Iowa Basis Forecasting

To illustrate the comparative basis forecasting performance of each method in a way potentially useful to Iowa farmers, grain merchandisers and extension specialists, we examine the out-of-sample forecast performance for Northeast Iowa in more depth. Each forecasting model is compared to the bench mark three-year-average forecasts.

Forecasts of three-year-average

The naive 3-year-average model approach to basis forecasting appeals to market participants mainly for its simplicity. Figure 4.2 is an example of its forecasts for Northeast Iowa corn basis. It is very clear that its forecasts are close to actual bases when the bases are stable over years such as the basis from 1990 to 1994. If basis has lots of variations from year to year (e.g. from late 1986 to the end of 1989), then this naive forecast often has large errors. Figure 4.3 plots the forecast errors. There are systematic positive errors in the early 80's, followed by systematic negative errors in the late 80's. But in the 90's, both the level and variance of the forecast errors decreased. This pattern is evident in the corn basis forecasts for all the market locations studied. For the soybean basis forecasts using this 3-yr-average approach, a similar pattern is also found.

The out-of-sample forecasts of the approaches discussed below focus on the period November 1990-October 1995, during which the naive approach had much better performance compared to the 1980s. It is difficult for other approaches to outperform this naive approach for this out-of-sample period.

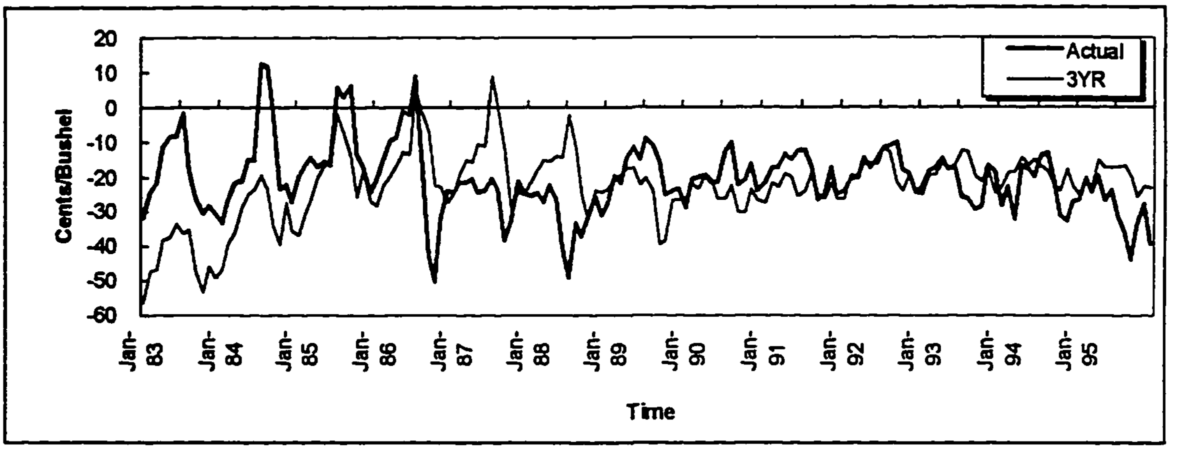

Figure 4.2. 3-year-average corn basis forecasts, Northeast Iowa

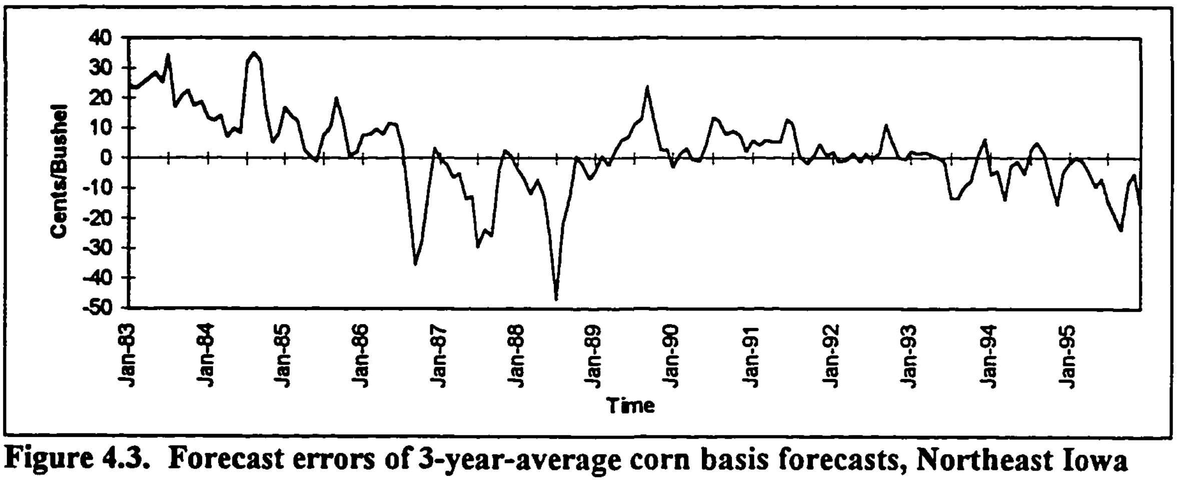

Figure 4.3. Forecast errors of 3-year-average corn basis forecasts, Northeast Iowa

Forecasts of 3-year-average plus model

All five-sets of 1-12 month ahead basis forecasts are obtained from the estimated coefficients and October USDA projections for total supply and export volume. Figure 4.4 plots three-year-average-plus and simple 3-year-average corn basis forecast errors for Northeast Iowa. It shows that basis forecasts errors generated by 3-year-average plus model follows closely the errors of 3-year-average forecasts. The plus model is a little better for the first set for forecasts from November 1990 to October 1991, and it is very close to three-year-average forecasts for the second sets of forecasts: November 1991 October 1992. For the last set of forecasts, November 1994 to October 1995, both forecasts are off a lot, but the plus model is even worse. The large forecast errors in recent years for both forecast methods are due to the facts that the basis in these months of the year deviated from usual basis patterns. The flooding in the summer of 1993 resulted in unusually wide basis and large forecast errors in the summer of 1993. The forecasts for subsequent years utilizing that wide basis led to

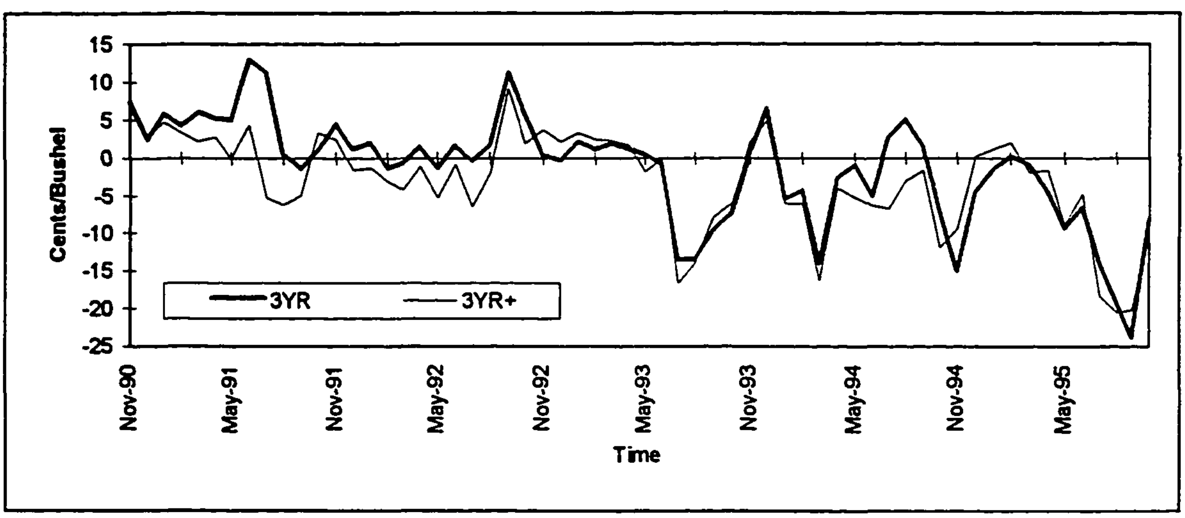

Figure 4.4. Northeast Iowa corn basis forecast errors: 3-yr-average-plus vs. 3-yr-average

large errors in subsequent years too.

Structural basis model forecasts

The estimated coefficients from basis behavior models will be used to forecast the out-of-sample corn and soybean basis. The ancillary forecasts for independent variables are obtained by the methods specified in Table 4.1. Each of the five sets of 1-12 months ahead basis forecasts then are derived by plugging these ancillary forecasts into the estimated equations. Figure 4.5 shows the basis forecasts for Northeast Iowa. This example clearly shows that structural basis behavior forecasting model does a much worse job than simple three-year-average forecasts. This forecasting method tends to overshoot the simple three-year-average forecasts, and has much larger errors.

However, the forecast errors of these structural basis behavior models are caused in part by the forecasting errors from the simple ancillary forecasting methods used. A better

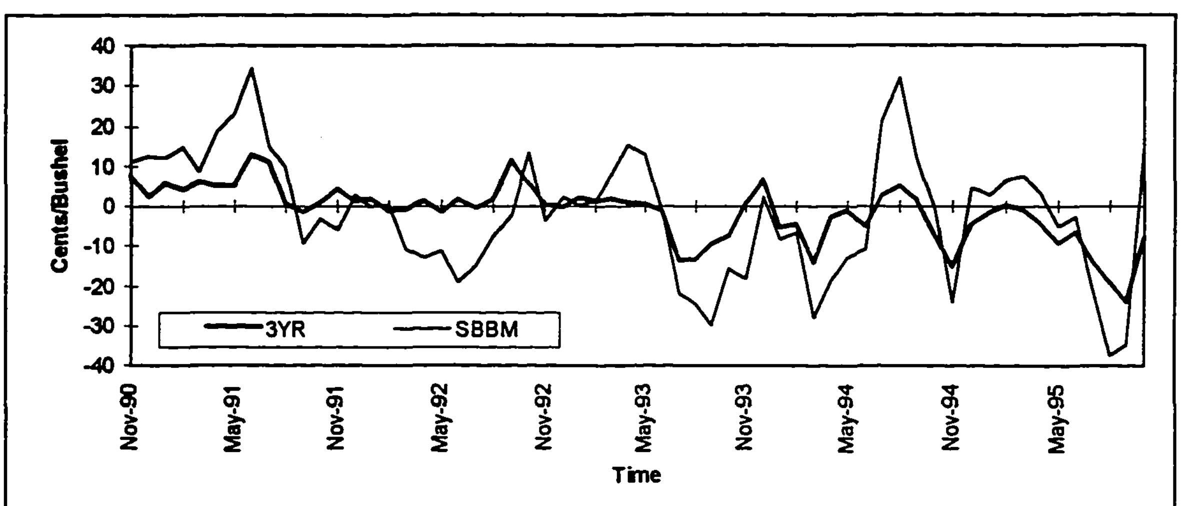

Figure 4.5. Northeast Iowa corn basis forecast errors: structural basis model vs. 3-yr-average

forecasting model for each independent variable could substantially improve the results, and is vital in this approach of basis forecasting. Yet the additional time and expense of developing ancillary forecasts makes it less practical than many other methods in forecasting corn and soybean basis.

ARIMA model forecasts

An example of the modeling procedure used to identify, estimate and forecast the seasonal ARIMA model is included in Appendix C. The seasonal ARIMA forecast errors of the nearby corn basis for Northeast Iowa and the naive forecast results are plotted in Figure 4.6 and Figure 4.7. It is clear that the Seasonal ARIMA forecasts are more accurate in the first year than in 1993 and 1994. This is due to the ARIMA using more recent basis information, one to three months lagged basis and one seasonal lagged basis in this case (the model selected is (3,1,3)(1,0,0), see Table 4.4 for models selected for all the markets), while

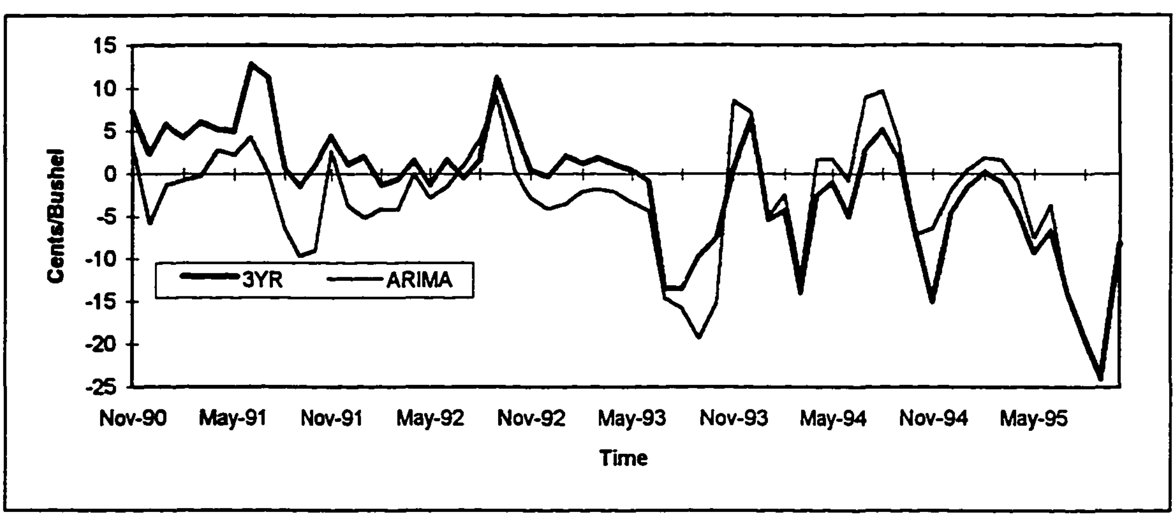

Figure 4.6. Northeast Iowa corn basis forecast errors, seasonal ARIMA vs. 3-year-average

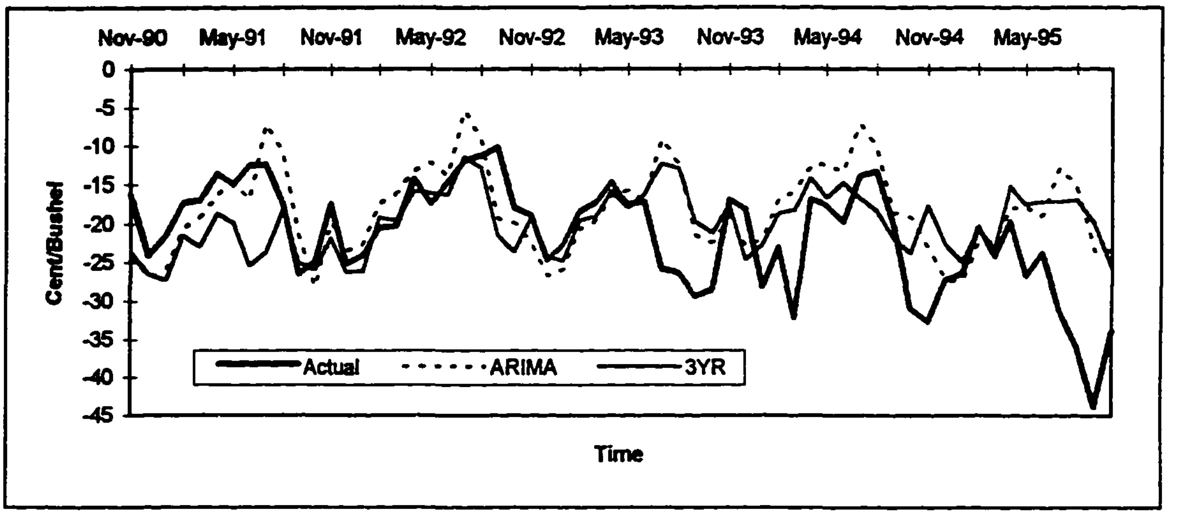

Figure 4.7. Northeast Iowa corn basis forecasts, seasonal ARIMA vs. 3-year-average

simple three-year-average forecasts do not utilize this most recent basis information.

State space model forecasts

Figure 4.8 plots the univariate state space model forecast and simple 3-year-average basis forecasts, Figure 4.9 compares multivariate (two-crops) state space model forecasts with simple three-year-average basis forecasts, and Figure 4.10 plots seven-market multivariate state space model forecasts against three-year-average forecasts. Figure 4.11 shows that all three state space modeling approaches are similar in the early years. These forecasts are almost straight line for each 1-12 month ahead forecast set. Only the two-crop multivariate approach, which considers both corn and soybean basis together in the same location, shows some deviations from straight line in the last two years of forecasting. However, the plots show that their forecasts could be more accurate than simple three-year average forecasts in some of the forecasting periods.

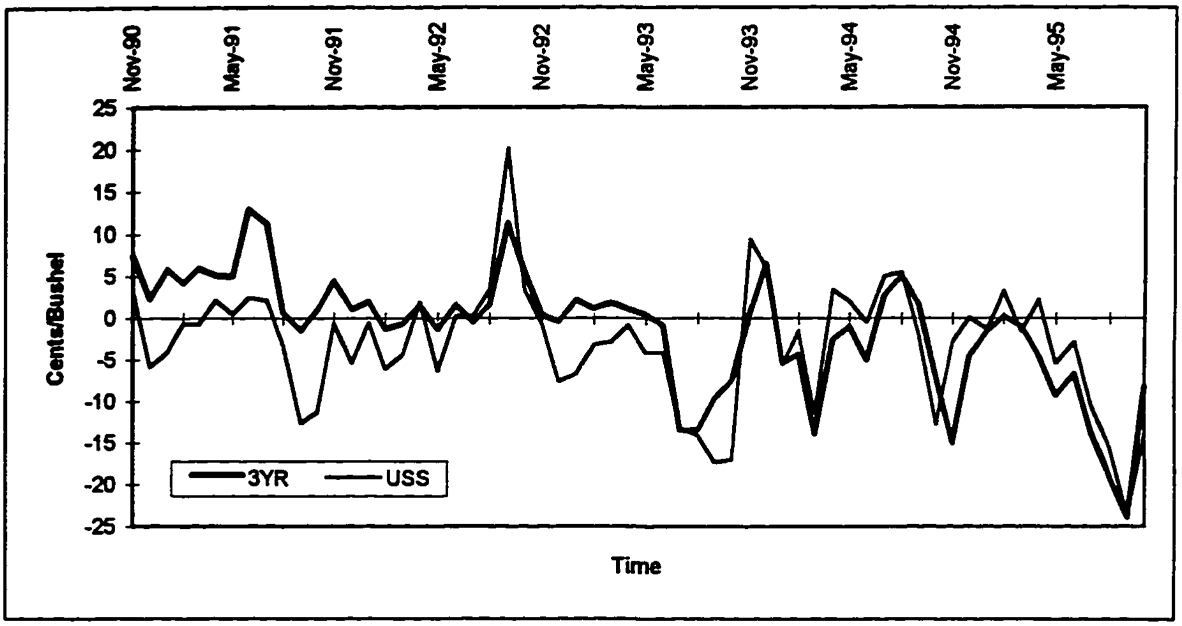

Figure 4.8. **Northeast Iowa corn basis forecast errors, univariate State Space models vs. three-year-average.**

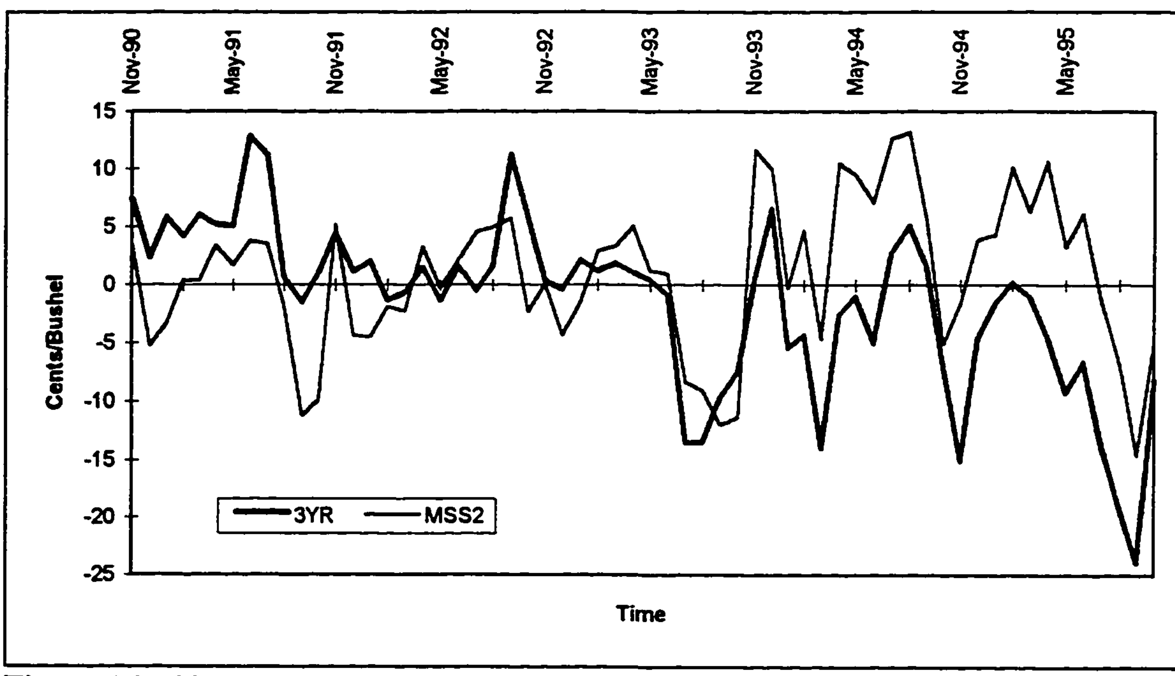

Figure 4.9. **Northeast Iowa corn basis forecasts, multivariate State Space model (2-crops) vs. three-year-average.**

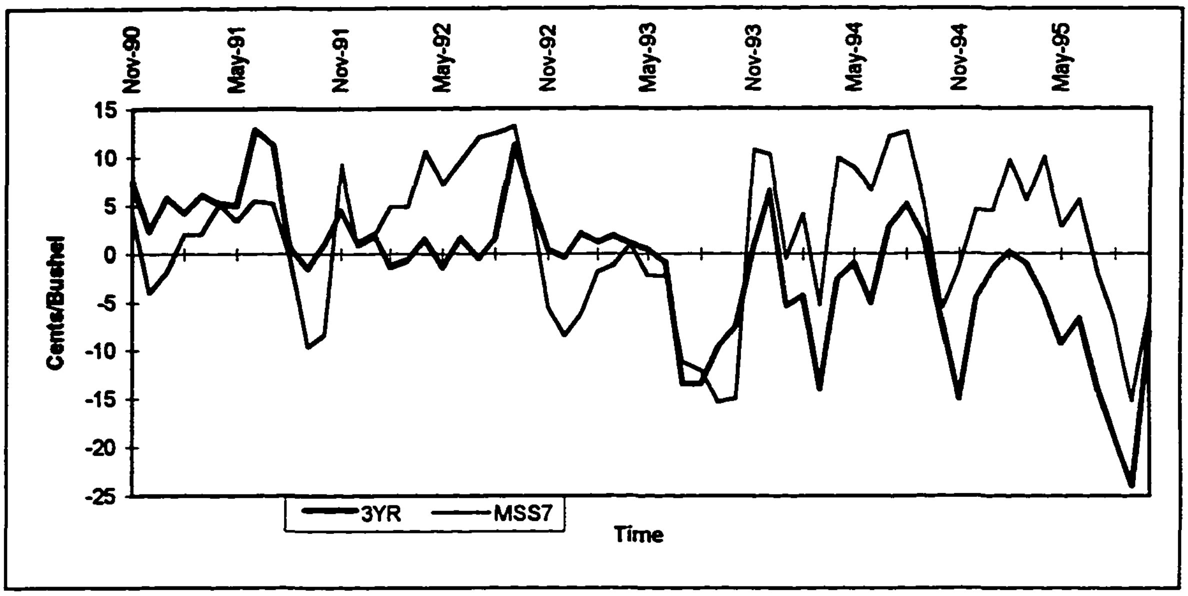

Figure 4.10. Northeast Iowa soybean basis forecasts, multivariate State Space model (7-markets) vs. three-year-average.

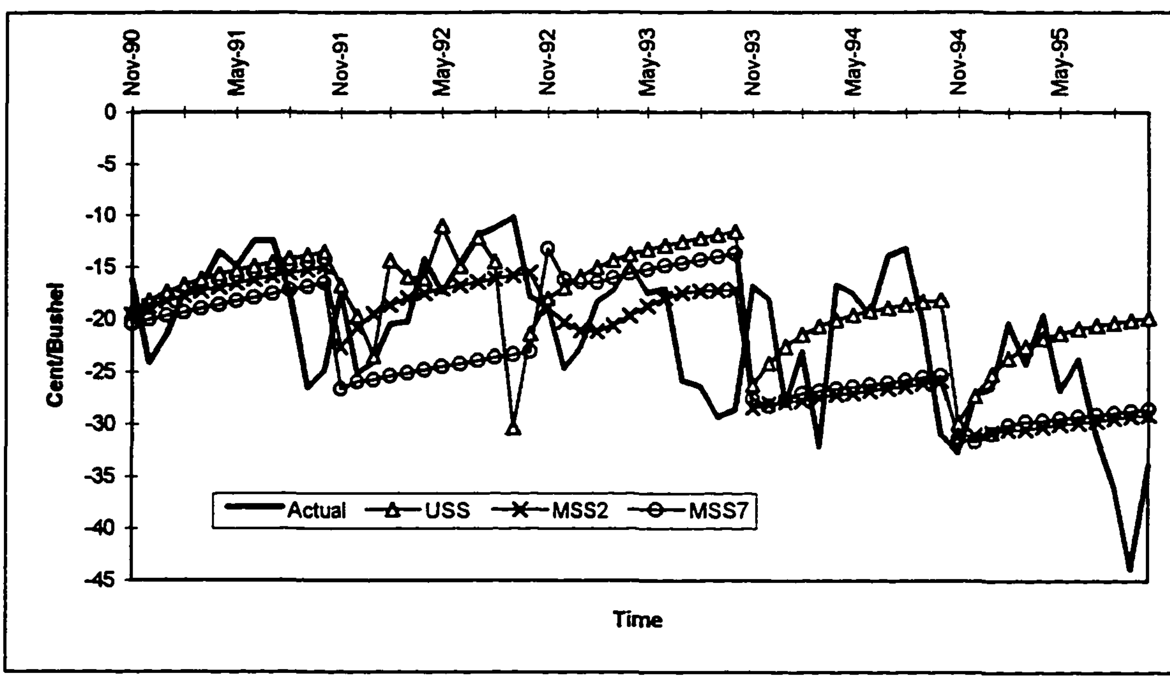

Figure 4.11. Northeast Iowa corn basis forecasts, state space approaches

Artificial neural network forecasts

Figure 4.12 plots the ANN with NFD algorithm forecasting for corn in Northeast Iowa. The ANN forecasts track the actual basis pretty well for the first two years in the corn model, but is off in 1995. Compared to three-year-average forecasts, ANN has an obvious advantage in the first set of the 1-12 months ahead forecasts (for Nov., 1990 to Oct., 1991) in the corn model, and three-year-average forecasts are better for the second set of forecasts from Nov., 1991 to Oct., 1992). Neither forecast performs well for the later years.

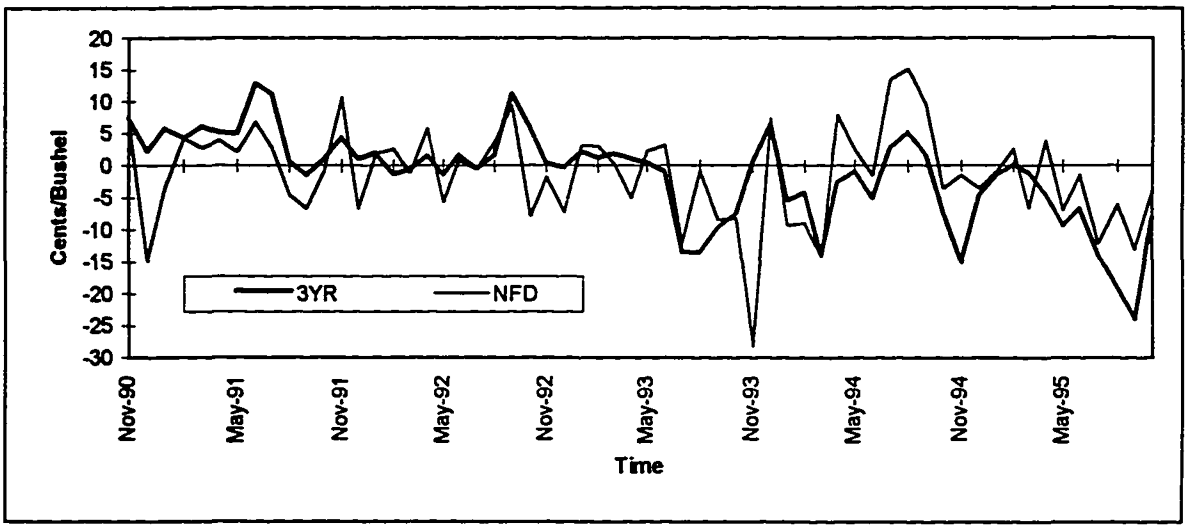

Figure 4.12. Northeast Iowa corn basis forecasts, neural networks vs. three-year-average.

Composite forecasts

A composite forecasting model is constructed as the simple average of selected forecasts: three-year-average-plus model, seasonal ARIMA model, univariate state space model, structural basis behavior model and neural network approach with noise-feedback descent algorithm. One model is chosen within similar approaches; that is, one out of three state space models and one out of two neural network approaches. Compared to three-year-average forecasts, the composite forecasts have smaller forecast errors for the first set of 1-12 month ahead forecasts (see Figure 4.13), but larger for the second and third years of corn basis forecasts in Northeast Iowa. They are close for the last two years (both are poor).

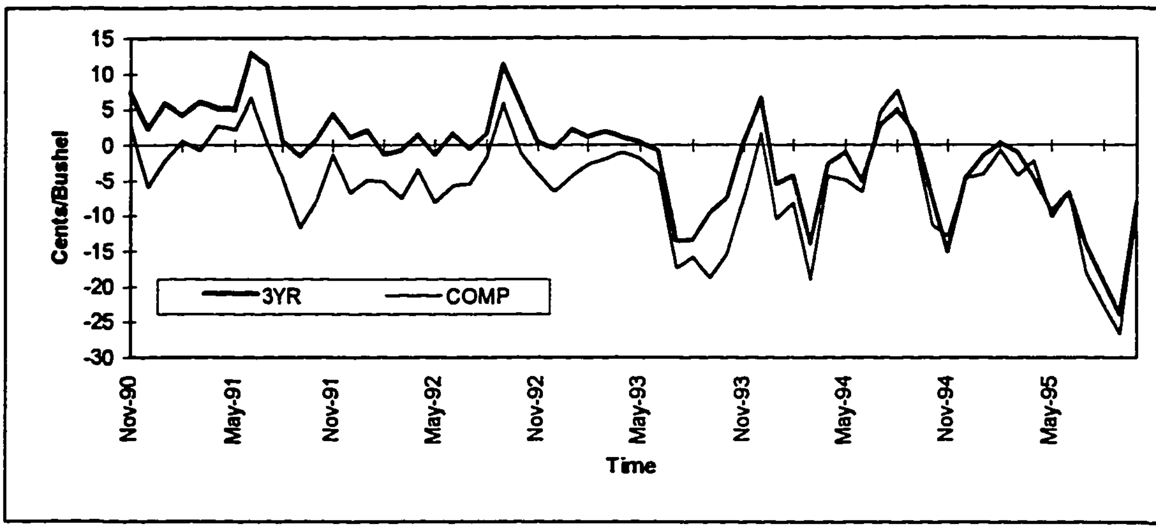

Figure 4.13. Northeast Iowa corn basis forecast errors, composite vs. 3-year average

Most forecasting models discussed above outperform the simple three-year-average corn basis forecasts in Northeast Iowa for the first and last sets of 1-12 months ahead (November to October) forecasts. Two models, the seasonal ARIMA and 2-crop multivariate state space models, have lower forecast errors than the three-year-average for the second forecasting period (11/91-10/92). One reason for these models not being able to perform well in third and fourth sets of forecasting (11/92-10/93 and 11/93-10/94) is due to unusual basis patterns before the forecast is made. The Northeast Iowa corn basis was unusually strong from June to October of 1992. These forecasting models tend to utilize that basis in their forecasting models, and result in bad forecasts for the next crop year. The unusually wide basis in the summer of 1993 (partially due to the flooding) also causes inaccurate forecasts by these models for the basis from November 1993 to October 1994. On the other hand, some of these unusual basis do not affect the simple three-year-average model forecasts as much because the latter consider only basis in t-12, t-24 and t-36.

Overall, we can conclude that three-year-average-plus and seasonal ARIMA are preferred to forecast Northeast Iowa corn basis when "normal" conditions prevail before the forecasting time (especially few months before October in this case). When unusual basis pattern occurs during the prior summer, the best forecast model will be the simple-three-year-average forecasts. The composite forecast has large forecast errors because of the extremely poor performance of the structural basis forecasting model. If the composite forecast were dropped, the composite of simple three-year-average, three-year-average-plus, and seasonal ARIMA models would have had improved performance, and could be a good potential forecasting model choice for Northeast Iowa.

CHAPTER 5. CONCLUSIONS

Corn and soybean nearby basis behavior in Northeast Iowa, Central Illinois, Chicago, St. Louis, Toledo, Gulf, Richmond, and Pacific Northwest (for corn only) are analyzed in this study. Theoretically the basis is determined by storage cost (opportunity cost), transportation cost (barge rates), supply factor (supply relative to storage capacity), demand factors (port exports, animal units consuming grain, and soybean crushing), and time to maturity of futures contract. Because of basis seasonal patterns and variance differences among contracts, this structural basis behavior model is estimated for each contract.

An autoregressive estimation procedure is selected in the presence of autocorrelation; otherwise Ordinary Least Squares is employed. Seemingly unrelated regression (either 2-crops at the same market or 7-markets for the same crop) is also estimated, but it does not provide improved results over Ordinary Least Squares. This is not consistent with Kahl and Curtis, who found that seemingly unrelated regression was superior than ordinary least square estimates. It might be that their independent variables for the models in South Carolina and Illinois were less correlated than the independent variables in the structural basis behavior models estimated here for seven or eight locations.

Individual models for each contract in this study provide more information about the differences in monthly corn and soybean basis relationships throughout the year, and generally explain 50-80% of the monthly variation in corn and soybean nearby basis. Kahl and Curtis' basis behavior model's R^2 was 0.48, which they argued was consistent with models from similar studies. Power and Johnson's Wisconsin corn basis behavior model had an R-square higher than 0.8, but their basis model for the whole storage season is a single July contract model and is a stacked model with a monthly trend dummy, which partially explains the higher R^2.

The results of these corn and soybean basis behavior models are generally consistent with previous studies, though a few interesting differences are found. The behavioral effects of the explanatory variables vary seasonally, and both slope and intercept differences are significant. This is consistent with, but more comprehensive than, prior studies (Kahl and Curtis, Garcia and Good, Power and Johnson).

Storage cost (or opportunity cost, equal to prime interest rate times cash price) is important in the early storage season, and is less significant as the storage season progresses. This is not consistent with the findings of Garcia and Good who showed that interest rate, which differs from our "opportunity" cost variable, was significant throughout the year. Powers and Johnson also found that interest rate was significant in determining Wisconsin corn basis for the whole storage season.

Barge rates significantly affect corn and soybean basis in the spring and fall. When the upper Mississippi river is closed during the winter, the grain bases in St. Louis and the Gulf respond to barge rate changes positively, different than other times of the year. In contrast, Garcia and Good found that barge rates only matter during May to July.

Production relative to storage capacity is found to significantly affect basis only for months during and immediately after harvest. This is consistent with the results of Garcia and Good, who found that production relative to storage capacity had strong effects on basis during harvest period. In their study of Ontario corn basis, Martin et al. found that both U.S.

and Canada crops are important for the fall months (November and December), but only Canada production relative to demand mattered for January through April.

Demand factors have mixed results in the present study. Export volumes usually do not affect basis. Animal units consuming grain are only significant for few contracts and few markets. This is consistent with Kahl and Curtis who found that animal units to be significant in South Carolina but not in Illinois.

Significant basis convergence is found in December to February for corn and November to December for soybeans. No evidence in this study suggests behavioral differences between CBOT delivery and non-delivery markets, and delivery and non-delivery months. This is in contrast with the results of Thompson et al., who showed that cash and futures price behavior differed between delivery and non-delivery months.

Regional basis behavior differences are found between production areas and port. The opportunity cost coefficient is found to be negative as expected for production market locations. But in Richmond and Pacific NW, the opportunity cost coefficient become positive in some spring and fall periods. For the production relative to storage capacity, a expected negative coefficient is found for all the markets except Pacific NW where a significant positive coefficient is found.

Basis variability was larger in 1980-86 than in the following decade. There was a upward trend before 1986, and a more stable pattern after 1986. Structural change tests provide mixed results. Further exploration with different model specifications suggests that some parameters vary over time, others are consistently important. In general, the estimation results of alternate specifications are consistent with that of the original basis structure model analyzed.

For basis forecasting, in addition to the conventional approaches--simple three-year-average forecasts, ARIMA, and econometric structural models, alternative forecast methods are also utilized. They are three-year-average-plus model, state space modeling approaches (univariate and multivariate), artificial neural networks and composite forecasts. For each model, out-of-sample basis forecasts of 1-12 months ahead (made in October each year) are made for five years, 1991-1995. With simple three-year-average forecasts as bench mark forecasts, the forecast performances of other models are compared with five forecast accuracy measures: mean absolute errors (MAE), root mean squared errors (RMSE), two Theils' U coefficients and a Henriksson-Merton test.

Basis forecasting performance comparisons show that the simple three-year-average forecasts can be outperformed by alternative models, but the marginal improvement is small and the overall ability of any method to forecast basis 1-12 months ahead is not good. Overall comparisons show 3-year-average-plus and seasonal ARIMA models are the best among the alternative methods studied in this paper. Other more complicated approaches, the state space and structural econometric models, have inconsistent and often poor performance. For example, the structural behavior model performs very well in 1-4 month ahead corn basis forecasting, but not for longer term forecasts (contradicting what is expected from prior studies). This may suggest that there is some room for improvement in the area of ancillary forecasts. Our results were similar to Taylor and Tomek who utilized an econometric forecasting model, and showed that the basis forecasts based on ancillary forecasts differed a lot from the in-sample forecasts. Based on this analysis, econometric models have serious limitations for basis forecasting. The composite forecast also does a reasonably good job, but it has the least practical value because it involves so many forecasting methods. In the future, a composite forecast of a few simple forecast models could be considered.

In conclusion, 3-year-average-plus and seasonal ARIMA models outperform simple 3-year-average forecasts and are most practical and easier to use than other alternative models. But, the reductions in forecast errors (RMSE) are in the range of 0.5-2 cents per bushel for corn (0.5-4 cents per bushel for soybeans); the value of this marginal improvement may justify the cost of developing and using these forecast models for some large volume grain merchandisers, but not be great enough for others.

APPENDIX A. SKETCH OF BASIS MODEL DERIVATION

Local basis, the price difference between a local spot price and futures price, has two components: temporal and spatial price differences. The temporal price difference is defined as the storage cost (price of storage) over a time interval, and spatial price relationship is mainly the transportation cost. The temporal price relationship can be derived from Tomek's (1996) simple intrayear two-period model. The profit maximization rule is used to derive the supply of storage equation from a short-run profit function:

$$R_2 = F_1 I_1 - P_1 I_1 - C(I_1)$$

where R, F, P and I represents revenue, futures price, cash price and inventory, respectively. $C(I_1)$, the cost function, is specified as:

$$C(I_1) = i P_1 I_1 + dI_1 - v \ln I_1$$

where i is the interest rate. The right side is composed of the opportunity cost of carrying stock, short-run costs, and the convenience yield of carrying stocks. First order condition of the revenue function, with the cost function being substituted in, gives the supply of storage equation:

(EQ A.1) $$P_1 - F_1 = v/I_1 - (iP_1 + d)$$

where $P_1 - F_1$ is the price of storage over a time interval.

The demand for storage equation is derived from consumption demands in two periods. Assuming that production, S_1, is fixed and consumed in two periods completely, he arrived at the demand for storage equation:

(EQ A.2) $$I_1 = (1/2)S_1 + (1/2)(a_2 - a_1) - (b/2)(P_1 - F_1)$$

where S_1[a] is the production, $(a_2 - a_1)$ is the expected demand for consumption in period two relative to the demand in period one for current consumption.

---

[a] The production term, S_1, is missing in Tomek's derivation

Equating supply of and demand for storage, we get the following equation:

$$-(b/2)(P_1 - F_1)^2 - (1/2)[(S_1 + a_2 - a_1 - (iP + d)b](P_1 - F_1) +$$

$$(1/2)(iP - d)(S_1 + a_2 - a_1) - v = 0$$

The price of storage $(P_1 - F_1)$ reduced form can be derived from this quadratic equation:

$$(P_1 - F_1) = f(S_1, iP, a_2 - a_1, d, v)$$

That is, the price of storage is a function of production, opportunity cost, relative demand in two periods, direct storage cost, and convenience yield. Assuming the second component, transportation cost as given as exogenous, then the local basis can be expressed as:

Basis = f (storage cost, transportation cost, production, stock, consumption demand)

where stock is in the place of convenience yield which can not be measured, and it is believed that convenience yield depends on the level of stock.

APPENDIX B. ADDITIONAL TABLES

Table B.1. RMSE and R squares of three stacked model specifications

	(1)Month Intercept Dummies	(2)Contract Intercept Dummies	(3) Contract Intercept and Slope Dummies
Corn			
		Root MSE	
Chicago	7.42	7.34	7.31
Gulf Port	9.52	9.45	8.55
Richmond	11.23	11.44	11.28
Northeast IA	9.48	9.46	9.16
St Louis	9.21	9.09	8.90
Toledo	7.91	7.92	7.88
Pacific NW	6.66	6.57	5.99
Central IL	8.38	8.29	8.31
		R Square	
Chicago	0.46	0.46	0.52
Gulf Port	0.27	0.25	0.47
Richmond	0.41	0.37	0.47
Northeast IA	0.62	0.61	0.68
St Louis	0.43	0.42	0.51
Toledo	0.56	0.54	0.60
Pacific NW	0.65	0.65	0.76
Central IL	0.53	0.53	0.58
Soybean			
		Root MSE	
Chicago	6.77	7.30	7.37
Gulf Port	10.53	10.56	8.79
Richmond	9.32	10.10	10.10
Northeast IA	9.71	10.03	8.86
St. Louis	8.54	8.55	8.09
Toledo	7.17	7.63	6.94
Central IL	7.48	7.49	6.63
		R Square	
Chicago	0.64	0.56	0.66
Gulf Port	0.38	0.34	0.69
Richmond	0.62	0.50	0.75
Northeast IA	0.73	0.70	0.83
St. Louis	0.57	0.56	0.69
Toledo	0.55	0.47	0.67
Central IL	0.64	0.63	0.78

Table B.2. Parameter estimates of corn basis contract behavior model for Northeast Iowa and Central Illinois

		Northeast IA		Central IL	
Contract	Variable	Par Est	P_value	Par Est	P_value
March	R^2	0.80		0.82	
	Intercept	9.534	0.784	-14.437	0.629
	Opp. cost	-0.800	0.000	-0.502	0.000
	Barge rate	-0.003	0.954	0.036	0.069
	Production	-7.691	0.026	-7.623	0.008
	Export	0.546	0.314	0.151	0.681
	AUC_G	-0.097	0.827	0.277	0.462
	TTM	-3.422	0.001	-2.865	0.000
May	R^2	0.78		0.82	
	Intercept	56.014	0.232	40.572	0.271
	Opp. cost	-0.558	0.001	-0.550	0.000
	Barge rate	-0.104	0.067	-0.024	0.502
	Production	-0.285	0.935	-1.092	0.688
	Export	-0.197	0.657	-0.273	0.411
	AUC_G	-0.567	0.325	-0.346	0.445
	TTM	-3.733	0.035	-2.092	0.104
July	R^2	0.72		0.62	
	Intercept	46.950	0.354	27.481	0.589
	Opp. cost	-0.296	0.053	-0.280	0.059
	Barge rate	-0.210	0.001	-0.107	0.037
	Production	-3.099	0.445	-1.794	0.671
	Export	-0.369	0.484	-0.659	0.221
	AUC_G	-0.356	0.571	-0.103	0.871
	TTM	-1.207	0.532	-0.307	0.871
September	R^2	0.82		0.72	
	Intercept	104.540	0.224	65.566	0.447
	Opp. cost	-0.423	0.055	-0.367	0.089
	Barge rate	-0.162	0.005	-0.090	0.061
	Production	-13.020	0.035	-6.272	0.331
	Export	0.661	0.491	0.756	0.426
	AUC_G	-1.017	0.362	-0.620	0.580
	TTM	0.261	0.907	1.084	0.598
December	R^2	0.64		0.61	
	Intercept	-75.328	0.213	-117.138	0.018
	Opp. cost	-0.501	0.069	0.002	0.991
	Barge rate	-0.184	0.000	-0.145	0.000
	Production	5.692	0.281	0.695	0.874
	Export	-0.320	0.573	-0.171	0.702
	AUC_G	1.054	0.166	1.654	0.009
	TTM	1.225	0.525	-0.032	0.983

Table B.3. Parameter estimates of corn basis contract behavior model for Chicago, St. Louis and Toledo

		Chicago		St. Louis		Toledo	
Contract	Variable	B Value	P_value	B Value	P_value	B Value	P_value
March	R^2	0.70		0.66		0.70	
	Intercept	11.010	0.690	43.217	0.203	50.520	0.102
	Opp. cost	-0.306	0.001	-0.469	0.000	-0.385	0.000
	Barge rate	0.015	0.437	0.090	0.039	0.052	0.168
	Production	-4.862	0.041	-8.745	0.012	-11.606	0.001
	Export	-9.149	0.304	0.349	0.492	3.237	0.355
	AUC_G	0.075	0.828	-0.324	0.450	-0.457	0.239
	TTM	-2.531	0.004	-2.893	0.003	-3.344	0.002
May	R^2	0.64		0.78		0.66	
	Intercept	28.888	0.423	54.401	0.131	61.711	0.056
	Opp. cost	-0.150	0.207	-0.409	0.001	-0.272	0.018
	Barge rate	-0.056	0.102	-0.019	0.644	-0.050	0.233
	Production	-4.473	0.113	-1.855	0.492	-5.780	0.055
	Export	-7.541	0.100	-0.154	0.629	0.859	0.775
	AUC_G	-0.047	0.916	-0.393	0.370	-0.573	0.150
	TTM	-3.220	0.048	-1.709	0.166	-2.418	0.186
July	R^2	0.56		0.55		0.43	
	Intercept	22.236	0.595	53.004	0.292	58.014	0.119
	Opp. cost	-0.042	0.754	-0.214	0.135	0.010	0.949
	Barge rate	-0.170	0.003	-0.139	0.026	-0.188	0.014
	Production	-2.564	0.509	-3.053	0.486	-4.281	0.344
	Export	5.007	0.569	-0.363	0.508	-4.035	0.477
	AUC_G	0.105	0.845	-0.252	0.686	-0.351	0.461
	TTM	-2.905	0.105	-0.955	0.643	-3.529	0.168
September	R^2	0.62		0.60		0.61	
	Intercept	48.119	0.553	140.357	0.110	133.679	0.091
	Opp. cost	-0.139	0.457	-0.336	0.128	-0.162	0.431
	Barge rate	-0.084	0.105	-0.065	0.283	-0.001	0.980
	Production	-6.253	0.300	-10.218	0.154	-6.800	0.354
	Export	4.191	0.534	0.933	0.390	-11.888	0.124
	AUC_G	-0.272	0.795	-1.436	0.206	-1.517	0.141
	TTM	1.657	0.437	0.554	0.831	5.095	0.109
December	R^2	0.55		0.57		0.60	
	Intercept	-23.166	0.519	-100.007	0.067	-15.881	0.709
	Opp. cost	-0.022	0.896	0.084	0.696	-0.103	0.547
	Barge rate	-0.085	0.004	-0.186	0.000	-0.107	0.001
	Production	-0.860	0.787	-0.210	0.968	0.886	0.834
	Export	-5.983	0.244	-0.095	0.861	-4.199	0.133
	AUC_G	0.525	0.250	1.706	0.015	0.302	0.565
	TTM	-1.603	0.166	-2.184	0.244	2.187	0.164

Table B.4. Parameter estimates of corn basis contract behavior model for Gulf ports, Pacific NW and Richmond

		Gulf		Pacific NW		Richmond	
Contract	Variable	B Value	P_value	B Value	P_value	B Value	P_value
March	R^2	0.72		0.58		0.84	
	Intercept	54.394	0.147	-59.634	0.016	97.517	0.028
	Opp. cost	-0.536	0.000	0.217	0.212	-0.266	0.075
	Barge rate	0.214	0.000	0.036	0.392	0.058	0.129
	Production	-6.663	0.148	23.631	0.019	-25.042	0.000
	Export	0.130	0.790	2.009	0.032	2.312	0.141
	AUC_G	-0.394	0.406	1.087	0.002	-0.584	0.306
	TTM	-3.255	0.001	-2.166	0.007	-5.484	0.000
May	R^2	0.62		0.66		0.57	
	Intercept	65.088	0.101	-62.048	0.208	104.509	0.015
	Opp. cost	-0.383	0.002	0.303	0.331	0.270	0.032
	Barge rate	0.074	0.121	0.046	0.630	-0.059	0.170
	Production	2.403	0.546	33.856	0.076	-4.265	0.400
	Export	-0.158	0.671	-0.003	0.999	-4.541	0.009
	AUC_G	-0.505	0.301	1.034	0.152	-1.066	0.061
	TTM	-1.507	0.291	-3.636	0.310	2.029	0.284
July	R^2	0.42		0.45		0.81	
	Intercept	67.730	0.190	-91.354	0.004	108.059	0.177
	Opp. cost	-0.274	0.054	0.319	0.104	0.298	0.208
	Barge rate	-0.026	0.664	-0.163	0.008	-0.174	0.057
	Production	0.640	0.910	26.811	0.036	5.609	0.550
	Export	-0.037	0.945	2.169	0.038	-8.739	0.055
	AUC_G	-0.445	0.489	1.593	0.001	-1.053	0.325
	TTM	-1.285	0.522	-0.540	0.811	-2.820	0.504
September	R^2	0.60		0.49		0.69	
	Intercept	113.019	0.197	-87.347	0.054	198.153	0.036
	Opp. cost	-0.384	0.055	1.387	0.002	-0.395	0.145
	Barge rate	0.045	0.403	-0.063	0.276	-0.031	0.663
	Production	-15.820	0.054	48.034	0.090	-6.258	0.540
	Export	0.994	0.297	-1.770	0.461	-4.011	0.572
	AUC_G	-0.951	0.402	1.131	0.099	-2.210	0.075
	TTM	2.532	0.253	-3.617	0.305	9.478	0.032
December	R^2	0.47		0.24		0.80	
	Intercept	-100.223	0.065	-108.945	0.002	69.753	0.300
	Opp. cost	0.011	0.958	0.732	0.002	-0.172	0.519
	Barge rate	-0.108	0.006	-0.033	0.271	0.018	0.679
	Production	6.118	0.356	23.689	0.070	-13.369	0.121
	Export	0.157	0.770	1.587	0.287	-0.120	0.968
	AUC_G	1.649	0.019	1.651	0.001	-0.569	0.519
	TTM	0.799	0.662	-1.884	0.255	-2.030	0.405

Tabel B.5. Parameter Estimates of soybean basis contract model (1)

		Noreast IA		Central IL	
Contract	Variable	B Value	P Value	B Value	P Value
January	R^2	0.79		0.77	
	Intercept	24.990	0.660	5.680	0.900
	Opp. cost	-0.322	0.006	-0.324	0.001
	Barge rate	-0.111	0.069	0.007	0.853
	Production	-42.725	0.309	-3.233	0.915
	Export	0.448	0.846	-2.685	0.171
	AUC_G	0.084	0.892	0.520	0.321
	Crushing	-0.026	0.518	-0.022	0.536
	TTM	-3.051	0.247	-7.689	0.000
March	R^2	0.86		0.84	
	Intercept	-14.900	0.784	-2.593	0.942
	Opp. cost	-0.441	0.000	-0.280	0.001
	Barge rate	0.093	0.205	0.027	0.458
	Production	-98.589	0.037	-18.247	0.503
	Export	-1.810	0.317	-0.703	0.591
	AUC_G	1.091	0.112	0.863	0.061
	Crushing	-0.073	0.074	-0.087	0.009
	TTM	-3.915	0.279	-1.115	0.699
May	R^2	0.60		0.83	
	Intercept	-30.018	0.631	-34.705	0.285
	Opp. cost	-0.105	0.291	-0.104	0.043
	Barge rate	-0.168	0.100	-0.047	0.206
	Production	-24.042	0.611	-63.969	0.004
	Export	0.767	0.666	0.609	0.373
	AUC_G	1.274	0.134	0.954	0.033
	Crushing	-0.111	0.036	-0.032	0.157
	TTM	-1.350	0.788	-1.534	0.405
July	R^2	0.65		0.67	
	Intercept	6.189	0.928	-14.265	0.782
	Opp. cost	-0.043	0.640	-0.096	0.152
	Barge rate	-0.347	0.000	-0.199	0.001
	Production	5.011	0.926	-3.982	0.915
	Export	-1.097	0.623	-0.479	0.774
	AUC_G	0.719	0.431	0.816	0.243
	Crushing	-0.090	0.050	-0.056	0.104
	TTM	-1.125	0.743	-1.121	0.680
Aug&Sep	R^2	0.75		0.66	
	Intercept	72.591	0.245	100.321	0.029
	Opp. cost	-0.208	0.049	-0.102	0.116
	Barge rate	-0.262	0.000	-0.012	0.736
	Production	-6.676	0.895	-97.321	0.005
	Export	-2.775	0.216	-2.143	0.197
	AUC_G	-0.527	0.498	-0.740	0.191
	Crushing	-0.003	0.898	0.002	0.930
	TTM	6.145	0.050	1.884	0.371
November	R^2	0.84		0.73	
	Intercept	51.625	0.314	53.558	0.290
	Opp. cost	-0.452	0.000	-0.287	0.005
	Barge rate	-0.141	0.002	-0.062	0.142
	Production	15.135	0.670	-10.051	0.744
	Export	-2.069	0.099	-1.976	0.096
	AUC_G	0.077	0.906	0.164	0.801
	Crushing	-0.047	0.076	-0.053	0.060
	TTM	-6.140	0.145	-7.029	0.125

Tabel B.6. Parameter Estimates of soybean basis contract model (1)

		Chicago		Toledo		St. Louis	
Contract	Variable	B Value	P Value	B Value	P Value	B Value	P Value
January	R^2	0.75		0.73		0.78	
	Intercept	92.295	0.046	69.185	0.208	89.561	0.082
	Opp. cost	-0.254	0.003	-0.275	0.003	-0.206	0.019
	Barge rate	-0.037	0.282	0.025	0.682	-0.059	0.244
	Production	2.325	0.945	-19.019	0.465	-2.871	0.922
	Export	-0.751	0.680	-9.539	0.137	0.531	0.786
	AUC_G	-0.399	0.420	-0.546	0.571	-0.505	0.372
	Crushing	-0.047	0.156	-0.027	0.810	-0.022	0.513
	TTM	-8.619	0.000	-7.919	0.004	-9.975	0.000
March	R^2	0.82		0.38		0.82	
	Intercept	115.063	0.004	84.581	0.055	108.294	0.007
	Opp. cost	-0.334	0.000	-0.050	0.525	-0.217	0.002
	Barge rate	0.038	0.231	-0.012	0.816	0.096	0.040
	Production	7.469	0.795	-46.500	0.058	-44.167	0.093
	Export	-1.463	0.184	-1.111	0.106	-0.188	0.871
	AUC_G	-0.789	0.087	0.142	0.121	-0.499	0.274
	Crushing	-0.057	0.030	-9.468	0.000	-0.072	0.013
	TTM	-3.686	0.111			-1.165	0.627
May	R^2	0.83		0.66		0.77	
	Intercept	47.474	0.024	111.169	0.035	20.400	0.573
	Opp. cost	-0.133	0.000	0.018	0.831	-0.021	0.678
	Barge rate	-0.054	0.039	-0.103	0.079	-0.056	0.205
	Production	1.859	0.901	-61.905	0.016	-74.138	0.002
	Export	0.165	0.737	2.106	0.693	1.190	0.106
	AUC_G	-0.261	0.330	-1.456	0.076	0.297	0.532
	Crushing	-0.025	0.108	0.192	0.047	-0.016	0.487
	TTM	-2.930	0.036	-4.832	0.035	-1.705	0.381
July	R^2	0.61		0.60		0.61	
	Intercept	52.132	0.247	89.001	0.014	62.580	0.290
	Opp. cost	-0.063	0.227	0.027	0.645	-0.032	0.668
	Barge rate	-0.144	0.001	-0.183	0.007	-0.269	0.001
	Production	10.970	0.739	-19.898	0.373	-16.232	0.697
	Export	-0.879	0.488	-5.301	0.434	-0.142	0.947
	AUC_G	-0.210	0.721	-1.112	0.045	0.077	0.921
	Crushing	-0.032	0.205	0.137	0.023	-0.056	0.153
	TTM	-2.197	0.247	-4.835	0.050	-1.158	0.701
Aug&Sep	R^2	0.62		0.60		0.61	
	Intercept	-53.308	0.418	148.431	0.009	145.814	0.030
	Opp. cost	0.038	0.670	-0.140	0.116	-0.252	0.021
	Barge rate	-0.140	0.005	-0.068	0.206	-0.119	0.056
	Production	-24.556	0.625	-44.418	0.293	48.343	0.322
	Export	0.105	0.960	2.458	0.852	-5.069	0.063
	AUC_G	1.029	0.228	-1.648	0.054	-1.377	0.093
	Crushing	-0.006	0.773	0.068	0.340	-0.010	0.703
	TTM	-0.543	0.881	1.689	0.634	6.131	0.081
November	R^2	0.45		0.59		0.79	
	Intercept	18.289	0.808	265.225	0.009	116.610	0.025
	Opp. cost	-0.161	0.242	-0.200	0.106	-0.289	0.002
	Barge rate	-0.131	0.037	-0.058	0.303	-0.092	0.023
	Production	1.574	0.976	-0.544	0.991	11.692	0.660
	Export	-0.848	0.636	-9.430	0.189	-1.716	0.116
	AUC_G	0.720	0.463	-4.309	0.011	-0.533	0.401
	Crushing	-0.042	0.288	0.400	0.009	-0.054	0.023
	TTM	-13.164	0.041	8.385	0.102	-7.665	0.055

Tabel B.7. Parameter Estimates of soybean basis contract model (3)

Contract	Variable	Gulf Ports B Value	Gulf Ports P Value	Richmond B Value	Richmond P Value
January	R^2	0.73		0.51	
	Intercept	65.542	0.304	13.551	0.857
	Opp. cost	-0.404	0.002	-0.238	0.078
	Barge rate	0.089	0.165	0.013	0.812
	Production	12.699	0.813	-72.564	0.083
	Export	-1.686	0.415	13.267	0.182
	AUC_G	0.521	0.581	0.019	0.990
	Crushing	-0.013	0.372	0.041	0.847
	TTM	-12.160	0.000	-11.852	0.005
March	R^2	0.72		0.49	
	Intercept	69.728	0.090	56.428	0.322
	Opp. cost	-0.437	0.000	0.011	0.929
	Barge rate	0.268	0.000	-0.002	0.972
	Production	-24.699	0.541	-71.184	0.051
	Export	-0.639	0.667	3.917	0.409
	AUC_G	0.529	0.342	-0.684	0.456
	Crushing	-0.027	0.004	0.131	0.371
	TTM	3.085	0.375	-11.088	0.003
May	R^2	0.14		0.55	
	Intercept	-14.076	0.808	44.370	0.378
	Opp. cost	-0.107	0.285	0.270	0.006
	Barge rate	0.047	0.628	-0.178	0.003
	Production	-11.462	0.805	-52.791	0.047
	Export	0.815	0.519	-1.595	0.796
	AUC_G	1.077	0.304	-1.149	0.169
	Crushing	-0.016	0.270	0.363	0.001
	TTM	1.678	0.609	-3.422	0.289
July	R^2	0.42		0.55	
	Intercept	-22.644	0.765	58.183	0.322
	Opp. cost	-0.028	0.822	0.040	0.663
	Barge rate	-0.171	0.115	-0.202	0.004
	Production	-16.421	0.828	43.517	0.250
	Export	-2.295	0.348	-8.156	0.464
	AUC_G	1.132	0.383	-1.466	0.115
	Crushing	-0.003	0.852	0.268	0.007
	TTM	-3.089	0.574	1.071	0.755
Aug&Sep	R^2	0.46		0.56	
	Intercept	113.026	0.234	3.375	0.958
	Opp. cost	-0.255	0.056	0.242	0.020
	Barge rate	0.078	0.262	-0.032	0.601
	Production	-60.165	0.501	-120.306	0.011
	Export	-4.497	0.146	-7.901	0.748
	AUC_G	-0.909	0.574	0.190	0.840
	Crushing	0.007	0.646	0.008	0.931
	TTM	4.588	0.372	1.167	0.812
November	R^2	0.55		0.31	
	Intercept	34.090	0.640	134.963	0.080
	Opp. cost	-0.347	0.006	-0.109	0.391
	Barge rate	0.026	0.658	-0.058	0.302
	Production	17.339	0.746	35.951	0.350
	Export	-1.150	0.529	-31.486	0.017
	AUC_G	0.837	0.474	-1.760	0.149
	Crushing	-0.017	0.199	0.092	0.507
	TTM	-6.679	0.311	-14.571	0.123

Table B.8. Parameter Estimates of three-year-average-plus forecasting model for Chicago

	Data set: 80-Oct/90		Data set: 80-Oct/91		Data set: 80-Oct/90	
VARIABLE	Est.Coef.	P-value	Est.Coef.	P-value	Est.Coef.	P-value
BS3CHG	-0.126	0.386	0.006	0.965	0.068	0.580
CNSUPPLY	-0.368	0.000	-0.345	0.000	-0.309	0.000
CNEXPORT	-1.018	0.008	-0.689	0.045	-0.374	0.209
DM1	5.523	0.170	4.593	0.219	4.133	0.226
DM2	9.594	0.024	7.890	0.043	6.971	0.050
DM3	8.564	0.043	5.674	0.142	4.782	0.171
DM4	9.869	0.020	8.035	0.039	7.347	0.038
DM5	11.591	0.008	8.158	0.040	7.090	0.048
DM6	11.362	0.011	8.895	0.027	7.875	0.031
DM7	19.222	0.000	15.328	0.001	13.411	0.001
DM8	13.683	0.003	10.823	0.009	9.586	0.011
DM9	1.404	0.719	0.736	0.840	0.732	0.827
DM11	6.533	0.118	5.900	0.129	5.519	0.119
DM12	3.705	0.369	2.714	0.478	2.357	0.499
CONSTANT	49.395	0.000	42.296	0.000	33.359	0.000
R^2	.47		.43		.40	

	Data set: 80-Oct/90		Data set: 80-Oct/90	
VARIABLE	Est.Coef	P-value	Est.Coef	P-value
BS3CHG	0.085	0.459	0.115	0.287
CNSUPPLY	-0.312	0.000	-0.279	0.000
CNEXPORT	-0.340	0.229	-0.091	0.682
DM1	4.187	0.179	4.454	0.125
DM2	6.811	0.035	7.488	0.013
DM3	4.619	0.145	5.160	0.080
DM4	7.302	0.024	8.124	0.007
DM5	6.854	0.035	7.573	0.012
DM6	7.649	0.021	7.718	0.012
DM7	12.142	0.001	12.120	0.000
DM8	8.927	0.009	9.005	0.004
DM9	0.560	0.854	1.030	0.718
DM11	4.995	0.121	6.000	0.046
DM12	2.282	0.471	2.881	0.329
CONSTANT	33.151	0.000	24.840	0.000
R^2	.39		.39	

Table B.9. Parameter estimates of structural basis behavior model for Chicago corn

Variable	Par Est	P-value	Par Est	P-value	Par Est	P-value	Par Est	P-value	Par Est	P-value
Data set	80-10/90		80-10/91		80-10/92		80-10/93		80-10/94	
					March					
R^2	.59		.73		.80		.42		.49	
INTER	123.98	0.07	71.53	0.11	62.17	0.08	53.35	0.06	40.42	0.09
RP_CHG	-0.34	0.00	-0.33	0.00	-0.33	0.00	-0.35	0.00	-0.34	0.00
BR_ILR	0.02	0.67	0.00	0.93	0.00	0.99	0.00	0.91	0.00	0.99
CN_CHG	-3.85	0.17	-3.14	0.24	-3.10	0.22	-2.84	0.22	-3.20	0.16
EX_CHG	-19.24	0.41	-24.97	0.24	-24.46	0.22	-17.98	0.28	-17.49	0.28
AUC_G	-1.43	0.11	-0.73	0.20	-0.60	0.18	-0.49	0.17	-0.30	0.29
TTM	-2.07	0.27	-1.75	0.28	-1.65	0.27	-1.80	0.20	-1.92	0.15
					May					
R^2	.58		.68		.77		.42		.48	
INTER	92.62	0.12	126.78	0.01	102.62	0.01	85.38	0.01	70.36	0.01
RP_CHG	-0.08	0.40	-0.07	0.48	-0.09	0.36	-0.10	0.27	-0.08	0.37
BR_ILR	-0.17	0.00	-0.15	0.00	-0.15	0.00	-0.15	0.00	-0.16	0.00
CN_CHG	-5.39	0.06	-5.48	0.06	-5.18	0.07	-4.27	0.11	-4.80	0.06
EX_CHG	-1.33	0.83	-2.49	0.68	-3.26	0.58	-4.33	0.45	-4.88	0.38
AUC_G	-0.80	0.31	-1.26	0.03	-0.93	0.05	-0.70	0.07	-0.48	0.13
TTM	1.72	0.52	0.11	0.97	-0.23	0.92	-0.84	0.71	-1.10	0.60
					July					
R^2	.58		.66		.78		.41		.49	
INTER	105.43	0.11	139.38	0.01	124.26	0.00	115.56	0.00	111.96	0.00
RP_CHG	0.01	0.93	0.02	0.84	0.02	0.88	0.01	0.92	0.01	0.93
BR_ILR	-0.29	0.00	-0.27	0.00	-0.28	0.00	-0.28	0.00	-0.28	0.00
CN_CHG	-4.36	0.14	-4.63	0.12	-4.58	0.11	-4.30	0.11	-4.39	0.09
EX_CHG	-1.43	0.86	-1.46	0.86	-1.47	0.85	-1.42	0.85	-1.22	0.87
AUC_G	-0.73	0.41	-1.19	0.07	-0.98	0.07	-0.88	0.04	-0.84	0.02
TTM	-4.27	0.08	-5.02	0.04	-5.04	0.02	-4.65	0.02	-4.04	0.04
					September					
R^2	.57		.64		.76		.40		.48	
INTER	100.27	0.61	35.97	0.79	77.51	0.42	48.98	0.53	46.60	0.45
RP_CHG	-0.02	0.92	-0.06	0.80	0.00	0.99	-0.01	0.97	-0.01	0.97
BR_ILR	-0.22	0.08	-0.23	0.04	-0.20	0.02	-0.19	0.01	-0.19	0.01
CN_CHG	-0.50	0.95	0.53	0.94	-0.33	0.96	0.52	0.93	0.49	0.94
EX_CHG	13.38	0.50	14.13	0.45	8.41	0.56	5.69	0.61	5.58	0.60
AUC_G	-0.88	0.74	0.00	1.00	-0.60	0.63	-0.26	0.80	-0.23	0.76
TTM	-0.75	0.90	-0.98	0.86	-0.67	0.90	0.15	0.97	0.39	0.92
					December					
R^2	.57		.65		.75		.40		.49	
INTER	-75.89	0.38	-59.91	0.37	-48.35	0.35	-22.03	0.61	-20.78	0.57
RP_CHG	0.12	0.55	0.09	0.62	0.10	0.59	0.09	0.60	0.09	0.58
BR_ILR	-0.11	0.05	-0.10	0.05	-0.09	0.04	-0.09	0.04	-0.08	0.01
CN_CHG	-3.23	0.43	-3.38	0.38	-3.56	0.30	-3.58	0.27	-3.48	0.25
EX_CHG	-9.15	0.23	-9.34	0.18	-9.87	0.13	-10.22	0.11	-10.50	0.06
AUC_G	1.30	0.25	1.06	0.22	0.90	0.18	0.55	0.32	0.52	0.26
TTM	-2.21	0.29	-1.72	0.36	-1.69	0.33	-1.79	0.27	-1.69	0.26

Table B.10. St. Louis barge rates forecasting equations

Data set: Jan.80-Oct.90
Model: SARIMA $(2,0,2)(2,2,0)_{12}$

	PAR EST	STD ERROR	T-STAT
AR(1)	0.201	0.173	1.160
AR(2)	0.609	0.133	4.579
MA(1)	-0.603	0.184	-3.275
MA(2)	0.271	0.132	2.049
SAR(1)	-0.909	0.082	-11.110
SAR(2)	-0.930	0.033	-27.930
CONSTANT	-1.596	2.969	-0.537
R^2	0.729		
Adjusted R^2	0.712		

Data set: Jan.80-Oct.92
Model: SARIMA $(2,0,1)(2,2,0)_{12}$

	PAR EST	STD ERROR	T-STAT
AR(1)	-0.052	0.107	-0.489
AR(2)	0.645	0.104	6.226
MA(1)	-0.929	0.060	-15.550
SAR(1)	-0.773	0.070	-11.100
SAR(2)	-0.901	0.037	-24.440
CONSTANT	-2.415	3.958	-0.610
R^2	0.710		
Adjusted R^2	0.699		

Data set: Jan.80-Oct.91
Model: SARIMA $(2,0,1)(2,2,0)_{12}$

	PAR EST	STD ERROR	T-STAT
AR(1)	-0.007	0.143	-0.048
AR(2)	0.631	0.132	4.771
MA(1)	-0.904	0.098	-9.246
SAR(1)	-0.820	0.076	-10.820
SAR(2)	-0.920	0.035	-26.460
CONSTANT	-2.118	4.061	-0.522
R^2	0.695		
Adjusted R^2	0.682		

Data set: Jan.80-Oct.93
Model: SARIMA $(2,0,2)(2,2,0)_{12}$

	PAR EST	STD ERROR	T-STAT
AR(1)	-0.072	0.095	-0.761
AR(2)	0.648	0.093	7.005
MA(1)	-0.935	0.051	-18.210
SAR(1)	-0.739	0.065	-11.330
SAR(2)	-0.899	0.036	-25.100
CONSTANT	-2.283	3.712	-0.615
R^2	0.706		
Adjusted R^2	0.696		

Data set: Jan.80-Oct.94
Model: SARIMA $(2,0,1)(2,2,0)_{12}$

	PAR EST	STD ERROR	T-STAT
AR(1)	0.213	0.250	0.853
AR(2)	0.379	0.214	1.776
MA(1)	-0.691	0.216	-3.191
SAR(1)	-0.736	0.059	-12.390
SAR(2)	-0.856	0.041	-20.870
CONSTANT	-2.167	3.168	-0.684
R^2	0.678		
Adjusted R^2	0.667		

Table B.11. Parameter estimates of Northeast Iowa corn production regressed on U.S. production

Data set	Regression Statistics	R^2s	Variables	*Coefficients*	*P-value*
80-89[a]					
	Multiple R	0.98	Intercept	-446.20	0.12
	R Square	0.97	U.S. Prod	0.56	0.00
	Adjusted R Square	0.97			
80-90					
	Multiple R	0.99	Intercept	-449.59	0.09
	R Square	0.97	U.S. Prod	0.56	0.00
	Adjusted R Square	0.97			
80-91					
	Multiple R	0.99	Intercept	-449.68	0.08
	R Square	0.97	U.S. Prod	0.56	0.00
	Adjusted R Square	0.97			
80-92					
	Multiple R	0.98	Intercept	-326.05	0.17
	R Square	0.97	U.S. Prod	0.54	0.00
	Adjusted R Square	0.97			
80-93					
	Multiple R	0.98	Intercept	-453.49	0.10
	R Square	0.96	U.S. Prod	0.55	0.00
	Adjusted R Square	0.95			

[a]The data set is used to estimate the relationship of regional production and U.S. production before in 1980-89, then use the estimated coefficient and the USDA WASDE production projection for 1990 to calculate the state production prediction for 1990.

Table B.12. Northeast Iowa corn and soybean multivariate state space model

Information Criterion for Autoregressive Models:

Lag=1	Lag=2	Lag=3	Lag=4	Lag=5	Lag=6	Lag=7	Lag=8	Lag=9	Lag=10
1166.74	1168.16	1164.82	1164.13	1165.45	1169.26	1165.68	1170.20	1158.85	1160.90

Yule-Walker Estimates for the Min AIC

	Lag=1		Lag=2		Lag=3		Lag=4	
	CBS_NIA	SBS_NIA	CBS_NIA	SBS_NIA	CBS_NIA	SBS_NIA	CBS_NIA	SBS_NIA
CBS_NIA	-0.10	0.02	-0.15	0.01	-0.35	0.17	-0.39	0.29
SBS_NIA	0.35	-0.45	0.09	-0.31	0.07	-0.23	-0.02	0.05

	Lag=5		Lag=6		Lag=7		Lag=8	
	CBS_NIA	SBS_NIA	CBS_NIA	SBS_NIA	CBS_NIA	SBS_NIA	CBS_NIA	SBS_NIA
CBS_NIA	-0.33	0.10	-0.17	0.01	-0.29	0.01	-0.11	0.05
SBS_NIA	0.05	-0.16	0.01	-0.25	0.01	-0.20	0.11	-0.24

	Lag=9	
	CBS_NIA	SBS_NIA
CBS_NIA	-0.30	0.10
SBS_NIA	-0.06	-0.21

Selected Statespace Form and Fitted Model

State Vector: CBS_NIA(T;T) SBS_NIA(T;T) CBS_NIA(T+1;T)

Estimate of the Transition Matrix

0	0	1
0.312	-0.303	0.307
-0.152	0.006	0.857

Input Matrix for the Innovation

1	0
0	1
-0.079	0.053

Variance Matrix for the Innovation

73.231	46.938
46.938	129.439

Parameter Estimates

Parameter	Estimate	Std. Err.	T value
F(2,1)	0.31	0.11	2.72
F(2,2)	-0.30	0.09	-3.55
F(2,3)	0.31	0.36	0.86
F(3,1)	-0.15	0.11	-1.40
F(3,2)	0.01	0.09	0.07
F(3,3)	0.86	0.10	8.32
G(3,1)	-0.08	0.10	-0.82
G(3,2)	0.05	0.08	0.70

APPENDIX C: AN EXAMPLE OF A SARIMA MODEL FORECAST

Soybean basis in Northeast Iowa is chosen for this illustration. The acf plot shows that it does not drop off quickly, this suggests the series is non-stationary (Figure C.1), so a first difference is calculated. Because the nonstationary series, pacf does not matter for identification. A seasonal pattern also can be seen in this acf plot (spikes at 12 and 24). Figure C.2 shows the acf of differenced series, which drops off to insignificance quickly at the regular lags, but not at seasonal lags (lag 24 has a bigger spike than lag 12). This suggest a nonstationary seasonal series and a seasonal difference of the first degree is needed.

Figure C.3 shows ACF and PACF plots for the series with both regular and seasonal first differences.[b] The acfs drops off quickly to insignificance after lag one for the nonseasonal lags, and the pacfs die down to not different than zero at lag five. This suggests a non-seasonal moving average term. The acf of seasonal lag is significant at 12 and not for lag 24, with the pacf decays from lag 12 to lag 24, which indicates a seasonal MA component, or both seasonal AR and MA terms.[c] So the tentative model is ARIMA $(0,1,1)(1,1,1)_{12}$, or the multiplicative model is:

$$(1-B)(1-B^{12})\,(1-\Phi_{12}B^{12})BS\_NIA = (1-\theta_1 B)(1-\Theta_{12}B^{12})a_t.$$

---

[b] The big spikes at lags 4 and 8 can be ignored. Pankratz argues that "strong seasonal variation can sometimes produce large (and misleading) autocorrelations at fractional multiples of the seasonal lag. These values can be misleading because they often become statistically insignificant after the realization is differenced by length s, or (in the residual acf) when AR and MA coefficients are estimated at seasonal lags (p.282).

[c] The PACF plot of longer data sets such as from 1980 to 1994, has pacf calculated for seasonal lags of 36, which is not significant. That is another indicator that seasonal component would include both AR and MA terms.

```
                                               0    0 0
AUTOCORRELATION FUNCTION OF THE SERIES    (1-B) (1-B )  BS_NIA

 1 0.79 .                           +     RRRRRRRRRRRRRRRRRRRRRRRRRRRR      .
 2 0.67 .                        +        RRRRRRRRRRRRRRRRRRRRRRRR          .
 3 0.60 .                      +          RRRRRRRRRRRRRRRRRRRRR             .
 4 0.57 .                     +           RRRRRRRRRRRRRRRRRRRR              .
 5 0.45 .                    +            RRRRRRRRRRRRRRRR                  .
 6 0.37 .                    +            RRRRRRRRRRRRRR+                   .
 7 0.33 .                   +             RRRRRRRRRRRR  +                   .
 8 0.31 .                   +             RRRRRRRRRRR   +                   .
 9 0.29 .                   +             RRRRRRRRRRR   +                   .
10 0.32 .                   +             RRRRRRRRRRRR  +                   .
11 0.33 .                   +             RRRRRRRRRRRR  +                   .
12 0.35 .                  +              RRRRRRRRRRRRR  +                  .
13 0.30 .                  +              RRRRRRRRRRR    +                  .
14 0.25 .                  +              RRRRRRRRR      +                  .
15 0.25 .                  +              RRRRRRRRRR     +                  .
16 0.22 .                 +               RRRRRRRR        +                 .
17 0.15 .                 +               RRRRRR          +                 .
18 0.09 .                 +               RRRR            +                 .
19 0.11 .                 +               RRRRR           +                 .
20 0.14 .                 +               RRRRRR          +                 .
21 0.09 .                 +               RRRR            +                 .
22 0.09 .                 +               RRRR            +                 .
23 0.09 .                 +               RRRR            +                 .
24 0.11 .                 +               RRRRR           +                 .
25 0.04 .                 +               RR              +                 .
26 -.02 .                 +              RR               +                 .
27 -.07 .                 +             RRR               +                 .
28 -.08 .                 +            RRRR               +                 .
29 -.13 .                 +           RRRRR               +                 .
30 -.14 .                 +          RRRRRR               +                 .
31 -.11 .                 +           RRRRR               +                 .
32 -.07 .                 +             RRR               +                 .
33 -.07 .                 +            RRRR               +                 .
                                                     0     0 0
PARTIAL AUTOCORRELATION FUNCTION OF THE SERIES  (1-B) (1-B  )  BS_NIA

 1 0.79 .                           +     RRRRRRRRRRRRRRRRRRRRRRRRRRRR      .
 2 0.12 .                           +     RRRRR +                           .
 3 0.08 .                           +     RRRR  +                           .
 4 0.15 .                           +     RRRRRR+                           .
 5 -.19 .                           RRRRRRR     +                           .
 6 -.03 .                           +    RR     +                           .
 7 0.07 .                           +     RRR   +                           .
 8 0.01 .                           +     R     +                           .
 9 0.09 .                           +     RRRR  +                           .
10 0.19 .                           +     RRRRRRR                           .
11 0.00 .                           +     R     +                           .
12 0.08 .                           +     RRRR  +                           .
13 -.14 .                           +RRRRRR     +                           .
14 -.16 .                           +RRRRRR     +                           .
15 0.14 .                           +     RRRRRR+                           .
16 -.10 .                           +  RRRR     +                           .
17 -.06 .                           +   RRR     +                           .
18 0.05 .                           +     RRR   +                           .
19 0.07 .                           +     RRR   +                           .
20 0.10 .                           +     RRRR  +                           .
21 -.11 .                           + RRRRR     +                           .
22 -.03 .                           +    RR     +                           .
23 -.03 .                           +    RR     +                           .
24 0.01 .                           +     R     +                           .
25 -.19 .                           RRRRRRR     +                           .
26 -.05 .                           +   RRR     +                           .
27 -.11 .                           + RRRRR     +                           .
28 0.09 .                           +     RRRR  +                           .
29 0.01 .                           +     R     +                           .
30 0.01 .                           +     R     +                           .
31 0.15 .                           +     RRRRRR+                           .
32 0.01 .                           +     R     +                           .
33 -.04 .                           +    RR     +                           .
```

Figure C.1. ACF and PACF of original series

```
                                                1     0 0
AUTOCORRELATION FUNCTION OF THE SERIES     (1-B) (1-B  )  BS_NIA

 1 -.20 .                           RRRRRRRR     +                            .
 2 -.12 .                            + RRRRR     +                            .
 3 -.09 .                            +  RRRR     +                            .
 4 0.21 .                            +     RRRRRRRR                           .
 5 -.11 .                           +  RRRRR      +                           .
 6 -.13 .                           +  RRRRR      +                           .
 7 -.04 .                           +     RR      +                           .
 8 -.03 .                           +     RR      +                           .
 9 -.13 .                           + RRRRRR      +                           .
10 0.05 .                           +      RRR    +                           .
11 0.01 .                           +      R      +                           .
12 0.15 .                           +      RRRRRR +                           .
13 0.06 .                           +      RRR    +                           .
14 -.13 .                           +  RRRRR      +                           .
15 0.11 .                           +      RRRRR  +                           .
16 0.08 .                           +      RRRR   +                           .
17 -.03 .                           +     RR      +                           .
18 -.16 .                           + RRRRRR      +                           .
19 -.03 .                           +     RR      +                           .
20 0.13 .                           +      RRRRR  +                           .
21 -.12 .                           +  RRRRR      +                           .
22 -.01 .                           +     RR      +                           .
23 -.01 .                           +      R      +                           .
24 0.20 .                           +      RRRRRRRR                           .
25 -.05 .                          +     RRR       +                          .
26 0.08 .                          +       RRRR    +                          .
27 -.11 .                          +   RRRRR       +                          .
28 0.11 .                          +       RRRRR   +                          .
29 -.07 .                          +     RRR       +                          .
30 -.09 .                          +    RRRR       +                          .
31 -.05 .                          +     RRR       +                          .
32 0.11 .                          +       RRRRR   +                          .
33 -.06 .                          +     RRR       +                          .

                                                      1     0 0
PARTIAL AUTOCORRELATION FUNCTION OF THE SERIES  (1-B) (1-B  )  BS_NIA

 1 -.20 .                           RRRRRRRR     +                            .
 2 -.17 .                            RRRRRRR     +                            .
 3 -.17 .                            RRRRRRR     +                            .
 4 0.14 .                            +     RRRRRR+                            .
 5 -.07 .                            +   RRR     +                            .
 6 -.14 .                            +RRRRRR     +                            .
 7 -.11 .                            + RRRRR     +                            .
 8 -.18 .                            RRRRRRR     +                            .
 9 -.24 .                          RRRRRRRRR     +                            .
10 -.09 .                            +  RRRR     +                            .
11 -.11 .                            + RRRRR     +                            .
12 0.09 .                            +     RRRR  +                            .
13 0.16 .                            +     RRRRRR+                            .
14 -.14 .                            +RRRRRR     +                            .
15 0.06 .                            +     RRR   +                            .
16 0.02 .                            +     RR    +                            .
17 -.06 .                            +   RRR     +                            .
18 -.07 .                            +   RRR     +                            .
19 -.11 .                            + RRRRR     +                            .
20 0.08 .                            +     RRRR  +                            .
21 -.02 .                            +    RR     +                            .
22 0.02 .                            +     RR    +                            .
23 -.05 .                            +   RRR     +                            .
24 0.15 .                            +     RRRRRR+                            .
25 0.02 .                            +     RR    +                            .
26 0.15 .                            +     RRRRRR+                            .
27 -.09 .                            +  RRRR     +                            .
28 -.01 .                            +     R     +                            .
29 0.01 .                            +     R     +                            .
30 -.15 .                            +RRRRRR     +                            .
31 -.02 .                            +    RR     +                            .
32 0.05 .                            +     RRR   +                            .
33 0.03 .                            +     RR    +                            .
```

Figure C.2. ACF and PACF of regular differenced series

```
                                                 1     12 1
AUTOCORRELATION FUNCTION OF THE SERIES     (1-B) (1-B   )  BS_NIA

 1 -.28 .                        RRRRRRRRRRR     +                            .
 2 -.11 .                           +  RRRRR      +                           .
 3 -.10 .                           +   RRRR      +                           .
 4 0.26 .                           +      RRRRRRRRRR                         .
 5 -.11 .                           +  RRRRR      +                           .
 6 0.03 .                           +      RR     +                           .
 7 0.03 .                           +      RR     +                           .
 8 -.25 .                         RRRRRRRRRR      +                           .
 9 0.00 .                          +       R       +                          .
10 0.15 .                          +       RRRRRR  +                          .
11 0.11 .                          +       RRRRR   +                          .
12 -.52 .                RRRRRRRRRRRRRRRRRRR       +                          .
13 0.21 .                         +        RRRRRRRR +                         .
14 -.10 .                         +     RRRR        +                         .
15 0.21 .                         +        RRRRRRRR +                         .
16 -.13 .                         +    RRRRR        +                         .
17 0.10 .                        +         RRRR      +                        .
18 -.07 .                        +      RRRR         +                        .
19 0.04 .                        +         RR        +                        .
20 0.10 .                        +         RRRR      +                        .
21 -.03 .                        +        RR         +                        .
22 0.01 .                        +         R         +                        .
23 -.04 .                        +        RR         +                        .
24 0.02 .                        +         RR        +                        .
25 -.04 .                        +        RR         +                        .
26 0.08 .                        +         RRRR      +                        .
27 -.13 .                        +     RRRRR         +                        .
28 0.06 .                        +         RRR       +                        .
29 -.07 .                        +       RRR         +                        .
30 0.02 .                        +         RR        +                        .
31 -.02 .                        +        RR         +                        .
32 0.06 .                        +         RRR       +                        .
33 -.03 .                        +        RR         +                        .

                                                         1     12 1
PARTIAL AUTOCORRELATION FUNCTION OF THE SERIES     (1-B) (1-B   )  BS_NIA

 1 -.28 .                        RRRRRRRRRRR     +                            .
 2 -.20 .                           RRRRRRRR     +                            .
 3 -.22 .                           RRRRRRRR     +                            .
 4 0.16 .                            +     RRRRRRR                            .
 5 -.02 .                            +    RR     +                            .
 6 0.06 .                            +     RRR   +                            .
 7 0.11 .                            +     RRRRR +                            .
 8 -.32 .                       RRRRRRRRRRRR     +                            .
 9 -.17 .                            RRRRRRR     +                            .
10 0.01 .                            +     R     +                            .
11 0.11 .                            +     RRRRR +                            .
12 -.41 .                    RRRRRRRRRRRRRRR     +                            .
13 -.02 .                            +    RR     +                            .
14 -.30 .                        RRRRRRRRRRR     +                            .
15 0.05 .                            +     RRR   +                            .
16 0.07 .                            +     RRR   +                            .
17 -.02 .                            +    RR     +                            .
18 0.13 .                            +     RRRRR +                            .
19 0.03 .                            +     RR    +                            .
20 -.14 .                            +RRRRRR     +                            .
21 -.12 .                            + RRRRR     +                            .
22 0.05 .                            +     RRR   +                            .
23 0.06 .                            +     RRR   +                            .
24 -.22 .                          RRRRRRRRR     +                            .
25 0.01 .                            +     RR    +                            .
26 -.21 .                           RRRRRRRR     +                            .
27 0.05 .                            +     RRR   +                            .
28 0.06 .                            +     RRR   +                            .
29 -.01 .                            +     R     +                            .
30 0.01 .                            +     R     +                            .
31 0.01 .                            +     R     +                            .
32 -.05 .                            +   RRR     +                            .
33 -.13 .                            +RRRRRR     +                            .
```

Figure C.3. ACF and PACF of regular and seasonal differenced series.

The results of estimation is:

```
MEAN OF SERIES =  0.5218E-01
VARIANCE OF SERIES =    241.4
STANDARD DEVIATION OF SERIES =     15.54

NET NUMBER OF OBS IS   119
DIFFERENCING: 1 CONSECUTIVE, 1 SEASONAL WITH SPAN 12
CONVERGENCE AFTER 16 ITERATIONS
INITIAL SUM OF SQS=      27193.444          FINAL SUM OF SQS=        11860.999

OR-SQUARE =  0.5837      R-SQUARE ADJUSTED =  0.5728
VARIANCE OF THE ESTIMATE-SIGMA**2 =   92.563
STANDARD ERROR OF THE ESTIMATE-SIGMA =   9.6210
AKAIKE INFORMATION CRITERIA -AIC(K) =   4.5951
SCHWARZ CRITERIA- SC(K) =   4.6885

                       PAR. EST.         STD. ERROR       T-STAT
              MA( 1)    0.35634          0.8718E-01        4.087
             SAR( 1)   -0.25950          0.8844E-01       -2.934
             SMA( 1)    0.87667          0.3285E-01        26.69
            CONSTANT   -0.86541E-01      0.1963           -0.4409

                                                  0      0 0
AUTOCORRELATION FUNCTION OF THE RESIDUALS  1-B) (1-B  )   R3

  1 0.02 .                              +    RR    +                      .
  2 -.17 .                              RRRRRRR    +                      .
  3 -.02 .                              +   RR     +                      .
  4 0.10 .                              +    RRRR  +                      .
  5 -.06 .                              +  RRR     +                      .
  6 -.02 .                              +   RR     +                      .
  7 -.02 .                              +   RR     +                      .
  8 -.21 .                             RRRRRRRR    +                      .
  9 -.12 .                              + RRRRR    +                      .
 10 0.08 .                             +     RRRR    +                    .
 11 0.04 .                             +     RR      +                    .
 12 0.02 .                             +     RR      +                    .
 13 0.06 .                             +     RRR     +                    .
 14 -.13 .                             +  RRRRR      +                    .
 15 0.14 .                             +     RRRRRR  +                    .
 16 0.02 .                             +     RR      +                    .
 17 -.01 .                             +     R       +                    .
 18 -.10 .                             +  RRRRR      +                    .
 19 0.08 .                             +     RRRR    +                    .
 20 0.06 .                             +     RRR     +                    .
 21 -.05 .                             +   RRR       +                    .
 22 0.06 .                             +     RRR     +                    .
 23 0.06 .                             +     RRR     +                    .
 24 -.06 .                             +   RRR       +                    .
```

All the coefficients (except constant term) are significant, and both the inevitability and stationary conditions are satisfied, $|\theta^*_1| < 1$, $|\Theta^*_{12}|<1$ and $|\Phi^*_{12}| <1$. But the model is not statistically adequate because residuals are not independent. The residual acf plot shows that there is a significant spike in lag 2; it has a t value larger than the warning levels (1.25 for lags 1, 2, and 3 and 1.6 for elsewhere except on seasonal lags). Going back to the estimated ACF and PACF plots of the differenced (both regular and seasonal) series, we may argue that both acf and pacf decay rather than one cuts off and another decays. This suggests another tentative model, adding an AR term into the above model. Other tentative models could include additional seasonal terms[d] because the residual acf plot show large spikes at quarter seasonal lags (4 and 8). Therefore the next two tentative models for estimation are:

$$(1-B)(1-B^{12})(1-\phi_1 B)(1-\Phi_{12}B^{12})BS\_NIA = (1-\theta_1 B)(1-\Theta_{12}B^{12})a_t \quad \text{and}$$

$$(1-B)(1-B^{12})\,(1-\phi_1 B)(1-\Phi_{12}B^{12}-\Phi_{24}B^{24})BS\_NIA = (1-\theta_1 B)(1-\Theta_{12}B^{12})a_t.$$

In terms of residual acfs, the second model performs better than the first one. The estimated coefficients and other results of the second model are listed as following:

```
R-SQUARE =  0.6008      R-SQUARE ADJUSTED =  0.5831
VARIANCE OF THE ESTIMATE-SIGMA**2 =    92.370
STANDARD ERROR OF THE ESTIMATE-SIGMA =    9.6109
AKAIKE INFORMATION CRITERIA -AIC(K) =   4.6266
SCHWARZ CRITERIA- SC(K) =   4.7668

                    PAR EST              STD ERROR          T-STAT
         AR( 1)     0.51189              0.1529              3.349
         MA( 1)     0.81679              0.1022              7.989
        SAR( 1)    -0.41182              0.9602E-01         -4.289
        SAR( 2)    -0.28712              0.9678E-01         -2.967
        SMA( 1)     0.73719              0.6278E-01          11.74
       CONSTANT    -0.99413E-01          0.8120E-01         -1.224

                                              0     0 0
AUTOCORRELATION FUNCTION OF THE SERIES    (1-B) (1-B   )  R3

 1 -.03 .                     +   RR    +                      .
 2 -.04 .                     +   RR    +                      .
 3 0.04 .                     +    RR   +                      .
 4 0.16 .                     +   RRRRRR+                      .
 5 -.04 .                     +  RRR    +                      .
 6 -.02 .                     +   RR    +                      .
 7 0.04 .                     +    RR   +                      .
 8 -.14 .                     +RRRRRR   +                      .
 9 -.08 .                     + RRRR    +                      .
10 0.03 .                     +    RR   +                      .
11 0.06 .                     +    RRR  +                      .
12 0.05 .                     +    RRR  +                      .
13 0.07 .                     +    RRR  +                      .
14 -.12 .                     + RRRRR   +                      .
15 0.15 .                     +   RRRRRR+                      .
16 -.05 .                    +   RRR     +                     .
17 0.00 .                    +     R     +                     .
18 -.09 .                    +  RRRR     +                     .
19 0.05 .                    +     RRR   +                     .
20 0.03 .                    +     RR    +                     .
21 -.09 .                    +  RRRR     +                     .
22 0.00 .                    +     R     +                     .
23 0.01 .                    +     R     +                     .
24 0.08 .                    +     RRRR  +                     .
```

All the coefficients are significant except the constant term. Again, the inevitability and stationary conditions are satisfied. No residual acf has t_values greater than the warning levels for the first three lags and seasonal lags. Though there are three out of 24 acfs with t values larger than the 1.6, these are at lags 4, 8 and 15. That's the best possible model for this

[d] Different seasonal terms are considered, such as second seasonal MA term, second AR term, or both. The second AR term is found significant and the overall model results in a better model over other additional seasonal terms.

data series (Northeast Iowa soybean basis), and it is used to forecast one to twelve months ahead bases. The following is the out-of-sample forecasts of Jan.-Dec. of 1991 for Northeast Iowa soybean basis based on data set 1980-1990:

FROM ORIGIN DATE 132, FORECASTS ARE CALCULATED UP TO 12 STEPS AHEAD

FUTURE DATE	LOWER	FORECAST	UPPER
133	-53.3926	-34.5528	-15.7130
134	-50.7457	-27.7985	-4.85126
135	-60.5679	-35.4740	-10.3801
136	-48.3391	-21.7983	4.74254
137	-56.6903	-29.0079	-1.32558
138	-61.9175	-33.2452	-4.57291
139	-58.5616	-28.9833	0.594992
140	-55.1027	-24.6702	5.76235
141	-66.8294	-35.5783	-4.32716
142	-75.0140	-42.9713	-10.9286
143	-75.0443	-42.2322	-9.42007
144	-72.8081	-39.2457	-5.68334

STD.DEV. OF ONE-STEP-AHEAD ERRORS-SIGMA = 9.612

REFERENCES CITED

Akaike, Hirotugu. "Canonical Correlation Analysis of Time Series and Use of the Information Criterion." *System Identification: Advances and Case Studies*, ed. Mehra, R.K. and D.G.Laintiotis. New York: Academic Press, 1976.

Aoki, Masanao. *State Space Modeling of Time Series*. Berlin: Springer-Verlag, 1987.

Aoki, Masanao and Arthur Havenner. "State Space Modeling of Multiple time Series." *Econometrics Reviews* 10(1991):1-59.

Azoff, E. Michael. *Neural Network Time Series Forecasting of Financial Markets*. Chichester, England: John Wiley & Son, 1994.

Benirschka, Martin and James K. Binkley. "Optimal Storage and Marketing Over Space and Time." *America Journal of Agricultural Economics* 77(1995):512-524.

Box, George and Gwilym Jenkins. *Time Series Analysis, Forecasting, and Control*. San Francisco, California: Holden Day, 1976.

Brennan, Donna, Jeffrey Williams, and Brian D. Wright. Brennan, Donna, Jeffrey Williams, and Brian D. Wright. "Convenience Yield without the Convenience: A Spatial-Temporal Interpretation of Storage under Backwardation." Forthcoming in *The Economic Journal* (1997).

Brennan, Michael J. "The Supply of Storage." *American Economic Review* 47(1958):50-72.

Brorsen, B. Wade, Jean-Paul Chavas, Warren R. Grant, and A. W. Ngenge. "Spatial and Temporal Relationships among Selected U.S. Grain Markets." *North Central Journal of Agricultural Economics* 7(1985):1-10.

Cheng, Bing and D.M. Titterington. "Neural Networks: A review from a statistical perspective." *Statistical Science* 9(1994):2-54.

Chicago Board of Trade (CBOT). *Understanding Basis--The Economics of Where and When*. Chicago: CBOT, 1990.

Claussen, Kent L. and J.William Uhrig. "Cash Soybean Price Prediction with Neural Networks." *Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Proceedings of the NCR-134 Conference* (1994): 56-65.

Cramer, Gail L, and Eric J. Wailes, eds. *Grain Marketing*. Boulder, Colorado: Westview Press, Inc., 1993.

Cumby, Robert E. and D. M. Modest. "Testing for Market Timing Ability." *Journal of Financial Economics* 19(1987):169-189.

Dasgupta, Chanda Ghose, Gary S. Dispensa, and Sanjoy Ghose. "Comparing the Predictive Performance of a Neural Network Model with Some Traditional market Response Models." *Journal of Forecasting* 10(1994):235-244.

Diebold, Francis X. "State Space Modeling of Time Series: A Review Essay." *Journal of Economic Dynamics and Control* 13(1989):597-612.

Fama, Eugene F. and Kenneth R. French. "Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the theory of Storage." *Journal of Business* 60(1987):55-73.

Fruin, Jerry E., Daniel W. Halbach and Lowell D. Hill. "Corn movements in the United States: Interregional flow patterns and transportation requirements in 1985". *Northern Central Regional Research Bulletin* 326, Agricultural Experiment Station, University of Illinois, Urbana-Champaign, 1990

Garcia, Philip and Darrel L. Good. "An Analysis if the Factors Influencing the Illinois Corn Basis, 1971-1981." *Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Proceedings of the NCR-134 Conference* (1983):306-326.

Garcia, Philip, Robert J. Hauser and Alan D. Tumblin. "Corn and Soybean Basis Behavior: an Intertemporal, Cross-Sectional Analysis." *Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Proceedings of the NCR-134 Conference* (1986):128-142.

Gorr, Wilpen L. "Research Perspective on Neural Network Forecasting." *International Journal of Forecasting* 10(1994):1-3.

Gorr, Wilpen L., Daniel Nagin, and Janusz Szczypula. "Comparative Study of Artificial Neural Network and Statistical Models for Predicting Student Grade Point Averages." *International Journal of Forecasting* 10(1994):17-34.

Greene, William H. *Econometric Analysis.* New York: Macmillan Publishing Company, 1993.

Granger, C.W.J. and G. McCollister. *Comparison of Forecasts of Selected Series by Adaptive, Box-Jenkins and State Apace Methods.* Los Angeles, California: ORSA/TIMS, 1979.

Grundnitski, Gary and Larry Osburn. "Forecasting S&P and Gold Futures Prices: An Application of Neural Networks." *Journal of Futures Markets* 13(1993):631-643.

Hamm, Lennie, B. Wade Brorson, and Ramesh Sharda. "Futures Trading with a Neural Network." *Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Proceedings of the NCR-134 Conference* (1993):286-296.

Harvey, Andrew C. *Forecasting Structural Time Series Models and the Kalman Filter*. New York: Cambridge University Press, 1989.

Hauser, Robert J., Philip Garcia and Alan D. Tumblin. "Basis Expectations and Soybean Hedging Effectiveness." *North Central Journal of Agricultural Economics* 12(1990):127-136.

Heifner, Richard G. "The Gains from Basing Grain Storage Decisions on Cash-Future Spreads." *Journal of Farm Economics* 48(1966):1490-1495.

Henriksson, Roy D. and Robert C. Merton. "On Market Timing and Investment Performance II: Statistical Procedures for Evaluating forecasting Skills." *Journal of Business* 54(1981):513-533.

Hieronymus, T. A. "Basis Patterns." *Views from the Trade: Readings in Futures Marketings*. Edited by A.E. Peck. Chicago: CBOT, Book III(1978):45-56. This paper was originally presented in 1963 at the Hedging Symposium for Country Grain Elevators at the Chicago Board of Trade.

Hill, Tim and Leorey Marquez, Marcus O'Connor, and William Remus. "Artificial Neural Network Model for Forecasting and Decision Making." *Journal of Forecasting* 10(1994):5-15.

Jain, Anil K. and Jianchang Mao. "Artificial Neural Networks: A Tutorial." *Computer* 29(1996):31-44.

Jenkins, Gwilym M. *Practical Experiences with Modeling and Forecasting Time Series*. St. Helier: Gwilym Jenkins & Partners (Overseas) Ltd. 1979.

Kaastra, T and M.S. Boyd. "Forecasting Futures Trading Volume Using Neural Networks." *Journal of Futures Markets* 15(1995): 953-970.

Kahl, Kandice H. "Changes in the Chicago Corn Basis, 1960-1975." *Agricultural Economics Research* 34(1982): 25-29.

Kahl, Kandice H. and Charles E, Curtis, Jr. " A Comparative Analysis of the Corn Basis in Feed Grain Deficit and Surplus Areas." *Review of Research in futures Markets* 5(1986):220-232.

Kaldor, Nicholas. "Speculation and Economic Stability." *Review of Economic Studies* 7(1939):1-27.

Kaldor, Nicholas. "A Note on the Theory of the Forward Market." *Review of Economic Studies* 7(1940):196-201.

Karlson, Nicholas; Brad Anderson and Reynold Dahl. *Cash-Futures Price Relationships: Guide to Corn Marketing*. Dept. of Agricultural and Applied Economics, University of Minnesota, Staff Paper Series P93-1 (1993).

Kennedy, Peter. *A Guide to Econometrics*. Cambridge, Massachusetts: The MIT Press, 1989.

Kenyon, David E. and Steven E. Kingsley. "An Analysis of Anticipatory Short Hedging Using Predicted Harvest Basis." *Southern Journal of Agricultural Economics* 5(1973):199-203.

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

Koehler, Anne B. and Emily S. Murphree. "A Comparison of Results from State Space Forecasting with Forecasts from the Makridakis Competition." *International Journal of Forecasting* 4(1988):45-55.

Kohzadi, N., M.S. Boyd, I. Kaastra, B.S. Kermanshahi, and Iebeling Kaastra. "Forecasting Livestock Prices with an Artificial Neural Network versus Linear time Series Models." *Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Proceedings of the NCR-134 Conference* (1994):131-143.

Kohzadi, N., M.S. Boyd, I. Kaastra, B.S. Kermanshahi, and D. Scuse. "Neural Networks for Forecasting: an Introduction." *Canadian Journal of Agricultural Economics* 43(1995):463-474.

Kuan, C.M. and T. Liu. "Forecasting Exchange Rates Using Feedforward and Recurent Neural Networks." *Journal of Applied Econometrics* 10(1995):347-364.

Larson, Donald W., Thomas R. Smith and E. Dean Baldwin. "Soybeans movements in the United States: Interregional flow patterns and transportation requirements in 1985". *Northern Central Regional Research Bulletin* 323, Agricultural Experiment Station, University of Urbana-Champaign, Sep., 1990

Leuthold, Raymond M., Joan C. Junkus and Jean E. Cordier. *The Theory and Practice of Futures Markets*. Lexington, Massachusetts: Lexington Books (1989): chapter 3 and chapter 7.

Marshall, John F. *Futures and Option Contracting: Theory and Practice*. Cincinnati, Ohio: South-Western Publishing CO., 1989.

Martin, Larry, John L. Groenewegen and Edward Pidgeon. "Factors Affecting Corn Basis in Southwestern Ontario." *American Journal of Agricultural Economics* 62(1980):107-112.

McCarthy, Joseph and Mohammd Najand. " State Space Modeling of Price and Volume Dependence: Evidence from Currency Futures." *Journal of Futures Markets* 13(1993):335-344.

McIntosh, Christopher S. and Jeffrey H. Dorfman. "A Comparison of Two Performance Measures." *American Journal of Agricultural Economics* 74(1992):209-214.

Mehra, R.K. "Identification in Control and Econometrics." *Current Development in the Interface: Economics, Econometrics, Mathematics*, ed. Hazewinkel, M. and A.H.G.Rinnooykan. Dordrecht, Holland: D. Reidel Publishing Company, 1982.

Mittnik, S.. "Forecasting with Balanced State Space Representations of Multivariate Distributed Lad Models." *Journal of Forecasting* 9(1990):207-218.

Mittnik, S.. "Macroeconomic Forecasting Experience with Balanced State Space Models." *International Journal of Forecasting* 6(1990):337-348.

Naik, Gopal and Raymond M. Leuthold. A Note on the Factors Affecting Corn Basis Relationships. *Southern Journal of Agricultural Economics* 23(1991):147-153.

Pankratz, Alan. *Forecasting with Univariate Box-Jenkins Models*. New York: Wiley-Interscience Publications, 1991.

Patterson, K. D.. "A State Space Model for Reducing the Uncertainty Associated with Preliminary Vintages of Data with an Application to Aggregate Consumption." *Economics Letters* 46(1994):215-222.

Patterson, K. D.. "A State Space Approach to Forecasting the Final Vintage of Revised Data with an Application to the Index of Industrial Production." *Journal of Forecasting* 14(1995):337-350.

Powers, Nicholas and Aaron Johnson, Jr. "Forecasting the Storage-season Wisconsin Basis for Corn." *Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Proceedings of the NCR-134 Conference* (1983):327-341.

Ravallion, Martin. "Testing Market Integration." *American Journal of Agricultural Economics* 68(1986):102-109.

SAS Institute. *SAS/ETS® User's Guide*, v6. Cary, North Carolina: SAS Institute Inc., 1993.

Sorenson, L. Orlo. "Grain Transportation." in Gail L. Cramer and Eric J. Wails eds. *Grain Marketing*. Boulder, Colorado: Westview Press, Inc., 1993.

Strobl, Maximilian, T. Randall Fortenbery, and Paul L. Fackler. *Forecasting Futures Basis: Model Interpretation and Evaluation*. Unpublished paper. North Carolina State University, 1996.

Taylor, Patricia D. and William G. Tomek. "Forecasting the Basis for corn in Western New York." *Journal of Northeast Agricultural Economic Council* 13(1984):97-102.

Telser, Lester G. "Futures Trading and the Storage of Corn and Wheat." *Journal of Political Economy* 66(1958):1-22.

Tomek, William G.. *Commodity Futures Prices as Forecasts*. Working Paper, Department of Agricultural, Resource, and Managerial Economics, Cornell University. WP96-07. August 1996.

Tomek, William G. and Robert J. Myers. "Empirical Analysis of Agricultural Commodity Prices: A Viewpoint." *Review of Agricultural Economics* 15 (1993):181-201.

Thompson, S. R., J. S. Eales, and R. J. Hauser. "An Empirical Analysis of Cash and Futures Grain Price Relationships in the North Central Region." *North Central Journal of Agricultural Economics* 12(1990):241-254.

Uhrig, J. William, Bernard A. Engel and W. Lance Baker. "An Application of Neural Networks: Predicting Corn Yields." *Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Proceedings of the NCR-134 Conference* (1992):407-417.

U.S. Department of Agriculture, Foreign Agricultural Service. "Transportation: Vital link in Exporting to the World." *AgExporter* 7(1995):4-5.

Vinod, H.D. and Parantap Basu. "Forecasting Consumption, Income and Real Interest Rates form Alternative State Space Models." *International Journal of Forecasting* 11(1995):217-231.

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

_______. "The theory of Price of Storage." *American Economic Review* 39(1949):1254-1262.

Wright, B. D., and J.C. Williams. "A Theory of Negative Prices for Storage." *Journal of Futures Markets* 9(1989):1-13.